Essays on Entrepreneurship and Innovation

Essays on Entrepreneurship and Innovation

A dissertation presented

by

Francisco Queiró

to

The Committee for the PhD in Business Economics

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Business Economics

Harvard University

Cambridge, Massachusetts

July 2015

Dissertation Advisor: *Author:*
Professor Andrei Shleifer **Francisco Queiró**

Essays on Entrepreneurship and Innovation

Abstract

These essays investigate the role of entrepreneurial human capital as a driver of innovation and growth. In the first chapter, I estimate the effect of manager education on firm employment growth using administrative panel data on the universe of firms in Portugal between 1995 and 2009. I exploit manager changes and switches between management and other occupations to account for unobserved firm and manager characteristics as well as selection into management. I find that a year of manager schooling increases firm growth by around 0.25 percentage points per year, and also increases survival. In addition, manager education is a highly persistent firm characteristic. These findings imply that manager education can lead to large differences in firm size over the lifecycle. On average, a firm with a college educated manager starts out just nine percent larger than a firm with a primary school educated manager. In a simple simulation, I find that it grows to two thirds larger by age 12 and three times larger by age 30.

Entrepreneurial human capital can increase growth through different mechanisms. One of them is the ability to incorporate advances in the technological frontier into production. In the second chapter I explore the importance of this mechanism by analyzing the relationship between local demand for knowledge and city growth using a new database of 5.5 million books published in Europe from 1450 to 1800. The database consists of individual book data drawn from over 72,000 library catalogs around the world, including most major national and research libraries. Exploiting within-city variation, I find that book production is a strong predictor of subsequent population growth. I then distinguish between possible interpretations of this relationship using information on book subjects. I find that the results

are robust for books on technology, finance, medicine and history, with technology and finance having the largest coefficients. In addition, although science books as a whole are insignificant, books on chemistry and geology also increase growth, which is consistent with the important roles of chemistry and coal mining during the Industrial Revolution. Books on other topics, such as religion or literature, are not associated with growth, suggesting that the findings reflect the diffusion of knowledge rather than literacy or consumption.

Contents

List of Tables

List of Figures

Acknowledgments

It has been eight years since I first entered Andrei Shleifer's office and declared my intention to pursue a PhD. Having failed to disabuse me of my "starry-eyed" view of academia – his words – Andrei took me under his wing. I have been enormously fortunate to benefit from his mentorship ever since, first as a research assistant and then as a student. Andrei's influence over my development has been wide-ranging, and it is difficult to do it justice in a few words. Perhaps above all I learned to cultivate a relentless curiosity about the world.

I am also indebted to Larry Katz, Josh Lerner and Alberto Alesina. Larry challenged me to sharpen my analysis and exposition and offered invaluable feedback throughout. Josh welcomed me when I became interested in entrepreneurship and has been more than generous with encouragement and help. Alberto offered enthusiastic support and pushed me to explore the implications of my findings to their full extent. In addition, I have benefited from conversations and feedback from Raj Chetty, Claudia Goldin, Simon Jaeger, Danial Lashkari, Tom Zimmerman and Eric Zwick, and I thank Joao Leão and Teresa Feliciano for the opportunity to work at the Gabinete de Estratégia e Estudos (GEE).

Finally, I am grateful to the wonderful friends I made in graduate shcool, in particular Joana Naritomi, Ian Tomb, Clara Zverina, Arash Nekoei, Ariel Stern, Ruchir Agarwal, Pedro Bordalo and Heather Schofield. Beyond their generous feedback, they have been a source of joy, inspiration and support.

Para a Elvira

Chapter 1

The Effect of Manager Education on Firm Growth

1.1 Introduction

Firms experience highly heterogeneous lifecycle growth.[1] New businesses tend to start small, and a large fraction exits within a few years after entry. Most survivors remain small throughout their lifecycle, but other firms exhibit significant growth. These differences matter. Entrepreneurship and the reallocation of workers across firms are key sources of economic growth,[2] and high growth firms account for a disproportionate share of reallocation and job creation.[3] There is also increasing evidence that firms in more developed countries exhibit stronger average growth, and that this has a significant impact on aggregate productivity.[4] In line with these findings, firm heterogeneity has become a central feature in standard models in various fields, such as economic growth (Aghion and Howitt, 1992),

[1] Dunne *et al.* (1989) and Davis and Haltiwanger (1992).

[2] Kortum and Lerner (2000), Foster *et al.* (2001) and Glaeser *et al.* (2014)

[3] In the U.S., according to Decker *et al.* (2014), an average of 18 percent of jobs are created every year, and 50 percent of these are accounted for by just 15 percent of firms

[4] Hsieh and Klenow (2014) and Bento and Restuccia (2014).

development accounting (Hsieh and Klenow, 2009) or trade (Melitz, 2003). Yet, little is known about why some firms grow more than others.

One possible source of heterogeneity is managerial human capital, an idea going back to the work of Penrose (1959) and Chandler (1977). This paper investigates the relationship between manager education and firm growth using administrative panel data on the universe of firms in Portugal. Portugal is a particularly attractive setting for four reasons. First, the data cover firms of all sizes and at all stages of the lifecycle. Second, there are detailed occupational codes for all workers, which enables me to identify managers directly based on their function.[5] Third, and unlike in samples of large firms, there is substantial variation in managerial education.[6] Finally, the large sample size and panel structure for both firms and workers enable the implementation of a research design that accounts for unobservable firm and manager characteristics.

I start by presenting evidence on the relationship between firm size, age and manager education to motivate the analysis. In the cross section of firms, firm size is positively associated with manager education, and not with non-manager education.[7] The extent of the relationship, however, depends strongly on firm age. For example, a firm with a college-educated manager is on average just 9 percent larger at entry than a firm with a primary-school educated manager. At 12 years old, which is the average age in the sample, the firm with the college-educated manager is nearly twice as large. And at 30 years old it is four times larger. To get a sense of magnitude, these differences are similar to the average size differences between U.S. and Indian manufacturing establishments at the same ages (Hsieh and Klenow, 2014). The same patterns hold when following a given cohort over

[5]Other studies have used indirect proxies like employment status or presence at entry to identify entrepreneurs (e.g. Nanda and Sorensen, 2010).

[6]In samples of large firms, most top managers are college graduates. For example, La Porta and Shleifer (2014) report that 76 percent of managers in the World Bank Enterprise Surveys of formal firms have a college degree. In the sample used in this paper the median firm's manager has only completed ninth grade, and the standard deviation of average manager education across firms is 4.3 years of schooling.

[7]This finding is consistent with La Porta and Shleifer (2008), who report that informal firms are smaller and have less educated managers than their formal counterparts in the World Bank Enterprise Surveys, and with Bloom and Van Reenen (2007), who find that good management practices are positively correlated with firm size among medium-sized manufacturing firms.

time and sorting firms by manager education at entry. The evidence clearly suggests that manager education increases firm growth. Investigating this hypothesis is the main goal of my paper.

I structure the analysis around two questions. First, do more educated managers increase firm growth? Second, is education itself the cause or a proxy for unobserved manager characteristics like talent, ambition or family background? Each question has independent interest. The first asks whether management matters, and in that sense tests the ideas of Penrose and Chandler using education as a measure of managerial quality. The second asks whether manager education itself matters, and if management is a channel through which education could affect firm dynamics and aggregate productivity.

I exploit within-firm variation in manager education to answer the first question. Using event studies of manager changes, I find that firm growth increases sharply for firms that hire college-educated managers relative to firms that hire managers who completed the 12^{th} grade or less, controlling flexibly for pretreatment firm characteristics. I then extend the analysis to the entire sample and use a firm fixed effects model to estimate the effect of manager education on growth. I find that a year of manager education increases firm growth by around 0.25 percentage points conditional on survival, and 0.35 when I account for exit by assigning a growth rate of -100 percent to exiting firms.

The identifying assumption in this design is that firms making different changes to manager education would have followed similar growth trends in the absence of these changes. One concern is that more educated managers might sort into firms that experience positive growth shocks. I present several pieces of evidence to address this concern and support the parallel trends assumption. First, the pre-treatment growth trends for the treatment and control groups in the event studies are very similar, with a clear break in the year of treatment. Second, a series of placebo event studies show that increases in education in other occupation groups – professionals, office workers, service workers and blue-collar workers – have no effect on firm growth. Third, the firm fixed effects estimates are stable across specifications including controls for a rich set of firm characteristics, sector-by-year fixed effects and firm-specific growth trends. Fourth, 77 percent of firm-years are managed by owners, suggesting that in practice the market for managers is limited, presumably due to agency costs. Managerial changes at owner-managed firms are likely to be confined to small networks around owners, namely their families, rather than reflect the type of sorting that would bias the results. I show that the results are unchanged when the sample is restricted to firms that were owner-managed both before and after the changes. Finally, the results are also stable when the sample is restricted to managerial changes where at least one of the exiting managers leaves the sample permanently. Since the data cover the universe of firms, this implies that with high probability this manager exited the labor force, and therefore that the change is more plausibly orthogonal to shocks to the firm's growth prospects.

I then turn to the second question. The results so far show that more educated managers increase firm growth, but do not reveal if the cause is education itself. For example, education could be correlated with factors such as ability, ambition or family background that might affect firm growth. Even if these characteristics were uncorrelated with education in the population, there could be a selection bias: workers with higher schooling could have a better outside option and therefore only choose to become managers when they have higher ability.

If omitted ability, ambition or background affect firm growth, they presumably also affect the outside option of managers in other occupations. For a sub-sample of managers that have switched between management and other occupations during the sample period, I observe income in non-managerial occupations at other points in time. I show below that under reasonable assumptions this information about the manager's outside option can be used to account for these unobserved characteristics. While this strategy addresses the bias from omitted manager characteristics, it introduces another bias since non-managerial income is partly determined by education. Unlike the original bias, however, this bias is plausibly negative, which implies that I estimate a lower bound on the true coefficient. In addition, the coefficient on non-managerial income can be used to recover the bias and obtain an estimate of the true coefficient. I report two findings. First, the estimated lower bound is positive and significant, and about twenty percent lower than the baseline estimate without the outside option control. Second, the bias-corrected estimate turns out to be very similar to the baseline estimate, suggesting the bias from omitted manager characteristics is negligible. These results hold both in the full sample and the owner-managed sample.

I conclude by evaluating the implications of these results for firm lifecycle growth. I first show that manager education is a highly persistent firm characteristic, with one and ten year autocorrelations of 0.97 and 0.82 respectively. I then perform a simple simulation and find that the estimated effect on growth translates into substantial differences in firm size over time. I calculate that a firm with a college educated manager grows to about two thirds larger than a firm with a primary school educated manager by age 12, and about three times larger by age 30, explaining most of the observed size difference between firms with similar differences in manager education in the cross-section.

This paper contributes to the literatures on management, entrepreneurship and firm dynamics. The relationship between management and firm performance is an old topic. Penrose (1959) and Chandler (1977) focused on the role of management as a constraint on the firm's ability to expand, an idea formalized in Lucas (1978). Difficulties in the measurement of managerial quality have traditionally limited progress in this literature, but recent research has made important strides. Bertrand and Schoar (2003) estimate manager fixed effects by tracking CEO movements across firms. Bloom and Van Reenen (2007) use detailed surveys to develop an index of management practices. Gennaioli *et al.* (2013) use the education of managers as a measure of their quality, as in this paper. These studies have found large effects of management on firm performance, but have focused on productivity and profitability rather than growth. In fact, because they use samples of large firms, these studies are conditional on having experienced the type of growth whose determinants we would like to understand.

In entrepreneurship studies, there is a growing emphasis on the fact that entrepreneurial characteristics and performance vary widely (Schoar, 2010; Hurst and Pugsley, 2011).

Understanding what distinguishes the entrepreneurs that drive economic growth is an open challenge.[8] The literature on firm dynamics has studied several hypotheses. Besides management quality, these papers have developed theories based on experimentation and learning (Jovanovic, 1982; Ericson and Pakes, 1995; Foster *et al.*, 2013), risk aversion (Kihlstrom and Laffont, 1979) and financial constraints (Cooley and Quadrini, 2001; Cabral and Mata, 2003; Albuquerque and Hopenhayn, 2004). But empirical evidence on the effects of these different factors has been limited.[9] This paper is among the first to find evidence of an entrepreneurial characteristic that drives significant differences in growth across firms.

Another strand in the firm dynamics literature studies differences in average firm growth across industries or countries. These papers have analyzed the effect of tax and regulatory enforcement (Harris and Todaro, 1970; Rauch, 1991), financial constraints (Desai *et al.*, 2003), contract enforcement (Bloom *et al.*, 2013; Akcigit *et al.*, 2015), culture (La Porta *et al.*, 1997; Bloom *et al.*, 2012) and trade costs (Melitz, 2003). Although the focus of this paper is on differences across firms, the findings suggest that manager education can play a role in explaining differences across countries as well. In that sense, the paper also highlights a mechanism through which education can account for cross-country income differences (Mankiw *et al.*, 1992), and complements the findings of Gennaioli *et al.* (2013).

The rest of the paper is organized as follows. Section 1.2 describes the data and analysis sample, and reports summary statistics. Section 1.3 presents graphical evidence on the correlation between firm size and manager education and how it varies with firm age. Section 2.4 explains the research design and presents the main findings. Section 1.5 evaluates implications for lifecycle growth. And section 2.5 concludes.

---

[8] A recent paper by Guzman and Stern (2015) uses information at the time of registration such as the firm's name or filling for a patent to predict right tail outcomes such as IPOs and acquisitions.

[9] Using a sample from the same dataset as this paper, Cabral and Mata (2003) argue for an important role for financial constraints, which they proxy for with the owner's age. But Angelini and Generale (2008) reach the opposite conclusion using direct measures of financial constraints for a sample of Italian firms.

1.2 Data

The paper uses data from *Quadros de Pessoal*, a matched employer-employee administrative panel data set collected annually by the Ministry of Employment in Portugal that covers the universe of firms with at least one employee and their workers, including owners and unpaid family workers. The survey combines firm-level information, such as total employment, sales and date of incorporation, with a wide range of worker characteristics. Over the sample period from 1995 to 2008 it contains 36 million worker observations corresponding to 9 million individuals, and 3.8 million firm observations corresponding to 701 thousand firms. This section defines the variables and sample used in the analysis and provides summary statistics.

1.2.1 Variable Definitions

Managers The concept of manager in this paper is that of a top decision maker, what is commonly referred to as top management. Typical top management decisions include resource allocation and high level coordination and evaluation (Chandler, 1977). A key step in the analysis is the identification of managers in the data. Studies of entrepreneurship using large scale employer-employee matched datasets have identified entrepreneurs as workers who report being self-employed or present at the time of incorporation (e.g. Nanda and Sorensen, 2010). These procedures exclude professionally managed firms as well as firms where entrepreneurs choose to become employees of the firm. A unique strength of the data I use is the 6-digit occupational classification system that was introduced in 1995 (CNP 94), which identifies detailed managerial occupations and in particular accounts for managers of small firms. This occupational classification system enables me to consistently define top managers across firms directly based on their function.

I use the occupational data to identify a firm's managers as follows. CNP 94 groups top management positions into two sections according to firm size: section 1.2 Firm Directors and section 1.3 Small Firm Directors and Managers, where small firms are defined as having fewer than 10 workers. For small firms, section 1.3 does not provide additional functional detail. I therefore define small firm managers as all workers reported under section 1.3. For larger firms, section 1.2 classifies managers according to a hierarchy. At the top are General Managers, whose description most closely resembles that of top decision makers. The next category are Operations Managers, who also participate in high level decision making but report to General Managers when they exist. And the last group are Other Managers, who lead narrower functional areas. Other Managers are further sub-divided by functions, such as Administrative, Financial and Sales. I define managers for section 1.2 firms as the workers at the top managerial position that the firm reports: General Managers if they exist, Operations Managers if there are no General Managers, and Other Managers if there are no General or Operations Managers. The results in the paper are robust to other procedures such as defining all workers under section 1.2 as the firm's managers.

This procedure identifies managers for 54 percent of firm-years and 77 percent of employment in the data. For the firm-years that do not report any workers under sections 1.2 and 1.3, I proceed as follows. If the firm reports any owners among its workers, I define the owners as the managers. This covers an additional 6 percent of firm-years and 3 percent of employment. If the firm does not report any owner-workers, I check whether the firm had any managers in the previous year. If it did, I assign the previous year's managers that are still working at the firm (but are now classified as workers) as managers. I then check whether the firm had any managers in the following year. If it did, I assign the following year's managers that were already working at the firm (but are still classified as workers) as managers. This procedure is meant to correct occupational classification errors, and covers another 4 percent of firm-years and 6 percent of employment. The results are robust to excluding managers identified through these additional procedures. In total, I identify managers for 64 percent of firm-years and 85 percent of employment in the sample.[10]

Other variables The outcome of interest in the paper is the firm's annual growth rate, and I also use data on average manager and non-manager education and age, the firm's age,

[10] The remaining 36 percent are very small firms, with an average of 3.6 workers. Managerial tasks at these firms are presumably minimal and not the primary focus of any of the firm's workers.

size, revenue and two digit sector. In some specifications I also use information on manager employment status and on non-manager income.

Firm growth is defined as the annual percentage change in employment, winsorized at the 99^{th} percentile.[11] In specifications where I account for firm exit, firm growth is defined as -100 percent in the year that a firm exits. All results are robust to winsorizing at the 95^{th} percentile instead. Firm size is defined as the firm's total employment at the beginning of the year.

Education is measured as years of schooling completed. The data report the highest level of schooling attained by each worker, where the levels are: no schooling, 4th grade, 6th grade, 9th grade, 12th grade, *bacharelato* and *licenciatura*. The *bacharelato* and *licenciatura* are higher education degrees typically lasting three and five years, respectively.[12] The distinction is similar to that between associate and bachelor's degrees in the U.S.. When a worker reports different levels of educational attainment in different years I take the mode of these reports as that worker's attainment, to reduce measurement error. If there is more than one mode, I take the lowest.

Firm age is constructed using the firm's reported year of incorporation. When a firm reports different years of incorporation over time I take the mode as I do with education. One issue with the revenue data is that it corresponds to calendar years, whereas all other variables are measured as of the survey's reference week in October.

1.2.2 Analysis Sample

The data are available for the period from 1985 to 2009, which implies that I can measure firm growth up to 2008. As described above, the occupational classification system used to identify a firm's managers was introduced in 1995. Firm age is also available starting in

[11] An advantage of defining firm growth as a percentage change is that it provides a natural way to account for firm exit. A disadvantage is that the distribution is significantly right-skewed, as compared to the distribution of log growth. I winsorize the data to reduce the influence of outliers on the right tail. All results hold using log growth as well.

[12] The higher education system changed in 2006 with the EU's Bologna Accords, but these changes are too recent to affect the sample used in this paper.

1995. I therefore restrict the analysis to the period between 1995 and 2008. In addition, data on worker characteristics is not available in 2001.[13]

The focus of the paper is on private-sector firms. I exclude state-owned firms, defined as those that take the legal form of *Empresa Publica* (state-owned company) or where the state has an equity stake of at least 50 percent. I also exclude government agencies, which are covered when they employ workers under private sector labor law, and non-profits. A number of large privatizations occurred during the sample period, involving significant mergers, breakups and downsizings. I exclude these firms by also dropping all private firms that were state-owned at any point in time.[14] Altogether, I exclude 2.6 percent of firms with these filters.

In addition, I use a set of firm-level controls in most specifications which includes non-manager characteristics. To ensure results are comparable across specifications with and without controls I restrict the sample to firm-years that have at least one manager and one non-manager and non-missing controls.[15] The final sample consists of 1.7 million observations and 327,823 firms.

Table 1.1 presents summary statistics. Average firm growth is 1.5 percent, with a standard deviation of 31 percent. Two additional facts stand out. First, there is substantial variation in manager education. The median firm's manager has the ninth grade, and the standard deviation across firms is 4.3 years of schooling. Non-manager education has a slightly lower median and displays less variation. Second, the sample is dominated by small firms. Average employment is 14 workers, with a median of five. Even the firm at the 90^{th} percentile employs only 23 workers. Firms employ 1.5 managers on average, and the

[13]The data set consists of three databases: a firm-level database (covering firm-level information such as firm age and total employment), an establishment-level database (e.g. location, employment) and a worker-level database (e.g. education, occupation). The worker-level database is not available in 2001.

[14]In some cases the privatized firms were reincorporated and show up as new firms in the data. To identify these cases, I follow the procedure in Braguinsky *et al.* (2011): I take all entering firms with over 50 employees and identify those where a majority of workers worked at state-owned firms in the previous year. This procedure identifies an additional 49 firms that I exclude.

[15]Missing controls is not a significant issue. Education and age are available for 99 percent of worker-years; firm age is available for 99 percent of firm-years and firm size is available for all firm-years.

majority employs just one. This reflects a key advantage of this dataset for the study of firm growth: the sample covers the universe of firms, rather than just firms that have already grown beyond a certain threshold.

Table 1.1: *Summary Statistics*

Variable	Mean	Std. Dev.	P10	P50	P90
Firm Growth (%)	1.48	30.99	-33.33	0	33.33
Manager Education	8.71	4.34	4	9	17
Non-Manager Education	7.56	3.01	4	7	12
Manager Age	45.02	10.27	32	45	59
Non-Manager Age	35.74	8.42	25.4	35	46.67
Number of Workers	14.18	86.51	2	5	23
Firm Age	12.57	12.42	2	9	27
Number of Managers	1.53	1.71	1	1	2

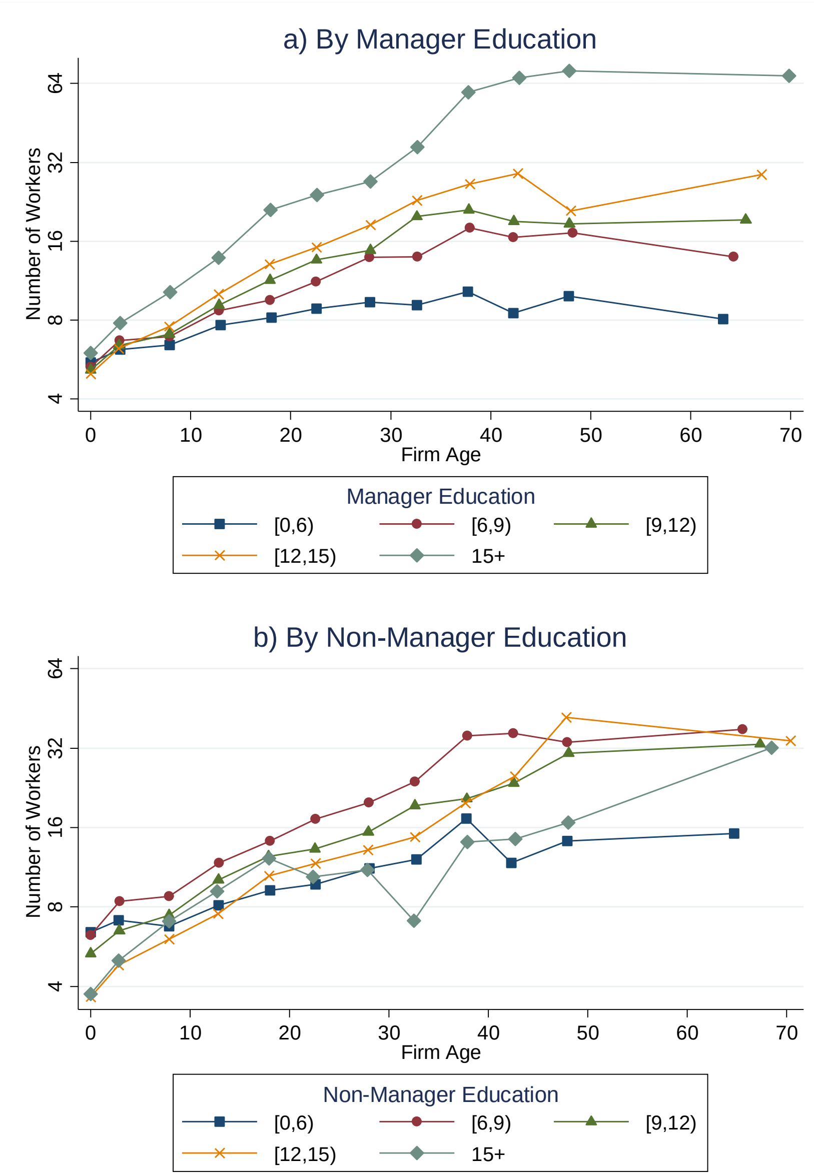
Notes: This table presents summary statistics for the main variables used in the analysis. Firm growth is the annual growth rate in employment, winsorized at the 99^{th} percentile. Manager education is the average years schooling of managers at the start of the year. Non-manager education, manager and non-manager age are defined analogously. The number of workers includes all firm workers regardless of employment status, including unpaid workers. Managers are defined in section 1.2 of the text. Firm age is based on the firm's reported year of incorporation.

1.3 Firm Size, Age and Manager Education

I start by presenting graphical evidence on the relationship between firm size, age and manager education to motivate the analysis.

Figure 1.1a presents the cross sectional relationship between firm size and manager education, broken down by firm age, all measured in 2009. To construct it, I first sort firms by average manager education into five groups – zero to less than six years of manager schooling, six to less than nine, nine to less than twelve, twelve to less than fifteen and fifteen and over. Within these groups, I divide firms into five-year age bins, including a separate bin for entrants and grouping firms over 50 years old into a > 50 bin. For each manager education group, I plot average firm size and average firm age for each age bin, where size is winsorized at the 99^{th} percentile within each age bin.

Figure 1.1: *Firm Size by Age, 2009 Cross-Section*



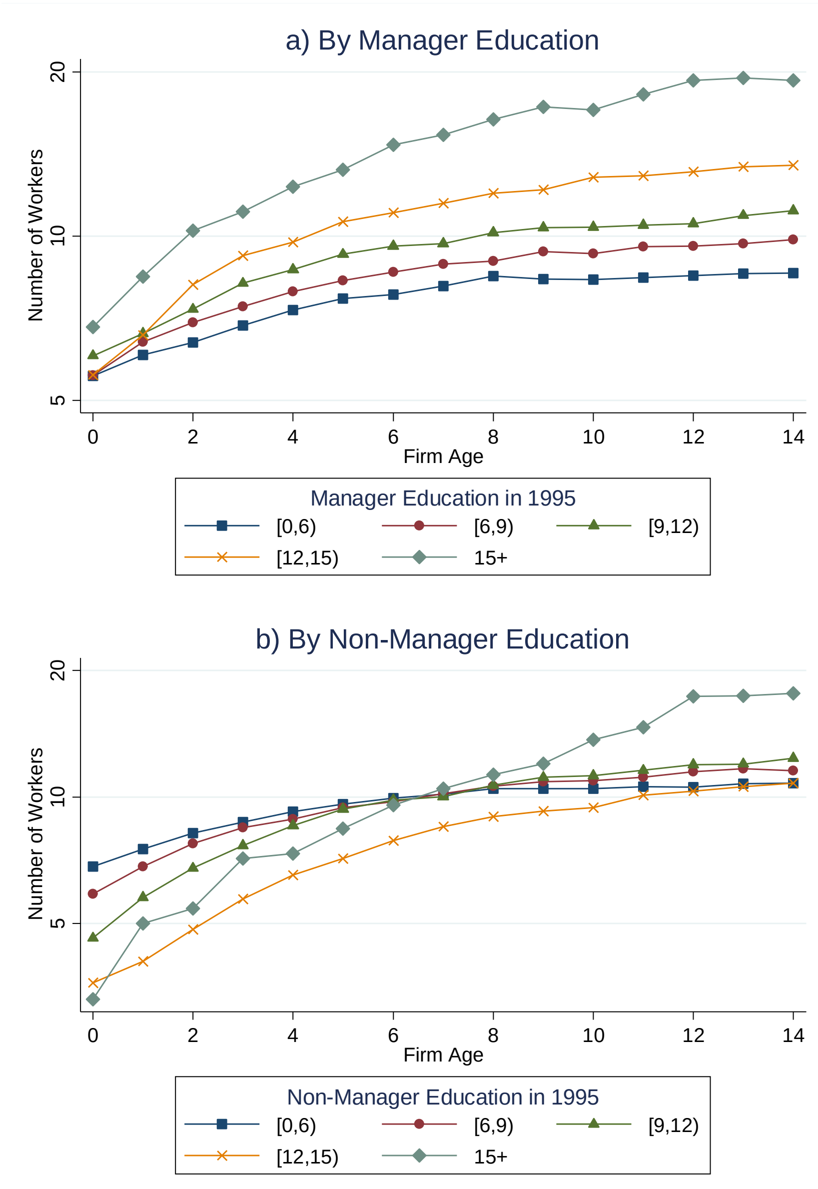
Notes: These figures plot average firm size (number of workers) by firm age bins in the 2009 cross section. The first bin includes entrants and the last includes firms over 50. The remaining firms are grouped into 5-year bins. Firm size is winsorized at the 99^{th} percentile within each age bin, The top graph sorts firms into five groups by average manager years of schooling, while the bottom graph sorts firms by average non-manager education instead.

The figure shows that firm size is positively correlated with manager education at all ages, but the extent of the relationship increases strongly with firm age. At entry, firms are small regardless of manager education. For example, firms with managers in the top group (15+) are on average just 9 percent larger at entry than firms whose managers have less than six years of schooling. For older firms, the difference across education groups becomes significantly larger. For the group aged 11 to 15, which corresponds to average age in the sample, firms in the top group are nearly twice as large as those whose managers have less than six years of schooling. Between 31 and 35, they are over four times larger, and beyond age 40 they are eight times larger. To get a sense of magnitude, Hsieh and Klenow (2014) report similar average size differences between U.S. and Indian manufacturing establishments at the same ages. Figure 1.1b shows that this relationship is specific to management by repeating the same exercise but sorting firms by non-manager education instead. If anything, firm size is negatively correlated with non-manager education at entry, and the two seem largely unrelated beyond age 10.

Figures 1.1a and 1.1b examine a cross section of firms. One explanation for these patterns could be that firms hire more educated managers as they grow. To address this possibility, I repeat the exercise but tracking a cohort of firms over time. I take the oldest cohort of entrants in the sample, the 1995 cohort, sort them by average manager schooling at entry and track these firms until age 14, in 2009. Figure 1.2a presents the results. I include only survivors at each age, but the pattern is the same when I assign a size of zero to firms that exit. Again, the differences across manager education groups are small at entry and become larger with age. By age 14, firms in the top group at entry are on average over twice as large as firms in the bottom group, the same magnitude as in the cross-section. Finally, figure 1.2b sorts the same firms by non-manager education at entry. As in the cross-section, there is a negative correlation between non-manager education and firm size at entry, and firm size across groups seems to converge over time.

Table 1.2 shows that the variation in firm size explained by manager education also increases substantially with firm age. The R^2 from a regression of log firm size on manager

Figure 1.2: *Firm Size by Age, 1995 Cohort*



Notes: These figures plot average firm size (number of workers) by firm age for the 1995 cohort. Firm size is winsorized at the 99^{th} percentile within each age bin. The top graph sorts firms into five groups by average manager years of schooling at entry, while the bottom graph sorts firms by average non-manager education at entry instead.

education in the 2009 cross-section rises to 20 percent for firms over 40 years old. For non-manager education, it is below 3 percent at all ages.

Table 1.2: *Firm Size Variation Explained by Manager and Non-Manager Education*

	R^2 From Univariate Regressions of Log Firm Size		
Firm Age Bin	Manager Education	Non-Manager Education	Observations
0	0.000	0.015	5113
1 to 5	0.001	0.007	39559
6 to 10	0.007	0.000	42095
11 to 15	0.013	0.000	24997
16 to 20	0.037	0.002	20003
21 to 25	0.063	0.003	12895
26 to 30	0.082	0.003	7899
31 to 35	0.121	0.007	4758
36 to 40	0.137	0.002	2527
41 to 45	0.203	0.026	2081
46 to 50	0.175	0.020	1069
51 and over	0.217	0.016	3074

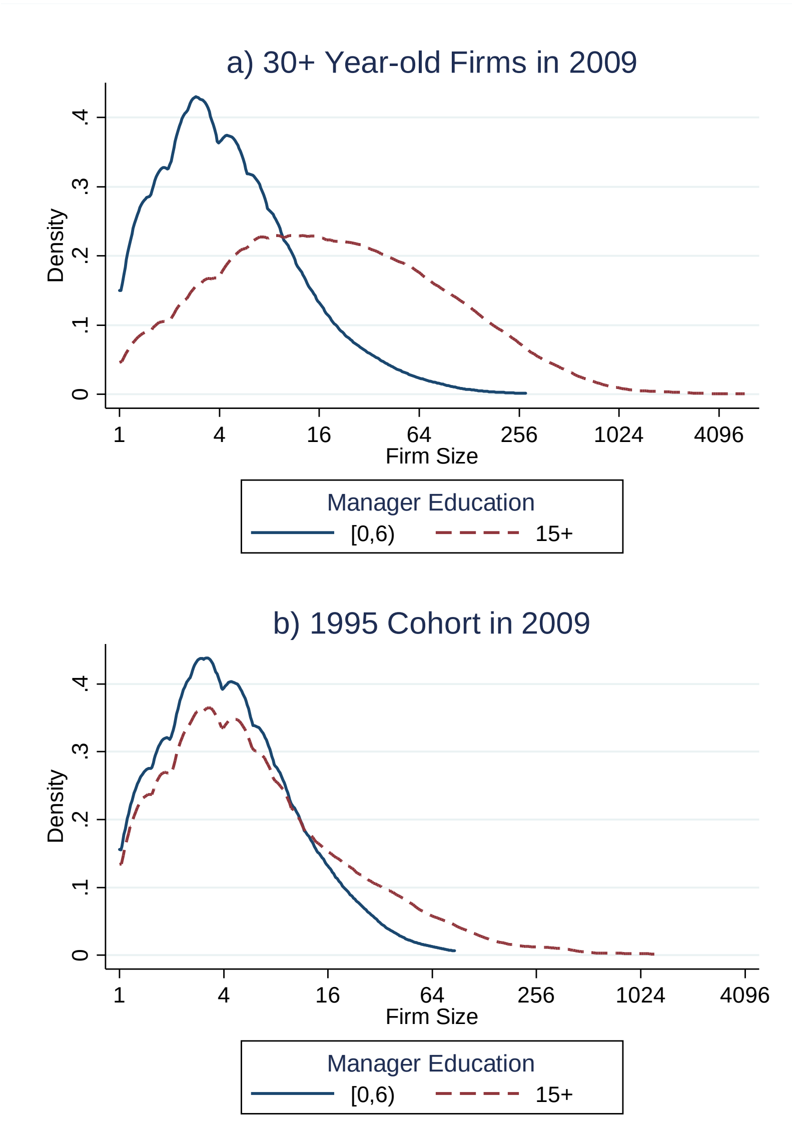
Notes: This table reports the R^2 from unvariate regressions of log firm size on manager education in the 2009 cross-section, by firm age bins (second column) and the R^2 from the corresponding regressions of log firm size on non-manager education (third column). The fourth column reports the number of observations in the regressions in each age bin.

These differences are not driven by superstars. Figure 1.3a shows the firm size distribution for 30+ year-old firms in the top and bottom groups in the 2009 cross-section.[16] Figure 1.3b does the same for firms from the 1995 cohort in 2009, again using manager education at entry. In both cases the difference in average size between the two groups is driven a mass of firms in the ∼10-250 size range, and not just the right tail.

Overall, the data show that there is a positive relationship between firm size and manager education and that this relationship strengthens significantly as firms age, which suggests that manager education increases firm growth. In the following analysis I explore this hypothesis.

[16] For figures 1.3a and 1.3b I use non-winsorized values for firm size.

Figure 1.3: *Firm Size Distribution*



Notes: These figures plot kernel density estimates of the firm size distribution for firms with different levels of average manager education. The solid line is the distribution for firms with managers that have between 0 and 6 years of schooling, and the dashed line is the distribution for firms with managers with 15 or more years of schooling. The top figure plots these two distributions for firms over 30 years old in the 2009 cross-section, with manager education measured contemporaneously (i.e. in 2009). The bottom one plots them for firms from the 1995 cohort in 2009, with manager education measured at entry.

1.4 Effect of Manager Education on Firm Growth

I structure the analysis around two questions. First, do more educated managers increase firm growth? And second, is education itself the cause or a proxy for unobserved manager characteristics like talent, ambition or family background? This section addresses each question in turn.

1.4.1 Do More Educated Managers Increase Firm Growth?

Methodology The challenge in answering the first question is isolating variation in manager education from other firm characteristics that affect firm growth. In order to account for unobserved drivers of growth, I exploit within-firm variation in manager education. I assume that growth g_{jt} for firm j at time t takes the form

$$g_{jt} = \alpha_j + \alpha_1 s_{jt} + `\mathbf{F}_{jt} + \delta_t + \varepsilon_{jt} \tag{1.1}$$

where s_{jt} is average manager education and $\mathbf{F}_{jt}$ is a vector of observable firm characteristics. α_j is a firm fixed effect and ε_{jt} is a random shock.

The key assumption underlying this design is that firms making different changes to manager education would have followed similar growth trends in the absence of these changes, conditional on changes to other observable characteristics. One concern is that better managers might sort into firms that experience positive growth shocks. In practice, 77 percent of firm-years in the sample are managed by firm owners, which suggests that there is a limited market for managers, presumably due to high agency costs. It is likely that most managerial changes are confined to small networks around firm owners, namely their families. I show below that the results hold when the sample is restricted to owner-managed firms. In addition, I present several pieces of evidence to support this parallel trends assumption, as well as results where I relax it by allowing for firm-specific growth trends.

Another issue is how to handle observations for years when manager education changes. In the data, I observe manager education at the start and end of the year, and if the two are different I infer that a change happened at some point in between. When a change occurs, the extent to which each management team is responsible for growth in that year is therefore unclear. In order to avoid this ambiguity and the measurement error it introduces in manager education, I restrict the sample to years when manager education does not change. For example, if a firm is present for five years in the sample and changes manager education at some point in year 3, I exclude year 3 from the analysis and identify the effect of the change by comparing growth in years 1 and 2, on one side, with growth in years 4 and 5 on the other. In the robustness analysis below I also report results removing this restriction and using average manager education at the start and end of the year instead.

Event Studies of Manager Changes I start by presenting graphical evidence from event studies around manager changes. I define an event as a change in manager education holding the number of managers constant. I further require that manager education be constant for at least one year before and one year after the change. This implies that the management teams before and after the change were present at the firm for at least a full year, and is meant to identify real changes in management rather than temporary absences in the survey reference week as well as reduce measurement error in manager education. Finally, and letting $t = 0$ denote the event year, I define the event window as the three years before and after the event, that is from $t = -3$ to $t = 2$. If a firm experiences multiple events, I include each event and the corresponding three-year window, regardless of any possible overlap across event windows.

To analyze the impact of manager education on firm growth I define two groups: a treatment group, consisting of firms that hire college-educated managers (545 events),[17] and a control group, consisting of firms that hire managers who have on average completed 12 years of schooling or less (2699 events). At each moment in event time, I regress the outcome variable on an indicator for treatment and a set of controls: quartics in pretreatment manager and non non-manager education and age, log firm size and log number of managers, as well as year, sector and firm age fixed effects. All controls are measured at the end of year

[17] College here means the equivalent of a bachelor's degree in Portugal, which typically corresponds to a total of 17 years of schooling.

$t = -1$, immediately prior to the manager change. I then plot average growth for treatment and control groups such that at each moment the difference between the two equals the coefficient on treatment and the weighted average of the two groups equals the sample average. The coefficients on treatment in the pretreatment period provide both a placebo test and a visual test of the parallel trends assumption.

Note that by construction an event requires the firm to be present in the sample between $t = -1$ and $t = 1$. In order to obtain a balanced sample, I additionally require that the firm be present in the sample in years $t = -3$, $t = -2$ and $t = 2$. This implies that I exclude manager changes before age 3 and also that I exclude firms that exit before the end of year $t = 2$. Average survival across all events equals 82 percent at $t = 1$ and 65 percent at $t = 2$. Requiring survival until $t = 2$ could bias the results if treated and control firms experience different survival rates, conditional on controls. To evaluate this possibility I regress survival at $t = 1$ and $t = 2$ on a treatment indicator and the same set of controls. In both cases, the coefficient on treatment is between 0 and 1 percent and insignificant.

As a first step, figure 1.4a shows the effect of treatment on year-end manager education.[18] Average manager education in the two groups is very similar in the pretreatment period, increases strongly in the event year for the treatment group and remains unchanged in the control group. Figure 1.4b shows the effect of treatment on firm growth. In the pretreatment period, there are no differences in trends or levels between the two groups, which supports the validity of the design. In the event year, growth in the treatment group increases sharply, while growth in the control group follows the pretreatment trend, and the difference between the two groups persists in the years after treatment. To provide a measure of the effect, I average firm growth in the pre and post treatment periods for each event,[19] calculate the

[18] I drop pretreatment manager education from the set of controls in this first step.

[19] I exclude the event year $t = 0$ from this calculation, for the same reason that I exclude years when manager education changes in the full sample analysis. Since I do not know at what moment in the event year the manager change occurred it is unclear which management team is responsible for growth in that year. Including the event year in the post-treatment period does not change the results.

change in average growth Δg_j and then estimate the following equation:

$$\Delta g_j = \alpha_0 + \alpha_1 D_j + `\mathbf{F}_j \tag{1.2}$$

where D denotes treatment and $\mathbf{F}$ the vector of pretreatment controls described above. I find that treatment increases firm growth by 4.15 percentage points ($p < 0.01$). Using the same procedure, I find that treatment increases manager education by 7.29 years of schooling. Dividing the two coefficients implies that a year of manager education increases firm growth by 0.57 percentage points.

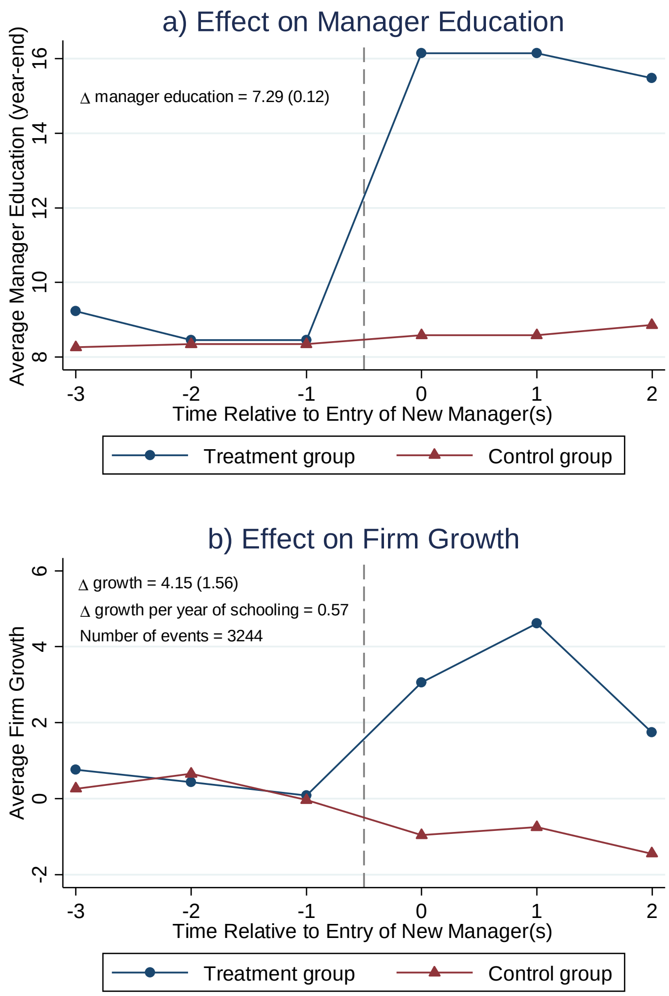
To reinforce the validity of these results, I perform a series of placebo tests by repeating the exercise with changes in average education for occupation groups other than managers. I use the standardized occupational codes in CNP 94 to form four occupation groups besides managers – professionals, office workers, service workers and blue collar workers.[20] I use the same methodology to construct event studies around worker changes in each of these four groups, modifying the set of pretreatment controls accordingly.[21] For professionals I use the same thresholds to define treatment and control groups as for managers: college versus 12th grade or less. For the other groups I adjust them downwards to reflect the level of education in each group. For office and service workers I use 12th grade or more versus 9th grade or less, and for blue collar workers 9th grade or more versus 6th grade or less. Figure 1.5 presents the results. In all four cases treatment leads to a large and significant increase in average education for the respective occupation group, but there is no effect on firm growth.

Full Sample Estimates The event studies examine only a subset of management changes in the data. I now extend the analysis by estimating the fixed effects model in (1.1) for the entire sample. Table 1.3 presents the results. All regressions include firm and year fixed

[20]Professionals correspond to sections 2 and 3 in CNP94, office workers to section 4, service workers to section 5 and blue collar workers to sections 6 through 9.

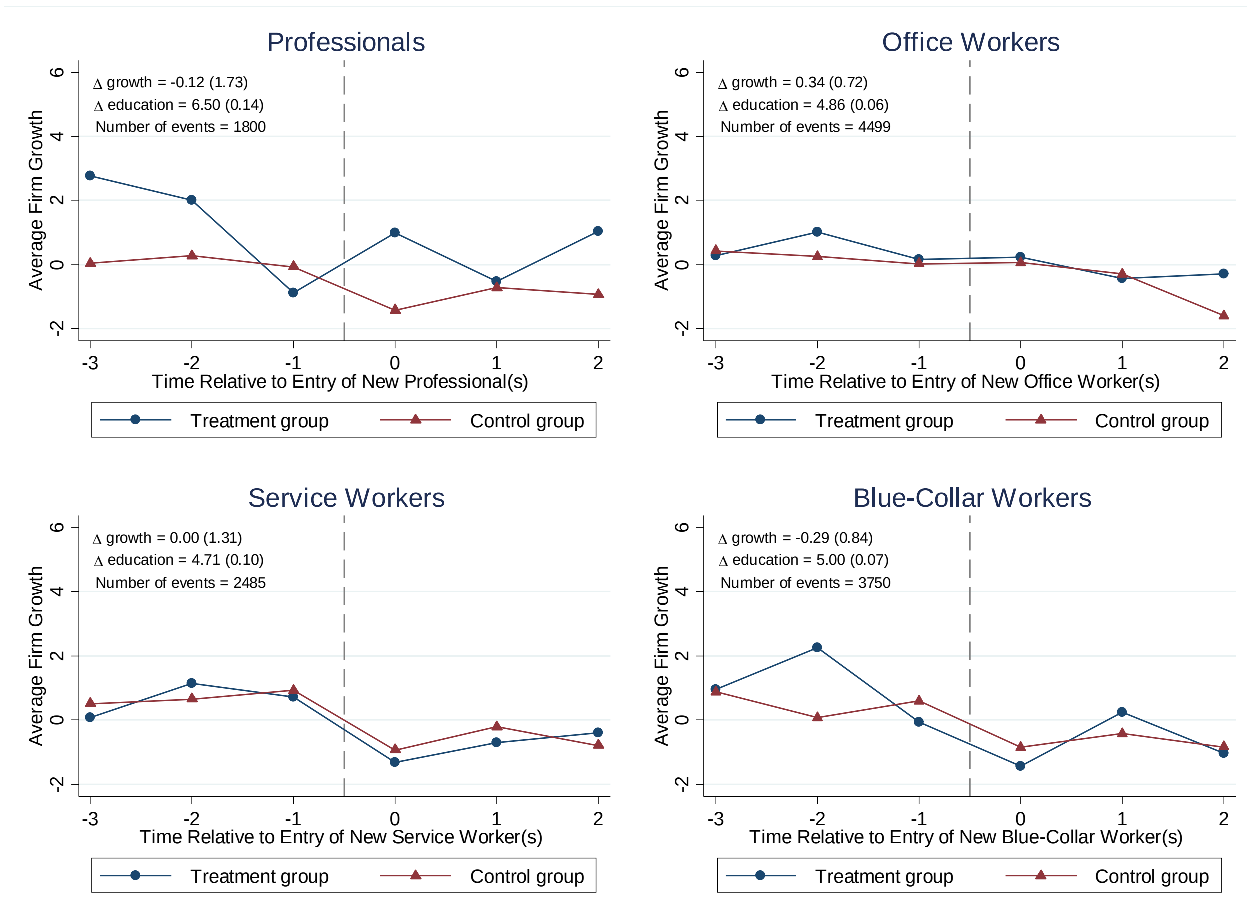
[21]For example, for the event study on professionals I split each firm's workers into professionals and non-professionals, and then change the set of pretreatment controls to include average education and age for professionals and non-professionals, instead of for managers and non-managers.

Figure 1.4: *Event Studies of Manager Changes*



Notes: These figures plot event studies of manager changes. An event is defined as a change in manager education holding the number of managers constant. t=0 denotes the event year and the event window is defined as the three years before and after the event. I further require that manager education be constant for at least one year before and one year after the change, and that the firm be present in the sample between t=-3 and t=2, in order to obtain a balanced sample. The treatment group is defined as firms that hire college-educated managers (545 events) and the control group as firms that hire managers with an average of 12 years of schooling or less (2699 events). The top graph plots the effect of treatment on manager education, and is constructed by regressing manager education at each moment in event time on an indicator for treatment and the pretreatment controls identified in the text, all measured at the end of t=-1. The bottom graph plots the effect on firm growth, and is constructed analogously, adding a quartic in manager education to the set of pretreatment controls. In both graphs, I plot the average outcome for treatment and control groups such that at each moment the difference between the two equals the coefficient on treatment and the weighted average of the two groups equals the sample average. I also report the coefficient on treatment from regressions of the change in manager education (top graph) and the change in firm growth (bottom graph) on an indicator for treatment and the set of pretreatment controls (standard errors in parenthesis). The changes are measured by averaging the pre and post treatment outcomes, excluding t=0, and taking the difference.

Figure 1.5: *Event Studies of Non-Manager Changes*



Notes: These figures plot event studies of non-manager changes. I split non-managers into four occupation groups using the standardized occupational codes in CNP 94 – professionals, office workers, service workers and blue-collar workers. Each graph plots the effect of treatment on firm growth from event studies for the corresponding group, replicating the methodology used for managers (see notes to figure 1.4) and adapting the set of pretreatment controls accordingly (e.g. in the graph for professionals, manager and non-manager education and age are replaced with professional and non-professional education and age). For professionals I use the same thresholds to define treatment and control groups as for managers: college vs. 12th grade or less. For the other groups I adjust them downwards to reflect the level of education in each group. For office and service workers I use 12th grade or more vs. 9th grade or less, and for blue collar workers 9th grade or more vs. 6th grade or less.

effects and standard errors are clustered at the firm level. Starting with the top panel, the baseline specification in column one has no controls. In this specification, a year of manager education increases firm growth by 0.26 percentage points per year. Column two adds a set of firm level controls consisting of quartics in average non-manager education, average manager and non-manager age, firm age, log firm size and log number of managers. The coefficient on manager education increases marginally to 0.27. Column three keeps the controls and replaces year effects with two-digit sector-by-year fixed effects, to account for industry-level shocks, and the coefficient remains 0.27. Column four adds linear firm-specific time trends to the controls from column two, relaxing the parallel trends assumption, and the coefficient drops slightly to 0.23. Finally, column five also keeps the controls and restricts the sample to owner-managed firms, to reduce the scope for endogenous sorting of managers to firms, and the coefficient also drops to 0.23.

All results so far are conditional on survival. The bottom panel in Table 1.3 replicates the specifications in the top panel but accounts for firm exit by assigning a growth rate of -100 percent when a firm exits. In expectation, the effect of manager education becomes larger. In the baseline specification the effect rises to 0.52 percentage points. The addition of controls matters more here, and the coefficient in the second column falls to 0.36. Adding sector-by-year fixed effects makes little difference, yielding a coefficient of 0.37. Including linear firm-specific time trends decreases the coefficient to 0.28. And restricting the sample to owners leaves the effect nearly unchanged, at 0.35. These stronger effects imply that manager education also increases firm survival.

Robustness Checks Table 1.4 presents a series of additional robustness checks. As in Table 1.3, the results in the top panel are conditional on survival while the results in the bottom panel account for firm exit, and all regressions include firm and year fixed effects.

In columns one and two in the top panel, I restrict the analysis to management changes where at least one of the exiting managers leaves the sample permanently.[22] Since the data

[22] I only consider exits up to 2007, so that a manager has to be absent from the sample for at least two years to be considered a permanent exit.

Table 1.3: *Effect of Manager Education on Firm Growth*

	Conditional on Survival				
	(1)	(2)	(3)	(4)	(5)
Manager Education	0.258***	0.270***	0.272***	0.233***	0.233***
	(0.028)	(0.039)	(0.038)	(0.067)	(0.057)
Observations	1383206	1383206	1383204	1383206	1051429
Number of Firms	304622	304622	304622	304622	265452
	In Expectation				
	(1)	(2)	(3)	(4)	(5)
Manager Education	0.523***	0.359***	0.374***	0.280***	0.346***
	(0.035)	(0.043)	(0.043)	(0.073)	(0.063)
Observations	1483279	1483279	1483277	1483279	1129234
Number of Firms	337617	337617	337617	337617	294197
Controls		Yes	Yes	Yes	Yes
Industry-by-Year FEs			Yes		
Firm-specific Trends				Yes	
Owner-Managed Only					Yes

Notes: This table reports regressions of annual firm growth, measured by employment, on average manager years of schooling. The top panel presents results conditional on firm survival, while the bottom panel accounts for firm exit by assigning a growth rate of -100% when a firm exits. All specifications include firm and year fixed effects, except if otherwise noted, and errors are clustered at the firm level. The sample is restricted to firm-years where manager education does not change between the start and end of the year, to minimize measurement error. Column (1) presents a baseline estimate with no controls. Column (2) adds controls consisting of quartics in average non-manager education, average manager and non-manager age, firm age, log firm size and log number of managers. Column (3) keeps the controls and replaces year fixed effects with industry-by-year fixed effects. Column (4) adds firm-specific linear time trends to the specification from Column (2), and Column (5) also keeps the controls and restricts the sample to owner-managed firms.

cover the universe of firms, this implies that with high probability the manager exited the labor force,[23] and therefore that the change is more plausibly orthogonal to shocks to the firm's growth prospects. This strategy parallels the literature that uses deaths to identify the effect of leaders (Johnson *et al.*, 1985; Bennedsen *et al.*, 2007). To isolate the change in manager education that resulted from these exits, I further restrict the sample to the year immediately before and the one immediately after the exit. Despite the much smaller

[23]It is also possible that the exiting manager became a public servant, self employed without any employees, joined the informal economy or was simply unemployed for a long period. In these cases they would no longer be covered by the data.

Table 1.4: *Effect of Manager Education on Firm Growth: Robustness*

	Conditional on Survival							
	Permanent Exits		Owner Education		Revenue Growth		Average Education	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Manager Education	0.265***	0.259***	0.305***	0.321***	0.461***	0.574***	0.162***	0.197***
	(0.084)	(0.095)	(0.053)	(0.042)	(0.084)	(0.099)	(0.025)	(0.034)
Owner Education			-0.039	-0.063**				
			(0.029)	(0.029)				
Observations	148513	148513	1086672	1086672	1108575	1108575	1480203	1480203
Number of Firms	104172	104172	270315	270315	260093	260093	298395	298395
	In Expectation							
	Permanent Exits		Owner Education		Revenue Growth		Average Education	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Manager Education	0.476***	0.337***	0.375***	0.576***	0.474***	0.538***	0.475***	0.553***
	(0.097)	(0.111)	(0.061)	(0.051)	(0.077)	(0.092)	(0.030)	(0.038)
Owner Education			-0.021	-0.021				
			(0.035)	(0.034)				
Observations	152553	152553	1167242	1167242	1108657	1108657	1589529	1589529
Number of Firms	105526	105526	300072	300072	260136	260136	332229	332229
Controls		Yes		Yes		Yes		Yes

Notes: This table presents regressions of annual firm growth on average manager years of schooling. The top panel presents results conditional on firm survival, while the bottom panel accounts for firm exit by assigning a growth rate of -100% when a firm exits. All specifications include firm and year fixed effects. Errors are clustered at the firm level. In columns (1) to (6) the sample is restricted to firm-years where manager education does not change between the start and end of the year, to minimize measurement error. Columns (1) and (2) restrict the sample to the year before and the year after a management change where at least one manager leaves the sample permanently. Columns (3) and (4) add owner education. Columns (5) and (6) use revenue rather than employment growth. Columns (7) and (8) remove the restriction to firm-years where manager education does not change, and measure manager education as the average of start and year-end values. Columns (1), (3), (5) and (7) use the baseline specification without controls. Columns (2), (4), (6) and (8) add the set of firm-level controls from table 1.3.

sample, the coefficient on manager education remains stable and highly significant: 0.27 without controls in column one and 0.26 with controls in column two.

A possible interpretation of the results is that manager education is proxying for lower financial constraints, a prominent alternative explanation for differences in firm growth. If this were the case, then in firms where the owner and manager are different people it should be the owner's education that matters. Columns three and four test and reject this hypothesis by adding owner education to the regressions. The coefficient on manager education rises slightly, to 0.32 without controls and 0.31 with controls, and the coefficient on owner education is small and actually negative.

Columns five and six show that the results are robust to using revenue growth as the outcome, also winsorized at the 99^{th} percentile, rather than employment growth. In fact the magnitudes are larger. In column five, without controls, a year of manager education increases revenue growth by 0.46 percentage points. Adding controls in column six increases the coefficient to 0.57.

All regressions so far restrict the sample to years when manager education does not change between the start and end of the year, to minimize measurement error. In columns seven and eight I remove this restriction and measure manager education as the average of start and year-end values.[24] For consistency, I do the same for non-manager education and for manager and non-manager age. As expected, the coefficient on manager education falls in this specification, to 0.16 without controls in column seven and 0.20 with controls in column eight, but remains highly significant.

The bottom panel replicates these specifications while accounting for firm exit, and finds the same patterns.

1.4.2 Is Education the Cause?

The results presented up to this point show that more educated managers increase firm growth, but not whether the cause is education itself or unobserved manager characteristics like talent, ambition or family background that education could be proxying for. Even if these characteristics were uncorrelated with education in the population, there could be a selection issue: agents with more education may have a better outside option in other occupations and may only choose to become managers if they are especially talented.

Methodology Suppose that besides schooling s, agents are endowed with managerial ability a^m and non-managerial ability a^l, where "ability" is shorthand for unobserved characteristics like talent, ambition or family background. If a^m also affects firm growth

[24]When a firm exits during the year, I use the starting value only.

then (1.1) becomes

$$g_{jt} = \alpha_j + \alpha_1 s_{jt} + \alpha_2 a^m_{jt} + \ell \mathbf{F}_{jt} + \delta_t + \varepsilon_{jt} \tag{1.3}$$

In the background, agents sort into management if their income as managers exceeds their outside option as non-managers. I assume that this outside option is given by the standard Mincerian wage equation

$$\ln w_{jt} = \beta_0 + \beta_1 s_{jt} + \beta_2 a^l_{jt} \tag{1.4}$$

where $\ln w_{jt}$ is the log of the outside option for the manager of firm j and time t, and where I omit experience to save on notation but will account for it below. Let the linear projection of a^m on a^l be given by

$$a^m_{jt} = \rho_0 + \rho_1 a^l_{jt} + \eta_{jt} \tag{1.5}$$

Using (1.4) and (1.5), (1.3) can be rewritten as

$$g_{jt} = \alpha'_j + (\alpha_1 - \frac{\alpha_2\rho_1\beta_1}{\beta_2}) s_{jt} + \frac{\alpha_2\rho_1}{\beta_2} \ln w_{jt} + \ell \mathbf{F}_{jt} + \delta_t + \alpha_2\eta_{jt} + \varepsilon_{jt} \tag{1.6}$$

where $\alpha'_j \equiv \alpha_j - \frac{\alpha_2\rho_1\beta_0}{\beta_2} + \alpha_2\rho_0$. This expression shows that the outside option $\ln w_{jt}$ can be used as a proxy control for a^m_{jt}. This addresses the bias from omitted a^m_{jt} while introducing another bias, since $\ln w_{jt}$ is partly determined by schooling s_{jt}. But unlike the original bias the new bias is plausibly negative. It is captured by the $\frac{\alpha_2\rho_1\beta_1}{\beta_2}$ term in the coefficient on s_{jt}. α_2 and β_2 are the effects of a^m and a^l on firm growth and wages, respectively, which are positive by definition. ρ_1 will be positive if the two abilities are positively correlated, as can reasonably be expected. A test of this condition is that the coefficient on $\ln w_i$ in (1.6) should be positive. Finally, β_1 is the return to schooling in the wage equation, which a large literature has shown is also positive (e.g. Card, 1999). Intuitively, conditional on $\ln w_{jt}$, higher s_{jt} implies lower a^l_{jt}, which in turn should imply lower a^m_{jt} and therefore lower g_{jt}. It follows that the coefficient on schooling in this specification represents a lower bound on the true coefficient α_1. In addition, the bias term is equal to the coefficient on $\ln w_{jt}$ multiplied by β_1. Since β_1 is a parameter that has been well identified in the literature, I can draw on reasonable estimates of β_1 to recover the bias and obtain an estimate of the true coefficient α_1.

The key assumption in this approach is that $corr(s_{jt}, \eta_{jt}) = 0$, meaning that there are no components of ability that increase firm growth and are uncorrelated with the manager's outside option but are correlated with schooling. While this might not hold exactly, it seems reasonable to argue that any such residual components of ability should not lead to first order biases. For example, it is likely that important unobserved characteristics like talent, ambition or family background affect both firm growth and the outside option. Moreover, since the outside option represents the manager's best alternative occupation it is likely that the mix of abilities required in that particular occupation is not too distant from the mix required in the managerial position that the agent chose.

In order to implement this design, data on managers' outside option is required. For this purpose I use a sample of switchers – people who have worked as managers and non-managers within the sample period.[25] In this sample, I observe a manager's income when working as a non-manager at other points in time, and I take the average of observed non-managerial income, residualized on year and experience dummies, as the manager's outside option. One concern with this procedure could be measurement error in the outside option, which might attenuate the coefficient on $\ln w_{it}$ and amplify the coefficient on s_i. But measurement error would attenuate the bias correction for the schooling coefficient in the same proportion, so that the bias-corrected estimate of α_1 would be unaffected.[26] As long as measurement error is not correlated with schooling, it should not affect the results.

[25] This sub-sample comprises about 20 percent of the sample of single-manager firms. I restrict the analysis to single-manager firms in order to simplify interpretation.

[26] With measurement error in $\ln w_{jt}$ given by u_{jt}, the estimated coefficient on manager schooling in (1.6) would equal $\alpha_1 - \frac{\alpha_2\rho_1\beta_1}{\beta_2} + \beta_1\frac{\alpha_2\rho_1}{\beta_2}(1 - \frac{\sigma_e^2}{\sigma_e^2+\sigma_u^2})$, where σ_e^2 is the variance of the residual in a regression of $\ln w_{jt}$ on the remaining covariates in (1.6). The coefficient on $\ln w_{jt}$ would equal $\frac{\alpha_2\rho_1}{\beta_2}\frac{\sigma_e^2}{\sigma_e^2+\sigma_u^2}$. Applying the bias correction to the coefficient on manager schooling would therefore lead to

$$\alpha_1 - \frac{\alpha_2\rho_1\beta_1}{\beta_2} + \beta_1\frac{\alpha_2\rho_1}{\beta_2}(1 - \frac{\sigma_e^2}{\sigma_e^2+\sigma_u^2}) + \beta_1\frac{\alpha_2\rho_1}{\beta_2}\frac{\sigma_e^2}{\sigma_e^2+\sigma_u^2} = \alpha_1 \quad (1.7)$$

Results Table 1.5 presents the findings. The first column adds the log of the manager's outside option to the specification from column two of table 1.3, which includes firm and year fixed effects as well as firm-level controls. As expected, the coefficient on the outside option is positive and significant, which validates the assumption of a positive ρ_1. The coefficient on manager schooling is 0.46 percentage points, and also remains significant. This represents a lower bound for the true effect in this sample. The last row in the table presents the bias-corrected coefficient α_1, which equals 0.57 percentage points. I assume a conservative estimate of 6 percent for the returns to schooling parameter β_1, but the results are not very sensitive to this choice. Assuming a value of 10 percent, in the upper end of the typical range in the literature, increases α_1 to 0.65 percentage points. For comparison, the second column uses the same specification and sample without the outside option control. Again as predicted, the coefficient on schooling is higher in this specification. In fact, it rises to 0.53 percentage points, close to the lower end for the bias adjusted coefficient, suggesting that the bias from omitted ability in the results described above might be minimal.

Table 1.5: *Effect of Manager Education on Firm Growth: Outside Option*

	(1)	(2)	(3)	(4)
Manager Education	0.456***	0.530***	0.526*	0.668**
	(0.163)	(0.160)	(0.317)	(0.314)
Log Outside Option	1.940**		4.161**	
	(0.876)		(1.761)	
Implied α_1 ($\beta_1 = 6\%$)	0.5721***		0.7751**	
	0.1616		0.3158	
Observations	181395	181395	119937	119937
Number of Firms	76359	76359	55635	55635
Owner-Managed Only			Yes	Yes

Notes: This table presents regressions of annual firm growth on average manager years of schooling and the manager's outside option, as measured by log wages while working in different occupations at other points in time. The implied α_1 is the bias-corrected coefficient on manager education, assuming a labor market return to schooling of 6%. See main text for details. All specifications include firm and year fixed effects and the set of firm-level controls from table 1.3. Errors are clustered at the firm level. In columns (3) and (4) the sample is restricted to owner-managed firms.

As a simple robustness check, the third and fourth columns restrict the sample to owner-managed firms, as in column five of table 1.3, and find the same pattern. The coefficient on manager schooling equals 0.53 when the manager's outside option is added, and the coefficient on the outside option is positive. The bias adjusted coefficient is 0.78, while the coefficient without the outside option control equals 0.67.

1.5 Implications for Lifecycle Growth

Does the effect of manager education on annual firm growth estimated in section 2.4 translate into large differences in firm size over the lifecycle? And can it account for the size differences observed in the data and described in section 1.3?

First, in order to account for the large size differences in the data, manager education must be a persistent firm characteristic. It turns out to be extremely persistent, with one and ten year autocorrelations of 0.97 and 0.82, respectively. Non-manager education is also strongly autocorrelated, but to a lesser extent, with one and ten year autocorrelations of 0.90 and 0.62.

I perform a simple simulation to evaluate the effect of manager education on lifecycle growth. I consider two firms, one whose manager has a college degree (17 years of schooling) and one with a primary-school-educated manager (4 years of schooling). This choice facilitates comparison with the top and bottom groups in figure 1.1a, which have an average of 16.7 and 4.1 years of manager schooling respectively. Firms in the top group are just 9 percent larger at entry than firms in the bottom group. I take this difference in starting sizes and use the estimated effect of manager education on firm growth from section 2.4 to simulate the relative size difference between the two firms at different ages. The coefficient on manager education in the top panel of Table 1.3 ranges from 0.23 to 0.27. I assume 0.25 for the purposes of this simulation, which implies a difference in annual growth rates between the two firms of $(17 - 4) \times 0.25 = 3.25$ percentage points per year. Table 1.6 shows the results. For firms aged 11 to 15, which corresponds to the average age in the sample, the simulation yields a 1.64 size difference versus 1.81 in the data. For firms aged 31 to 35, the simulation predicts a three-fold difference versus a four-fold difference in the data. Overall the simulation shows that the estimated effect of manager education on annual firm growth can lead to substantial differences in size over the lifecycle, and that it can explain most of the variation observed in the cross-sectional relationship between manager education and firm size.

Table 1.6: *Relative Firm Size: College vs Primary-School Management*

Firm Age Bin	Simulation	Data	Percent Explained
1 to 5	1.19	1.26	94
6 to 10	1.40	1.59	88
11 to 15	1.64	1.81	91
16 to 20	1.92	2.57	75
21 to 25	2.22	2.72	82
26 to 30	2.64	2.88	92
31 to 35	3.06	4.00	77
36 to 40	3.59	5.76	62
41 to 45	4.18	7.89	53
46 to 50	4.95	7.25	68
51 and over	9.00	8.49	106

Notes: This table compares simulated firm sizes using the estimated effect of manager education on firm growth to observed differences in firm size. The effect of manager education on annual growth is set at 0.25 percentage points per year of schooling. I simulate firm size over the lifecycle for two firms, one with college-educated managers (17 years of schooling) and another with primary-school-educated managers (4 years of schooling). These two firms are chosen for comparison with the top and bottom groups of manager education in figure 1.1a, which have an average of 16.7 and 4.1 years of manager schooling respectively. The simulation column presents the ratio of simulated sizes (college/primary-school) in different age bins. The age used for the simulation in each bin is the average age for firms in the top and bottom groups in figure 1.1a in that age bin. The data column presents the same ratio in the data, comparing the top and bottom groups in figure 1.1a.

1.6 Conclusion

Understanding the sources of firm growth is an open puzzle. Several theories have been offered in the literature, but empirical evidence to support them has been limited. This paper presents evidence in support of one of these theories. It finds that management quality, and specifically manager education, has a large effect on firm growth over the lifecycle.

What mechanisms drive the effect of more educated managers on growth is an important question for future research. An understanding of what skills are acquired through education would also be relevant, as well as an analysis of the role of business-specific vs. general education.

These results capture the effect of manager education at the firm level. At the aggregate level, increases in manager education may also strengthen selection, that is, the extent to which more productive firms grow and drive less productive firms to exit. Such an effect could help explain why more developed countries have larger firms on average, and play a key role in explaining differences in aggregate productivity across countries. Investigating this hypothesis is also a priority for future research.

Chapter 2

Knowledge and Growth: Evidence from Early Modern Europe

2.1 Introduction

Since the 1950s economists have attributed a central role to technology as a driver of economic growth (Abramovitz, 1956; Kendrick, 1956; Solow, 1957). Technology in this context takes on a broad meaning that includes any knowledge useful in production, from ideas for new products and services, to more efficient production processes, to better management. Given the difficulty of measuring knowledge in this encompassing sense, direct evidence on this relationship has mostly come from studies of corporate research and development expenditures (R&D) and productivity growth.[1] It is generally acknowledged, however, that R&D captures only a fraction of knowledge creation.[2] Moreover, these studies have found that the relationship between R&D and productivity is robust in the

[1]Seminal contributions in this field include Mansfield (1965), Griliches (1980, 1986) and Griliches and Mairesse (1984). Patents and patent citations have also been extensively used in the study of knowledge flows and spillovers (Jaffe *et al.*, 1993), but patent counts are not strong predictors of economic performance (Pakes, 1985; Griliches *et al.*, 1987).

[2]Measured R&D is typically the preserve of large corporations. Entrepreneurship and academic research are two major sources of innovation not well captured by R&D. Denison (1985) estimates that R&D accounts for about 20 percent of knowledge accumulation.

cross-sectional dimension only.[3]

One idea for a broader measure of knowledge, introduced by Baten and van Zanden (2008), is book production. But while books are a key tool in the transmission of economically useful knowledge, they also serve other purposes. Obvious examples include the diffusion of religious and cultural norms, political debate, artistic expression or entertainment. The aggregate book counts used by Baten and van Zanden (2008) are therefore open to different interpretations. They could be a better proxy for the level of literacy or consumption than for the type of knowledge economists would describe as technology.

This paper investigates the relationship between knowledge and growth at the city level using a new database of 5.5 million books published in ten European countries[4] from 1450 to 1800. The database consists of individual book data as recorded by WorldCat, a network of over 72,000 library catalogs around the world. The WorldCat network includes the national libraries of the ten countries in the sample, among other major national and research libraries, which implies that its combined catalogs plausibly contain all surviving book titles published in these countries. Crucially, the data contain information on the book's subject for a sample of 12 percent of book titles, enabling me to distinguish economically useful knowledge – e.g. books on technology – from literacy or consumption – e.g. religious books and literature.

The database covers the so-called hand-press period, between Gutenberg's invention of the printing press and the introduction of industrial printing presses in the early 19th century, and the historical context is important for several reasons. First, unlike in today's digital age, book production was arguably the main medium for knowledge transmission in this period. Second, local book production was a reasonable proxy for locally available

[3]R&D and productivity are positively related across firms, but within-firm variation in R&D is not robustly associated with productivity growth. Klette and Kortum (2004) label these findings as their stylized fact number one in a review of the evidence on firm-level innovation. See, for example, Mairesse and Sassenou (1991) for a survey.

[4]The data covers western European countries whose national library is part of the WorldCat network: Denmark, England, France, Germany, Ireland, Netherlands, Scotland, Spain, Sweden and Switzerland. Countries are defined by their modern borders.

knowledge. Dittmar (2011) presents evidence that transport costs for books were high in this period, and that it was more common for books to spread through reprints than inter-city trade. Grendler (1988) adds that lack of copyright protection kept print runs small and also encouraged local reprints as a diffusion mechanism.[5] Third, the data include separate records for different "manifestations" of a book, which include new editions and reprints. This implies that differences in diffusion across books are well captured by the data.[6] Finally, the use of an extended time period allows me to account for long lags in the effect of knowledge on output while still observing significant variation in knowledge diffusion over time, which allows me to identify this effect from within-city variation alone.[7]

I start by presenting basic trends in book production over time, both across countries and subjects. Switzerland and Germany were the early leaders in printing, both playing an important role in the diffusion of the Reformation in the early 16th century. Germany experienced a sharp fall in production during the Thirty Years' war (1618-1648), and England and the Netherlands rose to the top in the 17th century. France was behind the leading countries throughout most of the period, and caught up at the onset of the French revolution, while Spain had low levels of production throughout the entire period. Scotland, Denmark and Sweden were somewhere between the leading countries and France, and Ireland had virtually no production until the 17th century but grew steadily after that. Across subjects, early printing was dominated by religion and literature. Over time, some subjects – history, science, medicine, law – saw their share increase steadily, while other subjects only took off in the 18th century, such as technology, business and finance, agriculture and social science.

I then turn to estimating the effect of book production on city growth. I use population

[5] An an example, Grendler (1988) notes that "from 1509 to 1520, Erasmus' *Moriae encomium* appeared in at least thirty-five editions in nine different cities, printed by fourteen or more publishers. The places of publication were Antwerp, Basle (two publishers), Cologne, Florence, Mainz, Paris (four publishers), Sélestat, Strasburg and Venice (two publishers). Reprinting was almost routine at a time when the concept of literary property or binding commercial restrictions hardly existed".

[6] This point is also made by Grendler (1988), who writes that "the number of reprints, if any, presents a truer measure of a book's diffusion [than initial print runs]".

[7] Short timeframes coupled with the strong autocorrelation of R&D expenditures at the firm level have been problematic in studies that try to identify the effect of R&D on firm productivity from within-firm variation (Mairesse and Sassenou, 1991).

growth as a measure of economic growth at the city level, following standard practice in both historical and modern contexts (De Long and Shleifer, 1993; Glaeser *et al.*, 1995).[8] Exploiting differential within-city variation, and accounting for shocks at the country level, I find that book production has a strong effect on subsequent population growth. The results are robust to a variety of specifications, including alternative assumptions about the rate of incorporation of knowledge into output and the inclusion of city-specific time trends. The coefficient is also stable across time periods and robust to splitting the sample by country, with the exception of England where publishing was exceptionally concentrated in London.

Next, I distinguish between different interpretations of this effect by repeating the estimation by subject. Books on technology, finance, medicine and history are robustly related with growth, with technology and finance having the strongest effect. The first three subjects clearly capture the production and diffusion of economically useful knowledge. History includes books about current events at the time they were written, and one interpretation is that it captures the existence of a literate elite involved in or at least informed about public life. In addition, I show that although science books as a whole are insignificant, books in chemistry and geology in particular also increase growth in this period, with a coefficient similar in magnitude to that of technology and finance. This is consistent with the importance of the 18th-century Chemical Revolution (Clow and Clow, 1952) and coal mining as drivers of the Industrial Revolution. Books on all other topics, such as religion or literature, are not robustly associated with growth. Overall, these findings suggest that the effect of book production reflects the diffusion of knowledge rather than literacy or consumption.

This paper contributes to two literatures. Within the literature on innovation and growth, it helps bridge the gap between the broad measure of knowledge emphasized in growth theory (Romer, 1990; Grossman and Helpman, 1991; Aghion and Howitt, 1992) and the narrower but quantifiable proxies that have been used in empirical studies, chiefly R&D

[8]The underlying logic is that productivity shocks trigger migration towards more productive cities to the point that potential migrants remain indifferent across locations, with rising house prices limiting the extent of migration (Roback, 1982; Glaeser and Gottlieb, 2009)

(Mansfield, 1965; Griliches, 1980, 1986; Griliches and Mairesse, 1984), and also patents (Pakes, 1985; Griliches *et al.*, 1987; Caballero and Jaffe, 1993) and scientific papers (Adams, 1990). By exploiting the long run of history, it also accounts for long lags in the effect of knowledge on growth and finds robust effects within cities rather than in a cross-sectional design.

The other literature the paper contributes to is that on the historical role of knowledge and human capital before and during the Industrial Revolution. Dittmar (2011) shows that early adoption of the printing press had a sizable effect on city growth in the 16th century. Cantoni and Yuchtman (2014) find that the creation of universities increased growth, as measured by the establishment of markets, in medieval Germany. The predominant view, however, is that scientific knowledge (Mathias, 1969; Hall, 1974) and human capital (Mitch, 1993; Allen, 2003; Galor, 2005) did not become engines of growth before the second phase of industrialization, in the second half of the 19th century. Another view (Musson and Robinson, 1969; Mokyr, 2005a) argues that, while scientific breakthroughs might have played a limited role in this period, the scientific mindset that spread among elites in the 18th century set off an accumulation of useful knowledge, consisting of "catalogs of facts, based on experience and experiment rather than on understanding or careful analysis and testing" (Mokyr, 2005a), which in turn drove industrialization. Mokyr (2005b) also distinguishes the roles of average and upper-tail human capital, and argues that the latter – scientifically knowledgeable entrepreneurs and engineers – was key in the diffusion and application of useful knowledge in production. Squicciarini and Voigtländer (2014) provide evidence in favor of this second view by showing that the density of encyclopedia subscriptions in French cities, a proxy for the presence of upper-tail human capital, increased city growth from 1750 to 1850, but find no effect on earlier growth.

This paper also offers support for the second view. First, it confirms that there was a sharp increase in books on science and technology in the 18th century. Second, it shows that while some scientific fields – notably physics and biology – did not increase growth, in line with the predominant view, books on technology, chemistry and geology did have a strong effect on growth. Third, it finds that knowledge increased growth long before the

Industrial Revolution, in line with Dittmar (2011) and unlike Squicciarini and Voigtländer (2014). And fourth, it is the first paper to my knowledge to provide evidence on the effect of knowledge and upper-tail human capital on city growth by exploiting within-city rather than cross-sectional variation as in Dittmar (2011) or Squicciarini and Voigtländer (2014).

Lastly, the paper also relates specifically to the literature on the history of book production. Relative to Baten and van Zanden (2008) and Buringh and Van Zanden (2009), who report a total of just under 630 thousand books published in all of Europe in this period, this paper introduces significantly more comprehensive data on book production, covering 5.5 million books published in the ten countries in the sample alone. In addition, I use a single source of data, whereas Baten and van Zanden (2008) and Buringh and Van Zanden (2009) use multiple sources and make important adjustments to book totals based on their perception of the comprehensiveness of each source.

The rest of the paper is organized as follows. Section 2.2 describes the data. Section 2.3 presents trends in book production. Section 2.4 outlines the empirical methodology and reports the main findings. Section 2.5 concludes.

2.2 Data

2.2.1 Book Records

The paper uses a new data set constructed from book records drawn from WorldCat, a collection of over 72,000 library catalogs in 170 countries, including most major libraries around the world.[9] WorldCat is produced and maintained by the Online Computer Library Center cooperative, and it is the world's largest bibliographic database, with over 330 million records of books, periodicals, visual materials and sound recordings, among other document types.

Each record corresponds to a "manifestation" of a work. A manifestation could be an

[9] WorldCat covers 17 of the 19 largest libraries in the world, including all top 10, as listed in Wikipedia. (https://en.wikipedia.org/wiki/List_of_largest_libraries)

original title, a new edition or a reprint. WorldCat assigns a unique identifier to each record, and implements matching algorithms to eliminate duplicates. Among document types, books are defined as "books, pamphlets, technical reports, typescripts, theses, dissertations, manuscripts and other written works".

Book records contain information on the book's title, author, place and year of publication and language among other details. Importantly for the purposes of this paper, some book records contain information on the book's subject, which will be described in greater detail below. For illustration, figure 2.1 displays the record for the first major book printed in Europe, Gutenberg's Bible. The record shows that this book was printed in Mainz in either 1454 or 1455. Whenever a record indicates a range rather than a specific year of publication, as in this case, I assign an equal fraction of the book to each year in the range when computing annual book totals.

2.2.2 City Population

City population comes from Bairoch *et al.* (1988), henceforth the Bairoch data set, following standard practice in studies of city growth in early modern Europe. This data set covers 2,204 European cities that had a population of at least five thousand at any point between 800 and 1800, and it reports city population, when available, every 100 years up to 1700 and every 50 years after that until 1850.

2.2.3 Sample Definition

I focus the analysis on books published between 1450 and 1800 in Western European countries whose national library is part of WorldCat. This implies that the combined WorldCat catalogs plausibly contain all surviving book titles published in these countries. The ten countries that meet this criteria are Denmark, England, France, Germany, Ireland, Netherlands, Scotland, Spain, Sweden and Switzerland.[10] This leads to a set of just over six

[10]One important omission is Italy, an important center of early printing (Dittmar, 2011) whose national libraries of Rome and Florence are absent from WorldCat at present.

Figure 2.1: *WorldCat Book Record Example*

Title:	**Biblia latina.**
Author(s):	Gutenberg, Johann,; 1397?-1468.
Publication:	[Mainz], [Printer of the 42-line Bible (Johann Gutenberg)]
Year:	1454-1455?
Description:	[643] leaves (blank leaves [642] & [643] lacking), bound in 2 volumes 38-40 cm
Language:	Latin
Contents:	[1] Leaf [396], containing pts. of chapters III and IV of Jeremiah. Part of columns at outer margin cut off. Rubricated. Bound in ms. leaves.--[2] Leaf [454], containing pts of chap. XLVII and chap. XLVIII of Ezekiel. Part of columns at inner margin and first line of each column cut off.--[3] Leaf [461], containing pts. of chapters VII and VIII of Daniel. Part of columns at inner margin cut off. Gift of Mr. Batchelder.
Standard No:	**LCCN:** 52-51757
	SUBJECT(S)
Genre/Form:	Incunabula -- Germany -- Mainz -- 1454.
Note(s):	On paper. Leaves 1-34, 129-158, and 261 are of the first setting. Initial strokes supplied in red on leaves [1]-[130] and on a few other scattered leaves, headlines in alternate red and blue letters, initials in red or blue, some of the larger ones in red and blue, a few with ornaments extending to the margin./ V. 1: 37.4 x 27.3 cm. v. 2: 39.1 x 27.4 cm./ Binding 16-17th century brown polished blind-stamped calf over wooden boards, v. 1 dated 1600, both volumes rebacked in leather approximately of the same color./ Provenance: James Perry, Duke of Sussex, Bishop of Cashel, Earl of Crawford, Earl of Carysfort, Carl H. Pforzheimer./ De Ricci 8. Lazare 40. Norman 41.
Class Descriptors:	**LC:** BS75
Other Titles:	Bible. Latin. Mainz. Gutenberg (42 lines). 1454.
Document Type:	Book
Entry:	19790427
Update:	20140803
Accession No:	**OCLC:** 4904654

Notes: This figure displays the WorldCat record for the first major book printed in Europe, Gutenberg's Bible. Data on book production are constructed from book records such as this one. The fields used from each record are Publication, which includes the city where the book was printed, Year and Class Descriptors, which includes codes from standard library classification systems that I use to identify the book's subject, when available. In this case the Class Descriptors field includes the Library of Congress Classification BS75, which corresponds to class B "Philosophy, Psychology, Religion" and subclass BS "The Bible". Country of publication is not displayed in book records but is available as a search field, and I obtain it by performing country-specific searches when constructing the database.

million book records. From each record, I collect information on the book's country, place and year of publication, and on its subject, when available.

2.2.4 Matching Books and Cities

I then match these records with the cities in the Bairoch data set. There are two main challenges in the matching process. First, many cities appear under multiple names in book records, including variations in spelling and names in different languages. Latin city names, for example, are common. The Bairoch data includes a set of alternative city names which I also use in the matching process, but these only mitigate the problem slightly. I overcome this issue by using variant city names from the Consortium of European Research Libraries (CERL) Thesaurus, a dataset that contains variations in the names of print locations up to the mid-nineteenth century. To illustrate the severity of the problem, the city of Berlin has 58 alternative names in the CERL Thesaurus data, some of which are recognizable, like Berlinium, while others are less so, like Colonia Marchica. The second challenge is that city names are often misspelled, or appear within longer strings of text in the record's place of publication field. In addition, a few records indicate multiple cities, which I take to mean that the book was published in all cities mentioned. To address these issues, I take every sequence of up to four words in the place of publication field and match it with city names using a fuzzy string matching procedure based on the Ratcliff-Obershelp algorithm[11]. If there are multiple city matches for a given book record, then I assign that book to each of the matched cities.

The matching procedure yields at least one city match for 5.5 million records, or 91 percent of all books published in the 10 countries in the sample, and these records constitute the analysis sample for this paper. The unmatched nine percent fall almost exclusively under two categories: either the place of publication is missing, or the city and country

[11]This algorithm computes the similarity of two strings as twice the number of matching characters divided by the total number of characters in each string. Matching characters are those in the longest common subsequence of characters plus, recursively, matching characters in the unmatched region on either side of the longest common subsequence. I implement this algorithm using a Python package named FuzzyWuzzy, which is a wrapper for the difflib module in Python's standard library.

reported do not match, e.g. the country of publication field reports "France" while the place of publication field reports "Amsterdam".[12] Overall, the outcome of the matching process confirms that book printing was highly concentrated in cities, as argued by Dittmar (2011).

2.2.5 Book Subjects

A subset of book records include classification codes from standard library classification systems. The classification systems present in the data are the Library of Congress Classification (LCC), the Dewey Decimal Classification (DDC), the Universal Decimal Classification (UDC), the National Library of Medicine classification (NLM) and the National Agricultural Library classification (NAL). For example, the book record in figure 2.1 reports the LCC code BS75, which corresponds to class B "Philosophy, Psychology, Religion" and subclass BSŸhe Bible". Availability of these codes is presumably related to whether the library contributing the record uses one of these systems or not. Academic libraries in the U.S. and U.K. tend to use the LCC system, while public and school libraries tend to use DDC. In continental Europe, libraries tend to use DDC or UDC, but the use of standard classification systems is significantly less widespread than in the U.S. and U.K.

Overall, 12 percent of book records in the sample include at least one code from either the LCC, DDC or UDC systems. The NLM and NAL would add an additional one percent, but I exclude these subject-specific systems to avoid biasing the subject composition in the sample.[13] The most commonly reported code is the LCC, covering 10 percent of records. The DDC is next with 2 percent, and the UDC contributes less than 1 percent. When more than one code is available for a given record I take the LCC if available, and the DDC otherwise.

[12]Regarding the second case, some controversial books in this period were published under fictitious publisher names from a different city, often abroad, in order to protect the real publishers. A famous case is that of the publisher Pierre Marteau from Cologne. It was invented by exiled French printers in the Netherlands in the 17th century, and used by publishers in France and the Netherlands, and later in Germany, mostly in order to publish political satire (Brouillant, 1888). In the 18th century a supposed line of family members kept the label flourishing, with books published by Marteau's widow and heirs. In light of this, it is possible that for some books the correct country of publication is known and reported, along with the fictitious publisher and city. A WorldCat search for books published by Pierre Marteau in the Netherlands and France yields several examples of such cases.

[13]Including them does not change any of the results in the paper.

I use these codes to define 14 subjects: religion, literature, arts, history, language, philosophy, general, science, medicine, technology, agriculture, social science, business and finance, and law. These subjects correspond broadly to LCC, DDC and UDC first or second level divisions. Table 2.1 provides the correspondence between subjects and LCC, DDC and UDC codes. Subject labels are self-explanatory with a couple of exceptions. The most important one is history. Many books under this label were accounts of current events when they were published. As will be seen below, there were spikes in publications classified as history during major historical events like civil wars and revolutions. It is therefore more appropriate to think of this label as a combination of history and current events. The general label includes reference works, such as dictionaries, encyclopedias, indices, almanacs, bibliographies, and collections of periodicals.

Subject coverage is not even across countries. In particular, because libraries in the U.S. and U.K. are more likely to use standard classification systems, books published in England, Scotland and Ireland are more likely to have subject information than others. For example, 21 percent of books published in England have subject information, while in France and Germany that number falls to 11 and 9 percent respectively. Year of publication, on the other hand, is only weakly related to subject availability,[14] but books published during particular historical episodes, such as the English civil war, appear to be more likely to have subject information. In order to account for these biases in subject coverage, I adjust all subject-specific book counts by multiplying the raw count by the ratio of total books to books with subject information at the country-year level.

2.3 Trends in Book Production

In order to summarize the data, I start by presenting trends in book production, first across countries and then across subjects.

[14]In some countries the relationship is positive and in others it is negative, but in all cases its magnitude is small.

Table 2.1: *Classification of Book Subjects*

Subject	Library of Congress	Dewey Decimal Class.	Universal Dec. Class.
Religion	Subclasses BL to BX	Class 2	Class 2
Literature	Subclasses AC, PA3000 to PA8999, PN to PZ	Class 8 and division 08	Class 8 (except divisions 80, 81) and division 08
Arts	Classes M, N	Class 7	Class 7
History	Classes C (except subclass CE), D, E, F and subclass G	Class 9	Class 9
Language	Subclasses P to PM (except subclass PA3000 to PA8999)	Class 4	Divisions 80, 81
Philosophy	Subclasses B to BJ	Class 1 except division 15	Class 1 except division 15
General	Class Z and subclasses AE to AZ	Class 0 except division 08	Class 0 except division 08
Science	Class Q and subclass CE	Class 5 (except section 526) and sections 611, 612	Class 5 (except subdivision 528) and subdivisions 611, 612
Medicine	Class R	Division 61 except sections 611, 612	Division 61 except subdivisions 611, 612
Technology	Class T	Class 6 except divisions 61, 63, 65	Class 6 except divisions 61, 63, 65
Agriculture	Class S	Division 63	Division 63
Social Science	Classes G (except subclass G), H (except subclasses HE to HJ), J, L, U, V	Class 3 (except divisions 34, 38 and sections 332, 336), division 15 and section 526	Class 3 (except division 34 and subdivisions 336, 339), division 15 and subdivision 528
Business and Finance	Subclasses HE to HJ	Divisions 38, 65 and sections 332, 336	Division 65 and subdivisions 336 and 339
Law	Class K	Division 34	Division 34

Notes: This table presents the correspondence between the 14 book subjects used in the paper and the underlying codes in the Library of Congress Classification (LCC), the Dewey Decimal Classification (DDC) and the Universal Decimal Classification (UDC).

2.3.1 Cross-Country Trends

Figure 2.2 displays annual per capita book production by country for the entire sample period.[15,16] For visibility, I split the ten countries in the sample into two graphs. The top graph includes the five largest countries by population, and the bottom graph includes the remaining five countries. I use a log scale for books, both to reduce the visual impact of outliers, such as England in the early 1640s, and to emphasize differences across countries at different points in time.

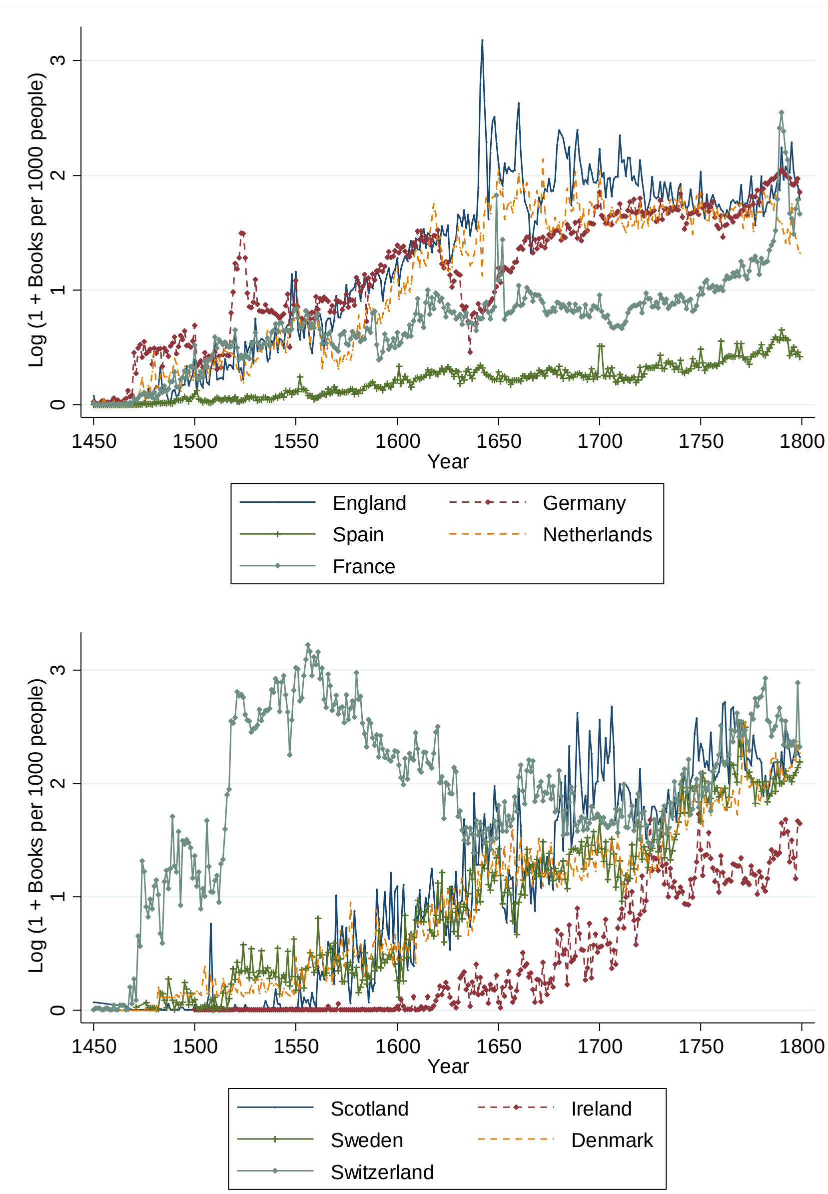
Switzerland had an early leading role as a printing center for the Reformation, which persisted until the early 17th century. Calvin and Zwingli were both based in Switzerland and made heavy use of the printing press to disseminate their ideas. A clear spike in book production is visible in 1517, the year Luther's Ninety-Five Theses were published. Germany, where printing technology was invented, was also particularly active in the early period. The same spike in production is visible in 1517, although there it was short-lived, but Germany was among the leaders until the onset of the devastating Thirty Years' war (1618-1648). Book production experienced a sharp drop in this period, and only fully recovered in the 18th century. As Germany collapsed, England rose to the top with a printing explosion during the civil war in the early 1640s. Book production remained high and volatile throughout the revolutionary period, and then declined somewhat in the 18th century, at which point other countries caught up with England.[17] The Netherlands experienced a significant increase after independence from Spain in the late 16th century, and remained just below England

[15] The numerator includes all books produced in each country, including the nine percent of books I am unable to match with a city. It is highly likely that these books were also published in cities in the Bairoch data, see section 2.2 for a discussion of unmatched books. Including only matched books makes little difference for the patterns observed in the graphs.

[16] The denominator includes only urban population, as measured in the Bairoch data, rather than the population of the entire country. Since printing was almost exclusively urban, this seems to be the appropriate choice conceptually. Using total country population only makes a substantial difference for the Netherlands, which had a much higher urbanization rate than other countries in the sample, and a smaller difference for Germany, which had a lower urbanization rate. Annual population numbers for these graphs are obtained by interpolating the Bairoch data and assuming constant population outside the interval of interpolation.

[17] It has been argued that England had lower levels of literacy than Sweden or Germany at the onset of the Industrial Revolution (Mitch, 1993), as measured by the ability to sign documents. This was not true in the case of book production, however, as Baten and van Zanden (2008) and this paper both show.

Figure 2.2: *Per Capita Book Production Across Countries*



Notes: This figure presents annual per capita book production for the ten countries in the sample. The numerator includes all book records associated with each country, including the nine percent of records without a city match. The denominator includes urban population only, as measured by the Bairoch data set. When a book record indicates multiple cities of publication, I count one book per city. When the record indicates a range for the year of publication I assign an equal fraction of the book to each year in the range.

until the very end of the sample, with book production dropping in the 1790s. France was initially among the leaders but fell behind in the 1560s. The brief spike around 1650 corresponds to the Fronde, a series of civil wars between 1648 and 1653. It remained behind until the revolution in 1789, which led to a burst of publications. Spain, unlike the rest of Europe, had very low levels of book production throughout the period.

Among the remaining countries, Scotland, Denmark and Sweden were somewhere between France and the leading countries for most of the sample period, and caught up with the leaders in the 18th century. Scotland had a period of intense publishing in the decades around 1700, which coincides with the start of the Scottish Enlightenment. Ireland had virtually no book production until the 17th century; in the 18th century its production grew substantially but it remained behind the leading countries.[18]

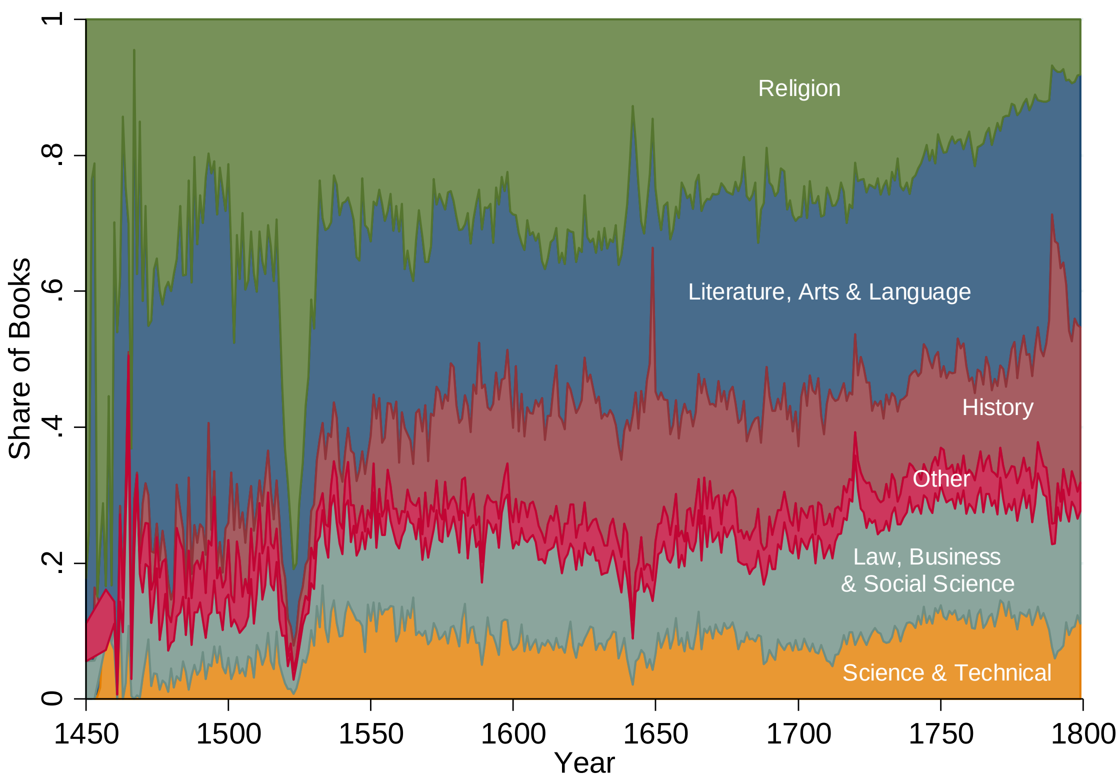
2.3.2 Subject Trends

A key advantage of these data is the ability to classify a sample of books by subject. Figure 2.3 offers a visual overview of subject trends in this sample by showing the share of books by subject over time, where I aggregate the 14 subjects defined in section 2.2 into six broader categories. Early printing was dominated by religion and literature. Religious books comprised up to 80 percent of books in the first decade of printing. Between the late 15th century and the first half of the 17th century, they represented about 35 percent of books, with the exception of the 1520s, at the onset of the Reformation, when they rose to 80 percent again. After this the share of religious books dropped markedly, falling to below 10 percent by 1800. Literature, arts and language also thrived in the 15th and 16th centuries,

[18] Baten and van Zanden (2008) and Buringh and Van Zanden (2009) also present data on the evolution of book production across countries in the same period. They report only 50-year aggregates rather than annual data, so a detailed comparison cannot be made. Still, the findings presented here appear to be broadly in line with theirs. The major exception is that they report much higher book production per capita in the Netherlands than in all other countries from the 17th century onwards. This difference is at least partially explained by the fact that they use total population in the denominator, whereas I use urban population only, and the Netherlands had a higher urbanization rate than the rest of Europe in this period. They also find that production dropped sharply in Switzerland in the 17th century, to the level of Spain, and that Sweden experienced a large increase in production in the second half of the 18th century, making it the second highest book producer in per capita terms. In the WorldCat data Switzerland remained close to the leaders; Sweden grew steadily over time and was among the leaders at the end of the 18th century, but did not stand out.

driven by the publication of Greek and Roman classics. Roughly 40 percent of books fell into this category between the 1460s and 1510. After the 1520s, this fraction remained relatively constant at around 30 percent throughout the sample period.

Figure 2.3: *Share of Books by Subject*



Notes: This figure presents the share of books by subject over time. I combine the 14 subjects defined in table 2.1 into six broader groups. "Science & Technical" includes Science, Medicine, Technology and Agriculture. "Other" includes General and Philosophy. Book counts for each subject are adjusted by multiplying the raw counts by the ratio of total books to books with subject information at the country-year level, to correct for different levels of subject availability across countries and time.

Other subjects gained importance over time. History, which includes books on current events as explained in section 2.2, took off in the mid-16th century and represented about 20 percent of publications from then onwards, with large spikes during major historical episodes such as the English and French civil wars of the mid-17th century and the French revolution in 1789. The share of books on law, business and social science grew steadily throughout the sample period, from about five to 20 percent. Science and technical books, which include books on science, medicine, technology and agriculture, are closely associated with economically useful knowledge. These books represented a small but growing share of publications throughout the sample period, going from about five percent in the 15th century to about 12 percent in the late 18th century. Finally, the "other" category includes philosophy and general books, and remained below five percent of books throughout the period.

Figures 2.4 and 2.5 offer a more detailed view by showing per capita book production for each of the 14 subjects. Importantly for the analysis below, they offer a breakdown of science and technical subjects from figure 2.3. Within this category, books on science and medicine grew from the 16th century onwards, while books on technology and agriculture only took off in the second half of the 18th century, at the onset of the Industrial Revolution. Science publications also accelerated markedly in the 18th century, as did those in social science, business and finance, and arts, while religious books declined. This is consistent with the Age of Enlightenment associated with this period.

How correlated was the publication of books across subjects? Table 2.2 presents the correlation matrix for log books per capita in each of the 14 subjects at the city-year level. Publications were positively but not strongly correlated across subjects, with most pairwise correlations falling in the 0.15-0.45 range.

2.4 The Effect of Book Production on City Growth

2.4.1 Methodology

Population is a widely used measure of economic development at the city level, both historically and in modern contexts (De Long and Shleifer, 1993; Acemoglu *et al.*, 2005; Glaeser *et al.*, 1995). The underlying logic is that differences in productivity growth across cities trigger migratory flows such that the utility of potential migrants remains equalized across cities, with rising house prices limiting migration to more productive cities (Roback,

Figure 2.4: *Per Capita Book Production by Subject (1/2)*

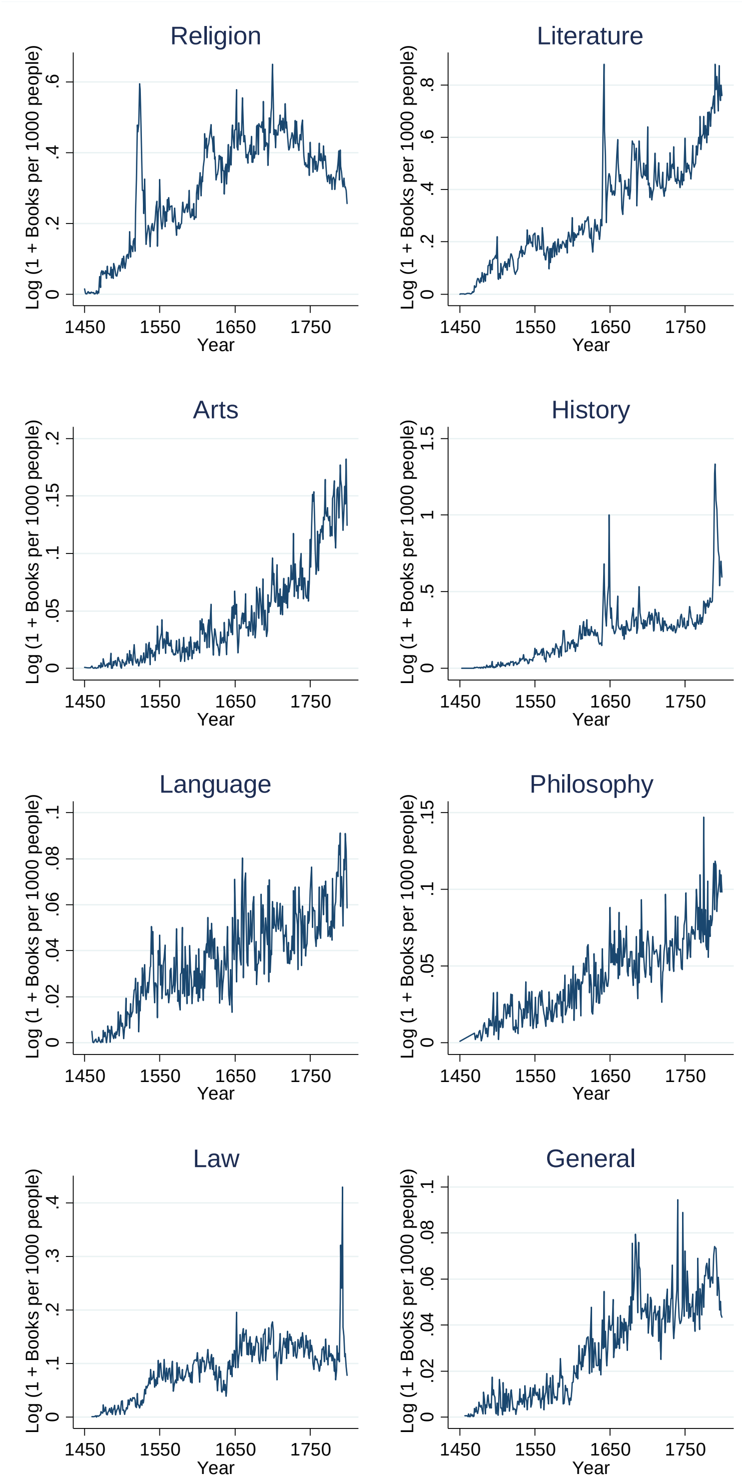
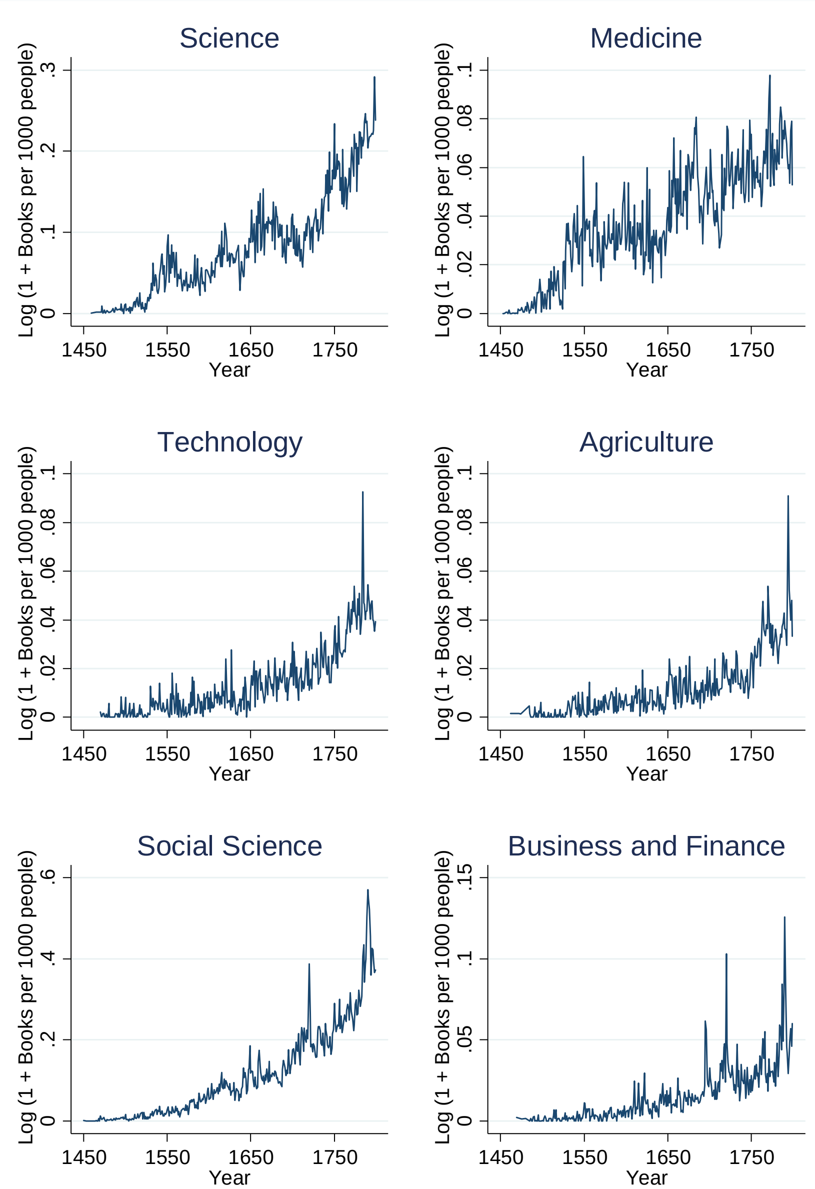


Figure 2.5: *Per Capita Book Production by Subject (2/2)*



Notes: These figures present annual per capita book production by subject. See notes to figure 2.2 for details on the measurement of per capita book production. Book counts for each subject are adjusted by multiplying the raw counts by the ratio of total books to books with subject information at the country-year level, to correct for different levels of subject availability across countries and time.

Table 2.2: *Per Capita Book Production by Subject - Correlation Matrix*

	Religion	Literature	Arts	History	Language	Philosophy	Law
Religion	1						
Literature	0.534	1					
Arts	0.266	0.326	1				
History	0.459	0.526	0.283	1			
Language	0.379	0.396	0.200	0.336	1		
Philosophy	0.413	0.431	0.219	0.362	0.323	1	
Law	0.419	0.385	0.187	0.370	0.316	0.320	1
General	0.304	0.340	0.183	0.305	0.270	0.271	0.247
Science	0.409	0.442	0.262	0.418	0.367	0.403	0.336
Medicine	0.348	0.383	0.194	0.352	0.325	0.339	0.374
Technology	0.191	0.224	0.175	0.220	0.182	0.195	0.156
Agriculture	0.165	0.228	0.148	0.188	0.154	0.164	0.144
S. Science	0.407	0.471	0.243	0.454	0.315	0.381	0.370
Bus. Fin.	0.147	0.203	0.123	0.191	0.117	0.132	0.160
	General	Science	Medicine	Technology	Agriculture	S. Science	Bus. Fin.
General	1						
Science	0.296	1					
Medicine	0.262	0.411	1				
Technology	0.164	0.270	0.200	1			
Agriculture	0.155	0.231	0.164	0.153	1		
S. Science	0.292	0.418	0.341	0.254	0.217	1	
Bus. Fin.	0.135	0.159	0.140	0.0963	0.101	0.220	1

Notes: This table presents the correlation matrix per capita book production across subjects. Observations are at the city-year level.

1982; Glaeser and Gottlieb, 2009).[19] Following this approach, let city growth be given by

$$\Delta \ln L_{it} = \Delta \ln A_{it} + \alpha_i + \eta_{jt} + \varepsilon_{it} \tag{2.1}$$

where Δ denotes the change from t to $t+1$, L_{it} is population, A_{it} is the level of technology, α_i captures fixed city-level drivers of growth, such as geographic features, persistent institutions or culture, η_{jt} are country-level shocks, such as revolutions or wars, and ε_{it} are city-level shocks.

Technological change reflects the incorporation of new knowledge, which will be measured by per capita book production $\frac{B}{L}$, into the existing technological stock. New knowledge is presumably not incorporated instantly into production processes, but instead diffuses over time. I assume that new knowledge is incorporated into technology at the rate γ and that the growth rate of technology can be written as

$$\Delta \ln A_{it} = F\left[\gamma \sum_{s=0}^{t} \frac{B_{is}}{L_{is}} (1-\gamma)^{(t-s)}\right] \tag{2.2}$$

The sum corresponds to the stock of knowledge that has not been incorporated into technology at time t, and F is an increasing function capturing the productivity of new knowledge. This equation parallels the technological change equation in knowledge-based models of endogenous growth (Romer, 1990; Grossman and Helpman, 1991; Aghion and Howitt, 1992), with new knowledge explicitly measured by books per capita instead of by the amount of human capital allocated to research.

Assuming a log form[20] for F and plugging (2.2) into (2.1), city growth can be expressed as a function of the stock of unincorporated knowledge at t and unobserved factors α_i, η_{jt} and ε_{it}:

$$\Delta \ln L_{it} = \beta \ln\left[\sum_{s=0}^{t} \frac{B_{is}}{L_{is}} (1-\gamma)^{(t-s)}\right] + \alpha_i' + \eta_{jt} + \varepsilon_{it} \tag{2.3}$$

[19] Blanchard and Katz (1992) show that this mechanism is also active across U.S. states

[20] The distribution of per capita book production is considerably right-skewed. Figure 2.6 suggests that a linear relationship between log city growth and log unincorporated knowledge is a good approximation. I use the log of one plus unincorporated knowledge in order to include city-year observations without any book production in the analysis.

where $\alpha_i' \equiv \alpha_i + \beta \ln \gamma$. Equation (2.3) will be the main estimating equation in the analysis below.

In the Bairoch data, population after 1450 is available in 100-year intervals between 1500 and 1700 and in 50-year intervals between 1700 and 1850. I therefore measure $\Delta \ln L_{it}$ as $L_{it+100} - \ln L_{it}$ in the earlier period and $2 \times (\ln L_{it+50} - \ln L_{it})$ in the later period, so that all growth rates are expressed in the same units.

The measurement of unincorporated knowledge $\sum_{s=0}^{t} \frac{B_{is}}{L_{is}} (1-\gamma)^{(t-s)}$ requires assuming a value for γ and a value for the initial stock in 1450. For my main specification I choose $\gamma = 0.1$, which implies that over 90 percent of knowledge is incorporated into technology within 25 years (i.e. $1 - 0.9^{25} = 0.928$), but I also show that the results are not very sensitive to other choices within a reasonable range for γ. The value of the initial stock in 1450 is unimportant given reasonable values of γ, and I set it to 0.[21] In addition, the measurement of $\frac{B}{L}$ requires annual population estimates, which I obtain by linear interpolation of log population using the Bairoch data.[22]

2.4.2 Results for All Books

I start by presenting results from estimating (2.3) for all books, without distinguishing between subjects. Throughout the rest of this section and in the corresponding graphs and tables I refer to unincorporated knowledge, as defined in equation (2.2), as books or book production for simplicity. In addition to the city and country-by-year fixed effects in (2.3), the set of controls for all regressions includes a quartic in log population at t, and standard errors are clustered at the city level.

Figure 2.6 presents the relationship between city growth and log book production in my

[21] For $\gamma = 0.1$, less than one percent of the knowledge created before 1450 was unincorporated into technology by 1500, the time of the first growth observation in the data. In any case, book production before printing was extremely low. According to Buringh and Van Zanden (2009) there were more books produced in Europe between 1450 and 1500 than in the preceding one thousand years.

[22] Outside the interval of known population points I assume constant population equal to the closest known point. For example, if the Bairoch data reports population between 1600 and 1750 only, I interpolate population between 1600 and 1700 and assign the 1600 population to years before 1600 (assigning population after 1700 is not necessary, since the last city growth observation is the one between 1700 and 1750).

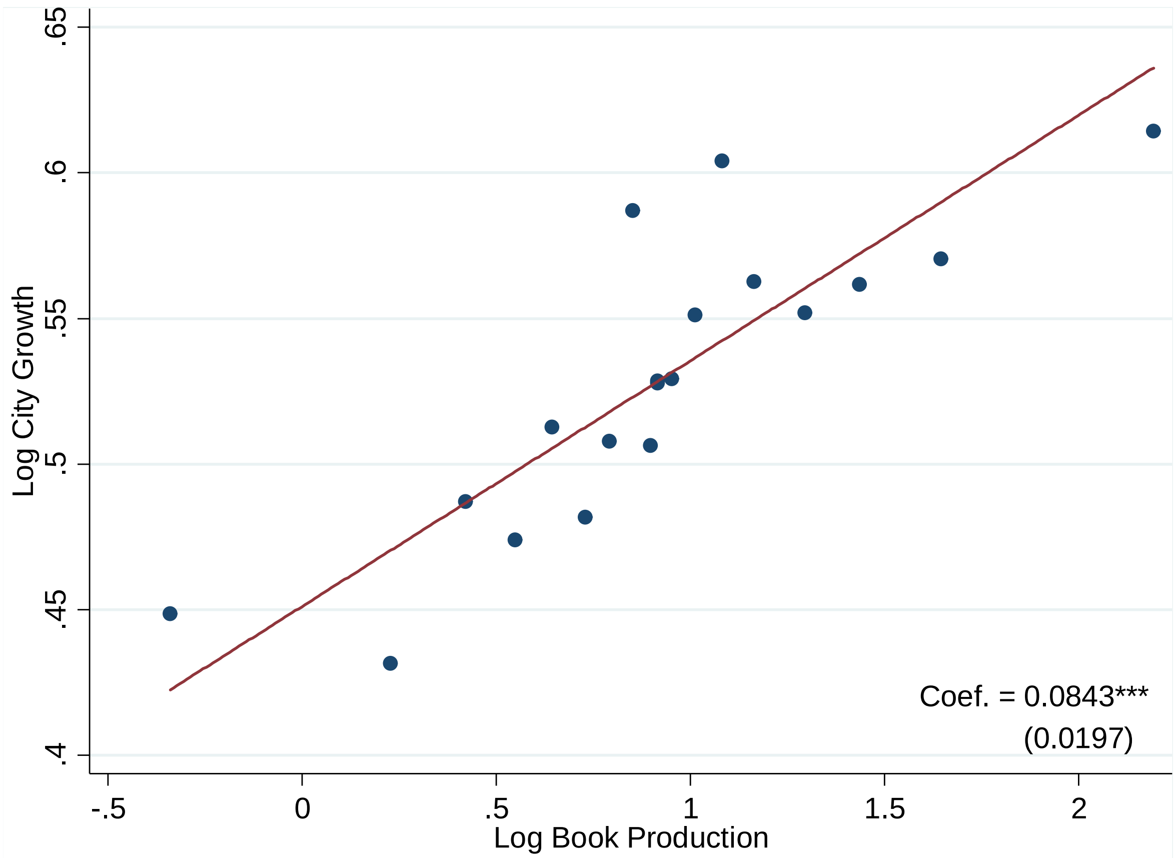
baseline specification, which sets $\gamma = 0.1$. The figure shows a binned scatter plot of city growth and log book production, both residualized on the set of controls, along with the corresponding regression line, coefficient and standard error.[23] The coefficient on log book production equals 8.43 log points, and is highly significant. In addition, the plot shows that the relationship is approximately linear, suggesting the log form for book production is an appropriate choice.

Table 2.3 presents the corresponding regression in column one, along with variations on the baseline specification in the remaining columns. Columns two and three show that the coefficient is not very sensitive to different values of γ. In column two I use $\gamma = 0.2$, which implies that over 90 percent of knowledge is incorporated in around 12 years instead of 25, and the coefficient rises slightly to 9.65 log points. In column two I use $\gamma = 0.05$, in which case it takes around 50 years for the same 90 percent of knowledge to be incorporated into technology, and the coefficient falls slightly to 7.14 log points. Column four adds linear city-specific time trends, to further account for unobserved drivers of growth. This is a particularly demanding robustness test since there are only five observations per city at most, with an average of three. The linear trends therefore absorb a considerable amount of variation in growth and book production and amplify the effect of measurement error. In this specification the coefficient falls to 5.56 log points and remains significant at the 10 percent level. Column five restricts the sample to cities with positive book production, and the coefficient is essentially unchanged at 8.64 log points. Column six weights observations by average city population across the sample period, which increases the coefficient to 11.12 log points. Finally, column seven shows that the results are robust to using levels instead of logs in book production.

Next, I examine how this relationship varies across time and geographies. The role of knowledge and human capital as drivers of economic growth before industrialization is

[23] The plot shows mean residual city growth and log book production for 19 equal-sized bins for all cities with positive book production at any point during the sample period plus one bin including all cities without book production. The regression line, coefficient and standard error are estimated on the underlying data, not the binned averages. All other binned scatter plots in the paper follow the same procedure.

Figure 2.6: *Effect of Book Production on City Growth*



Notes: This figure presents a binned scatter plot of log city growth and log book production. City growth is measured as the change in log population from t to $t + 100$ between 1500 and 1700 and two times the change in log population from t to $t + 50$ between 1700 and 1850. Book production is measured as the stock of unincorporated knowledge at t (see section 2.4.1 for details on the construction of this variable). Both variables are first residualized on city and country-by-year fixed effects and on a quartic in log population at t, and then grouped into 19 equal-sized bins for all cities with positive book production at any point during the sample period plus one bin including all cities without book production. The regression line, coefficient and standard error are estimated on the underlying data, not the binned averages. All other binned scatter plots in the paper follow the same procedure.

Table 2.3: *Effect of Book Production on City Growth*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
					Books>0	Weighted	
Log Books, $\gamma = 0.1$	0.0843***			0.0556*	0.0864***	0.1112***	
	(0.0197)			(0.0300)	(0.0211)	(0.0292)	
Log Books, $\gamma = 0.2$		0.0965***					
		(0.0220)					
Log Books, $\gamma = 0.05$			0.0714***				
			(0.0184)				
Books, $\gamma = 0.1$							0.0008**
							(0.0003)
Log Initial Population	Y	Y	Y	Y	Y	Y	Y
Country x Year FE	Y	Y	Y	Y	Y	Y	Y
City FE	Y	Y	Y	Y	Y	Y	Y
City Time Trends				Y			
Observations	3060	3060	3060	3060	2217	3060	3060
Number of Cities	1119	1119	1119	1119	877	1119	1119

Notes: This table presents regressions of city population growth on book production. City growth is measured as the change in log population from t to t+100 between 1500 and 1700 and two times the change in log population from t to t+50 between 1700 and 1850. Book production is measured as the stock of unincorporated knowledge at t (see section 2.4.1 for details on the construction of this variable). Columns one to three use alternative assumptions about the rate of knowledge incorporation into output. Column four adds city-level time trends. Column five restricts the sample to observations with positive book production, and column six weights cities by their average population in the sample period. Column seven uses the level rather than the log of book production.

disputed. While Dittmar (2011) and Cantoni and Yuchtman (2014) show that adoption of the printing press and the creation of universities, respectively, increased growth in late medieval and early modern Europe, Squicciarini and Voigtländer (2014) argue that the presence of knowledge elites did not increase growth before 1750. Column one in table 2.4 interacts log book production with time indicators and shows that book production increased growth both before and during industrialization. The coefficient on books is stable and highly significant across periods, ranging from 8.95 to 12.44 log points, with the exception of the coefficient for 1750, which is small and insignificant.[24] These results therefore do not reject the idea that knowledge and human capital were unimportant right at the onset of industrialization, in the second half of the 18th century, but suggest that

[24]The same pattern holds in cross sectional regressions of city growth and book production by time period, although the coefficient in the regression for 1750 is slightly larger and closer to being significant.

if anything this period was exceptional in that regard, rather than representative of the pre-industrial age.

Table 2.4: *Effect of Book Production on City Growth by Time Period and Country*

	(1) All	(2) France	(3) Germany	(4) England	(5) Netherlands	(6) Spain	(7) Other
Log Books		0.1564***	0.0922***	-0.0102	0.1307**	0.1351*	0.0736**
		(0.0446)	(0.0322)	(0.0883)	(0.0606)	(0.0799)	(0.0359)
Log Books x 1500	0.1085***						
	(0.0360)						
Log Books x 1600	0.1244***						
	(0.0243)						
Log Books x 1700	0.1096***						
	(0.0302)						
Log Books x 1750	0.0113						
	(0.0262)						
Log Books x 1800	0.0895***						
	(0.0257)						
Log Initial Population	Y	Y	Y	Y	Y	Y	Y
Country x Year FE	Y	Y	Y	Y	Y	Y	Y
City FE	Y	Y	Y	Y	Y	Y	Y
Observations	3060	790	700	364	171	769	266
Number of Cities	1119	339	237	156	60	252	75

Notes: This table presents additional regressions of city population growth on book production. See notes to table 2.3 for details on the construction of the variables. Column one interacts book production with time indicators. Columns two through seven split the sample by geography.

The remaining columns in table 2.4 split the sample by geography. Columns two to six present regressions for each of the five largest countries by population – France, Germany, England, Netherlands and Spain. Column seven groups the smaller countries – Denmark, Ireland, Scotland, Sweden and Switzerland – into one regression.[25] The coefficient on book production is significant in all cases except England, and ranges from 7.36 in the smaller countries to 15.64 log points in France. The fact that England is not significant is unsurprising, because publishing in England was uniquely concentrated. Printing was legally restricted to London, with the exception of Oxford and Cambridge Universities, from the 16th to the early 18th century. A Royal Charter in 1557 gave the Stationers' Company, a

[25]There are too few observations in the smaller countries to estimate the effect by country.

guild of printers in London, a monopoly on printing and the legal power to enforce it. The Statute of Ann in 1710 instituted copyright protection for authors and ended the Stationers' monopoly, but London's dominance persisted. The book production data show that these restrictions were highly effective: 93 percent of books published in England during the sample period were published in London, followed by Oxford and Cambridge with 2 and 1 percent respectively. This suggests that local book production was a poor proxy for local consumption in English cities, and that the English data probably have a low signal to noise ratio. Other capitals also had a prominent role as publishing centers, but not as dominant.[26] All results are unchanged when capital cities in all ten countries are excluded.

To get a sense of the magnitude of these findings, a city in the 75^{th} percentile of book production among all city-years with positive book production grew approximately 17.4 log points faster than a city without book production. This compares with an average growth of 52.8 log points and a standard deviation of 87.7 log points.

2.4.3 Results by Subject

The findings presented so far could be interpreted in different ways. Book production could measure knowledge diffusion, but it could also measure the level of literacy or consumption. In fact, section 2.3 showed that the majority of books published in this period were either on religion, literature, language or arts. This section distinguishes between these interpretations by estimating the effect of book production on growth by subject, using the 12 percent sample of book records with subject information. All estimates assume $\gamma = 0.1$, although again the results are not very sensitive to this choice.

Column one in table 2.5 reports the coefficients from estimating equation (2.3) separately for each of the 14 subjects defined in section 2.2. Given that book production is positively correlated across subjects, as shown in table 2.2, this specification minimizes bias from

[26]The highest level of concentration after England was in France, where 74 percent of books were published in Paris. In Spain, 47 percent of books were published in Madrid and in the Netherlands 45 percent were published in Amsterdam. Among the five largest countries, Germany had the least concentrated book production, probably because it was not politically integrated in this period. The top producers in Germany were the two cities that held major book fairs, Leipzig and Frankfurt, with 20 and 9 percent respectively.

measurement error but suffers from omitted variable bias, and is therefore loaded in favor of each subject. Seven subjects have positive and significant coefficients in column one. Books on business and finance and books on technology have the highest coefficients: 21.5 and 17.2 log points respectively. Next come books on medicine and history, with 10.0 and 9.3 log points, followed by arts with 8.6, social science with 7.2 and finally religion with 4.2. The coefficients on technology and history are significant at the 1 percent level, while the coefficients on arts and religion are barely significant at the 10 percent level.

Column two reports the corresponding coefficients when all subjects are included in the same regression. This specification is less vulnerable to omitted variable bias but might be more vulnerable to measurement error. Sampling error is presumably larger for less common subjects, which suggests that the coefficients on these subjects are more likely to be biased towards zero. Of the seven significant coefficients in column one, four remain significant in column two. Technology increases to 21.7 log points, business and finance drops marginally to 20.5 log points and medicine and history both increase slightly to 11.6 and 10.6 log points, respectively. The only change in conventional thresholds of significance is medicine, which becomes significant at the 10 percent level only. The coefficients on the remaining three subjects that were significant by themselves in column one – social science, arts and religion – fall substantially and become insignificant in column two. Since these are among the most common subjects in the sample, while technology and business and finance are among the least common, it is likely that their significance in column one was driven by omitted variable bias.[27]

Overall, only technology, business and finance, medicine and history are robustly associated with city growth across the two columns. Figure 2.7 presents binned scatter plots for each of these four subjects, constructed analogously to figure 2.6 and controlling

[27] In addition, two coefficients that were insignificant in column one become negative and significant in column two, agriculture and general. It is plausible that agricultural books were associated with rural growth, which may in turn have been negatively associated with urban growth. The coefficient on general books is harder to interpret, but it should be noted that it is significant at the 10 percent level only. In any case, strong conclusions should not be drawn for these two subjects given the inconsistency between the results in columns one and two.

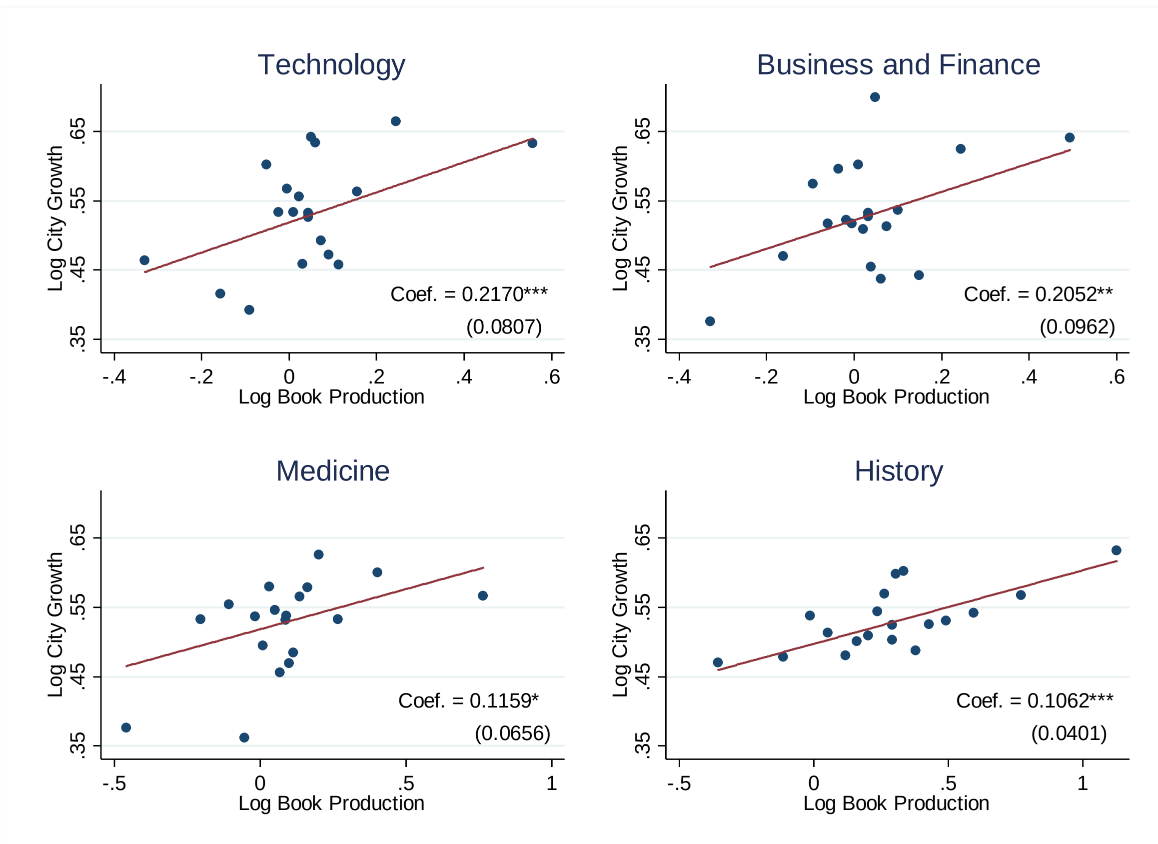
Table 2.5: *Effect of Book Production on City Growth by Subject*

	(1) Univariate	(2) Multivariate
Religion	0.042*	0.006
	(0.025)	(0.035)
Literature	0.036	-0.007
	(0.024)	(0.035)
Arts	0.086*	0.053
	(0.051)	(0.063)
History	0.093***	0.106***
	(0.029)	(0.040)
Language	0.052	-0.019
	(0.043)	(0.062)
Philosophy	0.035	-0.054
	(0.044)	(0.069)
Law	0.031	-0.034
	(0.032)	(0.037)
General	0.007	-0.111*
	(0.049)	(0.064)
Science	0.039	-0.069
	(0.031)	(0.050)
Medicine	0.100**	0.116*
	(0.041)	(0.066)
Technology	0.172***	0.217***
	(0.056)	(0.081)
Agriculture	-0.006	-0.215**
	(0.071)	(0.093)
Social Science	0.072**	0.015
	(0.035)	(0.057)
Business and Finance	0.215**	0.205**
	(0.087)	(0.096)
Log Initial Population	Y	Y
Country x Year FE	Y	Y
City FE	Y	Y
Observations	3060	3060
Number of Cities	1119	1119

Notes: This table presents regressions of city population growth on book production by subject. See notes to table 2.3 for details on the construction of the variables. In column one each coefficient is estimated in a separate regression where the corresponding subject enters alone, while column two reports coefficients when all subjects are included in the same regression.

for all other subjects, i.e. corresponding to the specification from column two of table 2.5. The larger standard errors translate into greater dispersion around the regression line, especially for technology and business and finance, the least common among the four subjects, but the relationship is clearly visible in all four cases. Technology, business and finance and medicine have a clear interpretation as economically useful knowledge. History, as explained in section 2.2, includes books about current events at the time they were written, which makes it harder to interpret. One possibility is that it captures the existence of a literate elite involved in or at least informed about public life.

Figure 2.7: *Effect of Book Production on City Growth by Subject*



Notes: This figure presents binned scatter plots of log city growth and log book production for books on technology, business and finance, medicine and history. See notes to figure 2.6 for details on the construction of the variables. The regression line, coefficient and standard error are estimated on the underlying data, not the binned averages.

It is interesting to further distinguish between books on business and books on finance. The former includes books on management (e.g. accounting, business arithmetic), commerce, transport and communications, while the latter includes books on personal, corporate and public finance. Figure 2.8 presents separate binned scatter plot for business and finance, controlling for all other subjects, and shows that the relationship with growth is entirely driven by finance books.[28] One interpretation of this finding is that while financial markets, and in particular banking, developed in Europe from the late middle ages onwards, managerial knowledge did not become relevant before the rise of the modern corporation in the second half of the 19th century (e.g. Chandler, 1977).

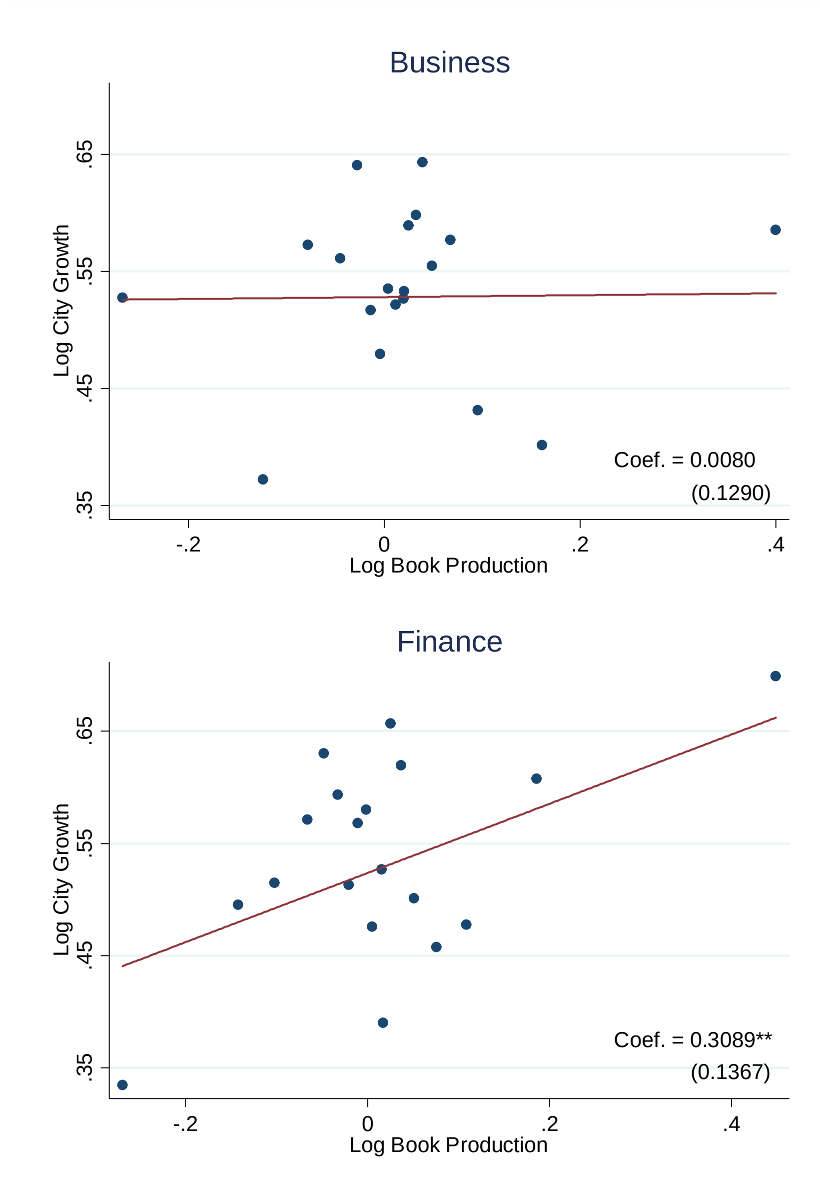
Turning to the remaining subjects, books on religion, literature and language, which are presumably associated with the level of literacy in the population, are not associated with growth. Legal books, which might be interpreted as a proxy for the rule of law, are also not significant. Books on social science are significant only when other subjects are omitted. Finally, one finding that seems surprising in light of the economically useful knowledge interpretation is the insignificance of science books, which I turn to next.

The Role of Science

The role of science as a driver of historical innovation and growth is debated. One view (Mathias, 1969; Hall, 1974) is that the Scientific Revolution of the 17th century had little impact on technology in the Industrial Revolution, which was driven by practical innovators devoid of scientific training. In this view, major breakthroughs in physics or chemistry with industrial applications only occurred later, and science only became pivotal in the second stage of industrialization, after 1850. Another view (Musson and Robinson, 1969; Mokyr, 2005a) is that the Scientific Revolution led to the diffusion of the scientific mindset associated with the Enlightenment, a belief that material progress can be attained through the accumulation of knowledge. This in turn set off an accumulation of useful knowledge

[28]The data do not have enough power to further distinguish between personal and corporate finance on one hand and public finance on the other: both coefficients are positive and large but insignificant.

Figure 2.8: *Effect of Book Production on City Growth - Business and Finance*



Notes: This figure presents binned scatter plots of log city growth and log book production for books on business and finance. See notes to figure 2.6 for details on the construction of the variables. The regression line, coefficient and standard error are estimated on the underlying data, not the binned averages.

consisting of "catalogs of facts, based on experience and experiment rather than on understanding or careful analysis and testing" (Mokyr, 2005a) which drove industrialization, even if the corresponding theoretical breakthroughs only occurred later. The results in the previous section provide some support for both views. Science books are not associated with growth, but books on technology, whose production increased sharply in the 18th century, are. However, the science category analyzed so far combines fields that are typically associated with economic growth – e.g. physics, chemistry or geology – with fields that are less so – e.g. astronomy, botany or zoology, and this aggregation might conceal important heterogeneity. In this section I analyze the relationship between science and growth by field.

I use the LCC subclasses within the Science class Q to divide science into 12 fields: mathematics, astronomy, physics, chemistry, geology, biology, physiology, anatomy, zoology, botany, microbiology and other science. Table 2.6 shows the correspondence with the DDC and UDC. As for all other subjects above, I estimate two specifications, one where each field enters alone and another where I control for books in all other science fields and subjects. Column one in table 2.7 reports the results when each field enters alone, and two fields are positive and significant. Chemistry has a coefficient of 18.2 log points, significant at the 5 percent level, and geology has a coefficient of 19.3 log points, significant at the 1 percent level. Both coefficients increase slightly and remain significant in column two, which adds controls for all other science fields and subjects. Chemistry rises to 20.0 log points, significant at the 10 percent level, and geology to 23.1 log points, significant at the 5 percent level. Figure 2.9 shows the corresponding binned scatter plots. In addition, two fields become negative and significant at the 10 percent level in the second column, anatomy and zoology. In the case of anatomy, including medicine in the regression is likely to be a bad control, in the sense that any effect of anatomy books on growth may work at least partly through the production of medical books. In line with this hypothesis, dropping medicine raises the coefficient on anatomy, which becomes insignificant.[29]

[29]The same reasoning can be applied to physics and chemistry, on one side, and technology on the other, but dropping technology in the second column only increases the coefficient on physics marginally.

Table 2.6: *Classification of Science Books*

Field	Library of Congress	Dewey Decimal Class.	Universal Dec. Class.
Mathematics	Subclass QA	Division 51	Division 51
Astronomy	Subclasses QB, CE	Division 52 except section 526	Division 52 except subdivision 528
Physics	Subclass QC	Division 53	Division 53
Chemistry	Subclass QD	Division 54	Division 54
Geology	Subclass QE	Divisions 55, 56	Divisions 55, 56
Biology	Subclass QH	Division 57 except sections 571, 572, 573, 575, 579	Division 57 except subdivisions 577, 578, 579
Physiology	Subclass QP	Sections 571, 572, 573, 575, 612	Subdivisions 577, 612
Anatomy	Subclass QM	Section 611	Subdivision 611
Zoology	Subclass QL	Division 59	Division 59
Botany	Subclass QK	Division 58	Division 58
Microbiology	Subclass QR	Section 579	Subdivisions 578, 579
Other Science	Subclass Q	Division 50	Division 50

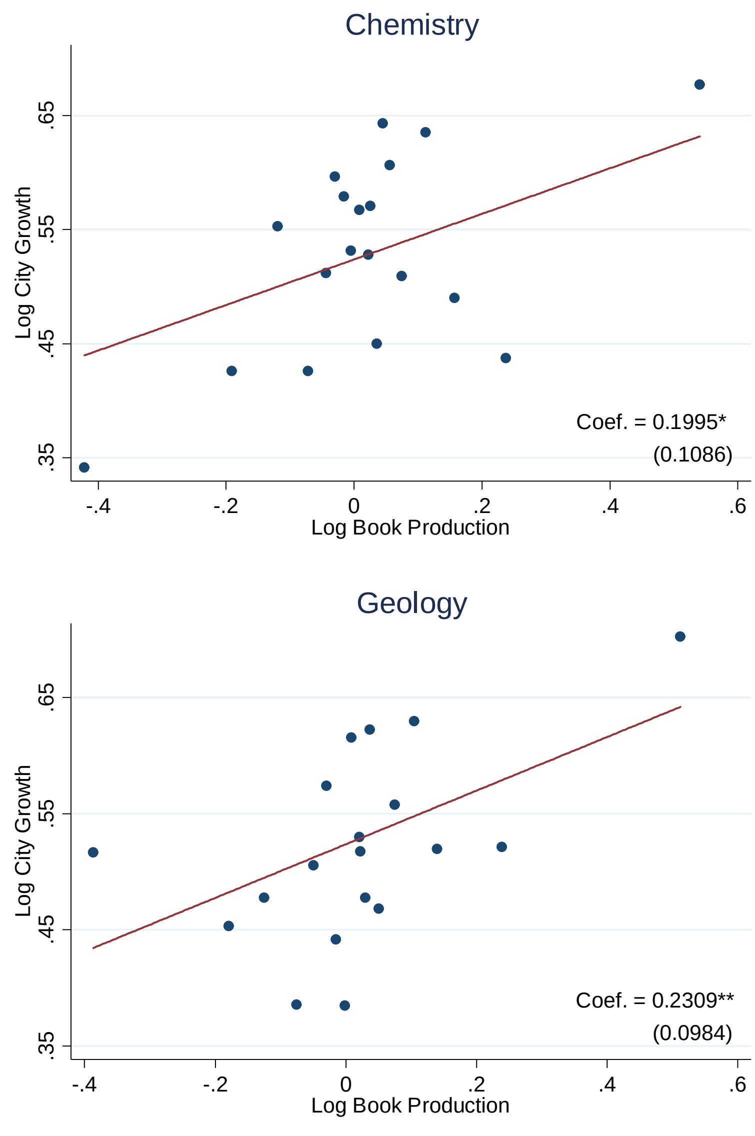
Notes: This table presents the correspondence between the 12 scientific fields used in the paper and the underlying codes in the Library of Congress Classification (LCC), the Dewey Decimal Classification (DDC) and the Universal Decimal Classification (UDC).

Table 2.7: *Effect of Book Production on City Growth by Science Field*

	(1) Univariate	(2) Multivariate
Mathematics	0.005	-0.059
	(0.063)	(0.077)
Astronomy	0.089	0.003
	(0.061)	(0.079)
Physics	-0.008	-0.127
	(0.062)	(0.088)
Chemistry	0.182**	0.200*
	(0.078)	(0.109)
Geology	0.193***	0.231**
	(0.062)	(0.098)
Biology	-0.037	-0.071
	(0.053)	(0.098)
Physiology	-0.054	-0.115
	(0.092)	(0.106)
Anatomy	-0.053	-0.189*
	(0.075)	(0.102)
Zoology	0.011	-0.120*
	(0.057)	(0.070)
Botany	0.086	0.060
	(0.055)	(0.087)
Microbiology	0.207	0.332
	(1.408)	(1.311)
Other Science	0.069	-0.010
	(0.060)	(0.096)
Log Initial Population	Y	Y
Country x Year FE	Y	Y
City FE	Y	Y
Observations	3060	3060
Number of Cities	1119	1119

Notes: This table presents regressions of city population growth on book production by scientific field. See notes to table 2.3 for details on the construction of the variables. In column one each coefficient is estimated in a separate regression where the corresponding field enters alone, while column two reports coefficients when all scientific fields, as well all other subjects apart from science, are included in one regression.

Figure 2.9: *Effect of Book Production on City Growth by Science Field*



Notes: This figure presents binned scatter plots of log city growth and log book production for books on chemistry and geology. See notes to figure 2.6 for details on the construction of the variables. The regression line, coefficient and standard error are estimated on the underlying data, not the binned averages.

These findings reveal that science did drive growth before and during the Industrial Revolution. While fields like physics and biology may have only developed later, chemistry and geology played an important role in this period. The importance of chemistry in early industrial development is highlighted by Clow and Clow (1952), and that of geology is consistent with the key role of coal mining in the Industrial Revolution.

2.5 Conclusion

This paper introduces a new database of 5.5 million books published in ten European countries between 1450-1800, and uses it to document historical patterns of book production and to estimate the effect of book production on city growth. It then uses information on book subjects to distinguish between interpretations of this effect. The effect is large and robust for books on technology, finance, medicine, history and, within science, chemistry and geology. Other topics, such as religion or literature, are not associated with growth. This suggests that the effect of book production reflects the diffusion of economically useful knowledge.

Relative to other measures of knowledge such as corporate R&D, book production has the advantage of being comprehensive as well as available over a long period of time, which enables me to estimate its effect from within-city variation alone.

These findings confirm the importance of knowledge accumulation and diffusion as a driver of long-run economic growth, in line with the central role it plays in growth theory.

References

Abramovitz, M. (1956). Resource and output trends in the united states since 1870. *The American Economic Review*, p. 5–23.

Acemoglu, D., Johnson, S., Robinson, J. A. and Yared, P. (2005). *From education to democracy?* Tech. rep., National Bureau of Economic Research.

Adams, J. D. (1990). Fundamental stocks of knowledge and productivity growth. *Journal of Political Economy*, **98** (4), 673–702.

Aghion, P. and Howitt, P. (1992). A model of growth through creative destruction. *Econometrica*, **60** (2), 323–351.

Akcigit, U., Alp, H. and Peters, M. (2015). *Lack of Selection and Imperfect Managerial Contracts: Firms Dynamics in Developing Countries*. Tech. rep., Working paper, University of Pennsylvania.

Albuquerque, R. and Hopenhayn, H. A. (2004). Optimal lending contracts and firm dynamics. *The Review of Economic Studies*, **71** (2), 285–315.

Allen, R. C. (2003). Progress and poverty in early modern europe. *The Economic History Review*, **56** (3), 403–443.

Angelini, P. and Generale, A. (2008). On the evolution of firm size distributions. *The American Economic Review*, p. 426–438.

Bairoch, P., Batou, J. and Pierre, C. (1988). *Population des villes européennes de 800 à 1850: banque de données et analyse sommaire des résultats (la)*. Librairie Droz.

Baten, J. and van Zanden, J. L. (2008). Book production and the onset of modern economic growth. *Journal of Economic Growth*, **13** (3), 217–235.

Bennedsen, M., Nielsen, K. M., Perez-Gonzalez, F. and Wolfenzon, D. (2007). Inside the family firm: The role of families in succession decisions and performance. *The Quarterly Journal of Economics*, **122** (2), 647–691.

Bento, P. and Restuccia, D. (2014). *Misallocation, Establishment Size, and Productivity*. Tech. rep., Working Paper, University of Toronto.

Bertrand, M. and Schoar, A. (2003). Managing with style: The effect of managers on firm policies. *The Quarterly Journal of Economics*, **118** (4), 1169–1208.

Blanchard, O. J. and Katz, L. F. (1992). Regional evolutions. *Brookings papers on economic activity*, p. 1–75.

Bloom, N., Brynjolfsson, E., Foster, L., Jarmin, R., Saporta Eksten, I. and Van Reenen, J. (2013). Management in america. *US Census Bureau Center for Economic Studies Paper No. CES-WP-13-01*.

—, Sadun, R. and Reenen, J. V. (2012). The organization of firms across countries. *The Quarterly Journal of Economics*, **127** (4), 1663–1705.

— and Van Reenen, J. (2007). Measuring and explaining management practices across firms and countries. *The Quarterly Journal of Economics*, p. 1351–1408.

Braguinsky, S., Branstetter, L. G. and Regateiro, A. (2011). *The incredible shrinking Portuguese firm*. Tech. rep., National Bureau of Economic Research.

Brouillant, L. J. d. (1888). *Histoire de Pierre du Marteau: imprimeur à Cologne (XVIIe-XVIIIe siècles)*. Maison Quantin.

Buringh, E. and Van Zanden, J. L. (2009). Charting the "Rise of the west": Manuscripts and printed books in europe, a long-term perspective from the sixth through eighteenth centuries. *The Journal of Economic History*, **69** (02), 409–445.

Caballero, R. J. and Jaffe, A. B. (1993). How high are the giants' shoulders: An empirical assessment of knowledge spillovers and creative destruction in a model of economic growth. In *NBER Macroeconomics Annual 1993, Volume 8*, MIT press, p. 15–86.

Cabral, L. M. and Mata, J. (2003). On the evolution of the firm size distribution: Facts and theory. *American economic review*, p. 1075–1090.

Cantoni, D. and Yuchtman, N. (2014). Medieval universities, legal institutions, and the commercial revolution. *The Quarterly Journal of Economics*, **129** (2), 823–887.

Card, D. (1999). The causal effect of education on earnings. *Handbook of labor economics*, **3**, 1801–1863.

Chandler, A. D. (1977). *The visible hand: The managerial revolution in American business*. Harvard University Press.

Clow, A. and Clow, N. L. (1952). *The Chemical Revolution: A Contribution to Social Technology*. Gordon and Breach Science Publishers.

Cooley, T. F. and Quadrini, V. (2001). Financial markets and firm dynamics. *American Economic Review*, p. 1286–1310.

Davis, S. J. and Haltiwanger, J. (1992). Gross job creation, gross job destruction, and employment reallocation. *The Quarterly Journal of Economics*, **107** (3), 819–863.

De Long, J. B. and Shleifer, A. (1993). Princes and merchants: European city growth before the industrial revolution. *Journal of Law and Economics*, p. 671–702.

Decker, R., Haltiwanger, J., Jarmin, R. and Miranda, J. (2014). The role of entrepreneurship in US job creation and economic dynamism. *Journal of Economic Perspectives*, **28** (3), 3–24.

Denison, E. F. (1985). *Trends in American Economic Growth, 1929-1982*. Brookings Institution Press.

Desai, M., Gompers, P. and Lerner, J. (2003). *Institutions, capital constraints and entrepreneurial firm dynamics: Evidence from Europe*. Tech. rep., National Bureau of Economic Research.

Dittmar, J. (2011). The welfare impact of a new good: The printed book.

Dunne, T., Roberts, M. J. and Samuelson, L. (1989). The growth and failure of u. s. manufacturing plants. *The Quarterly Journal of Economics*, **104** (4), 671.

Ericson, R. and Pakes, A. (1995). Markov-Perfect industry dynamics: A framework for empirical work. *The Review of Economic Studies*, **62** (1), 53–82, ArticleType: research-article / Full publication date: Jan., 1995 / Copyright © 1995 The Review of Economic Studies, Ltd.

Foster, L., Haltiwanger, J. C. and Krizan, C. J. (2001). Aggregate productivity growth. lessons from microeconomic evidence. In *New developments in productivity analysis*, University of Chicago Press, p. 303–372.

—, — and Syverson, C. (2013). The slow growth of new plants: Learning about demand? *National Bureau of Economic Research*.

Galor, O. (2005). From stagnation to growth: unified growth theory. *Handbook of economic growth*, **1**, 171–293.

Gennaioli, N., La Porta, R., Lopez-de-Silanes, F. and Shleifer, A. (2013). Human capital and regional development. *The Quarterly Journal of Economics*, **128** (1), 105–164.

Glaeser, E. L. and Gottlieb, J. D. (2009). The wealth of cities: Agglomeration economies and spatial equilibrium in the united states. *Journal of Economic Literature*, **47** (4), 983–1028.

—, Kerr, S. P. and Kerr, W. R. (2014). Entrepreneurship and urban growth: An empirical assessment with historical mines. *Review of Economics and Statistics*.

—, Scheinkman, J. A. and Shleifer, A. (1995). Economic growth in a cross-section of cities. *Journal of Monetary Economics*, **36** (1), 117–143.

Grendler, P. (1988). The conditions of enquiry: Printing and censorship. In C. B. Schmitt, Q. Skinner, E. Kessler and J. Kraye (eds.), *The Cambridge History of Renaissance Philosophy*, Cambridge: Cambridge University Press.

Griliches, Z. (1980). R & d and the productivity slowdown. *The American Economic Review*, p. 343–348.

— (1986). Productivity, r and d, and basic research at the firm level in the 1970's. *The American Economic Review*, p. 141–154.

— and Mairesse, J. (1984). Productivity and R&D at the firm level. In Z. Griliches (ed.), *R & D, patents, and productivity*, Chicago: University of Chicago Press, pp. 339–374.

—, Pakes, A. and Hall, B. H. (1987). The value of patents as indicators of inventive activity. In *Economic Policy and Technological Performance*, Citeseer.

Grossman, G. M. and Helpman, E. (1991). *Innovation and Growth in the Global Economy*. MIT Press.

Guzman, J. and Stern, S. (2015). Nowcasting and placecasting: Entrepreneurial quality and performance. In *Measuring Entrepreneurial Businesses: Current Knowledge and Challenges*, University of Chicago Press.

Hall, A. R. (1974). What did the industrial revolution in britain owe to science? In N. McKendrick (ed.), *Historical perspectives: studies in English thought and society, in honour of J. H. Plumb*, Europa, p. 129–151.

Harris, J. R. and Todaro, M. P. (1970). Migration, unemployment and development: a two-sector analysis. *The American economic review*, p. 126–142.

Hsieh, C. and Klenow, P. J. (2009). Misallocation and manufacturing TFP in china and india. *The Quarterly Journal of Economics*, **124** (4), 1403–1448.

— and — (2014). The life cycle of plants in india and mexico. *The Quarterly Journal of Economics*, **129** (3), 1035–1084.

Hurst, E. and Pugsley, B. W. (2011). What do small businesses do? *Brookings Papers on Economic Activity*, **2011** (2), 73–118.

Jaffe, A. B., Trajtenberg, M. and Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics*, **108** (3), 577–598.

Johnson, W. B., Magee, R. P., Nagarajan, N. J. and Newman, H. A. (1985). An analysis of the stock price reaction to sudden executive deaths: Implications for the managerial labor market. *Journal of Accounting and Economics*, **7** (1), 151–174.

Jovanovic, B. (1982). Selection and the evolution of industry. *Econometrica*, **50** (3), 649–670, ArticleType: research-article / Full publication date: May, 1982 / Copyright © 1982 The Econometric Society.

Kendrick, J. W. (1956). Productivity trends: Capital and labor. *The Review of Economics and Statistics*, p. 248–257.

Kihlstrom, R. E. and Laffont, J. (1979). A general equilibrium entrepreneurial theory of firm formation based on risk aversion. *The Journal of Political Economy*, p. 719–748.

Klette, T. and Kortum, S. (2004). Innovating firms and aggregate innovation. *Journal of Political Economy*, **112** (5), 986–1018.

Kortum, S. and Lerner, J. (2000). Assessing the contribution of venture capital to innovation. *RAND journal of Economics*, p. 674–692.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A. and Vishny, R. W. (1997). Trust in large organizations. *The American Economic Review*, p. 333–338.

— and Shleifer, A. (2014). Informality and development. *The Journal of Economic Perspectives*, p. 109–126.

Lucas, J., Robert E. (1978). On the size distribution of business firms. *The Bell Journal of Economics*, **9** (2), 508–523.

Mairesse, J. and Sassenou, M. (1991). R&D and productivity: a survey of econometric studies at the firm level.

Mankiw, N. G., Romer, D. and Weil, D. N. (1992). A contribution to the empirics of economic growth. *The Quarterly Journal of Economics*, **107** (2), 407–437.

Mansfield, E. (1965). Rates of return from industrial research and development. *The American Economic Review*, p. 310–322.

Mathias, P. (1969). Who unbound prometheus? science and technical change, 1600–1800. *Bulletin of Economic Research*, **21** (1), 3–16.

Melitz, M. J. (2003). The impact of trade on Intra-Industry reallocations and aggregate industry productivity. *Econometrica*, **71** (6), 1695–1725.

Mitch, D. (1993). The role of human capital in the first industrial revolution. *The British Industrial Revolution: An Economic Perspective*, p. 267–307.

Mokyr, J. (2005a). The intellectual origins of modern economic growth. *The Journal of Economic History*, **65** (02), 285–351.

— (2005b). Long-Term economic growth and the history of technology. In *Handbook of Economic Growth*, vol. 1, Elsevier, pp. 1113–1180.

Musson, A. E. and Robinson, E. (1969). *Science and Technology in the Industrial Revolution*. Manchester University Press.

Nanda, R. and Sorensen, J. B. (2010). Workplace peers and entrepreneurship. *Management Science*, **56** (7), 1116–1126.

Pakes, A. (1985). On patents, r & d, and the stock market rate of return. *Journal of Political Economy*, **93** (2), 390–409.

Penrose, E. T. (1959). *The Theory of the Growth of the Firm*. New York: John Wiley.

Rauch, J. E. (1991). Modelling the informal sector formally. *Journal of development Economics*, **35** (1), 33–47.

Roback, J. (1982). Wages, rents, and the quality of life. *Journal of Political Economy*, **90** (6), 1257–1278.

Romer, P. (1990). Endogenous technological change. *Journal of Political Economy*, **98** (5), 2.

Schoar, A. (2010). The divide between subsistence and transformational entrepreneurship. In *Innovation Policy and the Economy, Volume 10*, University of Chicago Press, p. 57–81.

Solow, R. M. (1957). Technical change and the aggregate production function. *The review of Economics and Statistics*, p. 312–320.

Squicciarini, M. P. and Voigtländer, N. (2014). *Human capital and industrialization: evidence from the age of enlightenment*. Tech. rep., National Bureau of Economic Research.