The French coal curse
Coal extraction and productivity in the long run*

Jean-Baptiste Chaniot

CREST and Collège de France

March 2025

Abstract

This paper argues that there is a negative and sustained impact of exposure to natural resources on productivity, education and industry specialization in the long-run. I first develop a highly stylized model of innovation and industry specialization where the availability of natural resources influences sectoral composition. This model predicts that an exogenous increase in the availability of natural resources is likely to induce over-specialization in the resource-intensive industry which leads to an inefficiently low level of education, and, in turn, to an inefficiently low level of productivity. I then leverage a French dataset spanning the whole XXth century to document a negative and sustained effect of exposure to coalfields on both patenting quantity and patenting quality. Meanwhile, coal extraction also exerted adverse impacts on upper-tail education as well as on industry specialization. I conclude the paper by outlining the main mechanisms which explain the persistence of this negative impact over time: I notably argue that specialization in the coal industry triggered negative feedback loops featuring technological and cultural lock-in effects.

Keywords: Patents, Coal, Education, Science **JEL Codes:** O33, J24, O15, J11, N34, N14

*I thank Philippe Aghion, Pierre Boyer, Roland Rathelot, Florian Grosset, Benoît Schmutz, Xavier d'Haultfoeuille, and Clément Malgouyres for helpful comments. I acknowledge financial support from the Farhi Innovation Lab. Address: Jean Baptiste Chaniot: CREST and Collège de France (jean-baptiste.chaniot@hec.edu)

1 Introduction

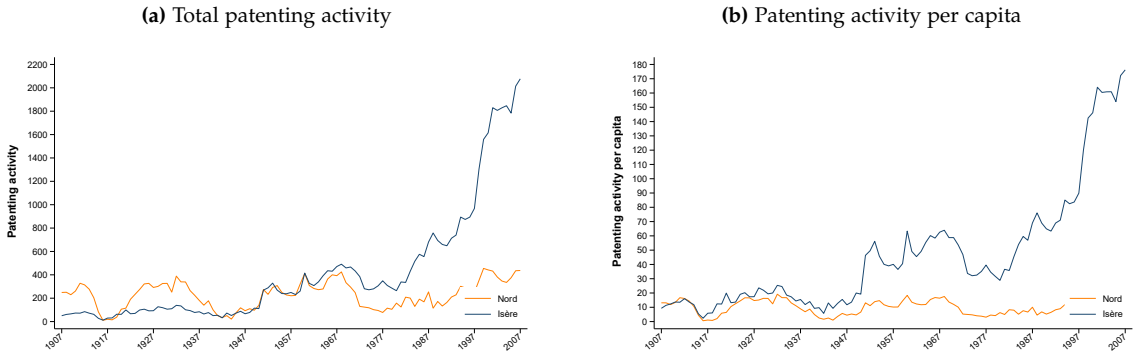
Coal extraction started in the beginning of the XVIIIth century in France as the first concession was granted in 1717 in the city of Anzin in the Nord department. This authorization paved the way for almost three centuries of coal extraction which made of France one of the most important coal exporters in Europe. By the end of the XIXth century, roughly 1,090 coalfields were actively run throughout the French territory and largely contributed to the industrialization process as well as to economic growth. Nord and Pas de Calais, which were the departments with the highest coal production throughout the first half of the XXth century, ranked respectively 3^{rd} and 8^{th} in terms of wealth creation between 1922 and 1939, suggesting that coal contributed substantially to industry capacity, at least in the beginning of the XXth century. This successful story however came to a halt in the middle of the XXth century, as the Jeanneney plan paved the way for the dismantling of the French coal industry. From then on, former coal basins have witnessed severe economic flaws: to mention but a few of them, unemployment in 2023 was much larger in the Nord department (9.5%) and in the Pas de Calais department (8.5%) compared to the national average (7.3%); similarly, a larger batch of the population was eligible to RSA in Nord (8.8%) and in Pas de Calais (8.1%) compared to the national average (5.8%).

In that respect, coal extraction pertains to ambiguous impacts on economic outcomes at the local level. On the one hand, coal resources are sometimes thought to be a key ingredient of economic development, at least since the pioneer analysis by Pomeranz (2014), who argues that living standards in Europe and in the most populated areas of China were not significantly different at the end of the XVIIIth century. In his opinion, the dramatic increase in European prosperity compared to Chinese levels started in the XIXth century, as European countries cashed in on their abundant coal resource to trigger the industrial revolution. On the other hand, natural resources may also induce adverse impacts on economic, social and human development. Among others, Marx (1867) was one of the first and most vivid opponents to the extractive industry which, to his mind, epitomized the exploitation of workers by capitalists. Over the past decades, a blooming literature has rekindled the interest regarding the effect of natural resources on long-run prosperity: to mention but a few of them, Sachs and Warner (1995) showed that economies with abundant natural resources tend to grow less rapidly than natural-resource-scarce economies; from a theoretical standpoint, Corden and Neary (1982) developed a model in which a boom in the natural resource industry may squeeze profitability in other traded good sectors. More recent work have focused on the negative long-run impact of coal on prosperity (Franck and Galor (2021), Esposito and Abramson (2021)) or on individual trajectories (Rud et al. (2024),

Allcott and Keniston (2017)).

Yet, in spite of the rising interest related to the effects of coal extraction in the long run, little attention has been paid to the impact of natural resources endowment on innovation. An understanding of the effect of resource endowment on innovation seems nevertheless crucial in light of Figure 1 which illustrates the long-lasting gap between patenting rates in Nord, a department which, as said earlier, was heavily involved in coal extraction compared to Isère, a department that barely mined any coal. While coal extraction is said to have lent a substantial support to early industrialization which, in turn, is often associated with innovation activities, there is no clear evidence that patenting activity rose quicker in departments abundantly endowed with coal resources; on the contrary, it looks like Nord (pictured in yellow), performed poorly with respects to Isère (pictured in navy blue). This finding holds irrespectively of whether I consider total patenting activity (Panel (a)), or patenting activity per capita (Panel (b)).

FIGURE 1: Patenting activity in Nord and in Isere, 1907 - 2007



Notes: Figure (a) plots the yearly total number of patents in Nord and in Isère from 1907 to 2007. Figure (b) plots the number of patents per 100,000 inhabitants in Nord and in Isère. Source: Bergeaud and Verluise (2024)

In order to assess the effect of natural resource on patenting activity, I first elaborate a highly-stylized innovation and industry-specialization model, whereby firms chose to work either in a resource-intensive industry using natural resources as an input or in a knowledge-intensive industry in which investments in knowledge are required. This model yields valuable predictions regarding the influence of natural resources availability on productivity, education and industry specialization. I then test these predictions using a newly assembled dataset spanning French departments throughout the XXth century. This dataset combines information on the number, the quality and the specialization of patents, as well as on the location of historical coalfields, population, geographical aspects and biological features either at the department or at the city level. To evaluate the incidence of coal on patenting activity, I complement OLS regressions with an IV strategy which, as is often done in this literature, exploits

the random heterogeneity in soil surfaces to instrument for the location of coalfields.

Overall, and consistently with the theoretical argument summarized in Section 2, my findings indicate that coal had a negative and lasting impact on French patenting activity at the local level. This effect was magnified for breakthrough inventions in the sense of Kelly et al. (2021), meaning that coal not only affected the quantity but also the quality of inventions. Said differently, coal reduced the ability of departments to invest in new technologies such as planes, computing sciences, communication or biotechnology. In terms of magnitude, my estimates suggest that a one standard deviation increase in coal exposure translated in a sustained reduction of innovation ranging between -17% and -28% for total patenting activity, and between 20% to 40 % for high-quality patenting activity. By contrast, inventions related to heavy industry sectors were almost unaffected by coal extraction. I further show that this negative impact reached other key economic parameters, including upper-tail education as well as industry specialization. I finally provide evidence that the persistence of such effects over time is accounted for by various mechanisms including path dependence, technological myopia induced by coal and a negative human capital feedback loop, whereby within-family transmission and brain drain patterns jointly slowed down the pace of human capital accumulation in coal basins.

Related literature My paper thus contributes to the literature by examining the effect of abundant resources on a wide set of economic outcomes, with a specific focus on innovation. The understanding of the effect of natural resource on economic outcomes has garnered a rising interest over time. The first contributions in that respect were that of Corden and Neary (1982) who developed a theoretical argument whereby a boom in a tradable sector, such as the natural resource industry, might squeeze productivity in other tradable sectors. On the empirical side, Sachs and Warner (1995) showed that countries endowed with natural resources tend to have slower growth rates compared to countries where natural resources are scarce. More recently, Esposito and Abramson (2021) found an adverse impact of coal on prosperity in the long run. Similarly, Franck and Galor (2021) provide evidence that early industrialization negatively affected economic development in the long run. To explain these findings, both papers insist on similar though slightly different mechanisms: while Esposito and Abramson (2021) investigate the education channel, Franck and Galor (2021) bring forward the concept of a dual technological and cultural lock-in, meaning that departments which experienced early industrialization were stuck in low-value added industries and were therefore reluctant to embrace economic progress. My paper contributes to this literature in a threefold way. First, I specifically focus on

patenting activity, both from a quantitative and a qualitative point of view, leveraging the richness as well as the granularity of the Bergeaud and Verluise (2024) dataset. Second, my findings suggest that the effect of coal on patenting activity is not akin to a reversal of fortune story, but rather pertains to a lasting negative impact. Third, I try to bring a theoretical contribution to the literature by building a simple two-sector model in which accessibility of natural resources drive an over-specialization in the mining sector, ending in poor human capital accumulation and hence innovation.

Additionally, this paper speaks to the blooming literature related to the determinants of economic growth and innovation. Important contributions in this field have been proposed, among others, by Aghion and Howitt (1992) who argue that growth results exclusively from technological progress which, itself, stems from competition among research firms that generate innovations. However, the innovation effort under laissez faire might either be too low or too high depending on the magnitude of two counteracting effects: on the one hand, inventors cannot fully reap the benefits of their innovation which reduces their incentives to create new ideas; on the other hand, inventors do not internalize the destruction of rents generated by their innovations. Whenever the former effect outweighs the latter, government intervention should support private R&D through better patent enforcement or subsidies to R&D. Other contributions in this field have stresses the role of historical and political factors in fostering long-run prosperity. Mokyr and Voth (2010) notably points the role of the Enlightenment period in shaping a specific European culture prone to science and progress; in particular, the European political fragmentation in the modern era paved the way for accrued competition between countries to attract major scientists which, in turn, stimulated inventions and productivity. From a political standpoint, Acemoglu et al. (2001) mention the role of institutions in long-run prosperity: leveraging the differences in institutional settings in former European colonies, they show that colonies where inclusive institutions, which encourage investments through better enforcement of property rights, were settled are now wealthier than those where extractive institutions, which concentrate power in the hands of a small elite and create a high risk of expropriation, prevailed. In that respect, Acemoglu et al. (2001) question the influence of geographical conditions on economic development and prosperity. Among others, Sachs et al. (2001) draws upon the intuition of Smith (1776) to contend that differences in soil characteristics as well as in accessibility can explain lasting economic patterns including the persistence of poverty across some countries in the world. In a complementary analysis, Sachs and Warner (2001) shows that countries endowed with abundant natural resources display poorer economic performances mostly because of price effects. I modestly contribute to this literature by providing empirical evidence of the negative influence of natural resources on productivity and

long run prosperity; given that the analysis is led at the department scale, meaning at a relatively disaggregated level which precludes differences in institutional settings among treated units, my results lend credence to the argument that natural resources, even at the local level, can result in adverse effects on prosperity and economic well being. Perhaps more importantly, this paper is, to the best of my knowledge, the first to document extensively the implications of coal mining on the pace of growth of innovation, as well as on its quality and specialization which ultimately drives economic prosperity.

Finally, this paper speaks to the voluminous literature on path dependence understood as the phenomenon where the current state or outcome of a system is influenced by its historical trajectory. David (1985) was the first to delve in that topic by showing the persistence of the QWERTY keyboard in spite of its poor performance with respects to other designs in terms of typing speed; he contends that this suboptimal persistence mostly stems from a lock-in effect, whereby both demand and supply for typing skills are used to a specific keyboard so that changing it would entail a significant cost. More recently, this literature has experienced a rekindle in the context of global climate change; among others, Aghion et al. (2012a) develop a theoretical model whereby firms with a larger track record in clean technologies are more likely to patent clean technologies in the future; conversely, firms involved in dirty technologies are less likely to develop green technologies, suggesting that government intervention might be required to comply with the Sustainable Development Goals (SDG). More specifically, they find upon calibrating their model with patent data that an increase of 40 percent of fuel prices would be required to allow clean innovation stocks to overtake dirty stocks in a 15 year span. Another important contribution in that field, though in a slightly different spirit, stems from Agrawal et al. (2009) who discuss the concept of company-towns, meaning cities where a small set of firms are overwhelmingly dominate the local labor market, implying that such cities tend to develop myopic behaviors, whereby they fail to diversity in various industries. Finally, recent papers have tried to bring evidence regarding the role of within-family transmission in lock-in mechanisms, whereby individuals fail to escape the education level or the job occupation of their ascendants: typically, Rud et al. (2024) look at the lasting impact of coalfield closures on wage and local employment; in a similar spirit, Brey and Rueda (2024) leverages the contraction of the coal industry in the UK to estimate the lifelong effects of being exposed to pit closures; they notably show that individuals exposed to such a shock as children have worse health throughout life, an effect which is transmitted over generations. My paper speaks to this path dependence literature in a twofold way: I first document a lasting impact of early exposure to coal industrialization on patenting activity, second, I examine the mecha-

nisms which led to the persistence of this effect, and notably show that specialization in the mining industry translated in a technological and cultural lock-in whereby departments abundantly endowed with natural resources suffered from self-reinforcing technological myopia and depletion of human capital.

The remainder of this paper is organized as follows: Section 2 drafts the theoretical argument. Section 3 details the data sources and outlines the methodology employed to measure exposure to coal mining. Section 4 presents the empirical strategy as well as the key results of this paper. Section 5 investigates the underlying mechanisms. Finally, Section 6 discusses potential threats to identification and suggests robustness checks.

2 The theoretical argument

2.1 Set-up

This set-up provides a static model of a two-sector economy with a continuum of firms of mass 1 and a representative consumer endowed with a revenue R. Firms can operate either in the resource-intensive sector i or in the knowledge-intensive sector j. The resource-intensive industry is one in which firms require natural resources as an input while the knowledge-intensive industry is one in which firms must acquire knowledge to generate output. The specificity of knowledge accumulated in the knowledge-intensive sector is that it can be transferred to the resource-intensive sector through between-industry spillovers. In short, the model essentially captures the trade-off between acquiring education and working in a knowledge-intensive sector which nurtures spillovers for other industries, and joining the resource-intensive sector, the profitability of which depends on the availability of natural resources. As will be explained later, this model builds on the theoretical framework elaborated by Aghion et al. (2015).

2.2 Environment

Consumption Industries i and j produce two goods, namely x_i and x_j. The representative consumer is endowed with a revenue R which she uses to maximizes her utility with respects to her budget constraint; formally, she optimizes:

$$\mathcal{U}(x_i, x_j) = \ln x_i + \ln x_j \tag{1}$$

Subject to the resource constraint:

$$p_x x_i + p_j x_j \leq R \tag{2}$$

Production There is a continuum of firms of mass 1 which join either the resource-intensive industry i or the knowledge-intensive industry j. Industry i is resource-intensive in the sense that it uses natural resources as an input; more specifically, firms joining the resource-intensive industry must pay a fixed cost which is increasing in the price of natural resources p_n. For simplicity, I assume that this fixed cost is linear in the price of natural resources so that profit for a firm k in the resource-intensive industry writes:

$$\Pi_{i,k}(p_i, p_n) = p_i D(p_i) \pi_{i,k} - p_n \tag{3}$$

Where p_i indexes the price of good x_i, $D(p_i)$ stacks total demand for good x_i and $\pi_{i,k}$ captures the profit made per unit sold by firm k. Equation 3 basically states that the profit of the firm depends on (i) her sales, (ii) her profit per unit sold and (iii) the price of natural resources. Symmetrically, firms may join the knowledge-intensive sector j in which case they need to pay a firm-specific knowledge cost κ_k where k indexes the firm. This is like saying that firms have differential abilities in terms of research and development which may depend on location, history, or workers. Similarly to equation 3, the profit made by a firm k in the knowledge-intensive industry writes:

$$\Pi_{j,k}(p_j, \kappa_k) = p_j D(p_j) \pi_{j,k} - \kappa_k \tag{4}$$

Innovation In both industries, firm may improve their profit margin per unit sold upon investing in R&D processes. More specifically, I follow Aghion et al. (2012b) who assume that a firm investing $\frac{q^2}{2}$ in R&D processes innovates successfully with probability q in which case she divides her marginal cost by a scale $(1 + \xi_l)$ where $l \in \{i, j\}$, so that ξ_l is industry-specific. As firms compete *a la Bertrand*, a firm who successfully innovates keeps her price at the marginal cost c_l of her closest competitor so that, with probability q_l, she derives a per-unit profit of:

$$\beta_l = (1 + \xi_l) c_l \tag{5}$$

This equation establishes that the per-unit sold profit margin β_l of a firm operating in sector l is increasing in the radicality of innovation as measured by ξ_l. Without loss

of generality, I assume that c_l is set to 1 which implies that:

$$\forall l \in \{i, j\}, \beta_l = 1 + \xi_l \tag{6}$$

With probability $1 - q_l$, innovation efforts in industry l are unsuccessful and the firm simply gets the competitive profit. While Aghion et al. (2012b) allow for a general formula, where competitive profit is negatively related to the within-industry degree of competition, I normalize profits in case innovation fails to 0. It is shown in the Appendices that this assumption is made without loss of generality, but it nevertheless proves useful to derive simple formulas.

Profits Wrapping-up, this implies that firms who join the resource-intensive sector maximize the following problem:

$$\Pi_{i,k}(q_i; p_i, p_n) = q_{i,k}\beta_i p_i D(p_i) - \frac{q_{i,k}^2}{2} - p_n \tag{7}$$

As price is exogenously fixed and is set to the initial marginal cost of each sector, which I normalized to be 1, this implies that:

$$\Pi_{i,k}(q_i; p_n) = q_{i,k}\beta_i D(1) - \frac{q_{i,k}^2}{2} - p_n \tag{8}$$

I introduce here a key assumption which departs from Aghion et al. (2015): I posit that the extent of cost reduction incurred by innovation effort in the resource-intensive industry is increasing in the innovation efforts achieved by the knowledge-intensive industry. Said differently, the resource-intensive industry benefits from spillovers emanating from the knowledge-intensive industry.

Assumption 1. *The radicality of innovation achieved by the resource-intensive industry i is linearly increasing in the innovation effort achieved by the knowledge-intensive sector j.*

Formally, this assumption writes:

$$\beta_i(\mathcal{Q}_j) = (1 + \xi_i)(\mathcal{Q}_j + 1) \tag{9}$$

Where:

$$\mathcal{Q}_j = \int_{k \in \mathcal{J}} q_{j,k} dk \tag{10}$$

And $\mathcal{J}$ indexes the set of firms who join the knowledge-intensive industry. This implies in turn that the profit of a firm operating in the resource-intensive industry i writes:

$$\Pi_{i,k}(q_{i,k}; p_n, \mathcal{Q}_j) = q_i \beta_i(\mathcal{Q}_j) D(1) - \frac{q_{i,k}^2}{2} - p_n \tag{11}$$

Meanwhile, I assume that there are no spillovers emanating from the resource-intensive industry, so that total profit achieved by a firm k operating in the knowledge-intensive industry simply writes:

$$\Pi_{j,k}(q_j; \kappa_k) = q_j \beta_j D(1) - \frac{q_j^2}{2} - \kappa_k \tag{12}$$

2.3 Solving the model

I then solve the model in three steps: first, the representative consumer optimizes her consumption given her budget constraint; second, firms choose the optimal innovation effort achieved in each industry and taking demand for granted; third and finally, I characterize the optimal specialization decisions made by each firm depending on the profits they may derive in each industry. These three steps pin down a unique equilibrium, the features of which are studied in the next subsection.

Consumption The representative consumer maximizes her utility under her resource constraint. Without loss of generality, I normalize R to 2 which trivially implies, as shown in Appendix A.1 that:

$$\forall l \in \{i, j\}, x_l^* = 1 \tag{13}$$

Innovation Plugging the optimal consumption level in the profit equations implies that firms, conditional on working in the resource-intensive industry i, optimize:

$$\Pi_{i,k}(q_{i,k}; p_n, \mathcal{Q}_j) = q_i \beta_i(\mathcal{Q}_j) - \frac{q_{i,k}^2}{2} - p_n \tag{14}$$

Then, as proved by Appendix A.2, the optimal innovation effort $q_{i,k}^*$ of firm k in the resource-intensive industry i writes:

$$q_{i,k}^* = (1 + \xi_i)(1 + \mathcal{Q}_j) \tag{15}$$

This equation basically states that the innovation effort achieved by a firm in the resource-intensive industry depends on the extent to which it may reduce its cost upon innovating. This, in turn, depends on two components: first, on the industry-specific parameter ξ_i, second on the spillovers emanating from the knowledge-industry as captured by $\mathcal{Q}_j$. Similarly, substituting $D(1) = 1$ implies that profit in the knowledge-intensive industry writes:

$$\Pi_{j,k}(q_j; \kappa_k) = q_{j,k}\beta_j - \frac{q_{j,k}^2}{2} - \kappa_k \tag{16}$$

As derived in Appendix A.2, the optimal innovation effort $q_{j,k}^*$ of a firm k in the knowledge-intensive industry j then writes:

$$q_{j,k}^* = \beta_j \tag{17}$$

This innovation effort only depends on the industry-specific parameter β_j which captures the extent to which firms in the knowledge-intensive industry may reduce their costs upon achieving successful innovation efforts. The key difference with respects to equation 15 is then that innovation in the knowledge-intensive industry does not depend on that in the resource-intensive industry. In other words, firms in the knowledge-intensive industry are generating positive externalities towards the resource-intensive one, while firms in the resource-intensive industry are generating no spillovers.

Specialization In order to choose which industry to operate in, a firm must compare its expected profits upon joining the resource-intensive industry with that obtained upon joining the knowledge-intensive industry. Given the optimal innovation decisions derived above, the profit of a firm joining the resource-intensive industry writes:

$$\Pi_{i,k}(p_n, \mathcal{Q}_j, \Delta_i) = \frac{1}{2}\left[(1+\xi_i)(1+\mathcal{Q}_j)\right]^2 - p_n \tag{18}$$

Reciprocally, the profit of a firm with ability κ_k joining the knowledge-intensive industry writes:

$$\Pi_{j,k}(\beta_j, \Delta_j, \kappa_k) = \frac{1}{2}\beta_j^2 - \kappa_k \tag{19}$$

For the sake of simplicity, I posit Assumption 2, which, without loss of generality, fixes the distribution of the knowledge cost κ_k :

Assumption 2. *The knowledge cost parameter κ_k is uniformly distributed over the interval* $[0, 1]$.

As will be demonstrated in Appendix A.3, and under assumption 2 and 4, there exists a unique value $\bar{\kappa}$ such that firms with a cost inferior to this cutoff value join the knowledge-intensive industry and firms with a cost in excess of this cutoff value join the resource-intensive industry. This yields Lemma 1:

Lemma 1. *There exists a unique cutoff value $\bar{\kappa}$ such that all firms with a cost of knowledge inferior to that cutoff join the knowledge intensive sector while firms with a cost larger than this cutoff join the resource-intensive sector.*

As $\bar{\kappa}$ directly pins down the share of firms who join the knowledge-intensive sector, it allows to compute innovation in both industries, total education effort, total innovation and, finally, total consumption. Hence, Lemma 1 ensures the existence of a unique general equilibrium which is uniquely characterized by the set of parameter $\langle \beta_i, \beta_j, p_n \rangle$.

2.4 Equilibrium properties

In this subsection, I briefly summarize the main characteristics of this equilibrium, by investigating the impact of the main parameter, namely the price of natural resources p_n, on industry specialization, education, total innovation effort and profits.

Industry specialization As shown in Appendix A.4, the price of natural resources positively affects the share of the knowledge-intensive sector. This finding is summarized in Proposition 1:

Proposition 1. *The larger the price of natural resources, the larger the share of the knowledge-intensive sector. Said differently, the larger the availability of natural resources, the lower the share of the knowledge-intensive sector.*

This proposition trivially stems from the trade-off that firms face upon choosing which industry to join. The larger the price of natural resources, the larger the fixed cost firms need to pay to enter the resource-intensive sector and, hence, the more profitable it becomes, all else equal, to invest in education and join the knowledge-intensive industry. Conversely, the lower the price of natural resources, meaning the larger their availability, the smaller the firm-specific knowledge cost must be to ensure that joining the knowledge-intensive industry remains profitable.

Knowledge level Knowledge is also a key variable in the model given that firms face a trade-off between investing in knowledge and joining the knowledge-intensive sector or joining the resource-intensive industry where no specific knowledge is required. Because this trade-off essentially funnels through two parameters, namely the price of natural resources on the one hand, and the firm-specific knowledge cost κ_k on the other hand, it proves quite intuitive to demonstrate Proposition 2 as will be done in Appendix A.5:

Proposition 2. *The overall knowledge level κ in the economy is decreasing in the availability of natural resources, which is like saying that it is increasing in the price of natural resources p_n.*

Proposition 2 is a direct corollary of Proposition 1 in the sense that firms need to pay a knowledge-fixed cost to join the knowledge-intensive industry. In that respect, the larger the price of natural resources, meaning the lower their availability, the larger the number of firms for which it is profitable to invest in knowledge and, in turn, to join the knowledge-intensive industry. By contrast, a lower price of natural resources, or, alternatively, a larger availability of such resources, discourages firms from investing in knowledge which translates in a larger share of the resource-intensive industry.

Innovation effort Innovation is the main outcome of this model and has garnered an increasing awareness since at least the pioneering work by Aghion and Howitt (1992). In the considered framework, innovation effort basically depends on two complementary forces: first, it depends on the individual effort achieved by each firm depending on the industry it works in; second, it depends on the sectoral composition of such an economy, for there are innovation spillovers between the knowledge-intensive industry and the resource-intensive one. Overall, total innovation $\mathcal{Q}$ writes:

$$\mathcal{Q} = s_i q_i^* + s_j q_j^* \tag{20}$$

Where s_i and s_j respectively index the share of the resource-intensive industry and that of the knowledge-intensive industry while q_i^* and q_j^* capture the optimal innovation effort of a firm operating respectively in the resource-intensive industry and in the knowledge intensive one. As derived in Appendix A.6, it can be proved that there exists a unique optimal value of the price of natural resources p_n which maximizes total innovation. This assertion is summarized in Proposition 3:

Proposition 3. *There is an inverted U relationship between the price of natural resources p_n and total innovation in the economy. In other words, total innovation first increases with the price of natural resources and then decreases once the optimal cutoff value p_n^* is met.*

This proposition naturally comes with Corollary 1:

Corollary 1. *There exists a unique value s_i of the share of the resource-intensive industry in the economy which maximizes total innovation.*

These propositions are quite straightforward to interpret: they suggest that a variation in the price of natural resources has ambiguous effects on total innovation depending on the initial price of such resources. Indeed, when the price of natural resources is excessively large, firms tend to specialize in the knowledge-intensive sector which then achieves large innovation efforts which benefit the resource-intensive sectors through between-industry spillovers; yet, as the fraction of firms specializing in the resource-intensive industry is small, spillovers are, to some extent, spoiled which lowers the overall innovation effort. Conversely, when the price of natural resources is excessively low, a large share of firms prefer to operate in the resource-intensive industry instead of investing in the knowledge-intensive industry which, in turn, lowers the innovation effort achieved in the knowledge-intensive industry and, finally, reduces the amount of spillovers emanating from the knowledge-intensive sector.

Profit A last and important prediction of the model speaks to the overall profit level in such an economy or, to state it differently, to the overall economic output. Let Π be total profit in such an economy; it is straightforward to show that:

$$\Pi = \int_{k \in \mathcal{I}} \Pi_{i,k} dk + \int_{k \in \mathcal{J}} \Pi_{j,k} dk \tag{21}$$

Where $\mathcal{I}$ and $\mathcal{J}$ respectively denote the set of firms joining the resource-intensive industry and that of firms joining the knowledge-intensive industry. In that respect, total profit Π depends both on the individual profit achieved by firms operating in each of those industries and on the share of each industry in the economy which, itself, varies with the expected profit recovered in each industry. Again, and as will be shown in Appendix A.7, there exists an optimal value of the price of natural resources which maximizes total profit in the economy. This is summarized by Proposition 4:

Proposition 4. *There exists an inverted-U relationship between the price of natural resources and total profits derived in such an economy.*

Proposition 4 is equivalent to its natural corollary 2:

Corollary 2. *There exists a unique growth-maximizing share of the resource-intensive in such an economy, such that total profits are maximized for this value.*

Proposition 4 and corollary 2 trivially state that the availability of natural resources exert ambiguous effects on total growth in such an economy. On the one hand, available natural resources are profitable for they lower the costs of the resource-intensive industry and hence maximize profit in this industry; meanwhile, an excessively low price of natural resources lead to an over-specialization in the resource-intensive sector, which, in turn, makes the knowledge-intensive sector shrink and thus curtails the spillovers emanating from that industry. In the end, there exists a unique optimal value of the share of the resource-intensive sector such that innovation is maximized in the economy. Given the negative relationship between the share of the resource-intensive sector and the price of natural resources, this is like saying that there exists an inverted U relationship between the price of natural resources and total innovation in the economy.

2.5 Welfare analysis

The theoretical argument suggests so far that the price of natural resources exerts an effect on total profits through two main channels, namely the composition effect and the knowledge effect. The composition effect encompasses the share of each industry in the economy while the knowledge effect captures the education effort achieved by the knowledge-intensive industry which benefits the resource-intensive one. I now turn to investigate the negative externalities stemming from this model. More specifically, this subsection shows that the share of the knowledge-intensive industry is likely to be inefficiently low in such an economy which, in turn, makes total profits and total innovation effort sub-optimally low.

Profits To know whether private decisions lead to the same outcome as that which would prevail in an economy ruled by a benevolent social planner, it suffices to derive the optimal share of the resource-intensive sector that would be picked by the central planner and compare it to that which stems from private firms decisions. Comparing those two values implies studying the sign of $\mathcal{G}$, where:

$$\mathcal{G} = s_i^a - s_i^* \tag{22}$$

Where s_i^a and s_i^* respectively capture the actual and the optimal value of the share of the resource-intensive industry in the economy. In an economy where private agents fully internalize the externalities, $\mathcal{G}$ would mechanically tend to 0. By contrast, if agents do not fully take into account the effect of their private decisions on overall welfare, a discrepancy would appear between the share of the resource-intensive

industry resulting from private decisions and that which would maximize overall welfare. To simplify computations, and derive tractable results, I posit the following assumption:

Assumption 3. *The value of the radicality parameter in the resource-intensive industry, namely ξ_i, is normalized to 0.*

This assumption imposes, without loss of generality, that the resource-intensive sector crucially depends on the innovation effort achieved by the knowledge-intensive industry to reduce its costs. Using assumption 4 then implies that $\mathcal{G}$ can be written:

$$\mathcal{G} = \frac{\beta_j(\beta_j + 1) + 2 + \phi(p_n)}{\beta_j^2} \tag{23}$$

Where ϕ is a decreasing function of the price of natural resources p_n. As will be demonstrated in Appendix A.8, it follows that $\mathcal{G}$ is positive for reasonable values of p_n and decreasing in the price p_n of natural resources; this is what Proposition 5 states:

Proposition 5. *There exists a positive wedge between the actual and the optimal share of the resource-intensive industry which decreases with the price of natural resources.*

In short, this proposition establishes that the actual share of the resource-intensive industry is excessively large with respects to the socially optimal level, which stems from the fact that private firms do not take into account the positive effect of their innovation in the knowledge-intensive industry on total innovation in the economy. In turn, firms tend to under specialize in the knowledge-intensive industry, which is all the truer when the price of natural resources is low, meaning that natural resources are easily available.

Education As a by-product of Proposition 5, it can be easily shown that the overall education effort of the economy is inefficiently low, and that this negative externality is all the larger as the price of natural resources is low. In concrete words, this means that a large availability of natural resources incentivize firms into joining the resource-intensive industry which, in turn, precludes them from joining the knowledge-intensive industry where they are led to invest in knowledge and R&D processes. This finding is summarized formally in Corollary 3:

Corollary 3. *There is a negative wedge between the actual knowledge level and the socially optimal one, in the sense that the actual knowledge level in the economy is lower than the optimal one that a central planner would achieve.*

Innovation In turn, innovation is inefficiently low in this economy given the spillovers emanating from the knowledge-intensive industry. Also, this negative wedge between the actual level of innovation and the socially optimal one is increasing in the availability of natural resources or, said differently, decreasing in the price of natural resources. This is what Corollary 4 states more formally:

Corollary 4. *The negative wedge between actual innovation and socially optimal innovation is increasing in the availability of natural resources or, said differently, decreasing in the price of natural resources.*

Corollary 4 is very similar to Proposition 5 in that it suggests that the overall innovation effort of the economy is inefficiently low, and that this negative externality is increasing in the availability of natural resources. Indeed, the greater the availability of natural resources, meaning, the lower the price of natural resources, the larger the share of the resource-intensive sector which, in turn, lowers the share of the knowledge-intensive industry. Yet, the former is generating positive spillovers in terms of productivity for the resource-intensive industry, implying that firms tend to over-specialize in the resource-intensive industry, and this all the more as a positive shock on natural resources availability occurs.

2.6 Summary of Empirical Expectations

Following the theoretical argument summarized above, it can be expected that:

1. *Excessively abundant natural resources lower the pace of innovation*: with a low price of natural resources, or, said differently, a large stock of natural resources, the economy delivers less inventions.

2. *Excessively abundant natural resources lower the incentive to acquire knowledge*: with a low price of natural resources, or, said differently, a large stock of natural resources, investment in knowledge becomes less profitable.

3. *Excessively abundant natural resources lead to an over-specialization in resource-intensive industries*: as the price of natural resources decreases, the share of resource-intensive industries increase up to an inefficiently large level, given that firms do not internalize the positive effects associated to specialization in high value-added industries.

Each of these predictions will be tested separately using different datasets; the remainder of this paper is organized as follows: section 3 discusses the French context

as well as data sources; section 4 summarizes the empirical strategy as well as the main results while section 5 elabores on plausible mechanisms that may explain persistence of such effects over time; finally, section 6 concludes and gives orientation for future research.

3 Background and data

This paper relies on a novel dataset at the department level, assembling information from various sources for each of the 95 French departments. Among others, it notably includes patent data, exposure to historical coalfields, education data, industry outcomes and demographic characteristics. I start by providing some context regarding coal extraction in France, and then move on to document data sources. Finally, I briefly explain how I construct the measure of coal exposure, the main regressor used in this paper.

3.1 Context

As highlighted, coal extraction started in the beginning of the XVIIIth century in France when the first coal concession was granted in the city of Anzin in 1717. Unlike in the United Kingdom, where coal output already reached 3 millions tons by 1700, it took decades for French coal industry to become sizable: at the end of the XVIIIth century, only 15 more coalfields had opened in France, and the output remained negligible. The number of concessions took off in the beginning of the XIXth century to reach a record level of 1,090 in 1907 before plateauing at around 800 for the subsequent decades. Improvements in the mining techniques however ensured that, in spite of a stagnation of the number of mine concessions, the output per capita increased steadily throughout the first half of the XXth century. This successful story however came to a halt in the middle of the XXth century, as the Jeanneney plan paved the way for the dismantling of the French coal industry. Step by step, coal extraction slowed down and the last French coalfield located in Creutzwald in the Moselle department closed in 2004, almost 110 years after its foundation. This process did not go without struggles: shutdowns of historical coalfields were often accompanied by protest movements from miners such as the famous 1963 strike which was followed by more than 90% of French miners.

3.2 Data sources

Coal Data To measure coal exposure at the department level, I rely on Camino, a recent dataset released by the French Ministry for Environment to compile information on all mining fields on the French territory since the early XVIIIth century. More specifically, Camino provides extensive details on the resource mined in each site, the location, the surface covered or the time span of the concession. Overall, I identify 1,092 licenses to run coalfields in France at the onset of the XXth century covering almost 7,000 square kilometers of the French territory. The main advantage of this dataset is thus twofold: first of all, its richness allows me to measure the heterogeneity in the exposure of each department to coal both spatially and across time; second, its granularity allows for a fine-grained identification strategy even at the city scale, a level of analysis which is sometimes leveraged in this paper. To ensure that my results are not driven by one specific measure of coal epxosure, I nevertheless complement this dataset with the location of the 16 largest historical French coalfields which I retrieve from the European Route of Industrial Heritage project. The project censuses historical industrial heritage sites across 44 European countries in Western Europe, and has already been used in the related literature (see, for instance, Esposito and Abramson (2021)). Later in this section, I show that the measures of coal exposure stemming from these alternative sources happen to coincide, and I show throughout the paper as well as in Appendices that my estimates are robust to using various measures of coal exposure.

Geographical data As will be explained later, coalfields location might not be totally random, in particular because coal extraction requires both a large labor force and specific technical skills. To mitigate the risk that local geographical characteristics are driving the estimate, as well as to implement the Instrumental Variable (IV) strategy detailed below, I hence retrieve various geological and geographical features at the department level in France. In terms of geographical characteristics, I mostly build on Luterbacher et al. (2004) who compiled long-run geolocated temperature data. I complement these details with land features collected by Ramankutty et al. (2002). Also, I retrieve from Fernihough and O'Rourke (2014) various measures of distance to important natural elements such as the Atlantic Ocean or the Mediterranean Sea. Finally, I exploit the well-known $1:5$ Million International Geological Map of Europe and Adjacent Areas from Asch (2005) to recover the location of carboniferous rock strata.

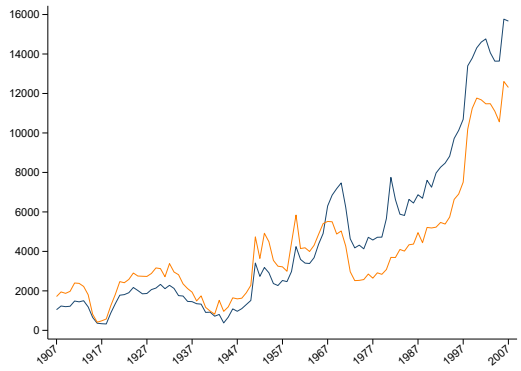
Patent Data The main dependent variable used in this paper is a measure of local innovation that I construct from counting the number of newly issued patents filed by either resident or international inventors every year. I recover patent data from the PatentCity dataset (Bergeaud and Verluise, 2024), which provides information on the location and nature of every patent filed in France since 1907, including the department and the name of the inventor and the technological class. In order to adequately reflect the innovation effort per department, I perform a minor adjustment to the count of patents: in case of multiple inventors in a single patent, I split the patent across the different departments from which its inventors come, so that each department is assigned a fraction of the patent. While my main results still hold when relaxing this adjustment, I shall only report the estimates of regressions using this measure in the main paper and postpone the estimates for the non-adjusted patent count to the Appendices. Patenting activity showed substantial variation during the XXth century; overall, and as pictured by 2 it followed an upward trend interrupted at times by major shocks such as both World Wars and the crises related to oil price surges in the late seventies. Perhaps more strikingly, what Figure 2 suggests is that patenting activity followed very different paths for coal and non coal regions throughout the XXth century; it indeed pictures the evolution of total patenting activity (Panel (a)) and patenting activity per capita (Panel (b)) in departments affected by coal mining (gold line) versus departments unaffected by coal mining (blue line) [1]. While the difference in innovation effort intensity it not salient before WWII, innovation trajectories in both kinds of departments diverge significantly in the aftermath of World War II, both in absolute and in relative values.

Education data Education ranks among the main outcome variables studied in this paper to explain the sustained difficulties experienced by coal departments in patenting activity. I exploit different sources spanning throughout the XXth century to reflect the lasting impact of exposure to coal industry on educational attainment. Among others, I hand collect the number of young men enrolled in the French Army who held the baccalauréat degree at the end of the XIXth century as well as in the beginning of the XXth century from the *Annuaires Statistiques de la France*. While baccalauréat rates now stand at record high levels (c. 90% in 2024 in France), they used to encompass a very high level of education achievement: typically, the average baccalauréat rate among youg men barely reached 2.1% in 1910, suggesting that hold-

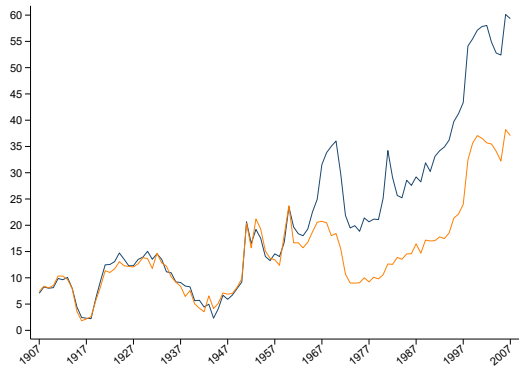
[1] For the sake of clarity, we removed Paris from the sample used to build those graphs: Paris indeed concentrated more than 50% of patenting activity throughout the XXth century, so that adding Paris in the sample would distort patenting activity in favor of non coal regions.

FIGURE 2: Patenting activity in coal and non coal regions

(a) Total patenting activity



(b) Patenting activity per capita



Notes: Figure (a) plots patenting activity (excluding Paris) in coal (gold) and non coal regions (blue) while Figure (b) pictures patenting activity per capita (excluding Paris) in coal (gold) and non coal regions (blue). Source: (Bergeaud and Verluise, 2024)

ing this degree was a signal of high educational abilities. To complement this early XXth century dataset, I obtain from censuses compiled by INSEE the breakdown of young people aged 18 to 25 according to their achieved level of studies; data is aggregated at the city level for seven years after WW2, respectively 1968, 1975, 1982, 1990, 1999, 2009 and 2016. Finally, I exploit administrative surveys conveying valuable and precise insights regarding the type of studies pursued by French citizens across the second half of the XXth century. To mention a few of them, I leverage the *Enquête Emploi en Continu* survey, which was born in the 50s to provide a regular measure of unemployment; later on, it expanded its coverage to include detailed data about the type of jobs occupied by each respondent as well as her birth place, her level of studies or her migration record. I also rely on the 1975 *Enquête Formation et Qualification Professionnelle* which records data on studies for roughly 35,000 respondents as well as the jobs occupied by their father and their grandfathers.

Employment data Industry specialization is also one of the key channels through which coal induced a negative impact on patenting activity in the long run. To evaluate the strength of this channel, I combine various sources spanning, again, the whole XXth century as well as the beginning of the XXIst century. I notably exploit the censuses of 1901, 1906, 1911, 1921, 1926, 1931 and 1936 to compute the share of some sectors of interest, such as engineering occupations, liberal jobs, chemistry or coal mining from 1901 to 1936. To perform a similar analysis later on in the XXth century, I mostly focus on the afore mentioned *Enquêtes Emploi en Continu* which cover more than 40 years starting 1980. The breakdown of occupation across sectors is very precise as it is made at the 4-digit NAF2 level, allowing for a fine-grained analysis. Finally, I again leverage the 1975 survey *Enquête Formation et Qualification Profession-*

nelle to investigate the transmission of job occupations within families, a key channel of path dependence in coal regions.

3.3 Construction of the coal exposure measure

Throughout the paper, and for the sake of robustness, I present my main results for different measures of exposure to coal industry. Most of the time, out of concision motives, I report coefficients for two to three main measures of coal exposure, but document the value of the coefficient attached to the other ones in the Appendices. As highlighted, the main measures of exposure to coal industry are built using two different data sources, respectively the *Cadastre minier* released by the French Ministry for Environment and the *European Route of Industrial Heritage* (ERIH) project. While the first dataset provides the localization of all coalfields licenses granted throughout French mining history, the second dataset informs on the main and historical coalfields located in France, meaning the most important coal mining sites in terms of coal output.

From the first dataset, I compute two measures of exposure to coal industry; respectively, I calculate $\tilde{s_i}$, the (log) surface of department i's territory that is covered with a coalfield, and the average proximity $\tilde{p_i}^1$ of all cities in the department to the closest coalfield. Formally, those measures are derived from the following equations:

$$\tilde{s_i} = \ln(1 + s_i) \tag{1}$$

Where s_i captures the total surface of coalfields actively run in department i. The logarithmic transformation is applied to keep departments where no coalfield is run in the sample; as such departments represent 44 out of the 95 French departments, this transform sounds necessary. A complementary measure is given by the average proximity $\tilde{p_i}^1$, which is computed as below:

$$\tilde{p_i}^1 = \ln(\frac{1}{1 + \tilde{d_i}^1}) \tag{2}$$

Where $\tilde{d_i}^1$ is the average minimal distance between cities of department i and the closest coalfield:

$$\tilde{d_i}^1 = \frac{1}{n_i} \sum_{j \in \Gamma_i} \min_{c \in \chi^1} d_{j,c} \tag{3}$$

Where n_i captures the number of cities in department i, Γ_i is the set of cities in department i and χ^1 indexes the set of coalfields. In concrete terms, $\tilde{d}_i$ is the average distance between each city of department i and the closest coalfield. Average proximity of department i to coalfields is then defined as the natural logarithm of this average distance scaled with unity. This measure is actually derived from the existing literature, namely from previous work by Esposito and Abramson (2021). It indeed features different advantages related to the log structure of the measure: because of the concavity of the log function, it ensures fast decreasing returns to scale in the exposure to coalfields, meaning that what crucially matters in terms of coal exposure is to have a large number of cities located in the direct vicinity of a coalfield. Additionally, and as emphasized by Esposito and Abramson (2021), proximity is a rather convenient indicator in terms of explanatory power: indeed, it seems more relevant from an interpretative point of view to consider the effect of an increased proximity towards coalfields rather than the effect of an increased distance.

From the second dataset, I compute $\tilde{p_i}^2$ which is recovered from the same formula as $\tilde{p_i}^1$ except for the fact that I substitute $\tilde{d_i}^2$ for $\tilde{d_i}^1$ in Equation 2, where $\tilde{d_i}^2$ is defined as follows:

$$\tilde{d_i}^2 = \frac{1}{n} \sum_{j \in \Gamma_i} \min_{c \in \chi^2} d_{j,c} \tag{4}$$

Where χ^2 indexes the set of coalfields registered in the ERIH project, meaning the set of the most important coalfields run throughout French history. As complementary robustness checks, I also show in the Appendix that the main results still hold when we define alternatively $\tilde{p_i}^k$ for $k \in 1, 2$ as:

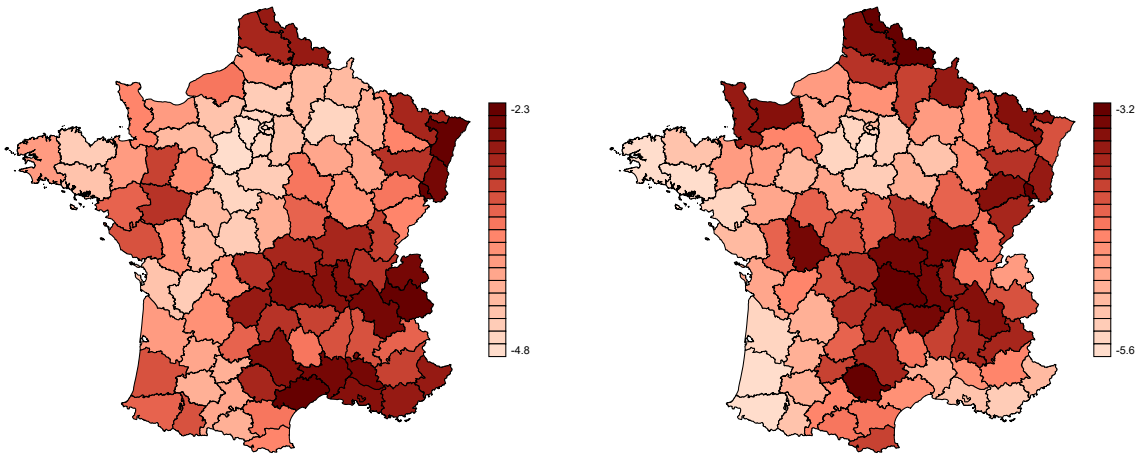
$$\tilde{p_i}^k = \frac{1}{n_i} \min_{c \in \chi^j} p_{j,c} \tag{5}$$

In words, this alternative measure takes the average of proximity indicators computed at the city level rather than applying the proximity transformation on the average distance which, itself, is measured at the city level. All in all, the various measures suggested in this subsection not only allow to test the robustness of the effect through a wide range of estimates, but it also provides valuable insights in terms of quantification by delivering a range of marginal effects rather than one single point estimate. In any case, the stability of the coal exposure measure across the different specifications is supported by Figure 3 which plots the exposure to coal industry at the department level for the two measures of $\tilde{p_i}^k$ discussed above. While the second measure plotted in Panel (b) is a bit more restrictive since it only focuses on the largest coalfields, it

FIGURE 3: Two continuous measures of coal exposure

(a) Using all French coalfields

(b) Using only major coalfields



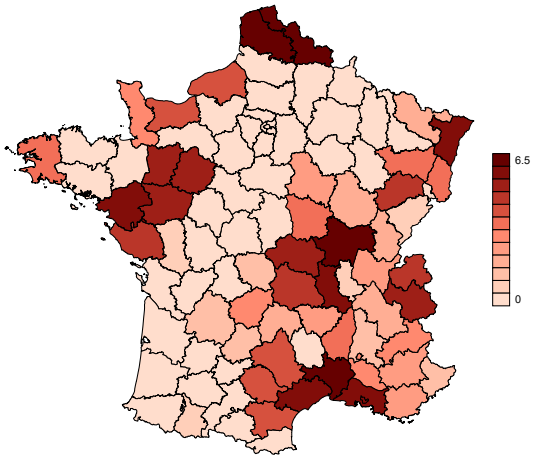
Notes: Figure (a) plots the coal proximity measure when taking all French coalfields operated by 1900 into account while Figure (b) pictures the coal proximity measure when restricting the set of coalfields to historical major coalfields documented in the ERIH project.

also better captures the major coal basins of Nord Pas de Calais, Alsace Lorraine and Saône et Loire. In the meantime, this specification downsizes the influence of coal extraction in the Southern part of France, in spite of a long mining tradition.

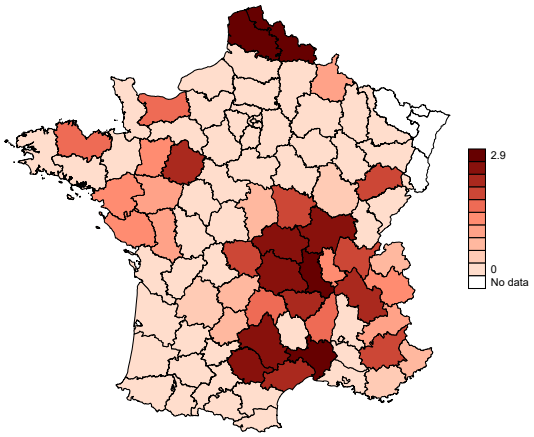
In spite of their subtle differences, those two specifications are similar in that they introduce a rather continuous measure of exposure to coal-related activities. In that respect, they may not fully account for the discrete difference that might exist between departments running coalfields and those not running any coalfield. As discussed previously, 44 out of the 95 French departments never hosted any coalfield prior to 1900, so that they might have remained untouched by coal extraction throughout the XXth century. In order to prevent potential concerns related to the continuity of the indicators $\tilde{p}_i^k$ plotted above, I also report in Figure 4 the values of coal exposure when measured by the log transform of the department's surface covered with a coalfield or the average log output of coal at the onset of the XXth century, which I recover from the *Annuaires Statistiques de la France*. Strikingly, coal exposure appears much more concentrated under these semi-discrete measures than under the rather continuous ones: this could represent both an advantage and a flaw. On the pro side, it enables to clearly distinguish between departments depending on whether they used to mine coal or not. On the con side, it makes the assumption that coal extraction did not generate any spillovers, which might be disputed in case of coalfields located close to the border with the next department. In light of the different pros and cons of the four measures introduced above, robustness checks will be performed for each regression to ensure that the estimates still hold under alternative measures of coal

FIGURE 4: Two semi-discrete measures of coal exposure

(a) Using all French coalfields



(b) Using only major coalfields



Notes: Figure (a) plots the (log) surface of coalfields while Figure (b) pictures the (log) average coal extraction output in tons at the onset of the XXth century. Sources: *Cadastre minier* and *Annuaires Statistiques de la France*

exposure. Finally, I also check in some of the main regression tables as well as in Appendix D that my results still hold when measuring coal exposure as a binary variable taking the value 1 if at least one coalfield was ran at some point in time in a given department.

4 Main results

4.1 Simple empirical specification

To measure the impact of exposure to coal industry on patenting activity, education and industry specialization throughout the XXth century and in the beginning of the XXIst century as well, I estimate pseudo poisson specifications and OLS regressions complemented with their IV strategy counterparts. My main dependent variable is either patenting activity, the education level or the share of specific industries in total employment. Formally, I estimate models of the following structure:

$$Y_{d,t} = \beta e_d + \chi \Gamma_d + \alpha_t + \varepsilon_{d,t} \tag{6}$$

Where $Y_{d,t}$ depicts the variable of interest, which can either be total patenting activity, both in counts or in logs, education level or the share of specific industries in total employment. Meanwhile, e_d is the coal exposure measure as described in Section 3 and Γ_d is a vector of geological, geographical and demographic controls usually

considered in that literature and detailed in Section 3: more specifically, they include distance to Paris, distance to the closest sea, suitability of land for agriculture activities, average value and standard deviation of temperatures since the XVth century, average altitude, ruggedness of soils, the size of agriculture land and the initial density of population at the onset of my period of analysis. Finally, α_t controls for year fixed effects that could influence patenting activity at time t. Two important points should be made clear: first, because of the time invariant structure of the main dependent variable, meaning coal exposure, I cannot control for department fixed effects; I however mitigate this issue in a twofold way: first, I control for a very wide range of local characteristics spanning from geographical to geological features as well as a demographic time varying component; I nevertheless show in the Appendices that the results are robust to allowing for region fixed effects, where region either refers to the old breakdown of French metropolitan territory in 22 entities, or to the new ventilation which started as of 2015 and only encompasses 15 entities.

Second, I sometimes allow for a time varying regressor, namely $p_{d,t}$ which is the density of the population recovered from the different censuses collected in France throughout the period under study; more specifically, 14 censuses were conduced at the national scale including 1901, 1906, 1911, 1921, 1926, 1931, 1936, 1946, 1954, 1962, 1968, 1975, 1982, 1990, 1999, 2006; later on, year on year assessments of the French population would be made. Controlling for the density of the population might prove crucial for many reasons; first of all, because, as highlighted, coalfields were likely to be mined in densely populated areas where labor force was abundant; this, however, should be mitigated by the introduction of the initial density of population as a control in the main specification. Second, because density is a key indicator of economic development which enables to compare departments that are similar in their economic development and that only differ according to their exposure to coal industry. As a robustness check, I show in the Appendices that the results still hold when substituting the time varying population density for the initial density.

The coefficient of interest, namely β, can be causally interpreted under the identifying assumption that the exposure to coal is orthogonal to the error term in equation 6 conditional on geographical and geological characteristics of each department and year fixed effects. Formally, this identifying assumption writes as:

$$\forall(d,t), \mathbb{E}[\varepsilon_{d,t} \times e_d | \alpha_t, \Gamma_d] = 0 \tag{7}$$

This identifying assumption states that, in the absence of coal extraction, departments would have followed similar innovation, education and industry patterns in the long

run conditional on all the observables included as controls in Equation 6. In order to support this assumption, I implement a twofold strategy: first of all, and as discussed previously, I allow for a rich set of initial geographical and geological characteristics to demonstrate that the effect of coal is not confounded by other parameters which could have affected simultaneously coal mining and patenting activity. Second, and most importantly, I leverage an Instrumental Variable strategy in the spirit of Esposito and Abramson (2021) to further prove that the estimates can be causally interpreted.

4.2 Instrumental variable strategy

As suggested, the exogeneity of coal mining with respects to department specifics is debatable. On the one hand, location of coal deposits could be endogenous to geological and geographical characteristics related to land suitability or to the accessibility of underground resources; if such characteristics were to influence the localization of economic activities and their intensity, my estimates would identify the impact of geographical and geological conditions on patenting activity rather than the sheer effect of exposure to coal industry. On the other hand, mining a coalfield requires technical skills as well as prospects of sufficiently high profits compared to other economic activities. Again, land suitability and weather conditions might have impacted the propensity to mine coal to the extent that variation in climatic outcomes may affect the profitability of the farming sector and, in turn, the likelihood to invest in industrial activities.

Therefore, the literature related to the impact of coal resources on long-run economic and political outcomes has often exploited a large set of instrumental variables. In particular, Esposito and Abramson (2021) make use of the $1:5$ Million International Geological Map of Europe and Adjacent Areas from Asch (2005) which provides a very detailed and fine-grained map of European soils geological content. More specifically, this map pictures the type of strata found at the surface of earth in each European country at a very detailed level. In particular, it exploits the location of carboniferous rock strata: coal is indeed a sedimentary rock that was deposited over 300 million years ago over a period named the Carboniferous Era, a geologic era that spans from the end of the Devonian Period (about 350 million years ago) to the beginning of the Permian Period (about 270 millions years ago). Because of their age, most of those deposits are found deep underground, but earth movements or soil erosion might have brought them close to the surface. In turn, it is more likely that regions where such movements have brought carbon strata to the surface find and mine coal. Hence, and following Esposito and Abramson (2021), I use the availability of carboniferous

surface geology in each department to predict the historical presence of coalfields in France throughout the period under study.

TABLE 1: Proximity to coal and distance to carbon rock strata

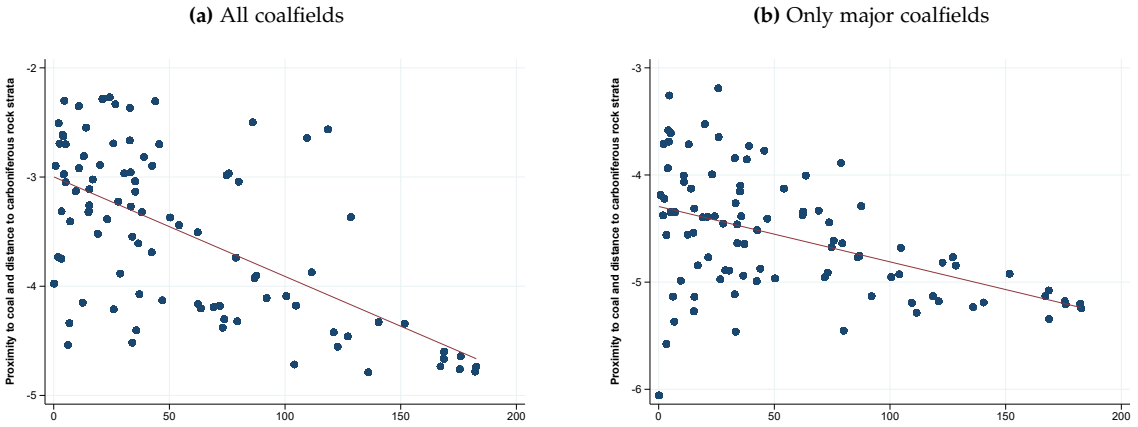
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Distance to carboniferous rock strata	-0.008***	-0.009***	-0.008***	-0.008***	-0.007***	-0.007***	-0.008***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Coal deposit		0.732***	0.548***	0.439***	0.459***	0.453***	0.340*
		(0.159)	(0.153)	(0.182)	(0.172)	(0.181)	(0.200)
Potato yield			-0.112	0.036	0.074	-0.064	0.064
			(0.163)	(0.190)	(0.185)	(0.188)	(0.165)
Wheat yield			-0.686	-3.129	-0.741	-0.040	-0.275
			(2.244)	(2.889)	(2.685)	(2.724)	(2.921)
Oat yield			0.474	2.597	0.368	0.222	0.175
			(2.122)	(2.665)	(2.538)	(2.562)	(2.786)
Proximity to Mediterranean sea				-0.117	0.205	0.118	0.489 **
				(0.137)	(0.174)	(0.175)	(0.237)
Proximity to Atlantic ocean				-0.195**	0.394**	0.439**	0.552***
				(0.082)	(0.178)	(0.178)	(0.159)
Median temperature					-0.626	-0.544	0.107
					(0.716)	(0.714)	(0.681)
Variation temperature					4.882***	5.798	6.899***
					(1.249)	(1.387)	(1.995)
Ruggedness of land						0.214	0.339**
						(0.142)	(0.155)
Land surface						-1.073*	-0.392
						(0.631)	(0.731)
Longitude							-0.037
							(0.097)
Latitude							0.197*
							(0.104)
(Log) Initial population							0.187*
							(0.187)
Observations	95	95	95	95	95	95	95
R-squared	0.371	0.472	0.520	0.563	0.649	0.672	0.704
F-Stat	83.44	104.81	43.92	23.50	33.69	36.88	42.45

Notes: The dependent variable is the proximity to coal measure as discussed in 3 while the explanatory variables are mainly the distance to the closest carbon rock strata, the existence of coal deposit within the department and alternative geological and geographic potential confounders. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

In concrete terms, I compute for each department the average distance of each city to the closest carbon sedimentary rock strata. The validity and strength of this instrument is clearly supported by Table 1 which summarizes the estimates of the coefficients attached to each regressor included in the first stage regression. Strikingly, the distance to carboniferous rock strata exerts a strong and negative impact over the proximity to coalfields as demonstrated by the highly significant coefficient which is very stable across the 7 different specifications. On the contrary, the coefficient on the binary variable capturing the location of a coal deposit is highly positive and significant at the 1% level in the first specification, but the value of the coefficient suddenly drops as I control for the initial population, suggesting that major coal deposits were not totally random to the initial settlement of the population. Additionally, the goodness of fit of the first model which exhibits a R-Squared of 37%, and which only incorporates the distance to carboniferous rock strata, is illustrative of the validity and

strength of my main instrument; adding the binary instrument, which proves slightly less significant in the seventh model, increases the R-Squared by 10% to reach 47.2%. Other geographical and geological controls do not significantly affect the goodness of fit of the model; actually, among the different covariates added as further controls, only 3 of them are significant at the 1% level in the extended specification, namely the proximity to the Atlantic ocean, the variation in historical temperatures and the (log of) the initial population.

FIGURE 5: Proximity to coal and distance to carboniferous rock strata



Notes: Figure (a) plots the coal proximity measure when taking all French coalfields operated by 1900 against the distance to carboniferous rock strata while Figure (b) plots the coal proximity measure when restricting the set of coalfields to historical major coalfields documented in the ERIH project against the distance to carboniferous rock strata.

The same analysis is performed for other measures of the exposure to coal, the results of which are displayed in Appendix C for the sake of concision. In any case, the coefficient on the distance to carboniferous rock strata is significant irrespectively of the numerous geographical and geological controls as potential confounding factors. This further supports the intuition that distance to carbon is a valid instrument as well as a strong predictor of coalfield extraction throughout the period under study. This argument is finally supported by Figure 5 which plots the two proximity measures of coal exposure discussed in Section 3 against the distance to carbon. The negative relation between proximity to coalfields, whether they are major and historical coalfields or only belonging to the overall set of coalfields run at the beginning of the XXth century, is clear and significant, further asserting the validity of my strategy to mitigate concerns related to potential endogeneity.

4.3 Coal mining and patenting activity

In this subsection, I briefly summarize the main results regarding the impact of coal exposure on patenting activity. The model elaborated in Section 2 indeed predicts that

abundant natural resources might induce an over-specialization in resource-intensive industries, which, in turn, could affect productivity growth. In the three following subsections, I back empirically these predictions by showing that coal exposure did not only affect patent quantity, but it also impacted patenting quality and specialization.

4.3.1 Coal mining and patenting quantity

I summarize in this subsection one of the key results of this paper, notably the negative and significant impact of exposure to coal industry on patenting activity throughout the period under study, meaning 1907 to 2016. More specifically, I report in Table 2 the main coefficients of interest when running Equation 6 for the most important specifications, namely the Pseudo Poisson count model, the OLS estimates and the IV estimates exploiting the instruments discussed in the previous subsection. While the overall regression model includes a dozen control variables, I only report here the estimates obtained for the most meaningful ones, namely coal exposure, distance to Paris, initial population density, longitude and latitude. These regressors are indeed the most likely to influence patenting activity and hence to be potential confounders to the analysis; numerous controls already detailed in Section 3 are however included under the Geographical controls and Geological controls parameters, including land suitability, distance to the sea, average altitude of the department and ruggedness of soils. It is also noteworthy that all results are reported for the two main measures of coal exposure, namely proximity to all coalfields on the one hand (henceforth measure 1), and proximity to the major coalfields on the other hand (henceforth measure 2). At times, I also present results using the semi-discrete measures plotted in Figure 4, but, most of the time, I postpone the summary of these results to Appendix D.

The evidence summarized in Table 2 clearly demonstrates the sustained impact of exposure to coal industry on patenting activity. In spite of many controls for either geographical, demographic or geological confounders, the coefficients on the two main measures of coal proximity are negative and significant at the 5% level and even at the 1 % level for most specifications. In terms of magnitude, they imply that a one standard deviation increase in coal exposure translates in a contraction of patenting activity ranging from -29.2% to 42% depending on the specification. Strikingly, the IV estimates are two to three times larger compared to the OLS estimates, which might be accounted for by different reasons. Among others, it should be noted that my coal proximity measure relies on coalfields that were actively mined at the end of the XIXth century, while it excluded those which only started to be run after 1900 or those that were already closed as of 1900. In that respect, my measure of coal expo-

TABLE 2: Effect of coal on patent quantity for various specifications

	Pseudo Poisson		**OLS**		**IV**	
	Measure 1	Measure 2	Measure 1	Measure 2	Measure 1	Measure 2
Coal exposure	-0.351**	-0.432***	-0.180**	-0.248***	-0.544***	-0.556***
	(0.148)	(0.134)	(0.079)	(0.086)	(0.166)	(0.168)
Initial density	1.075***	1.105***	0.519**	0.511***	0.695***	0.609***
	(0.119)	(0.120)	(0.147)	(0.157)	(0.191)	(0.198)
Distance to Paris	-0.006***	-0.006***	-0.004***	-0.004***	-0.002**	-0.003***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Longitude	-0.099	-0.069	0.173	0.167**	0.139	0.139*
	(0.169)	(0.144)	(0.090)	(0.076)	(0.117)	(0.076)
Latitude	0.174	0.195	-0.283**	-0.253**	-0.169	-0.147
	(0.267)	(0.230)	(0.0.136)	(0.116)	(0.183)	(0.123)
Observations	10,900	10,900	10,751	10,751	10,751	10,751
Year Fixed-effect	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Geological controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.850	0.852	0.614	0.615	0.351	0.387
F-Stat			13.31	14.14	17.04	14.87

Notes: The dependent variable is either patent count (Pseudo Poisson specification) or the (log) number of patents per capita (OLS and IV models) while the explanatory variables are mainly the coal exposure measure, distance to Paris, initial population density, longitude, latitude and other geographical and geological controls discussed in Section 3. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

sure neither captures the additional negative effect of coalfields that had not yet been discovered and exploited nor does it measure the negative effect related to formerly opened coalfields. Additionally, it is likely that locations which decided to mine coal, conditional on finding natural resources, were those with higher abilities; in technical terms, this means that the error term in equation 6 is likely to be positively correlated with the coal exposure measure which, in turn, might bias upwards the OLS estimates. Reassuringly however, the estimates for both measures of exposure to coal industry deliver very similar estimates irrespectively of the specification considered. The second measure provides slightly higher point estimates which is consistent with the fact that this measure only incorporates the largest coalfields which, by definition, must have exerted a larger negative impact than the average ones.

4.3.2 Coal mining and patenting quality

A recent strand in the literature has emphasized the importance of innovation quality rather than focusing on sheer quantity. The number of patents might actually not be indicative of the quality of such patents; it could be that some departments find a lot of inventions of poor significance while others restrict to game-changer inventions which bring a higher contribution to economic performance. In that case, my

TABLE 3: Effect of coal on citation weighted patents for various specifications

	Pseudo Poisson		**OLS**		**IV**	
	Measure 1	Measure 2	Measure 1	Measure 2	Measure 1	Measure 2
Coal exposure	-0.371**	-0.456***	-0.181***	-0.245***	-0.558***	-0.561***
	(0.167)	(0.152)	(0.079)	(0.089)	(0.174)	(0.177)
Initial density	0.983***	1.019***	0.449**	0.442**	0.615***	0.537**
	(0.127)	(0.121)	(0.144)	(0.193)	(0.213)	(0.229)
Distance to Paris	-0.007***	-0.007***	-0.004***	-0.004***	-0.002***	-0.003***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Longitude	-0.169	-0.130	0.125	0.123	0.078	0.093
	(0.178)	(0.154)	(0.087)	(0.077)	(0.114)	(0.080)
Latitude	0.146	0.156	-0.289**	-0.162*	-0.156	-0.147
	(0.287)	(0.255)	(0.141)	(0.131)	(0.181)	(0.135)
Observations	10,528	10,528	8,923	8,923	8,923	8,923
Year Fixed-effect	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Geological controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.805	0.807	0.568	0.453	0.194	0.219
F-Stat			8.24	7.85	9.70	9.39

Notes: The dependent variable is either patent count (Pseudo Poisson specification) weighted by the number of citations or the (log) number of patents weighted by the number of citations and divided by department population (OLS and IV models) while the explanatory variables are mainly the coal exposure measure, distance to Paris, initial population density, longitude, latitude and other geographical and geological controls discussed in Section 3. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

estimates could only reflect a quantity effect rather than showing a causal impact of coal exposure on the quality and productivity associated with inventions. Most of the time, quality is approximated by weighting patents with their number of citations to reflect the contribution of each patent to the related innovation literature. The underlying assumption behind this proxy is that the higher the number of citations received by a patent, the larger its contribution to the scientific corpus; conversely, receiving no citation might be indicative of a poor value with respects to scientific standards and achievements. Building on the existing literature, I thus transform the count variable of patents by their counterpart weighted with the number of citations received throughout time and check whether the estimates are still significant. The results are summarized in Table 3 which presents the coefficients on the main regressors when the dependent variable is either patent count weighted by the number of citations or the (log) number of patents weighted by citations and divided by the department population. Additional robustness checks are shown in Appendix D, especially to present the results of the same regression when using two alternative measures of coal exposure as explanatory variables.

Two lessons may be drawn from this table: first of all, coal exposure still has a negative and significant impact on patenting activity even when controlling for the number of

citations received by each patent. Apart from the Pseudo Poisson specification using the restrictive measure of coal proximity, where the coefficient on coal exposure is only significant at the 5% level, all coefficients on coal exposure are again significant at the 1% level, suggesting that not only patenting quantity, but also patenting significance and hence patenting quality were affected by coal exposure. Additionally, it is noteworthy that the estimates are slightly, though not significantly, larger than the estimates derived in the main regressions summarized in Table 2. Overall, the model presented in Table 3 suggests that a one standard deviation increase in coal exposure decreased quality-weighted innovation by c. 30% to 47% depending on the different specifications while the effect ranged from 29 to 42% in the general framework.

And yet, weighting patent by citations might not be fully indicative of the quality of each patent for different reasons. First of all, patent citations were much less common in the first half of the XXth century than in the second half of it. Second, some inventions that later on contributed significantly to economic progress and science as a whole only received a tiny number of citations: let's take the example of patent 174,465 issued by the United States Patent Trade Office (USPTO) to Graham Bell for the telephone in February 1876 and which totaled only 10 citations until March 1876 in spite of its large relevance for communication technologies. In order to correct for the potential flaws of citations-weighted patents measures, Kelly et al. (2021) elaborated a new measure of innovation quality over time by defining breakthrough inventions as innovations that are both novel and meaningful in the sense that they both differ from previous inventions and are useful to later research. Leveraging high-dimensional textual data from patent documents, they identify significant patents based on textual similarity of a given patent to previous and subsequent work. In other words, such patents are distinct from previous work and hence novel, but are related to subsequent innovations and hence meaningful. Given that their method is very computationally expensive, and that I only have access to patents titles - and not abstracts - in France before the early 80s, I exploit their dataset and collapse the number of breakthrough patents at the year $\times$ 3-digit IPC code; I then define for each year as a breakthrough technology the ten 3-digit IPC classes with the largest share of breakthrough patents. Finally, I aggregate the number of breakthrough inventions at the department $\times$ year level and run Equation 6 by substituting breakthrough inventions for standard patenting activity. Results of this exercise are summarized in Table 4 for the six usual specifications.

Again, and in line with the results discussed above regarding citations-weighted patents, two features clearly stand out. First, the estimates on coal exposure are still negative and significant, most of the time at the 1% level, further supporting the idea

TABLE 4: Effect of coal on breakthrough patents for various specifications

	Pseudo Poisson		OLS		IV	
	Measure 1	Measure 2	Measure 1	Measure 2	Measure 1	Measure 2
Coal exposure	-0.362**	-0.428***	-0.215***	-0.262***	-0.632***	-0.632***
	(0.158)	(0.153)	(0.078)	(0.092)	(0.172)	(0.175)
Initial density	0.996***	1.019***	0.552***	0.537***	0.746***	0.656***
	(0.119)	(0.121)	(0.161)	(0.170)	(0.195)	(0.211)
Distance to Paris	-0.006***	-0.007***	-0.004***	-0.004***	-0.003***	-0.003***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Longitude	-0.183	-0.147	0.210*	0.167**	0.165	0.177*
	(0.207)	(0.183)	(0.107)	(0.076)	(0.134)	(0.091)
Latitude	0.299	0.308	-0.337**	-0.316**	-0.199	-0.186
	(0.306)	(0.272)	(0.157)	(0.140)	(0.203)	(0.141)
Observations	9,678	9,678	9,022	9,022	9,022	9,022
Year Fixed-effect	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Geological controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.832	0.833	0.450	0.453	0.269	0.387
F-Stat			8.41	7.85	11.05	14.87

Notes: The dependent variable is either breakthrough patents count (Pseudo Poisson specification) or the (log) number of breakthrough patents per capita (OLS and IV models) while the explanatory variables are mainly the coal exposure measure, distance to Paris, initial population density, longitude, latitude and other geographical and geological controls discussed in Section 3. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

that innovation quality was negatively affected by coal exposure. Second, it is telling that the estimates are all larger than those derived in the general framework: while general estimates suggested that a one standard deviation increase in coal exposure reduced patenting activity by c. 29 to 42%, the estimates from Table 4 indicate that the magnitude would range from 30 to 47% which pertains to a sizable effect. On average, coefficients are about 10% larger when accounting for difference in inventions quality than in the general framework; all in all, there is robust evidence that coal exposure simultaneously reduced innovation quantity and quality.

4.3.3 Coal mining and patent specialization

Beyond its impact on innovation quantity and quality, coal exposure was also conducive of a specific specialization pattern in terms of innovative activities which might itself have affected the quality and quantity of patenting activity. This assertion is supported by Table 5 which summarizes the effect of coal exposure on a wide range of technologies. More specifically, textile patents encompass innovation which technology class begins with D0; mining patenting activity corresponds to patents belonging to IPC class E21; land vehicles refer to patents in IPC class B62; plane & space are patents in IPC class B64 while computing and information patents capture respec-

tively IPC class G06 and G16. Strikingly, the effect of coal exposure was heterogenous across technology classes. While I only report the estimates for 6 IPC classes out of concision, the overall pattern remains unchanged across specifications: the closer a technology class is to the industry specialization favored by coal proximity, the less negative the impact of coal exposure on this technology class. Conversely, the further an industry is from heavy industries related to early industrialization, the higher and the more negative the impact on long-run innovation activities.

TABLE 5: Effect of coal on innovation specialization

	Textile	Mining	Vehicles	Aeronautics	Computing	Information
Coal exposure	0.205	0.049	-0.220	-1.571***	-0.671**	-0.723**
	(0.136)	(0.239)	(0.243)	(0.318)	(0.319)	(0.303)
Observations	10,810	10,600	10,810	10,810	10,716	10,810
Year Fixed-effect	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Geological controls	Yes	Yes	Yes	Yes	Yes	Yes
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.673	0.550	0.660	0.709	0.759	0.745

Notes: The dependent variable is the number of patents related to each category mentioned while the explanatory variable is the proximity to coal according to the first measure discussed in Section 3; geographical and geological controls pertain to the covariates discussed in Section 3. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Typically, textile innovation does not seem to have been negatively impacted by exposure to coal. On the contrary, coal proximity seems to have positively affected textile related innovation though the coefficient is not significant due to statistical imprecision. Similarly ,the effect of coal exposure on mining techniques was rather positive in spite of a poor significance. The effect only starts to be negative for car-related inventions, but, again, the coefficient is not statistically significant, even at the 10% level. This trend is consistent with the sustained involvement of coal-exposed departments to mining and textile industries which were typical of the early industrialization that such departments experimented thanks to their proximity to natural resources. By contrast, patents related to newer sectors that bear no relation to coal extraction and the type of industries it may support have been negatively and significantly impacted by coal exposure. For instance, a one standard deviation increase in coal exposure corresponds to a 78% decrease in patenting activity related to plane industry, an effect which is more than twice as large as the overall effect derived for total patenting activity. While the estimates are not as large for computing and information technologies, they still suggest that the effect of coal exposure on them was much stronger than for overall patenting activity: a one standard deviation increase in coal exposure reduced patenting activity of technology classes G06 and G16 by respectively 49 and 51 %, two

estimates that are larger to the upper bound effect I find for total patenting activity.

To summarize, my main findings so far show that exposure to coal industry exerted a negative and significant impact on patenting activity throughout the XXth century [2]. Additionally, coal exposure negatively impacted the quality of innovations as measured either by a citation-weighted count of patents or by the more robust breakthrough component of innovation discussed by Kelly et al. (2021). This by-side finding suggests that, not only innovation quantity, but also innovation quality was affected by exposure to coalfields. Finally, and related to this second point, technology classes were heterogeneously affected by coal exposure: while technology class closely intertwined with the first industrial revolution, and hence with coal availability, were rather positively though not significantly affected, industries associated with the post-war period, and more precisely to the third industrial revolution, such as airplane industry, communication technologies, and computing power, were negatively and much strongly affected by coal exposure than standard patents.

4.4 Coal mining and industry specialization

The key force underpinning the results derived in the model exposed in Section 2 is that the price of natural resources directly influences the profitability of each industry, implying that a large and positive shock on the availability of natural resources, such as the discovery of a new coalfield, will likely trigger an increase in the share of the resource-intensive industry. Under the assumption that the former might be less efficient at the innovation process than other industries, such a behavior might lead to an inefficiently low level of innovation, a prediction that turned out to be confirmed by empirical facts summarized in the previous subsection. I now check whether exposure to coal-industry actually induced a lasting specialization in resource-intensive industries by contrast with knowledge-intensive ones. As I unfortunately lack a continuous and homogeneous dataset for the whole XXth century, I first investigate the impact of coal exposure on industry specialization in the early XXth century before turning to its sustained impact in the late XXth century and the beginning of the XXIst century.

TABLE 6: Effect of coal on share of *professions libérales*, 1901-1936

	Measure 1		**Measure 2**		**Measure 3**	
	OLS	IV	OLS	IV	OLS	IV
Coal exposure	-0.086**	-0.236***	-0.089	-0.230***	-0.105***	-0.181***
	(0.039)	(0.071)	(0.055)	(0.071)	(0.033)	(0.053)
Initial density	0.015***	0.022***	0.014***	0.018***	0.014***	0.016***
	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)
Distance to Paris	-0.001	-0.001	-0.001	-0.000	-0.001	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Longitude	-0.003	-0.035	0.012	0.007	-0.000	-0.012
	(0.063)	(0.073)	(0.059)	(0.063)	(0.054)	(0.052)
Latitude	0.057	0.078	0.060	0.082	0.104	0.147
	(0.095)	(0.100)	(0.096)	(0.101)	(0.091)	(0.092)
Observations	581	581	581	581	581	581
Year Fixed-effect	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Geological controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.334	0.202	0.329	0.221	0.346	0.267
F-Stat	5.36	8.33	5.63	7.07	6.14	6.22

Notes: The dependent variable is (log) share of people working as *profession libérale* while the main explanatory variable is the different measures of coal exposure, as discussed in Section 3. Other important covariates shown here include distance to Paris, initial population density, longitude, latitude. Other geographical and geological controls discussed in Section 3 are included. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

4.4.1 Coal mining and early XXth century industry specialization

Table 6 summarizes the effect of coal exposure on the share of high-value added jobs - as measured by *professions libérales* - in total population in the 1901-1936 period. In concrete terms, *professions libérales* is an occupation category defined by the *Annuaires Statistiques de la France* as encompassing doctors, teachers, engineers or artists. While these jobs encompass very different tasks and, somehow, education levels, they most of the time demand scientific or precise knowledge that, in turn, might contribute to innovation activities. In that respect, a high share of such occupations in the population is indicative of a superior ability to deliver new ideas and hence, patents. And it turns out that both OLS and IV estimates suggest a sizable and consistent effect across the three measures of coal exposure, indicating that coal industry adversely impacted the specialization of coal-dependent departments. While the OLS estimate for measure 2 is close to, but not significant at the 10% level, all other estimates, either

[2]Complementary checks performed in the Appendices provide support to the argument that the effect started as soon as the onset of the XXth century, precluding the explanation of a sheer reversal of fortune.

in the OLS or in the IV strategy are indeed negative and significant, either at the 5% or, most of the time, at the 1% level. In terms of magnitude, estimated coefficients imply that a one standard deviation increase in coal exposure contracted the share of people working as *professions libérales* from 8 to 21%.

These estimates are reinforced by Table 7 which depicts the effect of coal exposure on the (log) share of people classified as "technical experts" by the French census operated in 1921, 1926, 1931 and 1936. Unfortunately, my dataset for this dependent variable is restricted to those four censuses, given that the number of experts at the department level is not consistently reported for previous years. Allowing for both a wide range of confounding factors, and different empirical specifications, I however find a significant and negative effect of coal exposure on the share of expert jobs in total active population. This finding strengthens the previous statement made by Table 6 in so far as expert jobs is probably a better definition of occupations that were the most likely to capture the share of engineers or other science-based occupations in total population. In that respect, a lower presence of such jobs in total population might be indicative of lower abilities to contribute to scientific and research effort. Actually, coal exposure has substantially reduced the share of people working as experts: estimates are indeed very stable across the different measures of coal exposure, and point to a -14 to -35% reduction in the share of expert jobs following a one standard deviation increase in coal exposure. That the estimates are significantly larger than those derived for all *professions libérales* is no surprise: this second measure of high value added jobs better captures jobs that are the most likely to contribute to research efforts; given that some occupations among *professions libérales* were not necessarily related to science-based activities - such as, for instance, nurses - they may not have been affected by the lower investment in upper-tail education induced by coal exposure.

4.4.2 Coal mining and late XXth century industry specialization

As will also be shown for the upper-tail education channel, it is noteworthy and telling that such a negative effect of coal exposure on job specialization was sustained throughout the XXth century. Again, I unfortunately lack precise data for job classification in the 1936-1980 period; however, I am able to build a detailed dataset at the department level for the 40-year period between 1980 and 2020 which is informative of the share of each 4-digit NAF activity in total employment. More specifically, this dataset builds on the *Enquête emploi en continu*, A French administrative survey originally aimed at better understanding unemployment, and which later on started investigating the shape of employment at a local level. While these year-on-year sur-

TABLE 7: Effect of coal on share of expert jobs, 1921-1936

	Measure 1		Measure 2		Measure 3	
	OLS	IV	OLS	IV	OLS	IV
Coal exposure	-0.224***	-0.439***	-0.213***	-0.448***	-0.154***	-0.354***
	(0.049)	(0.113)	(0.067)	(0.117)	(0.051)	(0.091)
Initial density	0.594***	0.549***	0.637***	0.018***	0.506***	0.552***
	(0.116)	(0.125)	(0.117)	(0.159)	(0.114)	(0.143)
Distance to Paris	-0.001	-0.001	-0.001	-0.000	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Longitude	0.055	-0.005	0.101	0.081	0.098	0.071
	(0.085)	(0.098)	(0.090)	(0.102)	(0.081)	(0.078)
Latitude	0.144	0.216	0.135	0.209	0.146	0.247**
	(0.117)	(0.144)	(0.125)	(0.154)	(0.109)	(0.116)
Observations	356	356	356	356	356	356
Year Fixed-effect	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Geological controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.717	0.624	0.696	0.611	0.672	0.600
F-Stat	21.88	15.71	22.27	18.75	23.13	18.97

Notes: The dependent variable is (log) share of people working as experts while the main explanatory variable is the different measures of coal exposure, as discussed in Section 3. Other important covariates shown here include distance to Paris, initial population density, longitude, latitude. Other geographical and geological controls discussed in Section 3 are included. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

veys were not exhaustive, they nevertheless rely on samples of roughly 35,000 individuals allowing for both individual-level analysis and accurate aggregation at the department level. Building on these valuable insights, I show a strong persistence in the effect of coal on industry specialization at the local level: namely, I document that high-value added jobs such as finance occupations, scientific and technical advisory or occupations with advanced technological content were adversely impacted by coal exposure while low-skilled and labor intensive occupations such as iron industries were positively and significantly impacted. For concision reasons, I only report the estimates for two measures of coal exposure, namely proximity to coalfields and proximity to large coalfields as discussed in Section 3 and postpone the presentation of results for the third coal exposure measure to Appendix D. Also, and for the same reasons, I only show the point estimates for the coal exposure measure; derivation of the coefficients for initial density, distance to Paris, longitude and latitude are shown in the Appendix D as well.

The effect of coal exposure on late XXth century industry specialization is summarized in Table 8 which confirms that the effect of coal exposure on industry specialization was sustained over the long run. More specifically, coal exposure induced a

TABLE 8: Effect of coal exposure on post-war industry specialization, 1980-2020

	Metallurgy		Technical advisory		Finance		Technology	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Coal exposure	0.252*	0.455**	-0.102*	-0.418***	-0.073*	-0.328***	-0.099	-0.494***
	(0.148)	(0.200)	(0.056)	(0.113)	(0.039)	(0.078)	(0.065)	(0.131)
Observations	1,596	1,596	3,667	3,667	3,679	3,679	3,677	3,677
Year Fixed-effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geological controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.407	0.347	0.325	0.143	0.221	0.128	0.295	0.104
F-Stat	7.71	8.84	9.05	8.74	10.06	6.55	8.46	6.84

Notes: The dependent variable is (log) share of people working as experts while the main explanatory variable is the different measures of coal exposure, as discussed in Section 3. Other important covariates shown here include distance to Paris, initial population density, longitude, latitude. Other geographical and geological controls discussed in Section 3 are included. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

lock-in effect in low-value added specialization such as the metallurgy sector, while lastingly affecting the involvement in high-value added sectors such as technical advisory, finance industry or technology. In terms of magnitude, a one standard deviation increase in coal exposure is conducive of a 29 to 58% higher share of textile industry. By contrast, a one standard deviation increase in coal exposure lowers the share of technical advisory, finance industry and technology by respectively 52, 39 and 64%. While these magnitudes are substantial, it should be recall that they apply on industry shares, meaning that they indicate percentage point increases. All in all, the evidence brought in this subsection lends credence to the argument that exposure to coal extraction exerted a sustained and negative impact on industry specialization, in the sense that high-value added occupations were significantly curtailed while low-skilled intensive industries remained active up until the beginning of the XXIst century.

4.5 Coal mining and upper-tail education

A by-product of the model elaborated in Section 2 is that education is inefficiently low in places endowed with abundant natural resources, given that firms have an incentive to invest in resource-intensive industries which require poor education levels. I now investigate empirically this prediction, by looking at both early XXth education and late XXth education. The impact of either natural resources or early industrialization on education is indeed a contentious issue. Franck and Galor (2017) notably argue that early industrialization was conducive of larger literacy rates, flying in the face of the conventional wisdom that industry is rather associated with poorer human capital. And yet, their study mostly focuses on intermediate education, meaning average

levels of knowledge rather than the upper-tail part of it which might be the most important one to proxy for invention abilities. Indeed, Squicciarini (2020) demonstrates the differential impact of secular schooling with respects to catholic schooling in terms of contribution to the industrialization process. The former actually supported significantly the involvement in new industries while the latter barely exerted any effect. This finding, she argues, mostly stems from the different programs that were adopted by those distinct schooling systems: while secular schools rather focused on technical and scientific knowledge, catholic schools were reluctant to teach science-based content. Furthermore, Franck and Galor (2017) study early industrialization rather than exposure to coal industry, which might lead to different conclusions.

4.5.1 Coal mining and early XXth century education

TABLE 9: Effect of coal on upper-tail education, 1885-1911

	(Log) surface of coalfields		**Proximity to coalfields**		**Proximity to large coalfields**	
	OLS	IV	OLS	IV	OLS	IV
Coal exposure	-0.120**	-0.491**	-0.124**	-0.371***	-0.063	-0.208***
	(0.053)	(0.206)	(0.071)	(0.128)	(0.042)	(0.065)
Initial density	-0.009	-0.018	-0.006	-0.007	-0.000	0.014
	(0.019)	(0.018)	(0.021)	(0.024)	(0.022)	(0.026)
Distance to Paris	-0.001	-0.001	-0.001	-0.000	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Longitude	0.116	0.089	0.116	0.098	0.123*	0.119*
	(0.073)	(0.086)	(0.070)	(0.065)	(0.073)	(0.072)
Latitude	-0.126	-0.045	-0.131	-0.087	-0.127	-0.067
	(0.119)	(0.139)	(0.123)	(0.126)	(0.120)	(0.118)
Observations	1,789	1,789	1,789	1,789	1,789	1,789
Year Fixed-effect	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Geological controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.442	0.06	0.438	0.170	0.424	0.167
F-Stat	4.88	3.98	5.25	5.94	4.52	4.31

Notes: The dependent variable is the (log) ratio of soldiers who held the baccalauréat degree over the 1885-1911 period while the main explanatory variable is coal exposure, as measured by the three specifications discussed in Section 3. Other important covariates shown here include distance to Paris, initial population density, longitude, latitude. Other geographical and geological controls discussed in Section 3 are included. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

In order to elaborate on the impact of coal mining on upper-tail education in the early XXth century, I hence assemble a dataset at the department level which records the share of young men enrolled in the French army that held the *baccalauréat* degree from 1885 to 1911. I then investigate the impact of coal exposure using the usual measures of exposure to coal industry as well as the controls for geographical, geological and demographic potential confounders. The results of this analysis are conveyed in Table

9 which supports the idea that coal exposure induced a negative and significant effect on the share of young men holding the *baccalauréat* degree. Both OLS and IV estimates suggest for the three measures of education that exposure to coal induced a negative and significant impact on upper-tail education. In terms of magnitude, IV estimates suggest that a one standard deviation increase in coal exposure decreased the upper-education rate by 19% to 39% depending on the estimates. Strikingly, the impact of large coalfields seem a bit lower, given that the OLS estimates is negative but only significant at the 13% level while the IV estimate is negative and significant at the 1% level but substantially smaller than the two other estimates. This might be explained by the fact that large coalfields only encompass 16 major extraction sites while local coalfields were in the proportion of c. 900 at that time. In that respect, proximity to large coalfields is a much more stringent measure of coal exposure which might leave some departments with low values of coal exposure despite a potentially high level exposure to smaller coalfields. In any case, that the results are robust to both IV strategy and the inclusion of many confounding parameters is suggestive of the large impact coal extraction exerted on upper-tail education in the early XXth century.

4.5.2 Coal mining and education in the late XXth century

This negative impact was actually sustained throughout the XXth century. I unfortunately lack data for the 1914-1968 period in terms of education data, so that my later estimates only encompass the second half of the XXth century as well as the beginning of the XXIst century. Thanks to the richness of the INSEE databases corresponding to that period, I however leverage granular data, in fact at the city level, to document the persistent negative impact of coal exposure on upper-tail education. More specifically, I retrieve for each city the share of the population having no diploma, having the baccalaureate degree, or having a higher education degree. While the results are robust for the whole sample of French cities, I rather present the main estimates for cities with an initial population greater or equal to 1,000 inhabitants out of fear that my results may be driven by smaller cities where access to high schools is much more reduced. Robustness checks for the whole sample of French cities are presented in Appendices D. In the end, my sample is made of roughly 7,000 cities for which I have seven points in time, spanning from 1968 to 2020. In terms of descriptive statistics, the share of population aged 16 or more that held at most the baccalauréat degree raised sharply from 2.7% in 1968 to 16.8% in 2016; symmetrically, the share of population aged 16 or more that held a higher education degree improved at an even quicker pace from 1% in 1968 to 23.5% by the end of 2016.

Table 10 summarizes the effect of coal exposure on higher education in France from

TABLE 10: Effect of coal on lower and upper-tail education, 1968-2016

	(Log) No diploma		(Log) Baccalauréat degree		(Log) Higher education	
	OLS	IV	OLS	IV	OLS	IV
Coal exposure	0.012***	0.036*	-0.029***	-0.107***	-0.059***	-0.214***
	(0.004)	(0.017)	(0.009)	(0.040)	(0.016)	(0.075)
Initial density	-0.023***	-0.026***	0.084***	0.090***	0.148***	0.161***
	(0.004)	(0.005)	(0.007)	(0.008)	(0.014)	(0.015)
Distance to Paris	0.000*	0.000	-0.000	0.000	-0.001**	0.000
	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)
Longitude	-0.007	-0.015*	0.028**	0.043****	0.026	0.059**
	(0.005)	(0.007)	(0.012)	(0.014)	(0.020)	(0.027)
Latitude	0.016**	0.014	-0.047***	-0.030	-0.065**	-0.035
	(0.007)	(0.009)	(0.015)	(0.019)	(0.025)	(0.035)
Observations	49,345	49,345	49,195	49,195	48,393	48,393
Year Fixed-effect	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Geological controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.634	0.612	0.656	0.641	0.725	0.695
F-Stat	8.66	216.96	28.78	215.89	24.81	710.99

Notes: The dependent variable is respectively the (log) ratio of no diploma in total population, the (log) ratio of baccalauréat degree in total population and the (log) ratio of higher education degree in total population while the main explanatory variable is coal exposure, as measured by the proximity measure discussed in Section 3, but at the city level, rather than at the department level. Other important covariates shown here include distance to Paris, initial population density, longitude, latitude. Other geographical and geological controls discussed in Section 3 are included. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

1968 to 2016. Higher education is measured in a threefold way: first, I take the share of population which holds no diploma at all; second, I retain the share of population for which the highest diploma is the baccalauréat exam; finally, I consider the share of population which holds a higher education degree. In that respect, I consider measures which are increasingly restrictive in terms of the measure of upper-tail education. While having no diploma is the norm for most of the French population born before the post-war educational reforms, the baccalauréat exam progressively spread throughout the population; however, it took until recently for the success rates to reach the record levels now observed: in that respect, baccalauréat has long been a measure of upper-tail education in the French population; finally, higher education is the best proxy for upper-tail education and, to that extent, best encompasses the increase in the upper-tail education level in France. Turning to the results, two features clearly stand out.

First of all, coal exposure had a negative and significant impact over upper-tail education for the whole second half of the XXth century. Indeed, both OLS and IV estimates suggest that exposure to coal significantly increased the number of people holding no diploma while reducing sharply the share of people with the baccalauréat degree, or with a higher education degree. In terms of magnitude, the estimates again point to a

sizable effect: a one standard deviation increase in coal proximity increases the share of people having no diploma by 1 to 3.6%; it further reduces the share of people holding at most the baccalauréat degree by 3.5 to 9.5%; finally, and most strikingly, a one standard deviation increase in coal exposure is conducive of a 5.8 to 18.9% decline in higher education attainment. Second, and as highlighted above, the effect is all the larger as the measure of education better captures upper-tail rather than average education. While the effect of coal exposure is modest for the lack of diploma, it is large and substantial for higher education, suggesting that coal industry deprived cities from the opportunity of escaping a poor specialization through upper education.

TABLE 11: Effect of coal on share of *grandes écoles*, 1980-2020

	Proximity to all coalfields		**Prooximity to large fields**		**(Log) surface**	
	OLS	IV	OLS	IV	OLS	IV
Coal exposure	-0.148***	-0.522***	-0.165***	-0.528***	-0.238***	-0.538***
	(0.045)	(0.137)	(0.063)	(0.137)	(0.065)	(0.140)
Initial density	0.372***	0.637***	0.347***	0.523***	0.343***	0.439***
	(0.097)	(0.132)	(0.109)	(0.153)	(0.111)	(0.157)
Distance to Paris	-0.003***	-0.002***	-0.002***	-0.000	-0.002**	-0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Longitude	0.058	-0.044	-0.084	0.052	0.076	0.049
	(0.079)	(0.116)	(0.073)	(0.095)	(0.067)	(0.071)
Latitude	-0.226**	-0.115	-0.221**	-0.111	-0.188*	-0.085
	(0.105)	(0.153)	(0.108)	(0.151)	(0.095)	(0.106)
Observations	3,093	3,093	3,093	3,093	3,093	3,093
Year Fixed-effect	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Geological controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.525	0.204	0.487	0.176	0.501	0.252
F-Stat	10.52	6.27	8.93	7.31	8.54	6.67

Notes: The dependent variable is (log) share of surveyed people having completed their higher education in a *grande école* or in a engineer school while the main explanatory variable is the different measures of coal exposure, as discussed in Section 3. Other important covariates shown here include distance to Paris, initial population density, longitude, latitude. Other geographical and geological controls discussed in Section 3 are included. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

This statement is further supported by an in-depth analysis of education specialization at the department level. Relying on *Enquêtes emploi en continu* for the whole 1980-2020 period, I indeed retrieve for each department the share of people surveyed year on year which achieved their higher education in either *grandes écoles* or an engineer school not belonging to the somehow fuzzy *grandes écoles* label. Again, coal exposure negatively and substantially affected the upper tail of the education distribution. The estimates indeed suggest that a one standard deviation increase in coal exposure reduced the (log) share of people attending a *grande école* by 14 to 42%. Two facts deserve attention: first, the estimates are very stable across the different measures of coal exposure and only vary depending on the specification (either OLS

or IV). Second, the estimates do not contradict the values obtained in the city-level analysis: the former indeed suggested that a one standard deviation increase in coal exposure reduced the (log) share of higher education by 6 to 19%; given that this measure of higher education encompassed all kinds of upper education degree, it is not as restrictive as the *grande école* measure, a difference which might account for the higher magnitude of the coefficients obtained in Table 11. In any case, these estimates further lend credence to the argument that coal exposure had a negative and sustained impact over upper-tail education at the local level in France, a finding that might account for the lasting discrepancies in innovation abilities across French departments.

5 Mechanisms

In this section, I address the main mechanisms underpinning the negative and lasting impact of exposure to coalfields on patenting activity, education and industry specialization. While the standard economic framework would predict that investment in unprofitable industries would vanish as alternative opportunities emerge, externalities as those described in the model summarized in Section 2 might account for a sustained and negative effect of coal resources. More importantly, this section will argue that exposure to coal triggered a negative feedback loop, whereby firm and individual decisions jointly concurred to lower productivity efforts by altering human capital accumulation and by myopically investing in restricted industries.

5.1 Technological lock-in

One reason which could account for the sustained negative impact of coal exposure on productivity would be the existence of a technological lock-in, whereby departments heavily reliant on coal failed to diversify out of specific industries associated to coal mining. This subsection will document evidence on that channel in a three-step way: first, I will bring evidence that coal is an absorbing industry from a technological standpoint, in the sense that coal patentees are more myopic with respects to other technological fields; second, I will shed light on the fact that innovation was poorly diversified in coal basins; finally, I will demonstrate that coal basins tend to be more myopic in the sense of Agrawal et al. (2009), meaning that they disproportionately draw from a restricted set of knowledge upon inventing new ideas.

5.1.1 Is coal a myopic resource?

It is first noteworthy that coal technology is not similar to other scientific fields. Indeed, it turns out that coal-related inventions exhibit a systematic myopic pattern in the sense of Agrawal et al. (2009). Drawing on the example of "company town", the authors define creative myopia as a measure of the degree to which inventors are "nearsighted" in drawing disproportionately from prior inventions that are in some way close to them. They show that such behavior is likely to occur in industries where firm concentration is large. As coal has been historically associated to large industries, it could be that inventions related to coal are more myopic. To test this assumption, I assemble a dataset at the country x IPC3 x year level spanning the whole XXth century and compute technological myopia $M^f_{c,t}$ for each technology field f in country c at time t as follows:

$$M^f_{c,t} = \frac{1}{N^{\mathcal{A}}_{c,t}} \sum_{a \in \mathcal{A}} \frac{\tilde{C^f_{c,t}}}{C^f_{c,t}} \tag{8}$$

Where $M^f_{c,t}$ is the average technological myopia in technological field f in country c at time t, $N^{\mathcal{A}}_{c,t}$ indexes the number of assignees in technological field f in country c at time t, $\mathcal{A}$ is the set of assignees in technological field f in country c at time t, $\tilde{C^f_{c,t}}$ counts the number of citing patents of country c and technological field f at time t for which the technology class of the cited patent is the same as the technology class of the patent that makes the citations, and $C^f_{c,t}$ is the number of citations made by patents of technology field f in country c at time t. Intuitively, the larger $M^f_{c,t}$, the more assignees in a given technological field within a country tend to cite patents that belong to the same technological class as the patent they are making, meaning that inventors tend to focus on a narrow set of technological fields to make progress in their own knowledge effort.

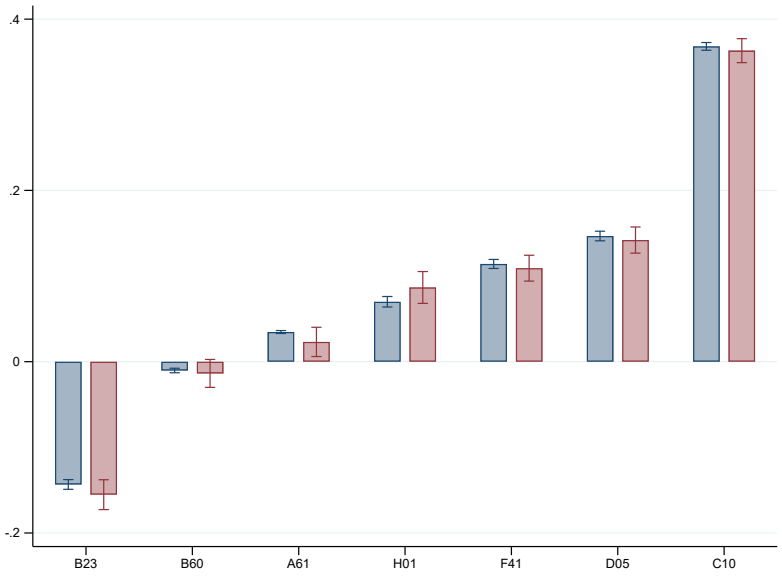
I then run the following regression:

$$M^f_{c,t} = \alpha_t + \eta_c + \sum_{k \in \mathcal{T}} \beta_k \mathbb{1}[k = f] + \varepsilon_{f,c,t} \tag{9}$$

Where α_t indexes year fixed-effects, η_c captures country fixed effects and β_k for $k \in \mathcal{T}$, where $\mathcal{T}$ indexes the set of all technological fields, are the parameters of interest. Namely, they capture the individual effect of each 3-digit IPC class on the level of technological myopia. In other words, they capture fixed effects specific to technology fields once country and year fixed effects are removed. Results from this specification

are shown in Figure 6 which clearly supports the idea that coal-induced inventions exhibit a substantially larger myopia compared to their technological counterparts. For the sake of clarity, I only report here the estimates obtained for 7 technological classes, namely B23 (machine-tools), B60 (transportation), A61 (health and leisure), H01 (electric elements), F41 (weapons and ammunition), D05 (sewing and embroidering) and finally C10 (coal, coke, gas and petroleum). This choice was made on the basis of the representativeness of such technological classes which have been peculiarly significant throughout the XXth century; yet, it is noteworthy that, in any case, coal technology (as captured by IPC3 "C10") exhibits the larger coefficient by far; the second largest technological-field specific coefficient is that of field D05 (sewing and embroidering) which is also reported in this figure.

FIGURE 6: Creative myopia by technological field



Notes: This Figure plots the results from equation 9 and shows the coefficients obtained for 7 IPC classe detailed above. Standard errors are clustered at the 3-digit IPC level. Two models are presented: in blue, equation 9 is tested with country fixed effects and year fixed effects while the red specification allows for country x year fixed effects.

In short, recovered estimates suggest that coal-related inventions tend to be thrice as myopic as sewing-related inventions which, themselves, are the second most myopic inventions according to the measure discussed above. These estimates are the more striking as they are derived using agricultural inventions from IPC class "A01" as the baseline, which itself is a highly myopic technological field. In that respect, it turns out that coal inventions might contribute to short-sighted inventions or, at least, to inventions that build on a narrow subset of past innovations. This, in turn, might explain why former coal-basins experienced a hard time at delivering valuable inventions throughout the XXth century: they might have disproportionately drawn from a restricted set of technological fields, an hypothesis that the following subsections will

reinforce.

5.1.2 Innovation diversity in coal basins

Previous paragraphs suggest that coal industry is myopic in the sense that it disproportionately draws knowledge from a restricted set of technical fields. This, in turn, could induce over-specialization of departments that once invested in the mining industry; in particular, it could be that innovation in coal departments was largely concentrated in a few technical fields. This assumption can be tested by defining technological concentration $C^T_{d,t}$ in department d at time t according to the metric specified by Agrawal et al. (2009):

$$C^T_{d,t} = \left[1 - \sum_{f \in \mathcal{T}} \left(\frac{{P_{f,d,t}}^2}{P_{d,t}}\right)\right] \frac{P_{d,t}}{P_{d,t} - 1} \tag{10}$$

Where $P_{f,d,t}$ counts the number of patents filed in technological field f in department d at time t while $P_{d,t}$ counts the number of patents filed in department d at time t. In words, this measure captures the concentration of patenting activity in department d at time t within a few technological class (first term of the RHS), operating an adjustment depending on the total number of patents within that department in a given year. The larger this ratio, the more diverse innovation is. Conversely, the lower this ratio, the more innovation tends to cluster within a few industries in a given department. To understand whether coal exposure had an impact over the innovation structure and, more specifically, over the degree of innovation concentration within a given department, I run equation 6 by taking $C^T_{d,t}$ as the dependent variable. The results are shown in Table 12 which clearly supports the idea that coal exposure triggered an over-specialization within a few technological fields. On average, a one standard deviation increase in coal exposure decreased technological diversity by 2 to 4% according to the IV estimates. This suggests that coal exposure strongly impacted, not only patenting quantity or even quality, but also patenting structure. Coal basins were more contingent on a few industries; as those industries started to fade in the beginning of the 60s, coal basins were locked-in a tight and poor industry specialization, precluding them from smoothly reconverting in frontier sectors.

5.1.3 Creative myopia in coal basins

Consistently with the evidence accumulated so far, coal basins seem to have adopted a myopic behavior in terms of technological progress throughout the second half of

TABLE 12: Effect of coal on technological diversification

	Proximity to all coalfields		**Proximity to large fields**		**Binary measure**	
	OLS	IV	OLS	IV	OLS	IV
Coal exposure	-0.007*	-0.021***	-0.022***	0.037***	-0.009	-0.047***
	(0.004)	(0.006)	(0.003)	(0.006)	(0.006)	(0.015)
Initial density	0.015	0.022***	0.023***	-0.017	0.015**	0.029**
	(0.006)	(0.008)	(0.006)	(0.008)	(0.006)	(0.012)
Distance to Paris	-0.000	0.000	0.000	0.000	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)
Longitude	0.001	-0.001	-0.000	-0.001	0.000	-0.002
	(0.004)	(0.011)	(0.004)	(0.005)	(0.004)	(0.008)
Latitude	0.002	0.007	0.005	0.007	0.001	0.003
	(0.006)	(0.008)	(0.005)	(0.008)	(0.006)	(0.008)
Observations	10,070	10,070	10,070	10,070	10,070	10,070
Year Fixed-effect	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Geological controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.180	0.048	0.173	0.048	0.169	0.031
F-Stat	8.84	8.89	8.15	8.89	8.52	5.61

Notes: The dependent variable is technological diversification as defined in equation 10 above while the main explanatory variable is the different measures of coal exposure, as discussed in Section 3. Other important covariates shown here include distance to Paris, initial population density, longitude, latitude. Other geographical and geological controls discussed in Section 3 are included. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

the XXth century. More specifically, former coal basins exhibited a systematically higher technological myopia in the sense of Agrawal et al. (2009) which was detailed in subsection 5.1.1. To ensure that the results obtained for coal as a technological class are similar when turning to a geography-based analysis, I compute technological myopia $M^T_{d,t}$ for each department d at time t as follows:

$$M^T_{d,t} = \frac{1}{N^{\mathcal{A}}_{d,t}} \sum_{a \in \mathcal{A}} \frac{\tilde{C^T_{d,t}}}{C^T_{d,t}} \tag{11}$$

Where $N_{d,t}$ indexes the number of inventors in department d at time t, $\mathcal{A}$ indexes the set of inventors in department d at time t, $\tilde{C^T_{d,t}}$ counts the number of citations made by patents filed in department d at time t which refer to patents belonging to the same 3-digit IPC class as the citing patent, and $C^T_{d,t}$ captures the total number of citations made by patents filed in department d at time t. In short, $M^T_{d,t}$ is very similar to the measure at the technology-class level discussed in the previous subsection, apart from it is computed at the department level rather than at the technological class level. I then simply test equation 6 by taking technological myopia $M^T_{d,t}$ as the dependent variable, which yields the results summarized in Table 13. In summary, exposure to coal extraction substantially increased technological myopia; in terms of magnitude,

a one standard deviation in coal exposure allegedly translated in a 1 to 6% increase in technological myopia, suggesting that the effect is sizable. Probably more important is the economic meaning of such a myopia: by building on a narrower set of ideas and knowledge compared to other departments, former coal basins lowered their chance of reaching the technological frontier as well as the likelihood that they could escape their overspecialization in the coal industry in the long run.

TABLE 13: Effect of coal on technological myopia

	Proximity to all coalfields		**(Log) coal surface**		**Binary measure**	
	OLS	IV	OLS	IV	OLS	IV
Coal exposure	0.010*	0.030**	0.007**	0.030**	0.065***	0.064**
	(0.006)	(0.013)	(0.003)	(0.006)	(0.012)	(0.026)
Initial density	-0.020**	-0.037***	-0.015**	-0.029**	-0.051**	-0.035***
	(0.010)	(0.013)	(0.007)	(0.014)	(0.023)	(0.015)
Distance to Paris	-0.000	0.000	-0.000	0.000	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)
Longitude	-0.003	-0.001	-0.001	0.003	0.000	-0.002
	(0.004)	(0.011)	(0.006)	(0.009)	(0.022)	(0.008)
Latitude	-0.005	-0.010	-0.003	-0.009	-0.006	0.003
	(0.006)	(0.013)	(0.009)	(0.014)	(0.031)	(0.008)
Observations	3,131	3,131	3,131	3,131	3,131	3,131
Year Fixed-effect	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Geological controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.173	0.054	0.173	0.051	0.162	0.031
F-Stat	5.43	4.51	5.62	4.51	5.82	5.61

Notes: The dependent variable is technological myopia as defined in equation 11 above while the main explanatory variable is the different measures of coal exposure, as discussed in Section 3. Other important covariates shown here include distance to Paris, initial population density, longitude, latitude. Other geographical and geological controls discussed in Section 3 are included. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

5.2 The burden of path dependence

Another mechanism which could account for the persistence in the negative impact of coal industry on productivity and patenting levels could be some form of path dependence. In the wake of Paul David's pioneering research (David (1985)), path dependence has garnered a rising interest in economic literature; important contributions in that respect include Acemoglu et al. (2001) and Aghion et al. (2012b). Given the time-span during which coalfields were actively ran in France, coal is likely to have lastingly affected industry and innovation structures. This subsection will investigate that plausible channel by first looking at graphical evidence, before showing that coal departments exhibit substantially larger similarities with respects to their former industry and innovation structure.

5.2.1 Graphical evidence

Before turning to empirical results, it is noteworthy to take a look at Figure 7 which summarizes the evolution of innovation structure in departments heavily invested in the coal industry compared to departments barely affected by coal extraction. Figure 7 simply plots the evolution of the year on year proximity with respects to the initial patenting structure of departments most heavily exposed to coal (gold) compared to departments least heavily exposed to coal (navy blue). Formally, this value is obtained by computing:

$$\forall t, \mathcal{S}_{d,t} = \mathcal{X}_{d,t} \cdot \mathcal{X}_{d,1907} \tag{12}$$

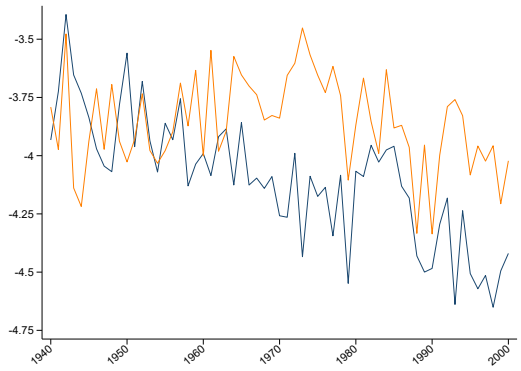
Where $\mathcal{X}_{d,t}$ is the vector that stacks the share of each 3-digit IPC category in total patenting activity of department d at time t while $\mathcal{X}_{d,1907}$ stacks the share of each 3-digit IPC category in total patenting activity of department d back in 1907. In words, the larger this product, the more similar patenting activity in department d at time t is to the patenting activity of that department back in 1907. Said differently, the larger this value, the more the department resembles its former innovation type, implying that it is locked in a given specialization. Clearly, departments heavily exposed to coal exhibit a sharper path dependence as measured by equation 12: throughout the 1940-2010 period, such departments indeed exhibit substantially larger levels of similarity with respects to their past innovation structure compared to departments less intensively exposed to coal extraction. This finding is robust to excluding Paris out of the sample, a finding which is important to the extent that Paris accounts for almost half of total French patenting activity in the XXth century.

5.2.2 Industry path dependence

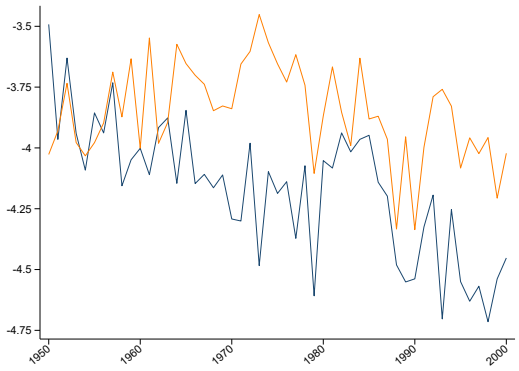
I now explore whether this graphical evidence pertains to a causal adverse effect of coal exposure on path dependence. I first do this by checking whether coal departments were more likely to suffer from industrial path dependence, in the sense that their industrial structure evolved less with respects to its baseline compared to non coal departments. Unfortunately, I lack homogeneous data on industrial employment throughout the XXth century; the best I can do is to leverage the *Enquêtes Emploi en Continu* which spans from 1980 to 2021 and provides detailed information on employment at the NAF 2-digit level in each department. I am thus able to compute an indicator of similarity of employment structure in each department compared to its 1980 level; formally, I compute industrial similarity $\mathcal{S}^I_{d,t}$ as:

FIGURE 7: Graphical evidence on path dependence in coal departments

(a) Including Paris



(b) Excluding Paris



Notes: Figure (a) plots the evolution of similarity with respects to 1907 innovation structure in the top 10% departments in terms of exposure to coal (gold) versus the bottom 1°% departments in terms of exposure to coal (navy blue), including Paris in the sample, while Figure (b) replicates this graph but excluding Paris as a robustness check.

$$\forall t \geq 1980, \mathcal{S}^{I}_{d,t} = \mathcal{X}^{I}_{d,t} \cdot \mathcal{X}^{I}_{d,1980} \tag{13}$$

Where $\mathcal{X}^{I}_{d,t}$ is the vector that stacks the share of each 2-digit NAF sector in total employment in department d at time t and $\mathcal{X}^{I}_{d,1980}$ is that same vector but for year 1980. In words, and as in the specification introduced in equation 12, $\mathcal{S}^{I}_{d,t}$ measures how close a department is from its initial industry structure. The closer this value is to 1, the more the department resembles its former industry structure; by contrast, the closer this value is to 0, the further away the department is from its baseline industry structure, and, hence, the less it exhibits path dependence in employment and industry specialization. I then run equation 6 using $\mathcal{S}^{I}_{d,t}$ as the dependent variable, and summarize the results obtained in Table 14. Strikingly, coal exposure induced a stronger dependence on past specialization as evidenced by the positive and significant coefficient obtained for 5 out of the 6 specifications considered.

On average, a one standard deviation increase in coal exposure raised dependence to past industrial structure by 10 to 20%. Alternatively, having been exposed to coal exposure, as measured by the binary variable, allegedly increased path dependence in terms of industry structure by 20 to 40% compared to departments that never mine any coal. Such estimates suggest that the effect of coal on industry specialization in the long run was lasting and sizable, which, in turn, could partly explain the persistent slowdown in productivity experienced by those departments throughout the XXth century, a slowdown which accelerated in the beginning of the 60s as suggested by the graphical evidence introduced in Figure 1.

TABLE 14: Effect of coal on industrial path dependence

	Proximity to all coalfields		**(Log) coal surface**		**Binary measure**	
	OLS	IV	OLS	IV	OLS	IV
Coal exposure	0.077	0.204***	0.103**	0.215***	0.174***	0.405**
	(0.048)	(0.070)	(0.046)	(0.073)	(0.062)	(0.143)
Initial density	-0.197**	-0.245***	-0.187**	-0.209***	-0.206***	-0.258***
	(0.075)	(0.081)	(0.097)	(0.075)	(0.073)	(0.087)
Distance to Paris	0.000	-0.000	0.000	-0.000	0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)
Longitude	-0.003	0.013	-0.000	0.013	0.007	0.032
	(0.004)	(0.073)	(0.057)	(0.055)	(0.064)	(0.068)
Latitude	-0.067	-0.114	-0.079	-0.125	-0.059	-0.087
	(0.099)	(0.110)	(0.084)	(0.089)	(0.088)	(0.097)
Observations	3,563	3,563	3,563	3,563	3,563	3,563
Year Fixed-effect	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Geological controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.745	0.405	0.748	0.424	0.751	0.407
F-Stat	12.68	9.45	11.53	9.85	11.99	7.66

Notes: The dependent variable is industrial path dependence as defined in equation 13 above while the main explanatory variable is the different measures of coal exposure, as discussed in Section 3. Other important covariates shown here include distance to Paris, initial population density, longitude, latitude. Other geographical and geological controls discussed in Section 3 are included. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

5.2.3 Innovation path dependence

In turn, industry path dependence also translated in innovation path dependence, in the sense that coal departments were more likely to keep the same innovation structure throughout the XXth century, precluding them from exploring more recent and contemporaneous technological fields. This finding can be supported by first defining technological path dependence as summarized in equation 12. I then leverage the patent sample spanning from 1945 to 2017 to test equation 6 where I take $\mathcal{S}_{d,t}$ as the dependent variable. Restricting the sample to the post-WW2 period aims at better capturing path dependence, in the sense that dependence to former industry structure may only be measured in the long run; testing for dependence of patenting activity in 1908 with respects to innovation structure in 1907 would, in that respect, not be meaningful; additionally, this allows me to exclude war periods during which innovation may have been substantially altered in its structure owing to government intervention. Results of this specification are summarized in Table 15.

Consistently with the estimates reported in Table 14, coal exposure seemingly strengthened reliance on past innovation. On average, a one standard deviation increase in coal exposure translated in an increase in dependence to past innovations by 12 to 38%, a range of magnitude that is roughly comparable to that derived in the previous

TABLE 15: Effect of coal on industrial path dependence

	Proximity to all coalfields		(Log) coal surface		Binary measure	
	OLS	IV	OLS	IV	OLS	IV
Coal exposure	0.123*	0.382***	0.175**	0.384***	0.234**	0.718***
	(0.062)	(0.118)	(0.075)	(0.117)	(0.105)	(0.241)
Initial density	-0.311***	-0.387***	-0.298***	-0.326***	-0.206***	-0.409***
	(0.075)	(0.107)	(0.100)	(0.103)	(0.116)	(0.156)
Distance to Paris	-0.000	-0.002	-0.000	-0.001	-0.000	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)
Longitude	0.036	0.101	0.023	0.044	0.052	0.150
	(0.078)	(0.088)	(0.080)	(0.080)	(0.078)	(0.102)
Latitude	-0.052	-0.190	-0.061	-0.150	-0.040	-0.152
	(0.119)	(0.138)	(0.119)	(0.119)	(0.117)	(0.134)
Observations	5,530	5,530	5,530	5,530	5,530	5,530
Year Fixed-effect	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Geological controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.097	0.025	0.091	0.424	0.088	0.011
F-Stat	2.40	2.80	3.03	3.56	2.63	1.83

Notes: The dependent variable is innovation path dependence as defined in equation 12 above while the main explanatory variable is the different measures of coal exposure, as discussed in Section 3. Other important covariates shown here include distance to Paris, initial population density, longitude, latitude. Other geographical and geological controls discussed in Section 3 are included. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

subsection. Alternatively, departments that mined coal exhibit, on average, a path dependence in terms of innovation which is greater by 25 to 70% compared to departments that never mined any coal. All in all, evidence accumulated in the three last subsections pertain to an adverse effect of coal exposure on capital and ideas reallocation: more specifically, departments which invested more heavily in coal did not manage to reallocate their funds and technical structure to alternative industries and sectors over time which might bring an explanation to their persistently dampened innovation levels.

5.3 Human capital trap

A last and important mechanism that could account for the persistence in the dampened productivity level of coal-exposed departments lies in the cultural lock-in suffered by such departments. More specifically, as they invested in low-skilled industries, these departments did not need to rely on a large basin of highly skilled workers which, in turn, lowered human capital accumulation as was evidenced in Section 5. Yet, it remains to be proved that coal basins suffered a peculiarly hard time at escaping this suboptimal equilibrium. This subsection will hence discuss the feedback loop whereby poor education levels were transmitted within family of coalfield workers

while gifted individuals had no choice but to leave their birth department in order to achieve better career outcomes in more dynamic places.

5.3.1 Within-family human capital transmission

I start this subsection by investigating the role of within-family transmission in the human capital trap suffered by coal departments. More specifically, I bring evidence that children whose father was a coalfield worker were less likely to achieve higher education or to achieve high career outcomes such as becoming a white collar, an engineer or even a teacher. I then rationalize these findings by showing that job transmission was peculiarly high within coalfield families compared to families with other job occupations. For that purpose, I leverage the 1977 survey *Formation et Qualification Profesionnelle* which collected data on more than 35,000 individuals coming from the whole French territory; in particular, this survey gathered information on career outcomes, educational trajectories, birth place, residence place as well as valuable insights on parents' career and achievements.

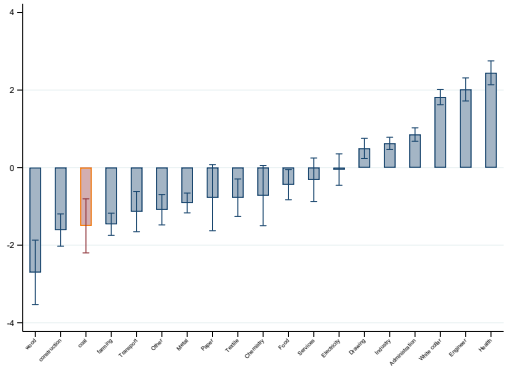
I then consider probit or logit models of the following form:

$$\mathcal{Y}_i = \Phi(\mathcal{X}_i\beta + \Gamma\mathcal{Z}_i + \varepsilon_i) \tag{14}$$

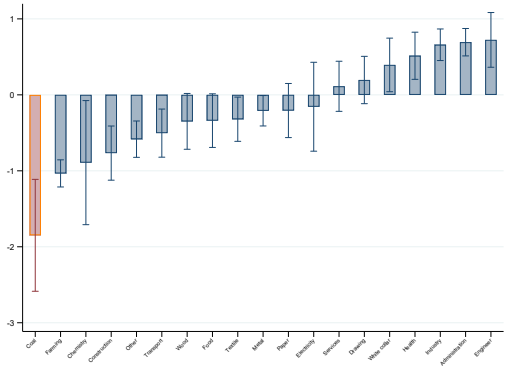
Where $\mathcal{Y}_i$ is a dichotomous outcome of interest for individual i taking the value 1 when the event of interest is fulfilled and 0 otherwise. Events of interest include having completed higher education, occupying a white collar position, having a high-skilled job or working as an intellectual. $\mathcal{X}_i$ is a vector of j dummy variables where line k takes value 1 if individual i's father occupied job k, while $\mathcal{Z}_i$ controls for other idiosyncratic factors such as the birth department, the residence department, the birth year or the sex. In concrete terms, the occupation of individuals' father is defined by the 1975 employment code following the typology of INSEE at that time. While this classification was as detailed as a 4-digit stratification, I focus on a broader classification of jobs which I detail in the Appendix; it basically consists in grouping consistent 2-digit NAF occupations together to end up with 20 job categories which are more tractable and easier to compare. I then leverage this specification for respectively four outcome variables: first, I investigate the role of father occupation on the probability of achieving higher education; second, I turn to the impact of father job on the probability of becoming a white collar; third, I explore the effect of father's job on the probability of becoming an engineer; I finally sum up by considering the impact of father's job on the probability of occupying a high-skilled job, where I define a high-skilled job as being a white collar, an engineer, a teacher or an intellectual.

FIGURE 8: Father's job effect on probability of achieving higher education and becoming a white-collar

(a) Upper education



(b) White collar jobs



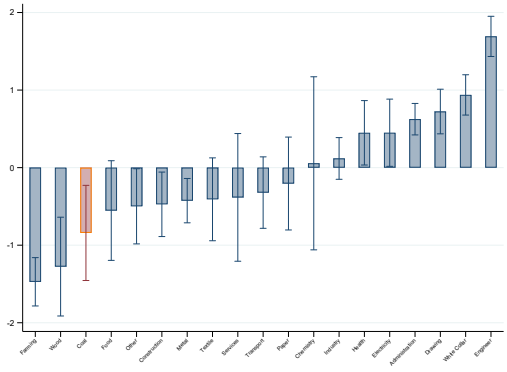
Notes: Figure (a) plots the coefficients when testing Equation 14 with a binary variable taking the value 1 if the individual achieved higher education and 0 otherwise while Figure (b) plots the coefficients when testing Equation 14 with a binary variable taking the value 1 if the individual works as a white collar and 0 otherwise. For the sake of readability, coal professions are pictured in orange while other professions are shown in blue.

Results of these specifications are shown respectively in Figure 8 and 9, where, for the sake of readability, I plot in orange the coefficients associated to the effect of having a father working in the coal mining industry, while blue coefficients are attached to other kinds of industries which include farming, metal, chemistry, electricity, administration, engineering, paper, wood, transportation, drawing, services, white collars, construction, textile, health, food and others. In summary, what these graphs suggest is that coal-professions, namely jobs which implied working in a coalfield, exerted a significant and negative effect on the probability of having kids achieving higher education, becoming white collars, engineers or, more generally, on having kids working in high-skilled positions. More importantly perhaps, these estimates suggest that, among the twenty job categories tested, the mining category ranks among the most negative ones in terms of the effect on kids' achievements. Indeed, coal industry ranks respectively 3^{rd}, 1^{st}, 3^{rd} and 2^{nd} in terms of negative impact on kids outcome when considering respectively higher education, occupying a white collar job, an engineer job and a high-skilled job.

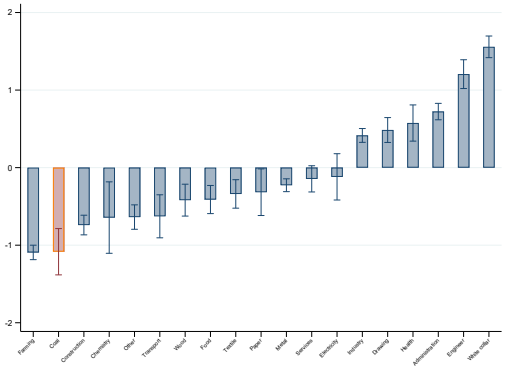
Jointly with farming occupations, coal jobs are the professions that exerted the most significant and negative impact on education and career outcomes of kids. In terms of magnitude, having a father working as a coalfield worker reduced the chance of achieving higher education by 5.2%, an outcome equivalent to that obtained for the kids of farming workers. Similarly, having a father working in coalfields translates in a 8% lower likelihood of working as a white collar, while decreasing the chance of working as an engineer by 1.9%. Overall, having a father working as a coalfield worker allegedly decreased the likelihood of occupying a high-skilled job (as defined above) by roughly 10.8%, an effect which is very close to that obtained for kids of farming

FIGURE 9: Father's job effect on probability of becoming an engineer of occupying a high-skilled job

(a) Engineer



(b) High-skilled job



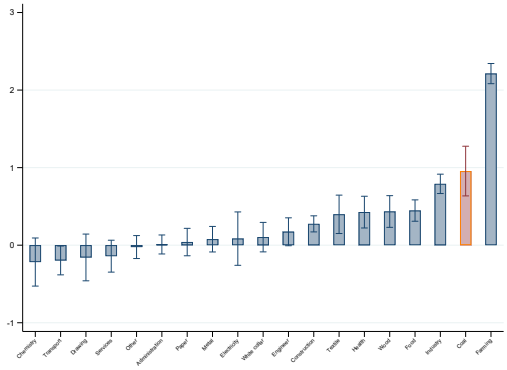
Notes: Figure (a) plots the coefficients when testing Equation 14 with a binary variable taking the value 1 if the individual works as an engineer and 0 otherwise while Figure (b) plots the coefficients when testing Equation 14 with a binary variable taking the value 1 if the individual occupies a high-skilled job and 0 otherwise. For the sake of readability, coal professions are pictured in orange while other professions are shown in blue.

workers (10.9%) and much larger than that corresponding to children of metal workers (-2.3%), construction workers (-7.4%) or even transportation workers (-6.3%). That the effect of having a father working as a coal worker is peculiarly large might be accounted for by the fact that kids of coal workers were more likely to work themselves as coal workers. This finding is indeed supported by testing equation **??** where the dichotomous variable is taken to be 1 if the father of surveyed individual occupied the same job as his own father, and 0 otherwise.

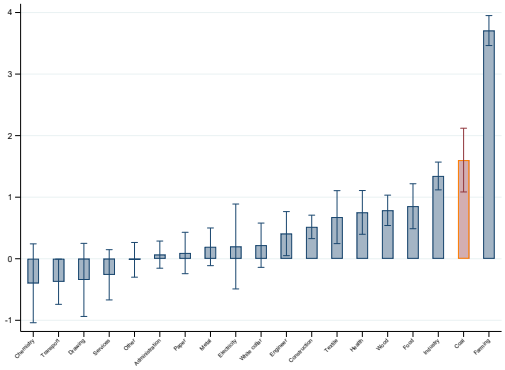
The results of this last check are summarized in Figure 10 which summarizes the probability of occupying the same job as one's father depending on the industry in which one works. As expected, farming jobs are those with the largest within-family transmission: being the son of a farmer increases chance of occupying the same position by 47%. While this effect is significantly larger than that derived for all other jobs, coal industry comes second and is associated to a 20% higher chance of intergenerational job transmission. To get a sense of the magnitude of this coefficient, it is telling to observe that the same coefficient for metal and construction workers correspond respectively t a 2.4% lower chance of working in the same field as one's father, and a 5.3% higher chance of working in the same field as one's father. All in all, this subsection lent credence in favor of the argument that within-family transmission was peculiarly strong among families related to coal. More specifically, being the kid of a coal worker substantially reduced the likelihood of building a strong human capital, or - which, to some extent, is equivalent - of occupying high-skilled positions; this finding mostly derives from the fact that such families were locked in this specific industry.

FIGURE 10: Probability of occupying the same job as one's father, by indsutry

(a) Probit model



(b) Logit model



Notes: Figure (a) plots the probit coefficients when testing Equation 14 with a binary variable taking the value 1 if the individual works in the same field as his father Figure (b) plots the same coefficients using a logit model. For the sake of readability, coal professions are pictured in orange while other professions are shown in blue.

5.3.2 High-skilled migration

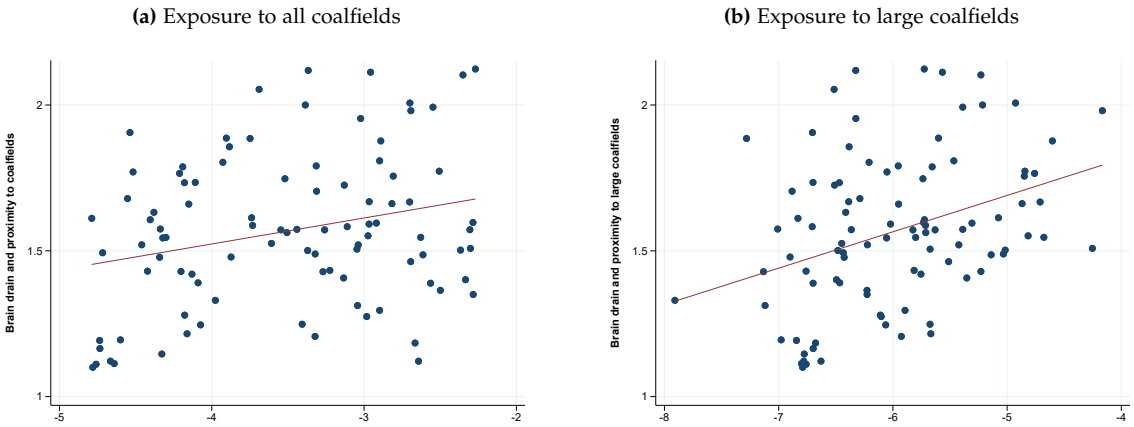
While families involved at some point of their history in coal mining seem to have been stuck in that industry, it could however be argued that innovation is a very concentrated process, so that only a few inventors suffice to drive productivity in a given place. In that respect, one could argue that being stuck in an industry is not necessarily detrimental from a patenting point of view, as long as a large enough pool of inventors may be maintained. Yet, there is empirical evidence that coal basins were peculiarly subject to a brain drain phenomenon, where by gifted individuals left their home department to achieve higher career or studies opportunities in another department. Migration is indeed a key component of the innovation network, which has been widely studied in the recent economic literature. To mention a few of them, Moser et al. (2014) leverage the Jewish emigration out of Nazi Germany to demonstrate how highly-education migrants revolutionized US science in the early 30s while Hornung (2014) exploits the Huguenot migration from France to Prussia in the early XVIIIth century to study the positive long-run impact of the Huguenot settlement on the productivity of textile manufactories in Germany. By contrast, I leverage the already mentioned *Enquête Emploi en Continu* to document a positive and sustained impact of coal exposure on out-migration of skilled labor force.

More specifically, I compute for each department d at time t the value $\mathcal{E}^S_{d,t}$ that captures excess representation of skilled individuals in migration out of a department. In concrete terms, for each available survey, I compute the share $s^S_{d,t}$ of skilled individuals among total individuals born in department d; similarly, I compute the share $s^{S,M}_{d,t}$ of skilled individuals among individuals who migrated out of their birth department d. I then define $\mathcal{E}^S_{d,t}$ to be:

$$\mathcal{E}^{S}_{d,t} = \frac{s^{S,M}_{d,t}}{s^{S}_{d,t}} \tag{15}$$

In words, this parameter captures the over-representation of skilled people in migration outflows out of any department d at time t. For all surveys, it relates the share of high skilled individuals who are no longer living in their birth department to the share of high-skilled individuals in that same department. The larger this ratio, the more high-skilled invididuals from a given department d tend to leave their birth department and find other opportunities in another French department. I then test equation 6 where I define the dependent variable to be $\mathcal{E}^{S}_{d,t}$, namely the over-representation of high-skilled individuals in out-migration patterns. Before turning to these empirical causal results, it is noteworthy that graphical evidence supports the idea of a strong brain drain effect out of coal departments.

FIGURE 11: Over-representation of high-skilled migration and coal exposure



Notes: Figure (a) plots excess migration outflows against proximity to all coalfields while Figure (b) plots excess migration outflows against proximity to large coalfields.

Figure 11 indeed states a clear and positive relationship between the exposure to coal, be it either measured as proximity to all coalfields (Panel (a)) or proximity to large coalfields (Panel (b)), on the brain drain measure. This positive correlation may be causally interpreted according to the results summarized in Table 16 which illustrates a positive and significant effect of coal exposure on the magnitude of brain drain out of French departments. All coefficients on the different coal exposure measures are indeed positive and significant at the 1% level, both in OLS and in IV specifications, suggesting that coal exposure induced a substantial brain drain out of former coal basins. In terms of magnitude, a one standard deviation increase in coal exposure translated in a 8 to 25% higher brain drain which suggests that the negative impact was sizable and sustained. Alternatively, departments which experienced coal mining in their history show 15 to 50% larger brain drain levels compared to departments

which never mine any coal. As will be shown in the Appendix, it is noteworthy that this effect still holds when adding unemployment rate as well as a variable coding for the number of universities per capita. In the end, the evidence summarized in this subsection add up to assert a negative and sustained feedback loop, whereby former coal basins were less likely to accumulate human capital, owing both to within-family transmission as well as to brain drain.

TABLE 16: Effect of coal exposure on brain drain

	Proximity to all coalfields		**(Log) coal surface**		**Binary measure**	
	OLS	IV	OLS	IV	OLS	IV
Coal exposure	0.104***	0.255***	0.082***	0.252***	0.150***	0.508***
	(0.033)	(0.062)	(0.028)	(0.056)	(0.051)	(0.135)
Initial density	0.037	-0.387***	0.029	-0.326***	0.043	-0.039
	(0.043)	(0.107)	(0.048)	(0.065)	(0.042)	(0.062)
Distance to Paris	0.001***	0.001	0.001***	0.001**	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
Longitude	-0.059	-0.025	-0.073**	-0.052	-0.059*	-0.002
	(0.036)	(0.049)	(0.033)	(0.035)	(0.035)	(0.059)
Latitude	0.138**	0.061	0.158***	0.089	0.163***	0.095
	(0.055)	(0.074)	(0.053)	(0.061)	(0.050)	(0.075)
Observations	3,697	3,697	3,697	3,697	3,697	3,697
Year Fixed-effect	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Geological controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.346	0.265	0.326	0.245	0.343	0.163
F-Stat	8.61	10.80	10.47	12.58	7.32	7.80

Notes: The dependent variable is brain drain as defined in 15 while the main explanatory variable is the different measures of coal exposure, as discussed in Section 3. Other important covariates shown here include distance to Paris, initial population density, longitude, latitude. Other geographical and geological controls discussed in Section 3 are included. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

6 Conclusion

In summary, this paper draws on a simplified theoretical framework to suggest the existence of a negative and sustained effect of exposure to extractive industries such as coal industry on productivity, education and industry specialization. Leveraging a unique French dataset, I provide evidence of a negative and significant effect of involvement in coal extraction on patenting activity at the local level in France throughout the XXth century as well as in the beginning of the XXIst century. Not only did coal exposure decrease the quantity of innovation; it also reduced its quality and sustainedly shaped its orientation for subsequent decades. This negative effect of coal exposure on patenting activity comes with a negative impact on industry special-

ization and education decisions, which persisted over time through creative myopia and a negative cultural lock-in.

References

Acemoglu, Daron, Simon Johnson, and James A. Robinson, "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution," NBER Working Papers 8460, National Bureau of Economic Research, Inc September 2001.

Aghion, Philippe and Peter Howitt, "A Model of Growth through Creative Destruction," *Econometrica*, March 1992, *60* (2), 323–351.

_ , Antoine Dechezleprtre, David Hemous, Ralf Martin, and John Van Reenen, "Carbon Taxes, Path Dependency and Directed Technical Change: Evidence from the Auto Industry," GRI Working Papers 102, Grantham Research Institute on Climate Change and the Environment November 2012.

_ , Jing Cai, Mathias Dewatripont, Luosha Du, Ann Harrison, and Patrick Legros, "Industrial Policy and Competition," *American Economic Journal: Macroeconomics*, October 2015, *7* (4), 1–32.

_ , Mathias Dewatripont, Luosha Du, Ann Harrison, and Patrick Legros, "Industrial Policy and Competition," NBER Working Papers 18048, National Bureau of Economic Research, Inc May 2012.

Agrawal, Ajay K., Iain M. Cockburn, and Carlos Rosell, "Not Invented Here? Innovation in Company Towns," NBER Working Papers 15437, National Bureau of Economic Research, Inc October 2009.

Allcott, Hunt and Daniel Keniston, "Dutch Disease or Agglomeration? The Local Economic Effects of Natural Resource Booms in Modern America," *The Review of Economic Studies*, 07 2017, *85* (2), 695–731.

Asch, Kristine, "The 1:5 Million International Geological Map of Europe and Adjacent Areas (IGME5000) (WMS)," 2005.

Bergeaud, Antonin and Cyril Verluise, "A new dataset to study a century of innovation in Europe and in the US," *Research Policy*, 2024, *53* (1), 104903.

Brey, B. and V. Rueda, "The Death of King Coal and the Scars of Deindustrialization," *CEPR Discussion Paper*, 2024, (19082).

Corden, W Max and J Peter Neary, "Booming Sector and De-Industrialisation in a Small Open Economy," *Economic Journal*, December 1982, *92* (368), 825–848.

David, Paul A, "Clio and the Economics of QWERTY," *American Economic Review*, May 1985, *75* (2), 332–337.

Esposito, Elena and Scott F. Abramson, "The European coal curse," *Journal of Economic Growth*, March 2021, *26* (1), 77–112.

Fernihough, Alan and Kevin Hjortshøj O'Rourke, "Coal and the European Industrial Revolution," NBER Working Papers 19802, National Bureau of Economic Research, Inc January 2014.

Franck, Raphaël and Oded Galor, "Technology-Skill Complementarity in Early Phases of Industrialization," NBER Working Papers 23197, National Bureau of Economic Research, Inc February 2017.

_ **and** _ , "Flowers of evil? Industrialization and long run development," *Journal of Monetary Economics*, 2021, *117* (C), 108–128.

Hornung, Erik, "Immigration and the Diffusion of Technology: The Huguenot Diaspora in Prussia," *American Economic Review*, January 2014, *104* (1), 84–122.

Kelly, Bryan, Dimitris Papanikolaou, Amit Seru, and Matt Taddy, "Measuring Technological Innovation over the Long Run," *American Economic Review: Insights*, September 2021, *3* (3), 303–20.

Luterbacher, Jürg, Daniel Dietrich, Elena Xoplaki, Martin Grosjean, and Heinz Wanner, "European Seasonal and Annual Temperature Variability, Trends, and Extremes Since 1500," *Science (New York, N.Y.)*, 04 2004, *303*, 1499–503.

Marx, Karl, *Le Capital*, Verlag von Otto Meisner, 1867.

Mokyr, Joel and Hans-Joachim Voth, "Understanding growth in Europe, 1700–1870: theory and evidence," *The Cambridge economic history of modern Europe*, 2010, *1*, 7–42.

Moser, Petra, Alessandra Voena, and Fabian Waldinger, "German Jewish Émigrés and US Invention," *American Economic Review*, October 2014, *104* (10), 3222–55.

Pomeranz, Kenneth, *The Great Divergence*, Princeton University Press, 2014.

Ramankutty, Navin, Jonathan Foley, John Norman, and Kevin Mcsweeney, "The Global Distribution of Cultivable Lands: Current Patterns and Sensitivity to Possible Climate Change," *geb*, 09 2002, *11*, 377–.

Rud, Juan-Pablo, Michael Simmons, Gerhard Toews, and Fernando Aragon, "Job displacement costs of phasing out coal," *Journal of Public Economics*, 2024, *236* (C).

Sachs, Jeffrey, Andrew Mellinger, and John Gallup, "The Geography of Poverty and Wealth," *Scientific American*, 04 2001, *284*, 70–5.

Sachs, Jeffrey D. and Andrew M. Warner, "Natural Resource Abundance and Economic Growth," NBER Working Papers 5398, National Bureau of Economic Research, Inc December 1995.

_ **and** _ , "The curse of natural resources," *European Economic Review*, May 2001, *45* (4-6), 827–838.

Smith, Adam, *The Wealth of Nations* 1776.

Squicciarini, Mara P., "Devotion and Development: Religiosity, Education, and Economic Progress in Nineteenth-Century France," *American Economic Review*, November 2020, *110* (11), 3454–3491.

ONLINE APPENDICES

Table of contents

A Theoretical framework **OA-1**

A.1 Proof of optimal consumption levels OA-1

A.2 Proof of optimal innovation decisions OA-2

A.3 Proof of uniqueness of cutoff value OA-3

A.4 Proof of Proposition 1 . . . . . OA-5

A.5 Proof of Proposition 2 . . . . . OA-6

A.6 Proof of Proposition 3 . . . . . OA-7

A.7 Proof of Proposition 4 . . . . . OA-9

A.8 Proof of Proposition 5 . . . . . OA-13

A.9 Proof of Corollary 3 . . . . . OA-17

B Data appendix **OA-18**

B.1 Coal data OA-18

C Instrumental variable **OA-19**

C.1 Graphical evidence OA-19

C.2 Regression evidence OA-19

D Additional robustness checks **OA-22**

D.1 Robustness on patent quantity OA-22

D.2 Robustness on patent quality OA-22

D.3 Robustness on specialization OA-24

D.4 Robustness on breakthrough patents OA-25

D.5 Are the results driven by Paris? OA-28

D.6 Robustness on education OA-30

D.7 Robustness on industry specialization OA-31

D.8 Robustness on public support OA-31
D.9 Allowing for region-fixed effects OA-34
D.10 Alternative clustering methods OA-35
D.11 Dynamic evolution of the effect over time OA-35

A Theoretical framework

A.1 Proof of optimal consumption levels

The representative consumer optimizes:

$$\mathcal{U}(x_i, x_j) = \ln x_i + \ln x_j \tag{A1}$$

Under the resource constraint:

$$x_i + x_j \leq 2 \tag{A2}$$

Given that the resource constraint is saturated, this yields:

$$x_i = 2 - x_j \tag{A3}$$

Substituting in equation A1 implies that:

$$\mathcal{U}(x_j) = \ln(2 - x_j) + ln(x_j) \tag{A4}$$

Taking the derivative with respects to x_j gives:

$$\frac{1}{2 - x_j^*} = \frac{1}{x_j^*} \tag{A5}$$

Which simplifies to:

$$2 - x_j^* = x_j^* \tag{A6}$$

So that:

$$x_j^* = 1 \tag{A7}$$

And:

$$x_i^* = 2 - x_j^* = 1 \tag{A8}$$

Hence:

$$\forall l \in \{i, j\}, x_l^* = 1 \tag{A9}$$

A.2 Proof of optimal innovation decisions

The innovation problem in the resource-intensive consists in maximizing:

$$\Pi_i(q_i; p_n, \mathcal{Q}_j) = q_i \beta_i(\mathcal{Q}_j) - \frac{q_i^2}{2} - p_n \tag{A10}$$

Taking the derivative with respects to q_i entails the following optimality condition:

$$\beta_i(\mathcal{Q}_j) - q_i^* = 0 \tag{A11}$$

Which trivially implies that:

$$q_i^* = \beta_i(\mathcal{Q}_j) \tag{A12}$$

Symmetrically, the innovation problem in the knowledge-intensive industry consists in maximizing:

$$\Pi_{j,k}(q_j; \kappa_k) = q_{j,k} \beta_j - \frac{q_{j,k}^2}{2} - \kappa_k \tag{A13}$$

Taking the derivative of the previous equation with respects to $q_{j,k}$ entails the following optimality condition:

$$\beta_j - q_{j,k}^* = 0 \tag{A14}$$

Which trivially yields:

$$q_j^* = \beta_j \tag{A15}$$

Where I removed the subscript k as the optimal innovation effort, conditionally on joining industry j, does not depend on the firm-specific individual ability as captured by κ_k.

A.3 Proof of uniqueness of cutoff value

As stated in the main text, firms join the knowledge-intensive industry if and only if the profit derived in the former is greater than that derived in the latter. Formally, this condition writes:

$$\frac{1}{2}\beta_j^2 - \kappa_k \geq \frac{1}{2}\left[(1+\xi_i)(1+\mathcal{Q}_j)\right]^2 - p_n \tag{A16}$$

Rearranging implies that the cost of education must abide by:

$$\kappa_k \leq \frac{1}{2}\beta_j^2 - \frac{1}{2}\left[(1+xi_i)(1+\mathcal{Q}_j)\right]^2 + p_n \tag{A17}$$

This formula basically states that only firms with a low enough cost of education will join the knowledge-intensive sector. As κ_k is uniformly distributed over the $[0,1]$ interval, this means that the fraction of firms with a sufficiently low level of education cost can be defined as:

$$\bar{\kappa} = \frac{1}{2}\beta_j^2 - \frac{1}{2}\left[(1+\xi_i)(1+\mathcal{Q}_j)\right]^2 + p_n \tag{A18}$$

Where:

$$\mathcal{Q}_j = \int_0^{\bar{\kappa}} q_{j,k}dk \tag{A19}$$

It then appears that equation A72 implicitly defines the equilibrium value $\bar{\kappa}$. More specifically, the left hand side of equation A72 is increasing in $\bar{\kappa}$ while the right hand side is decreasing in $\bar{\kappa}$ given that the derivative of the RHS with respects to $\bar{\kappa}$ writes:

$$-(1+\xi_i)\ \bar{\kappa}\left[(1+\xi_i)(1+\bar{\kappa}q_j)\right] \tag{A20}$$

The first term of this equation is negative while the second is positive. Hence, it follows that the right hand side of equation A72 is decreasing in $\bar{\kappa}$ while the left hand side is increasing. Let Θ be the following function:

$$\Theta(\bar{\kappa}; p_n, \beta_j) = \bar{\kappa} - \left(\frac{1}{2}\beta_j^2 - \frac{1}{2}\left[(1+\xi_i)(1+\mathcal{Q}_j)\right]^2 + p_n\right) \tag{A21}$$

And let assumption 4 write:

Assumption 4. $\lim_{\bar{\kappa}\to 0}\Theta(\bar{\kappa}; p_n, \beta_j) > 0$ *and* $\lim_{\bar{\kappa}\to 1}\Theta(\bar{\kappa}; p_n, \beta_j) \leq 0$

Then, it follows that there exists a unique solution to equation A72, meaning that there exists a unique value $\bar{\kappa}$ on the interval $[0, 1]$ such that firms with a cost inferior to $\bar{\kappa}$ invest in knowledge and firms with a cost in excess of $\bar{\kappa}$ join the resource-intensive industry. It is noteworthy that this demonstration could be made in a more general framework where no assumption is made on the within-industry degree of competition which, for the sake of clarity and concision, was set to 0 in this model.

A.4 Proof of Proposition 1

I start from the implicit equation which defines $\bar{\kappa}$:

$$\bar{\kappa} = \frac{1}{2}\left[\beta_j - (1-\Delta_j)\right]^2 - \frac{1}{2}\left[\beta_i(\mathcal{Q}_j) - (1-\Delta_i)\right]^2 - (1-\Delta_i) + p_n \tag{A22}$$

Taking the derivative on both sides with respects to p_n writes:

$$\frac{\partial \bar{\kappa}}{\partial p_n} = -\frac{1}{2}\frac{\partial\left[(1+\xi_i)(1+\bar{\kappa}q_j) - (1-\Delta_i)\right]^2}{\partial p_n} + 1 \tag{A23}$$

As:

$$\frac{1}{2}\frac{\partial\left[(1+\xi_i)(1+\bar{\kappa}q_j) - (1-\Delta_i)\right]^2}{\partial p_n} = q_j(1+\xi_i)\frac{\partial \bar{\kappa}}{\partial p_n}\left[(1+\xi_i)(1+\bar{\kappa}q_j) - (1-\Delta_i)\right] \tag{A24}$$

This implies:

$$\frac{\partial \bar{\kappa}}{\partial p_n} = 1 - \frac{\partial \bar{\kappa}}{p_n} q_j(1+\xi_i)\left[(1+\xi)(1+\bar{\kappa}) - (1-\Delta_i)\right] \tag{A25}$$

Rearranging yields:

$$\frac{\partial \bar{\kappa}}{\partial p_n}\left[1 + q_j(1+\xi_i)\left((1+\xi)(1+\bar{\kappa}) - (1-\Delta_i)\right)\right] = 1 \tag{A26}$$

Finally, I obtain that:

$$\frac{\partial \bar{\kappa}}{\partial p_n} = \frac{1}{\left[1 + q_j(1+\xi_i)\left((1+\xi)(1+\bar{\kappa}) - (1-\Delta_i)\right)\right]} \tag{A27}$$

Given that it was already shown before that:

$$(1+\xi)(1+\bar{\kappa}) - (1-\Delta_i) \geq 0 \tag{A28}$$

It is straightforward to notice that:

$$\frac{\partial \bar{\kappa}}{\partial p_n} \geq 0 \tag{A29}$$

Which proves Proposition 1.

A.5 Proof of Proposition 2

First, it should be recalled that total knowledge expenditures κ write:

$$\kappa = \int_0^{\bar{\kappa}} x dx \tag{A30}$$

Which trivially implies that:

$$\kappa = \bar{\kappa}^2 \tag{A31}$$

Hence,

$$\frac{\partial \kappa}{\partial p_n} = 2\frac{\partial \bar{\kappa}}{\partial p_n}\bar{\kappa} \tag{A32}$$

Since:

$$\frac{\partial \bar{\kappa}}{p_n} \geq 0 \tag{A33}$$

It naturally follows that:

$$\frac{\partial \kappa}{\partial p_n} \geq 0 \tag{A34}$$

Which is what Proposition 2 states.

A.6 Proof of Proposition 3

Let us first recall that total innovation $\mathcal{Q}$ writes:

$$\mathcal{Q} = s_i q_i^* + s_j q_j^* \tag{A35}$$

Replacing q_i^* and q_j^* with their equilibrium values write:

$$\mathcal{Q} = s_i \left[(1+\xi_i)(1+\mathcal{Q}_j)\right] + (1-s_i)\beta_j \tag{A36}$$

Also, let us recall that:

$$\mathcal{Q}_j = \int_0^{\bar{\kappa}} q_{j,k}^* \tag{A37}$$

Which yields that:

$$\mathcal{Q}_j = \bar{\kappa}\beta_j \tag{A38}$$

As it was demonstrated earlier that:

$$\bar{\kappa} = s_j \tag{A39}$$

I finally have:

$$\mathcal{Q}_j = s_j \beta_j \tag{A40}$$

So $\mathcal{Q}$ now writes:

$$\mathcal{Q} = s_i \left[(1+\xi_i)(1+(1-s_i)\beta_j\right] + (1-s_i)\beta_j \tag{A41}$$

Taking the derivative with respects to s_i, namely the share of the resource-intensive sector, implies:

$$(1+\xi_i)\left[1+(1-s_i^*)\beta_j\right] - s_i^*(1+\xi_i)\beta_j - \beta_j = 0 \tag{A42}$$

Rearranging and simplifying implies:

$$2s_i^*(1+\xi_i)\beta_j = 2(1+\xi_i) - \beta_j \tag{A43}$$

Which finally yields:

$$s_i^* = \frac{2(1+\xi_i) - \beta_j}{2\beta_j(1+\xi_i)} \tag{A44}$$

This shows that there exists a unique value s_i* which maximizes total innovation. Also, as $\mathcal{Q}$ is a quadratic function in s_i with a negative sign on the quadratic term, it necessarily follows that $\mathcal{Q}$ first increases and then decreases with s_i suggesting that innovation increases with the share of the resource-intensive sector before decreasing when the threshold s_i^* is reached. Interestingly enough, the optimal share of the resource-intensive industry from an innovation standpoint is decreasing in β_j, meaning that the better the knowledge-intensive is at generating inventions, the larger its share should be. Conversely, the optimal share of the resource-intensive industry is increasing in ξ_i which is straightforward as well for an increase in ξ_i corresponds to an increase in the ability of the resource-intensive industry to convert R&D processes into effective cost reduction.

Finally, as s_i is strictly decreasing in the price of natural resources, it is straightforward to show that Proposition 3 is equivalent to Corollary 1 which states that the total innovation effort is increasing and then decreasing with the price of natural resources p_n. Indeed, the larger the price of natural resources, the lower the share of the resource intensive industry. Starting from a very low level of the price of natural resources, increasing it can prove fruitful for it increases the share of the knowledge-intensive sector which, in turn, contributes to innovation in the resource-intensive industry through positive spillovers. At some point however, the resource-intensive industry shrinks by an excessive amount, which reduces the magnitude of the spillovers and, hence, the total innovation effort.

A.7 Proof of Proposition 4

First, let me recall that total profit in this economy writes:

$$\Pi = \int_{k\in\mathcal{I}} \pi_i dk + \int_{k\in\mathcal{J}} \pi_j dk \tag{A45}$$

Where π_i and π_j respectively index the firm-level profit in the resource-intensive industry and in the knowledge-intensive one. Also, s_i and s_j index the share of the resource-intensive industry and the share of the knowledge-intensive industry. This equation only asserts that total profits in this economy is equal to profits made by firm in each industry weighted by the share of each industry in total output. Using the optimal values of innovation, and the corresponding values of profit, it then follows that:

$$\Pi = \int_{k\in\mathcal{I}} \left[\frac{1}{2}\left[(1+\xi_i)(1+\mathcal{Q}_j)\right]^2 - p_n\right] dk + \int_{k\in\mathcal{J}} \left[\frac{1}{2}\beta_j^2 - \kappa_k\right] dk \tag{A46}$$

It should now be recalled that:

$$\mathcal{Q}_j = \int_{k\in\mathcal{J}} q_j^* dk = s_j\beta_j = (1-s_i)\beta_j \tag{A47}$$

This implies that total profit can be written in the following way:

$$\Pi = \int_{k\in\mathcal{I}} \left[\frac{1}{2}\left[(1+\xi_i)(1+\beta_j(1-s_i))\right]^2 - p_n\right] dk + \int_{k\in\mathcal{J}} \left[\frac{1}{2}\beta_j^2 - \kappa_k\right] dk \tag{A48}$$

Recalling that firms with an education cost inferior to $\bar{\kappa} = s_j$ join the knowledge-intensive sector enables to write:

$$\Pi = \int_0^{s_i} \left[\frac{1}{2}\left[(1+\xi_i)(1+\beta_j(1-s_i))\right]^2 - p_n\right] dk + \int_0^{1-s_i} \left[\frac{1}{2}\left[\beta_j\right]^2 - \kappa_k\right] dk \tag{A49}$$

For simplicity, I take the following notation for the remainder of this Appendix:

$$\gamma_i = 1 + \xi_i \tag{A50}$$

Which enables to rewrite total profit:

$$\Pi = \frac{s_i}{2}\left[\gamma_i(1+\beta_j(1-s_i))\right]^2 - s_i p_n + \frac{1-s_i}{2}\beta_j^2 - \frac{(1-s_i)^2}{2} \tag{A51}$$

Taking the derivative with respects to s_i implies the following optimality condition for s_i:

$$\frac{\partial \Pi}{\partial s_i} = \frac{\gamma_i^2}{2}\left[1+\beta_j(1-s_i)\right]^2 - \gamma_i^2 s_i\left[\beta_j(1+\beta_j(1-s_i))\right] - p_n - \frac{\beta_j^2}{2} + (1-s_i) \tag{A52}$$

Developing implies:

$$\frac{\partial \Pi}{\partial s_i} = \frac{\gamma_i^2}{2} + \beta_j\gamma_i^2 - \beta_j\gamma_i^2 s_i + \frac{(\gamma_i\beta_j)^2}{2}(1-s_i)^2 - \gamma_i^2\beta_j s_i - (\gamma_i\beta_j)^2 s_i + (\gamma_i\beta_j)^2 s_i^2 - \frac{\beta_j^2}{2} + (1-s_i) \tag{A53}$$

Or:

$$\frac{\partial \Pi}{\partial s_i} = \frac{3}{2}(\gamma_i\beta_j)^2 s_i^2 - s_i\left[2(\gamma_i\beta_j)^2 + 2\gamma_i^2\beta_j + 1\right] + \frac{1}{2}\left[\gamma_i^2 + 2\beta_j\gamma_i^2 + (\gamma_i\beta_j)^2 - \beta_j^2\right] - p_n \tag{A54}$$

It proves convenient to rearrange this expression in the following way:

$$\frac{\partial \Pi}{\partial s_i} = \frac{3}{2}(\gamma_i\beta_j)^2 s_i^2 - s_i\left[2\gamma_i\beta_j(\gamma_i\beta_j + \gamma_i) + 1\right] + \frac{1}{2}\left[(\gamma_i + \gamma_i\beta_j)^2 - \beta_j^2\right] - p_n \tag{A55}$$

Ths sign of this expression is ambiguous for it is quadratic with respects to s_i. This is a simple polynomial equation of degree 2 where I can compute Δ^*:

$$\Delta^* = \left[2\gamma_i\beta_j(\gamma_i\beta_j + \gamma_i) + 1\right]^2 - 4\left[\frac{3}{2}(\gamma_i\beta_j)^2\left(\frac{1}{2}(\gamma_i + \gamma_i\beta_j)^2 - \frac{1}{2}\beta_j^2 - p_n\right)\right] \tag{A56}$$

Developing entails the following expression:

$$\Delta^* = 4(\gamma_i\beta_j)^2(\gamma_i\beta_j + \gamma_i)^2 + 4\gamma_i\beta_j(\gamma_i\beta_j + \gamma_i) + 1 - 6(\gamma_i\beta_j)^2\left[\left(\frac{1}{2}(\gamma_i + \gamma_i\beta_j)^2 - \frac{1}{2}\beta_j^2 - p_n\right)\right] \tag{A57}$$

Simplifying then yields:

$$\Delta^* = (\gamma_i\beta_j)^2(\gamma_i\beta_j + \gamma_i)^2 + 4\gamma_i\beta_j(\gamma_i\beta_j + \gamma_i) + 1 + 3\beta_j^2(\gamma_i\beta_j)^2 + 6(\gamma_i\beta_j)^2 p_n \tag{A58}$$

It is straightforward to show that Δ is strictly larger than 0 for it is the sum of strictly positive components. Hence, equation A55 has 2 solutions, which I denote s_i^1 and s_i^2, such that:

$$s_i^1 = \frac{2\gamma_i\beta_j(\gamma_i\beta_j + \gamma_i) + 1 - \sqrt{\Delta}}{3(\gamma_i\beta_j)^2} \tag{A59}$$

And:

$$s_i^2 = \frac{2\gamma_i\beta_j(\gamma_i\beta_j + \gamma_i) + 1 + \sqrt{\Delta}}{3(\gamma_i\beta_j)^2} \tag{A60}$$

Yet, since $\gamma_i \geq 0$, it should be noticed that:

$$(\gamma_i\beta_j)^2(\gamma_i\beta_j + \gamma_i)^2 \geq (\gamma_i\beta_j)^2(\gamma_i\beta_j)^2 = (\gamma_i\beta_j)^4 \tag{A61}$$

Since:

$$4\gamma_i\beta_j(\gamma_i\beta_j + \gamma_i) + 1 + 3\beta_j^2(\gamma_i\beta_j)^2 + 6(\gamma_i\beta_j)^2 p_n \geq 0 \tag{A62}$$

It follows that:

$$\Delta^* \geq (\gamma_i\beta_j)^4 \tag{A63}$$

And, finally, that:

$$\sqrt{\Delta^*} \geq (\gamma_i\beta_j)^2 \tag{A64}$$

Meanwhile, we have that:

$$2\gamma_i\beta_j(\gamma_i\beta_j + \gamma_i) \geq 2(\gamma_i\beta_j)^2 \tag{A65}$$

Hence:

$$s_i^2 \geq \frac{3(\gamma_i \beta_j)^2}{3(\gamma_i \beta_j)^2} = 1 \tag{A66}$$

Yet, since s_i captures the share of the resource-intensive industry in the economy, it must hold that:

$$s_i \in [0, 1] \tag{A67}$$

This trivially ensures that there is one unique solution to equation A55 in the interval $[0, 1]$. As equation A55 is a polynomial equation of order 2 where the coefficient on the quadratic term is strictly greater than 0, it naturally follows that the derivative of profits with respects to s_i is decreasing in s_i, and is first positive and then negative on the $[0, 1]$ interval. This is like saying that Π has a unique maximum over the $[0, 1]$ interval which writes:

$$s_i^* = \frac{2\gamma_i \beta_j (\gamma_i \beta_j + \gamma_i) + 1 - \sqrt{\Delta}}{3(\gamma_i \beta_j)^2} \tag{A68}$$

Where $\Delta$is defined as above. This means that Π is first increasing, and then decreasing with the share of the resource-intensive industry s_i; as the latter is decreasing in the price of natural resources p_n this finally shows that there is an inverted U relationship between total profits and the price of natural resources in such an economy. Observe that, under assumption 4, s_i^* simply writes:

$$s_i^* = \frac{2\beta_j (\beta_j + 1) + 1 - \sqrt{\Delta^*}}{3\beta_j^2} \tag{A69}$$

A.8 Proof of Proposition 5

As stated in the main text, the gap between the actual and the optimal share of the resource-intensive industry writes:

$$\mathcal{G} = s_i^a - s_i^* \tag{A70}$$

It was shown in Appendix A.7 that the optimal share of the resource-intensive industry writes:

$$s_i^* = \frac{2\gamma_i\beta_j(\gamma_i\beta_j + \gamma_i) + 1 - \sqrt{\Delta}}{3(\gamma_i\beta_j)^2} \tag{A71}$$

In the meantime, and as shown in Appendix A.4, the share of the knowledge-intensive industry stemming from private decisions verifies:

$$s_j = \frac{1}{2}\beta_j^2 - \frac{1}{2}\left[(1+\xi_i)(1+\mathcal{Q}_j)\right]^2 + p_n \tag{A72}$$

Where, again:

$$\mathcal{Q}_j = \int_0^{s_j} q_j^* dk = s_j\beta_j \tag{A73}$$

Hence, using assumption 4, it must be that s_j verifies the following condition:

$$s_j = \frac{1}{2}\beta_j^2 - \frac{1}{2}(\beta_j s_j + 1)^2 + p_n \tag{A74}$$

Developing writes:

$$s_j = \frac{1}{2}\beta_j^2 - \frac{1}{2}(\beta_j^2 s_j^2 + 2\beta_j s_j + 1) + p_n \tag{A75}$$

Rearranging implies the following polynomial equation:

$$-\frac{1}{2}\beta_j^2 s_j^2 - (1+\beta_j)s_j + p_n + \frac{1}{2}\beta_j^2 - \frac{1}{2} = 0 \tag{A76}$$

The determinant Δ^a of this equation writes:

$$\Delta^a = (1+\beta_j)^2 + 2\beta_j^2(p_n + \frac{1}{2}\beta_j^2 - \frac{1}{2}) \tag{A77}$$

Clearly, Δ^a is strictly positive as the sum of two terms that are strictly greater than 0. It then follows that the above equation has two solutions, s_j^1 and s_j^2, such that:

$$s_j^1 = \frac{\sqrt{\Delta^a} - (1 + \beta_j)}{\beta_j^2} \tag{A78}$$

And:

$$s_j^2 = \frac{-\sqrt{\Delta^a} - (1 + \beta_j)}{\beta_j^2} \tag{A79}$$

Clearly, the second solution is negative which contradicts the condition $s_j \geq 0$ so that there is a unique equilibrium share of the knowledge-intensive industry in the economy which is:

$$s_j^a = \frac{\sqrt{\Delta^a} - (1 + \beta_j)}{\beta_j^2} \tag{A80}$$

As it is trivial that $s_i^a + s_j^a = 1$, it then naturally emerges that:

$$s_i^a = \frac{\beta_j(\beta_j + 1) + 1 - \sqrt{\Delta^a}}{\beta_j^2} \tag{A81}$$

Using Assumption 4 which normalizes γ_i to 1, it is now easy to rewrite $\mathcal{G}$ in the following way:

$$\mathcal{G} = \frac{\beta_j(\beta_j + 1) + 1 - \sqrt{\Delta^a}}{\beta_j^2} - \frac{2\beta_j(\beta_j + 1) + 1 - \sqrt{\Delta^*}}{3\beta_j^2} \tag{A82}$$

Rearranging implies that:

$$\mathcal{G} = \frac{3\beta_j(\beta_j + 1) + 3 - 3\sqrt{\Delta^a} - 2\beta_j(\beta_j + 1) - 1 + \sqrt{\Delta^*}}{3\beta_j^2} \tag{A83}$$

Simplifying entails that:

$$\mathcal{G} = \frac{\beta_j(\beta_j + 1) + 2 + \sqrt{\Delta^*} - 3\sqrt{\Delta^a}}{\beta_j^2} \tag{A84}$$

Hence, it proves quite easy to study the variation of $\mathcal{G}$ with respects to p_n as:

$$\frac{\partial \mathcal{G}}{p_n} = \frac{1}{\beta_j^2}\left[\frac{\partial \sqrt{\Delta^*}}{\partial p_n} - \frac{\partial \sqrt{\Delta^a}}{\partial p_n}\right] \tag{A85}$$

Which yields:

$$\frac{\partial \mathcal{G}}{\partial p_n} = \frac{1}{\beta_j^2}\left[\frac{3\beta_j^2}{\sqrt{\Delta^*}} - \frac{3\beta_j^2}{\sqrt{\Delta^a}}\right] \tag{A86}$$

It then suffices to compare Δ^* and Δ^a to know the effect of the price of natural resources on the wedge $\mathcal{G}$. Let $\mathcal{D} = \Delta^* - \Delta^a$.

$$\mathcal{D} = 4\beta_j^2 p_n + 3\beta_j^4 + 2\beta_j^3 + 2\beta_j - \beta_j^2 \tag{A87}$$

Given that $\beta_j \geq 1$ it naturally follows that:

$$\mathcal{D} \geq 0 \tag{A88}$$

Which implies that:

$$\frac{\partial \mathcal{G}}{p_n} \leq 0 \tag{A89}$$

It now remains to be proved that $\mathcal{G}$ is greater or equal to 0 for plausible values of p_n. Namely, let us recall that:

$$\forall k, \Pi_{i,k} = \frac{1}{2} s_j \beta_j^2 - p_n \tag{A90}$$

Given that s_j is a share, it must hold that:

$$\forall k, \Pi_{i,k} \leq \frac{1}{2}\beta_j^2 - p_n \tag{A91}$$

Hence, so that firms may achieve a positive profit in the resource-intensive industry, I need to make the following assumption, which is only a normalization assumption:

$$p_n \leq \frac{1}{2}\beta_j^2 \tag{A92}$$

Meanwhile, let us recall that:

$$\mathcal{G} = \frac{\beta_j^2 + \beta_j + 2 + \sqrt{\Delta^*} - 3\sqrt{\Delta^a}}{3\beta_j^2} \tag{A93}$$

Since it must hold that $p_n \leq \frac{\beta_j^2}{2}$, and given that $\mathcal{G}$ is a decreasing function of p_n, it must be that:

$$\forall p_n, \mathcal{G} \geq \frac{\beta_j^2 + \beta_j + 2 + \sqrt{7\beta_j^4 + 2\beta_j^3 - \beta_j^2 + 4\beta_j + 1} - 3\sqrt{2\beta_j^4 + 2\beta_j + 1}}{3\beta_j^2} \tag{A94}$$

The sign of the former expression is the same as that of the numerator, for the denominator is strictly greater than 0, as the square of a positive number. Hence, let us consider:

$$\mathcal{A} = \beta_j^2 + \beta_j + 2 + \sqrt{7\beta_j^4 + 2\beta_j^3 - \beta_j^2 + 4\beta_j + 1} - 3\sqrt{2\beta_j^4 + 2\beta_j + 1} \tag{A95}$$

It can be shown that:

$$\frac{\partial \mathcal{A}}{\partial \beta_j} = 2\beta_j + 1 + \frac{28\beta_j^3 + 6\beta_j^2 - 2\beta_j + 4}{2\sqrt{7\beta_j^4 + 2\beta_j^3 - \beta_j^2 + 4\beta_j + 1}} - \frac{24\beta_j^3 + 2}{2\sqrt{2\beta_j^4 + 2\beta_j + 1}} \tag{A96}$$

It can be formally shown that:

$$\frac{\partial \mathcal{A}}{\beta_j} \geq 0 \Leftrightarrow \beta_j \in [0, 3] \tag{A97}$$

Meaning that $\mathcal{G}$ is positive for plausible values of β_j in the sense that β_j captures the parameter by which marginal cost is decreased upon successful innovation. That this value lies in the $[1, 3]$ interval sounds a reasonable assumption, in the sense that it allows for large cost reductions in the model.

A.9 Proof of Corollary 3

Corollary 3 states that the knowledge effort in such an economy is inefficiently low, due to spillovers emanating from the knowledge-intensive industry to the resource-intensive one. This is easy to show; let us indeed recall that the overall knowledge effort $\mathcal{K}$ writes:

$$\mathcal{K} = \int_0^{s_j} \kappa_k dk \tag{A98}$$

As κ is uniformly distributed over the $[0, 1]$ interval, it follows that:

$$\mathcal{K} = \frac{s_j^2}{2} \tag{A99}$$

Let us denote $\mathcal{K}^a$ and $\mathcal{K}^*$ the actual and the optimal education efforts. It then follows that:

$$\mathcal{K}^a = s_j^a \tag{A100}$$

And:

$$\mathcal{K}^* = s_j^* \tag{A101}$$

As it was proved earlier that:

$$s_i^a \geq s_i^* \tag{A102}$$

Since $s_j = 1 - s_j$, it naturally follows that:

$$\mathcal{K}^* \geq \mathcal{K}^a \tag{A103}$$

Which is what Corollary 3 supported.

B Data appendix

B.1 Coal data

One potential limitation of the French *cadastre minier* leveraged to build the main coal exposure measure is that it only captures the licenses granted to run coalfields throughout the period under study; there is, however, no absolute certainty that all licenses were actively used throughout their authorization span. And yet, some mitigation points should be clarified. First of all, an official statement made on August, 3^{th} 1810 stipulates that applying for a mining license involves preliminary and thorough inquiries that prove the existence of resources in the underground. Second, among all resources that were mined in France throughout the XIXth and the XXth centuries, coal is the resource for which the number of abandoned projects was the lowest. [3] Third and finally, licenses were often withdrawn when the coalfield had not been actively run for a sufficiently large amount of time.

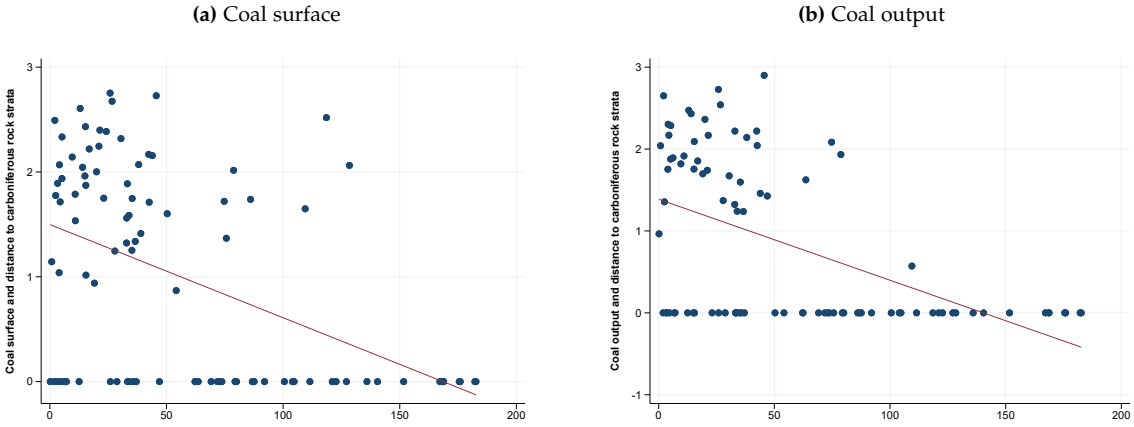
[3] While the proportion of licenses mines which were not actively run reaches 66% for iron resources, it barely reaches 40% for coalfields

C Instrumental variable

C.1 Graphical evidence

I discussed in Section 4 the plausibility of the measures of coal exposure used in the paper. In particular, four measures are introduced, out of which I mostly use two which I call the "continuous" ones to the extent that their values are continuous rather than discrete. Additionally, I introduced two semi-discrete measures based either on the (log) surface of coalfields in a given department or on the (log) coal output in a given department at the onset of the XXth century. I only reported graphical evidence for the validity of the suggested instrument for the two continuous measures of coal exposure and postponed this evidence to the Appendix for the semi-discrete measures. Hence, I report in Figure C1 graphical evidence of a negative relationship between the two semi-discrete measures of coal exposure and distance to carboniferous rock strata. As mentioned earlier, 44 out of the 95 French departments never mined coal over their history which explains the substantial share of zeros in the sample. In spite of this somehow puzzling fact, there exists a clear negative relation between coal exposure and the distance to carboniferous rock strata which, again, lends credence to the argument that coal was more likely to be mined where its deposits remained at the surface at the onset of industrialization.

FIGURE C1: Proximity to coal and distance to carboniferous rock strata



Notes: Figure (a) plots the coal surface when all French coalfields operated by 1900 against the distance to carboniferous rock strata while Figure (b) plots the average (log) coal output at the onset of the XXth century against the distance to carboniferous rock strata.

C.2 Regression evidence

Graphical evidence is indeed not sufficient to assert the validity of our considered instrument for the two semi-discrete measures of coal exposure; therefore, I show

in Table C1 that the instrument is still valid when I restrict my focus to the first semi-discrete measure of coal exposure, namely the (log) surface of coalfields in each department. Clearly, the coefficient associated to the distance to carboniferous rock strata remains negative and significant at the 1% level across all specifications. Also, the magnitude of the coefficient is similar to that derived in Table 1 which summarizes the regression results for the preferred measure of exposure. Adding potential confounders such as geographical variables, geological data or even demographic parameters do not change the estimates.

TABLE C1: Coal surface and distance to carbon rock strata

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Distance to carboniferous rock strata	-0.009***	-0.009***	-0.009***	-0.009***	-0.008***	-0.009***	-0.008***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Coal deposit		0.796***	0.627**	0.579*	0.527	0.437	0.128
		(0.261)	(0.311)	(0.331)	(0.335)	(0.346)	(0.364)
Potato yield			-0.017	-0.043	-0.096	-0.251	-0.088
			(0.298)	(0.305)	(0.319)	(0.323)	(0.284)
Wheat yield			-3.874	-5.821	-4.710	-3.774	-4.363
			(4.111)	(4.382)	(4.889)	(4.747)	(4.400)
Oat yield			3.747	5.339	4.698	4.540	4.699
			(3.829)	(4.382)	(4.516)	(4.312)	(4.067)
Proximity to Mediterranean sea				-0.261	0.085	0.101	0.568 **
				(0.199)	(0.318)	(0.328)	(0.429)
Proximity to Atlantic ocean				-0.092	0.419	0.439	0.190
				(0.123)	(0.305)	(0.316)	(0.309)
Median temperature					-1.272	-1.829	-1.720
					(1.049)	(1.183)	(1.089)
Variation temperature					2.578	4.616*	5.715
					(2.136)	(2.387)	(4.700)
Ruggedness of land						0.107	0.234
						(0.248)	(0.234)
Land surface						-1.901*	-1.905
						(1.082)	(1.236)
Longitude							-0.134
							(0.203)
Latitude							0.244
							(0.83)
(Log) Initial population							0.499**
							(0.233)
Observations	95	95	95	95	95	95	95
R-squared	0.210	0.277	0.297	0.301	0.334	0.363	0.490
F-Stat	37.49	34.08	14.70	11.78	10.28	10.20	12.64

Notes: The dependent variable is the log surface of coal as discussed in 3 while the explanatory variables are mainly the distance to the closest carbon rock strata, the existence of coal deposit within the department and alternative geological and geographic potential confounders. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

I finally perform the same robustness check for the second semi-discrete measure of coal exposure, namely the average (log) output of coal per department at the onset of the XXth century. Results of this complementary robustness check are reported in Table C2 which summarizes the effect of the distance to the closest carboniferous rock strata on the (log) average coal output measure at the onset of the XXth century. Again, the coefficient on the distance to carboniferous rock strata is negative and significant at the 1% level across all specifications, further asserting the validity

of distance to carboniferous rock strata as instrument for the various coal exposure measures. Adding various geological, geographical and demographic controls do not alter the estimates which also remain consistent with the other results derived in Table 1 and in Table C1.

TABLE C2: Coal surface and distance to carbon rock strata

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Distance to carboniferous rock strata	-0.010***	-0.010***	-0.010***	-0.010***	-0.010***	-0.010***	-0.011***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Coal deposit		0.326	0.427	0.360	0.303	0.311	-0.270
		(0.308)	(0.406)	(0.444)	(0.465)	(0.464)	(0.484)
Potato yield			-0.271	-0.216	-0.279	-0.208	-0.009
			(0.323)	(0.350)	(0.319)	(0.323)	(0.334)
Wheat yield			4.750	2.404	2.370	2.052	-1.313
			(4.228)	(5.053)	(5.242)	(5.275)	(4.867)
Oat yield			-4.300	-2.343	-1.981	-1.946	1.026
			(3.914)	(4.597)	(4.728)	(4.796)	(4.534)
Proximity to Mediterranean sea				-0.222	-0.029	0.101	0.390
				(0.208)	(0.351)	(0.391)	(0.423)
Proximity to Atlantic ocean				-0.138	-0.051	-0.074	0.154
				(0.127)	(0.246)	(0.255)	(0.310)
Median temperature					-.757	-0.783	-0.375
					(1.078)	(1.378)	(1.375)
Variation temperature					0.003	-0.423	4.460
					(1.939)	(2.439)	(5.037)
Ruggedness of land						-0.110	-0.093
						(0.294)	(0.263)
Land surface						0.530	3.360**
						(0.974)	(1.376)
Longitude							0.214
							(0.213)
Latitude							0.099
							(0.150)
(Log) Initial population							0.574**
							(0.218)
Observations	92	92	92	92	92	92	92
R-squared	0.261	0.272	0.283	0.297	0.300	0.300	0.490
F-Stat	57.25	28.87	11.67	8.80	10.28	10.20	12.64

Notes: The dependent variable is the log average coal output measure as discussed in 3 while the explanatory variables are mainly the distance to the closest carbon rock strata, the existence of coal deposit within the department and alternative geological and geographic potential confounders. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

D Additional robustness checks

D.1 Robustness on patent quantity

It has been documented in Section 4 that coal exposure induced a negative and significant effect on patenting quantity throughout the XXth century as well as in the beginning of the XXIst century. For the sake of concision, results were only derived using the two preferred measures of coal exposure. I nevertheless complement those results with the coefficients recovered from running Equation 6 where the coal exposure measure is alternatively one of the two semi-discrete measures as discussed in Section 3. Results of this specification are presented in Table D1 which clearly shows that my results are not contingent on one specific measure of coal exposure but rather hold across the different measures discussed in Section 3. In particular, the (log) surface measure and the (log) output measure deliver negative and significant estimates in terms of the effect of coal exposure on patenting activity; also, the estimates are of similar magnitudes with respects to my two preferred measures: they suggest that a one standard deviation increase in coal exposure was conducive of a -12.1 to a -41.7% decrease in patenting activity. For the sake of concision, I do not report the values obtained for the coefficients related to important covariates often mentioned in the main paper - such as, among others, initial demographics, longitude, latitude and distance to Paris - though there is not much to say on them here, given that the recovered estimates from this specification are highly similar to those discussed in the main paper. It is noteworthy that the coefficients obtained through the log output measure of coal exposure are slightly inferior to those recovered through the other specifications which might be explained by a selection issue: indeed, I lack output data for the three departments which still belong to Germany before World War I, namely Moselle, Bas Rhin and Haut Rhin; I thus need to exclude them from the sample when using that specific measure which, in turn, might bias the estimates downwards to the extent that those were important coal regions where the negative effects of exposure to coalfields activities might have been peculiarly large.

D.2 Robustness on patent quality

I showed in the third part of section 4 that not only did coal exposure affect patent quantity, but it also reduced its quality. To show that I ran several specifications of Equation 6 where my main explanatory variables were the two preferred measures of coal exposure. To further demonstrate the robustness of my findings to alternative measures of coal exposure, I report in Table D2 the results of the same regression

TABLE D1: Effect of coal on patent quantity for various specifications

	Pseudo Poisson		OLS		IV	
	Measure 1	Measure 2	Measure 1	Measure 2	Measure 1	Measure 2
Coal exposure	-0.318***	-0.212**	-0.200***	-0.134***	-0.539***	-0.409***
	(0.078)	(0.088)	(0.050)	(0.044)	(0.153)	(0.116)
Observations	10,900	10,465	10,685	10,340	10,685	10,340
Year Fixed-effect	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Geological controls	Yes	Yes	Yes	Yes	Yes	Yes
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.852	0.852	0.619	0.608	0.342	0.343
F-Stat			16.15	11.94	16.30	12.94

Notes: The dependent variable is either patent count (Pseudo Poisson specification) or the (log) number of patents per capita (OLS and IV models) while the explanatory variables are mainly the coal exposure measures, distance to Paris, initial population density, longitude, latitude and other geographical and geological controls discussed in Section 3. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

when using the two semi-discrete measures of coal exposure as discussed in Section 3, and where citations weighted patents are the measure of innovation quality. Three features again stand out: first, the estimates again point to a negative and sizable impact of coal exposure on patenting activity, with magnitudes ranging from -12 to -45% in terms of patenting reduction. Second, these magnitudes are consistent with those presented in the main paper, further asserting the validity of our analysis. Finally, the recovered estimates are, as in the main paper, slightly larger than those obtained when patents are taken as the dependent variable: this implies that the negative influence of coal on patenting activity was magnified for the highest quality patents.

TABLE D2: Effect of coal on citation weighted patents for various specifications

	Pseudo Poisson		OLS		IV	
	Measure 1	Measure 2	Measure 1	Measure 2	Measure 1	Measure 2
Coal exposure	-0.341***	-0.233**	-0.217***	-0.137***	-0.535***	-0.410***
	(0.152)	(0.096)	(0.048)	(0.044)	(0.152)	(0.115)
Observations	10,528	10,528	8,923	8,639	8,923	8,639
Year Fixed-effect	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Geological controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.808	0.807	0.579	0.567	0.201	0.195
F-Stat			10.65	7.85	11.29	8.69

Notes: The dependent variable is either patent count (Pseudo Poisson specification) weighted by the number of citations or the (log) number of patents weighted by the number of citations and divided by department population (OLS and IV models) while the explanatory variables are mainly the semi-discrete coal exposure measure, distance to Paris, initial population density, longitude, latitude and other geographical and geological controls discussed in Section 3. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

I perform the same robustness check when measuring patent quality with the Kelly et al. (2021) measure. Indeed and in line with what was already stressed in Section 4, citations are an imperfect proxy for the measure of innovation quality. Breakthrough

inventions, on the contrary, might be a better proxy for the relevance of an invention: by assessing both the novelty and the meaningful aspect of a patent, this measure allows to detect patents that not only differ of what was done in the past, but that also support further inventions and industrial development. Upon running Equation 6 where the dependent variable is breakthrough patenting activity as measured by Kelly et al. (2021) and using the two semi-discrete measures of coal exposure discussed in Section 3, I get the results summarized in Table D3 which supports the idea that coal exposure was detrimental not only to patent quantity but also to patent quality. Indeed, not only are the estimates on coal exposure measures negative and significant across all specifications, they also are larger than the coefficients derived for total patent activity for almost all specifications, indicating again that the effect of coal exposure was magnified for high-value innovation.

TABLE D3: Effect of coal on breakthrough patents for various specifications

	Pseudo Poisson		**OLS**		**IV**	
	Measure 1	Measure 2	Measure 1	Measure 2	Measure 1	Measure 2
Coal exposure	-0.358***	-0.192*	-0.223***	-0.133***	-0.605***	-0.457***
	(0.088)	(0.101)	(0.054)	(0.049)	(0.153)	(0.115)
Observations	9,678	9,366	8,923	8,716	9,022	8,716
Year Fixed-effect	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Geological controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.837	0.833	0.460	0.443	0.259	0.252
F-Stat			10.82	7.15	15.16	10.60

Notes: The dependent variable is either patent count (Pseudo Poisson specification) weighted by the number of citations or the (log) number of patents weighted by the number of citations and divided by department population (OLS and IV models) while the explanatory variables are mainly the semi-discrete coal exposure measure, distance to Paris, initial population density, longitude, latitude and other geographical and geological controls discussed in Section 3. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

D.3 Robustness on specialization

I discussed in the fourth subsection of Section 4 the impact of coal exposure on the specialization of departments across different technology classes. Namely, I showed that patenting activity was all the more negatively affected as technology classes were not related to industries that were favored in the early XXth century by abundant coal resources. And yet, out of concision, I only reported the estimates of Equation 6 for the main measure of proximity to coalfields, and postponed the presentation of results related to the measure of proximity indexed on large coalfields. Leveraging this second measure of coal exposure actually does not change the results substantially: while the estimates for plane, computing and information inventions are still negative, significant, and much larger than the coefficients derived for total patenting

activity, the coefficients on respectively textile, mining industry and car-related products are either positive or negative, but never statistically significant. This, in turn, suggests that the effect of coal exposure was heterogeneous across technology classes, depending both on their novelty and on their distance to industries related to coal.

TABLE D4: Effect of coal on innovation specialization

	Textile	Mining	Vehicles	Aeronautics	Computing	Information
Coal exposure	0.139	-0.326	-0.198	-1.188***	-0.634*	-0.634**
	(0.133)	(0.306)	(0.290)	(0.297)	(0.337)	(0.319)
Observations	10,810	10,600	10,810	10,810	10,716	10,810
Year Fixed-effect	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Geological controls	Yes	Yes	Yes	Yes	Yes	Yes
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.673	0.550	0.660	0.709	0.759	0.745

Notes: The dependent variable is the number of patents related to each category mentioned while the explanatory variable is the proximity to large coalfields as discussed in Section 3; geographical and geological controls pertain to the covariates discussed in Section 3. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

D.4 Robustness on breakthrough patents

My second key finding in terms of main results is that not only did coal exposure affect patent quantity, but it also decreased their quality. To support this statement, I notably leveraged the pioneering work by Kelly et al. (2021) who introduced a novel measure of patent quality by bringing forward the concept of breakthrough inventions. In their opinion, a patent is breakthrough provided that it is both novel and meaningful. Novel means that the patent semantically differs from past patents, while meaningful implies that its content will be used in future research and development. Drawing on their dataset and as explained in Section 4, I computed for each year and 3-digit IPC class the share of patents labeled as breakthrough by Kelly et al. (2021). I then defined as breakthrough inventions all patents belonging to the top 10 IPC class in terms of frequency of inventions classified as breakthrough. Two criticisms could, however, be heaped on that methodology: first, one could argue that patents which were breakthrough in the United States were not necessarily breakthrough in other countries at the same time; second, one could fear that the definition of breakthrough patenting activity is excessively sensitive to the number of technology classes which I retain as breakthrough. To alleviate these concerns, I perform different robustness checks the results of which are summarized in the following paragraphs.

As mentioned, the first concern could be that breakthrough patents are not necessarily the same at the same moment across the globe. One technology that is peculiarly ad-

vanced in the United States might already be largely implemented in other countries or the converse might also hold. In that respect, indexing breakthrough inventions in France at time t on breakthrough inventions in the US at the same moment might be misleading. To ensure that my baseline estimates adequately measure breakthrough inventions, and that my results are not driven by a measurement error, I allow for a timing difference in the diffusion of breakthrough techniques across the US and France. More specifically, I relax the assumption that breakthrough inventions spread at the same speed across the United States and France and allow for lags of 5, 10 and 20 years for such a diffusion. Hence, I define as a breakthrough technology class at time t in France an invention belonging to the 3-digit IPC class that were "breakthrough" in the United States respectively 5 or 10 years before and check that the estimates are still significant and larger than those derived in the baseline framework. The results of the robustness check when the lag if fixed at 5 years are pictured in Table D5. Clearly, my main estimates are not driven by a potential 5 years-lag between the timing of breakthrough inventions in the US and that in France; indeed, the coefficients on coal exposure are all significant either at the 5% or at the 1% level, and are rather magnified than reduced under this new specification with increases hovering around 5 p.p.

TABLE D5: Effect of coal on breakthrough patents, lag = 5

	Pseudo Poisson		**OLS**		**IV**	
	Measure 1	Measure 2	Measure 1	Measure 2	Measure 1	Measure 2
Coal exposure	-0.372**	-0.434***	-0.217***	-0.273***	-0.648***	-0.646***
	(0.177)	(0.167)	(0.080)	(0.092)	(0.176)	(0.178)
Observations	9,682	9,682	9,060	9,060	9,060	9,060
Year Fixed-effect	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Geological controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.821	0.822	0.462	0.467	0.266	0.219
F-Stat			8.93	8.35	11.17	9.39

Notes: The dependent variable is either the count of breakthrough patents as defined by the 5-year lag explained above (Pseudo Poisson specification) or the (log) number of patents weighted by the number of breakthrough patents as defined by the 5 year lag explained above and divided by department population (OLS and IV models) while the explanatory variables are mainly the coal exposure measure, distance to Paris, initial population density, longitude, latitude and other geographical and geological controls discussed in Section 3. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

I also test for the same robustness check, but applying an even more cautious 10 years lag, assuming that French innovation was, on average, lagging 10 years behind the US pace throughout the XXth century. Results of this complementary falsification test are summarized in Table D6 which further lends support to the idea that my main results are not driven by a framing issue, whereby I would not be capturing breakthrough innovation in France because of a timing issue. Lagging the classification of IPC categories in terms of their breakthrough propensity based on a 10-year lag with

respects to the US even gives slightly larger estimates across all specifications, with coefficients being significant at the 5% level or even at the 1% level in the majority of cases. Hence, this additional test strongly confirms that coal exposure not only affected innovation quantity but also innovation quality across various measures.

TABLE D6: Effect of coal on breakthrough patents, lag = 10

	Pseudo Poisson		**OLS**		**IV**	
	Measure 1	Measure 2	Measure 1	Measure 2	Measure 1	Measure 2
Coal exposure	-0.376**	-0.444***	-0.209***	-0.288***	-0.677***	-0.675***
	(0.185)	(0.171)	(0.080)	(0.096)	(0.181)	(0.180)
Observations	9,687	9,687	9,120	9,060	9,060	9,060
Year Fixed-effect	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Geological controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.812	0.814	0.447	0.462	0.253	0.296
F-Stat			8.79	8.32	10.53	9.10

Notes: The dependent variable is either the count of breakthrough patents as defined by the 10-year lag explained above (Pseudo Poisson specification) or the (log) number of patents weighted by the number of breakthrough patents as defined by the 10 year lag explained above and divided by department population (OLS and IV models) while the explanatory variables are mainly the coal exposure measure, distance to Paris, initial population density, longitude, latitude and other geographical and geological controls discussed in Section 3. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

A second criticism which could be formulated regarding my definition of breakthrough patenting activity is the threshold which I apply to detect breakthrough IPC classes. Namely, it has been said in the main paper that, for each year, I define as a breakthrough technology class one which belongs to the top 10 3-digit IPC categories in terms of frequency of the number of breakthrough patents. Yet, it could be argued that this ranking is somehow arbitrary and that my estimates might critically hinge on this 10 threshold. To speak to these issues, I implement two robustness checks in the way I define breakthrough technology classes, whereby I respectively make the classification either more stringent or more flexible. In concrete terms, rather than retaining only the 10 first technology classes as breakthrough ones, I respectively retain only the first 5 ones or the first 20 ones and show that the main findings derived in Section 4 are still valid. Table D7 first shows the results when restricting breakthrough inventions to the top 5 3-digit IPC categories in terms of yearly share of breakthrough inventions. While the estimates are slightly lower than those derived under the main specification, they are still all statistically significant either at the 5% or at the 1% level, and remain larger than the estimates obtained for total patenting activity. This lends credence to the argument that our results are not driven by a specific definition of breakthrough patenting activity, but that they rather hold under flexible assumptions.

As the last robustness check in this subsection, I replicate the table above using a wider definition of breakthrough inventions. Namely, I allow the top 20 IPC categories each year to proxy for breakthrough inventions and estimate equation 6 upon using

TABLE D7: Effect of coal on breakthrough patents, only top 5

	Pseudo Poisson		OLS		IV	
	Measure 1	Measure 2	Measure 1	Measure 2	Measure 1	Measure 2
Coal exposure	-0.349**	-0.425***	-0.226***	-0.261***	-0.617***	-0.607***
	(0.163)	(0.158)	(0.076)	(0.093)	(0.164)	(0.168)
Observations	9,678	9,678	8,137	8,137	8,137	8,137
Year Fixed-effect	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Geological controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.806	0.807	0.401	0.401	0.249	0.270
F-Stat			7.23	6.80	9.46	7.51

Notes: The dependent variable is either the count of breakthrough patents as defined by the top 5 measure explained above (Pseudo Poisson specification) or the (log) number of patents weighted by the number of breakthrough patents as defined by the top 5 measure explained above and divided by department population (OLS and IV models) while the explanatory variables are mainly the coal exposure measure, distance to Paris, initial population density, longitude, latitude and other geographical and geological controls discussed in Section 3. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

this new dependent variable. Recovered estimates are presented in Table D8 which show that exposure to coal translated in a negative and significant effect on patenting activity, even when allowing for a more flexible definition of breakthrough patenting. At the same time, it is noteworthy that the coefficients are of a smaller magnitude than those usually derived when applying other measures of breakthrough patenting activity. This finding deserves two comments: first, the estimates remain larger across almost all specifications than those obtained when using total patenting activity as the dependent variable, suggesting in turn that breakthrough inventions, even when defined with a very flexible way, were more negatively affected by coal exposure than global inventions. Second, it is no surprise that a broader definition of breakthrough inventions leads to lower estimates to the extent that such a larger measure might encompass technology classes which are further away from the technology frontier and hence less qualitative.

D.5 Are the results driven by Paris?

A natural question that arises in this kind of framework is whether the results are not driven by a specific unit in the sample. Indeed, it could be that the negative effect of coal on patenting activity is either driven by a member of the treatment group which would exhibit specific characteristics unrelated to coal extraction, or by a large enough member of the control group that, for some reasons, would exhibit abnormally large innovation performances. While I control for numerous geological, demographic and geographical factors which, normally, would reflect and reduce the importance of extreme values within either treated or control units, I cannot rule out such issues given that I don't allow for fixed effects since my treatment is a time-invariant vari-

TABLE D8: Effect of coal on breakthrough patents, only top 20

	Pseudo Poisson		OLS		IV	
	Measure 1	Measure 2	Measure 1	Measure 2	Measure 1	Measure 2
Coal exposure	-0.347**	-0.436***	-0.190**	-0.248***	-0.571***	-0.578***
	(0.146)	(0.142)	(0.079)	(0.091)	(0.170)	(0.171)
Observations	9,678	9,678	9,383	9,383	9,383	9,383
Year Fixed-effect	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Geological controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.847	0.849	0.489	0.494	0.230	0.325
F-Stat			9.13	8.97	11.92	10.38

Notes: The dependent variable is either the count of breakthrough patents as defined by the top 20 measure explained above (Pseudo Poisson specification) or the (log) number of patents weighted by the number of breakthrough patents as defined by the top 20 measure explained above and divided by department population (OLS and IV models) while the explanatory variables are mainly the coal exposure measure, distance to Paris, initial population density, longitude, latitude and other geographical and geological controls discussed in Section 3. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

able. A good candidate for such an outlier that could drive the estimates is Paris: the city-department indeed concentrated around 50% of French total patenting activity throughout the XXth century. I thus re-run all equations discussed above but excluding Paris out of the sample, and show that the estimates remain negative, significant, and of the same range of magnitudes as that obtained in the general framework. The results of this robustness check are indeed summarized in Table D9 for the two main measures of coal exposure, namely the average proximity to all French coalfields and the average proximity to the largest coalfields. Clearly, all coefficients remain significant either at the 10% level (for one specification, namely the OLS specification using the first measure of coal exposure) or at the 1% level (for all other estimates) implying that my baseline results are not driven by the specific case of Paris.

TABLE D9: Effect of coal on patenting activity, excluding Paris

	Pseudo Poisson		OLS		IV	
	Measure 1	Measure 2	Measure 1	Measure 2	Measure 1	Measure 2
Coal exposure	-0.416***	-0.535***	-0.150*	-0.228***	-0.536***	-0.524***
	(0.164)	(0.130)	(0.082)	(0.085)	(0.191)	(0.181)
Observations	10,695	10,695	10,570	10,570	10,570	8,716
Year Fixed-effect	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Geological controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.709	0.718	0.580	0.587	0.265	0.315
F-Stat			10.51	11.63	10.74	10.91

Notes: The dependent variable is either patent count (Pseudo Poisson specification) or the (log) number of patents and divided by department population (OLS and IV models) when excluding Paris out of the sample while the explanatory variables are mainly the semi-discrete coal exposure measure, distance to Paris, initial population density, longitude, latitude and other geographical and geological controls discussed in Section 3. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

D.6 Robustness on education

As highlighted in Section 5, education is a key mechanism accounting for the negative impact of coal exposure on patenting activity. More specifically, I have been arguing that coal extraction negatively affected the upper-tail of the education distribution as soon as the onset of the XXth century and used three measures of coal exposure to derive that prediction. As a robustness check, I document that this negative and causal effect still holds when considering the fourth measure of coal exposure, namely the (log) average output of coal at the onset of the XXth century; I complement this robustness check by showing that the negative impact of coal still holds when restricting the coal measure to a binary variable taking the value 1 in case at least one coalfield was to be found within the department at the beginning of the period under study in this sample, namely 1885. Results summarized in Table D10 clearly suggest that the negative effect of coal on the upper tail of the education distribution holds irrespectively of the measure of coal exposure I retain. Besides, the magnitude of the coefficients is similar to those derived in the main paper. It is however noteworthy that the gap between the OLS and the IV estimate is larger for the binary specification which might be explained by the substantially larger standard deviation associated with the binary IV outcome; further more, the discrepancy between OLS and IV estimates, which is discussed extensively in Appendix XX, might be attributed to the fact that coalfields still open in the course of XIXth century and even in the beginning of the XXth century; as coalfields were more likely to start being ran in departments where coalfields already were mined, the OLS estimates are likely to be biased towards zero.

TABLE D10: Effect of coal on baccalauréat rate, 1885-1914

	OLS		IV	
	Measure 1	Measure 2	Measure 1	Measure 2
Coal exposure	-0.103***	-0.120*	-0.159***	-0.451***
	(0.088)	(0.069)	(0.056)	(0.150)
Observations	1,789	1,789	1,789	1,789
Year Fixed-effect	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes
Geological controls	Yes	Yes	Yes	Yes
R-squared	0.440	0.424	0.212	0.151
F-Stat	5.49	4.38	5.56	4.43

Notes: The dependent variable is the log rate of baccalauréat in the 1885-1911 period while the explanatory variables are respectively the fourth measure of coal exposure (output based) or a binary variable taking the value 1 if at least one coalfield is located within the department, distance to Paris, initial population density, longitude, latitude and other geographical and geological controls discussed in Section 3. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

D.7 Robustness on industry specialization

In the second part of Section 5, I documented a negative and significant impact of coal exposure on the share of high-skilled jobs at the local level, as soon as the onset of the XXth century. To support these effects, I however only used three out of the five measures I have to proxy for coal exposure. Namely, I used neither the output-based measure nor the binary variable; to further show the impact of coal exposure on the likelihood of occupying a high-skilled job, I test the same equation as in Table 6 but substituting the last two measures of coal exposures for those used in the main paper. The results of this additional robustness check are presented in Table D11 which summarizes the effect of coal exposure on the share of *professions libérales* as measured by the seven censuses spanning the 1901-1936 period. Again, coal exposure induced a negative and significant effect on the extent of high-skilled occupations at the department level, lending further credence to the argument that coal extraction was detrimental to high-value added industries.

TABLE D11: Effect of coal on *professions libérales*, 1901-1936

	OLS		**IV**	
	Measure 1	Measure 2	Measure 1	Measure 2
Coal exposure	-0.080***	-0.137*	-0.195***	-0.494***
	(0.029)	(0.077)	(0.058)	(0.158)
Observations	581	581	581	581
Year Fixed-effect	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes
Geological controls	Yes	Yes	Yes	Yes
R-squared	0.335	0.325	0.223	0.165
F-Stat	5.69	5.19	6.79	7.38

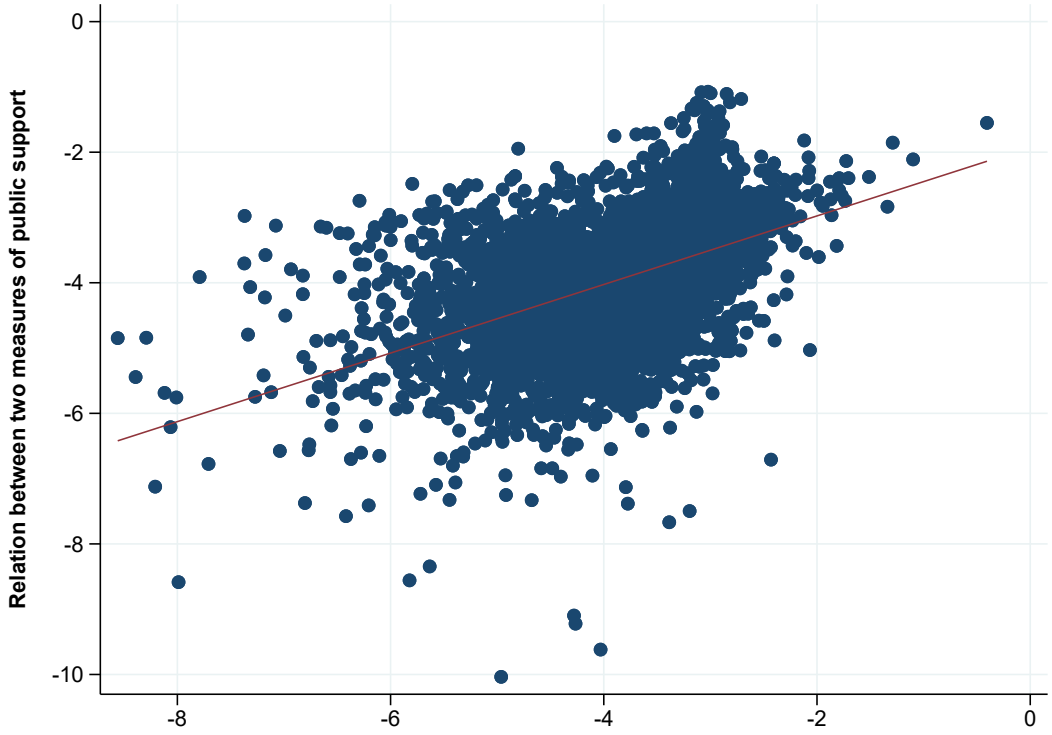
Notes: The dependent variable is the log share of liberal professions in the 1901-1936 period while the explanatory variables are respectively the fourth measure of coal exposure (output based) or a binary variable taking the value 1 if at least one coalfield is located within the department, distance to Paris, initial population density, longitude, latitude and other geographical and geological controls discussed in Section 3. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

D.8 Robustness on public support

I discussed in the fourth subsection of Section 5 the role played by public support in maintaining inequalities among departments in terms of innovation abilities. In particular, I built two measures of exposure of departments towards public investment either in R&D or in specific industrial sectors and leveraged those measures to show (i) that coal exposed departments were less likely to receive support on the part of the French government and (ii) that such a support was conducive of a significant increase

in patenting activity for targeted sectors and departments. Some robustness checks were however postponed to the appendices, and I briefly describe them here. In particular, Figure D1 provides evidence of the consistency between the two measures of exposure to public funding, suggesting that those measures adequately capture the support received by French departments throughout the period under study. Technically speaking, the coefficient of correlation between both variables is as high as 50%, providing strong support to the argument that both measures adequately capture the exposure of a department to public support.

FIGURE D1: Consistency between two measures of exposure to public support

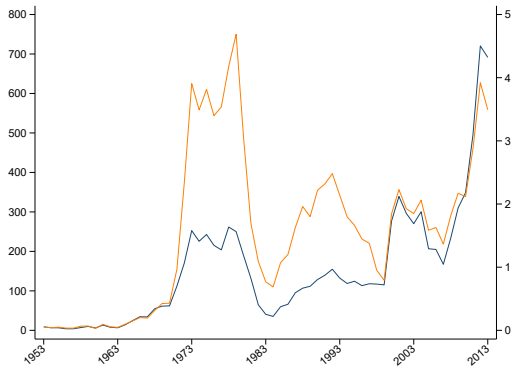


Notes: This graph plots the value of the (log) second measure of exposure to public support against the (log) first measure of exposure to public support.

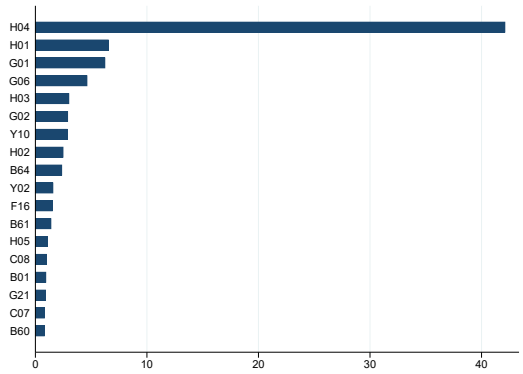
These two measures are all the more consistent as they seem to correspond to similar technology classes and industries. Among the three digit IPC classes most targeted by public investment as proxy by the *"grand champions"* policy, I indeed find that G01 and H01 also rank among the most prominent technology classes in terms of exposure to public funding, right after H04 which corresponds to electric communication techniques. In terms of evolution across time, I also find a sizable increase in the role of public support to specific industries, a role which peaked in the late 70s when the share of public-related firms in total patenting activity was as large as 5%, a finding, again, consistent with previous estimates derived in Section 5.

FIGURE D2: Public R&D support: time evolution and industry specialization

(a) Increase in public R&D support, 1953-2013



(b) Specialization of public support to R&D, 1953-2013

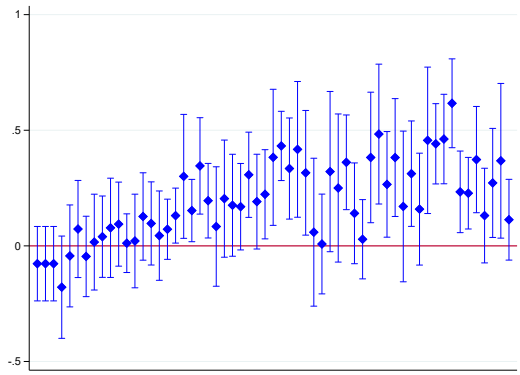


Notes: Figure (a) plots the evolution of public involvement in R&D while Figure (b) plots the technology classes most heavily targeted by public support to R&D.

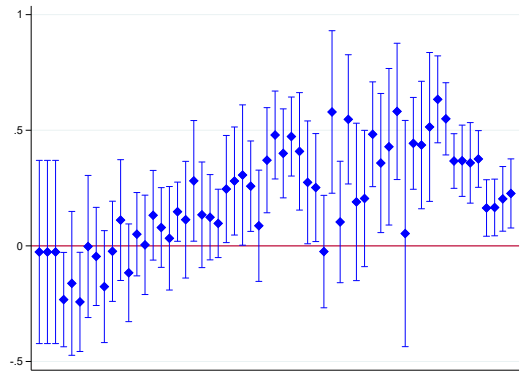
As an illustration of the increasing role played by public support in shaping the quantity and the direction of innovation, I report in Figure D3 the estimates obtained when running cross section regressions of patenting activity on the two measures of exposure to public support, allowing for the usual geographical, demographic and geological potential confounders. The interesting features of those estimates is that they show a substantial increase in the effect of industrial policy starting in the early 70s, an effet which pursued throughout the 2008 financial crisis. Not all coefficients are statistically significant, also because of the large standard deviations induced by the small sample size; overall however, these estimates are consistent with a sizable impact of public intervention in shaping the innovation effort after World War II.

FIGURE D3: Magnitude of public spillovers across time, 1953-2013

(a) Effect of public R&D support, 1953-2013



(b) Effect of *"grands projets"*, 1953-2013



Notes: Figure (a) plots the effect of public involvement in R&D on patenting activity while Figure (b) plots the effect of the *"grands projets"* policy on patenting activity. Estimates are derived from separate cross section regressions for all years from 1953 to 2013. Confidence intervals are shown at the 5% level.

D.9 Allowing for region-fixed effects

Across the main paper, I always presented equations where only year fixed effects were introduced and not department fixed effects. This empirical specification stems from the definition of the main covariate, namely the coal exposure measure, which is time-invariant and thus does not allow for department-level fixed effects. Instead, I complemented my dataset with many geographical, geological and demographic controls to ensure that the estimates were not driven by department-level idiosyncratic characteristics. Similarly, the IV strategy should ensure a neat identification strategy irrespectively of local confounding factors. Yet, it could be feared that I still omit a department-level important variable which would correlate the ability to invent new features. Beyond the structural reason for not including department fixed effects, it sounds noteworthy to recall that *region* is not a time invariant unit across time in France. More specifically, regions did not exist, at least from an administrative standpoint, until 1962, meaning roughly in the middle of my sample. Hence, no unified and homogeneous policy was implemented at that scale before at least 1962. Second, since then, the concept of region has not remained stable as a reform enacted in 2015 changed the boundaries of regions, agglomerating them in larger groups than before. In that respect, allowing for region fixed-effects is both a demanding robustness test and somehow a puzzling one, given that it pertains to entities that moved across time. To speak to concerns related to local idiosyncratic characteristics, I however run my main regressions, namely Equation 6, allowing for region fixed effects and summarize the results in Table D12. Again, the main estimates are unaffected by this exercise and all remain significant either at the 10% or, most of the time, at the 1% level. Also, all estimates are similar in terms of magnitude to those obtained in the general framework where no geography-based fixed effects are allowed.

TABLE D12: Effect of coal on patent quantity for various specifications, allowing for region F-E

	Pseudo Poisson		**OLS**		**IV**	
	Measure 1	Measure 2	Measure 1	Measure 2	Measure 1	Measure 2
Coal exposure	-0.258*	-0.372***	-0.293***	-0.249***	-0.747***	-0.490***
	(0.136)	(0.139)	(0.054)	(0.086)	(0.231)	(0.122)
Observations	10,810	10,810	10,685	10,685	10,685	10,685
Year Fixed-effect	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Geological controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.870	0.872	0.660	0.645	0.198	0.269
F-Stat			24.50	16.03	10.14	15.16

Notes: The dependent variable is either patent count (Pseudo Poisson specification) or the (log) number of patents per capita (OLS and IV models) while the explanatory variables are mainly the coal exposure measure, distance to Paris, initial population density, longitude, latitude and other geographical and geological controls discussed in Section 3. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

D.10 Alternative clustering methods

In the main paper, I cluster standard errors at the department level since my basic unit of observation is the year $\times$ department cell. It could however be that there exist economic spillovers within regions and across departments, in which case the nuisance terms could be correlated not only within departments across time but rather within region across time. Again, this assumption is not very likely to hold to the extent that department is a much more stable unit of treatment across time in France. To test for this assumption, I however again run Equation 6 where I cluster standard errors at the region level rather than at the department level and summarize the main findings of this robustness exercise in Table D13. Again, all estimates on the coal exposure measure remain negative and significant either at the 5% or at the 1% level, further supporting the idea that the results hold under a wide variety of specification and functional forms.

TABLE D13: Effect of coal on patent quantity for various specifications, clustering at the region level

	Pseudo Poisson		**OLS**		**IV**	
	Measure 1	Measure 2	Measure 1	Measure 2	Measure 1	Measure 2
Coal exposure	-0.350**	-0.432***	-0.179**	-0.248**	-0.544***	-0.556***
	(0.148)	(0.146)	(0.077)	(0.089)	(0.166)	(0.177)
Observations	10,810	10,810	10,685	10,685	10,685	10,685
Year Fixed-effect	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Geological controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.850	0.852	0.609	0.615	0.351	0.387
F-Stat			36.43	28.71	60.09	81.94

Notes: The dependent variable is either patent count (Pseudo Poisson specification) or the (log) number of patents per capita (OLS and IV models) while the explanatory variables are mainly the coal exposure measure, distance to Paris, initial population density, longitude, latitude and other geographical and geological controls discussed in Section 3. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

D.11 Dynamic evolution of the effect over time

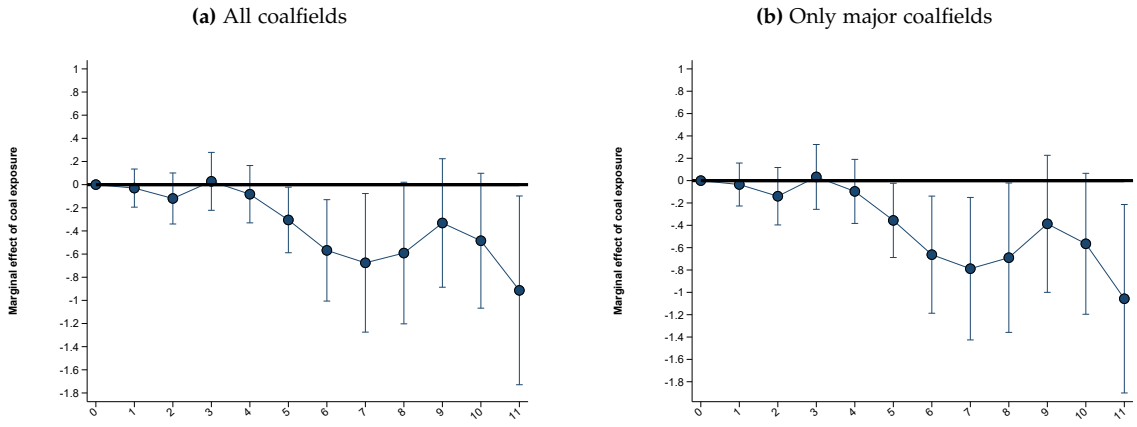
Figure 2 supported the idea that the effect of coal exposure on innovation was persistent over time; more importantly, it suggested that the magnified over time. A possible explanation for this finding will be brought forward by Section 5 which discusses plausible mechanisms for the negative impact of coal on patenting activity. Yet, it might prove useful to check whether coal exposure might account for the widening gap in innovation abilities throughout the XXth century and the beginning of the XXIst century. To evaluate this question, I collapse the yearly number of patents at the decade level for each department, and then run a new version of Equation 6 where I

interact decade fixed-effects with the coal exposure measure. This new specification, which I formally summarize in Equation D1, enables me to capture the time-varying effect of coal exposure on patenting activity:

$$\text{pat}_{d,t} = \sum_{k=1}^{11} \mathbb{1}_t(k) \times \beta_k e_d + \alpha_t + \eta_d + \varepsilon_{d,t} \tag{D1}$$

Where η_d captures department fixed effects; indeed, by contrast with Equation 6 which does not include department fixed effects, given that the main explanatory variable is time-invariant, I know have a time varying explanatory variable given that I interact the time-invariant coal exposure measure with year fixed effects. I then plot the results of this specification for the two main coal exposure measures in Figure D4 which suggests that the negative impact of coal exposure on patenting activity accelerated over time; more specifically, coal exposure exerted an increasingly negative effect over time with respects to its effect during the first decade of the XXth century. This, of course, does not imply that the effect did not already start prior to World War 2: indeed, running Equation 6 on the subset of patents before WW2 also yields negative and significant effects of coal exposure on patenting activity; yet, these two novel graphs support the idea that coal exposure was all the more negative starting in the second half of the XXth century, consistently with the mechanisms detailed in the next section.

FIGURE D4: Dynamic effect of coal exposure



Notes: Figure (a) plots the dynamic effect of coal exposure when coal exposure is defined as the proximity to all coalfields while Figure (b) plots the dynamic effect of coal exposure when coal exposure is computed restricting the set of coalfields to historical major coalfields documented in the ERIH project. For both panels, the coal exposure measure is instrumented with the carboniferous rock strata distance, as described in Section 4