Manuscript Number: JOEG271R1

Title: Book Production and the Onset of Modern Economic Growth

Article Type: Manuscript

Keywords: Book Production, Economic Growth, Human Capital

Corresponding Author: Prof. Dr. Joerg Baten, Dr.

Corresponding Author's Institution:

First Author: Joerg Baten, Dr.

Order of Authors: Joerg Baten; Joerg Baten, Dr.
Abstract

Endogenous growth theory suggests that human capital formation plays a significant role for the ‘wealth and poverty of nations.’ In contrast to some previous studies which denied the role of human capital as a crucial determinant of long-term growth, we confirm its importance. Indicators of human capital like literacy rates are lacking for the period of 1450-1913; hence, we use per capita book production as a proxy for advanced literacy skills. This study explains how, and to what extent, growth disparities are a function of human capital formation.

JEL: O14, O40, N10

Keywords: Book Production, Economic Growth, Human Capital
1. Introduction

There is considerable debate about the causes of welfare growth before the Industrial Revolution, and particularly about the role of human capital formation in that process (Acemoglu, Johnson, Robinson 2001, Cervellati and Sunde 2005, Galor and Weil 2000). One approach has focused on the demographic regime – i.e. a trade-off between the quantity and quality of children –, arguing that the Industrial Revolution must have been preceded by a switch towards a regime characterized by high human capital formation (Becker, Tamura and Murphy 1990; Lucas 2002). However, economic historians have not found evidence for such a demographic regime change in the centuries before 1800. Usually, the ‘demographic transition’ which inspired these theories is dated in the period after 1870, about one century after the onset of the Industrial Revolution in Great Britain (Mokyr 1990, Clark 2003; but see De Moor and Van Zanden (2005) for an alternative interpretation). In reaction to this, Galor and Weil (2002), Galor and Moav (2002), and Galor (2005) have modeled the initial growth spurt as driven by the positive effects of population growth and population density only, with demographically induced human capital formation having played a major role not before the second Industrial Revolution of the late 19th century. But also the evidence for a rise of human capital formation in the centuries before 1800 is disputed by economic historians, pointing out, for example, that levels of literacy stagnated during the Industrial Revolution (Mokyr 2002, Allen 2003, but see Boucekkine, de la Croix, and Peeters 2007a and 2007b discussed below).

These studies suffer from the problem that human capital formation is difficult to quantify, in particular for the pre-modern period. We propose that book production is better suited to measure a number of aspects of human capital formation than traditional measures such as literacy, i.e. the ability to sign a contract or register. This study is the first to present a time variant proxy for human capital as far back as the late Middle Ages for eight European countries. Moreover, little was previously known about the human capital of countries such as
India, China, Indonesia, and Japan, for which we present cross-sectional evidence in the last section of this paper. Our primary hypothesis is that human capital, as measured by our indicator of book production, can explain differences in economic growth over the period under study.

The evolution of movable printing sharply reduced the access costs to scientific knowledge, thereby raising the productivity and profitability of scientific activity dramatically. This, in turn, strongly stimulated the scientific branch of the economy. Qualitative interpretations of the long-term effects of the invention of printing have put forward that both the Reformation of the 16th century and, perhaps more importantly, the Scientific Revolution of the 17th century would not have occurred without the printing press which made it possible to distribute the heterodox ideas of reformers and scholars in spite of manifold attempts to suppress them (Eisenstein 1979). The increasing production of useful knowledge which resulted from these developments laid the basis for the ‘Industrial Enlightenment,’ which according to Mokyr (2002) was one of the main causes behind the Industrial Revolution of the 18th century. Moreover, it motivated processes of literacy creation, intensified reading activity (which might not be perfectly correlated with basic literacy skills), and a relatively high willingness to pay for books – at least in some countries.

We argue that it matters not only whether people produce ideas (as it is assumed in the Kremer 1993 model), but that ideas also need to be stored and be communicated to others. Since books used to be the single major medium for the storage and communication of knowledge, the production of books can be viewed as a measure of the production and accumulation of accessible knowledge over time. As Gutenberg’s innovation shows, technological developments can change the production, storage, and communication of knowledge, making it more profitable to devote time to the production of knowledge because (a) access costs to existing knowledge are lowered (since all ideas are to some extent combinations and variations of existing ones, access to the existing stock of ideas is a
prerequisite for the production of new ideas); and (b) the income which can be expected from producing new ideas rises (since these ideas can be better communicated to others than before). Cervelatti and Sunde (2005) have argued that these factors set in motion a cumulative process of knowledge growth and human capital formation characteristic of ’modern economic growth.’ Indeed, this is observable for Europe after 1454: there was, for example, a growing group of independent intellectuals and scholars who lived solely on their research and publications (Burke 2000).

The advantage of book production as a measure of human capital is that it is a truly economic variable which is closely linked to the demand for books in any given society and period. Hence, it reflects the degree to which a society has become literate and in which mass demand for books has emerged, which makes this index easier to interpret than the usual measures of literacy, which often define literacy as the mere ability to sign one’s name under a contract (such as a marriage contract). Van Zanden (2004) has found a high correlation of book production with other proxies of human capital, such as skill premium, or literacy as defined in the traditional way.

Of course, we are not the first to address the early human capital development. Boucekkine, de la Croix, and Peeters (2007a) have studied the microfoundations of early human capital creation in England and France (BCL from here). Following Cressy (1980), who found an almost constant literacy increase in England since 1540, BCL studied two booms of school foundation – in the late 16th and early 18th century – and compared the improvements in France between the late 17th and 18th centuries with the English development (on France, see Fleury and Valmary, 1957; more recently Houdaille, 1977). The perhaps most important contribution of BCL was to suggest a model with clear microfoundations how population density translated into early schooling achievements, based on the consideration that school children could not travel over longer distances at this time. The parents who expected higher income (or religious capabilities, one could add) of their children acted
rational in their decisions about school attendance and the best location of schools. The authors argued that this model provided microfoundations to the mechanism which the unified growth theory suggested for the switch from Malthusian stagnation to the First Industrial Revolution. They quantified the contribution of population density as being one third, life expectancy one sixth, and TFP growth one half, to literacy growth (Boucekkine, de la Croix, and Peeters 2007b). Below, we will compare the predictions of their model of human capital formation with the new evidence presented here.

Our paper is structured as follows. The newly created data set and its sources are presented in section 2. There we also present our measure of economic growth, the growth of real wages – a measure that extends from the 15th century to the 19th century for a number of countries (Allen 2001, 2003). In the third section we conduct a regression analysis where we estimate the impact of human capital on growth while controlling for a number of other variables. In particular, we allow for two competing explanations of early economic growth: intercontinental trade and growth enhancing institutions. In the fourth section, we conduct an out-of-sample test to see if the model from section 3 helps explain the early portion of the Great Divergence over the 19th century. The paper ends with summary of our findings.

2. Data: Book Production and Growth

Our estimates of the output of printed books are based on the number of titles or editions which appeared in Western Europe between 1454 and 1800. For the details about the data set, there is a longer working paper version (Baten and van Zanden 2006), where we also discuss the questions on book imports and exports, production and consumption. Our individual observation units, “titles “ include both first editions and re-editions, the latter being publications distinguished from previous editions by changes made in the contents (revised edition) or layout (new edition). To illustrate this, the first printing of Gutenberg’s bible is considered an individual title, as is any new edition of that bible, whereas a reprint of
exactly the same manuscript would not count as a title as such. The concept of edition and re-
edition is important for our study, since it partially solves the problem of weighing books by
importance. Books that were considered particularly important for the development of social
life (such as religious texts), for the functioning or improvement of institutions, or the transfer
of essential scientific knowledge were re-edited, sometimes again and again. Hence, we
measure not only the quantity of books, but also their implicit importance as indicated by their
re-editions.

We first discuss the general tendencies observable from the raw data. For the 15th
century, Belgium and Italy have the highest levels of book production per capita, overtaken at
around 1600 by the Netherlands and Great Britain (Figure 1). This is an important result in
itself, because literacy values in England during the Industrial Revolution were previously
estimated as relatively modest and stagnant (Mitch 1993). However, such a pessimistic view
of human capital formation during the 1750-1850 period tends to overlook the strong increase
which occurred in the centuries before 1750, when Great Britain became one of the most
literate countries of Europe. The strong growth of human capital formation in the Low
Countries and England is consistent with what we know from other studies (Hoeppner Moran
1985; Stephens 1990; Reis 2005). Thus, the higher number of books per capita could have
created substantially greater growth capabilities. Germany, the country where Gutenberg
invented the moveable type, belonged to the middle group as it had clearly less books per
capita than Great Britain or the Netherlands. Belgium displayed an interesting development,
starting out from a high level but falling off heavily in relative terms in the 18th century. This
is consistent with Vandenbroeke’s argument (1985) that the level of literacy may have
declined in Belgium in the late 18th and the first half of the 19th century. Spain started at
similarly low values as Great Britain in the 16th century, and arrived at the lowest overall
value in 1750-99. The biggest growth success was Sweden, rising from the lowest value in
1450-1599 to a position on par with Great Britain in the late 18th century. The Swedish and
Dutch gained enormously during that period, relative to the Germans, for example, who
destroyed their economy in religious wars.

Previous research charting the long term development of human capital formation
before 1800 has almost exclusively focused on the rise of literacy in this period. For a few
countries (England since the 15th, France since the 17th century) more or less consistent
estimates are available of the share of the population able to sign an act or register (see the
overview by Reis, 2005; on England: Stephens, 1990; Cressy, 1980). It is not always clear,
however, what the ability to sign one's name means – it may or may not imply being able to
read and write; as a first example, in Sweden, already at the end of the 18th century almost all
people signed their marriage register, but a large part of them was unable to read. Reis (2005)
estimates the level of overall literacy in Sweden at only 20-25%, whereas reading ability is
estimated as being close to 100% (Graff 1987, p. 226). Instead, we propose to use estimates
of the production and/or consumption of books as a measure of human capital. The
advantages of book consumption are that it measures the size of the market for books and
therefore represents actual use of reading skills, and is closely related to more advanced levels
of human capital formation beyond literacy. Moreover, it can be estimated in a consistent way
for a relatively large number of countries spanning a period of four and a half centuries.

How does our book indicator correlate with literacy rates? Available evidence for
England, France and Spain suggests a general correspondence (Figure 2). For England and
France, we found a relatively close correlation, but for Spain between 1500-49 and 1600-49,
only literacy increases rapidly, whereas our book indicator does not grow very much. Nalle
(1989) reports that literacy of ordinary peasants and craftsmen accused by the inquisition was
as high as in England. Even more convincing, she reports other samples – tax registers from
Santiago and wills from Madrid – which support this surprisingly high Spanish literacy. We would agree that in Spain basic literacy was growing, but not the more advanced components of human capital that we measure with our indicator. To sum up, the main trends known from the study of literacy are present in our estimates of book production, but there are also crucial differences, such as the Spanish case.

The model of Boucekkine, de la Croix, and Peeters (2007) mentioned above actually harmonizes quite nicely with the fact that Central, Eastern and Northern Europe were too sparsely developed for early school development. In contrast, the Netherlands developed the remarkable book per capita values due to their high population density. Moreover, the BCL model can explain why Spain had quite good literacy values and a relatively large number of schools in the late 16th and early 17th century. Spain had a peculiarly high urbanization rate, which stems from the fact that many farmers lived within city boundaries (Alvarez and Prados 2007). Hence the school commuting time could have been low for the Spanish school children, and literacy grew quickly over the 16th century. However, our human capital estimates deviate from those literacy estimates. In terms of book production per capita, Spain was among the lowest during the 16th and early 17th centuries. It might have been that while basic literacy was available, more advanced forms of human capital were lacking, which were also required for welfare growth – clearly Spain had one of the most disappointing real wage development records, jointly with Italy, during this period. This measurement of more advanced literacy might be a key contribution that our study provides to the literature. It might also be the reason why using simple signature counts as human capital proxy does not provide explanatory power for real wage growth in other studies (for example, in Allen 2003).

[Figure 2 and 3 around here]
Real wages are taken from the internet data archive set up and documented by Allen (2003). As exhibited in Figure 3, in most countries real wages declined during this period of rapid population growth. This stands in a certain contrast to Maddison’s estimates of GDP growth, which suggests that inequality was growing enormously (Hoffman et al. 2002). Only a small portion of the GDP-real wage divergence can be explained by the fact that population moved from the low real wage to high wage regions (such as from rural England to London). Another part of the divergence, some have argued, is caused by the fact that GDP measures the productivity in the modern and urban industrial and commercial sectors, rather than welfare of the whole population (among others, see Koepke and Baten 2005).

The Malthusian downward pressure (due to declining marginal agricultural product) kept real wages down, but some economies fared better in this situation, and we would argue that those countries had more advanced human capital. Only Great Britain and the Netherlands were able to recover to their late medieval level in the early 18th century, although they lost some ground again thereafter. The heaviest declines occurred in Spain and Italy. France and Germany formed a middle group with relatively modest secular trends, although Germany experienced a strong decline during the Thirty Years’ War but recovered again later. The Netherlands suffered during the struggle for independence from Spanish rule, but recovered already in the early 17th century. The net effect of these changes was a process of unconditional divergence, with the coefficient of variation almost doubling between the 15th and the 18th centuries.

Table 1 shows the descriptive statistics of the explanatory variables used in the following section. All of them are in logs, except for the dummy variables and intercontinental trade, which follows the original specification chosen in Allen (2003), partly because most values are zeros which would disappear in logarithmic form.
3. Regression analyses

We first run two fixed effects regressions, including dummies for all centuries under study (with the 17th century serving as the reference category), of the econometric form:

\[
\left( \ln(r_w_{it}) - \ln(r_w_{it-1}) \right) = \\
\beta_1 \ln(b_0_{it-1}) + \beta_2 \ln(r_w_{it-1}) + \beta_3 \text{tr}_{it-1} + \beta_4 \text{cap}_{it-1} + \sum_{k=1}^{p} \beta_k X_{kit-1} + \sum_{n=1}^{q} \beta_n D_{nit-1} + \mu_i + u_{it}
\]

where \( \left( \ln(r_w_{it}) - \ln(r_w_{it-1}) \right) \) is the increase of real wages in country i between the initial half of the century \( t-1 \) and the current half of the century \( t \); \( \ln(b_0_{it-1}) \) equals Log book production (the human capital proxy) in the initial half of the century, with \( \ln(r_w_{it-1}) \) indicating the level of real wages in the initial half of the century in log form, and \( \text{tr} \) and \( \text{cap} \) serving as our ‘alternative/complementary hypothesis’ variables for intercontinental trade and capital-protecting institutional constraints on the executive. \( X \) is a vector of \( p \) control variables such as land per capita, the Thirty Years’ War, and similar variables. \( D \) is a set of \( q \) century dummies which control for technological change in book production, the time variant impact of the growth impact per book between centuries, relative price changes, and similar factors. Finally, \( \mu_i \) are country fixed effects to control for unobservable country characteristics, and \( u_{it} \) is a stochastic error term.

Since the dependent variable is the real wage growth between the first and the second half of the century, it is unlikely that unit root problems arise (panel unit root tests with such small samples have very limited power). As all explanatory variables refer to the initial half of the century, we avoid contemporaneous correlation problems. Please note that we follow the standard procedure developed by Barro (1991, 1999, 2003) and many others, who regress growth rates on a set of “growth capabilities” measured in levels. The level of human capital
is precisely such a “growth capability”, since theory suggests that after controlling for the initial welfare level (which might also proxy a country’s capital stock, as Barro (1991) has argued), only countries with high human capital can achieve successive welfare growth.

Having discussed the international trade of books in the working paper version, we decided to test both the per capita production and consumption of books as indicators of human capital (Table 2). This dual approach allows us to assess the robustness to changes in concept. We find that the influence on real wage growth was statistically significant and substantial, even after controlling for a number of other factors. For example, the effect of an additional standard deviation of book production per capita is $1.3 \times 0.33 = 0.43$. This is quite a large value, given that the standard deviation of real wage growth is only 0.76. Hence, human capital as measured in books per capita had indeed a strong, positive, and economically significant impact on welfare growth.

Our proxy variable for intercontinental trade which encompasses Atlantic imports and exports per population is insignificant once we control for human capital, conditional convergence effects, and other variables. Moreover, this result is quite robust across various models (see also the regressions below).

The constraints on the executive powers to expropriate capital as defined in Acemoglu Johnson, Robinson (2002) turns out as an important determinant of growth. However, the positive influence of this variable does not change the other results significantly (in spite of some collinearity with our book production indicator, as indicated by the correlation coefficient of 0.56). Hence, human capital as proxied by book production might have been an additional factor in the facilitation of welfare growth, independent of the institutional factors (which were also important). We can therefore confirm our main hypothesis, and provide evidence for the hypothesis that institutions protecting capital mattered, whereas we find no confirmation for intercontinental trade to increase welfare levels.
The level of initial real wages is significantly negative, pointing to the existence of conditional convergence in early modern Europe. We would not put too much emphasis on this finding, given that fixed effects are somewhat biased towards indicating convergence effects (Durlauf et al. 2005). But to a certain extent, this might imply that technology transfer to initially less sophisticated countries (such as Sweden) took place.

Our control variables have mostly the expected sign, such as land per agricultural worker. Particularly this “Malthusian” variable has a remarkably high coefficient, and is statistically significant at the 10% level. Hence, Malthus was right when he argued that land was a limiting factor for early modern welfare growth, even if he did not take into account the positive counter-forces of human capital accumulation which had reached a critical level during the Industrial Revolution and in the end served to break up the premodern, Malthusian world, as will be shown in the following section.

Somewhat related to population size, we also tested the effects of the occurrence of the bubonic plague after 1450, which was more or less endemic in those countries until the late 17th century. The outbreaks of the bubonic plague were quite exogeneous, as opposed to the occurrence of other epidemic diseases – those were to a larger extent influenced by nutritional conditions, and hence by welfare levels. The bubonic plague is of special interest, because it might have disappeared earlier in countries with higher human capital such as England, whereas in France there were outbreaks until the early 18th century. The expected impact of plague on real wage growth might have a positive sign, as the reduction of labor force mechanically increases wages, and land and physical capital were less destroyed. On the other hand, one could imagine a negative correlation with human capital formation (and via this channel on wages), as lower life expectancies and greater uncertainty about survival might reduce the willingness to invest in learning. These effects with opposite signs might cancel each other out, as we find empirically only insignificant and small coefficients for the plague dummy (which is one if there was plague in a country and half century). Another explanation
of the insignificance might be that the quantitative effect of the plague and other epidemic
diseases was already reflected in our variable ‘land per agricultural worker’, which is closely
related to population losses or increase.

[Table 2 around here]

Among the control variables, we add a dummy variable for Germany during the Thirty
Years’ War for obvious reasons, and another for the post-war period (omitting them does not
change the other results). Hence, we are able to quantify the effect of the Thirty Year’s War
on Germany while controlling for the other relevant factors. Real wage growth was
substantially lower in the early 17th century, although this result is only economically, but not
statistically significant. The recovery effect after the war was even stronger: in the late 17th
century, i.e. after the peace treaty of 1648, conditional real wage growth increased by more
than one percent.

These results are very robust to various specifications (not shown). We also tested the
random effects model, but rejected it on the basis of the Hausman test. Nevertheless, even in
the random effects specification, books per capita have a significant impact on real wage
growth.

Kremer (1993) has argued that initial population size (or initial land area) might
impact on technological development in the long run, in particular that technology increases
more than proportionally with population size. He argues that the much larger land mass of
Eurasia led to its superiority over the Americas, and those over Australia, Tasmania and
Flinders island, respectively. As a proxy for technological development he uses the
availability of single technologies, such as “did those economies already have stone tools”,
“were they engaged in metallurgy” and so on, and population growth.\textsuperscript{6} We cannot
satisfactorily test his argument here, as the European countries were much more integrated
than Eurasia, the Americas, and Tasmania, which were divided by oceans (and his argument refers to the development over millennia). However, we performed the experiment to include (a) initial population size and (b) the agricultural land area in our regressions, in order not to neglect his influential reasoning. However, those variables turned out insignificant, and their inclusion did not change the significance of our human capital proxy (results available from authors).

[Table 3 around here]

We also need to address the issue of endogeneity. After all, books are a normal good, i.e. consumed more heavily with rising income. One strategy here which was already applied above is to take human capital as levels, and subsequent welfare growth in differences, in order to avoid contemporaneous correlation (and to measure growth capabilities). Moreover, we perform instrumental variable estimation (Table 3), with two different instruments. The lag in book production is our first instrument (corresponding to t-2 in the econometric model above, i.e. book production almost one century before the wage development to be explained). As a second instrument, we include Allen’s (2003) compilation of literacy rates around 1500 and 1800, and linear interpolations between those dates. How justified are those instruments? Good instrumental variables should be correlated with the potentially endogenous variable while not influencing the ultimate dependent variable, except via the potentially endogenous variable. The former is clearly true for both instruments, since both are correlated with initial books per capita. For example, in 1750-1799, for which Allen presents relatively reliable literacy data, the correlation between those variables is 0.80 (p-value 0.03), for all centuries the correlation is 0.73 (p-value 0.00). The latter is also the case here: since we regard our book variable as a proxy for advanced literacy (plus some storage of/access to knowledge functions), the effect of the interpolated literacy values should go through the same causal
channel. The same causality chain further applies to Allen’s literacy rates, hence we conclude that both instruments are justified. In order to preclude that our instruments are correlated with the error term, we performed the Sargan test for over-identifying restrictions in a regression with country dummies, and found that with a value of 0.11 (p-value 0.74), the Sargan test indicates that our instruments are valid.

4. Out-of-sample test: can 18\textsuperscript{th} century book production explain GDP growth between 1820 and 1913, as well as the Great Divergence?

In order to test whether book production can account for economic growth in the 19\textsuperscript{th} century, we perform a partial out-of-sample test of the relation between books per capita and GDP growth, using an enlarged data set which includes a number of non-European countries. Apart from this out-of-sample test, it is also interesting to assess whether the production of knowledge as proxied by book production can explain the ‘Great Divergence,’ i.e. the fact that Western Europe and North America grew rapidly in the 19\textsuperscript{th} century whereas (previously) highly developed regions such as China and India lagged behind (Pomeranz 2000). In other words, can we predict 19\textsuperscript{th} century growth using 18\textsuperscript{th} century book production estimates? In order to test this idea, the dataset is extended to include a number of countries for which data are available on the second half of the 18\textsuperscript{th} century – some of them European (Ireland, Switzerland, Poland, Russia), others non-European (United States, China, Indonesia, Japan and India).\textsuperscript{7} For the period of 1820-1913, relatively reliable GDP estimates are available.

Figure 4 shows the relationship between book production in the second half of the 18\textsuperscript{th} century, and the growth of GDP per capita over the 19\textsuperscript{th} century (the period of 1820-1913), according to Maddison’s (2001) estimates.

[Figure 4 around here]
The relationship is quite strong. Per capita book production in Asia was close to zero in the cases of India and Indonesia, and at about 3 per million inhabitants in China. In Japan, books per capita, at almost 7 per million inhabitants, were more than twice as high as in Russia, but still much lower than in Western Europe. Variation within Europe was also large, with the highest levels being attained by the Netherlands (538), Sweden (219), and Great Britain (198); not far behind came the U.S. with 141 books per million inhabitants per year. However, this large variation within Europe does not affect the overall results much. Only China is somewhat of an outlier in Figure 4, but has no greater influence on the relationship between the two variables. Both figures show that a close relationship exists between our measure of human capital formation for the 18th century, and economic performance in the period of the Industrial Revolution and its aftermath. Countries with a low level of human capital formation were unable to participate in the industrialization process which transformed the world economy, whereas countries with a better starting position managed to catch up with Great Britain – or even to overtake it. It is interesting to note that Japan invested heavily in schooling as early as the 18th century, which is evident from the existence of a mass market for books. The high level of schooling in pre-Meiji Japan is also confirmed by other evidence (Hayami and Kitô 1999; see also Van Leeuwen 2007) for the large difference between Japan on the one hand, and Indonesia and India on the other hand). Apparently, Japan’s high level of education alone was sufficient to make it a successful modernizer, whereas other Asian countries failed to industrialize in the 19th century.

5. Conclusion

In this study, we employed the number of books per capita in pre-industrial Europe as a proxy of advanced literacy. It was exactly in the countries in which book production increased fastest that real wages developed systematically faster over the centuries before the Industrial Revolution than in countries with lagging human capital formation. We performed a number
of tests to counter-check the validity of those results, such as robustness tests, instrumental variable estimations, and controlling for additional variables, but the results remained robust. Therefore, we may conclude that human capital formation as measured in this way had a strong and positive effect on economic performance in the centuries before 1800.

We also assessed two alternative hypotheses regarding the role of institutions and international (and in particular trans-Atlantic) trade in the growth process. Institutional patterns had a positive and independent effect, whereas an effect of intercontinental trade could not be established.

Finally, we assessed the movement of human capital formation and economic growth beyond 1800, until 1913, considering a larger sample of countries. Again, the number of book editions per capita allowed a forecast of countries’ subsequent growth capabilities in the century of the Great Divergence. Countries with high levels of human capital formation in the 18th century initiated or participated in the industrialization process of the 19th century, whereas countries with low levels of human capital formation were unable to do so, among them many of today’s Less Developed Countries such as India, Indonesia, and China.
References


Harvard University Press.

Literacy and Laicization in pre-Reformation York Diocese. Princeton, New Jersey:
Princeton University Press.


Kremer, M.J. (1993). ‘Population Growth and Technological Change: 1,000,000 B.C. to

Leeuwen, B. van (2007). Human capital and economic growth in Asia. An historical and
quantitative analysis of India, Indonesia and Japan, 1890-2000. Manuscript Ph D


Cambridge: Cambridge University Press.


Productivity in European Agriculture,’ Economic History Review XLIV, 215-239.

Zanden, J.L. van (2004). ‘Common workmen, philosophers and the birth of the European  
knowledge economy. About the price and the production of useful knowledge in  
Europe 1350-1800,’ paper for the GEHN conference on Useful Knowledge, Leiden,  
September 2004; revised 12 October 2004.
Endnotes

1 See also de la Croix, Boucekkine, and Peeters (2007b), and de la Croix, Boucekkine, and Licandro (2003) on the human capital effects on longevity.

2 On student attendance at the medical school of Montpellier, see Chartier, Compère and Julia (1976).

3 Although our results also indicate that after a strong growth during the 16th and 17th centuries, book consumption in England stagnated at a relatively high level during the 18th century; see also St. Clair (2004: 84-100) for an analysis of the slow growth of book production in this period.

4 For the Netherlands, the literacy of migrants to Amsterdam has been studied (see the review in van Zanden 2004), whereas many regions are lacking, because signatures on the marriage records were not customary in Central and Eastern Europe.

5 We thank an anonymous referee for this comment on the plague as potential exogenous force. See also Herlihy (1997) for the literacy effects of the bubonic plague of the 14th century.

6 Moreover, Kremer studies this question both theoretically and with time series data on the growth of world population, and finds that a larger population caused a stronger population growth.

7 Sources: for the European countries and the US the same as for the other European countries: USA and Ireland the English Short Title Catalogue, Switzerland, Poland and Russia: the German book fairs in combination with the Hand-pressed book file (for Russia also Marker 1982); Indonesia: Isa 1972; Japan: Hayami and Kitô 1999: 241; China: Tsien Tsuen-Hsiun 1985: 190; India: Darnton 2002; see also the discussion of global patterns of book production in Van Zanden 2004. Because the 1820 estimates of GDP levels for the following test are relatively weak and subject to much debate – with the revisionists claiming that the gap between Europe and China was much smaller than estimated by Maddison (Pomeranz 2002) – we also compare absolute levels of GDP per capita in 1913, and the same book production data for the period of 1750-1800, and found our results confirmed.
Figure 1: Book production in early modern Europe, 1450/99-1750/99 (number of new editions per million inhabitants, log scale of vertical axis)

Source: see text.
Figure 2: Book production and literacy in England and Spain, literacy as estimated by Cressy (1980) and Nalle (1989)

Note: UK excludes Scotland here (and Ireland).
Figure 3: Real wages in early modern Europe, 1450/99-1800/49

Source: see Table 1.
Figure 4: Book production per capita between 1750 and 1800, and GDP per capita growth in 1820-1913 (books on log scale)

GDP_pc_growth = 0.484 + 0.134 lnBooks_pc

R-squared = 0.5563, N=17
**Table 1: Descriptive statistics of the variables used in the regressions**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std.Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real Wage Change</td>
<td>54</td>
<td>-0.209</td>
<td>0.759</td>
<td>-1.662</td>
<td>1.734</td>
</tr>
<tr>
<td>Book Consumption</td>
<td>54</td>
<td>3.886</td>
<td>1.253</td>
<td>0.399</td>
<td>6.248</td>
</tr>
<tr>
<td>Book Production</td>
<td>54</td>
<td>3.870</td>
<td>1.300</td>
<td>0.399</td>
<td>6.296</td>
</tr>
<tr>
<td>Initial Real Wage</td>
<td>54</td>
<td>5.142</td>
<td>1.545</td>
<td>2.530</td>
<td>7.979</td>
</tr>
<tr>
<td>Land per Agric. Pop.</td>
<td>54</td>
<td>8.154</td>
<td>0.395</td>
<td>7.282</td>
<td>8.991</td>
</tr>
<tr>
<td>Intercont. Trade</td>
<td>54</td>
<td>128.209</td>
<td>446.534</td>
<td>0</td>
<td>2726.4</td>
</tr>
<tr>
<td>Bubonic Plague</td>
<td>54</td>
<td>0.556</td>
<td>0.502</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Post-war</td>
<td>54</td>
<td>0.019</td>
<td>0.136</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>30-Years-War</td>
<td>54</td>
<td>0.019</td>
<td>0.136</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Capital Protection</td>
<td>54</td>
<td>2.296</td>
<td>1.574</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>Literacy (est.)</td>
<td>54</td>
<td>24.611</td>
<td>13.982</td>
<td>6</td>
<td>68</td>
</tr>
</tbody>
</table>

Sources: book production/consumption: see text; capital protection: Acemoglu et al. (2002); other variables: Allen (2003), except were data were missing (agricultural land taken from van Zanden 1991). Real wages for Sweden were friendly provided by Christiaan van Bochove. 30-Years-War and Post-War relate to Germany. Postwar is 1 for Germany during the half century after the Thirty Years’ War. Book production and consumption, as well as intercontinental trade, are per capita. Bubonic plague is 1 if an outbreak is recorded for a country and half-century. We thank Leondro Prados de la Escosura, Paola Malanima, Carlos Álvarez-Nogal, and Sören Edvinsson for important information on the plague occurrences.
Table 2: Two fixed effects regressions of real wage changes, 1450-1849

<table>
<thead>
<tr>
<th></th>
<th>Model (1)</th>
<th>Model (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prod. or Cons.?</td>
<td>PROD.</td>
<td>CONS.</td>
</tr>
<tr>
<td>Books</td>
<td>0.33**</td>
<td>0.38**</td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>Initial Real Wage</td>
<td>-0.65***</td>
<td>-0.66***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Intercont. Trade</td>
<td>-0.00</td>
<td>-0.00</td>
</tr>
<tr>
<td></td>
<td>(0.69)</td>
<td>(0.76)</td>
</tr>
<tr>
<td>Capital Protection</td>
<td>0.21*</td>
<td>0.21**</td>
</tr>
<tr>
<td></td>
<td>(0.053)</td>
<td>(0.042)</td>
</tr>
<tr>
<td>Land per agric. Worker</td>
<td>1.48*</td>
<td>1.56*</td>
</tr>
<tr>
<td></td>
<td>(0.072)</td>
<td>(0.060)</td>
</tr>
<tr>
<td>Post-war</td>
<td>1.17*</td>
<td>1.11</td>
</tr>
<tr>
<td></td>
<td>(0.087)</td>
<td>(0.10)</td>
</tr>
<tr>
<td>30-Years-War</td>
<td>-0.94</td>
<td>-0.99</td>
</tr>
<tr>
<td></td>
<td>(0.15)</td>
<td>(0.13)</td>
</tr>
<tr>
<td>Bubonic Plague</td>
<td>-0.46</td>
<td>-0.47</td>
</tr>
<tr>
<td></td>
<td>(0.31)</td>
<td>(0.29)</td>
</tr>
<tr>
<td>D 16(^{th})</td>
<td>-0.05</td>
<td>-0.04</td>
</tr>
<tr>
<td></td>
<td>(0.88)</td>
<td>(0.89)</td>
</tr>
<tr>
<td>D 18(^{th})</td>
<td>-0.56</td>
<td>-0.61</td>
</tr>
<tr>
<td></td>
<td>(0.24)</td>
<td>(0.21)</td>
</tr>
<tr>
<td>D 19(^{th})</td>
<td>-0.88</td>
<td>-0.95*</td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.085)</td>
</tr>
<tr>
<td>Constant</td>
<td>-10.14</td>
<td>-10.90</td>
</tr>
<tr>
<td></td>
<td>(0.15)</td>
<td>(0.12)</td>
</tr>
</tbody>
</table>

N 54 54
<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number countries</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Adj. R-Square</td>
<td>0.29</td>
<td>0.29</td>
</tr>
</tbody>
</table>

P-values in parentheses. Notes: The dependent variable is always the first difference between the log real wage of one half century and the next. All explanatory variables refer to the initial half century. Sources: see Table 1.
Table 3: Controlling for endogeneity and institutions: three IV fixed effects regressions of real wage changes, 1450-1849

<table>
<thead>
<tr>
<th></th>
<th>Model (1)</th>
<th>Model (2)</th>
<th>Model (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prod. or Cons.?</td>
<td>Production</td>
<td>Consumption</td>
<td>Production</td>
</tr>
<tr>
<td>Books</td>
<td>0.52**</td>
<td>0.51*</td>
<td>0.45*</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.09)</td>
<td>(0.07)</td>
</tr>
<tr>
<td>Initial Real Wage</td>
<td>-0.90***</td>
<td>-0.90***</td>
<td>-0.93***</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>Intercont. Trade</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.26)</td>
<td>(0.18)</td>
<td>(0.76)</td>
</tr>
<tr>
<td>Capital Protection</td>
<td></td>
<td></td>
<td>0.23**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.035)</td>
</tr>
<tr>
<td>Land per Agric. Worker</td>
<td>0.93</td>
<td>0.82</td>
<td>1.01</td>
</tr>
<tr>
<td></td>
<td>(0.46)</td>
<td>(0.54)</td>
<td>(0.39)</td>
</tr>
<tr>
<td>Post-war</td>
<td>0.74</td>
<td>0.69</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>(0.30)</td>
<td>(0.34)</td>
<td>(0.31)</td>
</tr>
<tr>
<td>30-Years-War</td>
<td>-1.08</td>
<td>-1.13*</td>
<td>-1.15*</td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.07)</td>
</tr>
<tr>
<td>D 16th</td>
<td>0.12</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>(0.72)</td>
<td>(0.78)</td>
<td>(0.77)</td>
</tr>
<tr>
<td>D 18th</td>
<td>-0.32</td>
<td>-0.36</td>
<td>-0.27</td>
</tr>
<tr>
<td></td>
<td>(0.20)</td>
<td>(0.17)</td>
<td>(0.25)</td>
</tr>
<tr>
<td>D 19th</td>
<td>-0.67*</td>
<td>-0.73*</td>
<td>-0.83**</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.06)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>Constant</td>
<td>-5.17</td>
<td>-4.26</td>
<td>-5.93</td>
</tr>
<tr>
<td></td>
<td>(0.64)</td>
<td>(0.72)</td>
<td>(0.57)</td>
</tr>
<tr>
<td>Observations</td>
<td>48</td>
<td>48</td>
<td>48</td>
</tr>
<tr>
<td>Number of Countries</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Adj. R-Square</td>
<td>0.19</td>
<td>0.18</td>
<td>0.24</td>
</tr>
</tbody>
</table>

P-values in parentheses. Notes: see Table 1. The instrumental variables for book production and consumption are (a) lagged book production and consumption, and (b) literacy, assuming
for 1450-1499 the value of 1500, and interpolating linearly to 1800, for which direct estimates are available. We avoid using urbanisation proxies for the centuries between 1500 and 1800, as this could be seen as a separate growth determinant, whereas we want to measure the human capital effect as purely as possible. For Sweden, a value equal to that of Germany is assumed.
Response to referees

We accepted all the suggestions of the reviewers. We are very thankful to the reviewer for providing these important comments, they helped to make the paper much clearer. Below we explain how we included them, reporting the referee’s point in italics (often shortened or summarized), followed by a description of our changes to the text.

Referee 1
The authors should scurpulously refer to the history and economic papers already published on the issue addressed
Thank you very much for encouraging a much more detailed literature review. Although it might be impossible to review each single study on each the countries we included (for example, on Spain alone, there are at least six relevant studies we could cite, and many more of lesser contribution), we extended the literature review considerably and avoided the impression that the field is dominated by Mokyr and Allen.

I’d like to see a serious comparison work and a methodological discussion. For example please compare between Cressy’s estimates of literacy and your book-based indicator for Great Britain. Are there any other literacy studies based on book production? In such a case, what is the originality of your own work?
Yes, we compared at much greater length book production and literacy for England, France and Spain in a figure now, and Sweden (with conceptually different literacy rates) in the text. We highlighted the differences between the two, and their methodological implications. As far as we know, there are only country studies on book availability that do not claim to be comparable to other countries (plus a more descriptive working paper on Medieval manuscript numbers and early modern book numbers by one of us)

If I am not wrong, there is no (strictly speaking) demographic variable among them. I have in mind that even in the late middle ages, epidemics (not only the bubonic plagues) were frequently hitting the European (and non-European) populations. Admittedly, such epidemics have a mechanical upward effect on wages (because labor supply is terribly affected).
(...)May I suggest to include an epidemic-related variable among the controls just like the authors include war-related factors?
Unfortunately, there is no systematic data on the various sorts of epidemic disease for all these countries over time, but we were able to find out in which countries and half-centuries the bubonic plague was mentioned, and included it in the regressions (it turned out insignificant). It might have been that the quantitative effect of the plague and other epidemic diseases was already reflected in our “Malthusian” variable ‘agricultural population per arable land’, or that the potential wage-increasing and –decreasing effects of plague cancel each other out.

Referee 2
We accepted all the editorial remarks, which were very helpful indeed for improving the structure of arguments. Of course, we needed to restructure the text in order to respond to referee 1.

On, p.8, meaning of the second to last line of the first paragraph is not clear.
We omitted this part, as it might have been too narrowly focused on a previous comment by another colleague – and cutting it provided additional space to include the literature review which referee 1 suggested.
If you are going to refer to Kremer's "research equation," then you have to explain it.
Yes, perfectly right. After considering carefully the space and the interruption of the main line of the argument, we decided for deleting the hint to the research equation, but report Kremer’s results in somewhat greater detail (see the footnote on his theoretical and time series work).

In the current Figure 2, you have commas where you should have decimal points.
We corrected this.