Online Appendix

Human Capital and Industrialization: Evidence from the Age of Enlightenment

Mara P. Squicciarini — KULeuven and FWO
Nico Voigtländer — UCLA, NBER, and CEPR

A Model: Upper-Tail Knowledge and Industrialization

In this section, we provide a simple model of spatial technology diffusion that connects the historical evidence discussed above with the empirical analysis that follows below. The model distinguishes between worker skills and upper-tail knowledge of entrepreneurs. We present a mechanism where upper-tail knowledge enables local entrepreneurs to *improve* their technology, while worker skills raise the productivity for a *given* technology. This yields differential predictions for how the two types of human capital affect income and economic growth before and after industrialization.

The model features $n = 1, ..., N$ regions with given land endowment. In each region, there is a mass $L >> 1$ of workers who supply one unit of labor inelastically in each period. Worker skills h_n vary across regions. In addition, there are $i \in [0, 1]$ entrepreneurs who produce intermediate goods in manufacturing. A share s_n of entrepreneurs in region n disposes of upper-tail (scientific) knowledge.[1]

There is no saving, so that all income is consumed in each period. In any given period, workers optimally choose between working in two sectors: agriculture (A) and manufacturing (M). The latter is performed in cities, so that the manufacturing labor share also reflects urbanization. We keep the model tractable by following Hansen and Prescott (2002) in assuming that agricultural and manufacturing goods are perfect substitutes. In addition, we assume that workers and entrepreneurs are immobile, operating within their region of origin.[2]

Sector-specific wages in each region depend on both types of knowledge. First, average worker skills affect the efficiency of production in both sectors, but to a lesser degree in agriculture. Sec-

[1] To distinguish between their effects on development, we assume that h_n and s_n are independently distributed across regions. This also reflects the observation that our historical proxies for the two types of human capital, literacy and subscriber density, are uncorrelated across French departments.

[2] Relaxing this assumption by allowing for *costly* migration would yield very similar cross-sectional predictions. The historical literature documents substantial mobility, in particular of elites – merchants and legal officials were more mobile than artisans or unskilled workers (Benedict, 2005). Nevertheless, in Appendix D.3 we show that more than half (52%) of the "famous" people in scientific professions were born and died in the same city in France before 1887 (and this number is even higher when excluding Paris – 64%).

ond, highly skilled entrepreneurs can raise productivity in manufacturing. Because we focus on differential development in the cross-section, we take the aggregate technology frontier $\bar{A}$ as given. We then study the effects when $\bar{A}$ grows (exogenously) over time. Growing $\bar{A}$ has two interpretations that are both in line with the historical evidence: i) that France was a follower country, with most technological progress coming from Britain; and ii) more broadly, that the frontier of useful knowledge expanded during the period of Industrial Enlightenment and that this knowledge became more accessible due to the emergence of "open science" (Kelly, Mokyr, and Ó Gráda, 2014). The latter interpretation allows for the possibility that France also innovated (as suggested by the historical evidence in Section 2.1), instead of merely adopting existing technology.[3] Finally, all relevant cross-sectional predictions of our model can be derived in partial equilibrium, taking the price of output (in both sectors) as given and using it as the numeraire.

A.1 Production

Each worker supplies one unit of labor and chooses a sector of employment at the beginning of each period. Technology in all sectors exhibits constant returns, so that the scale of production is not important. We denote total labor in sector $j \in \{A, M\}$ in region n by $L_{j,n}$. In the following, we characterize the production technologies used by the two sectors. Agricultural output in region n is given by

$$Y_{A,n} = A_A h_n^{\beta_A} X_n^{\alpha_A} L_{A,n}^{1-\alpha_A} \ , \tag{A.1}$$

where X is land endowment, α_A is the share of land in production, and β_A reflects the sensitivity of agricultural productivity with respect to worker skills.[4] We assume that there are no property rights to land, so that wages in agriculture are given by the average product $y_{A,n} = Y_{A,n}/L_{A,n}$:

$$w_{A,n} = A_A h_n^{\beta_A} \left(\frac{X_n}{L_{A,n}} \right)^{\alpha_A} = A_A h_n^{\beta_A} \left(\frac{x_n}{l_{A,n}} \right)^{\alpha_A} \quad , \tag{A.2}$$

where $l_{A,n} = L_{A,n}/L$ is the agricultural labor share, and x_n is land per worker in region n.[5] Note that agricultural wages increase if the labor share in agriculture declines, because this leaves more land for each remaining peasant. Thus, growth in manufacturing indirectly raises wages in

[3] The second, broader, interpretation also reflects the historical account that upper-tail knowledge was crucial for both innovation and adoption. Consequently, distinguishing between these two dimensions (as for example in Vandenbussche, Aghion, and Meghir, 2006) is not crucial for our results.

[4] In growth models and development accounting, h typically multiplies L directly, reflecting the average impact of schooling on productivity via Mincerian returns (c.f. Bils and Klenow, 2000). By using different β_j for $j \in \{A, M\}$, we allow these returns to vary across sectors, i.e., we allow for sector-specific returns to worker skills.

[5] Alternatively, we could assume that workers earn their marginal product, while landlords earn rents from land but are not active as workers or entrepreneurs. Then, using $\widetilde{A}_A \equiv A_A/(1-\alpha)$ instead of A_A would leave the rest of the model unchanged.

agriculture.

Our modeling of the manufacturing sector builds on Acemoglu, Aghion, and Zilibotti (2006). The setup embeds a role for entrepreneurial skills in the manufacturing production process; it also has the advantage that it reduces to a simple aggregate production function. The final manufacturing good is produced under perfect competition by firms that use labor and a continuum of intermediates as inputs. The technology exhibits constant returns, so that we can focus on aggregate output in manufacturing, produced by a representative firm in the final sector:

$$Y_{M,n} = \xi \cdot \left(\int_0^1 A_{M,n}(i)^{1-\alpha_M} z_n(i)^{\alpha_M} di \right) \left(h_n^{\beta_M} L_{M,n} \right)^{1-\alpha_M} \tag{A.3}$$

where ξ is a constant, $z_n(i)$ is the flow of intermediate good i in final production, $L_{M,n}$ is total labor in manufacturing, β_M is the sensitivity of manufacturing production with respect to worker skills, and α_M denotes the share of intermediates in final production.[6] Intermediates are produced by entrepreneurs under monopolistic competition. Each entrepreneur $i \in [0, 1]$ produces a specific intermediate i by transforming one unit of the final good into one unit of the intermediate. Thus, the marginal cost is identical for all entrepreneurs. However, the productivity with which intermediates enter final production, $A_{M,n}(i)$, differs across entrepreneurs i.[7] We study the evolution of productivity as a function of entrepreneurial skills below.

Solving the entrepreneurs' optimization problem yields a simple expression for aggregate manufacturing output (see Appendix B.1 for detail):

$$Y_{M,n} = A_{M,n} h_n^{\beta_M} L_{M,n}, \quad \text{with} \quad A_{M,n} = \int_0^1 A_{M,n}(i) di \tag{A.4}$$

Thus, aggregate manufacturing productivity $A_{M,n}$ is a simple linear combination of individual entrepreneurial efficiencies $A_{M,n}(i)$. The first order condition of (A.3) with respect to $L_{M,n}$ implies

[6]We assume that the elasticity of manufacturing production with respect to worker skills β_M remains unchanged when productivity $A_{M,n}$ grows. In other words, we assume that technological change is neutral with respect to worker skills (but not with respect to entrepreneurial skills). The historical evidence discussed in the paper suggests that technological change during the first Industrial Revolution was biased towards unskilled workers (c.f. O'Rourke, Rahman, and Taylor, 2013). Introducing this into our model (e.g., via β_M depending negatively on $A_{M,n}$) would dampen the increase in industrial wages relative to agriculture as $A_{M,n}$ rises. However, unless taken to the extreme, this effect would not reverse the net income gains and thus not change our qualitative results. In fact, this extension would further diminish the role of worker skills during industrialization.

[7]Effectively, higher $A_{M,n}(i)$ raises the demand for intermediate i in final production, but it does not affect the unit cost of i. This approach ensures tractability. It can be motivated, for example, by interpreting $A_{M,n}(i)$ as the quality of intermediate i, so that more productive entrepreneurs produce higher quality intermediates at the same marginal cost. Note, however, that productivity can still be interpreted as the standard quantity-related concept: in equilibrium, entrepreneurs with high $A_{M,n}(i)$ sell more and make higher profits, so that our setup is akin to an alternative that loads productivity differences on marginal costs.

that a share $1-\alpha$ of final output is paid to labor. Combining this with (A.4) yields the wage rate in manufacturing:

$$w_{M,n} = (1-\alpha_M)A_{M,n}h_n^{\beta_M} \tag{A.5}$$

Finally, we assume $\beta_M > \beta_A$, i.e., that manufacturing production is relatively more sensitive with respect to worker skills. This assumption matters for cross-sectional predictions in our model, but it does not affect growth. In the following, we study the evolution of productivity, where entrepreneurial skills play a central role.

A.2 The Evolution of Productivity

The technological frontier at time t is given by $\bar{A}_t$, and it grows at an exogenous rate $\gamma_{\bar{A},t}$.[8] The frontier affects the productivity of individual entrepreneurs i at locations n, as represented by the productivity process

$$A_{M,n,t}(i) = \eta\bar{A}_t \; + \; (1-\eta)\left(1+\tau(i)\,\gamma_{\bar{A},t}\right)A_{M,n,t-1} \tag{A.6}$$

where $\eta \in (0,1)$, and $\tau(i)$ reflects two types of entrepreneurs: $\tau(i) = 1$ for those with upper-tail (scientific) knowledge, and $\tau(i) = 0$ for the remainder. $A_{M,n,t-1}$ is aggregate manufacturing productivity at location n in the previous period (described in more detail below). To interpret (A.6), consider first an entrepreneur with $\tau(i) = 0$. In this case, $\eta > 0$ guarantees that at least some innovation trickles through, and entrepreneurial productivity is the closer to the frontier the larger is η. We refer to this mechanism as (passive) catchup.

Next, consider highly skilled entrepreneurs with $\tau(i) = 1$. These also experience catchup, but in addition they actively improve their productivity, by the rate $\gamma_{\bar{A},t}$ relative to the initial local productivity $A_{M,n,t-1}$. We refer to this process as "knowledge effect" – highly skilled entrepreneurs improve local technology by keeping up with technical progress at the frontier. This reflects several dimensions of the historical evidence discussed in Section 2. First, more scientifically savvy entrepreneurs were more likely to know about the existence of new technologies, which reduced their search costs and raised the likelihood of *adoption*.[9] Second, they could *operate* modern technology more efficiently because of a better understanding of the underlying processes. Third, scientific knowledge made further *innovative* improvements more likely.[10] Importantly, the "knowledge ef-

[8]By taking the evolution of $\bar{A}_t$ as given, we abstract from the feedback mechanism in Unified Growth Theory whereby human capital drives *aggregate* technological progress (Galor, 2011). At the *local* level, however, our approach allows for upper-tail skills to accelerate productivity growth.

[9]As Mokyr (2000, p.30) put it: "Of course I do not argue that one could learn a craft just from reading an *Encyclopédie* article (though some of the articles in the *Encyclopédie* read much like cookbook entries). But ... once the reader knew what was known, he or she could look for details elsewhere."

[10]This interpretation is in line with Kelly et al. (2014), who argue that the Industrial Enlightenment generated ideas

fect" is the stronger the higher is $\gamma_{\bar{A},t}$. This reflects the argument by Nelson and Phelps (1966) that human capital – here in the form of its upper tail – is particularly useful in periods of rapid technological change.

We now turn to the evolution of aggregate manufacturing productivity at location n, $A_{M,n,t}$. This term corresponds to the average entrepreneurial productivities, as given by (A.4). Recall that at each location n, there is a share s_n of highly skilled entrepreneurs. Thus, integrating (A.6) over all entrepreneurs $i \in [0, 1]$ yields:

$$A_{M,n,t} = \eta \bar{A}_t + (1 - \eta) A_{M,n,t-1} \left(1 + s_n \cdot \gamma_{\bar{A},t}\right) \tag{A.7}$$

This equation illustrates three forces that drive manufacturing productivity at location n: First, passive catchup with the frontier, which depends on η. Second, the "knowledge effect," which is larger for regions with higher s_n, and larger when technological progress $\gamma_{\bar{A},t}$ is rapid. Third, there is also a spillover effect of entrepreneurs with upper-tail knowledge: they raise $A_{M,n,t}$, which is then reflected as $A_{M,n,t-1}$ in (A.6) in the following period, benefiting both entrepreneurs with and without scientific knowledge. Our setup also ensures that $A_{M,n,t} \leq \bar{A}_t$, which holds with equality in regions with $s_n = 1$. In other words, a region where all entrepreneurs have scientific knowledge will always be at the technological frontier.

Finally, we specify the productivity process in agriculture. We assume that upper-tail knowledge is not important in this sector.[11] However, some technologies from the frontier "trickle through" to agriculture, as well.[12] We model this process in the same fashion as for manufacturing, so that agricultural productivity in region n evolves according to

$$A_{A,n,t} = \eta \bar{A}_t + (1 - \eta) A_{A,n,t-1} \tag{A.8}$$

that were then implemented by entrepreneurs and scientists in the upper tail of the skill distribution. Similarly, Mokyr (2005) argues that technological progress often came in the form of micro-inventions by implementation of broader technological concepts.

[11] Compared to manufacturing, agriculture saw much less innovation that required advanced knowledge to be adopted. This pattern is clearly borne out by innovations exhibited at world fairs: Among the 6,377 exhibits at the 1851 Crystal Palace Exhibition in London, only 261 (or 4.1%) were agricultural machinery. At the other end of the spectrum, modern manufacturing sectors made up the large majority of innovations: textiles alone accounted for more than 26%, and engines and scientific instruments for another 15% (Moser, 2012, Table 3).

[12] This reflects the historical evidence that agricultural productivity also grew significantly during the industrial revolution (Crafts, 1985; Mokyr, 2010). Nevertheless, scientific knowledge did not play a role for technological progress in agriculture before the middle of the 19th century (see Appendix E.2). We note in passing that differential growth in agriculture and manufacturing is not essential for our results. Alternatively, the same productivity process in the two sectors, combined with non-homothetic demand, yields similar predictions due to high income translating into disproportionately more manufacturing demand.

Note that this equation corresponds to (A.7) with $s_n = 0$. Thus, in regions without upper-tail knowledge, productivity in agriculture and manufacturing are the same. This delivers a useful benchmark case for our analysis.

A.3 Equilibrium and Predictions

We now analyze how worker skills and upper-tail knowledge affect income, growth, and the sectoral allocation of labor. Importantly, we assume that (exogenous) technological progress at the frontier, $\gamma_{\bar{A},t}$, is initially slow and then accelerates. Growth in total factor productivity (TFP) was minuscule prior to industrialization – probably in the range of 0.1% per year (Galor, 2005) – and it then accelerated to approximately 1% in the mid-19th century (Crafts and Harley, 1992; Antràs and Voth, 2003). With low $\gamma_{\bar{A},t}$, equation (A.7) implies that upper-tail knowledge does not have important effects on regional productivity; it only matters when technology advances more quickly. This difference is crucial for our predictions before versus during industrialization.

Within each region n, labor mobility ensures that wages in agriculture and manufacturing are equalized: $w_{A,n} = w_{M,n} = w_n$. Using (A.2) and (A.5), this yields the employment share in agriculture:

$$l_{A,n} = \left(\frac{A_A}{(1-\alpha_M)A_M} h_n^{\beta_A - \beta_M} \right)^{\frac{1}{\alpha_A}} x_n \tag{A.9}$$

More land-abundant regions (higher x_n) have higher employment shares in agriculture. Since we assume that manufacturing production occurs in cities, the urbanization rate is given by $l_{M,n} = 1 - l_{A,M}$. In addition, equation (A.5) implies that wages grow at the same rate as local manufacturing productivity $A_{M,n}$.[13] The growth rate is thus given by

$$\gamma_{w,n,t} = \eta \left(\frac{\bar{A}_t}{A_{M,n,t-1}} - 1 \right) + (1-\eta)\, s_n \cdot \gamma_{\bar{A},t} \tag{A.10}$$

where the first term reflects the catchup effect, and the second term, the "knowledge effect."

We now present three predictions of the model. The first two analyze the cross-sectional effect of knowledge elites for the cases of relatively low $\gamma_{\bar{A}}$ (before the Industrial Revolution), and for high $\gamma_{\bar{A}}$ (in the industrial period). The third prediction highlights the role of worker skills. We discuss the intuition behind each prediction in the text.

Prediction 1. Pre-industrialization: *If the technological frontier expands slowly (low $\gamma_{\bar{A}}$), labor shares in manufacturing, wages, and economic growth are only weakly affected by local upper-tail knowledge.*

[13] Total income in region n also comprises entrepreneurial profits given by $\alpha_M(1-\alpha_M)Y_{M,n}$ (see Appendix B.1). Since profits are directly proportional to wages, we focus on the latter when discussing the model predictions.

Intuitively, if technological progress is slow, entrepreneurs with upper-tail knowledge enjoy only a tiny productivity advantage (or none at all, if $\gamma_{\bar{A}} = 0$). Thus, (A.7) implies that productivity is similar or identical in regions with high and low s_n. Consequently, wages and labor shares – given by (A.5) and (A.9) – are also similar in the cross-section. The same is true for income growth, given by (A.10). The left panel of Figure A.1 provides an illustration of Prediction 1. Under reasonable parameter choices, the percentage of entrepreneurs with scientific knowledge has only minuscule effects on development.[14] Finally, it is important to note that there were exceptions to the generally slow growth before Industrialization. One example is the Commercial Revolution in early modern Europe. In line with our model, there is historical evidence that advanced knowledge mattered during this period (Dittmar, 2013; Cantoni and Yuchtman, 2014).

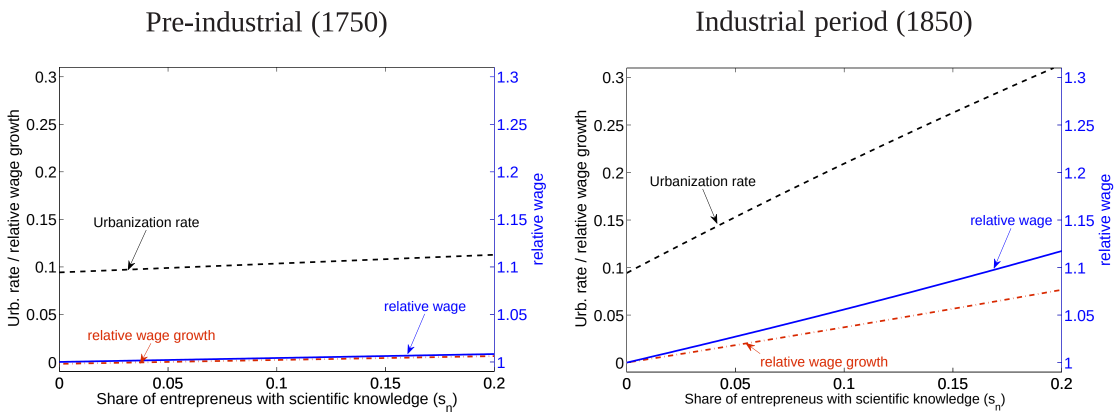

Figure A.1: Scientific knowledge and economic development

Notes: The figure illustrates how the share of entrepreneurs with scientific knowledge in region n, s_n, affects urbanization, wages, and economic growth. The left panel refers to the pre-industrial period (illustrating model Prediction 1), and the right panel illustrates Prediction 2, referring to the industrial period. The urbanization rate corresponds to the labor share in manufacturing. Wages (right axis) are reported relative to regions without scientific knowledge ($s_n = 0$). Relative wage growth (left axis) is measured as annual percentage growth in region n, net of growth in regions with $s_n = 0$.

Next, we turn to the industrialization period, when technological knowledge grew rapidly. The following prediction shows that, despite production knowledge being non-rival and available to all regions, it can have differential effects on economic development.

Prediction 2. During and after industrialization: *As the technological frontier expands more rapidly (high $\gamma_{\bar{A}}$), a larger local knowledge elite leads to higher wages, higher manufacturing employment, and faster economic growth.*

[14] The model calibration serves mainly illustrative purposes – we do not intend to precisely predict actual magnitudes of effects. Appendix B.3 explains our parameter choices in detail. We simulate the model for 250 periods, corresponding to 1600-1850.

The intuition for this prediction follows the same logic as above, but now with rapid technological growth at the frontier, so that upper-tail knowledge has sizeable effects on regional productivity. The right panel of Figure A.1 illustrates the prediction in the simple calibrated version of our model: both wages and manufacturing employment now grow hand-in-hand with the local density of scientific knowledge.

Finally, we describe the effect of worker skills on income and employment.

Prediction 3. Effect of worker skills: *Higher average worker skills h_n in region n lead to higher employment shares in manufacturing and higher regional wages, but not to faster growth. This holds irrespective of the rate of technological progress at the frontier.*

Figure A.2 illustrates this prediction. Regional wages in both sectors grow in worker skills, as given by (A.2) and (A.5). In addition, worker skills are more important in manufacturing than in agriculture ($\beta_M > \beta_A$). Thus, following equation (A.9), higher h_n leads to a concentration of employment in manufacturing – and thus in cities. Since these effects are independent of scientific knowledge, we can plot the pre-industrial and industrial periods together. Finally, wage growth as given by (A.10) is independent of worker skills. Intuitively, worker skills affect how productively a *given* technology is operated, but not which technology is used.

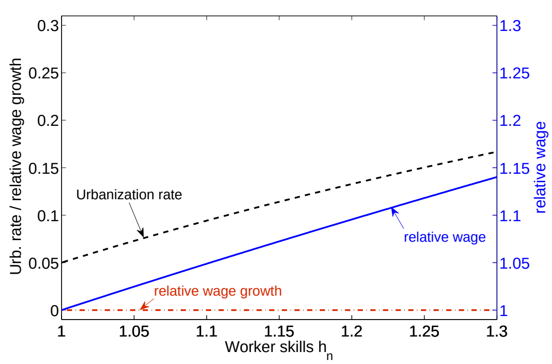

Figure A.2: The role of worker skills

Notes: The figure illustrates model Prediction 3, showing how worker skills in region n, h_n, affect urbanization, wages, and economic growth. Since the effect of worker skills does not change over time, the figure illustrates both the pre-industrial period and the industrial period. See Figure A.1 for a description of the three depicted variables.

Summing up, as compared to the previous theoretical literature, our model provides a more differentiated view on the role of human capital during industrialization. Distinguishing between worker skills and upper-tail (scientific) knowledge allows us to derive predictions that differ importantly for the two types of human capital.

B Model: Technical Results and Proofs

B.1 Aggregate Manufacturing Production Function

Demand for intermediate product i in region n follows from (A.3) and is given by:

$$z_n(i) = \xi A_{M,n}(i) h_n^{\beta_M} L_{M,n} \left(\frac{1}{p(i)} \right)^{\frac{1}{1-\alpha_M}} \tag{B.1}$$

Since each entrepreneur i faces a constant marginal cost of one unit of the final good, his total costs are given by $1 \cdot z_n(i)$. Consequently, entrepreneurs choose their price $p(i)$ so as to maximize profits $\pi_n(i) = (p(i) - 1) \cdot z_n(i)$, subject to (B.1). This yields:

$$p(i) = \frac{1}{\alpha_M}, \; \forall i \tag{B.2}$$

Using this in (A.3) and choosing the constant $\xi \equiv \alpha_M^{-\frac{2 \cdot \alpha_M}{1-\alpha_M}}$ implies that total manufacturing output in region n is given by

$$Y_{M,n} = A_{M,n} h_n^{\beta_M} L_{M,n}, \text{ with } A_{M,n} \equiv \int_0^1 A_{M,n}(i) di \tag{B.3}$$

Consequently, aggregate manufacturing productivity is a simple linear combination of individual entrepreneurial efficiencies. Note from (A.3) that a share $1 - \alpha_M$ of total output is paid to labor: $w_{M,n} L_{M,n} = (1 - \alpha_M) Y_{M,n}$, so that the wage rate in region n is given by

$$w_{M,n} = (1 - \alpha_M) A_{M,n} h_n^{\beta_M} \tag{B.4}$$

This completes the set of equations used in the main analysis. For completeness, we also present the total profits by entrepreneurs, as well as output net of intermediate inputs. Total profits are given by $\Pi_n = \int_0^1 \pi_n(i) di$, with $\pi_n(i)$ given above. Substituting (B.1) and (B.2), we obtain:

$$\Pi_n = \alpha_M (1 - \alpha_M) Y_{M,n} \tag{B.5}$$

Finally, note that the measure of total output given by (B.3) still includes the part that is used for intermediate production. Net output is given by $Y_{M,n}^{net} = Y_{M,n} - \int_0^1 z(i) di$. Using (B.1) and (B.2) this yields:

$$Y_{M,n}^{net} = \left(1 - \alpha_M^2\right) Y_{M,n} \tag{B.6}$$

It is straightforward to verify that total wage payments and profits add up to $Y_{M,n}^{net}$.

B.2 The Evolution of Productivity in Manufacturing

Aggregate productivity follows from (B.3), by integrating equation (A.6) from the paper over all entrepreneurs $i \in [0,1]$. Without loss of generality, we assume that entrepreneurs $i \leq s_n$ have scientific knowledge, and the remainder does not. This yields:

$$A_{M,n,t} = \int_0^{s_n} A_{M,n,t}(\tau(i)=1)di + \int_{s_n}^1 A_{M,n,t}(\tau(i)=0)di \tag{B.7}$$

Substituting for $A_{M,n,t}(i)$ from (A.6) then implies:

$$A_{M,n,t} = \eta\bar{A}_t + (1-\eta)A_{M,n,t-1}\left(1 + s_n \cdot \gamma_{\bar{A},t}\right) \tag{B.8}$$

B.3 Calibration

To calibrate our model, we choose $\beta_A = 0.3$, $\beta_M = 0.5$, and $\alpha_A = 0.6$ (α_M is a free parameter, chosen implicitly with x_n). For the catchup parameter, we use $\eta = 0.02$, which implies that regions with $s_n = 0$ lag about 35 years behind the frontier. We simulate the model for 250 periods, corresponding to 1600-1850. Over the first 100 periods, we use $\gamma_{\bar{A}} = 0.1\%$; thereafter, $\gamma_{\bar{A}}$ rises by 0.015% each year, so that it reaches 2.35% after 250 periods. This is consistent with the combined contribution of TFP and physical capital to growth during the mid-19th century (Antràs and Voth, 2003).[15] Our calibration of $\gamma_{\bar{A}}$ ensures that wages follow the trend of p.c. income shown in appendix Figure C.1. We normalize $\bar{A}_M = 1$ in 1700 and choose land x_n such that the labor share in (overall) manufacturing is approximately 10% – this corresponds to the urbanization rate in France in 1700.[16] In Figure A.1, we keep average worker skills constant at $h_n = 1.1$ and use the range of $s_n \in [0, 0.2]$. The latter reflects the historical accounts that only a small share of French entrepreneurs was progressive and possessed upper-tail knowledge. Changing either h_n or the interval for s_n does not affect our qualitative predictions.

When simulating the effect worker skills, we keep scientific knowledge constant at the benchmark level $s_n = 0$. The range $h_n \in (1.0, 1.3)$ on the horizontal axis of Figure A.2 is chosen as follows: the average return to schooling in a cross-section of countries is about 0.1 (Bils and Klenow, 2000). In our sample, literacy in 1686 varies between 0 and 60% across French departments, and we assume that literacy is equivalent to 5 years of schooling. With this, we obtain an upper bound of $\exp(0.1 \cdot 0.6 \cdot 5) \approx 1.3$, and a lower bound of $\exp(0) = 1$. Changing these values

[15] Since we abstract from physical capital in the model, we have to load its effect on TFP growth if we want to match historical data.

[16] This is computed using total city population from Bairoch, Batou, and Chèvre (1988), divided by French population from McEvedy and Jones (1978).

does not alter our qualitative results.

C Data

C.1 Growth of per Capita Income in France, 1650-1900

Figure C.1 shows GDP per capita in France over the period 1650 to 1900. Overall, the data confirm Roehl's (1976) assessment that there is no clear "take-off" point. Nevertheless, growth became steady in the mid-18th century and only slowed down temporarily during the decade after the French Revolution.

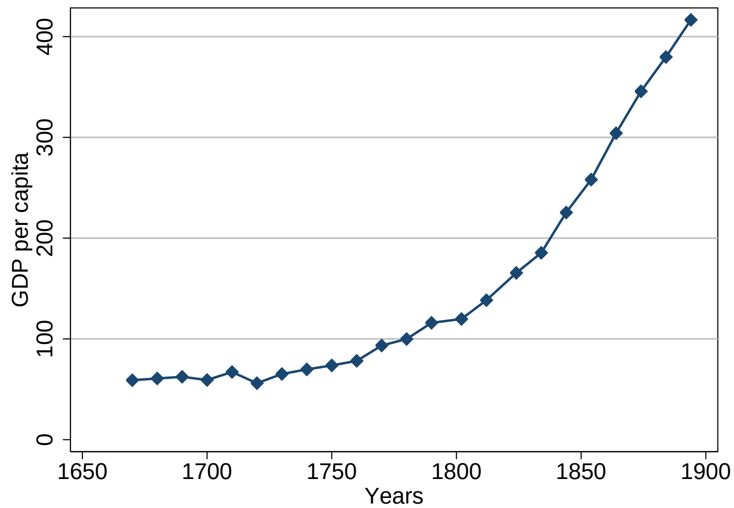

Figure C.1: Trend in per capita GDP in France, 1670-1890

Notes: The figure shows the evolution of French GDP per capita over the period 1670-1890. GDP per capita is computed using data on gross output (at current prices) and weighted price deflators from Marczewski (1961), combined with population data from McEvedy and Jones (1978). Linear interpolation is used for decades with missing data. The unit of measurement is Francs (in 1905-1913 prices); income in 1820 corresponds to approximately 1,100 International (1990) dollar, according to Maddison (2007).

Industrialization was not homogenous across France, as shown in Figure C.2. Both wages in industry (left panel) and employment shares (right panel) show substantial dispersion across departments in the mid-19th century. This is also evident in the corresponding histograms (Figure C.3). We exploit this spatial heterogeneity in our empirical analysis.

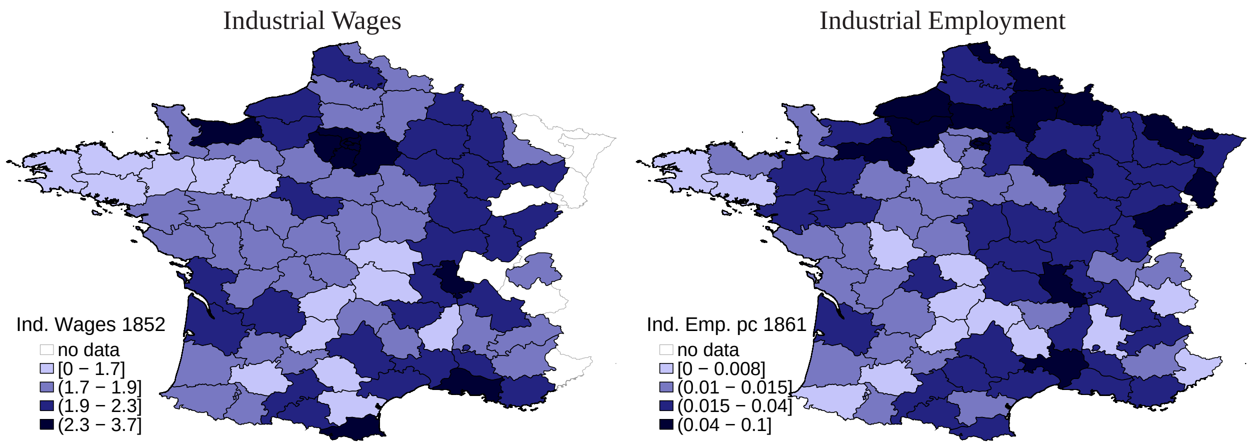

Figure C.2: Dispersion of Industrial Wages and Industrial Employment

Notes: The left panel shows the spatial distribution of industrial wages in 1852. The right panel shows the distribution of industrial employment per capita in 1861. Data on industrial wages are from Goreaux (1956), and data on industrial employment in 1861 are from the *Statistique Industrielle* of the Statistique Général de France (1861). Both variables are described in Section 3.2.

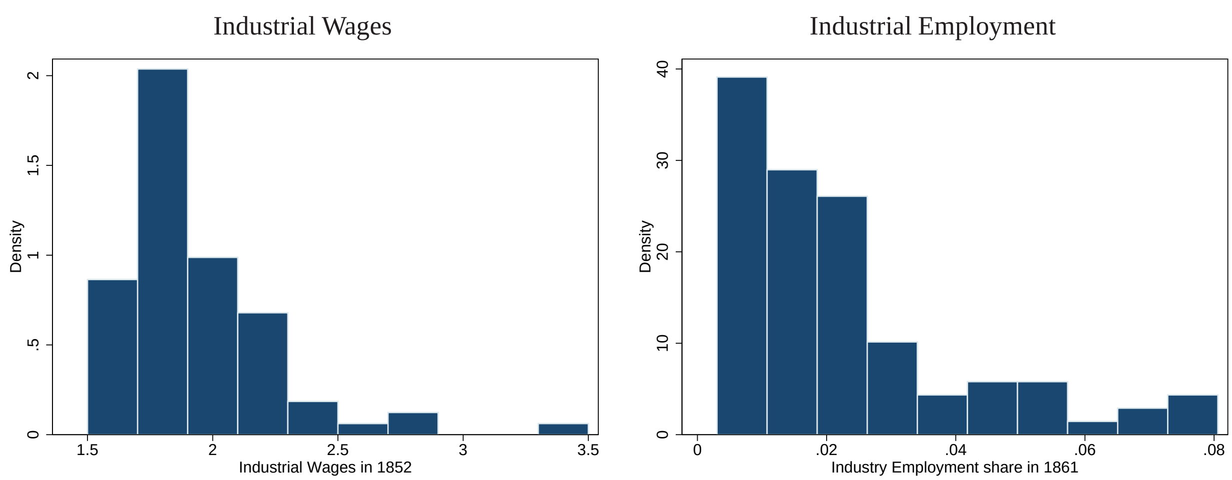

Figure C.3: Dispersion of Industrial Wages and Industrial Employment

Notes: The left panel shows distribution of industrial wages by department in 1852, and the right panel shows the distribution of industriaial employment per capita in 1861. See the note to Figure C.2 for detail.

C.2 Subscriber Density: Distribution, Alternative Definitions, and Historical Detail

Spatial distribution of subscribers to the Encyclopédie

Figure C.4 shows the correlation between subscriber density and literacy – in 1686 (left panel) and in 1786 (right panel). The two variables are uncorrelated.

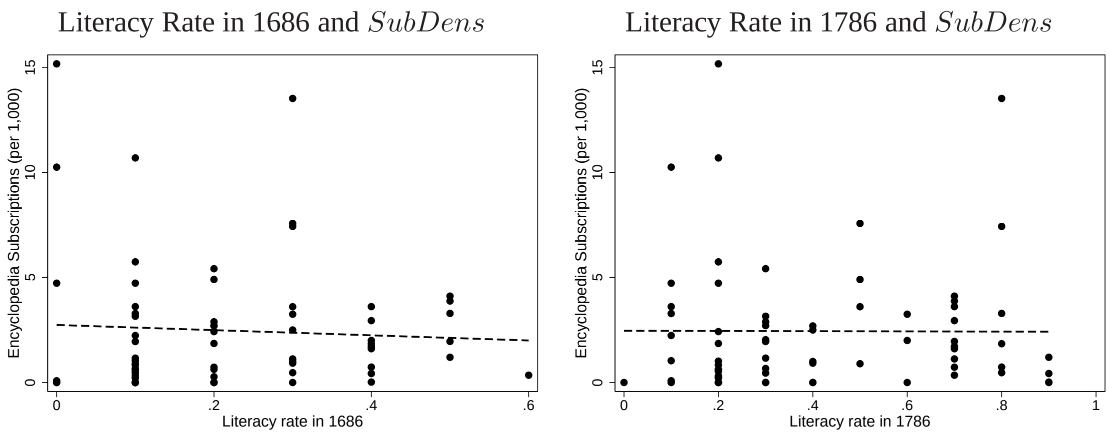

Figure C.4: Correlation between Literacy and $SubDens$

Notes: The figure shows the correlation between subscriber density and literacy in 1686 (left panel) and between subscriber density and literacy in 1786 (right panel). Subscriber density is defined as $SubDens = Subs/pop_{1750}$.

Next, Figure C.5 shows the distribution of our main explanatory variable, subscriber density, for all cities (panel A) and for cities with positive subscriptions (panel B). In the following, we introduce two alternative definitions. First, the left panel of Figure C.6 shows subscriptions per 1,000 city inhabitants without taking logs. Comparing this with our main measure illustrates that taking logs tightens the distribution, restricting the extent to which extreme values affect our results.

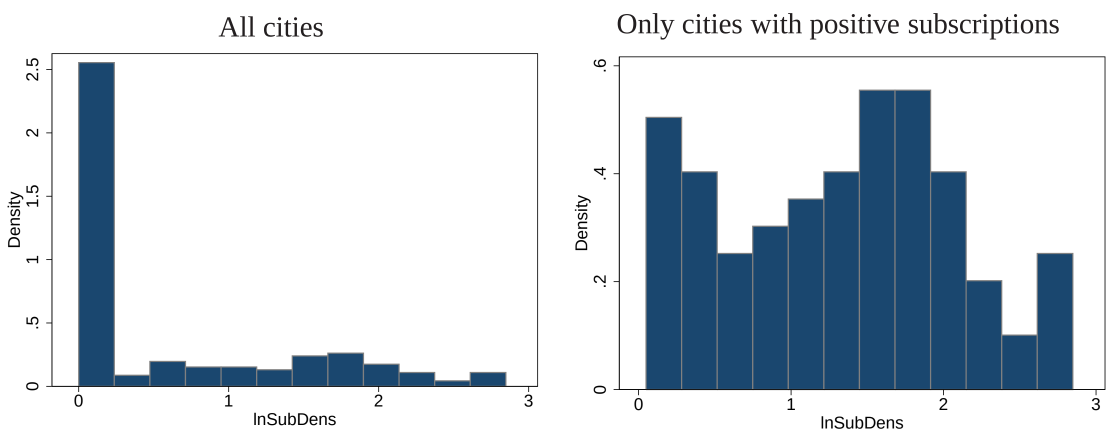

Figure C.5: Baseline measure of subscriber density: $lnSubDens$

Notes: The figure shows the distribution of our main explanatory variable, subscriber density, defined as $lnSubDens = ln(Subs/pop_{1750} + 1)$. The right panel plots the histogram for all cities, and the left panel plots the histogram only for cities with positive subscriptions.

Alternative measures of subscriber density

For cities without subscriptions, our baseline measure of subscriber density does not exploit all available variation: by assigning zero density throughout, it does not take into account city size. For example, the zero subscriptions for 22,000 inhabitants in Arles (in the South of France) arguably reflect a lower density of scientific knowledge than $Subs = 0$ in the town of Saverne (in the North-East) with 1,000 inhabitants in 1750. To exploit this additional information, we define an alternative variable as follows: $lnSubDens2 = \ln[(Subs + 1)/pop_{1750}]$, where pop_{1750} is city population in 1750. This introduces variation across cities with zero subscriptions: $lnSubDens2$ is the smaller the larger is city population. The resulting distribution is presented in the right panel of Figure C.6. Clearly, there is now more variation in the left part. Below in Appendix D.1, we show that the two alternative measures yield very similar results compared with our main measure $lnSubDens$.

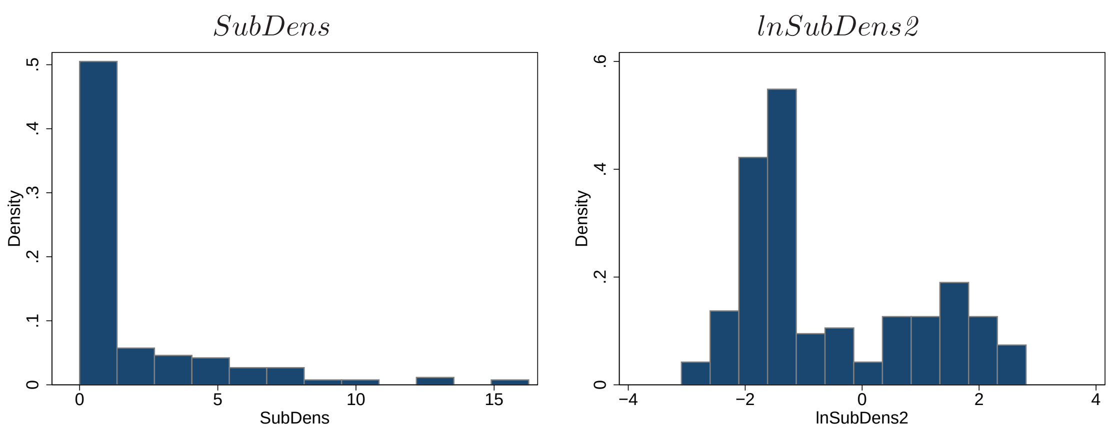

Figure C.6: Subscriber density: with and without variation across towns with $Subs = 0$

Notes: The left panel shows subscriber density defined as $SubDens = Subs/pop_{1750}$. In this setup, all cities with zero subscriptions have the value $SubDens = 0$. The right panel plots the histogram for $lnSubDens2 = \ln[(Subs + 1)/pop_{1750}]$, which allows for variation between cities with zero subscriptions, depending on population size.

Individual subscribers, book sellers, and cities with zero subscriptions

For a small subset of subscribers to the Quarto edition of the *Encyclopédie*, an individual list survived. The list was compiled by the bookseller Lépagnez and attached to the first volume of the *Bibliothèque Nationale* set of the Quarto edition in 1777. The list is reprinted in (Lough, 1968, pp.466-473). This list comprises customers of two booksellers, Lépagnez of Besançon (department Doubs) and Chaboz of Dole (department Jura – adjacent to Doubs). There are altogether 253 subscribers, with information on profession and social status, as well as place of residence. Out of the 253 subscribers, 175 (69.2%) lived in the same cities as the booksellers – Besançon or Dole. In addition, the vast majority of the remaining subscribers lived in the same departments (Jura and Doubs) in towns or villages that were too small to enter the Bairoch et al. (1988) dataset. This supports our convention to assign zero subscriptions to those cities in Bairoch et al. for which no sales are reported. While this may introduce measurement error, it is unlikely to systematically affect our results: in Appendix Table D.3 we show that our results are equally strong for small, medium, and large cities (for which this type of measurement error is increasingly unlikely). In addition, since those subscriptions that were sold outside the city boundaries went to buyers in the immediate neighborhood, our department-level analysis captures most of the remaining variation.

C.3 Soldier Height

Data on soldier height before 1750 are from Komlos (2005), who digitized height data and city of birth for more than 38,000 French soldiers from 1716 to 1784 – these include conscripts born between 1650 and 1770. In many cases, names of birthplaces are not reported, unrecognizable, or ambiguous. Among those that can be clearly identified, we matched soldiers to all towns and cities listed in Bairoch et al. (1988). In addition, for those cities not listed in Bairoch et al. (1988), we hand-coded the department for all towns with more than 10 soldiers in the dataset. Overall, this yields observations on about 22,000 soldiers in all 90 departments. Approximately 19,000 of these were drafted prior to 1750. From these data, we compute the average height at the department level before 1750 (i.e., we drop all drafts that occurred in 1750-84). Since average height may have changed over the period 1716-49, we filter out birth cohort fixed effects by decade. We also address the possibility that the age at recruitment may have differed across regions, which in turn may affect conscript height because body growth continues into the early 20s. We filter out age-specific patterns by controlling for age and age^2. Table C.1 reports the results of these regressions.

C.4 Control Variables: Sources and Descriptions

Baseline controls

Our baseline set of controls includes various geographic characteristics of cities, such as dummies for cities with ports on the Atlantic Ocean and on the Mediterranean Sea, as well as for cities located on navigable rivers. These data are from Dittmar (2011). Moreover, we use a set of "early knowledge controls": a dummy for cities that hosted a university before 1750 (Jedin, Latourette, and Martin, 1970; Darby and Fullard, 1970), a dummy for cities that had a printing press between 1450 and 1500 (Febvre and Martin, 1958; Clair, 1976), and the number of editions printed before 1501 (ISTC, 2008). We construct a dummy for cities located in non-French speaking departments using linguistic data from http://www.lexilogos.com/france_carte_dialectes.htm. There are three main groups of romance languages in France: langue d'oc, langue d'oil (the official French), and langue francoprovencal. We consider all three "French." By this definition, the following dialects are "non-French": Alsacien, Basque, Breton, Catalan, and Corsican.

STN books sales

Data on book sales to each French city are from the FBTEE (2012) project that reconstructed the book trade of the important Swiss publishing house STN (which also published the *Encyclopédie*) over the period 1769-1794. This source (available at http://chop.leeds.ac.uk/stn/) covers the sales of more than 400,000 copies, for which the location of the buyer (private or book traders, with the vast majority located in France) is available. Note that STN book sales are directly comparable to our data on *Encyclopédie* subscriptions, since both occurred during the same period, were shipped from one place (Switzerland and Lyon) towards all of France, and reflected local demand for books (sales) rather than supply (printing).

Pays d'élection

While France was a centralized state already before the French Revolution, in some regions – the so-called *pays d'élection* – the king exerted particularly strong power in fiscal and financial matters: in the *pays d'élection* a representative of the royal administration was responsible for the assessment and collection of taxes. In contrast, the *pays d'état* and the *pays d'imposition* enjoyed higher autonomy in terms of taxation. *Pays d'élection* correspond to *généralités* – an administrative region of the Ancien Régime that does not always overlap with department borders. We construct a dummy for cities located in *pays d'élection* based on the maps provided at http://fr.wikipedia.org/wiki/Généralité_(France), in combination with the city-*généralités* correspondence in the atlas by Brette (1904). When running regressions at the department level, we compute for each department the share of cities located in *pays d'élection*.

Pre-industrial activity

We follow Abramson and Boix (2013) to construct data on pre-industrial centers in France, based on the original sources Carus-Wilson (1966) and Sprandel (1968). These include the total number of mines, forges, iron trading locations, and textile manufactures. We sum the number of such locations per department and calculate the local density of pre-industrial activities (in the same fashion as subscriber density). About half of the departments have some type of pre-industrial activities, with the highest numbers in the departments of Isère, Nord and Pas de Calais.

Local density of the nobility

Data on the number of noble families is provided by the Almanach de Saxe Gotha.[17] More specifically, we use data on *marquises*. Entries also contain information on the departments of origin of these families, as well as the dates of creation and (if applicable) extinction of the dynasty. For each department, we compute the total number of *marquis* families existing in 1750. This number varies substantially, from zero in Corsica and less than 4 in the departments of Haut-Rhin, Vosges, Meuse, and Moselle (all located in the North East) to 32 and 39 noble families in the departments of Seine et Marne, and Seine et Oise. We then compute the local density of noble families $lnNoblesDens$ in the same way as subscriber density, but normalizing by department-level population in 1750.

C.5 From City-Level to Arrondissement- and Department-Level Data

In our arrondissement (Tables 11 and 12) and department level regressions (Tables 6 and 7), we aggregate city subscriptions to these geographic units. More specifically, we compute $SubDens_{arr}$ and $SubDens_{dep}$ as the average subscriptions per capita across all cities in a given arrondissement (department), where population corresponds to 1750 city inhabitants from Bairoch et al. (1988). We then derive $lnSubDens_{arr} = \ln(1 + SubDens_{arr})$ and $lnSubDens_{dep} = \ln(1 + SubDens_{dep})$.[18] Thus, arrondissements (and departments) without subscriptions receive a value of zero.

We also aggregate the city-level control variables listed in Table 1 to the arrondissement and department level. This procedure yields dummies for arrondissements (departments) having at least one Atlantic port, a Mediterranean port, or a navigable river, and a dummy for non-French speaking departments. Moreover, we also control for the potential confounding factors described in Section 3.3: the density of STN books traded in France is calculated in the same way as $lnSubDens_{dep}$, while the density of pre-industrial activities, the density of nobles in 1750, and the density of executed under the "Reign of Terror" are already defined at the department level. The aggregation for

[17] Available at http://en.wikipedia.org/wiki/List_of_French_marquisates#cite_note-1.

[18] Throughout the paper and appendix, we use $lnSubDens$ (i.e., without subscript) to refer to subscriber density at all levels of aggregation.

pays d'élection is less straightforward because these regions do not always overlap with department borders (see Appendix C.4 above). When running regressions at the department level, we thus compute for each department the *share* of cities located in *pays d'élection*.

C.6 Balancedness of the Sample

In the main text, we analyzed the correlation of subscriber density with other city characteristics. Here, we focus on the extensive margin, comparing cities with and without subscriptions. In Table C.2, columns 1-3 report the mean and standard deviations of subscriber density and our baseline controls, and column 4 shows the t-test for the difference in means between cities with and without subscriptions. On average, cities with subscriptions were significantly larger, with 26,000 vs. 6,300 inhabitants. This difference is not surprising, however, given that larger cities have many more potential subscribers. There is no significant difference with respect to location at Atlantic or Mediterranean ports, but with respect to location on navigable rivers. Cities with subscriptions are more likely to have a university prior to 1750, as well as a printing press and a higher number of books printed in 1500. These differences are statistically significant. Finally, cities with subscriptions also tend to have somewhat higher levels of literacy in 1786.

Because of its strong correlation with city population, the dummy for subscriptions that we used above will also capture other city characteristics that are related to size. Next, we restrict attention to cities with above-zero subscriptions and split the sample into those with above- and below-median subscriptions *per capita*. Columns 5-8 of Table C.2 show the results. For cities with above- and below-median subscriptions per capita, population is almost identical.[19] The same is true for all other controls, with a few exceptions that reflect local advanced knowledge: 34% and 39% of cities with above-median subscriptions are hosting a university and a printing press respectively, compared to 14% and 26% of cities with below-median subscriptions – the difference, however, is not significant for the "Printing Press" variable. Cities with below-median subscriptions tend to have slightly higher levels of literacy in 1786 (reversing the pattern from columns 2 and 3), but this difference is not statistically significant.

[19] We exclude Paris, because all city-level regressions include a dummy for the French capital. With its 570,000 inhabitants, Paris was five times larger in 1750 than the second-biggest city, Lyon. When including Paris cities with below-median subscriptions actually have a larger population on average, reversing the pattern from columns 1-3, while the remaining results are largely unchanged.

Table C.1: Controlling for the effect of age and birth decade on soldier height, 1660-1740

Dependent variable: Soldier height in cm	
Age	0.440***
	(0.021)
Age^2	-0.006***
	(0.000)
Birth Decade Dummy	Yes
R^2	0.07
Observations	29292

Notes: This regression is run by OLS. *Birth Decade Dummies* include dummies for all decades from 1660 to 1740. Robust standard errors in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table C.2: Subscriptions and control variables – descriptive statistics

	All cities				Cities with subs>0 (Paris excl.)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All	Subs>0	No Subs	t-test	All	Above median	Below median	t-test
Subscriptions per 1,000	1.73	3.94			3.97	6.81	1.27	10.26
	(3.14)	(3.71)			(3.72)	(3.40)	(0.99)	
Population in 1750	15	26.05	6.31	3.28	19.57	19.07	20.05	-0.24
	(42.57)	(62.52)	(3.05)		(18.68)	(21.45)	(15.85)	
Atlantic Port	0.06	0.08	0.05	1.03	0.08	0.07	0.09	-0.33
	(0.24)	(0.27)	(0.21)		(0.28)	(0.26)	(0.29)	
Mediterranean Port	0.03	0.04	0.03	0.3	0.04	0.02	0.05	-0.54
	(0.17)	(0.18)	(0.16)		(0.19)	(0.16)	(0.21)	
Navigable River	0.09	0.16	0.03	3.41	0.15	0.12	0.19	-0.81
	(0.28)	(0.37)	(0.16)		(0.36)	(0.33)	(0.39)	
University	0.11	0.25	0.01	5.53	0.24	0.34	0.14	2.21
	(0.32)	(0.43)	(0.10)		(0.43)	(0.48)	(0.35)	
Printing Press	0.18	0.33	0.06	5.01	0.32	0.39	0.26	1.32
	(0.39)	(0.47)	(0.25)		(0.47)	(0.49)	(0.44)	
ln(Books Printed 1500)	0.48	0.87	0.17	3.82	0.78	0.93	0.64	0.87
	(1.31)	(1.74)	(0.7)		(1.55)	(1.62)	(1.49)	
Literacy 1786	0.43	0.46	0.39	1.73	0.45	0.42	0.49	-1.32
	(0.25)	(0.25)	(0.25)		(0.25)	(0.24)	(0.25)	
Observations (Min/Max)	166/193	82/85	84/108		81/84	40/41	41/43	

Notes: This table compares the means of subscriber density, literacy, as well as our "Baseline" and "Early Knowledge" controls (as listed in Table 1). In columns 1-4 we use the full sample, including all cities (column 1), and then distinguish between cities with (column 2) vs. without subscriptions (column 3). In columns 5-8 we restrict the sample to all cities with positive subscriptions (column 5), and we distinguish between cities with above- (column 6) vs. below-median (column 7) subscriptions per capita. Columns 4 and 8 present t-tests for the difference in means between cities with vs. without subscriptions and cities with above- vs. below- median subscriptions, respectively. For details on *Encyclopédie* subscriptions and controls see the notes to Table 1, Section 3, and Appendix C.4.

D Additional Empirical Results

D.1 Additional Results for City Growth

We now perform a series of robustness checks and show that our baseline results on city growth (Table 3) are not driven by sample composition or by a particular specification of our variable of interest.

Comparing cities with and without subscribers – matching illustrated

Figure D.1 illustrates our matching results for cities with vs. without subscribers. We use propensity score matching to find, for each city *with* subscribers, the closest match in terms of population in 1750 among cities *without* subscribers. This gives two samples of equal size (each with 84 cities). We then plot the average population growth over the period 1750-1850 for the cities without subscribers (left bar), and for those with at least one subscriber (right bar). The 95% confidence intervals shown in the figure illustrate that cities with subscribers grew significantly faster than their counterparts with similar population size but without subscribers.

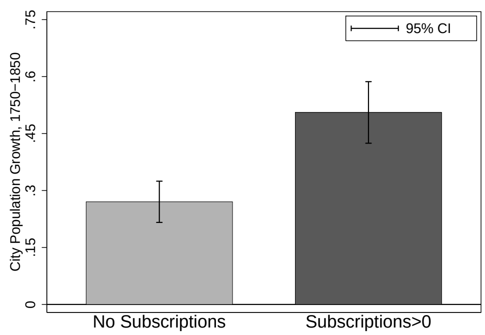

Figure D.1: *Encyclopédie* subscriptions and city growth, 1750-1850

Notes: The figure plots the average population growth over the period 1750-1850 for cities without *Encyclopédie* subscribers (left bar), and for those with at least one subscriber (right bar). The sample of cities without subscribers is obtained using propensity score matching based on population size: by computing for each city *with* subscribers the closest match in terms of population in 1750 among cities *without* subscribers. There are 84 cities in each subsample. Paris is excluded because it had 4 times the population of the second-largest city (Lyon) in 1750.

Splitting the main period 1750-1850

Table D.1 splits our main period of analysis into two sub-periods, 1750-1800 (column 1) and 1800-1850 (columns 2 and 3). This roughly captures the pre- and post-Revolution period, with major differences in institutions. We find that our results hold for both sub-periods (columns 1 and 2). In column 3, we also control for growth in the earlier period and find a negative and significant coefficient.[20] This suggests that unobserved factors that determined city growth prior to 1800 did not foster growth thereafter, which is in line with a structural break after the French Revolution. Nevertheless, cities with higher subscriber density grew faster under both regimes. This makes it unlikely that subscriber density reflects unobserved institutions that in turn drive growth, complementing our results for *pays d'élection*.

Table D.1: City growth over the sub-periods 1750-1800 and 1800-1850

Dependent variable: log city growth over the indicated period

	(1)	(2)	(3)
Period	1750-1800	1800-1850	1800-1850 Control for prior growth
lnSubDens	0.108***	0.059**	0.073***
	(0.035)	(0.028)	(0.027)
Growth 1750-1800			-0.187***
			(0.069)
Controls	✓	✓	✓
R^2	0.29	0.39	0.42
Observations	192	192	192

Notes: All regressions are run at the city level, include a dummy for Paris, and are weighted by population in 1750. The dependent variable is log city population growth over the period indicated in the table header. "Controls" include the baseline controls and early knowledge controls listed in Table 1 (columns 2 and 3 control for initial population in 1800 instead of 1750). For details on $lnSubDens$ and controls see the notes to Table 1. Robust standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Stable subsamples

In Table D.2, we run our growth regression for the 1750-1850 period, using four different subsamples, each including those cities for which population data are available in the year 1400, 1500, 1600, and 1700 respectively. For comparison, we also report the results for the full sample in column 1 (which yields the same coefficient as our baseline specification in column 3 in Table 3).

[20]This is unlikely to be driven by reversion to the mean, i.e., by fast initial growth being mechanically followed by slower subsequent growth, because the regression separately controls for initial log population in 1800.

The coefficients of $lnSubDens$ are very similar in magnitude to those of Table 3 and are always significant at the 1% level.

Table D.2: Restricting the sample to cities with available data for the pre-1750 period

Dependent variable: log city growth, 1750-1850

	(1)	(2)	(3)	(4)	(5)
Data on city size in	1750	1700	1600	1500	1400
lnSubDens	0.169***	0.154***	0.139***	0.157***	0.193***
	(0.033)	(0.036)	(0.042)	(0.037)	(0.052)
Controls	✓	✓	✓	✓	✓
R^2	0.36	0.39	0.57	0.55	0.57
Observations	193	148	58	62	50

Notes: All regressions are run at the city level, include a dummy for Paris, and are weighted by city population in 1750. The dependent variable is log city population growth in 1750-1850. We use cities where data on population from Bairoch et al. (1988) are available over for the year indicated in the header. "Controls" include the baseline controls and early knowledge controls listed in Table 1. For details on $lnSubDens$ and controls see the notes to Table 1. Robust standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Figure D.2 uses the consistent sample of cities for which population growth can be computed both in 1700-50 (pre-industrial) and in 1750-1850 (i.e., the sample used in column 1 of Table D.2. The figure confirms the emergence of a strong relationship between subscriber density and city growth after 1750.

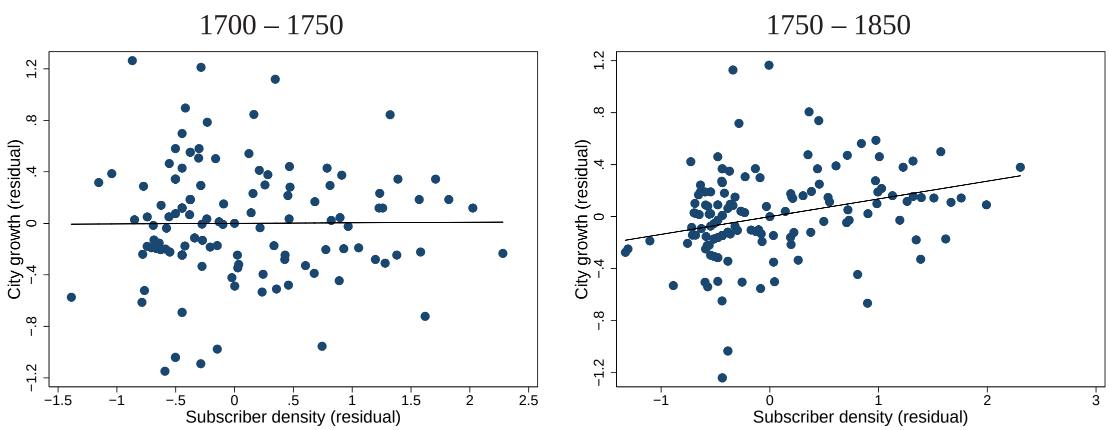

Figure D.2: *Encyclopédie* subscriptions and city growth – before and after 1750

Notes: The figure plots average annual city growth in France against *Encyclopédie* subscriber density ($lnSubDens$), after controlling for our baseline controls (listed in Table 1). The left panel uses the period before industrialization set in (1700-1750). The right panel examines the same cities over the period of French industrialization, 1750-1850. The sample is the same in both panels, including only cities for which growth can be computed over both periods. Among these, average annual city growth was 0.28% and 0.38% over the two periods, respectively.

Sample split by initial city size

In Table D.3 we show that our main result holds for small, medium, and large cities. We split the sample by terciles of initial city population in 1750. The coefficient on subscriber density is largest within the first tercile, and of similar magnitude for medium- and high-population cities. Figure D.3 presents the corresponding partial scatterplots, showing that the relationships in the subsamples are not driven by outliers.

Table D.3: Subscriptions for small, medium and large cities

Dependent variable: log city growth, 1750-1850			
	(1)	(2)	(3)
City size in 1750	Small	Medium	Large
lnSubDens	0.325***	0.149***	0.136**
	(0.079)	(0.053)	(0.052)
Controls	✓	✓	✓
R^2	0.36	0.25	0.51
Observations	69	64	60

Notes: All regressions are run at the city level, include a dummy for Paris, and are weighted by city population in 1750. The dependent variable is log city population growth in 1750-1850. The sample is split by terciles of city population in 1750. The first tercile (small) include 69 cities with a population between 1,000 and 5,000 inhabitants; the second tercile (medium) includes 64 cities with a population between 6,000 and 10,000 inhabitants; the third tercile (large) includes 60 cities with a population above 10,000 inhabitants. “Controls” include the baseline controls and early knowledge controls listed in Table 1. For details on $lnSubDens$ and controls see the notes to Table 1. Robust standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

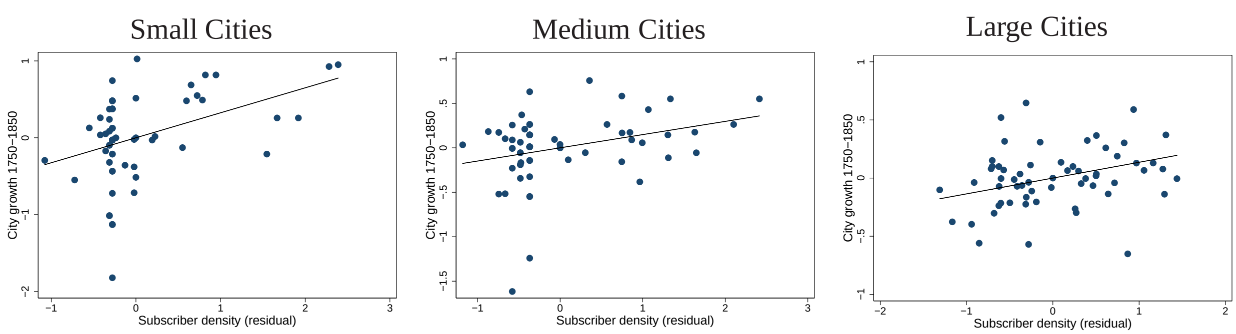

Figure D.3: Subscriber density and city growth for small, medium and large cities

Notes: The figure provides the partial scatterplots for *Encyclopédie* subscriber density ($lnSubDens$) corresponding to the regressions in columns 1-3 in Table D.3 City sizes correspond to population terciles in 1750.

Dummies for increasing subscriber density

In the main text, we have distinguished between the intensive and extensive margin of subscriber density by including the dummy $I_{Subs>0}$ for above-zero subscriptions to capture the former, and subscriber density $lnSubDens$ to reflect the latter. We found that the intensive margin was important, i.e., that additional subscriptions beyond the first are strongly associated with growth (see for example column 4 of Table 3 or column 4 of Table 5). These previous results were based on a linear relationship. In Table D.4, we provide additional evidence that the intensive margin matters, without imposing linearity. In addition to an indicator for zero subscriptions ($I_{Subs=0}$), we now include two dummies for cities with above-zero subscriptions, $I_{Subs>0,b.m.}$ for those with below-median subscribers per capita, and $I_{Subs>0,a.m.}$ for those with $SubDens > 0$, but above the median. We find that the coefficients increase for each step, and that these differences are statistically significant (as shown by the p-values in the bottom of the Table D.4).

Table D.4: Dummies for increasing subscriber density

Dependent variable: log city growth, 1750-1850		
	(1)	(2)
$I_{Subs=0}$	0.414***	0.468***
	(0.091)	(0.105)
$I_{Subs>0,\ b.m.}$	0.635***	0.711***
	(0.114)	(0.132)
$I_{Subs>0,\ a.m.}$	0.792***	0.855***
	(0.134)	(0.147)
Baseline Controls	✓	✓
Early Knowledge Controls		✓
R^2	0.73	0.73
Observations	193	193
p-values, test of difference in coefficients		
$I_{Subs=0} = I_{Subs>0,\ b.m.}$	[0.002]	[0.001]
$I_{Subs>0,\ b.m.} = I_{Subs>0,\ a.m.}$	[0.025]	[0.026]

Notes: All regressions are run at the city level, include a dummy for Paris, do not include a constant term and are weighted by city population in 1750. The dependent variable is log city population growth in 1750-1850. We use three indicator variables to classify cities based on their Encyclopdia subscriptions per capita: $I_{Subs=0}$ takes on value 1 for cities without subscriptions, $I_{Subs>0,\ b.m.}$ takes on value 1 for cities with positive subscription, but below the median for all cities with $Subs > 0$ (from 0.5 to 3.25 subscriptions per 1,000 inhabitants), and $I_{Subs>0,\ a.m.}$ takes on value 1 for cities with positive and above-median subscriptions per capita (from 3.35 to 16.25 per 1,000). There are 108 cities with zero subscriptions, 43 with positive and below-median subscriptions per capita and 42 cities with positive and above-median subscriptions per capita. "Baseline Controls" and "Early Knowledge Controls" are those listed in Table 1. For further detail see the notes to Table 1. Robust standard errors in parentheses.* p<0.1, ** p<0.05, *** p<0.01.

Restricting the sample to $Subs > 0$, and alternative definitions of subscriber density

In Table D.5 we check the robustness of our results on city growth to alternative definitions of subscriber density and different samples. Panel A uses our standard measure of subscriber density, $lnSubDens$, but restricts the sample to cities with positive subscriptions. Panel B uses subscriber density without logs ($SubDens$) in the full sample. Panel C employs the alternative definition $lnSubDens2$ which allows for variation across cities with zero subscribers (see Appendix C.2). In all cases, our results continue to hold: subscriber density is positively and significantly associated with city growth over the 1750-1850 period, and the relationship is substantially weaker over the pre-industrial period 1700-1750 (column 5).

Table D.5: *Encyclopédie* subscriptions and city growth: Alternative definitions and samples

Dependent variable: log city growth, 1750-1850

	(1)	(2)	(3)	(4)	(5)
				[unweighted]	1700-1750
Panel A: $lnSubDens = ln(Subs/pop_{1750} + 1)$. Only cities with $Subs > 0$					
lnSubDens	0.090	0.135***	0.117**	0.086*	-0.069
	(0.063)	(0.051)	(0.045)	(0.044)	(0.053)
Baseline Controls		✓	✓	✓	✓
Early Knowledge Controls			✓	✓	✓
R^2	0.08	0.46	0.48	0.38	0.32
Observations	85	85	85	85	76
Panel B: $SubDens = Subs/pop_{1750}$. Full sample					
SubDens	0.021**	0.040***	0.040***	0.048***	0.002
	(0.009)	(0.009)	(0.008)	(0.008)	(0.009)
Baseline Controls		✓	✓	✓	✓
Early Knowledge Controls			✓	✓	✓
R^2	0.10	0.35	0.35	0.26	0.17
Observations	193	193	193	193	148
Panel C: $lnSubDens2 = ln[(1 + Subs)/pop_{1750}]$. Full sample					
lnSubDens2	0.061***	0.088***	0.086***	0.107***	0.003
	(0.022)	(0.018)	(0.017)	(0.019)	(0.021)
Baseline Controls		✓	✓	✓	✓
Early Knowledge Controls			✓	✓	✓
R^2	0.13	0.36	0.37	0.28	0.17
Observations	193	193	193	193	148

Notes: All regressions are run at the city level, include a dummy for Paris, and are weighted (except column 4) by city population of the respective period. The dependent variable is log city population growth over the period 1750-1850. Panel A uses our standard definition of subscriber density $lnSubDens$ and restricts the sample to cities with positive subscriptions. Panel B uses the definition of subscriber density without logs ($SubDens = Subs/pop_{1750}$), and Panel C uses a log-based specification that introduces variation across cities with zero subscriptions ($lnSubDens2 = ln((1 + Subs)/pop_{1750})$) – see Appendix C.2 for further description. "Baseline Controls" and "Early Knowledge Controls" are those listed in Table 1 (column 5 controls for initial population in 1700 instead of 1750). For further detail see the notes to Table 1. Robust standard errors in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Next, we check the robustness of our results for literacy and additional controls. Table D.6 repeats the main regressions from Table 5: odd columns correspond to the specification with literacy only (as in column 1 in Table 5), and even columns present the full specification with all additional controls and literacy (as in column 3 in Table 5). We also vary the sample and the measure of subscriber density: columns 1-2 use our main measure, $lnSubDens$, but restrict the sample to cities with positive subscriptions. Columns 3-4 use subscriber density without logs ($SubDens$) in the full sample, and columns 5-6 use $lnSubDens2$, which allows for variation across cities with zero subscribers (see Appendix C.2). In all cases, subscriber density remains a strong predictor of city growth after 1750, while the coefficient on literacy is negative and in some cases significant, confirming the findings in Table 5.

Table D.6: Alternative specifications of Table 5

Dependent variable: log city growth, 1750-1850

Subs. Density:	$lnSubDens$		$SubDens$		$lnSubDens2$	
Sample:	Subs>0		Full sample		Full sample	
	(1)	(2)	(3)	(4)	(5)	(6)
Subs. Density†	0.125**	0.154***	0.042***	0.048***	0.096***	0.101***
	(0.051)	(0.051)	(0.009)	(0.010)	(0.021)	(0.021)
Literacy 1786	-0.315*	-0.430**	-0.187	-0.214	-0.219	-0.271*
	(0.167)	(0.182)	(0.144)	(0.142)	(0.138)	(0.153)
lnSTNBooksDens		-0.020		-0.029		-0.016
		(0.022)		(0.024)		(0.021)
lnPreIndDens		0.707		1.217***		0.861**
		(0.629)		(0.404)		(0.398)
lnDistanceCoal		0.078		0.015		0.036
		(0.061)		(0.045)		(0.046)
Pays d'Eléction		-0.127		-0.017		-0.081
		(0.105)		(0.072)		(0.074)
lnNoblesDens		-0.172		0.166		0.129
		(0.185)		(0.111)		(0.111)
Controls	✓	✓	✓	✓	✓	✓
R^2	0.51	0.56	0.37	0.41	0.39	0.41
Observations	82	81	166	164	166	164

Notes: All regressions are run at the city level, include a dummy for Paris, and are weighted by city population in 1750. The dependent variable is log city population growth in 1750-1850. Columns 1-2 restrict the sample to cities with positive subscriptions. "Controls" include the baseline controls and early knowledge controls as listed in Table 1. For details on $lnSubDens$ and the other explanatory variables see notes to Table 1. Standard errors (clustered at the department level) in parentheses . * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

† Subs. Density represents three measures of city-level subscriber density to the Quarto edition of the *Encyclopédie*: $lnSubDens$ in columns 1-2, $SubDens$ in columns 3-4, and $lnSubDens2$ in columns 5-6. For details on the three measures see the text, as well as Section 3.1, Appendix C.2, and the notes to Table D.5.

Discussion of Results for Control Variables

In Table 5 in the paper, we included all additional control variables together, so that multicollinearity may potentially be the reason why most of them showed insignificant coefficients. In Table D.7, we include these controls one-by-one. We confirm the results from the paper: only pre-industrial activity is positively and significantly associated with city growth (column 1). This confirms the findings in Abramson and Boix (2013), who show that in Europe overall, industrialization was more likely to take place in territories with a proto-industrial base.[21] Note that controlling for early industrial centers does not alter the coefficient on subscriber density, suggesting that the two are parallel, rather than competing explanations (recall also the weak negative correlation between the two measures in Table 1). Comparing their magnitude, the standardized beta coefficients are 0.16 for pre-industrial activity, and 0.38 for subscriber density. Distance to coal (column 2), the reach of central institutions (*pays d'élection*, column 3), and nobility density (column 4) are not significantly associated with city growth. The latter is in line with the historical evidence discussed in Section 2.2 that only a progressive subset of the nobility (which was more likely to read the *Encyclopédie*) was involved in industrial activity. Note also that when all additional controls are included together (column 4 of Table 5), the individual coefficients remain largely unchanged.

[21] This work builds on a large literature that has examined proto-industrialization – a change in the spatial organization of the pre-industrial economy, when the rural labor force increasingly engaged in market-oriented craft production. According to Mendels' (1972) seminal contribution, this process "preceded and prepared modern industrialization proper" [p. 241]. Mokyr (1976) rejected this view and instead argued that proto-industrialization provided cheap "surplus" labor that fueled industrialization. Coleman (1983) strongly criticized the literature on proto-industrialization, arguing that its hypotheses do not fit the facts.

Table D.7: Additional Controls: One-by-one

Dependent variable: log city growth, 1750-1850				
	(1)	(2)	(3)	(4)
lnSubDens	0.176***	0.178***	0.187***	0.176***
	(0.038)	(0.039)	(0.040)	(0.041)
Literacy 1786	-0.276**	-0.179	-0.190	-0.208
	(0.133)	(0.149)	(0.143)	(0.144)
lnPreIndDens	1.107***			
	(0.363)			
lnDistanceCoal		-0.020		
		(0.039)		
Pays d'Eléction			-0.076	
			(0.065)	
lnNoblesDens				0.085
				(0.135)
$\ln(pop_{initial})$	-0.067	-0.080*	-0.086*	-0.061
	(0.040)	(0.045)	(0.045)	(0.055)
Controls	✓	✓	✓	✓
R^2	0.40	0.38	0.39	0.38
Observations	166	166	164	166

Notes: All regressions are run at the city level, include a dummy for Paris, and are weighted by city population in 1750. The dependent variable is log city population growth in 1750-1850. "Controls" include the baseline controls and early knowledge controls listed in Table 1. For details on the explanatory variables see notes to Table 1. Standard errors (clustered at the department level) in parentheses . * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Literacy and city growth before 1750

We now shed more light on the relationship between literacy and city growth. First, we analyze the pre-industrial period 1700-1750 in Table D.8. For these regressions, we use the earlier literacy rates in 1686 (which are highly correlated with those in 1786, with a correlation coefficient of 0.84). Coefficients are clustered at the department level – the geographical unit of observation for literacy. Table D.8 shows that literacy is weakly positively (but only in column 3 marginally significantly) associated with city growth in 1700-50. While the relationship is too weak to lend itself to interpretation, one explanation is the larger importance of medium-level worker skills in traditional artisan manufacturing (c.f. de Pleijt and Weisdorf, 2014). Subscriber density is not associated with growth prior to 1750, confirming our previous results.

Table D.8: City growth and literacy in the pre-industrial period, 1700-1750

Dependent variable: log city growth, 1700-1750				
	(1)	(2)	(3)	(4) [unweighted]
lnSubDens		0.036	0.004	0.010
		(0.038)	(0.039)	(0.031)
Literacy 1686	0.274	0.265	0.327*	0.202
	(0.184)	(0.183)	(0.190)	(0.211)
Controls			✓	✓
R^2	0.05	0.06	0.22	0.13
Observations	126	126	126	126

Notes: All regressions are run at the city level, include a dummy for Paris, and are weighted (expect for column 4) by city population in 1700. The dependent variable is log city population growth in 1700-1750. "Controls" include the baseline controls and early knowledge controls listed in Table 1 (with the exception that here we control for population in 1700 instead of 1750). For details on $lnSubDens$, Literacy and controls see notes to Table 1. Standard errors (clustered at the department level) in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Check for complementarity between literacy and upper-tail knowledge

In Table D.9 we analyze whether there may have been a complementarity between literacy and upper-tail knowledge during the period of French industrialization. In column 1, we interact literacy rates with subscriber density and obtain a small, negative and insignificant coefficient. In columns 2 and 3 we split the sample into departments with below- and above-median literacy, respectively. We find almost identical coefficients on subscriber density in both subsamples, indicating that the relationship between upper-tail knowledge and growth did not depend on literacy. This is in line with our discussion in Section 2.4: literacy approximates medium-level worker skills that were not a limiting factor in the adoption of industrial technology. In contrast, spatial variation in the type of workers who *were* crucial – the small number of high-quality craftsmen at the very top of the worker skill distribution (engineers, instrument makers, and mechanics) – is unlikely to be reflected by literacy rates.[22] Consequently, we do not expect a complementarity between literacy and upper-tail knowledge.

Table D.9: Complementarity between literacy and upper-tail knowledge?

Dependent variable: log city growth, 1750-1850

	(1)	(2)	(3)
Sample:	All	Below Median Lit.	Above Median Lit.
lnSubDens	0.197***	0.186***	0.181***
	(0.060)	(0.044)	(0.067)
Literacy 1786	-0.173		
	(0.209)		
lnSubDens × Literacy	-0.040		
	(0.127)		
Controls	✓	✓	✓
R^2	0.38	0.40	0.38
Observations	166	87	79

Notes: All regressions are run at the city level, include a dummy for Paris, and are weighted by city population in 1750. The dependent variable is log city population growth over the period 1750-1850. “Controls” include the baseline controls and early knowledge controls listed in Table 1. For details on $lnSubDens$, Literacy and controls see notes to Table 1. Robust standard errors (clustered at the department level in column 1) in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

[22] Note also the discussion in Section 2.4 that France initially lacked these high-quality craftsmen, so that progressive entrepreneurs hired them from Britain. Thus, our proxy for the presence of enlightened elites may itself reflect some of the spatial variation in high-quality craftsmen.

D.2 Additional Results for Soldier Height

We now show further robustness checks on our income and industrialization regressions. Table D.10 reports results for average conscript height in the pre-1750 period. We perform a similar analysis as in Table 6, but we now separately regress conscript height on $lnSubDens$ (column 1, 4) and Literacy (columns 2, 5). Then, in columns 7 and 8, we weight regressions by the number of soldiers observed in each department. All our results still hold: soldier height prior to 1750 is positively associated with literacy, but not with $lnSubDens$.

Table D.10: Soldier height before 1750

Dependent variable: Soldier height in cm (controlling for age and birth decade)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
							Weighted	
Literacy 1686	1.170**		1.050*	1.135**		0.994*	1.007**	0.989**
	(0.540)		(0.527)	(0.558)		(0.550)	(0.433)	(0.441)
lnSubDens		-0.045	0.113		0.013	0.117	0.049	0.062
		(0.129)	(0.116)		(0.118)	(0.117)	(0.112)	(0.111)
Baseline Controls				✓	✓	✓		✓
R^2	0.05	0.00	0.06	0.14	0.15	0.16	0.08	0.13
Observations	77	87	75	77	87	75	75	75

Notes: All regressions are run at the department level and include a dummy for Paris (Department Seine). The dependent variable is average soldier height recorded over the period 1716-49 and collected by Komlos (2005). To account for variation in height and soldier age within this period, we control for age, age squared, and birth decade (see Appendix C.1). In 7-8 we weight regressions by the number of soldiers per department. "Baseline Controls" are those listed in Table 1. For details on Literacy, $lnSubDens$ and controls see the notes to Table 1. Original city-level variables are aggregated to the department level as described in Appendix C.5. Robust standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

D.3 Local Persistence and Alternative Proxies for Upper-Tail Knowledge: Sources and Detail

Scientific societies

Data on scientific societies are from McClellan (1985). These include the year of foundation and the number of ordinary members. Altogether, there were 30 cities in France hosting scientific societies – all of them were founded prior to 1784, and 22 of them prior to 1750. In the results presented in the paper and below, we use only those founded prior to 1750. Using all 30 scientific societies (i.e., including those that were officially registered after 1750, and before 1784) gives very similar results.

In Table D.11 we examine the relationship between early scientific societies (founded prior to

1750) and *Encyclopédie* subscriptions. Column 1 shows that 91% of cities with an early scientific society also had subscribers to the *Encyclopédie*, as compared to 26% of all other cities reported in Bairoch et al. (1988). Column 2 confirms this pattern: cities with scientific societies had on average 4.54 subscribers per 1,000 inhabitants, while cities without scientific societies had 1.23.

Table D.11: Subscriptions and Scientific Societies

	(1)	(2)
Scientific Society	Subs. Dummy	Subs pc
Yes	0.91	4.54
No	0.26	1.23

Notes: This table analyzes the relationship between early scientific societies and *Encyclopédie* subscriptions. Altogether, there were 22 French cities with scientific societies founded before 1750. Column 1 shows the percentage of cities with subscribers to the *Encyclopédie*, and column 2 shows subscriptions per 1,000 inhabitants, for cities with and without scientific societies. For details on scientific societies and *Encyclopédie* subscriptions see Section 3.1 and Appendix D.3.

Descriptions des Arts et Métiers

The *Descriptions des Arts et Métiers* was a multi-volume publication entirely devoted to the "useful arts." Its origin can be traced back to Jean-Baptiste Colbert, who in 1675 requested the French Academy of Sciences to write a detailed description of the mechanical arts. The idea was to connect artisans with scientists so that they could mutually benefit from this interaction. However, only when René-Antoine Ferchault de Réaumur became responsible for the project in 1709, significant progress was made, and the first volume was published in 1761 (Carpenter, 2011). Our data are from the Neuchâtel (STN) edition of the *Descriptions des Arts et Métiers* sold from 1771 to 1783.[23] The text covered a wide range of "useful arts" (from objects of daily use such as candles or soap to mathematical and astronomical instruments), and a large part was devoted to industrial activities (such as iron production, textiles, and various metal products). Similarly to the *Encyclopédie*, about 1,800 plates accompanied the text, illustrating the mechanical arts.[24]

There are 40 French cities for which the STN reports sales of the *Descriptions des Arts et Métiers*; 80% of these cities also had *Encyclopédie* subscribers, and subscriber density was almost four times higher than in the remaining cities (Table D.12).

A similarly strong relationship holds between per capita sales of the *Descriptions des Arts et*

[23] Data available at http://chop.leeds.ac.uk/stn/, accessed on 1 September 2014.

[24] The *Descriptions des Arts et Métiers* was also a source of inspiration for the *Encyclopédie*. In some cases, the latter were accused of copying from the former, resulting in claims of plagiarism. For more information on the relationship between the *Descriptions des Arts et Métiers* and the *Encyclopédie* see Watts (1952).

Table D.12: Subscriptions and *Descriptions des Arts et Métiers*

	(1)	(2)
Descr. Arts et Métiers	Subs. Dummy	Subs pc
Yes	0.8	4.23
No	0.24	1.08

Notes: This table analyzes the relationship between sales of the *Descriptions des Arts et Métiers* and *Encyclopédie* subscriptions. Altogether, there were 40 French cities with positive sales of the *Descriptions des Arts et Métiers* reported by the FBTEE (2012) project. Column 1 shows the percentage of cities with subscribers to the *Encyclopédie*, and column 2 shows subscriptions per 1,000 inhabitants, for cities with and without sales of the *Descriptions des Arts et Métiers*. For details on the *Descriptions des Arts et Métiers* and *Encyclopédie* subscriptions see Section 3.1 in the paper and Appendix D.3.

Métiers and subscriber density (columns 1-2 of Table D.13), and this holds even when we restrict the sample to the 40 cities with at least one registered sale of the *Descriptions des Arts et Métiers* (columns 3-4). Finally, columns 5-8 document the robustness of the relationship between sales of the *Descriptions des Arts et Métiers* and city growth when including our baseline controls and using different specifications: column 5 does not control for literacy and thus extends the sample; column 6 repeats our main regression from Table 8 (but does not include early knowledge controls); and columns 7 and 8 confirm the robustness of the city growth result to using sales per capita without logs.

Table D.13: *Descriptions des Arts et Métiers*, Subscriber Density, and City Growth

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var.:	*lnSubDens*				Log City Growth, 1750-1850			
	full sample		*ArtMétiers>0*				[no log]	[no log]
ArtsMét. Dens	2.071***	2.653***	1.824**	2.428***	0.449***	0.589***	0.280***	0.343***
	(0.694)	(0.770)	(0.745)	(0.779)	(0.145)	(0.157)	(0.081)	(0.093)
Literacy 1786		✓		✓		✓		✓
Baseline Controls	✓	✓	✓	✓	✓	✓	✓	✓
R^2	0.34	0.38	0.49	0.63	0.27	0.30	0.27	0.30
Observations	193	166	40	37	193	166	193	166

Notes: All regressions are run at the city level, include a dummy for Paris, and are weighted by city population in 1750. "Baseline Controls" are those listed in Table 1. For further detail see the notes to Table 1. Robust standard errors (clustered at the department level in columns 2, 4, 6, and 8) in parentheses. * p<0.1, ** p<0.05, *** p<0.01.
† "ArtMét. Dens" represents two measures of city-level sales of the *Descriptions des Arts et Métiers*: *ArtMétDens* in columns 7-8, and *lnArtMétDens* in all other columns. These are computed analogous to $lnSubDens$ and $SubDens$, respectively, as described in Section 3.1.

Famous scientists

Data on "famous" people born in 1000-1887 are from the *Index Bio-Bibliographicus Notorum Hominum* (IBN), as coded by de la Croix and Licandro (2012). For each person, this includes the city of birth, year of birth, profession, and city of death. There are 2,513 "famous" people that worked in scientific professions (science, mathematics, chemistry, physics, or medicine) whom we identify in our sample by city of birth and/or city of birth. 164 cities in our sample were the place of birth and/or the place of death of at least one "famous scientist." For each city, we compute the density of famous scientists as ln(1+famous scientists/pop_{1750}), where we divide by city population in 1750 because this is closest to the mean year of birth of the "famous" individuals (it also guarantees direct comparability with the way in which $lnSubDens$ is calculated).

In Table D.14 we examine the relationship between subscriber density and the presence of famous scientists before 1750 (columns 1-2) and after 1750 (columns 3-4). In both cases the two proxies for knowledge elites are strongly correlated, confirming our results from Table 9 in the paper. In addition, we find a strong positive relationship between famous scientists per capita born in 1750-1887 and those born in 1000-1749, with a coefficient of 0.78 (std error 0.16) when all baseline controls and a dummy for Paris are included. This provides further support for a relatively stable spatial distribution of scientific elites.

Table D.14: *Encyclopédie* subscriptions and famous scientists born before and after 1750

Dep. var.: Density of famous scientists

	(1)	(2)	(3)	(4)
Birth year	1000-1749		1750-1887	
lnSubDens	0.026**	0.024*	0.093***	0.116***
	(0.012)	(0.014)	(0.023)	(0.024)
Controls		✓		✓
R^2	0.03	0.04	0.09	0.20
Observations	193	193	193	193

Notes: All regressions are run at the city level and include a dummy for Paris. The dependent variable is (log) famous scientists per capita. These are people listed in the *Index Bio-Bibliographicus Notorum Hominum* whose profession is related to science, mathematics, chemistry, physics, or medicine. These data are from de la Croix and Licandro (2012). In total, there 614 famous scientists born in the period 1000–1749 and 1899 famous scientists born during 1750-1887. "Controls" include the baseline controls and early knowledge controls listed in Table 1. For details on $lnSubDens$ and controls see the notes to Table 1. Robust standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Mobility of knowledge elites

A potential concern is that differential mobility of knowledge elites could affect our results. For example, Benedict (2005) documents that French elites (such as merchants and legal officials) were more mobile than artisans or unskilled workers. If elites were attracted to cities that experienced more rapid growth after 1750, this could explain why *Encyclopédie* subscriptions there were higher in 1777-79. We implicitly addressed this issue by showing that earlier proxies for knowledge elites (pre-1750 scientific societies and Huguenots in 1670) confirm our main results, which makes reverse causality unlikely.

Here, we shed more light on the mobility of knowledge elites, using our data on "famous scientists." We begin by analyzing the average mobility of the 1,149 famous scientists in our sample whose places of birth and death are *both* known. Among these, 52% were born and died in the same city, and when excluding Paris, this number is as high as 64%. Consequently, there was a substantial degree of local persistence of knowledge elites over their lifetime – places that bred more scientists also tended to keep them throughout their lives. Next, we use the full sample of 2,513 "famous scientists." Figure D.4 plots the number of famous scientists deceased against the number of famous scientists born in each French city of our sample. Once Paris is excluded (right panel), a tight relationship emerges: all observations are close to the 45° line. Since all our regressions include a dummy variable for Paris, selective migration of elites is unlikely to affect our results.

Science Professionals in 1851

These data are from the 1851 *Recensement* published by the Statistique Général de France (1851). They include department-level information on the number of people in professions related to science (medicine and *hommes de lettres et savants*). For each department, we compute the density of "science professionals" as ln(1+science professionals/pop_{1851}), where pop_{1851} is the total department population in 1851.

Innovation exhibits in 1851

Moser (2005) coded data on innovations exhibited at the Crystal Palace World's Fair in London in 1851. The database includes information on both patented and not patented innovations from 30 different industries, together with the city of origin of the exhibitor. We match 1,261 exhibits with our city database (78 cities display a positive number of innovations). For each city we define the density of innovative activities as ln(1+number of exhibits/pop_{1850}).

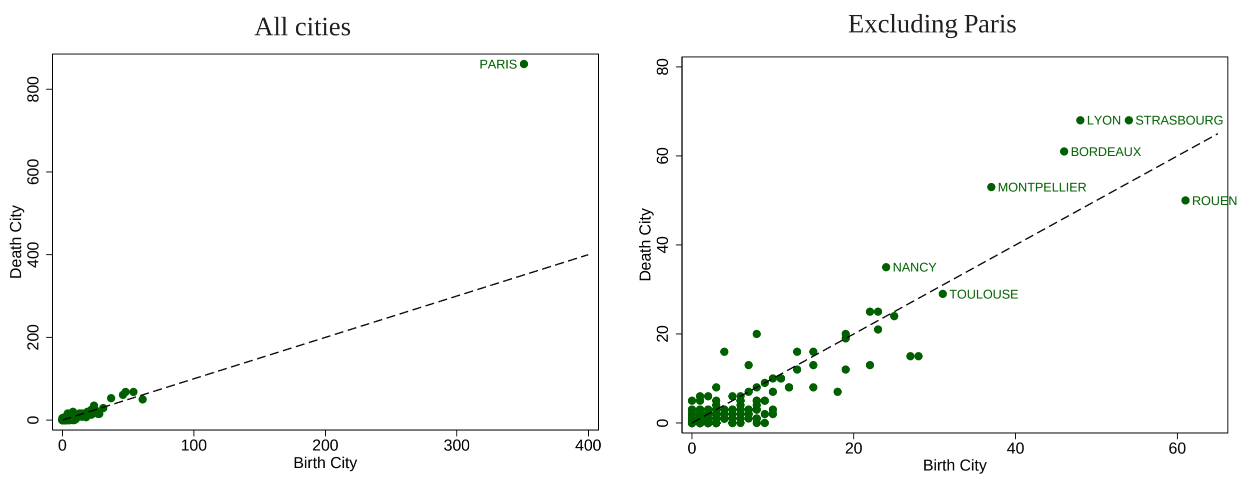

Figure D.4: Mobility of Knowledge Elites

Notes: The figures plot a 45° line together with the number of famous scientists by city of birth (horizontal axis) and the number of famous scientists by city of death (vertical axis). "Famous scientists" are people listed in the *Index Bio-Bibliographicus Notorum Hominum* whose profession is related to science, mathematics, chemistry, physics, or medicine. The data are from de la Croix and Licandro (2012). In total, there are 2,513 famous scientists listed for France over the period 1000–1887. For details on famous scientists see Section 5.1. The left panel includes all cities; the right panel excludes Paris.

D.4 Huguenots and Upper-Tail Knowledge: Sources and Detail

Data on Huguenots are from Mours (1958). We use information on the Huguenot population residing in 78 French department in 1670 and 1815. The highest numbers of Huguenots are observed in the departments of Charente-Maritime and Gard. In Table 10 we define $lnHugDens_{1670}$ in the same fashion as the other department-level "density" variables, i.e., relative to aggregate department-level city population in 1750. While this provides direct comparability to the other explanatory variables, total urban population in 1750 is clearly a rough proxy for overall department population in 1670. Unfortunately, the latter is not directly available, and departments did not yet exist. We thus provide the following alternative approximation: we extrapolate the department level population back to 1700, by using the population growth from total city population in each department between 1700 and 1850, together with department level population in 1831 (when it is first available). Table D.15 reports the results when using this alternative proxy for department population to compute $lnHugDens2_{1670}$. Our results still hold: Huguenot density is a strong predictor of upper-tail knowledge, but not of literacy (columns 1-2) and it is positively associated with city growth after 1750 (column 3), but not before (columns 5 and 6).

Finally, Figure D.5 shows that the department-level Huguenot population in 1815 is closely

Table D.15: City growth and historical determinants of upper-tail knowledge

Huguenots (alternative), subscriber density, and city growth

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.:	$SubDens$	Literacy	Log City Growth			
			1750-1850		1700-1750	
lnHugDens2$_{1670}$	2.264***	-0.098	0.723**	0.363	0.457	0.400
	(0.781)	(0.308)	(0.331)	(0.339)	(0.348)	(0.330)
lnSubDens				0.159***		0.024
				(0.041)		(0.032)
Baseline Controls	✓	✓	✓	✓	✓	✓
R^2	0.23	0.07	0.15	0.21	0.10	0.10
Observations	142	142	142	142	119	119

Notes: All regressions are run at the city level and include a dummy for Paris. "Baseline Controls" are those listed in Table 1 (columns 5-6 control for initial population in 1700 instead of 1750). $lnHugDens2_{1670}$ is calculated using the alternative approximation for 1670 department-level population described in the text above. For details on $lnSubDens$, Literacy and controls see the notes to Table 1. Standard errors (clustered at the department level) in parentheses. * p<0.1, ** p<0.05, *** p<0.01

related to its counterpart in 1670.[25] This implies that emigration (and conversion) after the revocation of the Edict of Nantes in 1685 was not systematically stronger in some regions than in others. Consequently, our use of Huguenot density in 1670 is a good proxy for their (clandestine) presence in the 18th century.

[25]After the French Revolution, Huguenots again gained equal rights as French citizens. The first year for which a head count is available is 1815. According to the data by Mours (1958), the average number of Huguenots per department declined from 11,310 to 5,640 between 1670 and 1815.

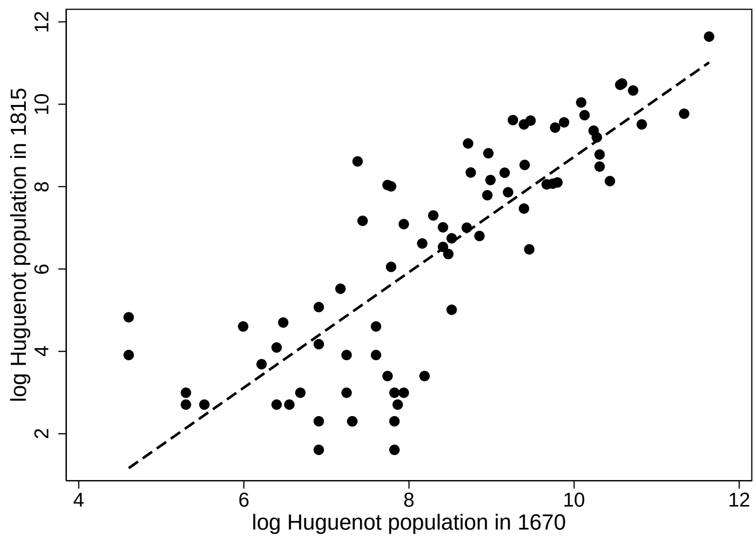

Figure D.5: Huguenot population in 1670 and 1815

Notes: Data from Mours (1958). The figure shows that department-level Huguenot population in 1815 is strongly related to Huguenot presence in 1670.

D.5 Innovation, Knowledge Elites, and Productivity: Sources and Detail

This appendix complements Section 5.3 in the paper, providing detail on sources, methodology, and additional results.

Firm-Level Data and Results

We use French firm level data from Chanut, Heffer, Mairesse, and Postel-Vinay (2000), who cleaned and digitalized a survey of more than 14,000 firms, originally conducted by the *Statistique Générale de la France* over the period 1839-1847. The data are collected at the arrondissement (sub-county) level, and categorize firms into 13 main manufacturing sectors.[26] Merging the 13 French and the 21 British sectors, we obtain 8 consistent sectors.[27] We then compute their innovation index as the weighted average "share of inventive output" from the British patent data.[28] Using this index, we classify French sectors into "modern" and "old," based on above- vs. below-median innovation index. Table D.16 lists the resulting 8 sectors together with their innovation index. The dataset has about 630 (800) sector-arrondissement observations for "modern" ("old") sectors.

In our analysis, we use male wages as dependent variable. These represent the average daily wages for men (recorded in centimes). We compute establishment size as the total number of workers, divided by the number of establishments for each firm. To control for agglomeration, we include the log of total population, as well as the urbanization rate (both measured at the department level in 1831). These data are from Statistique Général de France (1878). In Table 12, we use information on the number of steam engines and other engines (which include water, wind and animal engines), as proxies for an industry's dependence on energy-related up-front investment. These figures are reported in Chanut et al. (2000).

British Inventions and Encyclopédie *Plates*

Data on *Encyclopédie* plates are from *the Encyclopedia of Diderot and d'Alembert: collaborative translation project*[29] and from the *ARTFL Encyclopédie project*.[30] For each volume, these sources list all articles and plates of the *Encyclopédie* of Diderot and d'Alembert. Altogether, there are 2,575 plates, accompanying 326 entries. They contain reference numbers and characters to match

[26]Typically, arrondissements map one-to-one into cities in our sample; only 2 arrondissements include more than one city with observed population in 1750 and above-zero subscriptions. The departments of Corsica, Savoie, Haute Savoie and Territoire de Belfort are not included in the survey.

[27]Some categories overlap, so that a consistent match results in 8 sectors. Whenever there is more than one British sector corresponding to a French sector, we compute a weighted average of the share of innovative output, where the weights are the number of patents in the British sector.

[28]We use the average, rather than the sum of "inventive output" shares because otherwise aggregating many sectors would mechanically raise their innovation index.

[29]Available at http://quod.lib.umich.edu/d/did/index.html

[30]Available at http://encyclopedie.uchicago.edu/

Table D.16: Classifaction of individual industrial sectors into “modern” and “old”

	(1)	(2)
Sector Name	Innovation Index	Sector Type
Textile and Clothing	0.145	modern
Printing Technology, and Scientific Instruments	0.094	modern
Furniture and Lighting	0.045	modern
Transportation Equipment	0.040	modern
Metal and Metal Products	0.039	old
Leather	0.018	old
Mining	0.017	old
Ceramics and Glass	0.012	old

Notes: For each sector, column 1 reports the innovation index, obtained using data from Nuvolari and Tartari (2011); column 2 classifies sectors into “modern” or “old” manufacturing, based on the median of the share of the innovation index. For details on the innovation index and on the French industrial survey see Section 5.3 and Appendix D.5.

the various parts of the figure with the text and the legend. For all 326 entries, we computed the exact number of accompanying plates, and – where possible – we matched each entry with one of the 21 British Industrial sector from Nuvolari and Tartari (2011). In total, 156 entries accompanied by 1,272 plates describe manufacturing technologies. Among them, 103 entries (and 849 plates) are dedicated to the "modern" sector. The number of plates per entry varies substantially. For instance, among all plates describing manufacturing, 30 entries have only one plate, while the entries "Clock Making" and "Turner and Turning Lathe" have 51 and 87 plates respectively.

Table D.17 shows that the sectors that we classify as "modern" had a share of inventive output of 0.084 – more than five times higher than the figure for "old" sectors.[31] Column 2 shows that more than two thirds of all plates dedicated to manufacturing in the *Encyclopédie* described "modern" technologies.

Table D.17: Modern vs. old sectors: Share of inventive output and *Encyclopédie* plates

	(1) Share Invent. Output	(2) Share Plates *Encyclopédie*
Modern	0.084	0.67
Old	0.016	0.33

Notes: This table distinguishes between "modern" and "old" manufacturing sectors, based on the median share of total "inventive output" from Nuvolari and Tartari (2011) (see Section 5.3 for detail). Column 1 shows the average share of total "inventive output" for both manufacturing technologies. Column 2 compares the share of plates describing "modern" and "old" manufacturing technologies in all *Encyclopédie* plates dedicated to manufacturing (see Appendix D.5 for sources and detail).

Finally, Table D.18 shows that our results in Table 11 in the paper are very similar when we use the alternative measure of subscriber density $SubDens$ (without taking logs).

[31] Of course, this difference results from classifying these sectors according to above- and below-median share of inventive output, as described in Appendix D.5.

Table D.18: Alternative specification of subscriber density in Table 11 – $SubDens$ no logs

Dep. Var.: log wages (by sector and arrondissement)						
	(1)	(2)	(3)	(4)	(5)	(6)
SubDens	0.010**	0.010**	0.009***	0.007*	0.004	
	(0.004)	(0.004)	(0.003)	(0.004)	(0.006)	
$SubDens \times Modern$	0.019***	0.015***	0.012***	0.015***	0.016***	0.014***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
School Rate 1837	0.248***	0.233***	0.239***	0.167**		
	(0.070)	(0.072)	(0.071)	(0.069)		
$School \times Modern$	-0.027	-0.035	-0.032	-0.007	0.016	0.051
	(0.068)	(0.066)	(0.069)	(0.091)	(0.100)	(0.108)
Establishment Size	0.055***	0.045***	0.042***	0.046***	0.041***	0.039***
	(0.008)	(0.008)	(0.009)	(0.010)	(0.010)	(0.010)
$Size \times Modern$	-0.068***	-0.031***	-0.029**	-0.030*	-0.034**	-0.034**
	(0.011)	(0.011)	(0.014)	(0.016)	(0.015)	(0.016)
Modern Sector	0.133***					
	(0.036)					
Sector FE		✓	✓	✓	✓	✓
Baseline Controls			✓	✓	✓	✓
Additional Controls				✓	✓	✓
Department FE					✓	(✓)
Arrondissement FE						✓
R^2	0.13	0.22	0.35	0.37	0.49	0.58
Observations	1482	1482	968	844	844	844

Notes: All regressions are run at the arrondissement level and include a dummy for Paris (Department Seine). The dependent variable is the log of average male wages across all firms in a sector j in arrondissement n. There are more than 14,000 firms in the sample (see Appendix D.5). Firms a classified into 8 sectors, and the 4 most innovative ones are categorized as "modern" (see Appendix Section D.5 and Table D.16 for detail). *Establishment size* is the (log) average number of workers across all firms in j and n. "Baseline Controls" and "Additional Controls" are those listed in Table 1; we also control for (log) total department-level population and urbanization rates (both in 1831) to capture agglomeration effects. For each control variable, both its level and its interaction with "modern" is included. $SubDens$ is the number of *Encyclopédie* subscribers per 1,000 city inhabitants in 1750. For details on controls see the notes to Table 1. Original city-level variables are aggregated to the arrondissement level as described in Appendix C.5. Standard errors (clustered at the department level) in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

E Further Historical Background

E.1 French Institutions Before and After the Revolution

As we argued in the introduction, the fact that we use variation within France avoids many of the problems common to cross-country studies. In particular, since France was a highly centralized (and, until 1789, absolutist) state, regional variation in institutions or the rule of law was limited. In the paper, we used *pays d'élection* as a rough proxy for the king's reach in taxation. In the following, we provide a more detailed historical discussion of the judicial system, focusing on the centralization of the rule of law. Braudel (1982) argues that

> *[B]y the thirteenth century, France was already the major political unit of the continent...having all the requisite ancient and modern characteristics of a state: the charismatic aura, the judicial, administrative and above all financial institutions, without which the political unit would have been completely inert.* (Braudel, 1982, p.323)

The king with his *conseil* was at the top of the royal judicial system. At the next level, there were about 15 regional and provincial parliaments and councils. Below these followed approximately 400 local royal courts – called *bailliages* in northern France and *sénéchaussées* in southern France. These were responsible for appeals for non-nobles and ecclesiastics, and they were also the first instance for disputes concerning nobles. Finally, the bottom of the pyramid of *royal* jurisdictions was represented by about a thousand lower courts (Hamscher, 2012, p.11), whose names differed across regions (e.g., prévôtés, vicomtés, or châtellenies).

The judges of the royal justice system were generally fidéles to the king (Mousnier, 1979, p.354), received the same training, and were closely monitored. The same is true for those of the lower local courts – the *seigneurial* jurisdictions.[32] Seigneurial courts were ultimately subject to the king's justice, and their sentences could be appealed in royal tribunals (Hamscher, 2012). Graham (2011) observes that

> *[J]udges, lawyers and court officials received the same sorts of training during the eighteenth century, whether they ended up working in royal jurisdictions or seigneurial ones. Officials in the royal courts were reasonably effective in monitoring their seigneurial counterparts precisely because of their shared expertise and a concern to defend their mutual interests.* (Graham, 2011, p.16)

Similarly, Muessig (2012, p.212) concludes that *"by the 14th century at the latest, royal jurisdiction dominated, and the feudal justice seigneuriale became subordinate to royal jurisdiction."* Johnson and Koyama (2014) describe how centralization was reinforced after civil disturbances in

[32] The *seigneurial* courts were appointed by lords, or seigneurs, and exercised a civil and criminal jurisdiction within the limit of the seigneurie.

the early 1650s ("Fonde") had been put down. This was accompanied by both a vast increase in the number of government officials and efforts to align their incentives with the central government, rather than the provinces. All this points to the conclusion that the institutional setting was – in comparison to other countries – relatively homogenous across the different French regions. In Graham's (2011, p.16) words: *"by the eighteenth century, there was in essence just one system of justice in France: that of the king."*

During the French Revolution, the judicial system was largely reformed, and almost all *Ancien Régime* courts were abolished. The revolutionary program initially planned to institute a new system relying mostly on informal mediation, rather than on formal law. However, after the Reign of Terror, this distaste for legal formality faded away, and lawyers and formal courts were re-instituted. A crucial change occurred during the Napoleonic period when key reforms were implemented, establishing the institutions that still characterize today's French legal system.[33] At the bottom of the private judicial system were the *tribunaux de première instance*, which had general jurisdiction over all civil and criminal matters. The next level was represented by the *cours d'appel*, which handled appeals from the *tribunaux de première instance*. Finally, at the top of the judicial system was the *cour de cassation* (Kessler, 2010).

Most important for our study, if France was centralized before the Revolution, it became even more so thereafter. For example, Tilly states that *"revolutionaries installed one of the first systems of direct rule ever to take shape in a large state"* (Tilly, 1990, pp. 107-10). On the other hand, among the early observers, Alexis de Tocqueville emphasized that the French government was already "highly centralized" before 1789, and that in this dimension, the French Revolution resulted in "far fewer changes" than is usually supposed (de Tocqueville, 1856, pp.ix, 20, 32-41).

E.2 Agricultural Productivity and Scientific Knowledge

Increasing agricultural productivity was crucial for industrialization because it enabled rapid population growth and increasing urbanization, due to the increasing availability of food for nonfarm people. However, in this initial phase, the increase in food production occurred without the application of modern science. As Johnson (1997, p. 2) points out: "The application of scientific knowledge, both basic and applied, to agriculture is a recent event, dating from the middle of the 19th century." The beginning of agricultural research is typically dated from the time of the work of Justus von Liebig in agricultural chemistry in the 1840s.[34] By that time, laboratory experiments on agricultural research were founded around Europe (for instance in England Lawes established

[33] An important difference with the *Ancien Régime* was the introduction of a distinct public judicial system, where private individuals could challenge state actions.

[34] His pioneering work *Organic Chemistry in its Relation to Agriculture and Physiology* "launched the systematic development of the agricultural sciences" (Scheewe, 2000, p.17).

an experiment station and started the production of the superphosphate fertilizer (Huffman and Evenson, 2008). Innovations in the chemical industry led to large increases in yield per hectare through the understanding of the critical role of nitrogen in food production and the introduction of commercial fertilizer (Huffman and Evenson, 2008; Smil, 1997).

The pre-industrial period also witnessed significant increases in agricultural productivity. However, these did not require scientific knowledge. Instead, they were "the result of the activities of private individuals who had little formal research training" (Huffman and Evenson, 2008, p. 17). Johnson (1997, p. 7) reinforces this point: "[a] significant increase in food production in Western Europe started well before the application of modern science. Prior to two centuries ago, food production increased primarily as a result of more intensive use of land, rather than from increased yields per unit of sown area of the principal grains." Similarly, Grantham (1989, p. 44) suggests that until the mid-19th century, the growth of agricultural production depended more on "intensive use of known technology than on novel methods."

Increasing commercialization during industrialization triggered changes in agricultural productivity through different channels. For example, European farmers reacted to market opportunities by increasing labor input and investment, and by choosing more marketable crops. However, Clark (1987) shows that before 1850, work rates (rather than technological improvements) affected differences in agricultural output across European regions. Using data for Britain, Clark (1987) documents that the increase of agricultural productivity from 1661 to 1841 was mainly due to labor inputs, and that only 15% can be attributed to technical progress. Similarly, analyzing detailed data from France, Grantham (1989) suggests that technical innovation contributed to the improvement of agricultural production only after the 1840s, with the introduction of commercial fertilizers and mechanical harvesting.

From the 1830s and 1840s farmers started to buy inputs directly from the manufacturing industry. This dramatically increased access to inputs compared to the pre-industrial period when they would obtain seeds from their own harvest, manure from their livestock, and grass to feed animals from their own farm. In addition, the availability of cheap iron fostered the use of better equipment such as iron ploughs, drills, and reapers in agricultural production. However, none of these productivity-enhancing methods introduced before the 1840s required advanced technological or even scientific knowledge.

References

Abramson, S. F. and C. Boix (2013). The Roots of the Industrial Revolution: Political Institutions or (Socially Embedded) Know-How? APSA 2013 Annual Meeting Paper.

Acemoglu, D., P. Aghion, and F. Zilibotti (2006). Distance to Frontier, Selection, and Economic Growth. *Journal of the European Economic Association 4*(1), 37–74.

Antràs, P. and H.-J. Voth (2003). Factor Prices and Productivity Growth during the British Industrial Revolution. *Explorations in Economic History 38*, 52–77.

Aron, J. P., P. Dumont, and E. Le Roy Ladurie (1972). *Anthropologie du Conscrit Français.* Paris: Mouton De Gruyter.

Bairoch, P., J. Batou, and P. Chèvre (1988). *La Population des villes Europeenees de 800 à 1850: Banque de Données et Analyse Sommaire des Résultats*. Geneva: Centre d'histoire economique Internationale de l'Université de Genève, Libraire Droz.

Barraclough, G. (1978). *The Times Atlas of World History*. London Times Books.

Benedict, P. (2005). *Cities and Social Change in Early Modern France*. London and New York: Routledge.

Bils, M. and P. J. Klenow (2000). Does Schooling Cause Growth? *American Economic Review 90*(5), 1160–1183.

Braudel, F. (1982). *Civilization and Capitalism, 15th–18th Century: The perspective of the world*. University of California Press.

Brette, A. (1904). *Atlas des bailliages ou jurisicatins asimilées ayant formé unité électorale en 1789 dressé d'après les actes de la convocation conservés au Archives nationales*. Paris: Imprim. nationale.

Cantoni, D. and N. Yuchtman (2014). Medieval Universities, Legal Institutions, and the Commercial Revolution. *Quarterly Journal of Economics*. forthcoming.

Carpenter, K. (2011). *Physiocrats, Antiphysiocracy and Pfeiffer*, Chapter Manufactures in European Economic Literature of the Enlightenment: The Description des Arts et Métiers and the Schauplatz der Künste und Handwerke, pp. 5–21. New York, Springer, 2011.

Carus-Wilson, E. M. (1966). *The Woollen Industry. Cambridge Economic History of Europe. Volume II*. Cambridge University Press.

Chanut, J.-M., J. Heffer, J. Mairesse, and G. Postel-Vinay (2000). *L'Industrie franccaise au milieu du 19e siécle: les enquêtes de la statistique générale de la France*. Editions de l'école des hautes études en sciences sociales, Paris.

Clair, C. (1976). *A History of European Printing*. London: Academic.

Clark, G. (1987). Productivity Growth without Technical Change in European Agriculture before 1850. *Journal of Economic History 47*(2), 419–432.

Coleman, D. C. (1983). Proto-Industrialization: A Concept Too Many. *Economic History Review 36*(3), 435–448.

Crafts, N. F. R. (1985). *British Economic Growth During the Industrial Revolution*. Oxford: Oxford University Press.

Crafts, N. F. R. and K. Harley (1992). Output Growth and the British Industrial Revolution: A restatement of the Crafts-Harley View. *Economic History Review 45*, 703–730.

Darby, H. C. and H. Fullard (1970). *The New Cambridge Modern History*, Volume 14. Cambridge: Cambridge University Press.

Darnton, R. (1973). The Encyclopedie Wars of Prerevolutionary France. *The American Historical Review 78*(5), 1331–1352.

de la Croix, D. and O. Licandro (2012). The Longevity of Famous People from Hammurabi to Einstein. CORE Discussion Paper 2012/52.

de Pleijt, A. M. and J. L. Weisdorf (2014). Human Capital Formation from Occupations: The 'Deskilling Hypothesis' Revisited. CGEH Working Paper 57.

de Tocqueville, A. (1856). *The Old Regime and the French Revolution*. Garden City, N.Y.: Doubleday. Reprinted in 1955.

Delefortrie, N. and J. Morice (1959). *Les Revenus departmentaux en 1864 et en 1954*. Paris, France: Colin.

Dittmar, J. E. (2011). Information Technology and Economic Change: The Impact of the Printing Press. *Quarterly Journal of Economics 126*, 1133–1172.

Dittmar, J. E. (2013). New Media, Firms, Ideas, and Growth: European Cities After Gutenberg. Working Paper.

FBTEE (2012). *The French Book Trade in Enlightenemnt Europe Database*.

Febvre, L. and H.-J. Martin (1958). *L'Apparition du Livre*. Paris: Albin Michel.

Furet, F. and J. Ozouf (1977). *Lire et écrire. L'alphabétisation des Français de Calvin à Judes Ferry*. Minuit, Paris.

Galor, O. (2005). From Stagnation to Growth: Unified Growth Theory. In P. Aghion and S. Durlauf (Eds.), *Handbook of Economic Growth*, Volume 1, Chapter 4, pp. 171–293. Elsevier.

Galor, O. (2011). *Unified Growth Theory*. Princeton University Press.

Goreaux, L. M. (1956). Les migrations agricoles en France depuis un siécle et leur relation avec certains facteurs économiques. *Études et Conjoncture 11*, 327–376.

Graham, H. (2011). Humble Petitioners? Seigneurial Courts, Royal Justice and the Role of Litigants in the Eighteenth Century. *French History & Civilization 4*, 12–21.

Grantham, G. (1989). Agricultural Supply During the Industrial Revolution: French Evidence and European Implications. *Journal of Economic History 49*(1), 43–72.

Hamscher, A. N. (2012). *The Royal Financial Administration and the Prosecution of Crime in France, 1670–1789*. Newark: University of Delaware Press.

Hansen, G. and E. Prescott (2002). Malthus to Solow. *American Economic Review 92*(4), 1205–1217.

Huffman, W. and R. Evenson (2008). *Science for Agriculture: A Long-Term Perspective*. John Wiley & Sons.

ISTC (2008). *Incunabula Short-Title Catalogue*. London: British Library.

Jedin, H., K. S. Latourette, and J. Martin (1970). *Atlas Zur Kirchengeschichte*. Freiburg: Herder.

Johnson, D. (1997). Agriculture and the Wealth of Nations. *American Economic Review, Papers & Proceedings 87*(2), 1–12.

Johnson, N. D. and M. Koyama (2014, February). Taxes, Lawyers, and the Decline of Witch Trials in

France. *Journal of Law and Economics 57*, 77–112.

Kelly, M., J. Mokyr, and C. Ó Gráda (2014). Precocious Albion: A New Interpretation of the British Industrial Revolution. *Annual Reviews of Economics 6*. forthcoming.

Kessler, A. D. (2010). Marginalization and Myth: The Corporatist Roots of France's Forgotten Elective Judiciary. *American Journal of Comparative Law 58*, 679–720.

Komlos, J. (2005). *Height of French Soldiers, 1716–1784*. [Computer file] ICPSR04363-v1.

Lough, J. (1968). *Essays on the Encyclopédie of Diderot and dŠAlembert*. London: Oxford University Press.

Maddison, A. (2007). *Historical Statistics*. University of Groningen. Webpage: http://www.ggdc.net/maddison/.

Marczewski, J. (1961). Some Aspects of the Economic Growth of France 1660-1958. *Economic Development and Cultural Change 9*(3), 369–386.

McClellan, J. E. (1985). *Science Reorganized: Scientific Societies in the Eighteenth Century*. New York: Columbia University Press.

McEvedy, C. and R. Jones (1978). *Atlas of World Population History, Facts on File*. New York.

Mendels, F. F. (1972). Proto-industrialization: The First Phase of the Industrialization Process. *Economic History Review 32*(1), 241–261.

Mokyr, J. (1976). *Industrialization in the Low Countries, 1795–1850*. Yale University Press.

Mokyr, J. (2000). Knowledge, Technology, and Economic Growth During the Industrial Revolution. In B. V. Ark and G. Kuper (Eds.), *Technology and Productivity Growth*. Hague: Kluwert. Available at http://faculty.wcas.northwestern.edu/~jmokyr/Groningen.pdf.

Mokyr, J. (2005). Long-Term Economic Growth and the History of Technology. Volume 1 of *Handbook of Economic Growth*, pp. 1113–80. Elsevier, Amsterdam.

Mokyr, J. (2010). *The Enlightened Economy: An Economic History of Britain 1700-1850*. Yale University Press.

Moser, P. (2005). How Do Patent Laws Influence Innovation? Evidence from Nineteenth-Century World's Fairs. *American Economic Review 95*(4), 1214–1236.

Moser, P. (2012). Innovation without Patents: Evidence from World's Fairs. *Journal of Law and Economics 55*(1), 43–74.

Mours, S. (1958). *Les Églises Réformées en France. Tableaux et Cartes*. Librairie Protestante.

Mousnier, R. (1979). *The Institutions of France Under the Absolute Monarchy, 1598–1789*, Volume 2. University of Chicago Press.

Muessig, U. (2012). *Judges and Judging in the History of the Common Law and Civil Law*, Chapter Superior courts in early-modern France, England and the Holy Roman Empire, pp. 209–233. Cambridge University Press.

Murphy, T. E. (2010). Old Habits Die Hard (Sometimes): What can Département Heterogeneity Tell us about the French Fertility Decline? IGIER Working Paper 364.

Nelson, R. R. and E. S. Phelps (1966). Investment in Humans, Technological Diffusion, and Economic Growth. *American Economic Review 56*(2), 69–75.

Nuvolari, A. and V. Tartari (2011). Bennet Woodcroft and the value of English patents, 1617–1841. *Explorations in Economic History 48*, 97–115.

O'Rourke, K. H., A. S. Rahman, and A. M. Taylor (2013). Luddites, the Industrial Revolution, and the Demographic Transition. *Journal of Economic Growth 18*, 373–409. NBER working paper 14484.

Roehl, R. (1976). French Industrialization: A Reconsideration. *Explorations in Economic History 13*, 223–281.

Scheewe, W. (2000). *Nurturing the Soil-feeding the People*. Rex Bookstore, Inc.

Smil, V. (1997). Global Population and the Nitrogen Cycle. *Scientific American 277*(1), 76–81.

Sprandel, R. (1968). *Das Eisengewerbe im Mittelalter*. A. Hiersemann.

Statistique Général de France (1851). *Recensement 1851*.

Statistique Général de France (1861). *Statistique Industrielle*.

Statistique Général de France (1878). *Annuaire Statistique de la France*. Paris, Imprimerie Nationale.

Tilly, C. (1990). *Coercion, Capital, and European States, 990–1990*. London: Blackwell.

Vandenbussche, J., P. Aghion, and C. Meghir (2006). Growth, Distance to Frontier and Composition of Human Capital. *Journal of Economic Growth 11*(2), 97–127.

Watts, G. (1952). The Encyclopédie and the Descriptions des arts et métiers. *The French Review 25*(6), 444–454.

Overview of the variables used in the paper (1/2)

Variable Name	**Variable Description**	**Source**
	Explanatory Variables	
I_{Subs}	dummy equal to 1 for cities with a positive number of subscriptions	Darnton (1973)
lnSubDens	log of 1+ subscriptions per capita	Darnton (1973)
$I_{Scient.Society}$	dummy equal to 1 for cities hosting a scientific society before 1750	McClellan (1985)
lnMembDens	log of 1+ per capita members of scientific societies before 1750	McClellan (1985)
$I_{ArtsMetiers}$	dummy equal to 1 for cities with sales of the *Descriptions des Arts et Métiers*	McClellan (1985)
lnArtMétDens	log of 1+ sales per capita of the *Descriptions des Arts et Métiers*	FBTEE (2012)
Literacy	percentage of men able to sign their wedding certificate in 1686 or 1786	Furet and Ozouf (1977)
School Rate 1837	ratio of students to school-age population in 1837	Murphy (2010)
	Outcome Variables	
log city growth	log of city population growth over the indicated periods	Bairoch et al. (1988)
Soldier Height pre-1750	average soldier height in cm in 1716-1749 (dept level; cohort and age controls)	Komlos (2005)
Soldier Height 1819-1826	soldier height in cm (dept. level)	Aron, Dumont, and Le Roy Ladurie (1972)
ln (disposable income)	log disposable income in 1864 (dept. level)	Delefortrie and Morice (1959)
Industrial Output	industrial output per capita in 1861 (dept. level)	Statistique Général de France (1861)
Industrial Employment	industrial employment per capita in 1861 (dept. level)	Statistique Général de France (1861)
Wage Industry	wage in industry in 1852 (dept. level)	Goreaux (1956)
Wage Agric	wage in agriculture in 1852 (dept. level)	Goreaux (1956)
Wage modern/old	wage in “modern” and “old” sectors in 1839-1847 (arrond. level)	Chanut et al. (2000)

Overview of the variables used in the paper (2/2)

Variable Name	**Variable Description**	**Source**
	Baseline Controls	
$lnPop_{initial}$	log city population at the beginning of the respective period	Bairoch et al. (1988)
Atlantic Port	dummy equal to 1 for cities located on an Atlantic port	Dittmar (2011)
Mediterranean Port	dummy equal to 1 for cities located on a Mediterranean port	Dittmar (2011)
Navigable River	dummy equal to 1 for cities located on a navigable river	Dittmar (2011)
University	dummy equal to 1 for cities hosting a university before 1750	Jedin et al. (1970); Darby and Fullard (1970)
Paris	dummy equal to 1 for Paris or for the Seine department	
Non French Speaking	dummy equal to 1 for cities located in non-French speaking departments	http://www.lexilogos.com/france_carte_dialectes.htm
Printing press in 1500	dummy equal to 1 for cities who had a printing press in 1500	Febvre and Martin (1958); Clair (1976)
ln(Books Printed 1500)	log of editions printed before 1500	ISTC (2008)
Establishment size	log of number of workers per establishment in 1839-1847 (arrond. level)	Chanut et al. (2000)
Log population in 1831	log total population in 1831 (dept. level)	Statistique Général de France (1878)
Urbanization rate in 1831	urban population, divided by total population in 1831 (dept. level)	Statistique Général de France (1878)
Population in 1861	total population in 1861 (dept. level)	Statistique Général de France (1851)
	Additional Controls	
lnSTNBooksDens	log of 1+ STN book sales per capita	FBTEE (2012)
Pays d'élection	dummy equal to 1 for cities located in a *pay d'éléction*	http://fr.wikipedia.org/wiki/Généralité_(France)
lnPreIndDens	log of 1+ pre-industrial centers per capita (dept. level)	Carus-Wilson (1966); Sprandel (1968)
lnDistanceCoal	log distance (in km) from the closest coal field	Barraclough (1978)
lnNoblesDens	lot of 1+ noble families per capita (dept. level)	http://en.wikipedia.org/wiki/List_of_French_marquisates#cite_note-1
	Other Variables	
lnHuguenotsDens	log of 1+ huguenots per capita in 1670 (dept. level)	Mours (1958)
lnScientistsDens	log of 1 + famous scientists per capita in 1000-1887	de la Croix and Licandro (2012)
lnInnovationsDens	log of 1+ innovations (exhibited at the Crystal Palace Exhibition in London in 1851) per capita	Moser (2005)
lnScienceProf1851	log of 1 + people in scientific professions per capita in 1851 (dept. level)	Statistique Général de France (1851)
Share Inventive Output	measure based on reference-weighted patents, adjusted for the sector-specific frequency of patenting rates and citations	Nuvolari and Tartari (2011)
Share Plates *Encyclopédie*	share of *Encyclopédie* plates related to modern sectors, divided by total plates dedicated to manufacturing	http://quod.lib.umich.edu/d/did/index.html http://encyclopedie.uchicago.edu/