CrossMark

The role of human capital and innovation in economic development: evidence from post-Malthusian Prussia

Francesco Cinnirella[1] · **Jochen Streb**[2]

Abstract The effect of human capital on growth involves multiple channels. On the one hand, an increase in human capital directly affects economic growth by enhancing labor productivity in production. On the other hand, human capital is an important input into R&D and therefore increases labor productivity indirectly by accelerating technological change. In addition, different types of human capital such as basic and higher education or training-on-the-job might play different roles in both production and innovation activities. We merge individual data on valuable patents granted in Prussia in the late nineteenth-century with county-level data on literacy, craftsmanship, secondary schooling, and income tax revenues to explore the complex relationship between various types of human capital, innovation, and income. We find that the Second Industrial Revolution can be seen as a transition period when it comes to the role of human capital. As in the preceding First Industrial Revolution, "useful knowledge" embodied in master craftsmen was related to innovation, especially of independent inventors. As in the subsequent twentieth century, the quality of basic education

✉ Francesco Cinnirella
cinnirella@ifo.de

Jochen Streb
streb@uni-mannheim.de

[1] CESifo, CEPR, and CAGE, Ifo Institute, Poschingerstr. 5, 81679 Munich, Germany

[2] University of Mannheim, L 7, 3-5, 68131 Mannheim, Germany

Springer

was associated with both workers' productivity and firms' R&D processes. In a final step, we show that literacy had also a negative effect on fertility which increased with innovation. In general, our findings support the notion that the accumulation of basic human capital was crucial for the transition to modern economic growth.

Keywords Human capital · Growth · Literacy · Innovation · Patents · Prussia

JEL Classification N13 · N33 · O11 · O31

1 Introduction

Most growth economists probably agree on the notion that human capital is one of the decisive factors in explaining why some countries became rich and others remained poor in the past two centuries (Goldin and Katz 1998; Lucas 2002; Glaeser et al. 2004; Hanushek and Woessmann 2008; Jones and Romer 2010; Galor 2011; Acemoglu et al. 2014; Gennaioli et al. 2013). There is still dissent with regard to the channels through which human capital fosters economic growth, though. On the one hand, human capital is interpreted as an independent factor of production which increases productivity for a given level of technology (Lucas 1988; Mankiw et al. 1992). On the other hand, human capital is seen as an input in the innovation process and therefore as a complement to technology (Benhabib and Spiegel 1994; Romer 1990; Nelson and Phelps 1966). According to this view, higher levels of human capital lead to the generation or diffusion of new technologies or to a more efficient adoption of a given technology, thereby shifting the frontier of the production-possibility set outwards.

In this paper we empirically address the role of human capital by analyzing the economic development in Prussian counties in the late nineteenth century. This is the period in which Prussia as a national economy caught up successfully to the more advanced British economy and even took over the technological leadership in some high-tech sectors such as chemicals or electrical engineering (Pierenkemper and Tilly 2004; Broadberry and Burhop 2010). Since the academic debate about the relative importance of different types of human capital during industrialization is not yet settled, we consider the influence of three different measures, namely literacy rates, craftsmanship, and secondary school enrollment rates. Valuable patents are used as a measure for innovation. We first analyze which type of human capital was mostly related to innovation in the Second Industrial Revolution. We find that "useful knowledge" embodied in skilled craftsmen was still an important determinant for innovation activities of private inventors in more traditional industries such as steam boilers or mechanical metal processing. In contrast to the First Industrial Revolution, however, the average stock of basic knowledge measured by the literacy rate was positively related to firms' innovation in high-tech industries such as chemicals or electrical engineering.

The surprising result that the stock of literate people was associated with high-tech innovations can be explained by three different lines of argument. First, high literacy rates might indicate an above-average general investment in children's human capital that comprised also higher education, nutrition, and preventive health care (Baten et al. 2014; Kelly et al. 2015; Prados de la Escosura 2015; Mühlhoff 2016). Second, a high literacy rate might have facilitated transregional knowledge spillover and therefore a county's outstanding innovativeness because reading was the only way to learn about ideas occurred outside the own locality. Third, high-tech innovations and formal basic education might have been complements in

highly innovative industries where white-collar occupations increased exponentially during the Second Industrial Revolution (Goldin and Katz 1998).

After having established the relationship between different types of human capital and innovation, we explore the impact of the former on income. Because data on GDP per capita are not available for Prussian counties we use information on per capita income-tax revenues as a second-best measure.[1] In the late nineteenth century, a Prussian household's income tax was based on its taxable income from capital, immovable property, trade and commerce, and dependent employment. We therefore argue that Prussian income tax revenues reflect the productivity of the economic activity of the county. Our empirical analysis shows that literacy and patenting activities are independently correlated with income. This evidence lends support to the notion that basic education had an impact on Prussian growth through both the direct production channel and the indirect innovation channel.

The fact that our empirical analysis is based on county-level data within a single country should alleviate the problem of disentangling the effect of human capital from that of institutions (Acemoglu et al. 2001; Glaeser et al. 2004; Acemoglu et al. 2014). Yet we know that, for instance, the process of peasants' emancipation which started in Prussia in 1807 has been very heterogeneous, which means that land-tenure and labor market relations might still have been different across counties in the late nineteenth century (Cinnirella and Hornung 2016; Cinnirella and Schueler 2016). What is more, Acemoglu et al. (2011) have shown that Napoleonic institutional reforms significantly influenced the economic development in some German states that later became part of Prussia. More recent studies have focused on geography, natural resources, and access to markets to explain the location and the rise of manufacturing in the nineteenth century (Crafts and Wolf 2014; Fernihough and O'Rourke 2014; Gutberlet 2014). That is why we will account for differences in institutions, geography, and market access to ensure that the estimated effect of human capital on income is not mediated by these factors.

Finally, we study another channel through which human capital might have affected Prussian economic growth. An important characteristic of the transition to modern economic growth is the rapid decline in fertility rates. Unified Growth Theory postulates that parents' increased emphasis on children's education, triggered by higher levels of technological progress, determined lower fertility outcomes (Galor 2011, 2012). Recent studies have provided empirical evidence of such a mechanism (Becker et al. 2010, 2012; Murphy 2015). We extend this literature by showing that the negative impact of literacy on fertility is larger in more technologically advanced counties, where arguably the returns to education were higher.

The rest of the paper is organized as follows. Section 2 discusses the role of human capital in economic growth on the basis of the theoretical and empirical literature. Section 3 introduces the data and discusses our measures for innovation, human capital, and income. Section 4 analyzes the relationship between different types of human capital and innovation. Section 5 tests whether literacy had a persistent impact on Prussian counties' income tax revenues net of the effects of master craftsmen and formal secondary education. Section 6 analyzes to what extent literacy and innovation affect income independently of each other. Section 7 establishes that the negative impact of literacy on fertility increased with technological progress. Section 8 concludes.

---

[1] Prussia introduced a general income tax in 1891. See Schremmer (1994), pp. 151–155.

2 Related literature

In their influential paper, Mankiw et al. (1992) assume that a society's aggregated stock of human capital can serve as a substitute for technology. That is why they augment Solow's classical growth model (Solow 1956) with the factor human capital which is assumed to increase the workforce's productivity. Analyzing cross-country variation of GDP per capita in 1985 for 1998 non-oil producing countries Mankiw et al. (1992) show that the inclusion of the investment in human capital, measured as the fraction of the population aged 15–19 that is enrolled in secondary schools, increases the performance of their regression model considerably. As a result, countries with similar technologies, rates of savings, and population growth might still not converge because of enduring differences in national education policies.

Benhabib and Spiegel (1994) raise serious doubts on whether human capital should enter the macroeconomic production function as a separate factor. Following the seminal approach of Nelson and Phelps (1966) they assume instead that human capital is primarily needed to increase the level of technology by adopting foreign superior technology or by creating own innovations. In other words, human capital is not a substitute for but a necessary component to innovation. In their empirical approach, Benhabib and Spiegel assume that a country's level of technology is determined by both imitating and innovative activities. A country's capability to innovate is estimated by its human capital *stock* measured by average years of schooling in the labor force. Its potential to imitate is approximated by the existing gap between the technological level of the productivity leader and its own inferior technological level. It is important to note that they find a significant impact only of *levels* of human capital on growth of total factor productivity. In contrast, the *growth* of human capital does not affect economic performance. Our analysis will concentrate on levels as data on literacy and the proxy for income are only available for a single cross-section.

Vandenbussche et al. (2006) stress the importance of the *composition* of human capital between basic knowledge and higher education, where the former is relevant for imitation and the latter is important for innovation.[2] As a result, backward countries which want to catch up by imitating (foreign) technology should invest primarily in primary schooling, whereas countries at the technological frontier should concentrate on increasing the quality and quantity of higher education.[3] In our analysis we will also distinguish between the basic level of human capital measured by the literacy rate and higher education measured by enrollment rates in secondary education.

In historical perspective, the role of formal education during the First Industrial Revolution is generally considered as minor, especially in the case of the British industrial revolution (Mitch 1993, 1999; Clark 2005) for which relative prices of production inputs (Allen 2003) or cultural factors (Clark 2009; Mokyr 2009) have been regarded as the main determinants. O'Rourke and Williamson (1995) argue that the contribution of education to Sweden's catch-up to Britain and the United States was also modest. Similarly, Taylor (1999) finds only second order effects of school enrollment rates on the growth performance of seven countries in the period 1870–1914.

Recently, some scholars have started to stress the importance of a few highly skilled craftsmen and entrepreneurs as engines of the First Industrial Revolution, while downplaying the importance of average levels of literacy (Galor and Moav 2002; Mokyr 2005; Mokyr and

---

[2] Note that we do not have the data necessary to test whether human capital formation facilitated the imitation of (foreign) technology. Richter and Streb (2011), however, show for the German machine tool industry that it was primarily the former imitators that became successful innovators in the late nineteenth century.

[3] See also Acemoglu et al. (2006) for a model on firm's innovation and imitation activities and distance to the technological frontier.

Voth 2009). Explaining why the Industrial Revolution started in Britain, Meisenzahl and Mokyr (2011) argue that the main advantage was "the supply of highly skilled, mechanically able craftsmen who were able to adapt, implement, improve, and tweak new technologies and who provided the micro inventions necessary to make macro inventions highly productive and remunerative". If this was true, the institution of apprenticeship, and not the school system, provided the necessary skills for technological progress (Van der Beek 2014; Van der Beek and Feldman 2016).

Squicciarini and Voigtländer (2015) support this "upper-tail knowledge" hypothesis. They show for pre-industrial France that the local presence of a rather tiny "knowledge elite", proxied by the density of subscriptions to Diderot's and d'Alembert's Encyclopédie, had a significant positive effect on city growth between 1750 and 1850, whereas the effect of the average worker's skills, measured by the literacy rate, was negligible. Their main argument is that the knowledge elite fostered economic growth by transforming "useful knowledge" into profitable innovations. In contrast, better educated workers only raised the productivity of any given technology, but they did not help to adopt new technology.

In a recent paper, Kelly et al. (2015) point out that, if one looks at county level data, literacy rates still had a positive impact on the English industrialization. Showing that high literacy rates were positively associated with firms' patenting activities in innovative industries such as chemicals and electrical engineering, our analysis of nineteenth-century Prussian counties supports this revisionist view. This latter result is puzzling, however, because it is not in line with the theoretical assumption that successful innovation activities in science-based industries first and foremost require a sufficient amount of higher educated people (Vandenbussche et al. 2006) but are less affected by the quantity of basic education (here measured by literacy rates).

We offer three alternative explanations to resolve this apparent contradiction between theory and history. First, it is conceivable that high literacy rates capture also an above-average investment in children's human capital which comprised not only primary schooling but also higher education, nutrition, and preventive health care such as smallpox vaccination (Baten et al. 2014; Kelly et al. 2015; Prados de la Escosura 2015; Mühlhoff 2016). If this was true, high literacy rates would signal that a county's (younger) workforce had comparatively high cognitive skills. Second, it was after all literacy that provided access to a broader knowledge base needed to develop innovations. In a world without modern mass communication techniques the only way to learn more about the ideas, theories and experiences of people outside the own local network was to read books, personal letters, newspapers, patents or encyclopedias. That is why a high literacy rate might have facilitated transregional knowledge spillover and therefore a county's above-average innovativeness. Baten and van Zanden (2008), for example, have shown that book production, which is also driven by the demand of more and more literate people, was related to growth in real wages in the period 1450–1849. Third, high-tech innovations and basic formal education might have been complements during the Second Industrial Revolution. It is often argued that industrialization replaced skilled craftsmen with machinery and unskilled laborers. Goldin and Katz (1998), however, point out that skilled workers were still needed to install and maintain machinery and to supervise and direct the unskilled workforce. It was especially in innovative industries where more and more white-collar workers were hired to man the newly founded divisions for sales, accounting, and R&D. At the highly innovative chemical firm BASF, for example, the number of chemists, clerks, and technicians increased tenfold between 1876 and 1900 (von Hippel 2002, p. 106). As a result, innovative firms might have settled in counties with high education standards or have lobbied for an improvement of primary education. Both strategies would explain the positive correlation between literacy and patenting.

To identify the exact historical linkages between literacy and patenting activities in Prussian counties is beyond the scope of this study. Based on the data at hand, however, we provide ample evidence that these linkages must have existed. What is more, we explicitly take into account that different types of human capital (literacy, craftsmanship, and secondary schooling) might have had different effects on both innovation and income.

In a long-run perspective, Unified Growth Theory assumes that in the Malthusian period technological progress was mainly affected by the scale of the population, via a larger supply of innovative ideas, a higher demand for innovation, and accelerated technological diffusion. In later stages of development, when the technological progress had become more complex, human capital took a crucial role in advancing the technological frontier (Galor 2011, pp. 147–148). The accumulation of human capital is also seen as a main driver of the demographic transition, which is a necessary precondition for modern economic growth. By combining the neoclassical Solow model with elements of the demographic transition, Dalgaard and Strulik (2013) explain about two thirds of the cross-country variation in GDP per worker in 2000. Their study is of particular interest for our research as they also distinguish between a direct impact of human capital (measured by enrollment rates and school years) on growth and an indirect effect via technological innovation. Controlling for innovation with the help of the proxy R&D expenditures, they still find a sizable direct impact of human capital on growth (Dalgaard and Strulik 2013, Table 5).

3 Data

3.1 Valuable patents

Pure patent counts allocate the same weight to every patent, no matter whether it has a low or a high economic value for the patentee or the society. Using the number of patents as an indicator for innovative activity can therefore lead to a large measurement error.[4] One way to distinguish patents with a high value from those with a low value is to use a patent's lifespan as an indicator for its profitability (Schankerman and Pakes 1986; for other methods to identify valuable patents, see Streb 2016). The caveat is that this method works only for patent systems with periodical renewal fees. The German patent law introduced in 1877 met this precondition (Seckelmann 2006, pp. 170–179). It allowed both private persons and firms to apply for patent protection that could last up to 15 years. To keep a patent granted in force the patentee had to pay an annual renewal fee which grew stepwise from 50 Marks in the first and second year to 700 Marks in the fifteenth year. Facing the rising expenditures for holding a particular patent, a patentee was supposed to decide to renew his patent only when the cost of doing this was lower than the expected future returns of the patent. The assumption that this mechanism worked as intended is supported by the finding that, in the late nineteenth century, 70% of all German patents granted had already been given up after just 5 years. Only about 5% of all patents reached the maximum age of 15 years (Streb et al. 2006, p. 352). We therefore assume in the following that only those long-lived Prussian patents that survived at least 10 years can be interpreted as valuable innovations.[5]

---

[4] For a discussion of the shortcomings of patent statistics, see Griliches 1990.

[5] A basic question about the life span approach is how many years a patent had to be in force to be interpreted as a valuable patent. Streb et al. (2006) explain in great detail why they decided to choose the cut-off point of 10 years to distinguish patents with a low economic value from those with a high economic one.
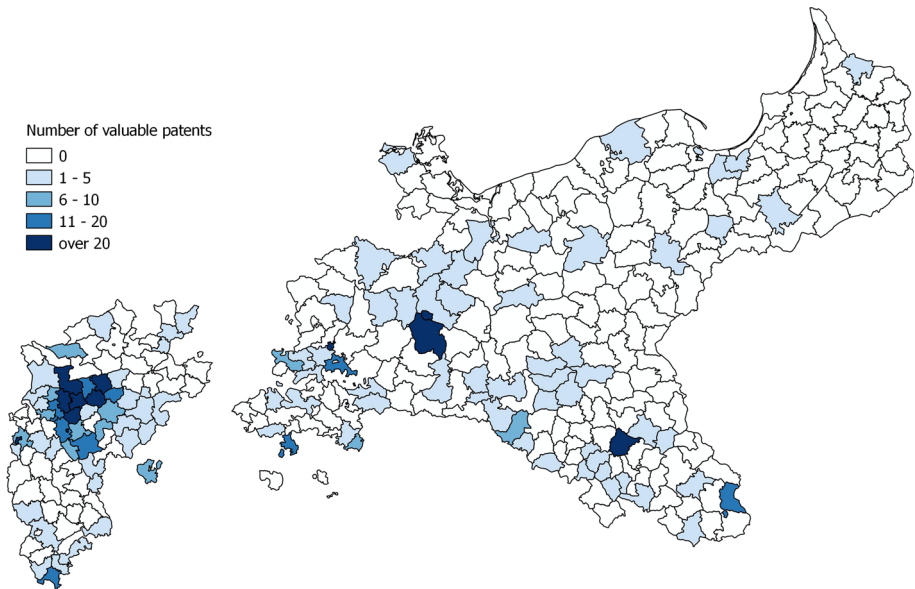
Fig. 1 The stock of valuable patents, 1877–1890. *Note* Prussian county borders as in 1849

By merging individual data on valuable patents granted in Prussia from 1877 to 1890 with county-level information our project makes use of two different, recently assembled historical data bases.[6] The prime source for the patent data is the annual *Verzeichnis der im Vorjahre erteilten Patente* published by the German patent office since 1877, which lists all patents granted in the preceding year. From these periodical listings, in 2005, Jörg Baten and Jochen Streb selected all patents that were held for at least 10 years and were therefore interpreted as valuable. In this selection, we identified 1740 valuable Prussian patents granted between 1877 and 1890, which will be used as a measure for the level of successful innovation. The information attached to each of these patents comprises the year in which the patent was granted, the technological class of the invention, as well as name and location of the patent holder. The name of the patent holder allows us to distinguish, admittedly with some margins of error, whether a patent was held by an independent inventor (or a small private company) or by a firm (private partnership or incorporated company), because in the latter case, information about the legal form of the company[7] (e.g. GmbH, KG or AG) was added to the name of the patent holder.[8] The information about the location of the patent holder is used to assign the origin of a particular patent to a geographical unit. In the original data base, the patents were allocated to administrative districts (*Regierungsbezirk*), the middle administrative level of the German states (Streb et al. 2006, p. 365). To get a more disaggregated picture of the geographical distribution of valuable patents we re-assigned them to the rural and urban counties ("*Land- und Stadtkreise*"), the lower administrative level of

---

[6] For the patent data, see Streb et al. (2006, 2007), Labuske and Streb (2008), Degner (2009), Richter and Streb (2011), Lehmann-Hasemeyer and Streb (2016). For the Prussian county level data see Becker et al. (2014).

[7] The different organizational forms among which firms could choose in late-nineteenth-century Germany are described in Guinnane et al. (2007).

[8] Differently from other patent systems, for example the one in the US, the German system reported only the patent holder and did not distinguish between original inventor and patent holder.
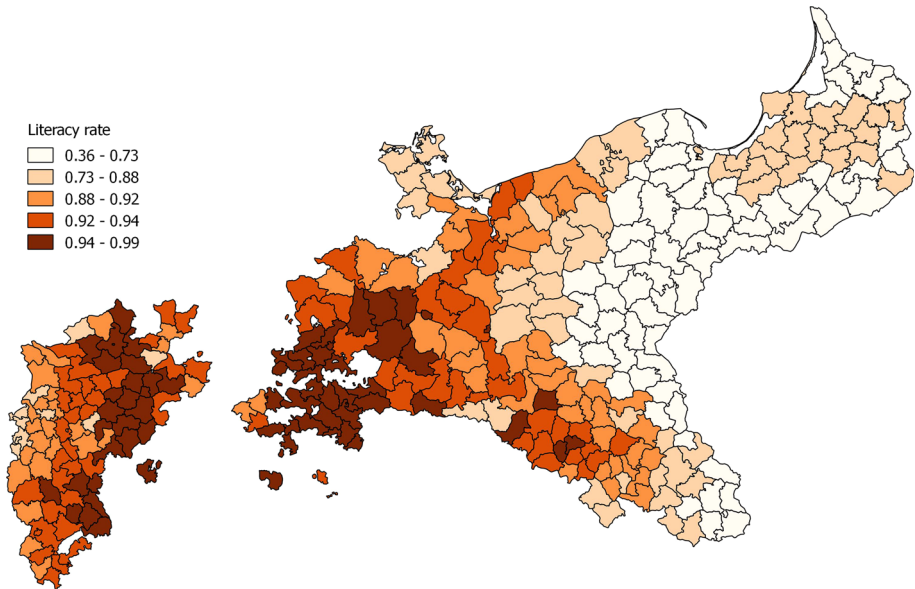
Fig. 2 Literacy rates in 1871. *Note* Quintiles. Prussian county borders as in 1849.

the Prussian state. Figure 1 shows the geographic distribution of valuable patents granted in Prussia between 1877 and 1890 in the borders of 1849.[9]

Figure 1 reveals that valuable patents were by no means equally distributed across Prussia. Innovative centers were the Ruhr area, the counties neighboring the rivers Rhine and Main, the Province of Saxony, Greater Berlin and parts of Silesia, whereas large parts of Poznán, Pomerania, West- and East Prussia display very low patenting activity. Firms' innovative activities were even more geographically concentrated.

3.2 Human capital

As explained above, the first part of our empirical analysis aims at studying which type of human capital is mostly related to innovation.[10] To capture the different skill categories we use the literacy rate, the enrollment rate in secondary education, and the share of master craftsmen in our analysis. The literacy rate refers to the ability to read and write of people older than 10 in 1871 as reported in the population census.[11] Following Benhabib and Spiegel (2005) we expect that the *levels* of literacy have an impact on innovation and economic growth. In Fig. 2 we show the geographic distribution of the literacy rate. As discussed in Becker and

---

[9] Note that we will display all maps with county borders as in 1849. In all regressions that include information about the distribution of master craftsmen in 1849, we also use the counties of 1849 as the units of observation. In the regressions that do not control for master craftsmen, we analyze the Prussian counties in the borders of 1871 thereby increasing the number of observations. The geographical distribution of patents granted in the US and Great Britain is presented in Sokoloff (1988) and Moser (2011), respectively.

[10] Proxies for human capital generally used in historical context are literacy, enrollment rates, or numeracy based on the extent of age heaping (Baten and Crayen 2010).

[11] Prussian censuses also provide information about the enrollment rate in primary schools. However, enrollment rates are only an imperfect measure for the stock of human capital. Indeed, in the second part of the analysis (Table 6) we will see that primary school enrollment rate has no measurable impact on income.
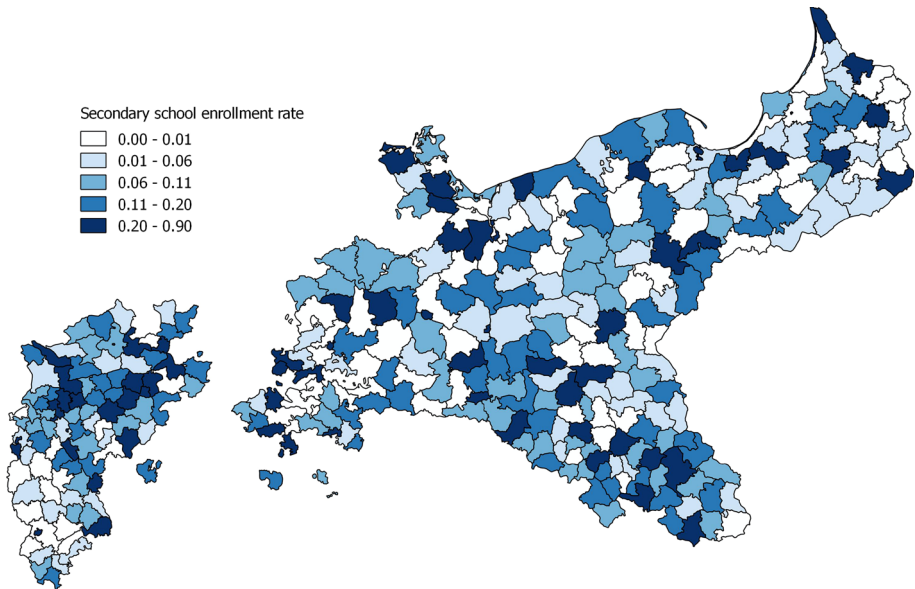
Fig. 3 Secondary school enrollment rates in 1864. *Note* Quintiles. Prussian county borders as in 1849.

Woessmann (2009), the literacy rate is high in the central counties of Prussia near Wittenberg, whereas it is lower in the peripheral counties, especially in the east.

Since we have no information about the *level* of secondary education in Prussian counties, we follow the example of Mankiw et al. (1992) and use as a proxy secondary school enrollment rates constructed for the years 1864 and 1873.[12] Secondary education in Prussia was heterogeneous in terms of curriculum as it included a humanistic track, the *Gymnasium*, and a more business-related track, the *Realschule*. Given that the majority of students in Prussia enrolled in the *Gymnasium*, it is not clear a priori whether we should expect a positive or a negative relationship between enrollment in secondary education and innovative activity.[13] Figure 3 shows that secondary schooling, which was usually located in urban districts, was more evenly distributed across Prussia than our other measures of human capital.

In the spirit of Meisenzahl and Mokyr (2011) as well as Squicciarini and Voigtländer (2015) we also use information about the relative number of master craftsmen to capture the extent of "useful knowledge" that was locally available. The Prussian census of 1849 reports data on employment by trade and explicitly distinguishes between masters and apprentices. Using this information we can construct the share of *master* craftsmen in the population.[14]

---

[12] Data on enrollment rates in secondary education in 1864 come from the census and refer to *Realschulen, Progymnasien, Gymnasien, Mittlere und höhere Privatschulen und Erziehungsanstalten aller Art, Provinzial-, Kunst- und Gewerbe, Ackerbau-, Navigations- und Handelsschulen, Handwerker-Fortbildungsanstalten, sog. Sonntagsschulen*. Information on secondary education in 1873 has been inferred from Wiese (1874) and refers to *Gymnasium, Progymnasium, Realschulen, and Bürgerschulen*. The two measures are therefore not directly comparable as they include different categories in their numerator. Nevertheless, the correlation between the two rates is 0.7. Both rates are computed on the male population aged 15–19 in 1864 and 1871, respectively, as information on the reference population in the year 1873 is not available.

[13] Semrad (2015) analyzes the effect of secondary education on economic outcomes in nineteenth century Bavaria.

[14] The trades reported in the census used to construct the share of master craftsmen are the following: soap boiler, lorimer, rope maker, sailmaker, firehosemaker, tailor, hatter, well sinker, cabinetmaker,
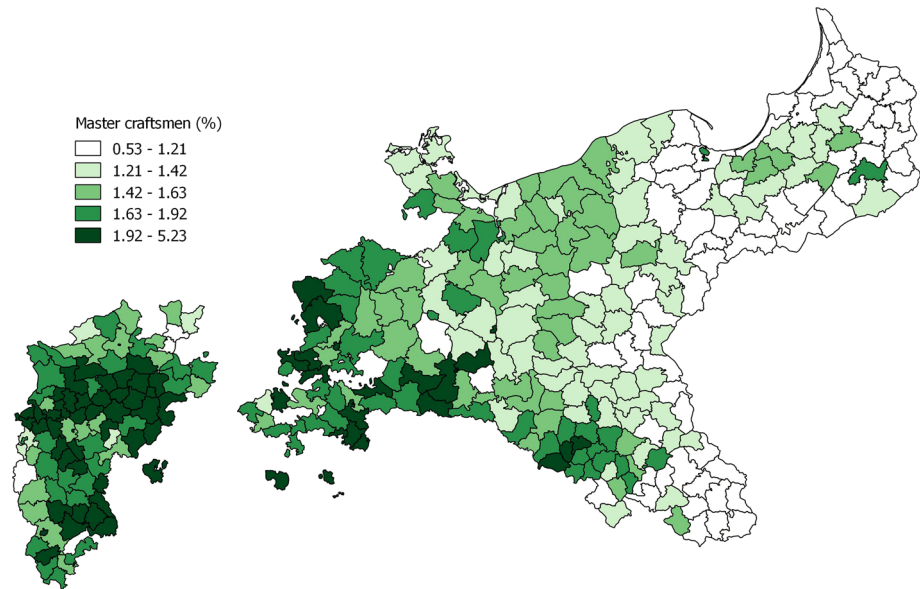
Fig. 4 Master craftsmen in 1849 (in percent). *Note* Quintiles. Prussian county borders as in 1849. Master craftsmen include the following trades: soap boiler, lorimer, rope maker, sailmaker, firehosemaker, tailor, hatter, well sinker, cabinetmaker, wheelwright, wainwright, cooper, turner, comb maker, brush maker, basket maker, slater, potter, glazier, blacksmith, locksmith, belt maker, coppersmith, bell founder, tin molder, plumber, wire drawer. We have also added self-employed producers of scientific instruments, musical instruments, clockmakers, goldsmith, and engravers

Figure 4 shows that master craftsmen are concentrated in the early industrial centers such as Silesia or the Ruhr region.

3.3 Income tax revenues

As already explained above, we will rely on information on income tax revenues when estimating the effect of human capital and innovation on income. The Prussian Statistical Office published data on income tax revenues and the number of income taxpayers averaged over the period 1899–1903 for all Prussian counties.[15] Since the tax reform of 1891, a Prussian household's income tax had been based on its taxable income (gross income minus income-related expenses) resulting from capital, immovable property, trade and commerce, and dependent employment. All households with an annual taxable income below 900 Marks were exempted from income tax. As a result, in 1900, only about 30% of the Prussian population was required to pay income tax. Therefore, we estimate only income of the prosperous part of Prussian counties' inhabitants who could arguably make especially good

---

Footnote 12 continued

wheelwright, wainwright, cooper, turner, comb maker, brush maker, basket maker, slater, potter, glazier, blacksmith, locksmith, belt maker, coppersmith, bell founder, tin molder, plumber, wire drawer. We have also added self-employed producers of scientific instruments, musical instruments, clockmakers, goldsmith, and engravers.

[15] The source of the 1901 income tax statistics is *Festschrift des Königlich Preussischen Statistischen Bureaus zur Jahrhundertfeier seines Bestehens* (1905).Berlin: Verlag des Königlich Preussischen Statistischen Bureaus.
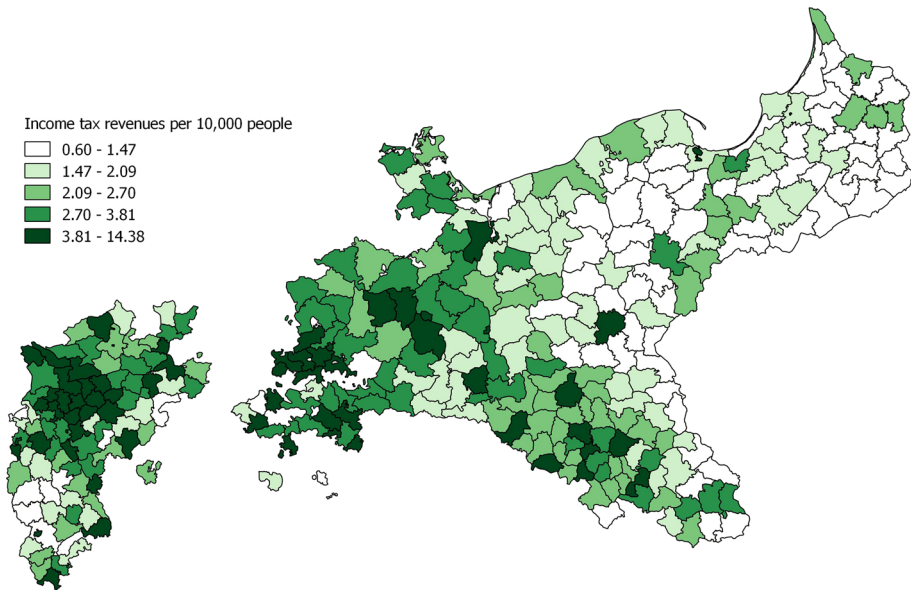
Fig. 5 Income tax revenues per 10,000 people. *Note* Quintiles. Average income tax revenues for the period 1899–1903 in German Marks. Prussian county borders as in 1849

use of their human capital stock. The progression of income tax rates was very moderate: Tax rates ranged from 1.13% for taxable incomes between 900 and 3000 Marks up to 3.95% for taxable incomes beyond 100,000 Marks (Schremmer 1994, pp. 152–155). Consequently, the estimation bias resulting from disproportionately high tax revenues in counties with an above-average number of people with very high income is limited. Using log tax revenues will also decrease the influence of observations with extremely high values. It is also important to note that our tax variable refers only to state income tax revenues *excluding* local surcharges. This ensures that we are not measuring different propensities to tax across counties. Figure 5 shows the geographic distribution of income tax revenues per 10,000 people in German Marks. Figure 6 displays the cross-sectional distribution of log income tax revenues per capita across Prussian counties which resembles the typical log-normal distribution of income.

We also considered the problem that income tax revenues were only reported in nominal terms. We collected provincial data on prices of grain, meat, and butter for the period 1899–1903[16] and used them to construct a province-specific cost of living index based on the weights presented by Allen (2009, Table 2.3). The potential issue of tax evasion was tackled by correcting our information about tax revenues with estimates of tax evasion in Prussian districts (*Regierungsbezirk*) around 1900 by Meisel (1911, p. 313). Both modifications did not change our main results and are therefore only reported in the online Appendix (see Table A9).

---

[16] Provincial prices are from the Preußisches Statistisches Landesamt (2008): M*arkt- und Kleinpreise ausgewählter Güter in Preußen 1816 bis 1928*. Available at GESIS Datenarchiv, Cologne, histat. Studiennummer 8302, Datenfile Version 1.0.0.
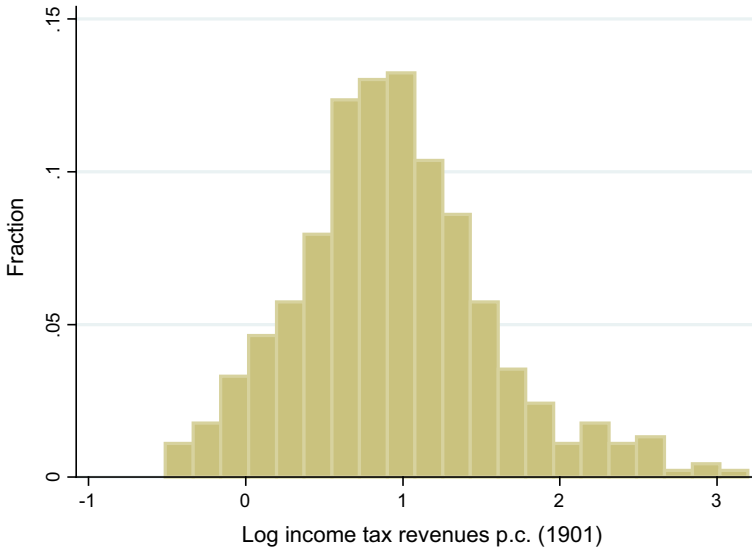
Fig. 6 The distribution of log income-tax revenues per capita across Prussian counties in the period 1899–1903. *Source*: see text

3.4 Control variables

Due to the abundance of information provided by the 1871 (and 1882) Prussian census we can include a rich set of control variables in our regression analysis.[17] First of all, we will account for scale effects by controlling for log population size (Galor 2011). In addition, we include urbanization rate in the set of control variables.

Religious precepts can influence investments in education. Becker and Woessmann (2009) have shown that counties that were dominated by Protestants invested more in primary education than Catholic counties. Botticini and Eckstein (2005, 2007) argue that religious norms induced Jewish communities to acquire higher levels of human capital. Indeed, data on the number of illiterate people *within* religious denominations reported in the 1871 census show that illiteracy rates were particularly low for Jews and Protestants: mean illiteracy rate among Jews was 5.6 and 6.8% among Protestants, whereas illiteracy rate among Catholic came to 11.3%.[18] In our regression analysis we therefore also control for the share of Protestant and Jews in the Prussian counties.[19]

Cinnirella and Hornung (2016) show that large landowners impeded the accumulation of human capital in nineteenth-century Prussia because they were primarily interested in securing a stock of unschooled and cheap farm laborers, and thus opposed the education of

---

[17] See Königliches Statistisches Bureau (1874): *Die Gemeinden und Gutsbezirke des Preussischen Staates und ihre Bevölkerung: Nach den Urmaterialien der allgemeinen Volkszählung vom 1. Dezember 1871*. Berlin.

[18] Note that overall literacy rate and illiteracy rate by religious denominations cannot be directly compared as the statistics rely on different reference populations.

[19] When using distance to Wittenberg as an instrument for literacy we will exclude the share of Protestants from the regression analysis as we do not expect to find any residual impact of Protestantism on income (Becker and Woessmann 2009). Anyway, when including the share of Protestantism, we obtain the same results.

Table 1 Descriptive statistics. *Source*: see text.

	Mean (1)	Std. Dev. (2)	Min. (3)	Max. (4)
Income tax revenues per capita (German Marks in 1901)	2.96	2.22	0.60	14.38
Literacy rate	0.84	0.14	0.36	0.99
Secondary school enrollment rate	0.12	0.14	0.00	0.90
Primary school enrollment rate	0.83	0.05	0.54	0.94
Master craftsmen (%)	1.58	0.49	0.53	5.23
Absolute number of patents	4.23	26.91	0	464
Patents per 10,000 people (1877–1890)	0.35	0.94	0.00	7.60
Firm patents per 10,000 people (1877–1890)	0.10	0.32	0.00	2.95
Independent patents per 10,000 people (1877–1890)	0.25	0.72	0.00	7.03
Log population size	10.88	0.44	9.60	13.62
Urban population (%)	27.85	20.04	0.00	100.00
Share of large landownership	0.31	0.23	0.00	0.82
Protestants (%)	59.98	38.87	0.26	99.80
Jews (%)	1.09	1.17	0.00	7.37
Labor force in manufacturing (%)	11.57	5.76	2.24	29.15
Born in municipality (%)	57.82	11.77	33.57	84.94
Share of non-German speaking people	0.15	0.29	0.00	0.94
Linguistic polarization	0.20	0.35	0.00	1.00
Distance to Wittenberg (km)	332.4	163.4	6.11	731.53
Distance to Berlin (km)	328.6	160.4	3.05	653.05

Descriptive statistics for county borders in 1849. Secondary school enrollment rates are observed in 1864 and are computed over the male population aged 15–19. Primary school enrollment rates are observed in 1886 and are computed over the population aged 6–14. Master craftsmen include the following trades: soap boiler, lorimer, rope maker, sailmaker, firehosemaker, tailor, hatter, well sinker, cabinetmaker, wheelwright, wainwright, cooper, turner, comb maker, brush maker, basket maker, slater, potter, glazier, blacksmith, locksmith, belt maker, coppersmith, bell founder, tin molder, plumber, wire drawer. We have also added self-employed producers of scientific instruments, musical instruments, clockmakers, goldsmith, and engravers

the peasants' children.[20] The occupational census in 1882 reports information on farm size. To account for the influence of large landowners we constructed the variable "share of large landholdings" as the area of landholdings larger than 100 hectares over the total area.

In the late nineteenth century many Prussians migrated to regions with high demand for labor, for example to the coal fields of the Ruhr area. To control for this inflow of people who might have also brought along new ideas and technologies, we make use of the variable "share of people born in the respective municipality".

Descriptive statistics of the variables used in our analysis are shown in Table 1.[21] The average level of literacy rate (84%) is one of the highest in Europe for the period under consideration. However, as also shown in Fig. 2, there is a large variation in literacy as in 10% of the counties one out of three individuals was not able to read and write.

---

[20] For a theoretical perspective on the relationship between land distribution and human capital formation see Galor et al. (2009).

[21] To be consistent with most of the empirical analysis, the descriptive statistics refer to the county borders in 1849.

There is also a large variation in the number of valuable patents in our sample. The average number of valuable patents in a county for the period 1877–1890 is about 4, with 65% of the counties having zero valuable patents. The city-county of Berlin reached the outstanding number of 464 valuable patents in the period under consideration, which is due to the fact that very innovative firms such as Siemens and Agfa were located there (Degner 2009). Berlin-based inventors might also have profited from their proximity to the Imperial Patent Office which facilitated communication with the members of this highly-centralized administration. To take this effect into account, we also control for the distance to Berlin. However, since Berlin constitutes a clear outlier we exclude the county from the first part of the analysis where we study the determinants of innovative activity.[22]

4 The relationship between different types of human capital and innovation

We begin our empirical analysis by examining the relationship between the different types of human capital and a county's innovative activity. We distinguish between basic knowledge measured by the literacy rate, higher formal education in the form of secondary school enrollment rate, and useful practical knowledge measured by the share of master craftsmen.[23] These three variables for human capital are standardized with zero mean and unit standard deviation to facilitate comparison. Since the share of master craftsmen is computed from the occupational census in 1849 we aggregate the variables at the county borders of 1849 which reduces the number of observations to 335. Because of its very large number of patents (464), we drop the city-county of Berlin when analyzing patenting activities. Thus the final number of observations is 334.

The dependent variable is the absolute number of valuable patents granted in the period between 1877 and 1890, a count variable that takes on nonnegative integer values. Given the particular nature of the dependent variable we will use different estimation methods from the class of count-data models. This will not only be a test of the robustness of our results but will also lead to a better understanding of the role of the different dimensions of human capital on innovation.

In Table 2 we employ different models to study the relationship between the absolute number of patents and the different types of human capital. We begin with estimating a standard OLS model (columns 1–2). Since the OLS model is not appropriate when dealing with a large number of zeros in the dependent variable, we constrain the sample to those counties with a positive number of patents ($n = 115$). The OLS estimates suggest that both literacy and master craftsmen are significantly and positively related to the number of valuable patents, conditional on having at least one valuable patent. The coefficients for secondary school enrollment rates are insignificant.

In columns 3 and 4 we estimate the zero-inflated negative binomial model which accounts for the excessive number of zeros and for over-dispersion, i.e. the fact that the (conditional) variance exceeds the (conditional) mean. In particular, this model assumes that the excess-zeros are generated by a separate binary process which is independent from the process that leads to the count data. Since most of the patents in our sample are related to manufactur-

---

[22]   Including Berlin in the analysis does not change the results.

[23]   In the online Appendix we show the distribution of the literacy rate in 1871 (Figure A1) and the distribution of patents p.c. in 1877–1890 (Figure A2). In Figure A3 we present, in the left panel, the relationship between patents p.c. (*10,000) and literacy rate and, in the right panel, the relationship between patents p.c. (*10,000) and the share of master craftsmen.

ing (even the patent class 45 "agriculture" mostly contained agricultural machinery), the geographical distribution of valuable patents follows rather mechanically the geographical distribution of the industrial centers. That is why we assume that the percentage of people working in manufacturing predicts the excess number of zeros. The hypothesis is that counties with a low level of manufacturing were less dependent on innovation and therefore had zero valuable patents. Hence, in the first step, a logit model is used to estimate whether or not a Prussian county was capable (or willing) to get at least one valuable patent. In the second step, a count model is employed to estimate the absolute number of valuable patents in those counties where the occurrence of a valuable patent is more likely. Note that in the second step of the estimation zero patents are also possible. Yet, by predicting the excess-zeros in this way we reduce the potential impact of human capital on zero innovation, i.e. we implicitly assume that human capital mostly predicts variation in patents in the positive domain.

Columns 3 and 4 of Table 2 reveal that the percentage of the labor force working in manufacturing is indeed a good predictor of the excess zeros: the higher the percentage of people in manufacturing, the lower is the probability of observing a county with zero valuable patents. The estimates show that literacy and master craftsmen are significantly related to the number of valuable patents.[24] In columns 5 and 6 we estimate a standard negative binomial (NB) model which does not explicitly model the excess zeros but accounts for over-dispersion. Again, the estimates suggest a positive relationship between both literacy and master craftsmen and a county's innovation, though the coefficient for master craftsmen in column 6 comes only close to significance with a p value of 0.107. In columns 7 and 8 we show the results of a Poisson model which is a special case of the negative binomial distribution assuming that the dispersion parameter is zero. These estimates also indicate that both literacy and master craftsmen are positively related to patenting. In sharp contrast, enrollment in secondary education has no significant impact on patenting activities regardless of the model we use. Finally, in column 9 we run a "horse-race" between literacy and master craftsmen in the Poisson model. This estimate implies that patenting activity was influenced rather by literacy than by master craftsmen. As the following regression results will show, however, the impact of the different types of human capital on innovation depends on both the type of patentee and the technology we look at.

With respect to the control variables, we find that both a larger population and a higher urbanization rate are associated with more innovation. However, we generally do not find any association of religions such as Protestantism or Judaism with patenting. This suggests that, once accounting for the human capital channel, Protestantism and Judaism had not residual impact on innovation.[25]

So far, the results suggest that both literacy and master craftsmen are positively related to patenting activities. The next subsection will show that we can sharpen our results by explicitly considering the new role of firms in the innovation process of the Second Industrial Revolution.

4.1 Human capital and innovation by type of patent holder

The rise of chemicals and electrical engineering in the late nineteenth century relied much more on the availability and use of scientific knowledge than the innovations of the First Industrial Revolution in the late eighteenth century (Moser 2012, 2013). That is why, during the nineteenth century, the proverbial ingenious inventor lost ground to a more systematic and

---

[24] Note that in all specifications we include the full set of control variables.

[25] Estimates showing the full set of controls are reported in the online Appendix, Tables A1–A5.

Table 2 Human capital and patenting activity

	OLS		ZINB		NB		Poisson		Poisson
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Literacy rate	4.995**		0.564**		0.659**		0.946***		0.813**
	(1.953)		(0.228)		(0.286)		(0.326)		(0.336)
Master craftsmen (share)		3.398**		0.273**		0.233		0.166**	0.095
		(1.631)		(0.131)		(0.144)		(0.073)	(0.075)
Secondary school e. r.	−2.224	−1.069	−0.194	−0.178	−0.053	−0.006	0.022	0.222	0.059
	(2.045)	(2.038)	(0.156)	(0.163)	(0.164)	(0.170)	(0.138)	(0.140)	(0.156)
Perc. labor force in manufacturing	0.003	0.001			0.149***	0.161***	0.085***	0.094***	0.087***
	(0.185)	(0.187)			(0.030)	(0.031)	(0.025)	(0.023)	(0.025)
Inflate perc. labor force in manufacturing			−0.414***	−0.432***					
			(0.076)	(0.079)					
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	115	115	334	334	334	334	334	334	334
R-squared	0.49	0.52							
Pseudo-R²					0.20	0.19	0.67	0.66	0.67

The dependent variable is the absolute number of valuable patents. OLS estimates in columns 1 and 2 are restricted to counties with the number of valuable patents >0. In columns 3–4 we estimate zero-inflated negative binomial (ZINB) models with the percentage of labor force in manufacturing as the predictor of zeros; in columns 5–6 we estimate a negative binomial (NB) model; in column 7–9 we estimate a Poisson model. Literacy rate, share of master craftsmen, and secondary school enrollment rates are standardized with zero mean and unitary standard deviation. The manufacturing sector includes mining, construction, and manufacture of metals, machinery, equipment, chemicals, textiles, paper, leather, food products, and wood. Control variable included are: log population, urbanization rate, share of large landownership, percent of Protestants, percent of Jewish, distance to Berlin. Berlin is excluded from the regression. Robust standard errors in parenthesis. Significance at * 10, ** 5, *** 1%

science-based research carried out in the newly-founded R&D departments of larger firms. This was especially true in Germany where firms like BASF, Bayer or Siemens profited from the advanced national education and research system.[26]

In Table 3 we estimate different count-data models for independent innovations (Panel A) and for firms' innovations (Panel B). In addition to our usual control variables we take into account the share of people employed in banking, insurance, and law in 1882. As patent fees were relatively expensive in Germany, easy access to credit and patent lawyers might have especially facilitated private inventors' patenting activity.

Our estimates now lead to a clearer result. Panel A shows that the coefficient for master craftsmen is significantly positive regardless of the model we use (columns 5–8, Panel A) whereas literacy is only weakly related to independent innovation (columns 1–4, Panel A). The horse race in column 9 provides additional support to the notion that master craftsmen and not literacy was related to independent innovation in the German Empire.

Panel B reveals that the opposite is true with respect to firms' innovation: Literacy is significantly and positively related to firms' innovation (columns 1–4, Panel B), whereas master craftsmen do not show any significant relationship with the innovation outcome (columns 5–8, Panel B). The horse race regression in column 9 confirms this finding.

Interestingly enough, when controlling for urbanization, the impact of secondary schooling is, in many cases, significantly negative with respect to firms' innovation. This result might partly be explained by the fact that larger German firms already hired R&D personnel and other white-collar employees on the supra-regional labor market (Baten et al. 2007). Another possible explanation refers to the different types of secondary schools with very different curricula. As already mentioned, our variable for secondary education includes enrollment in both *Gymnasium* and *Realschule*, where the former had a more humanistic curriculum, and the latter was more business-oriented.[27] The fact that the *Gymnasium* dominated numerically secondary education (on average 60% of the students were enrolled in the *Gymnasium*) could explain the negative relationship between secondary education and firms' patenting activity. Counties with a relatively large number of secondary schools might have had comparatively many public offices and private firms that were specialized in administrative or distributive aspects of manufacturing rather than in production activities where the patentable technological innovation occurred.[28]

Considered as a whole, the findings of Table 3 suggest that the practical knowledge of master craftsmen which played such a prominent role during the First Industrial Revolution was still useful for independent innovations in more "traditional" technological fields such as machine-building, but rather unimportant for developing the science-based innovations of the Prussian firms that were engaged in chemicals and electrical engineering.

We can test this sector-related hypothesis by differentiating innovation activities by technological classes (see Streb et al. 2006, p. 359). On the one hand, we look at "traditional" innovations ("machines") whose valuable patents were assigned to the technological fields "metal working", "machine parts", and "steam boilers". On the other hand, we consider

---

[26] See, for example, Murmann (2003) and Grupp et al. (2002). See also Burhop and Lübbers (2010) and Labuske and Streb (2008).

[27] According to the regulation of 1856, in the *Gymnasium*, 10 h out of 30 h per week went to the study of Latin and 6 h to the study of Greek. Only 3–4 h per week were devoted to the study of mathematics and 1 to 2 h to the study of physics or natural history (Wiese 1874, p. 76). In the *Realschule*, however, between 2 and 6 h out of 32 h per week were dedicated to the study of natural science, 5 or 6 h to mathematics, and "only" about 5 h to the study of Latin (Wiese 1874, p.78). The so-called "Realschule of second order" was even more scientifically oriented and completely without Latin and Greek (Wiese 1874, p. 79).

[28] We thank an anonymous referee for providing this interpretation.

Table 3 Human capital and patenting activity by patent holder

Panel A	Independent innovation				Independent innovation				
	(1) ZINB	(2) ZINB	(3) NB	(4) Poisson	(5) ZINB	(6) ZINB	(7) NB	(8) Poisson	(9) Poisson
Literacy rate	0.405*	0.263	0.443	0.571**					0.376
	(0.239)	(0.236)	(0.273)	(0.284)					(0.275)
Master craftsmen (share)					0.254*	0.219**	0.233**	0.201***	0.166***
					(0.136)	(0.106)	(0.112)	(0.059)	(0.063)
Secondary school e.r.	−0.065	−0.123	0.011	−0.009	−0.045	−0.107	0.067	0.099	0.008
	(0.168)	(0.169)	(0.196)	(0.168)	(0.177)	(0.172)	(0.201)	(0.144)	(0.162)
Perc. labor force in manufacturing			0.136***	0.113***			0.140***	0.119***	0.115***
			(0.026)	(0.025)			(0.025)	(0.024)	(0.025)
Inflate Perc. labor force in manufacturing	−0.383***	−0.436***			−0.402***	−0.439***			
	(0.083)	(0.095)			(0.089)	(0.092)			
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bank, Insurance, Law	No	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Observations	334	334	334	334	334	334	334	334	334
Non-zero obs	105	105	105	105	105	105	105	105	105

Table 3 continued

Panel B	Firm innovation				Firm innovation				
	(1) ZINB	(2) ZINB	(3) NB	(4) Poisson	(5) ZINB	(6) ZINB	(7) NB	(8) Poisson	(9) Poisson
Literacy rate	0.853*	0.586	1.312***	1.432***					1.394**
	(0.467)	(0.506)	(0.356)	(0.524)					(0.571)
Master craftsmen (share)					0.093	0.048	0.183	0.132	0.025
					(0.219)	(0.197)	(0.184)	(0.093)	(0.101)
Secondary school e.r.	−0.454**	−0.515**	−0.579**	−0.442	−0.494**	−0.536*	−0.582**	−0.134	−0.440
	(0.226)	(0.257)	(0.253)	(0.280)	(0.245)	(0.277)	(0.293)	(0.203)	(0.284)
Perc. labor force in manufacturing			0.170***	0.052*			0.188***	0.054**	0.052*
			(0.048)	(0.028)			(0.048)	(0.027)	(0.028)
Inflate Perc. labor force in manufacturing	−0.506***	−0.546***			−0.581***	−0.602***			
	(0.132)	(0.145)			(0.140)	(0.143)			
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bank, Insurance, Law	No	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Observations	334	334	334	334	334	334	334	334	334
Non-zero obs	56	56	56	56	56	56	56	56	56

In the upper Panel A, the dependent variable is the absolute number of valuable patents of independent individuals. In the bottom Panel B, the dependent variable is the absolute number of valuable patents of firms. Zero-inflated negative binomial (ZINB) estimates are presented in columns 1–2 and 5–6; negative binomial (NB) estimates in columns 3 and 7; Poisson estimates in columns 4, 8, and 9. Literacy rate, share of master craftsmen, and secondary school enrollment rates are standardized with zero mean and unitary standard deviation. The manufacturing sector includes mining, construction, and manufacture of metals, machinery, equipment, chemicals, textiles, paper, leather, food products, and wood. Control variables included are: log population, urbanization rate, share of large landownership, percent of Protestants, percent of Jewish, and distance to Berlin. Berlin is excluded from the regression. Controls for bank, insurance, and law include the share of people employed in banks and insurance, the share of people employed in legal occupations. Robust standard errors in parenthesis. Significance at * 10, ** 5, *** 1%

patents granted in the fields of "electrical engineering" and "chemicals" including "synthetic dyes" as science-based innovations (chemistry/electricity).

In Panel A of Table 4 we focus on innovation in the "machine" industry. The estimates of both the negative binomial and the Poisson model in columns 1–2 and 4–5 (Panel A) depict a strong pattern: Literacy is not related to "traditional" innovation in the machine industry, whereas a larger share of master craftsmen is strongly related to that type of innovation.[29] These results also hold when using an alternative outcome variable. Columns 6 shows a significant correlation between master craftsmen in 1849 and the share of people employed in metal working computed from the occupational census in 1882. Consistently, literacy is not related to occupation in metal working (column 3). This result is also corroborated by the horse race regressions presented in columns 7 and 8.[30]

The estimates presented in Panel B of Table 4 focus on innovation and occupation in the scientifically advanced sectors chemistry and electrical engineering.[31] The results are unambiguous: literacy is strongly related to innovation in chemistry and electrical engineering (columns 1–2, Panel B), whereas master craftsmen do not significantly predict that type of innovation (columns 4–5, Panel B). The results are similar when using as dependent variable the share of people employed in chemistry in 1882 (columns 3 and 6, Panel B).[32] The horse race regressions in columns 7 and 8 (Panel B) confirm the notion that literacy had a positive impact on innovation and employment in technologically advanced industries.[33]

Note that we have not identified the direction of causality. In line with our explanation presented in Sect. 2, innovations might have flourished in counties with a well-educated population because entrepreneurs decided to locate innovative and high-tech firms in highly productive areas characterized by high levels of literacy. With regard to the decision where to establish a firm, the local availability of literacy might have been as important as the access to near-by natural resources such as water or coal. On the other hand, an above-average economic growth of counties with many firm innovations might have generated the tax revenues that were needed to finance better education at the first place.

In sum, the analysis of the relationship between different types of human capital and innovation lends support to the hypothesis that the Second Industrial Revolution was a transition period when it comes to the role of human capital. As in the First Industrial Revolution, "useful knowledge" embodied in skilled craftsmen was still an important determinant for innovation activities of private inventors in more traditional industries such as steam boil-

---

[29] Here, we do not estimate a zero-inflated negative binomial model since the variation in valuable patents after modeling the excess number of zeros is too low. In the case of the machine industry the share of counties with zero patents is 88%. In the case of chemistry and electrical engineering 94% of the counties have zero patents.

[30] In the horse race specification in column 8 of Table 4 the coefficient for literacy is even significantly negative. This result is driven by the four counties with the highest employment share in manufacturing. By dropping these four counties the coefficient for literacy becomes virtually zero (beta −0.055, SD 0.073) whereas the coefficient for master craftsmen remains large and highly significant (beta 0.206, SD 0.098).

[31] Note that in Panel B we exclude the county of *Elberfeld* which is an outlier because of the many valuable patents the chemical company Bayer held.

[32] Unfortunately, the occupational census in 1882 did not consider the category electrical engineering.

[33] To provide further support for our hypothesis that master craftsmen are only related to traditional independent innovation, we regress the share of self-employed people in manufacturing in 1882 on our different human capital variables. This exercise is based on the assumption that a high share of self-employed people is a good proxy for the dominance of small workshops whereas a low share of self-employed people indicates the existence of many large companies which were able to finance their own R&D department. The estimates are reported in Table A6 in the online Appendix. Again, the horse race specification provides a clear result: with respect to self-employment in manufacturing literacy has a significantly negative coefficient whereas the variable for master craftsmen has a significantly positively coefficient.

Table 4 Human capital and innovative activity in traditional (Panel A) and frontier technology (Panel B)

Panel A	Machine industry							
	Patents (1) NB	Patents (2) Poisson	Sh. employed (3) OLS	Patents (4) NB	Patents (5) Poisson	Sh. employed (6) OLS	Patents (7) Poisson	Sh. employed (8) OLS
Literacy rate	0.511	1.143	−0.102				0.613	−0.382***
	(0.410)	(0.701)	(0.078)				(0.693)	(0.138)
Master craftsmen (share)				0.401**	0.340***	0.452**	0.308****	0.591***
				(0.193)	(0.070)	(0.196)	(0.084)	(0.210)
Secondary school e.r.	0.092	0.175	−0.027	0.125	0.257	−0.015	0.187	0.020
	(0.333)	(0.225)	(0.113)	(0.361)	(0.175)	(0.101)	(0.190)	(0.092)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bank, Insurance, Law	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	334	334	334	334	334	334	334	334
Pseudo R-squared	0.29	0.55		0.30	0.57		0.58	
R-squared			0.34			0.43		0.47

Table 4 continued

Panel B	Chemistry and Electrical engineering industry							
	Patents (1) NB	Patents (2) Poisson	Sh. employed (3) OLS	Patents (4) NB	Patents (5) Poisson	Sh. employed (6) OLS	Patents (7) Poisson	Sh. employed (8) OLS
Literacy rate	1.072**	1.146**	0.018**				1.002*	0.020**
	(0.545)	(0.560)	(0.008)				(0.552)	(0.010)
Master craftsmen (share)				0.245	0.279	0.004	0.202	−0.004
				(0.289)	(0.176)	(0.008)	(0.187)	(0.009)
Secondary school e.r.	−0.844*	−0.507	−0.034*	−0.776	−0.321	−0.032*	−0.565	−0.034*
	(0.507)	(0.440)	(0.019)	(0.549)	(0.343)	(0.019)	(0.443)	(0.019)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bank, Insurance, Law	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	333	333	333	333	333	333	333	333
Pseudo R-squared	0.31	0.44		0.30	0.43		0.45	
R-squared			0.14			0.13		0.14

In the upper Panel A, the dependent variable in columns 1–2, 4–5 and 7 is the absolute number of valuable patents in the following technological classes: steam boilers, machine parts, and mechanical metal processing. In columns 3, 6 and 8 the dependent variable is the share of people employed in metal working in 1882. In the bottom Panel B, the dependent variable in columns 1–2, 4–5 and 7 is the absolute number of valuable patents in the following technological classes: electrical engineering, dyeing, chemical process. In column 3, 6 and 8 the dependent variable is the share of people employed in chemistry in 1882. In panel B, we exclude the county of Elberfeld as it is an outlier with 53 patents in chemistry by the chemical firm Bayer. Literacy rate, share of master craftsmen, and secondary school enrollment rates are standardized with zero mean and unitary standard deviation. Control variables included are: share of people working in manufacturing, log population, urbanization rate, share of large landownership, percent of Protestants, percent of Jewish, and distance to Berlin. Berlin is excluded from the regression. Controls for bank, insurance, and law include the share of people employed in banks and insurance, the share of people employed in legal occupations. Robust standard errors in parenthesis. Significance at * 10, ** 5, *** 1%

ers or mechanical metal processing. Simultaneously, however, basic formal education, here measured by literacy, became more and more important for firms' patenting activities in high-tech industries such as chemicals or electrical engineering.

5 Different types of human capital and income

After having identified the most important correlations between the different types of human capital and innovation, we now turn our attention to the impact of human capital on income. Before accounting explicitly for patenting activity, we test whether literacy rate, master craftsmen and formal secondary education had a *persistent* impact on tax revenues. We estimate the following model:

$$\text{Log}(Y_{1898-1903}) = \beta L_{1871} + \gamma Z_{1849} + \delta S_{1864} + \mathbf{X}'\lambda + \varepsilon \tag{1}$$

where Y is the mean tax revenues per capita for the period 1898–1903, L is the literacy rate in 1871; Z is the share of master craftsmen in 1849, and S is the enrollment rate in secondary schools in 1864. Due to collinearity between the literacy rate and the share of master craftsmen, however, we avoid horse race specifications and include literacy and master craftsmen only separately in the various specifications.[34] The vector X again includes log population size (1871), the urbanization rate (1871), the share of large landownership (1882), the shares of Protestants (1871), the share of Jews (1871), the share of people born in the municipality (1871), and a binary variable for the counties west of the river Elbe. Robust standard errors are clustered at the district (*Regierungsbezirk*) level to allow for correlation within larger administrative units. The OLS estimates are reported in Table 5.

Column 1 reports a strong bivariate positive correlation between log tax revenues and literacy rates. Column 2 shows a similar result for master craftsmen. The coefficients of the two types of human capital are also comparable in magnitude. In column 3 we include secondary school enrollment rates plus our set of controls and the coefficient for literacy remains highly significant and of similar size. Secondary education is positively related to tax revenues. As expected, population size and urbanization are also significantly positively related to tax revenues. Interestingly, once accounting for literacy, the coefficient of the share of Protestantism is negative. A higher percentage of people born in the respective municipality, that is a low immigration rate, is negatively associated with tax revenues. Finally, counties west of the river Elbe have higher levels of income tax revenues. The estimated coefficients can be interpreted as semi-elasticities: a change in the literacy rate by one standard deviation is associated with a 30 % increase in tax revenues. The same change in the secondary school enrollment rate leads to an increase of tax revenues by 12%. In column 4 we run the same specification but for the share of master craftsmen. As one can see, the coefficient is now less than one third in comparison to the unconditional regression (column 2) and significant only at the 10% level. Secondary education keeps being significantly positively associated with income tax revenues.

Bils and Klenow (2000) point to the fact that most empirical studies only document correlations between human capital and economic development but do not establish causal effects. We have already discussed the problem of reverse causality: Economically more advanced counties might have had both the need and the financial means to build up a more efficient education system. In this case, economic growth would have fostered human

---

[34] However, horse race specifications are run in the instrumental variable approach in the online Appendix.

Table 5 Human capital and log income tax revenues

	(1)	(2)	(3)	(4)	(5)	(6)
Literacy rate (std)	0.362***		0.303***		0.234***	
	(0.037)		(0.023)		(0.027)	
Master craftsmen		0.301***		0.089*		0.045
		(0.048)		(0.049)		(0.037)
Secondary school e. r.			0.124***	0.177***	0.161***	0.200***
			(0.023)	(0.031)	(0.026)	(0.032)
Log population			0.282***	0.341***	0.165***	0.132**
			(0.041)	(0.052)	(0.053)	(0.063)
Perc. urban population			0.009***	0.010***	0.005***	0.004*
			(0.001)	(0.002)	(0.001)	(0.002)
Share large landownership			0.187	0.144	0.265**	0.273**
			(0.153)	(0.213)	(0.117)	(0.127)
Perc. Protestants			−0.002**	−0.000	−0.002***	−0.001
			(0.001)	(0.001)	(0.001)	(0.001)
Perc. Jewish			−0.013	−0.105***	−0.013	−0.076***
			(0.021)	(0.026)	(0.020)	(0.023)
Perc. born in municipality			−0.013***	−0.004	−0.011***	−0.004**
			(0.002)	(0.003)	(0.002)	(0.002)
West of river Elbe			0.355***	0.414***	0.255***	0.238***
			(0.063)	(0.089)	(0.047)	(0.059)
Log income school teachers					1.307***	2.020***
					(0.239)	(0.260)
Log income tax revenues p.c.					0.031*	0.032
					(0.017)	(0.020)
Observations	335	335	335	335	324	324
R-squared	0.35	0.24	0.78	0.70	0.77	0.72

OLS estimates. The dependent variable is the log of mean income tax revenues per capita for the period 1899–1903. Literacy rate, share of master craftsmen, and secondary school enrollment rates are standardized with zero mean and unitary standard deviation. Robust standard errors clustered at the district (*Regierungsbezirk*) level. Significance at * 10, ** 5, *** 1%

capital formation and not the other way around.[35] We have information on income of male elementary school teachers for 1886 and past income-tax revenues for 1877. The inclusion of such variables, although not addressing the issue of reverse causality, allows us to estimate the impact of the human capital variables on the subsequent *growth* of income.[36] In column

---

[35] Another problem is that omitted variables like a county's formal and informal institutions, its predominant religious beliefs, or its industrial structure might have influenced both economic growth and human capital formation. Under these circumstances, the correlation between human capital and income would just indicate that both variables were determined by the same unobserved variable.

[36] The inclusion of these controls is not unproblematic since these variables are postdated with respect to literacy which we measure in 1871. Also note that, due to differences in the tax-base and in the number of individuals who were subject to taxation, the revenues of the older Prussian income tax system differ significantly from those of the more modern income tax, both in size and distribution. See Schremmer (1994, p. 146).

5 the inclusion of both variables reduces the size of the coefficient for literacy slightly.[37] As expected, both the income of school teachers and tax revenues in 1877 are positively related to our dependent variable. The inclusion of these variables in column 6 renders the coefficient for the share of master craftsmen smaller and insignificant. It is important to note that secondary education remains always significantly positively related to income tax revenues.

This set of results suggests that improvements in formal education which led to higher literacy and higher enrollment in secondary schooling indeed fostered a Prussian county's economic growth during the Second Industrial Revolution.

Becker and Woessmann (2009) reanimated the debate about the role of religion for European economic growth. The authors claim that, at least in Prussia, Protestantism had an effect on economic outcomes only through comparatively higher levels of primary schooling (measured by literacy) that resulted from Protestants' precept that everyone should be able to read the Bible.[38] To address the problem of reverse causality—Protestants might have preferred richer counties with better schooling—and the omitted variable bias, Becker and Woessmann (2009) employ an instrumental variable approach instrumenting Prussian counties' literacy rates with the distance to Wittenberg, the supposed center of German Protestantism. We cannot apply an analogous identification strategy in our study because it is verging on the impossible to find instruments for three different types of human capital that are all orthogonal to income tax revenues. What we can do is to test whether the results of Becker and Woessmann (2009) still hold when introducing master craftsmen and enrollment in secondary schooling as additional controls.

Note that we do not make any claim of causality with respect to the share of master craftsmen or secondary education. In the online Appendix we show that the share of master craftsmen and enrollment in secondary schools are not correlated with distance to Wittenberg (Figure A4). This suggests that distance to Wittenberg identifies exogenous variation only in literacy but not in the other two types of human capital we analyze. The results of the extended instrumental variable approach are reported in Tables A7 and A8 in the online Appendix. The first-stage F-statistic indicates the strength of distance to Wittenberg as an instrument for literacy. The instrumental variable estimates still suggest a statistically and economically significant effect of literacy on tax revenues. We find that the effect of literacy on tax revenues is also robust to the introduction of geographic controls such as soil quality, slope, temperatures and precipitation.[39] For additional robustness checks see the online Appendix.

6 Literacy, innovation, and income

One of the main objectives of this paper is to find out to what extent literacy and innovation affect income independently of each other. The left panel of Fig. 7 shows the relationship

[37] Income tax revenues in 1877 are not available for city counties and that explains why we only have 324 observations in columns 7 and 8.

[38] Cantoni (2015) uses the same identification strategy but analyzes the development of cities in a longer time period. He finds no effect of Protestantism on city size. See also Cantoni (2012).

[39] We constructed variables for the soil quality and the slope of the counties using raster data at a $10\,km \times 10\,km$ resolution from the European Soil Database (Panos et al. 2012). The variable for texture increases with the quality of the soil and ranges from zero to five according to the combination of sand and clay (coarse, medium, medium fine, fine, very fine). The variable for slope ranges from zero to four (level (0–8%), sloping (8–15%), moderately steep (15–25%), steep (over 25%)). Data on annual mean temperature and annual precipitation are representative for the period 1950–2000 and are obtained from WorldClim–Global Climate Data (www.worldclim.org).
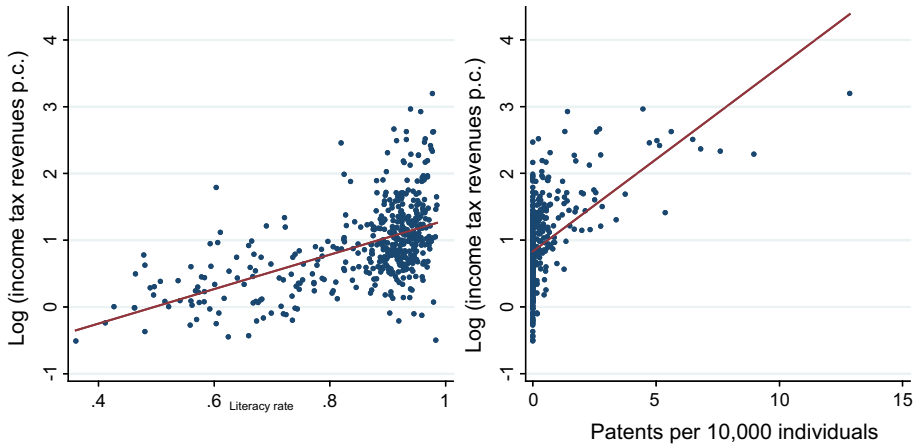
Fig. 7 The relationship of log income tax revenues p.c. with literacy rates and valuable patents *Source*: see text

between log income tax revenues per capita and literacy rate while the right panel presents the relationship between log income tax revenues and patents per 10,000 people. In both cases, a positive relationship is clearly visible.

In the first part of our empirical analysis we have shown that patenting activity in Prussian counties depended on literacy, in particular in the case of large firms and high-tech innovation. Because of this relationship and since we do not have an instrument for patenting activity that would need to be orthogonal to both literacy and income tax revenues, we can analyze this issue only descriptively by estimating the following OLS model:

$$\text{Log}(Y_{1898-1903}) = \beta L + \gamma P + \mathbf{X}'\lambda + \varepsilon \tag{2}$$

where the dependent variable is the log of the mean income tax revenues per capita for the period 1898–1903, L is the literacy rate, and P is the number of patents per 10,000 people in the period 1877–1890. The vector X includes our standard set of control variables. Note that we will also include enrollment rates in primary and secondary schools. To facilitate comparison, all human capital variables and the variable for patenting activity are standardized with zero mean and unit standard deviation. Standard errors are clustered at the district (*Regierungsbezirk*) level. Since in this section we are not interested in assessing the impact of master craftsmen which are observed at county borders as in 1849, we can revert to the county borders in place in 1871. That is why the number of observations increases to 452.

In the regression we include also geographic controls and "historical" controls to account for persistent differences in economic patterns at the county level. The geographic controls comprise variables for soil quality, slope, mean temperature and mean precipitation. The "historical" controls include dummies for having a university or a school in 1517, dummies for having a former Imperial or a Hanseatic city in 1500, respectively, and the level of city population in 1500.[40]

OLS estimates of Eq. (2) are presented in Table 6. In column 1 we regress log tax revenues on literacy and enrollment rates in primary schools in order to compare directly the effects of the *stock* and the *flow* of basic education. In addition, we control for secondary school enrollment. As one can see, literacy is still positively and significantly related to tax revenues

---

[40] For details and sources see the online Appendix.

Table 6 Human capital, patenting activity, and log income tax revenues

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Literacy rate	0.254***							0.251***	0.249***	0.249***
	(0.027)							(0.027)	(0.027)	(0.027)
2nd Quintile of literacy		0.415***								
		(0.073)								
3rd Quintile of literacy		0.481***								
		(0.078)								
4th Quintile of literacy		0.483***								
		(0.081)								
5th Quintile of literacy		0.666***								
		(0.084)								
Secondary school e.r. (1874)	0.054**	0.065**						0.053*	0.055*	0.050*
	(0.026)	(0.028)						(0.027)	(0.027)	(0.026)
Primary school e.r. (1886)	−0.030	−0.023						−0.031	−0.034	−0.030
	(0.046)	(0.045)						(0.049)	(0.047)	(0.047)
Patents p.c.			0.063***	0.039*				0.054***		
			(0.020)	(0.022)				(0.012)		
Patent dummy				0.193***						
				(0.037)						
Patents 1					0.127**					
					(0.051)					
Patents 2–4					0.262***					
					(0.057)					
Patents >5					0.258***					
					(0.054)					

Table 6 continued

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Firm patents p.c.						0.050**				
						(0.024)				
Independent patents p.c.							0.040**			
							(0.017)			
Number of tech. classes									0.018**	
									(0.007)	
Diversity of tech. classes										0.283***
										(0.057)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Historical controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	452	452	452	452	452	452	452	452	451	451
R-squared	0.72	0.71	0.66	0.67	0.67	0.66	0.65	0.72	0.72	0.72

OLS estimates. The dependent variable is the log of mean income tax revenues per capita for the period 1899–1903. Control variables are: log population, urbanization rate, share of large landownership, percent of Protestants, percent of Jewish, percent born in the relative municipality, west Elbe. The geographic controls include: soil quality, slope, mean temperature and mean precipitation. The historical controls include: dummies for having a university or a school in 1517, dummies for having a former Imperial city or a Hanseatic city in the county, city population in 1500. Robust standard errors clustered at the district (*Regierungsbezirk*) level. Significance at * 10, ** 5, *** 1%

while primary school enrollment rate has no significant impact. Enrollment in secondary schools has a significant positive coefficient which might be driven by reverse causality: The high demand of economically well-developed counties for skilled workers might have motivated investment in secondary schooling.

By including quintiles of the distribution of literacy, in column 2, we test for non-linearity in the relationship between literacy and tax revenues. The estimates reveal an interesting pattern: On the one hand, counties in the second, third, or fourth quintile of the distribution of literacy generated between 41 and 48% higher tax revenues than the counties in the bottom 20%. This suggests that Prussian counties needed to exceed a minimum threshold of literacy to prosper economically. On the other hand, counties at the top 20% of the distribution of literacy had significantly higher tax revenues than the preceding quintiles.[41] The interpretation of this result is consistent with our notion of a complementarity between technology and skills presented in Sect. 2, according to which innovative technologies with a comparatively high economic performance require high levels of human capital. Data on the occupational structure corroborate this interpretation. For example, the share of people employed in high-tech chemistry at the top quintile of the literacy distribution (0.13%) is significantly larger than the respective share employed in the fourth quintile (0.08%). On the contrary, mean employment in the traditional textile sector already peaked in the third quintile of the literacy distribution with 2.5% and decreased afterwards to 0.9% in the top quintile.

In column 3 we focus our attention on the relationship between innovation and log income tax revenues. The coefficient for valuable patents is positively and significantly related to log tax revenues. The magnitude of the coefficient, when compared to literacy, suggests a comparatively smaller impact of valuable patents on income: a standard deviation increase in patents per 10,000 people is associated with an increase in tax revenues by about 6%. In column 4 we analyze the relationship between innovation and income tax revenues both at the extensive and intensive margin. To have at least one valuable patent seems to be more relevant for income tax revenues than the absolute number of valuable patents.

In column 5 we further explore the possibility of non-linearity in the relationship between valuable patents and income tax revenues by including dummy variables for having 1, between 2 and 4, or 5 or more valuable patents, respectively. The results provide again some evidence for non-linearity. Income tax revenues in counties with five or more patents are not higher than those in counties with only a few patents (2–4).[42] In columns 6 and 7 we exploit our distinction between firms' and independent patents and estimate their separate relationship with income tax revenues. The coefficients of both types of patents are positive and significant and of similar size.

In columns 8 we run a horse race between different measures for human capital and innovation activities. The estimates suggest that literacy and valuable patents are independently of each other related to income tax revenues. Most importantly, the magnitude of the coefficients is almost identical in size to the coefficients in the regressions that consider both variables in isolation. This finding supports our hypothesis that the impact of human capital on growth involves multiple channels.

In columns 9 and 10 we provide further support to this interpretation by using alternative measures for the scope of innovation activities.[43] In most Prussian counties the inventors

---

[41] The coefficients for the second, third, and fourth quintiles are not statistically different from each other. The coefficient for the fifth quintile is instead statistically different from the other coefficients.

[42] However, having one valuable patent is associated with a 12% increase in income tax revenues in comparison to counties without any valuable patents.

[43] Given the high number of patents (464) and of technological classes (69) in Berlin, we exclude that observation from the analysis. The results are unaffected if we include Berlin.

concentrated on a very small variety of technological classes. In fact, the average number of different technological classes in a county is 1.5 (whereas the average quantity of valuable patents comes to 3.8). Column 9 shows that the number of technological classes is positively correlated with tax revenues. The coefficient for literacy (and secondary education) remains virtually unaffected. The data on technological classes can also be used to construct an index of technological fractionalization. This index measures the probability of drawing two different technological classes when randomly selecting two patents from a county's patent sample.[44] Column 10 reveals that this variable is also strongly positively related to tax revenues. Moving from a county with zero technological diversity to a county with a fully heterogeneous spectrum of technological classes is associated with an increase in tax revenues of 28%. This result holds when controlling for the number of patents per capita and for having at least one valuable patent.[45] Again, the coefficients for literacy and secondary education stay the same.

7 Human capital and the transition to modern economic growth

An important characteristic of the transition to modern economic growth is the rapid decline in fertility rates at the end of the nineteenth century. Theories about the so-called quantity-quality tradeoff argue that parents' increasing investment in their children's education during the Second Industrial Revolution was accompanied by decreasing levels of fertility (Galor 2011, 2012). The logic behind this hypothesis is that technological progress demanded higher levels of education which, in the presence of a household's binding budget constraint, were traded against the quantity of children.

We add to this literature by estimating the impact of literacy on fertility outcomes at different levels of technological progress. According to Unified Growth Theory, the negative effect of education on fertility should be more accentuated when the returns to education are higher (Galor 2011, p. 154). We can test this hypothesis by interacting the literacy rate with the number of patents that were held in a county. Table 3 had shown that literacy is more related to firms' innovation than to independent innovation. Therefore, it is reasonable to assume that the "returns to literacy" are higher in counties with many firms' innovations than in counties with many independent innovations. We test this assumption in Table 7.[46]

The fertility outcomes used in this analysis are the crude birth rates (CBR) in 1890 and 1900 and the marital fertility rates (MFR) in the same years.[47] The estimates show that the interaction between literacy and patents is only significant in the case of firms' innovation. Interestingly enough, firms' innovation is also positively correlated with the fertility outcomes whereas independent innovation shows no significant relationship with the CBR or the MFR. In sum, the estimates of Table 7 support the notion that the negative impact of literacy on fertility increases with technological progress that raises the returns to education. In general, we provide substantial evidence in favor of the hypothesis that literacy played a crucial and multifaceted role in Prussian's transition to modern economic growth.

---

[44] This index is equal to one minus the Herfindahl index and varies between zero and one. Counties with zero patents are assigned zero technological diversity.

[45] We did not include these results for the sake of space.

[46] In this case the variable for literacy rate is centered at the mean.

[47] The marital fertility rate is computed as the number of births over the number of married women aged 15–49.

Table 7 Literacy, patenting activity, and fertility outcomes

	Firm innovation				Independent innovation			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	CBR 1890	MFR 1890	CBR 1900	MFR 1900	CBR 1890	MFR 1890	CBR 1900	MFR 1900
Literacy rate (centered)	−0.023***	−0.194***	−0.024***	−0.257***	−0.024***	−0.206***	−0.025***	−0.267***
	(0.003)	(0.027)	(0.003)	(0.028)	(0.003)	(0.028)	(0.003)	(0.028)
Firm patents p.c.	0.004***	0.026**	0.004**	0.016*				
	(0.001)	(0.010)	(0.001)	(0.009)				
Literacy rate * firm patents p.c.	−0.043***	−0.294***	−0.038**	−0.195**				
	(0.015)	(0.094)	(0.015)	(0.095)				
Indep. patents p.c.					0.002*	0.005	0.002	0.001
					(0.001)	(0.003)	(0.001)	(0.004)
Literacy * indep. patents p.c.					−0.013	−0.042	−0.011	−0.013
					(0.010)	(0.036)	(0.010)	(0.037)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Historical controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	452	452	452	452	452	452	452	452
R-squared	0.56	0.52	0.48	0.56	0.56	0.51	0.48	0.55

OLS estimates. CBR and MFR indicate crude birth rate and marital fertility rate, respectively. Literacy rate is centered at the mean. Control variables are: log population, urbanization rate, share of large landownership, percent of Protestants, percent of Jewish, percent born in the relative municipality, west Elbe. The geographic controls include: soil quality, slope, mean temperature and mean precipitation. The historical controls include: dummies for having a university or a school in 1517, dummies for having a former Imperial city or a Hanseatic city in the county, city population in 1500. Robust standard errors clustered at the district (*Regierungsbezirk*) level. Significance at * 10, ** 5, *** 1%

8 Conclusion

We merged individual data on valuable patents granted in Prussia in the late nineteenth-century with county-level data on literacy, craftsmanship, secondary schooling, and income tax revenues to explore the complex relationship between different types of human capital, innovation, and income.

It is widely believed in economics that improving formal education of broad segments of the population is a necessary precondition for a country's transition from stagnation and backwardness to sustainable economic growth. In sharp contrast to this view, some scholars have claimed that average levels of literacy played no important role during the First Industrial Revolution in Britain and other European countries. They stress, instead, the importance of a few highly skilled craftsmen and entrepreneurs who were needed to transform useful knowledge into economically successful innovations. With this article we bridge the gap between these two perceptions by showing that the Second Industrial Revolution might have been a transition period when it comes to the role of human capital. As in the preceding First Industrial Revolution, "useful knowledge" embodied in master craftsmen was strongly related to innovation of independent inventors. But at the end of the nineteenth century the time of the famous independent inventor was mostly gone. The newly-founded industrial R&D departments took over the lead in developing more and more science-based innovations. At the same time, the quality of basic education became positively associated with both workers' productivity and firms' R&D processes.

To explain the surprising result that literacy was positively correlated with innovation activities we presented three different explanations that are not mutually exclusive. First, high literacy rates might indicate an above-average general investment in children's human capital that comprised also higher education, nutrition, and preventive health care. Second, a high literacy rate might have facilitated transregional knowledge spillover because reading was the only way to learn about ideas that have occurred outside the own locality. Third, high-tech innovations and formal basic education might have been complements in highly innovative industries.

Our empirical analysis also strongly suggests that literacy and patenting activities are independently correlated with income. This evidence lends support to the notion that basic education had an impact on Prussian growth through both the direct production channel and the indirect innovation channel. What is more, literacy paved the way to modern economic growth also by accelerating structural change and fertility decline. Regarding the latter, our results are consistent with the interpretation that technological advancements, by augmenting the returns to education, provided more incentives to invest in education and amplified the quantity-quality tradeoff.

References

Acemoglu, D., Aghion, P., & Zilibotti, F. (2006). Distance to Frontier, selection, and economic growth. *Journal of the European Economic Association*, *4*, 37–74.

Acemoglu, D., Cantoni, D., Johnson, S., & Robinson, J. A. (2011). The consequences of radical reform: The French revolution. *American Economic Review*, *101*, 3286–3307.

Acemoglu, D., Gallego, F. A., & Robinson, J. A. (2014). Institutions, human capital and development. *Annual Review of Economics*, *6*, 875–912.

Acemoglu, D., Johnson, S., & Robinson, J. A. (2001). The colonial origin of comparative development: An empirical investigation. *American Economic Review*, *91*, 1369–1401.

Allen, R. C. (2003). Progress and poverty in early modern Europe. *Economic History Review*, *56*, 403–443.

Allen, R. C. (2009). *The British Industrial Revolution in global perspective*. Cambridge: University Press.

Baten, J., & Crayen, D. (2010). Global trends in numeracy 1820–1949 and its implications for long-term growth. *Explorations in Economic History*, *47*, 82–99.

Baten, J., Crayen, D., & Voth, H.-J. (2014). Numeracy and the impact of high food prices in industrializing Britain, 1780–1850. *Review of Economics and Statistics*, *96*, 418–430.

Baten, J., Spadavecchia, A., Streb, J., & Yin, S. (2007). What made Southwest German Firms innovative around 1900? Assessing the importance of intra- and inter-industry externalities. *Oxford Economic Papers*, *59*, i105–i126.

Baten, J., & van Zanden, J. L. (2008). Book production and the onset of modern economic growth. *Journal of Economic Growth*, *13*, 217–235.

Becker, S. O., Cinnirella, F., Hornung, E., & Woessmann, L. (2014). iPEHD—The ifo Prussian Economic History Database. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, *47*, 57–66.

Becker, S. O., Cinnirella, F., & Woessmann, L. (2010). The trade-off between fertility and education: Evidence from before the demographic transition. *Journal of Economic Growth*, *15*, 177–204.

Becker, S. O., Cinnirella, F., & Woessmann, L. (2012). The effect of investment in children's education on fertility in 1816 Prussia. *Cliometrica*, *6*, 29–44.

Becker, S. O., Hornung, E., & Woessmann, L. (2011). Education and catch-up in the industrial revolution. *American Economic Journal: Macroeconomics*, *3*, 92–126.

Becker, S. O., & Woessmann, L. (2009). Was weber wrong? A human capital theory of protestant economic history. *Quarterly Journal of Economics*, *124*, 531–596.

Benhabib, J., & Spiegel, M. M. (1994). The role of human capital in economic development: Evidence from aggregate cross-country data. *Journal of Monetary Economics*, *34*, 143–173.

Benhabib, J., & Spiegel, M. M. (2005). Human capital and technology diffusion. In P. Aghion & S. N. Durlauf (Eds.), *Handbook of economic growth* (Vol. 1A, pp. 935–966). Amsterdam: North Holland.

Bils, M., & Klenow, P. (2000). Does schooling cause growth? *American Economic Review*, *90*, 1160–1183.

Botticini, M., & Eckstein, Z. (2005). Jewish occupational selection: Education, restrictions, or minorities? *Journal of Economic History*, *65*(4), 922–948.

Botticini, M., & Eckstein, Z. (2007). From farmers to merchants, conversions and diaspora: human capital and Jewish history. *Journal of the European Economic Association*, *5*, 885–926.

Broadberry, S., & Burhop, C. (2010). Real wages and labor productivity in Britain and Germany, 1871–1938: A unified approach to the international comparison of living standards. *Journal of Economic History*, *70*, 400–427.

Burhop, C., & Lübbers, T. (2010). Incentives and innovation? R&D management in Germany's chemical and Electrical Engineering Industries around 1900. *Explorations in Economic History*, *47*, 100–111.

Cantoni, D. (2012). Adopting a new religion: The case of protestantism in 16th century Germany. *Economic Journal*, *122*, 502–531.

Cantoni, D. (2015). The economic effects of the protestant reformation: Testing the weber hypothesis in the German lands. *Journal of the European Economic Association*, *13*, 561–598.

Cinnirella, F., & Hornung, E. (2016). Landownership concentration and the expansion of education. *Journal of Development Economics*, *121*, 135–152.

Cinnirella, F., & Schueler, R. (2016) The cost of decentralization: Linguistic polarization and the provision of education. *CEPR Discussion Paper* 11274.

Clark, G. (2005). The condition of the working class in England, 1209–2003. *Journal of Political Economy*, *113*, 1307–1340.

Clark, G. (2009). *A farewell to alms: A brief economic history of the world*. Princeton: University Press.

Crafts, N., & Wolf, N. (2014). The location of the UK cotton textiles industry in 1838: A quantitative analysis. *Journal of Economic History*, *74*, 1103–1139.

Dalgaard, C.-J., & Strulik, H. (2013). The history augmented solow model. *European Economic Review*, *63*, 134–149.

Degner, H. (2009). Schumpeterian firms before and after world war I: The innovative few and the non-innovative many. *Zeitschrift für Unternehmensgeschichte*, *54*, 50–72.

Franz, M. (1911). Moral und Technik bei der Veranlagung der preussischen Einkommensteuer. *Jahrbuch für Gesetzgebung, Verwaltung und Volkswirtschaft im Deutschen Reich*, *35*, 285–373.

Fernihough, A., & O'Rourke, K. H. (2014). Coal and the European Industrial Revolution. *NBER Working Paper No. 19802*.

Festschrift des Königlich Preussischen Statistischen Bureaus zur Jahrhundertfeier seines Bestehens (1905). Berlin: Verlag des Königlich Preussischen Statistischen Bureaus.

Galor, O. (2011). *Unified growth theory*. Princeton: University Press.

Galor, O. (2012). The demographic transition: Causes and consequences. *Cliometrica*, *6*, 1–28.

Galor, O., & Moav, O. (2002). Natural selection and the origin of economic growth. *Quarterly Journal of Economics*, *117*, 1133–1192.

Galor, O., Moav, O., & Vollrath, D. (2009). Inequality in landownership, the emergence of human-capital promoting institutions, and the great divergence. *Review of Economic Studies*, *76*, 143–179.

Gennaioli, N., La Porta, R., Lopez-de-Silanes, F., & Shleifer, A. (2013). Human capital and regional development. *Quarterly Journal of Economics*, *128*, 105–164.

Glaeser, E. L., Porta, R. L., de Silanes, F. L., & Shleifer, A. (2004). Do institutions cause growth? *Journal of Economic Growth*, *9*, 271–303.

Goldin, C., & Katz, L. F. (1998). The origins of technology-skill complementarity. *Quarterly Journal of Economics*, *113*, 693–732.

Griliches, Z. (1990). Patent statistics as economic indicators: A survey. *Journal of Economic Literature*, *33*, 1661–1707.

Grupp, H., Dominguez-Lacasa, I., & Friedrich-Nishio, M. (2002). *Das deutsche Innovationssystem seit der Reichsgründung*. Heidelberg: Springer.

Guinnane, T. W., Harris, R., Lamoreaux, N. R., & Rosenthal, J.-L. (2007). Putting the corporation in its place. *Enterprise and Society*, *8*, 687–729.

Gutberlet, T. (2014). Mechanization and the spatial distribution of industries in the German Empire, 1875 to 1907. *Economic History Review*, *67*, 463–491.

Hanushek, E. A., & Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of Economic Literature*, *46*, 607–668.

Jones, C. I., & Romer, P. M. (2010). The new kaldor facts: Ideas, institutions, population, and human capital. *American Economic Journal: Macroeconomics*, *2*, 224–245.

Kelly, M., Mokyr, J., & Gráda, C. (2015). Roots of the industrial revolution. *UCD Centre for Economic Research Working Paper Series* 15/24.

Königliches Statistisches Bureau (1874) *Die Gemeinden und Gutsbezirke des Preussischen Staates und ihre Bevölkerung: Nach den Urmaterialien der allgemeinen Volkszählung vom 1. Dezember 1871*. Berlin.

Labuske, K., & Streb, J. (2008). Technological creativity and cheap labour? Explaining the growing international competitiveness of German mechanical engineering before world war I. *German Economic Review*, *9*, 65–86.

Lehmann-Hasemeyer, S., & Streb, J. (2016). The Berlin stock exchange in imperial Germany: A market for new technology? *American Economic Review*, *106*, 3558–3576.

Lucas, R. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, *22*, 3–42.

Lucas, R. (2002). The industrial revolution: Past and future. In R. Lucas (Ed.), *Lectures on Economic Growth* (pp. 109–190). Cambridge: Harvard University Press.

Mankiw, G. N., Romer, D., & Weil, D. N. (1992). A contribution to the empirics of economic growth. *Quarterly Journal of Economics*, *107*, 407–437.

Meisenzahl, R., & Mokyr, J. (2011). The rate and direction of invention in the British industrial revolution: Incentives and institutions. *NBER Working Paper No. 16993*.

Mitch, D. (1993). The role of human capital in the first industrial revolution. In J. Mokyr (Ed.), *The British Industrial Revolution: An economic perspective* (pp. 267–307). Boulder: Westview.

Mitch, D. (1999). The role of education and skill in the British Industrial Revolution. In J. Mokyr (Ed.), *The British Industrial Revolution: An economic perspective* (2nd ed., pp. 241–279). Boulder: Westview.

Mokyr, J. (2005). The intellectual origin of modern economic growth. *Journal of Economic History*, *65*, 285–351.

Mokyr, J. (2009). *The Enlightened Economy*. New Haven: Yale University Press.

Mokyr, J., & Voth, H.-J. (2009). Understanding growth in Europe, 1700–1870: Theory and evidence. In S. Broadberry & K. O'Rourke (Eds.), *The Cambridge economic history of modern Europe volume 1: 1700–1870* (pp. 7–42). Cambridge: University Press.

Moser, P. (2011). Do patents weaken the localization of innovations? Evidence from World's fairs. *Journal of Economic History*, *71*, 363–382.

Moser, P. (2012). Innovation without Patents—Evidence from World's fairs. *Journal of Law and Economics*, *55*, 43–74.

Moser, P. (2013). Patents and Innovation: Evidence from economic history. *Journal of Economic Literature*, *27*, 23–44.

Mühlhoff, K. (2016). *Darwin beats Malthus: Medicalization, evolutionary anthropology and the demographic transition*. Unpublished Working Paper.

Murmann, J. P. (2003). *Knowledge and competitive advantage: The coevolution of firms, technology, and national institutions*. Cambridge: University Press.

Murphy, T. E. (2015). Old Habits Die Hard (Sometimes). Can Département Heterogeneity Tell Us Something about the French Fertility Decline? *Journal of Economic Growth*, *20*, 177–222.

Nelson, R. R., & Phelps, E. S. (1966). Investment in humans, technological diffusion, and economic growth. *American Economic Review*, *56*, 69–75.

O'Rourke, K. H., & Williamson, J. G. (1995). Education, globalization and catch-up: Scandinavia in the Swedish mirror. *Scandinavian Economic History Review*, *43*, 287–309.

Panos, P., van Liedekerke, M., Jones, A., & Montanarella, L. (2012). European soil data centre: Response to European policy support and public data requirements. *Land Use Policy*, *29*, 329–338.

Pierenkemper, T., & Tilly, R. H. (2004). *The German economy during the nineteenth century*. New York and Oxford: Berghahn.

Prados de la Escosura, L. (2015). World Human Development: 1870–2007. *Review of Income and Wealth*, *61*, 220–247.

Preußisches Statistisches Landesamt (2008) M*arkt- und Kleinpreise ausgewählter Güter in Preußen 1816 bis 1928*. GESIS Datenarchiv, Cologne, histat. Studiennummer 8302, Datenfile Version 1.0.0.

Richter, R., & Streb, J. (2011). Catching-up and falling behind: Knowledge spillover from American to German machine tool makers. *Journal of Economic History*, *71*, 1006–1031.

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, *98*, S71–S102.

Schankerman, M., & Pakes, A. (1986). Estimates of the value of patent rights in European Countries during the Post-1950 period. *Economic Journal*, *96*, 1052–1076.

Schremmer, E. (1994). *Steuern und Staatsfinanzen während der Industrialisierung Europas*. Berlin: Springer.

Seckelmann, M. (2006). *Industrialisierung, Internationalisierung und Patentrecht im Deutschen Reich, 1871–1914*. Frankfurt/Main: Klostermann.

Semrad, A. (2015). Modern secondary education and economic performance: The introduction of the Gewerbeschule and Realschule in nineteenth-century Bavaria. *Economic History Review*, *68*, 1306–1338.

Sokoloff, K. L. (1988). Inventive activity in early industrial America: Evidence from patent records, 1790–1846. *Journal of Economic History*, *48*, 813–850.

Solow, R. M. (1956). A contribution to the theory of economic growth. *Quarterly Journal of Economics*, *70*, 65–94.

Streb, J. (2016). The cliometric study of innovation. In C. Diebolt & M. Haupert (Eds.), *Handbook of cliometrics* (pp. 447–468). Heidelberg: Springer.

Streb, J., Baten, J., & Yin, S. (2006). Technological and geographical knowledge spillover in the German Empire 1877–1918. *Economic History Review*, *59*, 347–373.

Streb, J., Wallusch, J., & Yin, S. (2007). Knowledge spill-over from new to old industries: The case of German synthetic dyes and textiles 1878–1913. *Explorations in Economic History*, *44*, 203–223.

Squicciarini, M. P., & Voigtländer, N. (2015). Human capital and industrialization: Evidence from the age of enlightenment. *Quarterly Journal of Economics*, *130*, 1825–1883.

Taylor, A. M. (1999). Sources of convergence in the late nineteenth century. *European Economic Review*, *43*, 1621–1645.

Vandenbussche, J., Aghion, P., & Meghir, C. (2006). Growth, distance to Frontier and composition of human capital. *Journal of Economic Growth*, *11*, 97–127.

Van der Beek, K. (2014). England's eighteenth-century demand for high-quality workmanship: Evidence from apprenticeship, 1710–1770. In: A. Greif, L. Kiesling & J. Nye (Eds.), *Institutions, innovation, and industrialization: Essays in economic history and development, a Festschrift Volume in the Honor of Prof. Joel Mokyr.* (pp. 268–274) Princeton: University Press.

Van der Beek, K., & Feldman, N. (2016). Skill choice and skill complementarity in eighteenth century England: 1710–1770. *Explorations in Economic History*, *59*, 94–113.

von Hippel, W. (2002) Auf dem Weg zum Weltunternehmen (1865–1900). In: W. Abelshauser (ed.): *Die BASF. Eine Unternehmensgeschichte*. Munich: Beck, pp. 17–116.

Wiese, L. A. (1874). *Das höhere Schulwesen in Preussen* (vol. 3, pp 1869–1873). Berlin: Wiegandt Grieben.