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

Jean-Baptiste Chaniot

CREST and Collège de France

February 2026

Abstract

This paper argues that abundant natural resources exert a negative impact on productivity and innovation in the long-run. I first develop a stylized model of endogenous growth featuring an extractive industry using unskilled labor and an intermediary industry incorporating skilled labor. Innovation stems from private decision by skilled workers who capture monopoly rents in the intermediary industry. The model predicts that an exogenous increase in the availability of natural resources raises profitability of the extractive industry and, hence, lowers the incentive to innovate. This negative effect is reinforced by the negative impact of natural resources on human capital formation as well as on industry structure. I then use a French dataset spanning the entire twentieth century to document a negative and sustained effect of exposure to coalfields on both the quantity and the quality of patents. Consistently with the theoretical argument, I further show that this negative pattern might be accounted for by two main mechanisms, namely a negative impact on human capital as well as a technical trap.

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

*I thank Philippe Aghion, Pierre Boyer, Roland Rathelot, Florian Grosset, Etienne Fize, Simon Bunel, Arthur Guillouzouic, 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 (jeanbaptiste.chaniot@ensae.fr)

1 Introduction

As of 1922, the French department of Nord ranked 10 among French departments in terms of GDP per capita and 8 in terms of patenting activity per capita, a performance which mostly stemmed from its strong industrial basis. By the beginning of the XXIst century, the economic outlook of Nord had nevertheless dramatically deteriorated: as of 2016, Nord was ranking, respectively, 85 and 63 in terms of GDP per capita and of patenting activity per capita. Meanwhile, unemployment in Nord stood at a substantially higher level (8.8%) compared to the national average (7.3%). Similarly, Belfort used to be one of the most productive French departments: indeed, it systematically ranked in the top 5 of French departments in terms of patenting activity per capita at the onset of the XXth century; as of 2015 however, it only ranked 28 according to innovations per capita. Interestingly enough, Nord and Belfort are two departments which both relied on abundant coal resources and hence developed a large extractive industry in the course of the XIXth century. The pattern of reversal of fortune they seem to have experienced fits in the broader debate regarding the impact of natural resources on economic growth and productivity in the long run.

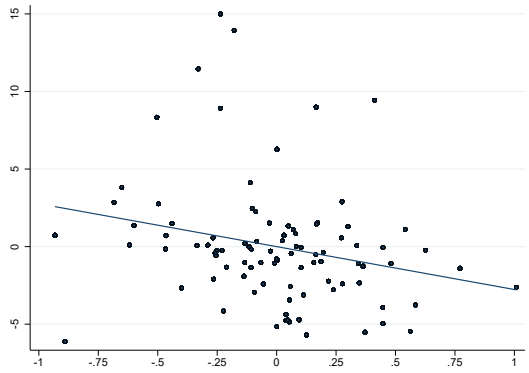
On the one hand, coal resources are thought to be a key ingredient of economic development, at least since the pioneer analysis by Pomeranz (2014), who argues that living standards in Europe and in the most populated areas of China were not significantly different at the end of the XVIIIth century. In 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. More recently, an extensive corpus of the literature has investigated the concept of "dutch disease" in reference to the negative effects of gas discovery in the Netherlands in the sixties.

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.

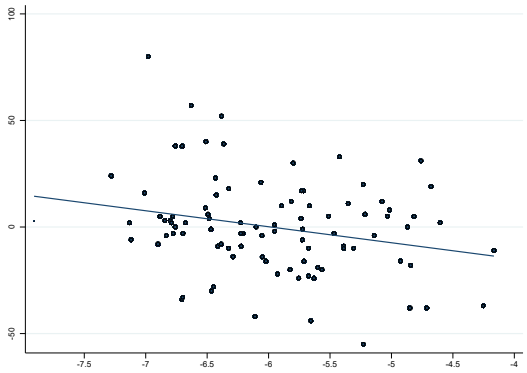
(2024), Allcott and Keniston (2017)). However, despite increasing interest related to the long-term effects of coal extraction, little attention has been paid to the impact of the endowment of natural resources on innovation in the long run. Yet, understanding the effect of resource endowment on innovation seems crucial in light of Figure 1 which suggests a negative correlation between productivity and exposure to coal throughout the XXth century. More specifically, Figure 1a shows the growth rate in patenting activity between 1914 and 2014 and exposure to coalfields, while Figure 1b captures the relation between the evolution of the nationwide ranking of each French department in terms of patenting activity and exposure to coalfields; intuitively, a negative value means that the relative innovative ability of a given department deteriorated over time, While these plots do not provide causal evidence, they both suggest that departments abundantly endowed with coal resources displayed lower productivity growth rates during the XXth century.

FIGURE 1: Evolution of productivity *versus* exposure to coal

(a) Patenting evolution vs. coal



(b) Patent ranking vs. coal



Notes: Figure (a) plots the evolution in patenting activity with respect to average proximity to coalfields while Figure (b) plots the evolution the ranking in terms of patenting activity among departments. Source: Bergeaud and Verluise (2024), Bonnet et al. (2021)

In order to assess the effect of natural resource on patenting activity, I first elaborate a highly stylized endogenous growth model, where a consumption good industry generates output using both an intermediary good and an extractive good. The intermediary good is produced by a monopolist using skilled labor while the extractive good is produced competitively using unskilled workers whose productivity level positively depends on the availability of natural resources. At each period in time, one skilled worker is randomly drawn to achieve an innovation which consists in introducing a new variety of the intermediary good; if the innovation effort achieved by the inventor is successful, the productivity level of the economy permanently increases. In equilibrium, the impact of natural resources on innovation depends on three contradictory forces; on the one hand, large natural resources positively affect incentives to innovate through a *market size* effect: the larger natural resources, the

larger the output of the consumption good industry which raises the demand for intermediary goods and, hence, the profitability of inventions. In the meantime, large natural resources increase the wage rate for unskilled workers and, hence, for skilled workers which reduces the profitability of the intermediary good industry and, hence, of achieving innovation effort, an effect which I call the *substitution* effect; meanwhile, large natural resources divert workers from the skilled industry which lowers the knowledge stock in the economy and, thus, the effectiveness of research, which might be termed the *human capital* effect. The model then predicts an ambiguous impact of natural resources on productivity, for the *market size* effect generates a positive effect while the other effects, namely the *substitution* and the *human capital* effects work in the opposite direction. Overall, the model suggests that there exists a cutoff value above which natural resources are detrimental to innovation; this cutoff value is all the lower as natural resources barely contribute to the production process of consumption goods, an assumption which is likely to hold for the case of coal in the French context throughout the XXth century.

These predictions are then tested empirically using a newly assembled dataset that spans French departments from 1907 to 2016. 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 characteristics 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 as-good-as-random heterogeneity in geological components of soil surfaces to instrument for the location of coalfields. 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); said differently, not only did extractive industries affect the quantity of patents, they also reduced their quality. 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 then examine plausible mechanisms that may explain such a long-run effect: more specifically, I bring evidence that extractive industries are conducive of lower human capital accumulation, in particular in the extreme upper-tail of the distribution; I also find support in favor of a path dependence mechanism, namely the fact that departments which heavily invested in extractive industries failed to diversify in alternative industries or technologies, a process which was reinforced by individual level behavior. I exam-

ine complementary mechanisms such as the reversal of fortune hypothesis, the early industrialization effect or the institutional argument.

Related literature My paper first 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 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; later on, Krugman (1987) further elaborated on the dutch disease framework to suggest that a comparative advantage in natural resources might be detrimental in the long run. 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. In the American context, Matheis (2016) argue that increased coal production results in positive effects on demographic and industry-level outcomes in the short-run, but adverse effects in the long run. These findings are consistent with the analysis by Black et al. (2005) who exploit booms in the coal industry to identify a negative impact of exposure to extractive industries on human capital; more specifically, they suggest that such booms drive the returns to school down which, in the long run, might negatively affect the stock of knowledge. Finally, a nascent strand of the literature focuses on the political economy implications of resources discovery; Mehlum et al. (2006) was the first to suggest that natural resources *per se* were not necessarily an evil, but that their impact on economy and society depended on the institutional framework in which they were extracted. Empirical evidence followed this seminal paper, delivering mixed evidence. Important contributions such as that from Vicente (2007), who uses a natural experiment in West African, and that from Caselli and Michaels (2013), who considers the case of Brasilian municipalities, suggest that natural resources may drive corruption or illegal activities. While this literature departs from my main goals, I will nevertheless explore potential mechanisms related to institutional development. Overall, my paper contributes to this literature in a threefold way. First, I focus specifically on patent activity, both from a quantitative and a qualitative point of view, leveraging the richness and 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 Romer (1990), and later, on, by Aghion and Howitt (1992) who argue that growth results from technological progress which, itself, stems from competition among research firms that generate innovations. 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. Other important contributions to assess the impact of human capital on long run innovation include those from Valero and Van Reenen (2019), Squicciarini (2020), Squicciarini and Voigtländer (2015) or Dittmar and Meisenzahl (2022). From a political standpoint, Acemoglu et al. (2001a) 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. (2001a) 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. A last strand of this rich body of literature emphasizes the role of governments in shaping industrial and innovative activities in the long run: recent contributions in that field include those from Gross and Sampat (2020), Dyevre (2023) and Bergeaud et al. (2022). I contribute to this literature by providing empirical evidence of the negative influence of natural resources on productivity; 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 extensively document 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 is due to 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. (2012) develop a theoretical model in which firms with a stronger 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). 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, namely cities where a small set of firms 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, so that 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 mechanisms 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 4 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 while Section 6 finally concludes.

2 The model

2.1 Set-up

The model is closely related to the canonical endogenous growth model by Aghion and Howitt (1992) with some simplifying assumptions, though. Basically, there is a continuum L of infinitely-lived individuals with identical inter-temporally additive preferences defined over lifetime consumption, and the constant rate of time preference $r > 0$. The marginal utility of consumption is assumed constant.

There is no dis-utility from supplying labor. Labor may be either skilled or unskilled; more specifically, individuals may decide in the beginning of each period whether to be a skilled worker in the intermediary good industry, which entails a fixed cost of education, or an unskilled worker in the extractive industry, which entails no fixed cost.

At each point in time, one skilled worker is randomly selected to become a potential innovator in the intermediate industry, in which case she chooses her optimal innovation effort to maximize her 1 period monopoly rent. In case she successfully innovates, the productivity level of the economy permanently improves.

2.2 Environment

The consumption good industry There is a single representative firm producing consumption goods using both natural resources c and the intermediary good x. More specifically, At each period, there is a flow output of consumption goods indexed by y^c produced according to the same process:

$$Y^c = Ax^{\alpha}c^{1-\alpha} - px - c \tag{1}$$

Where A captures the productivity level of the economy depending on the research effort achieved by the research industry which may vary over time, x is the flow input of intermediate goods, c is the flow input of natural resources, and α is a scalar in the $(0, 1)$ interval. Also, p denotes the price of the intermediary good while the price of natural resources is normalized to 1.

The intermediary industry There is a single monopolist firm in the economy which produces a flow x of intermediate good and fixes its price p depending on the demand it expects from the consumption good industry. The intermediate good is produced using only skilled labor according to the linear technology:

$$x = L_S \tag{2}$$

Where L_S is the flow of skilled labor, which itself depends on the amount of laborers who decide to get educated. Given the monopoly structure of this industry, the monopolist sets its price to maximize expected profits depending on the demand it expects from the consumption good industry.

The extractive industry Unskilled workers may also join the extractive industry which exploits natural resources according to a linear technology such that the flow output of goods produced by the extractive industry writes:

$$Y^E = \theta L_U \tag{3}$$

Where Y^E is the flow output of goods produced by the extractive industry, L_U is the flow of unskilled labor and θ is a parameter indexing the availability of natural resources. Intuitively indeed, the larger the availability of natural resources, the more productive the extractive industry is.

The research industry Finally, there is a research industry where only skilled workers may operate. At each period, one skilled worker is randomly drawn and gets an idea; it is like saying that there is an exogenous probability of getting a patentable idea, which might be converted in actual innovation by incurring a research effort. More specifically, upon achieving an effort of $\frac{n^2}{2}$, the researcher successfully innovates and finds a new variety of the intermediary good with a probability γn, where γ indexes the ease with which the randomly selected researcher converts a patentable idea into an actual innovation. Said differently, γ is a measure of the human capital in the economy. Whenever a researcher successfully implements an innovation, she replaces the current monopolist in the intermediary industry and enjoys the monopoly rent for one period; yet, the productivity level in the economy permanently increases from A to A^+. This is like saying that inventors are myopic, in the sense that they only consider their private marginal benefit in the innovation decision. Hence, profit in the research industry writes:

$$\Pi^R = n\gamma\Pi^I(A^+) - \frac{n^2}{2} \tag{4}$$

Where n indexes the research effort achieved in the research industry, γ captures the human capital parameter, A^+ is the updated productivity level enabled by the quality-improving innovation, $\Pi^I(A^+)$ captures the post-innovation profit level of the intermediary good industry conditional on a successful innovation which permanently increases the productivity level from A to A^+. Without loss of generality, I posit that:

$$A^+ = A + \lambda \tag{5}$$

Meaning that the improvement in the overall productivity enters follows an additive process depending on a scale parameter λ which captures the radicality of inventions, namely the ability of inventions to transform and improve the production function. I also assume that the human capital level γ is an increasing and concave function of the number of skilled workers; said differently, human capital increases less than 1 for 1 as the number of workers who get educated rises. This is like assuming that there is a *learning by doing* process, whereby the randomly selected inventor is all the more likely to achieve successful innovation as skills are important in the economy. Typically, I posit that:

$$\gamma = L_s^\beta \tag{6}$$

Where L_s captures the mass of skilled labor in the economy and β is a scalar in the $(0, 1)$ interval which captures the efficiency with which human capital is used in the production process of inventions.

2.3 Solving individual problems

The consumption industry's decision problem Consumption good industry must choose the optimal amount of intermediate good x^* and of natural resources c^* in order to maximize profit. These levels write:

$$\begin{cases} \alpha A x^{\alpha-1} c^{1-\alpha} = p \\ (1-\alpha) A x^{\alpha} c^{-\alpha} = 1 \end{cases} \tag{7}$$

Combining both equations trivially yields the usual results from Cobb-Douglass specifications, namely:

$$\frac{x}{c} = \frac{\alpha}{p(1-\alpha)} \tag{8}$$

Typically, the demand for the intermediary good is increasing in its importance α in the production function and decreasing in its price p.

The monopolist's decision problem The monopolist maximizes her profit with respect to x, the quantity of intermediary good she may sold. Trivially, the profit of the intermediary industry writes:

$$\Pi^I = (p - w_s)x \tag{9}$$

Where p indexes the price of the intermediary good, w_s is the wage rate for skilled workers, and x is the amount of intermediary good. Plugging the value of p obtained in Equation 7 implies that the monopolist's profit writes:

$$\Pi^I = \alpha A x^{\alpha} c^{1-\alpha} - w_s x \tag{10}$$

It then follows that x^* must solve:

$$\alpha^2 A x^{\alpha-1} c^{1-\alpha} = w_s \tag{11}$$

The extractive industry's decision problem The extractive industry produces natural resources competitively using a constant returns to scale function. More specifically, it produces natural resources using unskilled labor as the sole input 1 for θ, meaning that each unskilled worker generates an output of θ. Hence, the extractive industry necessarily pays workers their marginal productivity, namely θ, and produces exactly the amount c required by the consumption good industry.

Workers' decision problem The mass L of workers must choose between acquiring education and working as skilled workers in the intermediary good industry, or working as unskilled workers in the extractive industry. For the sake of tractability, let us assume that the fixed cost of education $\bar{e}$ is normalized to 0. It then follows that workers join the intermediate good industry as long as it is offering a wage rate greater or equal to that offered in the extractive industry. Marginal productivity is constant and equal to θ in the extractive industry; in the meantime, marginal productivity in the intermediate industry is given by Equation A7. In order for workers to be

indifferent between acquiring education and becoming skilled or staying unskilled, the wage rate for skilled workers must be such that:

$$w_s = \theta \tag{12}$$

The researcher's decision problem At each period, one skilled worker is randomly drawn to have a patentable idea that might be turned in actual innovation provided that she achieves innovation efforts. Her optimal effort must thus maximize expected monopoly rents obtained upon innovating; hence, using Equation A4, the optimal innovation effort n^* writes:

$$n^* = \gamma \Pi^I \tag{13}$$

2.4 Equilibrium decisions

Production of intermediary good Equation 12 states that the wage of skilled workers must be set in equilibrium to θ. Then, x^* must verify:

$$x^* = c \left(\frac{\alpha^2 A}{\theta} \right)^{\frac{1}{1-\alpha}} \tag{14}$$

Also, labor must be allocated either to skilled or to unskilled employment; since each unskilled worker produces θ units of output, it follows that:

$$L = x + \frac{c}{\theta} \tag{15}$$

Combining Equation 15 with Equation 14 implies that:

$$x^* = \frac{L \left(\alpha^2 A\right)^{\frac{1}{1-\alpha}} \theta^{\frac{-\alpha}{1-\alpha}}}{1 + \left(\alpha^2 A\right)^{\frac{1}{1-\alpha}} \theta^{\frac{-\alpha}{1-\alpha}}} \tag{16}$$

It is then straightforward to show Proposition 1:

Proposition 1. *The equilibrium output of intermediate good x is decreasing in the availability of natural resources θ*

The intuition for this proposition is that natural resources drive the wage rate of unskilled workers up; by an arbitrage condition, the wage rate of skilled workers

must raise as well; yet, marginal productivity of labor is decreasing in the intermediate good industry, so that it cannot fully offset the increase in the unskilled wage rate. Hence, larger natural resources reduce the employment in the intermediary good industry.

Profits in the intermediary good industry Profits are simply equal to sales times profit per unit sold. Unit profit can easily be recovered by combining Equations 7 and 14, so that profit in the intermediary industry write:

$$\Pi^I = \left(\frac{1-\alpha}{\alpha}\right) L \frac{\left(\alpha^2 A\right)^{\frac{1}{1-\alpha}} \theta^{\frac{1-2\alpha}{1-\alpha}}}{1 + \left(\alpha^2 A\right)^{\frac{1}{1-\alpha}} \theta^{\frac{-\alpha}{1-\alpha}}} \tag{17}$$

It is then straightforward to show that natural resources exert an ambiguous effect on profits in the intermediate good industry. As long as α is low enough, namely as long as α is inferior to $1/2$, the effect of natural resources on profits in the intermediary good industry is unambiguously positive. Yet, whenever α is greater than $1/2$, there exists a cutoff $\hat{\theta}$ such that any value of θ in excess of this cutoff reduces the profit level of the intermediary good industry. The value of this cutoff is given by Proposition 2:

Proposition 2. *There exists a cutoff value $\hat{\theta}$ such that any value of θ in excess of this cutoff reduces the profit rate of the intermediary good industry. Formally, $\hat{\theta}$ writes:*

$$\hat{\theta} = \left(\alpha^2 A\right)^{\frac{1}{\alpha}} \left(\frac{1-\alpha}{2\alpha - 1}\right)^{\frac{1-\alpha}{\alpha}} \tag{18}$$

In words, Proposition 2 stems from the fact that natural resources have a twofold effect on profits in the intermediary industry. On the one hand, natural resources generate a *market size* effect, in the sense that they increase the output of the consumption industry which, given that natural resources are complements to the intermediary good in the production function of the consumption good, raise the demand for the intermediate good and, hence, the profitability of the intermediate industry. In the meantime, natural resources entail a *substitution effect* in the sense that they increase the wage rate of both skilled and unskilled workers, which lower the profitability of the intermediate industry. As natural resources get excessively large, the *substitution* effect tends to dominate over the *market size* effect, which reduces the profitability of the intermediate industry.

Innovation effort in the research industry In equilibrium, the randomly selected researcher sets n^* according to Equation 13 in order to maximize her profit. It then

follows that the optimal research effort writes:

$$n^* = \left(\frac{1-\alpha}{\alpha}\right)\left(L(\alpha^2 A(1+\lambda))^{\frac{1}{1-\alpha}}\right)^{\beta+1}\left(\frac{\theta^{\frac{1-\alpha\beta-2\alpha}{(\beta+1)(1-\alpha)}}}{1+(\alpha^2(A+\lambda))^{\frac{1}{1-\alpha}}\theta^{\frac{-\alpha}{1-\alpha}}}\right)^{\beta+1} \tag{19}$$

It may then be shown that θ, again, generates an ambiguous effect on n^* for three forces are now playing, two of which were already mentioned for the profit level in the intermediary industry, and one of which only pops up in the research industry. Indeed, the *market size* effect positively affects the research effort through its impact on profits in the intermediate industry; similarly, the *substitution* effect negatively affects the research effect through its impact on profits in the intermediate industry. The new effect is the *human capital effect* whereby larger natural resources unambiguously reduce the share of workers for which it pays to acquire education and join the intermediary industry; in turn, larger natural resources unambiguously reduce the pool of potential inventors which, in turn, reduces the efficiency with which potential inventors may translate patentable ideas into actual inventions.

In the end, natural resources negatively impact innovation effort whenever natural resources are larger than a cutoff value given by Proposition 3:

Proposition 3. *Natural resources negatively affect innovation efforts whenever they are larger than the cutoff value $\tilde{\theta}$ which writes:*

$$\tilde{\theta} = \left(\alpha^2(A+\lambda)\right)^{\frac{1}{\alpha}}\left(\frac{1-\alpha}{2\alpha+\alpha\beta-1}\right)^{\frac{1-\alpha}{\alpha}} \tag{20}$$

In words, Proposition 3 suggests that the innovation effort is lowered whenever natural resources are in excess of a certain threshold which positively correlates the radicality of innovation λ as well as the initial productivity level A, but negatively correlates the efficiency β with which education may be used in delivering new inventions. Interestingly enough, it is straightforward to show that this cutoff is lower than that for profits in the intermediary industry, which implies that the effect of natural resources on innovation is all the more negative as they implicitly determine the number of skilled workers which, in turn, positively influences the effort rate. Finally, and as a last remark, it may be noticed that $\tilde{\theta}$ takes the value 0 when α is set to 1: in that case indeed, there is no market size effect for natural resources do not contribute to the production function of the consumption good industry: hence, they only exert a negative impact through both the *human capital* effect and the *substitution* effect.

2.5 Growth dynamics

In the end, it is easy to compute growth dynamics in this economy. As the number of workers and the production processes are stationary, growth only comes from improved productivity. As productivity improvements depend on the decisions made by randomly selected researchers, it follows that:

$$\forall t, A_{t+1} = A_t + n^* \lambda \tag{21}$$

Where n^* captures the optimal innovation effort incurred by the randomly selected researcher. Given that n^* is negatively affected by natural resources whenever the former are larger than the cutoff value $\tilde{\theta}$ it follows that growth, through its productivity component, is negatively affected by natural resources whenever the latter are excessively large. Additionally, it is noteworthy that λ also affects the pace of innovation in a twofold way; the larger λ, the larger the cutoff value $\tilde{\theta}$ above which natural resources start to exert a negative impact on innovation; also, the larger λ, the larger the step-size of innovation and, hence, the higher productivity growth is.

2.6 Extension

A natural extension that emerges in such a framework consists in endogenizing the value of λ, namely the radicality parameter. Indeed, productivity growth in this model depends on both the innovation effort and the radicality parameter. Yet, it could be that exposure to natural resources not only alters the extent of the innovation effort, but also its quality, in the sense that it could have an impact over the radicality of innovation as captured by the λ parameter. Formally, I encompass this assumption by assuming that each randomly drawn inventor may achieve a λ^+ innovation upon paying a fixed cost F. Thus, an inventor who achieved successful innovation efforts and paid the fixed cost F permanently moves the productivity level in the economy from A to $A + \lambda^+$ such that $\lambda^+ > \lambda$. This is like saying that, upon achieving some fundamental research, the skilled worker might open up the way for a new field or a new industry which proves drastically more efficient than previous ones.

Then, a skilled worker randomly chosen to achieve innovation effort at time t chooses to achieve a plain innovation rather than a radical innovation if and only if her profit rate in the former case is greater than in the latter; this writes formally:

$$\Pi^R(\lambda) > \Pi(\lambda^+) - F \tag{22}$$

It is then straightforward to prove, as will be done in Appendix A.9, that there exists a cutoff value $\tilde{F}$ such that achieving radical innovation effort never pays when the fixed cost to fundamental research is greater or equal to this threshold.

Proposition 4. *There exists a cutoff value $\tilde{F}$ such that it never pays to achieve radical innovation whenever the fixed cost to radical innovation is in excess of that threshold. More specifically, $\tilde{F}$ writes:*

$$\tilde{F} = \left(\frac{1-\alpha}{\sqrt{2}\alpha}\right)^2 \theta^{\frac{2-2\alpha\beta-4\alpha}{1-\alpha}} \left(\left(\frac{\left(\alpha^2\left(A+\lambda^+\right)\right)^{\frac{1}{1-\alpha}}}{1+\left(\alpha^2(A+\lambda^+)\right)^{\frac{1}{1-\alpha}}} \right)^{2\beta+2} - \left(\frac{\left(\alpha^2\left(A+\lambda\right)\right)^{\frac{1}{1-\alpha}}}{1+\left(\alpha^2(A+\lambda)\right)^{\frac{1}{1-\alpha}}} \right)^{2\beta+2} \right) \tag{23}$$

Interestingly enough, $\tilde{F}$ decreases with θ as long as θ is greater or equal to a certain threshold; this is what Corollary 1 states:

Corollary 1. *The incentive to achieve radical innovation, meaning the incentive to achieve innovation in a new industrial field, is decreasing in θ as long as θ is larger than $\breve{\theta}$, such that $\breve{\theta}$ writes:*

$$\breve{\theta} = \left(\frac{(1-\alpha)\left(\alpha^2(A+\lambda^+)\right)^{\frac{1}{1-\alpha}}}{\alpha\beta+2\alpha-1} \right)^{\frac{1-\alpha}{\alpha}} \tag{24}$$

Importantly enough, this cutoff value is increasing in λ^+: the larger the increase in permanent productivity induced by the radical innovation, the less likely it is that natural resources act as a deterrent to radical innovation. By contrast, $\breve{\theta}$ is decreasing in β, suggesting that the larger the return to education, the lower the value above which natural resources exert a negative role on radical innovation. Finally, the sign of α is rather ambiguous, though it proves interesting to consider the extreme case when α tends to 1: then, natural resources unambiguously generate a negative effect on radical innovation, in the sense that $\breve{\theta}$ then takes the value 0.

In words, what Corollary 1 and Proposition 4 imply is that natural resources may prevent local economies from adapting to new industries; indeed, investing in a new industry implies a fixed cost which may be large enough to offset the profit increase associated with a higher productivity level. Given that excessive natural resources reduce the profit rate in the intermediary industry, enabling modern industries to develop might be more difficult in industries abundantly endowed with natural resources for their wage rate is excessively large to allow for a substantial profit margin in case of adaptation to a new industry.

2.7 Summary of predictions

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

1. *Excessively abundant natural resources directly lower the pace of innovation*: excessively large natural resources drive the wage of skilled workers up which lowers the profitability of the intermediary industry and, hence, of inventions.
2. *The effect is all the larger as natural resources are of poor significance for the production of consumption goods*: whenever natural resources are barely useful in the production process of consumption goods, their negative effect on innovation tends to dominante.
3. *This negative impact may be amplified if natural resources affect education*: if natural resources affect the upper-tail education as measured by γ, their effect on productivity might be all the more detrimental.
4. *This negative impact may also be amplified if natural resources affect the step-size of innovation*: if natural resources induce a specialization such that this specialization implies lower values of λ, then the impact of natural resources on innovation is all the more negative.

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 elaborates 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

Coal mining 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 a wide range of detailed sources. I notably leverage 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 exposure, 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 lit-

erature (see, for instance, Esposito and Abramson (2021)). Finally, and following the existing literature (see, among others, Esposito and Abramson (2021) and Fernihough and O'Rourke (2014)), I gather information on the location of the most important European coalfields as surveyed by Châtel (1931) which proves useful in two respects: first, this source was already used in previous papers; second, it allows me to have a broader view of coal extraction at the European level, enabling robustness checks as well as testing alternative mechanisms. 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.

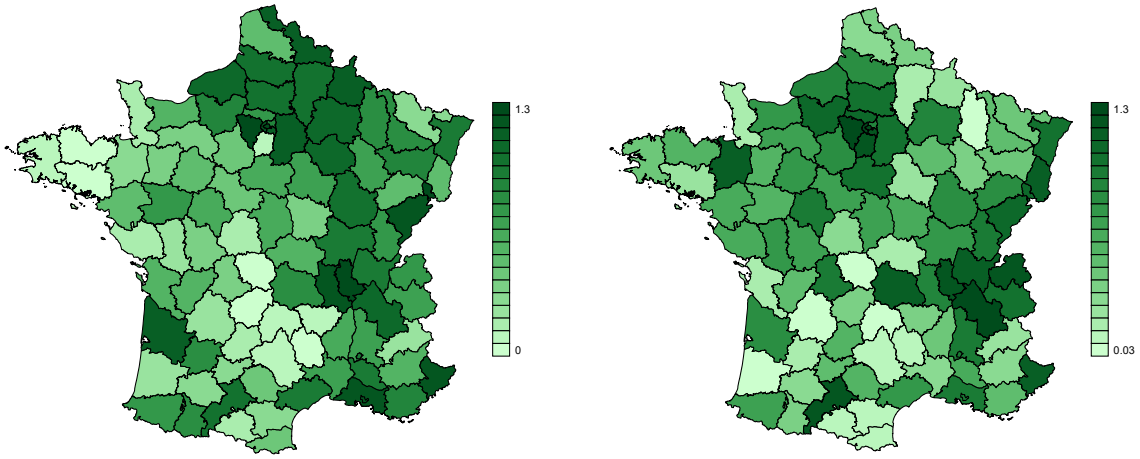
Geographical data As will be explained later, coalfields location might not be totally random, in particular because coal extraction requires both a large labor force and specific technical skills. To mitigate the risk that local geographical characteristics are driving the estimate, as well as to implement 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; in particular, it shows sizable variation in terms of its geographical dispersion, as suggested by Figure 2.Typically, the North-eastern part of France, which used to rank among the most productive departments in terms of innovation per capita, experienced a substantial decline relative to Western departments; to take but one compelling example, patenting activity per capita was 4 times larger in Nord compared to Ille et Vilaine as of 1914; a century later, Ille et Vilaine showed a 3.5 times larger patenting activity per capita compared to Nord, implying that variations in terms of prosperity and innovation abilities were not negligible in the course of the XXth century.

FIGURE 2: Evolution of patenting activity per capita in France: 2014 vs. 1914

(a) Patenting activity per capita in France, 1914 **(b)** Patenting activiy per capita in France, 2014



Notes: Figure 2a shows the ventilation of patenting activity per capita at the department level in France in 1914; darker shades of green indicate higher values of patenting activity per capita; meanwhile, Figure 2b shows the dispersion of patenting activity per capita at the department level in France in 2014. Source: (Bergeaud and Verluise, 2024)

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 holding 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. Finally, I add data from Ruggles et al. (2024) in order to investigate the impact of coal booms on education in the long run; indeed, the granularity of American censuses throughout the XIXth and the XXth century is peculiarly tailored to study dynamic effects of an exogenous increase in the size of extractive industries.

Employment data Industry specialization is also one of the key channels through which coal may have resulted in a negative impact on patenting activity in the long run. To evaluate the strength of this channel, I combine various sources including the *Enquêtes Emploi en Continu* which cover more than 40 years starting 1980. The breakdown of occupation across sectors is fine-grained as it is made at the 4-digit NAF2 level, allowing for an in-depth analysis. I also leverage the 1975 survey *Enquête Formation et Qualification Professionnelle* to investigate the transmission of job occupations within families, a key channel of path dependence in coal regions. Finally, I use granular data at the county level in the United States to unveil some of the mechanisms identified in this paper; more specifically, I leverage employment and industry data from Ruggles et al. (2024) in order to investigate the impact of booms in the extractive industry on employment concentration. The reason why I exploit American data at times in the paper is due to the fact that similar sources are either not always available in France, or not consistently reported throughout censuses, which prevents me from achieving a careful study.

3.3 Construction of the coal exposure measure

In order to quantify the effect of exposure to natural resources on productivity and innovation in the long run, I first need to assess the heterogeneity in the exposure of French departments to coal mining. As highlighted, the main measures of

exposure to coal industry are built using three different data sources, respectively the *Cadastre minier* released by the French Ministry for Environment, the *European Route of Industrial Heritage* (ERIH) project and the *Atlas des houillières d'Europe* by Châtel (1931). 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; finally, the third dataset summarizes the most important coalfields at the European level in the beginning of the XXth century. For the sake of concision, I only report here the construction of the main coal exposure measures, namely those which are built using the proximity indicator towards the closest coalfield. The discussion of alternative coal exposure measures is postponed to the Appendices. Importantly enough, most of the estimates derived in the paper will be based on two main measures of coal exposure detailed below; results obtained upon using the alternative measures will be postponed to the Appendices.

In short, coal exposure of department d is computed as the average proximity of each French city i located in department d with respect to the closest coalfield. Formally, let $\tilde{d}^d$ denote the average distance of any city located within department d with respect to the closest coalfield. By construction, $\tilde{d}^d$ writes:

$$\tilde{d}^d = \frac{1}{n_d} \sum_{j \in \Gamma_d} \min_{c \in \mathcal{C}} d_{j,c} \tag{1}$$

Where n_d captures the number of cities in department d, Γ_d is the set of cities in department d while $\mathcal{C}$ encompasses the set of French coalfields. It is then straightforward to compute the average proximity $\tilde{p}^d$ at the department level by applying the following log transformation to $\tilde{d}^d$:

$$\tilde{p}^d = \ln\left(\frac{1}{1 + \tilde{d}^d}\right) \tag{2}$$

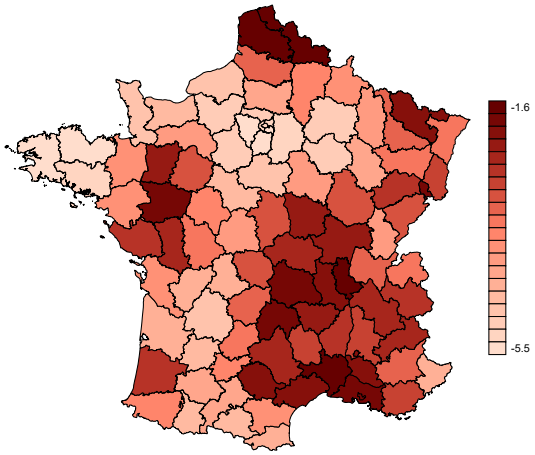
This log-transform is inspired from previous work including that by Esposito and Abramson (2021). It indeed features different advantages related to the log structure of the measure: because of the concavity of the log function, it ensures fast decreasing returns to scale in the exposure to coalfields, 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: 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 dis-

tance. I then compute the value of this metric considering different space of coalfields, namely, considering alternative sets for $\mathcal{C}$. I start by computing $\tilde{p}^d$ for the the set of the largest coalfields as reported in the aforementioned ERIH database, which allows me to capture not only the number of coalfields, but also their size. I then let $\mathcal{C}$ be the set of coalfields reported in the Châtel (1931) Atlas, a source which is frequently used in that literature (see, among others, Esposito and Abramson (2021)). I finally leverage the whole set of French coalfields that were open in the beginning of the XXth century according to the *cadastre minier*; this choice stems from the fact that the number of coalfields reached its peak in France in the early XXth century, implying that exposure to coal was at its highest value in the beginning of the XXth century.

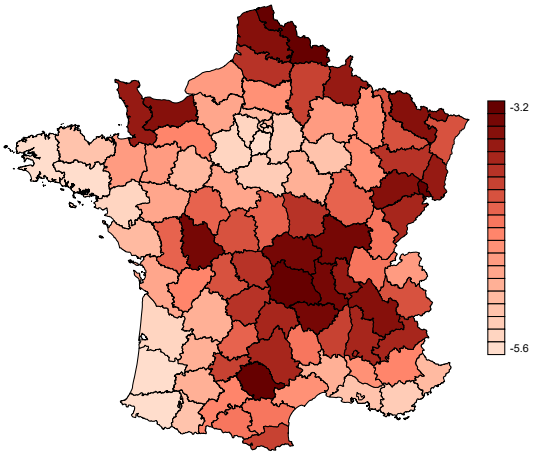
Out of concision, I only report the department level exposure to coal industry using the two first proximity measures in Figure 3 and postpone the discussion of the alternative measures to Appendix B; more specifically, Figure 3a shows the exposure to coal mining when the set $\mathcal{C}$ of coalfields is taken according to the Châtel (1931) map while Figure 3b shows the exposure to coal mining when the set $\mathcal{C}$ incorporates the largest coalfields as reported in the ERIH project. Two features deserve attention: first, it is noteworthy that both measures are broadly consistent, suggesting that exposure to mining activities is estimated precisely at the department level; second, three important coal basins emerge at the national level: the Northern coal basin, encompassing the aforementioned department of Nord; second, the eastern coal basin around the departments of Moselle and Bas-Rhin; third and finally, the large coal basin in the mid-south eastern part of France, including, among others, the departments of Saône et Loire, Puy de Dôme and Allier.

FIGURE 3: Two main measures of coal exposure

(a) Using the Châtel (1931) map



(b) Using the ERIH project



Notes: Figure (a) plots the coal proximity measure when taking the Châtel (1931) map as the basis to elaborate the set $\mathcal{C}$ of French coalfields while Figure (b) pictures the coal proximity measure when restricting the set of coalfields to historical major coalfields documented in the ERIH project.

4 Empirical results

4.1 Empirical specification

To measure the impact of exposure to coal industry on patenting activity, , I estimate repeated cross-section equations, using either Pseudo-Poisson or standard OLS regression frameworks. Technically, I am estimating the following equation for each available year in my sample:

$$Y_{d,t} = \alpha + \lambda e_d + \chi \Gamma_d + \varepsilon_{d,t} \tag{3}$$

Where $Y_{d,t}$ is the variable of interest, namely patenting activity in department d at time t, 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 various crops, 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. Such controls are especially important given that the time invariant structure of my main explanatory variable imposes the repeated cross-section functional form which, itself, prevents me from allowing department fixed effects.

The coefficient of interest, namely λ, can be causally interpreted under the identifying assumption that the exposure to coal is orthogonal to the error term in equation 3 conditional on geographical and geological characteristics of each department and year fixed effects. Formally, this identifying assumption writes as:

$$\forall(d,t), \mathbb{E}[\varepsilon_{d,t} \times e_d | \alpha_t, \Gamma_d] = 0 \tag{4}$$

This identifying assumption states that, in the absence of coal extraction, departments would have followed similar innovation trends in the long run conditional on all the observables included as controls in Equation 3. 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 respect 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. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

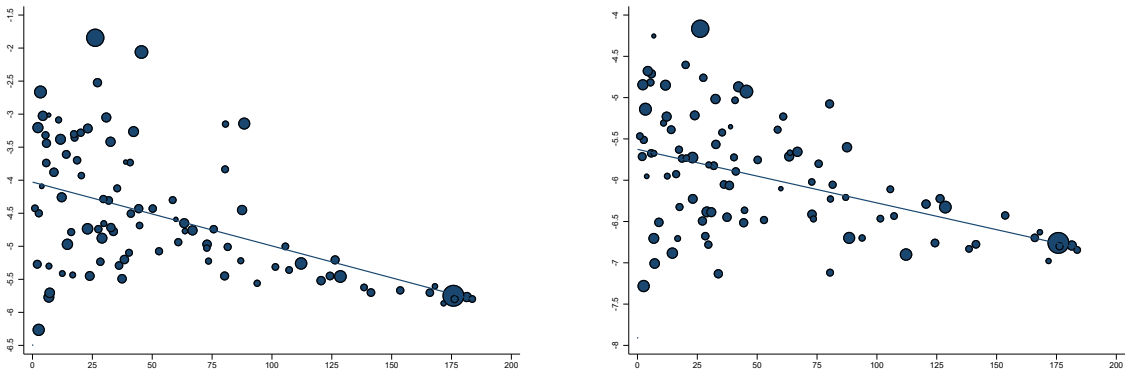
In concrete terms, I compute for each department the average distance of each city to the closest carbon sedimentary rock strata. The validity and strength of this instrument is clearly supported by Table 1 which summarizes the estimates of the coefficients attached to each regressor included in the first stage regression. Strikingly,

the distance to carboniferous rock strata exerts a strong and negative impact over the proximity to large 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 4: Proximity to coal and distance to carboniferous rock strata

(a) Coalfields from the Châtel (1931) map

(b) Coalfields from the ERIH project



Notes: Figure (a) plots the coal proximity measure when taking the Châtel (1931) database 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. Dot size is weighted according to the initial population prevailing in each department

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 4 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 Main results

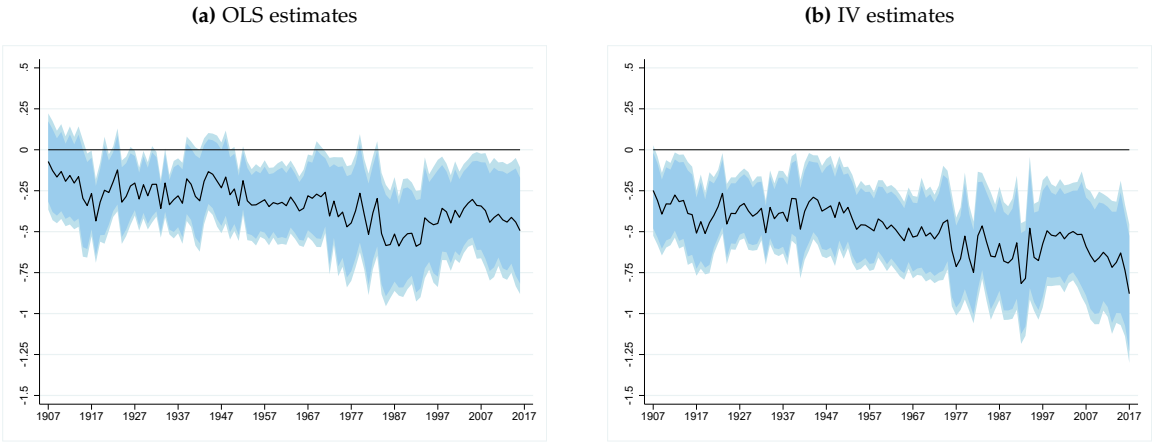
In this subsection, I introduce the main results regarding the impact of coal exposure on patenting activity. As a reminder, Section 2 suggests that natural resources might exert a detrimental impact on productivity growth when exceeding a specified threshold; more precisely, the model implies that the effect is all the more likely to be detrimental, as natural resources are of lower significance in the production function of consumption goods; given that coal was increasingly replaced by alternative energy sources in the course of the twentieth century, the effect might have deteriorated over time. In order to support these predictions, I start by showing that coal exposure induced a negative and significant impact on innovation activity throughout the XXth century in France (Section 4.3.1); I then provide evidence that this effect also holds for breakthrough patenting activity (Section 4.3.2). Finally, I emphasize that the effect varies across technology classes (Section 4.3.3).

4.3.1 Coal exposure and patenting activity

I start by bringing evidence of a negative and significant impact of coal exposure on total patenting activity at the department level. Thus, Equation 3 is run for all years in my sample which span 1907 to 2016. Results are summarized in Figure 5 which clearly suggests a negative, significant and sustained impact of exposure to coalfields proxy by the exposure to largest coalfields on patenting activity. Results are shown both for the OLS specification (Figure 5a) and for the IV specification (Figure 5b). Results will be derived for alternative coal exposure measures in Appendix D. It is first noteworthy that the estimates imply a large effect of exposure to coalfields on innovation activity: estimates are indeed significant both statistically and economically; irrespectively of whether we consider the OLS or the IV specification, almost all coefficients are significant at the 5% and even at the 1% level throughout the XXth century. More importantly, results imply that a one standard deviation increase in exposure to coalfields reduced innovation activity by roughly 34 to 43% depending on the retained measure of coal exposure. Interestingly enough, IV estimates are slightly larger than OLS estimates, which is consistent with the intuition that coal extraction might have correlated with covariates which, themselves, positively affected economic

growth in the long run. Typically, it could be that coal extraction required some technical skills which implies that departments could mine coal only to the extent that, as industrialization started, they had a large enough basin of skilled workers; in that respect, coal extraction would have induced a reversal of fortune, an assumption which will be tested later on in Section 5.3. Finally, estimates are significant as soon as the beginning of the XXth century and seem to worsen over time: while this finding will be discussed more extensively later on, it may mirror the predictions of the theoretical argument discussed in Section 2, namely that the effect of coal on innovation is all the more negative as natural resources are of poor significance in the production process of consumption goods (meaning that α is large enough). It could also encompass path dependence effects which may convey through industry specialization or education attainment.

FIGURE 5: Dynamic effect of coal on patenting quantity



Notes: Figure (a) plots the dynamic effect of coal exposure on total patenting activity using the OLS estimates obtained when running Equation **??** while Figure (b) plots the same dynamic effect when instrumenting coal exposure through distance to the closest carboniferous rock strata.

4.3.2 Coal exposure and breakthrough patenting activity

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 achieve numerous inventions of poor significance while others restrict to game-changer inventions which bring a higher contribution to economic performance. In that case, the estimates could reflect a quantitative rather than a qualitative effect. To explore that assumption, I substitute breakthrough patenting activity for total patenting activity as the dependent variable. Breakthrough patenting activity is a measure which was introduced by Kelly et al. (2021) to identify inventions that are both novel and meaningful. A patent is said to be novel whenever it semantically differs from previous

patent, while it is said to be meaningful if it semantically resembles future patents.

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 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; for each year, I then define as a breakthrough technology the top ten 3-digit IPC classes in terms of the share of breakthrough patents. Finally, I aggregate the number of breakthrough inventions at the department $\times$ year level and run Equation 3 by substituting breakthrough inventions for standard patenting activity.

Results of this exercise are summarized in Figure 6 which pictures estimates both for OLS (Figure 6a) and for IV estimates (Figure 6b). Results shown in this subsection rely on the coal exposure measure computed in Section 4 which uses the ERIH project data; robustness checks for alternative coal exposure measures are provided in Appendix D. Given that breakthrough patenting activity is a subset of total patenting activity, estimates are slightly noisier than those derived in the general framework. Results are nevertheless broadly consistent with those derived in the previous subsection. It is first noteworthy that estimates are almost systematically negative and significant either at the 1 or at the 5% level both in the OLS and in the IV specifications. Then, estimates again suggest that a one standard deviation in exposure to coal reduced patenting activity by roughly 30 to 40% depending on the measures of coal exposure. Importantly enough, IV estimates are also larger than OLS estimates, which further support the intuition that coal activities may not have located randomly with respect to initial conditions. Finally, the effect of coal on radical innovation worsened over time, a finding which is consistent with lessons drawn from Figure 5.

4.3.3 Coal mining and patent specialization

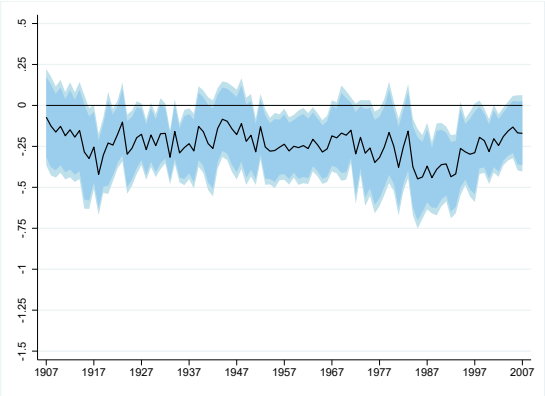
Beyond their impact on innovation quantity and quality, extractive industries may have influenced the orientation of innovative activities, which, in turn, may also influence both the quantity and the quality of innovation in the long-run. To investigate the potential heterogeneity of the effect of coal exposure on different technical categories, I consider the following model:

$$Y_{d,c,t} = \sum_{k=1}^{22} \mathbb{1}_c(k)\lambda_k e_d + \alpha_d + \gamma_c + \beta_t + \eta_{d,c,t} \tag{5}$$

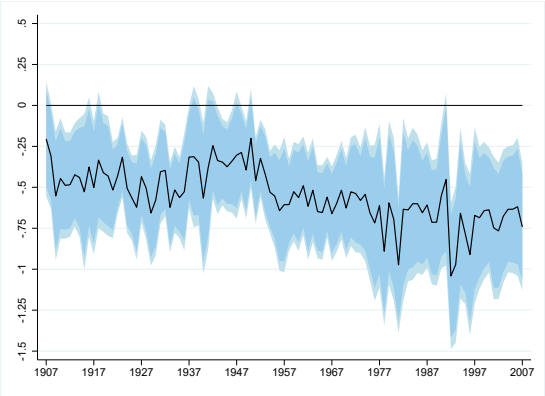
Where $Y_{d,c,t}$ captures an outcome variable of interest in department d, technology

FIGURE 6: Dynamic effect of coal on breakthrough patenting quantity

(a) OLS estimates



(b) IV estimates

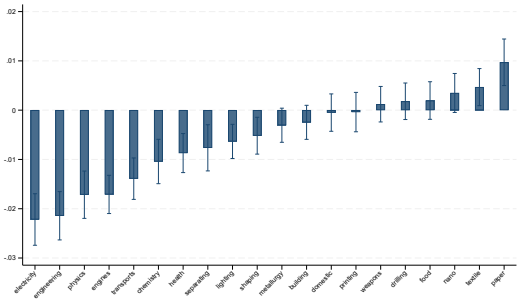


Notes: Figure (a) plots the dynamic effect of coal exposure on breakthrough patenting activity using the OLS estimates obtained when running Equation 3 while Figure (b) plots the same dynamic effect when instrumenting coal exposure through distance to the closest carboniferous rock strata.

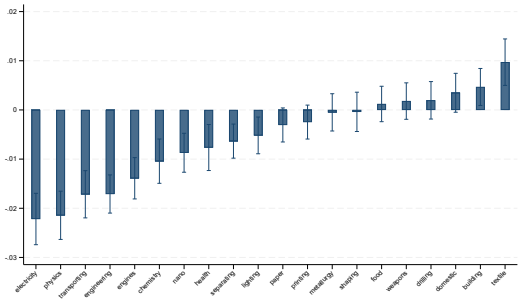
class c and time t, α_d captures department fixed effects, γ_c captures technology class fixed effects and β_t indexes year fixed effects. I consider two variants of this equation, namely one in which I take the (ln) number of patents as the dependent variable and one in which I take the share of patents that each technology class represents at the local level. I then plot the results for each of these specifications in Figure 7 which suggests that the impact of coal exposure on patenting activity was heterogeneous across technological classes. While the estimates of both specifications slightly vary, they nevertheless highlight some clear differences: typically, patents related to electricity, physics, transportation, engineering or chemistry were negatively and significantly impacted by coal exposure, while patents related to textile, building, food or drilling were rather positively affected though not always significantly.

FIGURE 7: Effect of coal on patent specialization

(a) Total patenting activity



(b) Share of patenting activity



Notes: Figure (a) plots the coefficients λ obtained for each technological category when running Equation 5 where the dependent variable is the (log) number of patents at the cluster x year level while Figure (b) plots the same coefficients when running Equation 5 where the dependent variable is the year on year share of each technology class in total patenting activity by any department.

This finding supports the idea that coal might not have been systematically detri-

mental to invention abilities; interestingly enough, technology classes associated with inventions from the past seem to have been either positively, if not always significantly, affected by coal extraction. Given that such industries tended to deliver fewer radical innovations in the long run, they might indicate that extractive departments were stuck at some point in low-value added industries which, in turn, may have reinforced the negative impact of coal extraction on patenting activity. Theoretically speaking, this finding may give support to the idea that exposure to natural resources, as proxy by θ in the model, may have correlated negatively with radicality of innovation λ which, in turn, might account for the negative and sustained impact of exposure to coalfields on innovation.

5 Mechanisms

The previous section documented a negative and significant impact of exposure to coalfields on innovation in the long run. Importantly enough, results imply that such an impact worsened over time, a statement that corroborates two predictions of the model, namely that the effect of extractive industries on innovation is all the more detrimental, as these industries negatively affect (i) human capital formation and (ii) industry specialization. I therefore bring evidence in this section showing the interplay between coal exposure and education (Section 5.1), the impact of coal exposure on industry specialization (Section 5.2) as well as complementary mechanisms that might have reinforced the results identified above (Section 5.3).

5.1 Human capital

The model elaborated in Section 2 lends credence to the idea that natural resources might generate a negative *human capital* effect, to the extent that abundant natural resources improve the wage rate of skilled workers which, in turn, lowers the profitability of the intermediary industry and, finally, the incentive of workers to acquire education. Given that innovation itself depends on the share of workers who decide to acquire education, it follows that innovation may be all the more negatively impacted by natural resources as the latter entail a decrease in the share of skilled workers. In order to identify this plausible channel, I start by offering city-level evidence that the impact of coal exposure on innovation was heterogeneous depending on the initial endowment of high human capital (Section 5.1.1); I then turn to city-level evidence of a negative and sustained impact of exposure to coal on higher education (Section

5.1.2) as well as to individual-level evidence (Section 5.1.3) before leveraging an event-study analysis to further assert the causal impact of extractive industries on long-run human capital formation (Section 5.1.4).

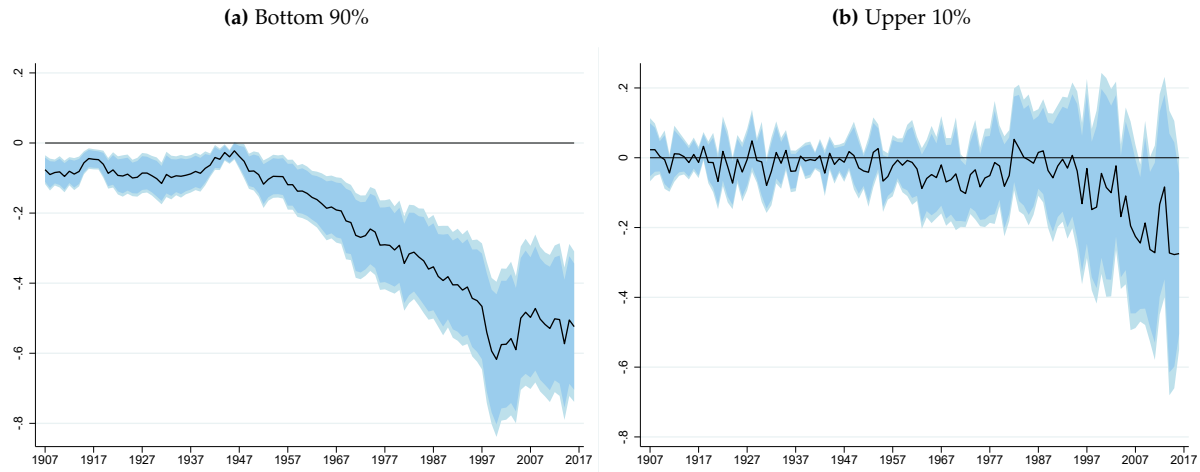
5.1.1 Heterogeneity of coal exposure effect depending on initial human capital

Before turning to causal evidence, it proves convenient to investigate the heterogeneity of the effect of coal exposure on innovation depending on the initial level of human capital. In order to identify this heterogeneity, I rely on the dataset assembled by Squicciarini and Voigtländer (2015) who argue that the number of subscribers to the French *Encyclopédie* is a good proxy of high-value knowledge in the end of the XVIIIth century. Key to their argument is the idea that one must distinguish between *average* human capital and *upper-tail* human capital. In particular, Mokyr (2005) stressed the importance of the density in the upper-tail education to foster industrialization and growth in the long-run. In turn, cities which exhibited higher initial level of upper-tail education *before* the start of industrialization and coal extraction might have been protected from detrimental impacts of such extractive activities in the long run. Technically speaking, a higher initial value of upper-tail human capital may have translated through a larger initial productivity endowment, which, in turn, increased the cutoff value $\tilde{\theta}$ above which natural resources start to exert a detrimental impact on productivity. In order to test this assumption, I leverage the upper-tail education measures discussed by Squicciarini and Voigtländer (2015): more specifically, I recover the number of subscribers to the French *Encyclopédie* in the more than 300 cities considered in their paper. The issue with their paper is that they restrict their focus on roughly 300 cities which might not be enough in order to study long-run heterogeneity in innovation trends. Hence, I build a measure of upper-tail education at the local level as follows:

$$h_i = \sum_{c \in \mathcal{C}} e^{-d_{i,c}} h_c \tag{6}$$

Where h_i is the measure of upper-tail human capital in city i, $\mathcal{C}$ is the set of cities incorporated in the dataset by Squicciarini and Voigtländer (2015) and $d_{i,c}$ is the distance between city i and city c. In words, h_i is a discounted measure of the exposure of city i to human capital of cities registered in the aforementioned database. Intuitively, a city which is located close to important subscription places of the French *Encyclopédie* might have enjoyed a higher level of human capital before the Industrial Revolution through local spillovers. I then simply run Equation 3 at the city level by contrating cities belonging to the upper 10% of the sample in terms of the value of h_i with cities

FIGURE 8: Heterogeneity of coal exposure effect on innovation depending on initial upper-tail human capital



Notes: These graphs show heterogeneity of the effect of coal exposure on innovation at the city level depending on whether a city belongs to the bottom 90% of upper tail human capital before industrial revolution (Figure 8a) or the the top 10% of upper tail human capital before industrial revolution (Figure 8b). Darker blue shaded areas indicate 90% confidence intervals while lighter blue shaded areas denote 95% confidence intervals.

belonging to the bottom 90% of the sample in terms of the value of h_i. Results of this heterogeneity analysis are displayed in Figure 8 which clearly supports the idea that the effect of coal exposure on innovation varied substantially with the intial endowment in terms of upper-tail education. As the scale is normalized on both sides, coefficients can be directly interpreted and suggest that the effect of coal exposure on innovation at the city level was already negative as soon as the beginning of the XXth century and worsened over time, with a rapid deterioration in the beginning of the 1950s, for cities poorly endowed in terms of initial upper-tail human capital; by contrast, coal exposure barely affected innovation in cities endowed with a strong initial endowment of upper-tail human capital, a finding which is consistent with the role played by education in shaping the impact of coal exposure on innovation activity.

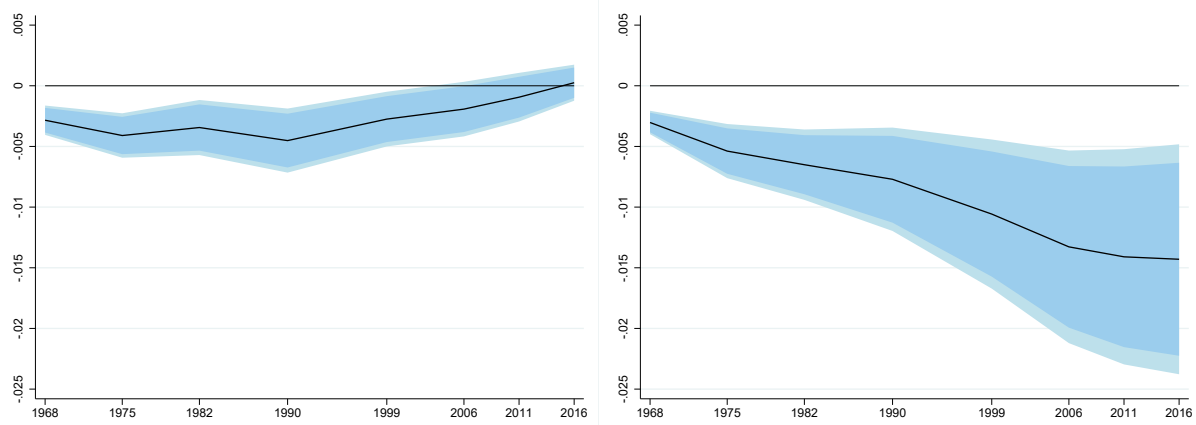
5.1.2 Upper-tail education at the city-level

The previous subsection provided evidence that education played a pivotal role in shaping the effect of coal exposure on innovation activity. Yet, it remains to investigate whether education itself was affected by exposure to coal, which, in turn, might have reinforced the negative impact which coal induces over innovation beyond the threshold specified in Section 2. To estimate this channel, I examine the relation between proximity to coalfields and the share of the population which achieved higher education either in the form of obtaining the *baccalauréat* or in the form of achieving college education. In order to carry out this analysis, I leverage census data from INSEE which details city-level information on the education level of inhabitants for

eight points after World War II, namely 1968, 1975, 1982, 1990, 1999, 2006, 2011 and 2016. I then simply run cross-sectional regressions for each of these 8 censuses where I take the dependent variable to be a schooling indicator, and I use the usual geological, geographical and demographic controls on top of my coal exposure measure. As was done for the main section (see Section 4), I report graphically the results of these regressions in a twofold purpose and postpone to the Appendices the summary of the main results obtained when running a pooled-OLS regression.

FIGURE 9: Coal exposure and long-run education: evidence from the post-war French city-level

(a) Share of population with *baccalauréat* **(b)** Share of population with college education



Notes: These graphs show the effect of coal exposure on city-level education after World War II; more specifically, these graphs plot the coefficients obtained when running an instrumental variable regression where the dependent variable is respectively taken to be the share of the population with *baccalauréat* as the highest degree (Panel a) and the share of the population with college education (Panel b). Results are plotted dynamically for the eight post war censuses, namely 1968n 1975, 1982, 1990, 1999, 2006, 2011 and 2016. Standard errors are clustered at the department level. Darker blue shaded areas indicate 90% confidence intervals while lighter blue shaded areas denote 95% confidence intervals.

Three features stand out from this exercise. First, exposure to coal industry induced a negative and significant impact on upper-tail education, even in the long run. Typically, a one standard deviation increase in exposure to coal industry entails an average -0.1 pp decrease in the share of upper-tail education. That this effect is lasting and significant suggests that the negative impact of extractive industries on education might account for a significant share of the overall negative impact of extractive industries on French local productivity. This is consistent with the theoretical argument made in Section 2, which suggests that abundant natural resources reduce the incentive for workers to acquire education and, hence, to improve the overall productivity of the economy. Second, the impact is substantially larger for college education than for *baccalauréat*; indeed, Figures 9a and 9b are plotted using the same scale to ensure comparability. Clearly, the estimates are significantly larger for college education compared to baccalauréat; since the former indicates a higher school attainment, results imply that exposure to extractive industries was peculiarly detrimental to the extreme upper tail of the education distribution. More importantly, given that innovation is

known to be a highly-skewed process (see, among others, Akcigit et al. (2016)), the findings derived in this subsection may suggest that the negative impact of extractive industries on patenting activity might have funneled through its effect on upper tail human capital.

5.1.3 Individual-level evidence

In order to further elaborate on the impact of extractive industries on human capital formation, I complement the city-level analysis with a careful study at the individual level. I therefore exploit American censuses across time to document a negative and significant correlation between extractive occupation and low level of literacy. More specifically, I use four censuses released by Ruggles et al. (2024) which comprise fine-grained information on more than 3 million American individuals in the XIXth century. Unfortunately indeed, I don't have access to such a granular dataset at the French level, and hence make the assumption that the impact of extractive industries on human capital formation at the individual level is roughly homogeneous acros countries. In terms of empirical strategy, I estimate the following model:

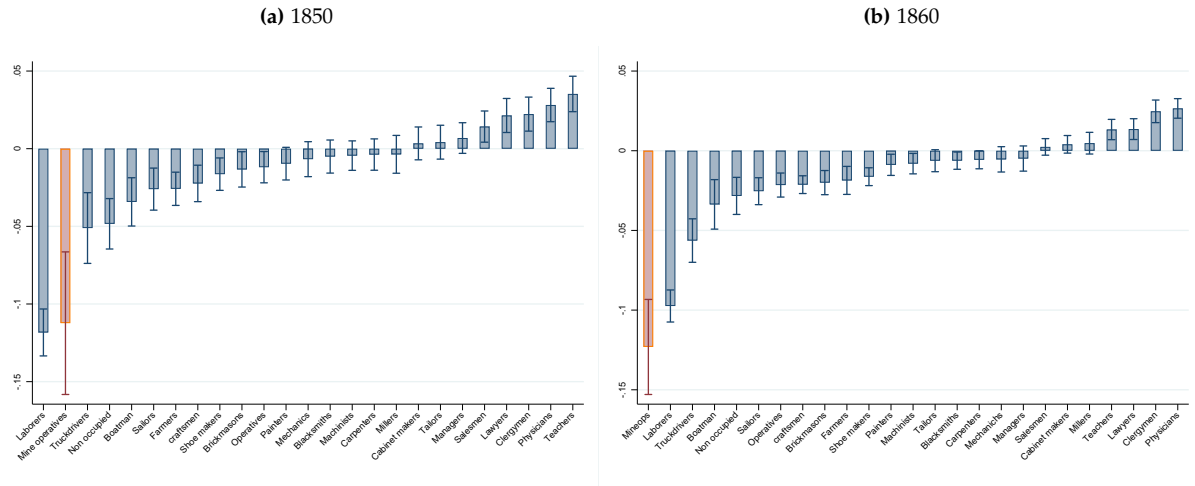
$$Y_i = \sum_{k \in \mathcal{I}} \mathbb{1}(k) + \sum_{j \in \mathcal{T}} \mathbb{1}(j) + \sum_{l \in \mathcal{C}} \mathbb{1}(l) + \varepsilon_i \tag{7}$$

Where $\mathbb{1}(k)$ is a binary variable taking the value 1 if individual i is working in occupation k, $\mathbb{1}(j)$ is a binary variable taking the value 1 if individual i was registered at time j and $\mathbb{1}(l)$ is a binary variable taking the value 1 if individual i is living in county l. In words, this model aims at estimating the coefficients attached to each categorical variable respectively capturing (i) the job occupied by individual i, the year when individual i was surveyed and (iii) the geographical origin of individual i. I am mostly interested in the coefficients attached to categorical variables related to occupations. Indeed, I want to estimate whether, all else equal, extractive jobs are (i) more likely to generate a negative impact over educational outcomes and (ii) how these effects compare with other similar occupations.

The interest of this new specification is twofold: first, the richness of the Ruggles et al. (2024) dataset allows me to track the effect of extractive industry as early as the second half of the XIXth century, a key period in the economic development of Western countries; second, it enables me to shed a more specific light on the impact of extractive industries by not only looking at a geographical approach, but rather an occupation-based approach. In concrete terms, I run equation 7 for four successive decades, namely 1850, 1860, 1870 and 1880 which allows me to cover roughly 3 million

American individuals. Results are graphically summarized in Figures 10a, 10b, 11a and 11b.

FIGURE 10: Extractive industries and XIXth century individual education, 1850-1860



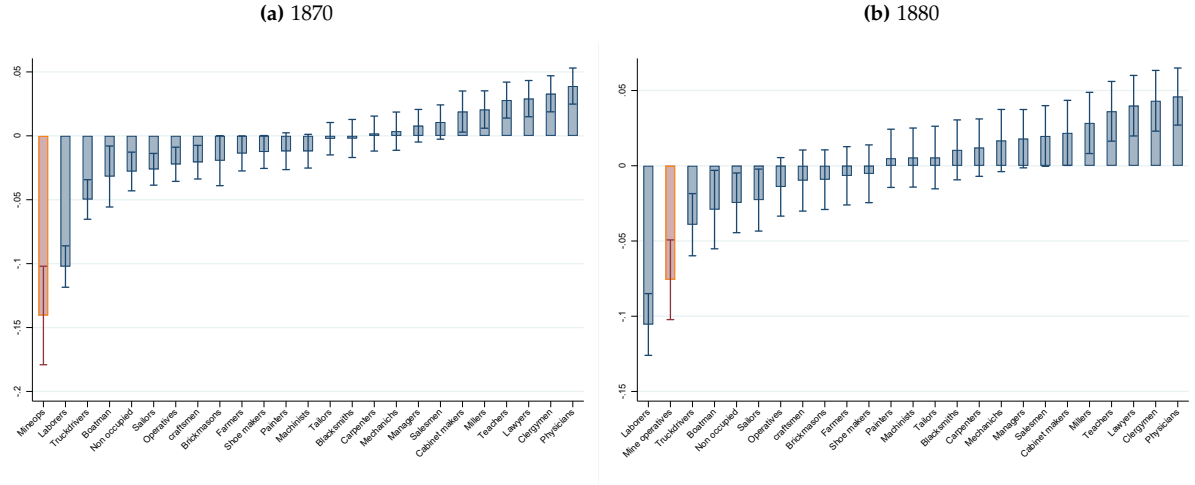
Notes: These graphs show the coefficients attached to categorical variables depicting individual occupations; while there are hundreds of categories in the American censuses, I only report here estimates for the 25 most important categories in numeric terms. Panel (a) plots the results obtained for the 1850 census while Panel (b) plots the same results for the 1860 census. Standard errors are clustered at the county level.

For the sake of clarity, I only report the estimated coefficients for the 25 most prevalent job occupations in the US at that time. Yet, the pattern of the results would appear very similar if I were to include all 1950-job occupations, in the sense that the negative and significant coefficient attached to extractive industries would still rank among the largest ones in absolute values, lending further credence to the argument that extractive industries were largely associated with a deskilling process. More specifically, the graphs clearly show that extractive industries, the coefficients of which are displayed in orange, systematically rank in the two largest job occupations in terms of negative impact over literacy rates. The dependent variable taken here is indeed a binary variable taking the value 1 if the individual both can read and write, and zero otherwise. Two features clearly stand out: first, extractive industries rank among the most detrimental ones in terms of educational abilities; second, the gap with other kinds of job occupations is large and significant, suggesting that, apart from the "laborers not elsewhere classified", extractive industries were largely associated with a negative and economically sizable effect on education.

5.1.4 An event study design

In order to further assess the relation between coal extraction and long-run educational outcomes, I further exploit the richness of the American censuses to conduct an event-study analysis by leveraging boom periods in the coal industry. Importantly

FIGURE 11: Extractive industries and XIXth century individual education, 1870-1880



Notes: These graphs show the coefficients attached to categorical variables depicting individual occupations; while there are hundreds of categories in the American censuses, I only report here estimates for the 25 most important categories in numeric terms. Panel (a) plots the results obtained for the 1870 census while Panel (b) plots the same results for the 1880 census. Standard errors are clustered at the county level.

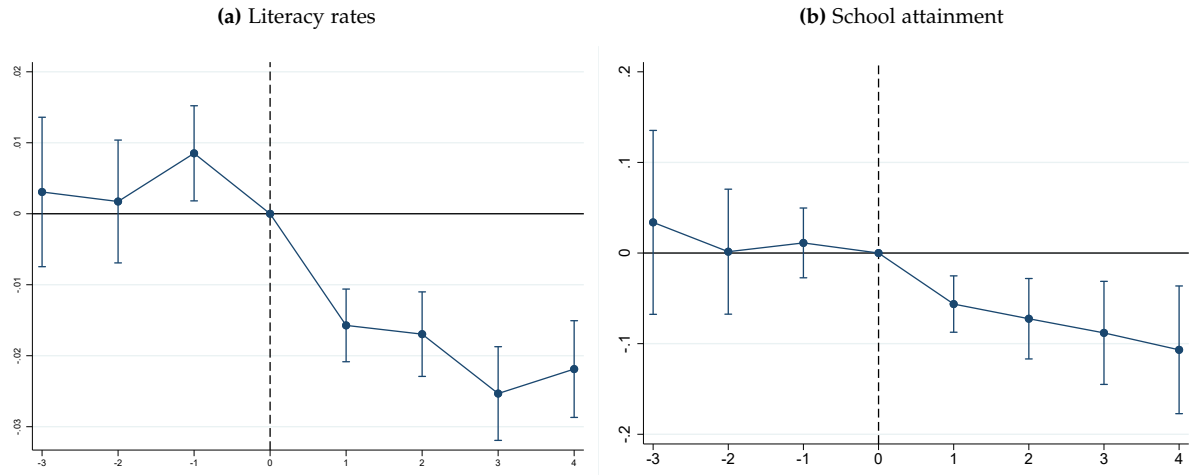
enough, this strategy also allows me to demonstrate the causal impact of coal exposure on innovation, given that I am relying on the estimator developped by De Chaisemartin and d'Haultfoeuille (2024). In the spirit of Black et al. (2005), I consider a classical event study framework of the following form:

$$Y_{c,t} = \sum_{h=-3}^{5} \lambda_h \mathbb{1}[K_{c,t} = h] + \alpha_c + \beta_t + \eta_{c,t}. \tag{8}$$

Where $K_{c_t} = t - E_c$ is the difference between the current year t and the year in which county c experiences a boom in extractive industries and $Y_{c,t}$ is a measure of education attainment in county c at time t. Formally, a county is said to experience a boom in the extractive industry whenever the decennial growth rate of the number of workers involved in the extractive industry exceeds 20%; alternatively, I allow a boom to be defined a situation where the growth rate in the coal industry exceeds the overall local growth rate by a certain threshold. As will be shown in the Appendices, these alternative definitions of a boom in the coal industry do not change the pattern of results. On average, Figures 12a and 12b suggest that coal booms are followed by a decline in the education levels of counties, as proxy by either the share of population who can both read and write or the number of individuals who attended school during childhood. Results are very similar for both measures of education and imply sizable impact of coal booms. Estimates imply that coal booms are associated with a c. 2% permanent decline in the education rates. That coefficients are very similar to one another is expected in the light of the fact that, without school attainment at that time, being able to read and write was barely possible. This complementary evidence brings

further credence to the idea that, by increasing the profitability of industries which were found to require little to no specific skills, coal booms substantially reduced the incentive to acquire education and knowledge in the long run.

FIGURE 12: Impact of a coal boom on educational outcomes



Notes: These graphs show the results of the event study analysis introduced in Equation 8. Panel (a) shows the impact of a coal boom on the share of the population who can both read and write while Panel (b) shows the same impact on the share of the population who has been attending school. Estimation is achieved using the De Chaisemartin and d'Haultfoeuille (2024) approach. Standard errors are clustered at the county level.

5.2 Industry specialization

An alternative reason that could explain the sustained negative impact of coal exposure on productivity would be the existence of a technological lock-in, whereby departments heavily dependent on coal failed to diversify out of specific industries associated with coal mining. In order to evaluate that channel, I start by bringing evidence that departments abundantly endowed with coal exhibited, on average, a higher concentration of innovation throughout the period under study (Section 5.2.1). I then complement this finding by showing that coal exposure induced a strong exposure to myopic industries, meaning industries that fail to incorporate knowledge stemming from alternative fields which reduced the innovation scope and, subsequently, the innovation activity of extractive departments (Section 5.2.2). Finally, I show that individual decisions reinforced the dependence on a small number of industries through within-family transmission patterns (Section 5.2.3).

5.2.1 Lack of indutrial diversity

It has been stressed at the end of Section 4 that the causal effect of coal extraction on innovation worsened throughout the XXth century with a noticeable increase in

the beginning of the sixties. A key mechanism which could account for both findings could stem from the over-specialization induced by extractive activities: as suggested by the theoretical argument made in Section 2, specialization in extractive industries may not be detrimental in the short-run, provided that the demand for extractive products is large enough. Yet, specialization in such industries entail two negative aspects in the long run: first, it exposes extractive basins to idiosyncratic shocks in the form of increased competition from alternative energy sources; second, it prevents them from diversifying in alternative industries which necessarily curtails innovation activity in the long run, given that mining activities soon became technologically mature.

In order to investigate that potential channel, I substitute two concentration measures of innovation for the dependent variable $Y_{d,t}$ in equation 3; more specifically, I start by computing a measure of technology concentration $\mathcal{C}^a_{d,t}$ in the spirit of Agrawal et al. (2009) by summing the square of the shares of each 3-digit IPC technology class at the department $\times$ year level. Formally, $\mathcal{C}^a_{d,t}$ writes:

$$\mathcal{C}^a_{d,t} = \sum_{c \in \mathcal{T}} \left(\frac{P_{d,c,t}}{P_{d,t}} \right)^2 \tag{9}$$

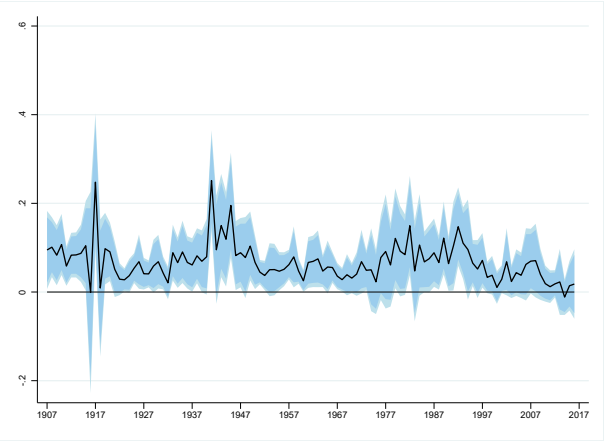
Intuitively, the larger the value taken by the RHS of equation 9, the more innovation tends to depend on a few industries. As a complementary measure of technological concentration, I also compute $s^3_{d,t}$, namely the share of top 3 3-digit IPC classes at the department $\times$ year level in order to ensure that my estimates are not driven by one specific definition of technology concentration. I then run equation 3 controlling for the usual department-level covariates. I summarize the main results obtained using the IV strategy in Figures 13a and 13b.

Exposure to the extractive industries induced a positive, significant and lasting impact over technology concentration according to both measures throughout the XXth century. While the estimates are somehow noisy for a few years in the sample, most of them are statistically significant, either at the 5% or even at the 1% level. Correspondingly, and as will be derived in the Appendices, pooled OLS estimates are also positive and statistically significant at the 1% level, suggesting that innovation tended to be more concentrated around a few technologies in departments heavily involved in extractive industries. In turn, those departments were also more exposed to idiosyncratic shocks, and probably less able to draw on inventions stemming from technology classes they were not involved in.

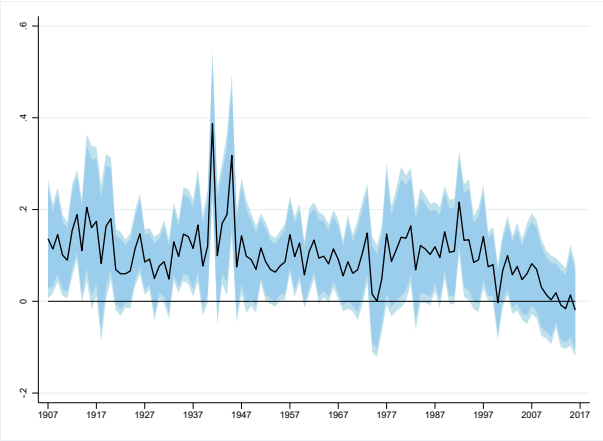
In order to further assert the role played by extractive industries in shaping technology concentration at the local level, I once again leverage the granularity of the

FIGURE 13: Coal extraction and patent concentration

(a) Technology concentration



(b) Top 3 IPC classes



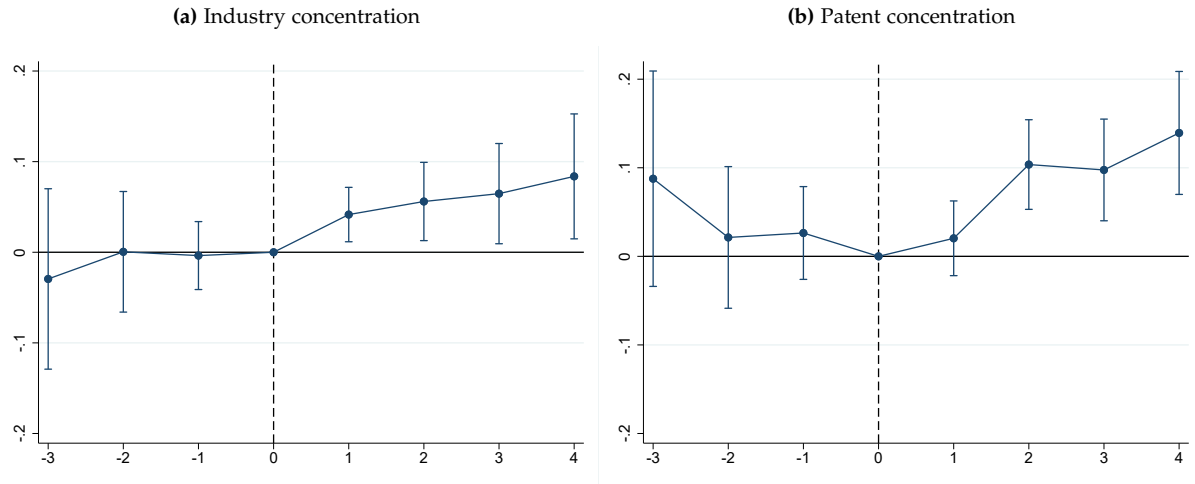
Notes: These graphs show the coefficients attached to the coal exposure variable when running equation **??** successively for all years between 1907 and 2016, and where the dependent variable is taken to be technological concentration as measured using equation 9 (Panel (a)) or using the share of the top 3 technology class in department year innovation. Standard errors are clustered at the department level.

Ruggles et al. (2024) in order to investigate the impact of a coal boom, as defined in section 5.1.4, on the concentration of industrial activity at the local level. More specifically, I look at the impact of a coal boom on the concentration index of both industry and patenting activity; this is like substituting a measure of industry or technology concentration for education measures in Equation 8. Results of this exercise are summarized in Figure 14, which clearly demonstrates a positive, significant and causal impact of boom mining on concentration of both industrial activity and patenting activity. Importantly enough, the effect is strong and lasting, as suggested by the fact that coefficients are still statistically significant at the 5% level more than fourty years after the shock occured. This is consistent with the intuition that extractive industries shape industry specialization in the long run and, hence, might exert a substantial impact on productivity growth.

5.2.2 Creative myopia

Innovation concentration may not be an evil *per se*. Many papers have indeed recently laid the emphasis on the rising concentration of industrial and innovative activities especially in the United States (see, for instance, Kerr and Robert-Nicoud (2020)) and documented positive spillovers emanating from such clusters, rather than negative impacts from increased specialization (see, among others, Glaeser et al. (1992), Gruber et al. (2023) and Moretti (2021)). In that respect, it might be argued that identifying a higher concentration of innovative activities across a few industries within coal basins may reflect a higher value of innovation rather than a lock-in effect. Innovation con-

FIGURE 14: Coal extraction and patent concentration: event-study



Notes: These graphs show the impact of a mining boom as defined in Section 5.1.4 on industry concentration (Figure 14a) and on patent concentration (Figure 14b) using the estimator by De Chaisemartin and d'Haultfoeuille (2024). Confidence intervals are shown at the 95% level and standard errors are clustered at the county level.

centration actually turns to a curse whenever, as pointed by Agrawal et al. (2009), it is associated to a propensity to disproportionately draw from a narrow set of inventions. Said differently, departments heavily involved in extractive industries might exhibit a tendency to focus excessively on some past inventions without trying to incorporate frontier knowledge. In order to investigate that channel, I assemble a dataset at the country, 3 digit IPC class, year level and, following Agrawal et al. (2009), I compute creative myopia in country c for technology class f at time t as:

$$M_{c,t}^{f} = \frac{1}{N_{c,t}^{\mathcal{A}}} \sum_{a \in \mathcal{A}} \frac{\tilde{C_{c,t}^{f}}}{C_{c,t}^{f}} \tag{10}$$

Where $M_{c,t}^{f}$ is the average technological myopia in technological field f in country c at time t, $N_{c,t}^{\mathcal{A}}$ 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, $\tilde{C_{c,t}^{f}}$ 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 $C_{c,t}^{f}$ is the number of citations made by patents of technology field f in country c at time t. Intuitively, the larger $M_{c,t}^{f}$, 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 consider the following regression:

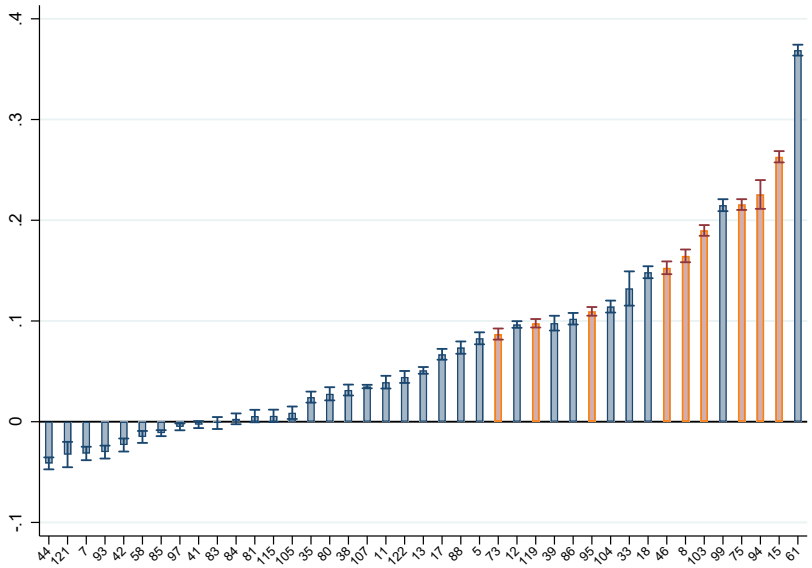
$$M_{c,t}^{f} = \alpha_t + \eta_c + \sum_{k \in \mathcal{T}} \beta_k \mathbb{1}[k = f] + \varepsilon_{f,c,t} \tag{11}$$

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 pertain to the individual effect of each 3-digit IPC class on the level of technological myopia. Results from this specification are shown in Figure 15 where I only plot coefficients which are larger than -0.1; said differently, this plot only pictures estimates which are large enough, so that the graph may still be read easily. Coefficients in orange pertain to technology classes that were over-represented in French coal basins throughout the XXth century; a technology i is said to be over-represented in a coal basin whenever its ratio $\mathcal{R}_i$ of over-representation belongs to the top 25% of the distribution across all technological classes. Formally, $\mathcal{R}_i$ writes:

$$\mathcal{R}_i = \frac{s_i^d}{s_i^c} \tag{12}$$

Where s_i^d captures the share of technology i in the innovation made by department d while s_i^c captures the share of technology i in total innovation made by country c. The larger this value, the more technology field i tends to be disproportionately represented in department d suggesting that department d specialized in that technology throughout the period under study. Strikingly, Figure 15 suggests that across technology classes which were most prevalent in coal basins, many of them exhibit large levels of creative myopia. Typically, categories D05, which refer to textile inventions, and F17, which comprise inventions related to gas storage, are among the 5 most myopic technology classes throughout the XXth century and are both over-represented in coal basins. More importantly perhaps, out of the 10 most myopic technology classes, 6 belong to the top 25% technology classes in terms of over-representation across coal basins. These results are indicative of a bias exhibited by departments involved in coal extraction towards technologies that fail to incorporate knowledge emanating from other fields. Theoretically speaking, this finding is consistent with the intuition that a larger value of θ in the model, namely largely available natural resources, may be negatively correlated with λ in the sense that highly concentrated and myopic industries are less likely to achieve breakthrough inventions compared to diversified and open industries.

FIGURE 15: Creative myopia by technological field



Notes: This Figure plots the results from equation 11 and shows the coefficients obtained for technology classes with the highest levels of technology myopia; bars pictured in orange pertain to technology classes which are over-represented in coal basins. Standard errors are clustered at the technology class level.

5.2.3 Individual-level path dependence

On top of the industry-level concentration and myopia evidenced in the previous subsection, I provide evidence of an individual level lock-in, in the sense that children born in mine operative families were less likely to (i) occupy a different job category than their parents' one and (ii) to achieve high-value added educational and career paths. In order to identify such channels, 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. Consistently with the strategy adopted in the previous subsection, I run models akin to Equation 13:

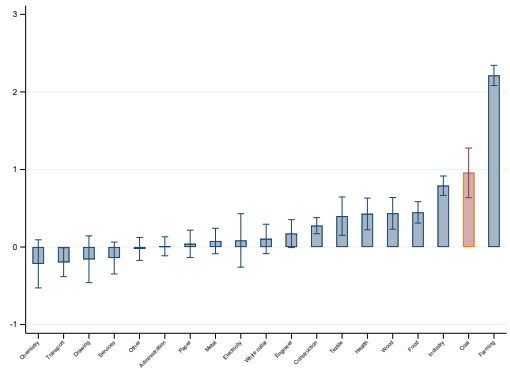
$$\mathcal{Y}_i = \Phi(\sum_{j \in C} \mathbb{1}(j) + \Gamma \mathcal{Z}_i + \varepsilon_i) \tag{13}$$

Where Y_i captures a binary outcome of interest for individual i, whether it be the fact of occupying the same job as one's father, achieving higher education or having a certain kind of job occupation, $\mathbb{1}(j)$ takes the value 1 if the father's respondent occupied job category j and $\mathcal{Z}_i$ is a vector of individual-level characteristics, such as sex, age or birth department. Technically, the occupation of a respondent's father is defined by the 1975 employment code following the typology of INSEE at that

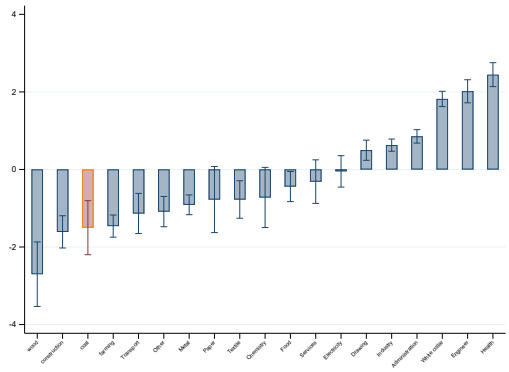
time. While this classification is detailed at the 4-digit level, I focus on a broader classification of jobs which I detail in the Appendix; it basically consists in grouping consistent 2-digit NAF occupations together to end up with 20 job categories which are more tractable and easier to compare.

FIGURE 16: Father'job, within-family job transmission and higher education

(a) Father's occupation and job transmission



(b) Father's occupation and higher education



Notes: Figure (a) plots the probit coefficients when testing Equation **??** 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 when the dependent variable is a binary variable taking the value 1 if respondent achieved higher education. For the sake of readability, coal professions are pictured in orange while other professions are shown in blue.

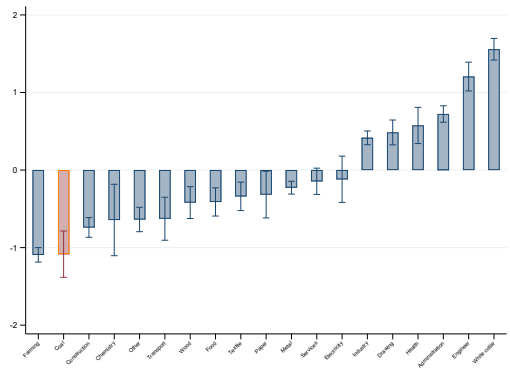
I then leverage this specification to consider different kinds of outcomes Y_i: I first define Y_i to take the value 1 if respondent i occupies the same job category as her father; I second define Y_i to take the value 1 if respondent i achieved higher education; third, I fix Y_i to be 1 when the respondent occupies a white collar job; fourth and finally, I set Y_i to 1 when respondent i occupies a high-value added job, meaning a job which fits in one of the following categories: engineer, teacher, scientists or white collar. In short, the first test captures within-family transmission of jobs; the second investigates the impact of the different father's occupations on the probability f achieving higher education; the third considers the impact of having a father working as a mine operative on the likelihood of becoming a white collar; the fourth finally explores the effect of having a father working as a mine operative on the probability of holding a high-value added job.

Results of the first and second tests are introduced in Figure 16 which shows (i) a positive and strong impact of having a father working as a mine operative on the likelihood of within-family job transmission and (ii) a negative and strong impact of having a father working as a mine operative on the likeihood of achieving higher education. 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-related jobs come second and are associated to a 20% higher chance of inter-

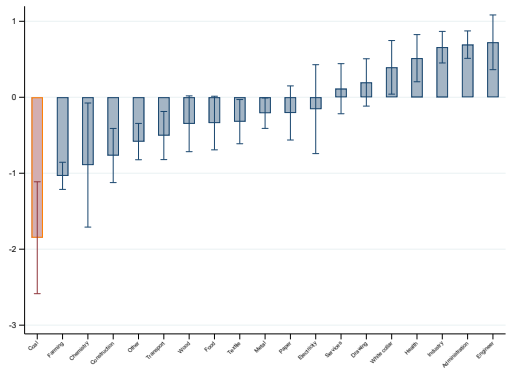
generational job transmission. To get a sense of the magnitude of this coefficient, it is telling to observe that the same coefficient for metal and construction workers correspond, respectively, to a 2.4% lower chance of working in the same field as one's father, and a 5.3% higher chance of working in the same field as one's father. In terms of the effect of father's job on higher education, having a father working as a mine operative reduced the chance of achieving higher education by 5.2%, an outcome equivalent to that obtained for the kids of farming workers, but much more detrimental than that derived for all other job categories.

FIGURE 17: Father's job effect on probability of occupying a high-value added position

(a) High-value added job



(b) White collar jobs



Notes: Figure (a) plots the coefficients when testing Equation **??** with a binary variable taking the value 1 if the individual occupies a high-value added job and 0 otherwise while Figure (b) plots the coefficients when testing Equation **??** 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.

Finally, results of the third and fourth tests are summarized in Figure 17 which conveys the ideas that kids of mine operatives were among the least likely to achieve a career either as a white collar or, more generally speaking, as a high-value added employee. According to the estimates, having a father working in extractive industries translates in a 8% lower likelihood of working as a white collar, while decreasing the chance of occupying a high-skilled position by roughly 10.8%, an effect which is very close to that obtained for kids of farming 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%). In turn, these estimates imply that, even if a given department was willing to shift its industry structure and invest in newer activities, it was still constrained by path dependence at the individual level.

5.3 Alternative mechanisms

Industry specialization and human capital have proved to be key mechanisms accounting for the sustained pattern in the effect of coal exposure on French innovation

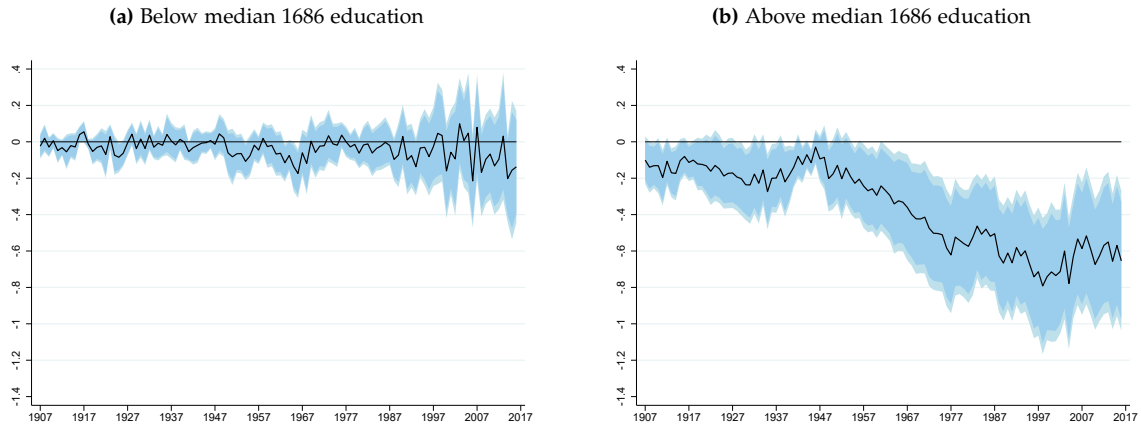
in the course of the XXth century. Alternative mechanisms may nevertheless have played an important role as well. In this subsection, I briefly investigate three complementary explanations, namely the reversal of fortune story, the early industrialization effect (see Franck and Galor (2021), Franck and Galor (2021)) and the institutional argument akin to the dutch disease story. As summarized below, I find some evidence in support of the reversal of fortune story (Section 5.3.1) as well as in support of the institutional argument (Section 5.3.2) but no sign that early industrialization played a key role (Section 5.3.3).

5.3.1 Reversal of fortune

A possible channel could resemble the intuition elaborated by Acemoglu et al. (2002), namely the concept of a *reversal of fortune*. Exposure to extractive industries *per se* might in fact not be sufficient to account for the persistence in the long-term in the observed innovation trajectories. In particular, depending on the initial position of cities in the wealth distribution at the beginning of the First Industrial Revolution, the effect of proximity to coal resources might have shown substantial heterogeneity. While this channel is difficult to test given the lack of precise and reliable data at the local level in France before the industrial revolution, I nevertheless draw on the recently assembled dataset by Piketty (2023) to provide an intuition of the reversal of fortune experienced by cities located nearby coal resources in the end of the XVIIIth century and which exhibited a high level of *average* education. The point has been made before that *average* and *upper tail* education may have played different roles in the course of Industrial Revolution. Namely, *average* skills may have been important in shifting from an agrarian economy to an industrial one, while *upper-tail* human capital was necessary to ensure a fast increase in productivity growth rates. I thus investigate whether a high level of *average* human capital before the industrial revolution translated in lower innovation activity in the long run; the intuition for this assumption would be that coal extraction was possible only to the extent that the average level of education at the local level was sufficiently large to allow for industrial development, by contrast with an agrarian economy. If a given city was located nearby coal resources, but lacked the average education to start extracting resources on a large scale, it could then be protected from the negative long-run impact of the extractive industry compared to a city located nearby coal resources as well, but showing a sufficient amount of skills to get involved in extractive activities. Reversal of fortune would then occur, in the sense that relatively more educated cities - from the standpoint of average education - would be negatively affected by coal exposure while less educated cities would not.

In order to evaluate this channel, I I run repeated cross section regressions as modeled by Equation 3 at the city level separately for two kinds of cities, respectively those below and above the median education level of 1686. In order to proxy for *average* education in the end of the XVIIth century, I exploit the share of spouses that could sign their wedding act which I obtain from the aforementioned Piketty (2023) database. Intuitively, cities with an above-median level of education before the industrial revolution were, all else equal, cities better equipped in order to deal with the challenges that the new technologies of that time opposed.

FIGURE 18: Reversal of fortune: heterogenous impact of coal exposure by initial education level



Notes: Figure (a) plots the impact of coal exposure on innovation in the long run using Equation **??** for cities with a below median education level as of 1686 while Figure (b) plots the same value but for cities exhibiting an above median level of education as of 1686. Confidence intervals at the 95% level are plotted in lighter blue while darker blue confidence intervals show 90% confidence intervals.

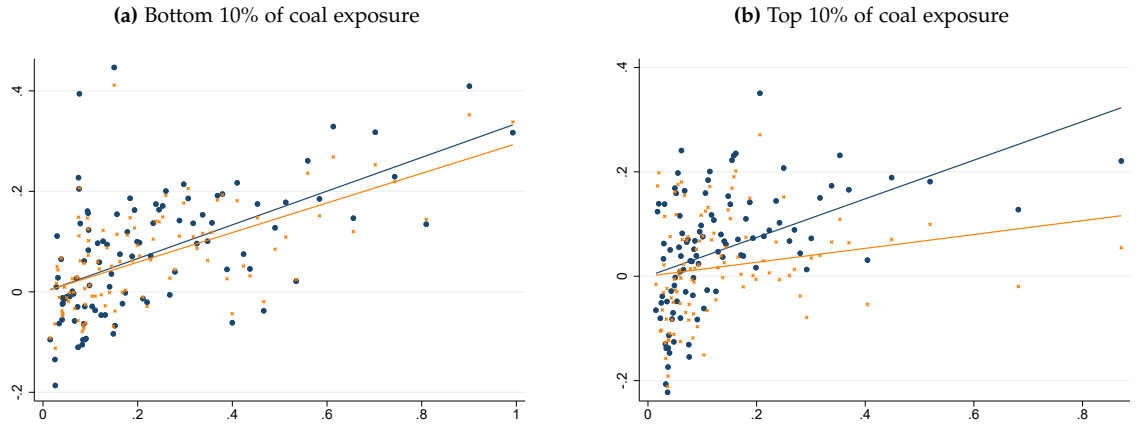
Figure 18 contrasts the effect of coal exposure depending on whether cities belonged to the below (Figure 18a) or the above median group of education distribution in 1686 (Figure 18b). The figures suggest that there was a strong heterogeneity in the effect of proximity to coal: while the below median group was barely affected, the above median group was negatively and significantly affected throughout the XXth century, with a marked deterioration in the beginning of the second half of the XXth century. As will be shown in the Appendices, those results remain valid if we split by the 1786 education level, or if we consider alternative quantiles of education distribution. It should nevertheless be noted that these results do not violate the usual results on the role played by education in fostering economic growth and innovation in the long run (see, among others, Romer (1990), Mokyr and Voth (2010), Squicciarini (2020) or Squicciarini and Voigtländer (2015)); indeed, it is important to distinguish between average and upper-tail education when it comes to understand the determinants of economic growth. While the latter is decisive, there is little evidence that the former played any role.

What the aforementioned results, however, suggest is not that education was detri-

mental to industrial progress, but rather that education before the industrial revolution might have adversely affected later economic performance *conditional* on being located nearby coal. A plausible interpretation for that finding could be that a higher level of education before the industrial revolution was a key advantage in order to make the most of the inventions elaborated at that time. More specifically, extracting natural resources in large quantities was demanding in terms of new techniques. Therefore, only cities with a sufficiently large stock of human capital could get actively involved in the new industries brought about by the industrial revolution. Given that these hovered around steam generation, which required coal extraction, it could be that only cities with a sufficiently large education level decided to extract coal on a large scale.

This interpretation is akin to assuming that such cities experienced a reversal of fortune at the onset of the industrial revolution: while they relied on a large stock of human capital before the industrial revolution, they still did not manage to transform it into a higher prosperity in the course of the XIXth century. A way to test this interpretation would be to study the evolution of the correlation between initial education and subsequent measures of economic progress. More specifically, should the assumption posited above be true, one would expect the correlation between initial education and economic prosperity to weaken over time for cities located nearby coal resources. By contrast, such a correlation is not doomed to fade for cities further away from coal resources. While such a test is tricky to run properly, both from a methodological and from a data point of view, I nevertheless provide intuitive evidence in support of this idea.

FIGURE 19: Weaker correlation between initial education and later prosperity for coal basins



Notes: Figure (a) plots the correlation between the initial education level and respectively the 1790 percentile of revenue in the national distribution (navy blue) and the 1900 percentile of revenue in the national distribution (orange) for cities located in the bottom 10% of coal exposure while Figure (b) plots the same correlation but for cities ranking in the upper 10% of the coal exposure measure.

The results of this test are shown in Figure 19 which supports the idea that the relation

between initial education and revenue did not evolve in the same way over time depending on whether we consider cities located close to coalfields (Panel (b)) or cities located further away from coalfields (Panel (a)). Indeed, there seems to exist no significant difference in terms of the relation between initial exposure and economic outcomes in the first panel; in words, this means that initial education exerted the same positive effect on local revenue in 1790 than in 1900 for cities further away from coalfields. By contrast, the positive impact of initial education on revenues of cities closed to coalfields, which was positive, significant and larger than the coefficient obtained for cities further way from coalfields in 1790, turns insignificant as of 1900, suggesting that, in the course of a century, such departments experienced a sudden decline in their growth rate.

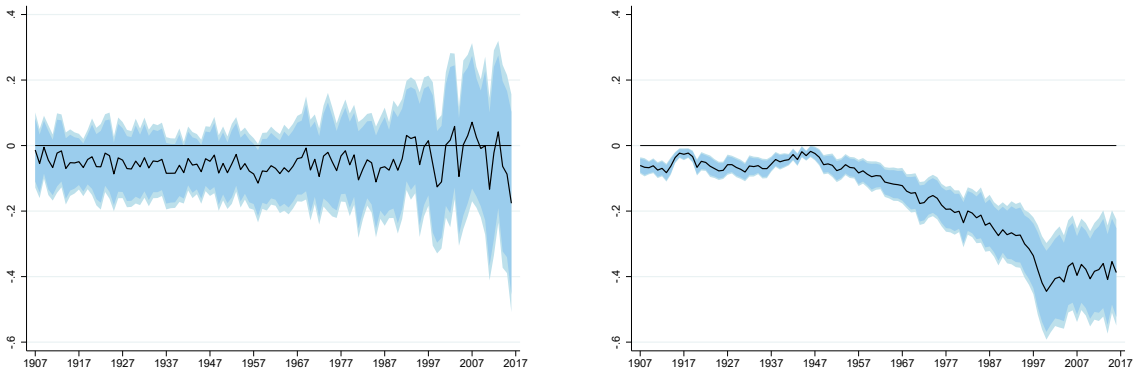
5.3.2 The role played by institutions

Since North (1991) at least, institutions have long been acknowledged by both historians and economists to play a significant role in shaping prosperity. Leveraging historical events, mounting evidence indeed suggests that differences in institutions explain a substantial part of long-run growth patterns (see, among others, Acemoglu et al. (2001b), Acemoglu et al. (2002), Dell (2010), Nunn (2007), Glaeser and Shleifer (2002)). More related to our topic is the literature which studies the interplay between natural resources and institutions: in particular, Mehlum et al. (2006) observes that countries abundantly endowed with natural resources exhibit both "winners" and "losers", in the sense that some of these countries manage to achieve high economic performance while others fail to experience economic development. Fundamentally, countries with larger institutional development are expected to better manage natural resources and hence transform them from a curse to a blessing.

However, testing this channel in the French context proves intricate, at least for two reasons. First, France is a highly centralized country, so that it is unlikely that there exists substantial institutional differences at the local level. Second, exhaustive data collection on French economic and political development started in the beginning of the XIXth century, a period when coal mining had already started. In order to mitigate these concerns, and bring tentative evidence of the potential heterogeneity in the effect of coal mining depending on local institutions, I once again leverage the dataset by Piketty (2023). More specifically, I compute for each French city the share of citizens who participated in the 1795 referendum and use it as a proxy for institutional development. I then split the sample of French cities in two subsets, namely the group of French cities belonging to the top 10% cities in terms of participation rate in the 1795 referendum and the group of French cities belonging to the bottom 90%. Participation

FIGURE 20: Early institutionalization and long run patenting - 1

(a) Top 10% **(b)** Bottom 90%



Notes: Figure (a) plots the effect of exposure to coal on city-level patenting activity for cities belonging in the top 10% of the distribution in terms of participation rate to the 1795 referendum while Figure (b) plots the same effect but for cities belonging in the bottom 90% distribution in terms of participation rate to the 1795 referendum. Standard errors are clustered at the department level.

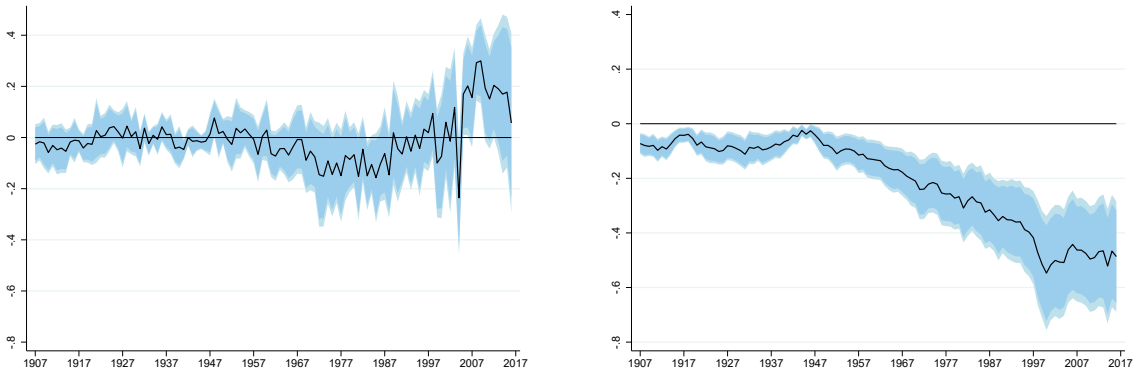
in the electoral debate is indeed a key ingredient of institutional development; the larger the initial level of participation before the industrial revolution, the larger the institutional development before coal mining started must have been.

I then simply run Equation 3 at the city level, and plot the results obtained for the two groups of cities in Figure 20 which pertains to a different effect of coal exposure on cities depending on whether they had a relatively high or low initial level of institutional development. Although the coefficients hover around 0 and are insignificant for the group of cities with an initial high level of institutional development, they are negative and significant throughout the XXth century for cities with a relatively lower level of institutional development. For the sake of comparability, the same scale is used for the two graphs, strongly lending credence to the idea that the patterns in both groups were substantially different throughout the XXth century. In terms of interpretation, those estimates suggest that two cities with the same level of proximity with respect to a large coalfield, but exhibiting different initial levels of institutional development, did not react in the same way to coal mining. This evidence would, however, require a more careful examination which exceeds the scope of this analysis: in particular, it could be that the participation rate in the 1795 referendum reflects other economic or development components which I fail to capture.

In order to ensure that these estimates are not driven by anything peculiar to the 1795 referendum, I investigate the institutions channel in a second framework whereby I leverage the newly released dataset by Gay (2025) who compiled all social events in France spanning 1661 to 1789. Interestingly enough, the dataset allows to weigh each event - whether it be a riot, an assassination of an official - by its importance relative to other events of the sample. Following a process detailed in the Appendices, I thus

FIGURE 21: Early institutionalization and long run patenting - 2

(a) Top 50% **(b)** Bottom 50%



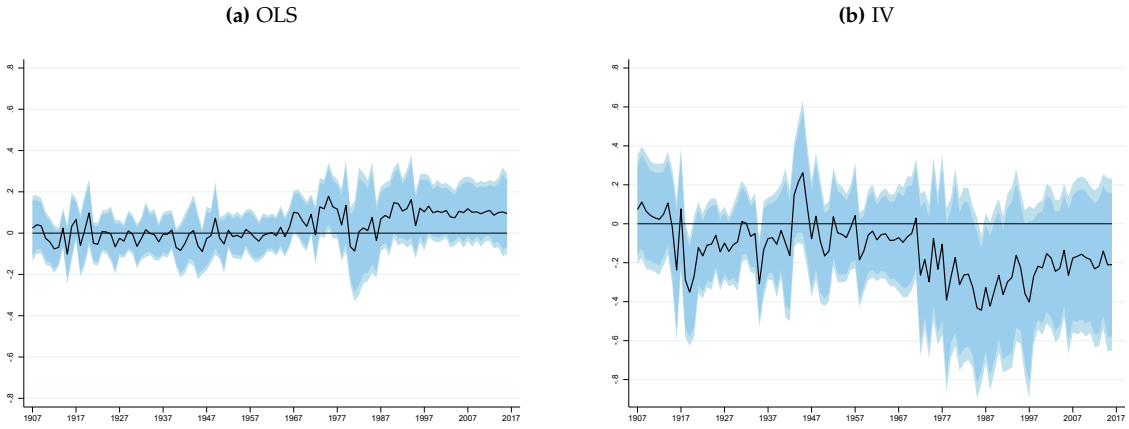
Notes: Figure (a) plots the effect of exposure to coal on city-level patenting activity for cities belonging in the bottom 50% of the distribution in terms of social unrest while Figure (b) plots the same effect but for cities belonging in the top 50% distribution in terms of social unrest, the intuition being that a higher value of social unrest exemplifies a lower level of local institutional development. Standard errors are clustered at the department level.

compute for each French city an indicator of exposure to social unrest from 1661 to 1789: the interest of that measure is that it enables me to capture social unrest *prior* to the expansion of French coal industry. I then simply split the group of French cities depending on whether they belong to the upper or the lower tail of the distribution in terms of exposure per capita, and check whether the results of this specification are consistent with those summarized above. This exercise is summarized in Figure 21 which brings credence to the argument that the negative effect of natural resources on innovation and industrialization is magnified whenever institutional development is poor, while it is largely curtailed when initial institutional development is strong enough. Again, while the effect of exposure to natural resources is noisy and neutral throughout the XXth century for cities with a larger initial institutional development (left Panel), the effect of natural resources on innovation in cities with a higher exposure to social unrest seems to have been negative and significant throughout the XXth century, with a marked deterioration in the after-war period. Given that the same scale is applied on both graphs, comparability is perfectly ensured and, again, brings convincing evidence in favor of the argument that institutions are key to economic development in the long run. One caveat of this approach, nevertheless, is to consider institutions rather as an initial ingredient of economic development, instead of assuming that it could be a variable endogenous to the economic process of rent extraction.

5.3.3 Early industrialization

In a pioneering paper mentioned previously, Franck and Galor (2021) suggest a negative causal relationship between early industrialization and long run development. They instrument the number of steam machines per capita with the distance to Fresnes Sur Escaut, the first French city where a steam engine was to be found in the XVIIIth century, and find that a larger number of steam engines in the second half of the XXth century was conducive to lower levels of economic development in the long run. In their view then, early industrialization turned from a blessing to a curse in the course of the XXth century. Rather than coal exposure, it could be that early industrialization, as proxy by the initial number of steam machines per capita, is the true explanatory variable behind our results. In order to test this channel, I simply run Equation **??** by adding the number of steam machines per department as computed in their paper and test whether this variable is a key predictor of long run patenting activity.

FIGURE 22: Early industrialization and long run patenting



Notes: Figure (a) plots the effect of early industrialization as proxy by the number of steam machines per capita in the 1860-1865 period on total patenting activity while Figure (b) plots the same effect but when the number of steam engines is instrumented by the distance to Fresnes sur Escaut, the first French city to have a steam engine during the industrial revolution.

I derive results both for OLS and for IV estimates, using the same instrument they apply in their paper, and controling for the distance to coalfields. Results of this alternative mechanism are displayed in Figure 22 which rejects the assumption that early industrialization rather than coal resources might have been driving productivity down in the long run. Indeed, none of the coefficients obtained on exposure to steam engines are found to be significant in the course of the XXth century. The estimates are larger and more negative in the IV framework, yet coefficients are not significant, neither at the 5% nor at the 10% level; on top of that, and as will be shown in the Appendices, the estimates on the effect of coal exposure remain very stable in spite of the inclusion of this further control. Hence, there is no positive evidence

of a role played by early industrialization, rather than coal exposure. Interestingly enough, the negative (though not significant) effect in the IV specification appears in the aftermath of World War II, more specifically in the 60s, a period when de industrialization appeared in formerly industrialized places; in that respect, the exposure to steam engines seem rather to reflect the industrial decay than the long run negative impact associated to extractive industries.

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 specialization 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.

_ , _ , and _ , "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review*, December 2001, *91* (5), 1369–1401.

_ , _ , and _ , "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution*," *The Quarterly Journal of Economics*, 11 2002, *117* (4), 1231–1294.

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.

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.

Akcigit, Ufuk, Salomé Baslandze, and Stefanie Stantcheva, "Taxation and the International Mobility of Inventors," *American Economic Review*, October 2016, *106* (10), 2930–81.

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.

_ , Arthur Guillouzouic, Emeric Henry, and Clement Malgouyres, "From public labs to private firms: magnitude and channels of R&D spillovers," CEP Discussion Papers dp1882, Centre for Economic Performance, LSE October 2022.

Black, Dan A., Terra G. McKinnish, and Seth G. Sanders, "Tight Labor Markets and the Demand for Education: Evidence from the Coal Boom and Bust," *ILR Review*, October 2005, *59* (1), 3–16.

Bonnet, Florian, Hippolyte d'Albis, and Aurélie Sotura, "Income Inequality across French Departments over the Last 100 Years [Les inégalités de revenu entre les départements français depuis cent ans]," PSE-Ecole d'économie de Paris (Postprint) hal-03374099, HAL October 2021.

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

Caselli, Francesco and Guy Michaels, "Do Oil Windfalls Improve Living Standards? Evidence from Brazil," *American Economic Journal: Applied Economics*, January 2013, *5* (1), 208–38.

Chaisemartin, Clément De and Xavier d'Haultfoeuille, "Difference-in-differences estimators of intertemporal treatment effects," *Review of Economics and Statistics*, 2024, pp. 1–45.

Châtel, Dollfus, *Atlas des Houillières européennes*, Société de Documentation Industrielle, 1931.

Colella, Fabrizio, Rafael Lalive, Seyhun Sakalli, and Mathias Thoenig, "ACREG: Stata module to perform Arbitrary Correlation Regression," Statistical Software Components, Boston College Department of Economics Dec 2020.

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.

Dell, Melissa, "The Persistent Effects of Peru's Mining Mita," *Econometrica*, November 2010, *78* (6), 1863–1903.

Dittmar, Jeremiah and Ralph R. Meisenzahl, "The research university, invention and industry: evidence from German history," CEP Discussion Papers dp1856, Centre for Economic Performance, LSE June 2022.

Dyevre, Arnaud, "Public R&D Spillovers and Productivity Growth," 2023. Mimeo LSE.

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

Gay, Victor, "The Jean Nicolas Database. The French Rebellion, 1661-1789," 2025.

Glaeser, Edward L. and Andrei Shleifer, "Legal Origins," *The Quarterly Journal of Economics*, 2002, *117* (4), 1193–1229.

Glaeser, Edward L, Hedi D. Kallal, Jose A. Scheinkman, and Andrei Shleifer, "Growth in Cities," *Journal of Political Economy*, December 1992, *100* (6), 1126–1152.

Gross, Daniel P. and Bhaven N. Sampat, "America, Jump-started: World War II R&D and the Takeoff of the U.S. Innovation System," NBER Working Papers 27375, National Bureau of Economic Research June 2020.

Gruber, Jonathan, Simon Johnson, and Enrico Moretti, "Place-Based Productivity and Costs in Science," *Entrepreneurship and Innovation Policy and the Economy*, 2023, *2* (1), 167–184.

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.

Kerr, William R. and Frederic Robert-Nicoud, "Tech Clusters," *Journal of Economic Perspectives*, August 2020, *34* (3), 50–76.

Krugman, Paul, "The narrow moving band, the Dutch disease, and the competitive consequences of Mrs. Thatcher : Notes on trade in the presence of dynamic scale economies," *Journal of Development Economics*, October 1987, *27* (1-2), 41–55.

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.

Matheis, Mike, "Local Economic Impacts of Coal Mining in the United States 1870 to 1970," *The Journal of Economic History*, December 2016, *76* (4), 1152–1181.

Mehlum, Halvor, Karl Moene, and Ragnar Torvik, "Institutions and the Resource Curse," in Roger D. Congleton, Kai A. Konrad, and Arye L. Hillman, eds., *40 Years of Research on Rent Seeking 2*, Springer Books, Springer, July 2006, pp. 245–264.

Mokyr, Joel, "Long-Term Economic Growth and the History of Technology," in Philippe Aghion and Steven N. Durlauf, eds., *Handbook of Economic Growth*, Vol. 1, Amsterdam: Elsevier, 2005, pp. 1113–1180.

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

Moretti, Enrico, "The Effect of High-Tech Clusters on the Productivity of Top Inventors," *American Economic Review*, October 2021, *111* (10), 3328–75.

North, Douglass C., "Institutions," *Journal of Economic Perspectives*, March 1991, *5* (1), 97–112.

Nunn, Nathan, "The Long-Term Effects of Africa's Slave Trades," MPRA Paper 4134, University Library of Munich, Germany July 2007.

Piketty, Julia Cagé Thomas, *Une histoire du conflit politique: élections et inégalités sociales en France, 1789-2022*, Le Seuil, 2023.

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–.

Romer, Paul M, "Endogenous Technological Change," *Journal of Political Economy*, October 1990, *98* (5), 71–102.

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

Ruggles, Steven, Matt A. Nelson, Matthew Sobek, Catherine A. Fitch, Ronald Goeken, J. David Hacker, Evan Roberts, and J. Robert Warren, "IPUMS Ancestry Full Count Data," 2024.

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.

Squicciarini, Mara P and Nico Voigtländer, "Human capital and industrialization: Evidence from the age of enlightenment," *The Quarterly Journal of Economics*, 2015, *130* (4), 1825–1883.

Valero, Anna and John Van Reenen, "The economic impact of universities: Evidence from across the globe," *Economics of Education Review*, 2019, *68*, 53–67.

Vicente, Pedro C., "Does Oil Corrupt? Evidence from a Natural Experiment in West Africa," Economics Series Working Papers 317, University of Oxford, Department of Economics April 2007.

ONLINE APPENDICES

Table of contents

A Theoretical framework **OA-1**

A.1 Proof of consumption good demand OA-1

A.2 Proof of optimal research effort OA-1

A.3 Proof of intermediary good quantity OA-1

A.4 Proof of Proposition 1 . . . . . OA-3

A.5 Proof of profit in the intermediary industry OA-4

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

A.7 Proof of the optimal innovation effort OA-7

A.8 Proof of Proposition 3 . . . . . OA-7

A.9 Proof of the extension OA-8

B Coal measure appendix **OA-11**

B.1 Alternative measures OA-11

B.2 Coal exposure under these alternatives OA-12

C Instrumental variable **OA-13**

C.1 Graphical evidence OA-13

C.2 Regression evidence OA-14

D Additional robustness checks **OA-18**

D.1 Main results under alternative coal exposure measures OA-18

D.2 Flexible functional forms OA-23

D.3 Alternative clustering methods OA-27

A Theoretical framework

A.1 Proof of consumption good demand

Let us recall that the representative firm in the consumer good industry solves the following program:

$$\max_{x} Ax^{\alpha}c^{1-\alpha} - px \tag{A1}$$

Taking the first order condition with respect to x and c trivially yield:

$$\begin{cases} \alpha A x^{\alpha-1} c^{1-\alpha} = p \\ (1-\alpha) A x^{\alpha} c^{-\alpha} = 1 \end{cases} \tag{A2}$$

Combining both equations finally implies that:

$$\frac{x}{c} = \frac{\alpha}{p(1-\alpha)} \tag{A3}$$

A.2 Proof of optimal research effort

Let us recall that the research effort made by the randomly selected skilled worker writes:

$$\Pi^{R} = n\gamma\Pi^{I}(A^{+}) - \frac{n^2}{2} \tag{A4}$$

Taking the first order condition with respect to n implies:

$$\gamma\Pi^{I}(A^{+}) - n = 0 \tag{A5}$$

Which is equivalent to:

$$n^{*} = \gamma\Pi^{I}(A^{+}) \tag{A6}$$

A.3 Proof of intermediary good quantity

The optimality condition for the monopolist in the intermediary good industry writes:

$$\alpha^2 A x^{\alpha-1} c^{1-\alpha} = w_s \tag{A7}$$

In equilibrium, workers must be indifferent between becoming skilled workers and getting the wage rate w_s and staying unskilled and getting the wage rate prevailing in the extractive industry, namely θ. This implies that:

$$\alpha^2 A x^{\alpha-1} c^{1-\alpha} = \theta \tag{A8}$$

This may be rewritten as:

$$x^* = c \left(\frac{\alpha^2 A}{\theta} \right)^{\frac{1}{1-\alpha}} \tag{A9}$$

In the meantime yet, the market clearing condition imposes that:

$$L = L_U + L_S \tag{A10}$$

Meaning that skilled labor and unskilled labor must sum to the overall labor supply in the economy, for labor is inelastically supplied. Given that each unskilled worker produces extractive resources 1 for θ, it follows that:

$$c = \theta L_U \tag{A11}$$

Given that each skilled worker produces intermediary goods 1 for 1, we may write:

$$x + \frac{c}{\theta} = L \tag{A12}$$

This can be expressed as:

$$c = \theta(L - x) \tag{A13}$$

Combining this last equation with Equation A9 implies that:

$$x^* = \theta(L - x^*) \left(\frac{\alpha^2 A}{\theta} \right)^{\frac{1}{1-\alpha}} \tag{A14}$$

Which, after rearranging, implies that x^* writes:

$$x^* = \frac{L\left(\alpha^2 A\right)^{\frac{1}{1-\alpha}} \theta^{\frac{-\alpha}{1-\alpha}}}{1+\left(\alpha^2 A\right)^{\frac{1}{1-\alpha}} \theta^{\frac{-\alpha}{1-\alpha}}} \tag{A15}$$

A.4 Proof of Proposition 1

Proposition 1 states that skilled labor, meaning labor which is allocated to the intermediary good industry, is decreasing in the availability of natural resources θ. Let us first observe that:

$$x^* = \frac{Lz}{1+z} \tag{A16}$$

Where:

$$z = \left(\alpha^2 A\right)^{\frac{1}{1-\alpha}} \theta^{\frac{-\alpha}{1-\alpha}} \tag{A17}$$

It now suffices to compute the variation of f with respect to z, where:

$$f(z) = \frac{Lz}{1+z} \tag{A18}$$

Obviously we have:

$$\frac{\partial f}{\partial z} = \frac{L(1+z) - Lz}{(1+z)^2} \tag{A19}$$

Which simplifies to:

$$\frac{\partial f}{\partial z} = \frac{L}{(1+z)^2} \geq 0 \tag{A20}$$

Then, f is an increasing function of z. Meanwhile, it is straightforward to show that:

$$\frac{\partial z}{\partial \theta} < 0 \tag{A21}$$

Meaning that x^* is a decreasing function of θ. Hence, the larger the availability of natural resources θ, the lower the mass of skilled workers in the economy. Intuitively indeed, the larger θ, the more productive the extractive industry is, and hence, the larger the wage rate for unskilled workers. By arbitrage, this must raise the wage rate in the intermediary industry, meaning for skilled workers; the increase in the

wage rate of skilled workers nevertheless reduces the profitability of the intermediary industry, which reduces its desired output level x^* and correspondingly increases its price level p.

A.5 Proof of profit in the intermediary industry

Profit in the intermediary industry simply writes:

$$\Pi^I = (p - w_s)x^* \tag{A22}$$

Given that w_s is fixed to θ by an arbitrage argument, it follows that:

$$\Pi^I = (p - \theta)x^* \tag{A23}$$

It then suffices to compute the value of p in equilibrium. Equation A2 simply states that:

$$\alpha A x^{\alpha-1} c^{1-\alpha} = p \tag{A24}$$

After rearranging we obtain:

$$p = \alpha A \left(\frac{x}{c}\right)^{\alpha-1} \tag{A25}$$

Using Equation A9, we also have:

$$\frac{x}{c} = \left(\frac{\alpha^2 A}{\theta}\right)^{\frac{1}{1-\alpha}} \tag{A26}$$

Which rearranges to:

$$\left(\frac{x}{c}\right)^{\alpha-1} = \frac{\theta}{\alpha^2 A} \tag{A27}$$

We then have:

$$\frac{p}{\alpha A} = \frac{\theta}{\alpha^2 A} \tag{A28}$$

This means that:

$$p = \frac{\theta}{\alpha} \tag{A29}$$

Plugging this price value in the profit equation finally implies that:

$$\Pi^I = \left(\frac{\theta}{\alpha} - \theta\right) x^* \tag{A30}$$

Which writes:

$$\Pi^I = \left(\frac{\theta}{\alpha} - \theta\right) \theta(L - x^*) \left(\frac{\alpha^2 A}{\theta}\right)^{\frac{1}{1-\alpha}} \tag{A31}$$

Which simplifies to:

$$\Pi^I = \left(\frac{1-\alpha}{\alpha}\right) L \frac{\left(\alpha^2 A\right)^{\frac{1}{1-\alpha}} \theta^{\frac{1-2\alpha}{1-\alpha}}}{1 + \left(\alpha^2 A\right)^{\frac{1}{1-\alpha}} \theta^{\frac{-\alpha}{1-\alpha}}} \tag{A32}$$

A.6 Proof of Proposition 2

Proposition 2 states that the effect of natural resources on the profit rate of the intermediary industry is ambiguous. More specifically, it posits that, whenever α is lower or equal to $\frac{1}{2}$, the effect of natural resources on innovation is unambiguously positive; by contrast, whenever α is greater or equal to $\frac{1}{2}$, there exists a cutoff value $\hat{\theta}$ such that, beyond this value, natural resources lower the profit rate in the intermediary industry. To see that, it suffices to examine the sign of the derivative of Π^I with respect to θ. To simplify computations, it is useful to observe that this derivate will be of the sign of the derivative of:

$$g(\theta) = \frac{\theta^{\frac{1-2\alpha}{1-\alpha}}}{1 + \left(\alpha^2 A\right)^{\frac{1}{1-\alpha}} \theta^{\frac{-\alpha}{1-\alpha}}} \tag{A33}$$

It is now straightforward to compute the derivative of g with respect to θ which writes:

$$\frac{\partial g}{\partial \theta} = \frac{\left(\frac{1-2\alpha}{1-\alpha}\right)\theta^{\frac{-\alpha}{1-\alpha}}(1+y\theta^{\frac{-\alpha}{1-\alpha}}) - \left(-\frac{-\alpha}{1-\alpha}y\theta^{\frac{-1}{1-\alpha}}\right)\theta^{\frac{1-2\alpha}{1-\alpha}}}{(1+y\theta^{\frac{-\alpha}{1-\alpha}})^2} \tag{A34}$$

Where y is defined by:

$$y = \left(\alpha^2 A\right)^{\frac{1}{1-\alpha}} \tag{A35}$$

Given that the denominator of Equation A34 is unambiguously positive, it follows that the derivative of g has the same sign as its numerator, which we call N; simplifying N then yields:

$$N = \theta^{\frac{-\alpha}{1-\alpha}}\left(\left(\frac{1-2\alpha}{1-\alpha}\right) + y\theta^{\frac{-\alpha}{1-\alpha}}\right) \tag{A36}$$

Clearly, whenever $\alpha \leq \frac{1}{2}$, the first term in the bracket on the RHS is positive, while the second is unambiguously positive as the product of two positive terms. Hence, whenever α is low enough, N is unambiguously positive, meaning that Π^I increases with θ. By contrast, whenever α is greater or equal to $\frac{1}{2}$, the sign of N becomes ambiguous given that the second term in the brackets unambiguously decreases with θ while the first is constant. Hence, it is straightforward to show that the term in the brackets decreases with θ; also, given that:

$$\lim_{\theta \to \infty} \theta^{\frac{-\alpha}{1-\alpha}} = 0 \tag{A37}$$

It follows that the limit when θ goes to infinity of the term in the brackets is negative whenever α is greater or equal to $\frac{1}{2}$; given that N is a continuous and well behaved function it is trivial to show that there exists a unique cutoff value $\hat{\theta}$ such that N is positive below that value, and negative above. It suffices to solve: $N = 0$ to recover this value, which trivially yields:

$$\hat{\theta} = \left(\frac{y(1-\alpha)}{2\alpha - 1}\right)^{\frac{1-\alpha}{\alpha}} \tag{A38}$$

Which is precisely what Proposition 2 demonstrated. Hence, whenever α is large enough, meaning whenever the intermediary good is key in the production process of the consumption good, natural resources negatively affect the profit rate of the intermediary industry as long as they are in excess of a certain threshold, which unambiguously increases with the initial productivity level.

A.7 Proof of the optimal innovation effort

It was shown earlier that the optimal innovation effort n^* must solve for:

$$n^* = \gamma \Pi^I(A^+) \tag{A39}$$

Given that we have recovered the values of x^* and of Π^I earlier in the Appendices, and given that $\gamma = x^\beta$, it follows that:

$$n^* = \left(\frac{1-\alpha}{\alpha}\right)\left(L(\alpha^2 A(1+\lambda))^{\frac{1}{1-\alpha}}\right)^{\beta+1}\left(\frac{\theta^{\frac{1-\alpha\beta-2\alpha}{(\beta+1)(1-\alpha)}}}{1+(\alpha^2(A+\lambda))^{\frac{1}{1-\alpha}}\theta^{\frac{-\alpha}{1-\alpha}}}\right)^{\beta+1} \tag{A40}$$

A.8 Proof of Proposition 3

While the optimal value of n^* seems to be rather intricate in terms of its variation with respect to θ, the problem may actually be much simplified by observing that n^* will vary the same way as h, where h writes:

$$h = \left(\frac{\theta^{\frac{1-\alpha\beta-2\alpha}{(\beta+1)(1-\alpha)}}}{1+(\alpha^2(A+\lambda))^{\frac{1}{1-\alpha}}\theta^{\frac{-\alpha}{1-\alpha}}}\right)^{\beta+1} \tag{A41}$$

Taking the derivative of h with respect to θ writes:

$$\frac{\partial h}{\partial \theta} = \left(\frac{1-\alpha\beta-2\alpha}{(\beta+1)(1-\alpha)}\right)\theta^{\frac{1-\alpha\beta-2\alpha}{(1+\beta)(1-\alpha)}-1}(1+y\theta^{\frac{-\alpha}{1-\alpha}}) - \left(\left(\frac{-\alpha}{1-\alpha}\right)y\theta^{\frac{-\alpha}{1-\alpha}-1}\right)\theta^{\frac{1-\alpha\beta-2\alpha}{(\beta+1)(1-\alpha)}} \tag{A42}$$

Which after rearranging simplifies to:

$$\frac{\partial h}{\partial \theta} = \theta^{\frac{1-\alpha\beta-2\alpha}{(1+\beta)(1-\alpha)}}\left(\theta^{-1}\left(\frac{1-\alpha\beta-2\alpha}{(\beta+1)(1-\alpha)}\right)+\frac{y}{\beta+1}\theta^{\frac{-1}{1-\alpha}}\right) \tag{A43}$$

This function depends on the sign of the RHS bracket, which I call B:

$$B = \left(\theta^{-1}\left(\frac{1-\alpha\beta-2\alpha}{(\beta+1)(1-\alpha)}\right)+\frac{y}{\beta+1}\theta^{\frac{-1}{1-\alpha}}\right) \tag{A44}$$

Multiplying by θ, it is straightforward to show that B is of the sign of C:

$$C = \left(\frac{1 - \alpha\beta - 2\alpha}{(\beta + 1)(1 - \alpha)} \right) + \frac{y}{\beta + 1} \theta^{\frac{-\alpha}{1-\alpha}} \tag{A45}$$

Which is clearly a decreasing function of θ which tends to a negative value when θ goes to infinity, as long as:

$$\frac{1 - \alpha\beta - 2\alpha}{(\beta + 1)(1 - \alpha)} \leq 0 \tag{A46}$$

Meaning as long as (i) the intermediary good is important in the production process of the consumption good and (ii) education exhibits large returns to scale. In other words, and as was shown for the cutoff value regarding the profit of the intermediary industry, it is straightforward to show that there exists a cutoff $\tilde{\theta}$ such that n^* increases with θ whenever θ is lower than $\tilde{\theta}$ and decreases with θ whenever θ is larger than this cutoff. Trivially, $\tilde{\theta}$ is the value such that C is set to 0, which writes:

$$\tilde{\theta} = \left(\alpha^2 (A + \lambda) \right)^{\frac{1}{\alpha}} \left(\frac{1 - \alpha}{2\alpha + \alpha\beta - 1} \right)^{\frac{1-\alpha}{\alpha}} \tag{A47}$$

Intuitively, this value is decreasing in β, meaning that, as education of skilled workers exhibits larger returns, the effect of natural resources on innovation becomes all the more negative. Said differently, the larger the returns to education, the lower the cutoff value $\tilde{\theta}$. Interestingly enough indeed, the cutoff value $\tilde{\theta}$ is clearly lower than $\hat{\theta}$ which is straightforward given that, overall, θ exerts three effects on the research effort while it only exerts two on the profit rate in the intermediary industry. Indeed, it generates a positive *market size* effect on the profit rate of the intermediary industry through increased demand for intermediary products whenever natural resources are large enough; in the meantime, it also exerts a substitution effect, whereby higher natural resources raise the wage rate for skilled workers and thus affect the profit rate in the intermediary industry. But a third effect occurs regarding the optimal research effect, which is unambiguously negative: given that larger natural resources reduce the size of the intermediary industry, meaning the number of workers operating in the intermediary industry, it unambiguously lowers human capital formation in the economy which, in turn, reduces the growth of productivity.

A.9 Proof of the extension

Proposition 4 and Corollary 1 discussed the impact of exposure to natural resources on the probability of radical innovation, meaning on the probability that the interme-

diary industry impulses a shift towards never industries. More precisely, these results suggested that whenever θ is larger than a certain cutoff labeled $\breve{\theta}$, radical innovation may never prove profitable. This naturally follows from the comparison of profit reaped by the randomly chosen skilled worker in case she achieves *plain* innovation and in case she achieves *radical* innovation. As recalled by Proposition 4, profit in the former case is greater than in the latter case if and only if:

$$F > \left(\frac{1-\alpha}{\sqrt{2}\alpha}\right)^2 \theta^{\frac{2-2\alpha\beta-4\alpha}{1-\alpha}} \left(\left(\frac{\left(\alpha^2\left(A+\lambda^+\right)\right)^{\frac{1}{1-\alpha}}}{1+\left(\alpha^2(A+\lambda^+)\right)^{\frac{1}{1-\alpha}}} \right)^{2\beta+2} - \left(\frac{\left(\alpha^2\left(A+\lambda\right)\right)^{\frac{1}{1-\alpha}}}{1+\left(\alpha^2(A+\lambda)\right)^{\frac{1}{1-\alpha}}} \right)^{2\beta+2} \right) \tag{A48}$$

While Equation A48 seems rather intricate, it proves convenient to reformulate it in order to derive a sufficient condition, such that profit under *plain* innovation is larger than profit under *radical* innovation. This sufficient condition simply consists in setting the second term of the Right Hand Side of Equation A48 to 0, which now implies that a sufficient condition on F to imply that the randomly chosen skilled worker prefers to achieve plain innovation writes:

$$F > \left(\frac{1-\alpha}{\sqrt{2}\alpha}\right)^2 \theta^{\frac{2-2\alpha\beta-4\alpha}{1-\alpha}} \left(\frac{\left(\alpha^2\left(A+\lambda^+\right)\right)^{\frac{1}{1-\alpha}}}{1+\left(\alpha^2(A+\lambda^+)\right)^{\frac{1}{1-\alpha}}} \right)^{2\beta+2} \tag{A49}$$

Rearranging this formula yields:

$$F > \frac{1}{2}\left(\frac{1-\alpha}{\alpha}\right)^2 \left(\alpha^2(A+\lambda^+)\right)^{\frac{2(\beta+1)}{1-\alpha}} \left(\frac{\theta^{\frac{1-\alpha\beta-2\alpha}{(1+\beta)(1-\alpha)}}}{1+z\theta^{\frac{-\alpha}{1-\alpha}}} \right) \tag{A50}$$

Where z is defined by:

$$z = \left(\alpha^2(A+\lambda^+)\right)^{\frac{1}{1-\alpha}} \tag{A51}$$

It then trivially follows that the way the cutoff value $\tilde{F}$ varies depends on the variation of the function h defined by:

$$h(\theta) = \left(\frac{\theta^{\frac{1-\alpha\beta-2\alpha}{(1+\beta)(1-\alpha)}}}{1+z\theta^{\frac{-\alpha}{1-\alpha}}} \right) \tag{A52}$$

The derivative of which, with respect to θ, is given by:

$$\frac{\partial h}{\partial \theta} = \frac{\theta^{\frac{1-\alpha\beta-2\alpha}{(1-\alpha)(1+\beta)}-1}\left(z\theta^{\frac{-\alpha}{1-\alpha}}(1-\alpha)+1-\alpha\beta-2\alpha\right)}{\left(1+z\theta^{\frac{-\alpha}{1-\alpha}}\right)^2} \tag{A53}$$

Where the denominator is clearly a positive function while the sign of the numerator is rather ambiguous. Actually, it is straightforward to show that the numerator first increases with θ and then decreases with θ. More specifically, θ starts to exert a negative effect on h and then on $\tilde{F}$ as long as θ is greater or equal than the value which neutralizes the numerator of h, which writes:

$$\breve{\theta} = \left(\frac{(1-\alpha)\left(\alpha^2(A+\lambda^+)\right)^{\frac{1}{1-\alpha}}}{\alpha\beta+2\alpha-1}\right)^{\frac{1-\alpha}{\alpha}} \tag{A54}$$

Which is precisely what Corollary 1 demonstrated.

B Coal measure appendix

Section 4 introduced two important measures of coal exposure which are then leveraged in the remainder of the paper to investigate the relation between coal mining and innovation in the long run. For the sake of concision, only two main measures were discussed in this section. Yet, it could be that alternative measures of coal exposure would yield different estimates. In order to alleviate these concerns, I start by introducing alternative measures of coal exposure before plotting the exposure of each French department to coal when using these alternative measures.

B.1 Alternative measures

The gist of the two main measures of coal exposure discussed in Section 4 is to compute the average exposure of French cities in any given department d to the closest coalfield. More specifically, the two main measures were computed using either the map by Châtel (1931) or the data compiled by the ERIH project in order to proxy for $\mathcal{C}$, namely the set of coalfields considered in the paper. An alternative estimation could consist in taking, not only the largest coalfields, but all coalfields that were open in France in the beginning of the XXth century. This approach is, nevertheless, less cautious, to the extent that some coalfields opened for a short period, given that they soon proved to be not profitable. I however make this robustness exercise, by computing $\tilde{p}^d$ when $\mathcal{C}$ is taken to be the whole set of coalfields open as of the beginning of the XXth century.

I then complement this new measure by three alternative measures. The first alternative measure consists in computing the (log) surface of coalfields at the department level in the beginning of the XXth century, as recorded in the French *cadastre minier*. Technically, I have no choice but to compute this new measure $\tilde{e}^s$ as:

$$\tilde{e}_d^1 = \ln(s_d + 1) \tag{B1}$$

Where s_d indexes the surface of coalfields in department d in the beginning of the XXth century. Indeed, this measure is null for some departments, implying that the log transform is necessary in order to ensure that no department is dropped out of the sample. In a similar way, I define e_d^2 as the log transform of the average coal output in department d in the onset of the XXth century:

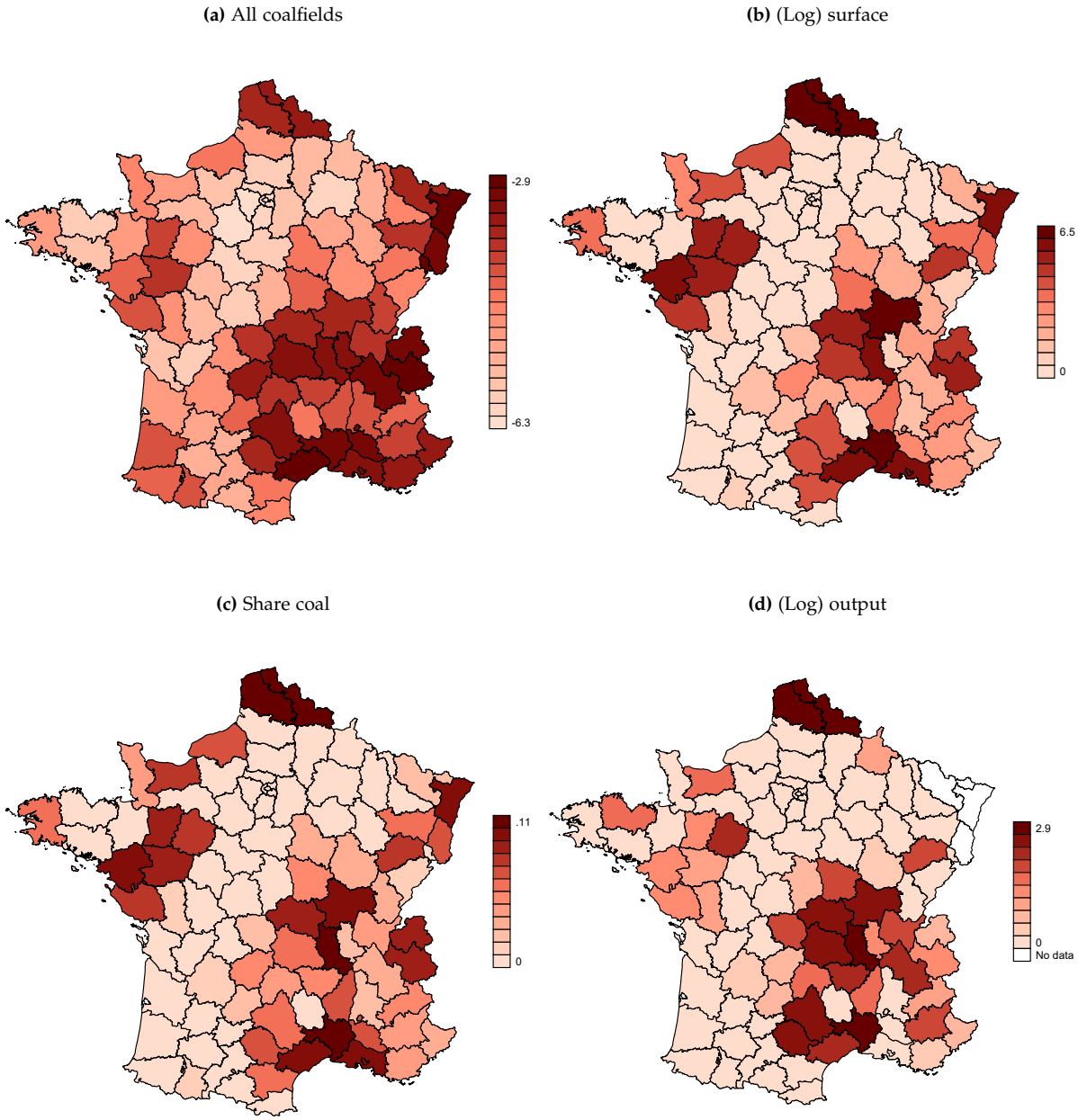
$$\tilde{e}_d^2 = \ln(o_d + 1) \tag{B2}$$

Where o_d is the average coal output in department d from 1900 to 1907. This output-based metric enables me to check whether coalfields as registered either in the ERIH, or in the Châtel (1931) map, actually deliver large coal output and are not only abandoned coalfields. Finally, I let e_d^3 be a measure of the share of the surface of department d which is covered with a coalfield; it could indeed be that a given department hosts a large number of coalfields, but that these coalfields nevertheless represent a tiny share of its overall surface, in which case the effect on coal on local economic outcomes might be mitigated.

B.2 Coal exposure under these alternatives

Figure B1 plots coal exposure when it is proxy with the alternative measures introduced in Section B.1. Importantly enough, the three last alternative measures, namely the surface and the output based measures are semi-discrete measures, in the sense that the value of coal exposure is zero for around half of the sample, namely for all departments which never mined any coal and which, by construction, had neither coalfields nor coal output. Also, it should be noted that data is missing for three departments, namely Moselle, Haut Rhin and Bas Rhin, in the fourth context, namely when coal exposure is averaged by the coal output in the beginning of the XXth century; this is to be explained by these three departments were, at that time, attached to the German Empire, implying that I lack data on coal output for these departments prior to World War 1.

FIGURE B1: Coal exposure under alternative measures



Notes: Figure B1a plots the exposure to coal when $\mathcal{C}$ is taken to encompass all coalfields which were open in France in the beginning of the XXth century, irrespectively of whether they were large coalfields or not. Figure B1b captures exposure to coal using the (log) surface of coalfields; Figure B1c depicts exposure to coal depending on the share of territory covered with a coalfield while Figure B1d shows the exposure to coalfields depending on average output of coal in the beginning of the XXth century.

C Instrumental variable

C.1 Graphical evidence

Throughout the paper, I mostly considered two measures of coal exposure which were discussed in Section 4; as a result, I only showed the robustness of the instrument for

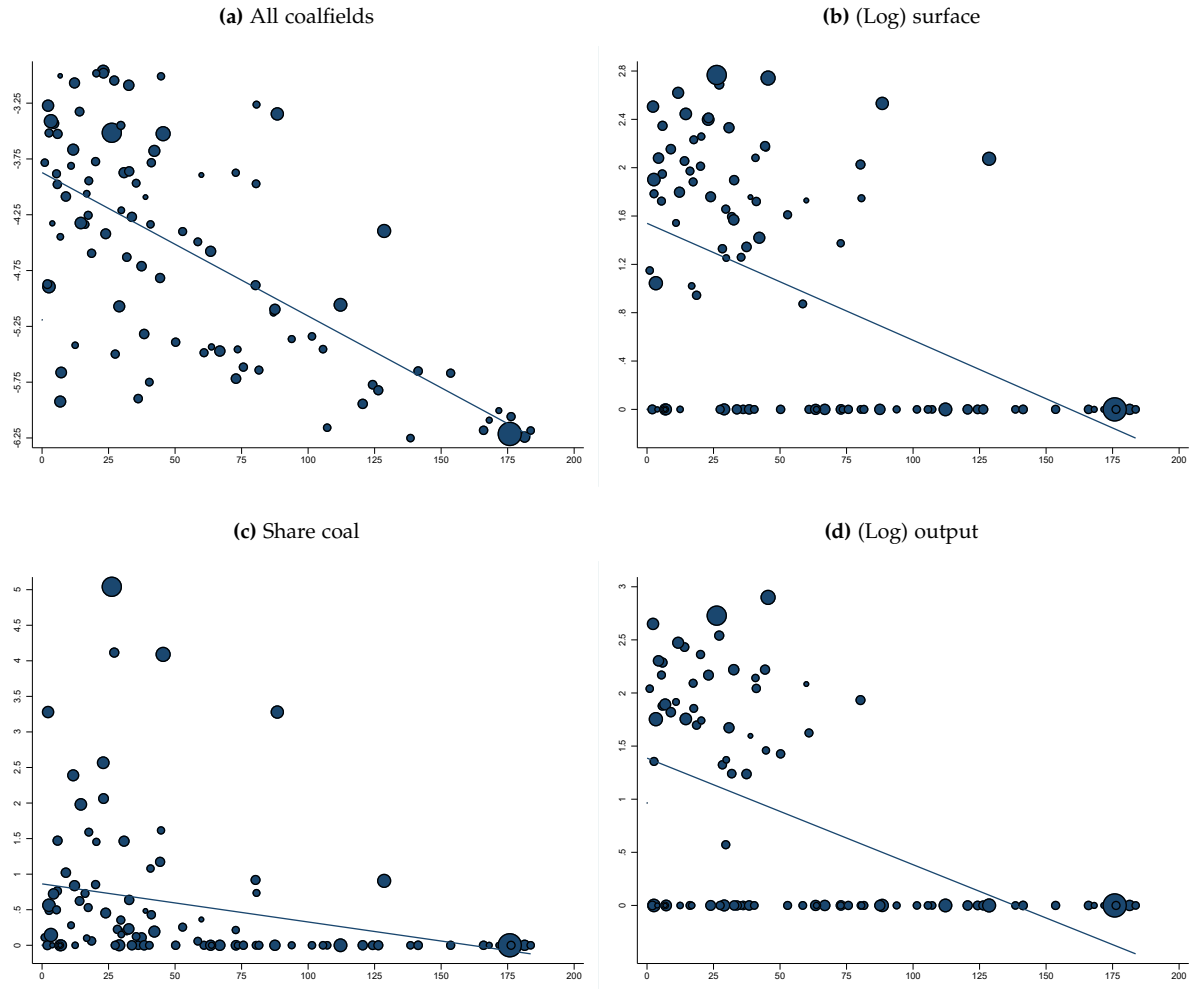
the main measure of coal exposure, namely the exposure to large coalfields, without further asserting the validity of the instrument for alternative measure of coal exposures. I nevertheless show in this subsection that the instrument proves valid when measuring coal exposure with the alternative measures introduced in Section B. That the instrument proves valid for all kinds of measure of coal exposure used in this paper is further asserted by Figure C1 which brings further credence to the intuition of a negative relation between exposure to coal and distance to the closest carboniferous rock strata. All alternative measures of coal exposure indeed correlate negatively and significantly with the instrument; in particular, exposure to all coalfields exhibits a peculiarly large and negative coefficient, while other correlations are slightly noisiers, due to the fact that the three last measures are semi-discrete ones, for which around half of the sample automatically has the value 0 in terms of exposure to coal.

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.

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.

FIGURE C1: Coal exposure under alternative measures



Notes: Figure B1a plots exposure to coal when $\mathcal{C}$ is taken to encompass all coalfields which were open in France in the beginning of the XXth century, irrespectively of whether they were large coalfields or not, against distance to carboniferous rock strata. Figure B1b plots the same relation but when exposure to coal is defined as the (log) surface of coalfields; Figure B1c plots the same relation but when exposure to coal is proxy by the share of territory covered with a coalfield while Figure B1d shows the same relation when exposure to coalfields depends on average output of coal in the beginning of the XXth century.

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. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

TABLE C2: Coal output 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

In this subsection, I present the main results when allowing for either alternative measures of coal exposure or alternative functional forms. Main results were indeed derived using mainly the proximity to large coalfields as explanatory variable for exposure to coal. Yet, I documented above that other measures could proxy for the exposure to coalfields. I then test whether OLS estimates and their IV counterparts are changed whenever I am using alternative measures of coal exposure, or whenever I am trying alternative functional forms. I start by focusing on alternative measures of coal exposure, and then turn to flexible functional forms.

D.1 Main results under alternative coal exposure measures

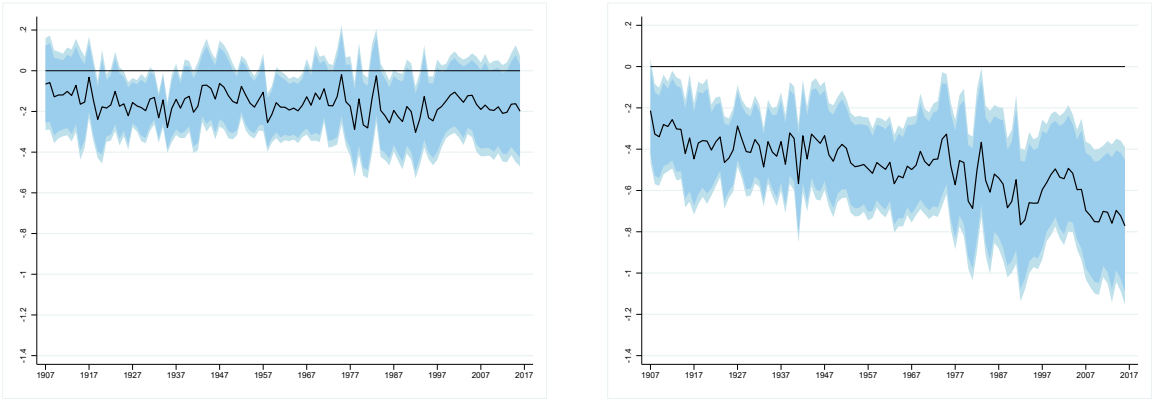
This subsection is itself divided in different components corresponding to the main estimates derived in this paper. More specifically, I start by looking at the impact of alternative coal exposure measures on total patenting activity (Section D.1.1)before investigating the impact of these alternative coal exposure measures on patent quality (Section D.1.2).

D.1.1 Patenting quantity

All in all, the evidence summarized below brings strong support in favor of the argument established in Section 4.3 which suggested the idea that coal exposure induced a negative, significant and increasing effect on innovation at the department level in France throughout the XXth century. Indeed, the coefficients on the coal exposure measure are systematically negative and significant when using the IV specification, suggesting that the causal effect was large and detrimental to innovation in the long run. Interestingly enough, for all specifications, the coefficient on the IV specification are larger than under the OLS model, and by a large amount. That these estimates are significantly larger than their IV estimates are consistent with the fact that three of the four alternative specifications introduced here are specifications where the coal exposure measure is a semi-discrete one, which leads to a higher bias of OLS estimates. It is also noteworthy that the estimates are broadly consistent in terms of magnitude with those recovered in the main section of results.

FIGURE D1: Coal exposure measure and patenting activity, Châtel (1931) measure

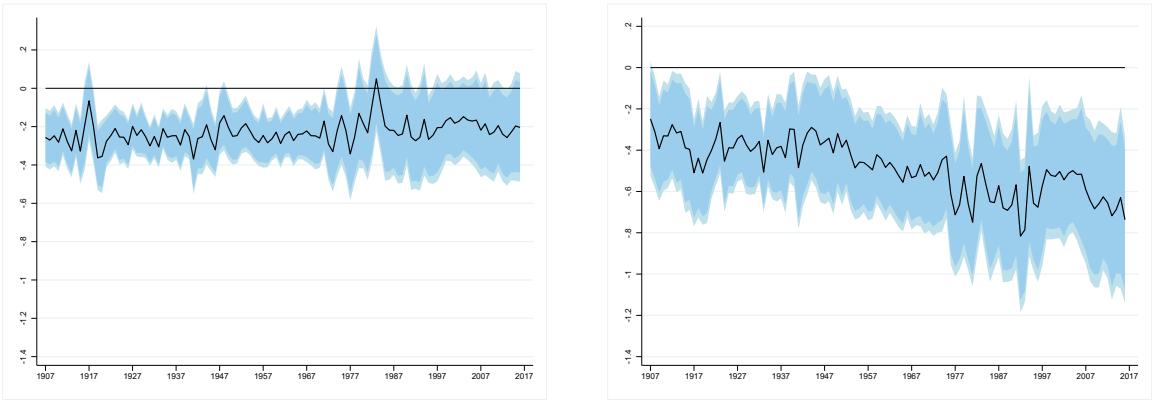
(a) OLS **(b)** IV



Notes: Figure D1a the effect of exposure to coal as measured by the Châtel (1931) map on total innovation activity using the OLS specification while Figure D1b plots the same effect when instrumenting the coal exposure measure with distance to carboniferous rock strata.

FIGURE D2: Coal exposure measure and patenting activity, (ln) surface measure

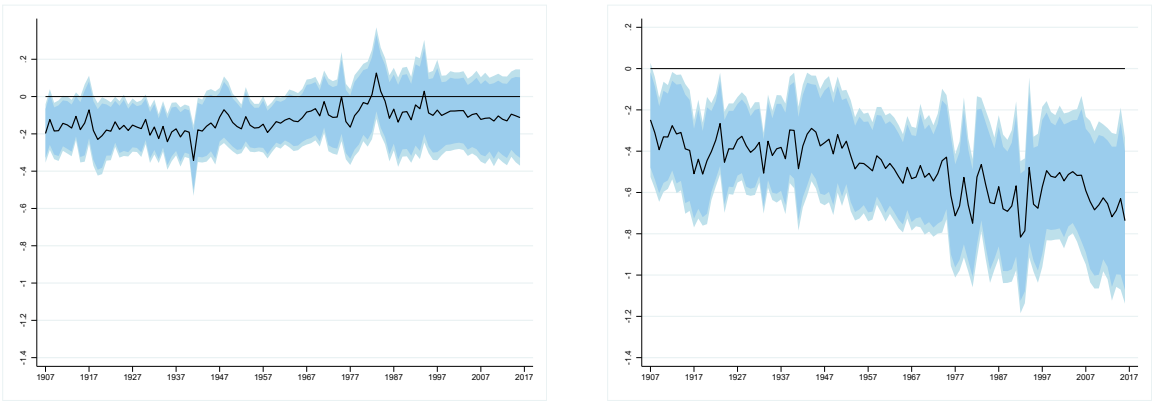
(a) OLS **(b)** IV



Notes: Figure D2a the effect of exposure to coal as measured by the (ln) surface of coalfields on total innovation activity using the OLS specification while Figure D2b plots the same effect when instrumenting the coal exposure measure with distance to carboniferous rock strata.

FIGURE D3: Coal exposure measure and patenting activity, share surface measure

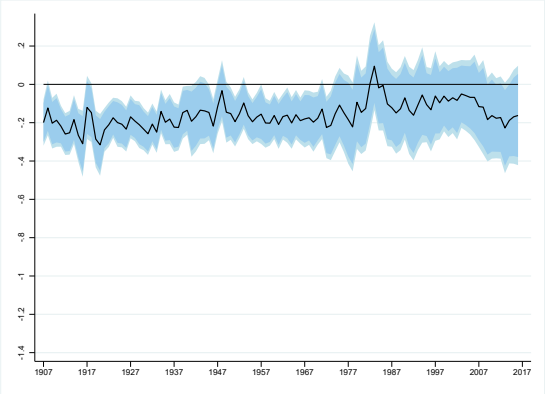
(a) OLS **(b)** IV



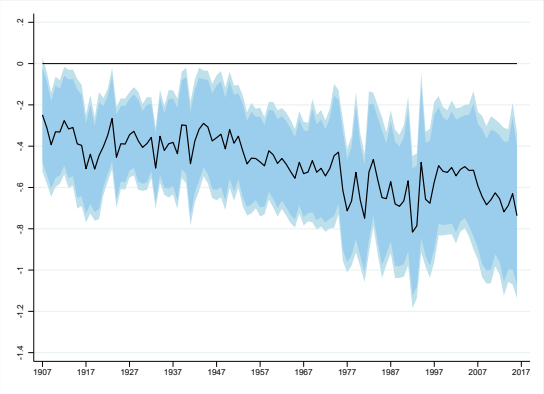
Notes: Figure D3a the effect of exposure to coal as measured by the share of the department surface covered with coalfields on total innovation activity using the OLS specification while Figure D3b plots the same effect when instrumenting the coal exposure measure with distance to carboniferous rock strata.

FIGURE D4: Coal exposure measure and patenting activity, output-based measure

(a) OLS



(b) IV

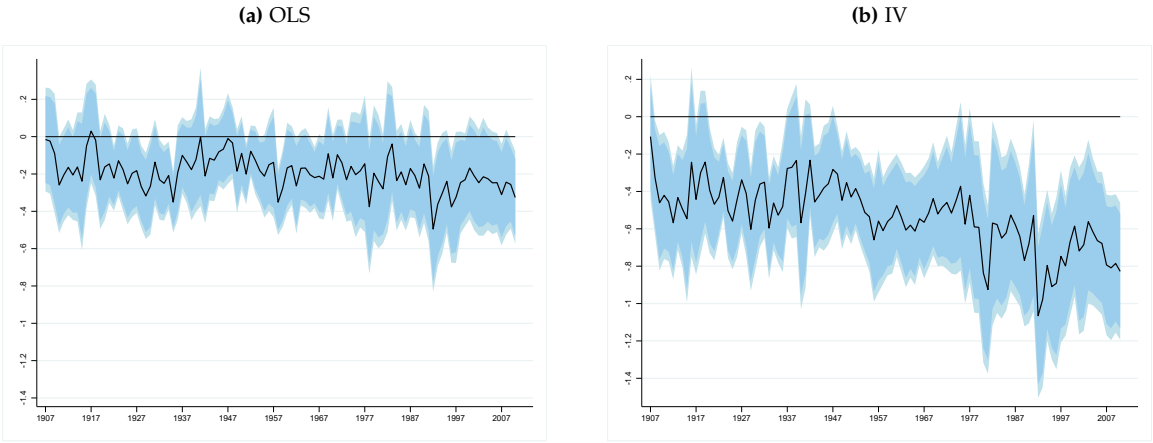


Notes: Figure D4a the effect of exposure to coal as measured by the average output of coal in the onset of the XXth century on total innovation activity using the OLS specification while Figure D4b plots the same effect when instrumenting the coal exposure measure with distance to carboniferous rock strata.

D.1.2 Patenting quality

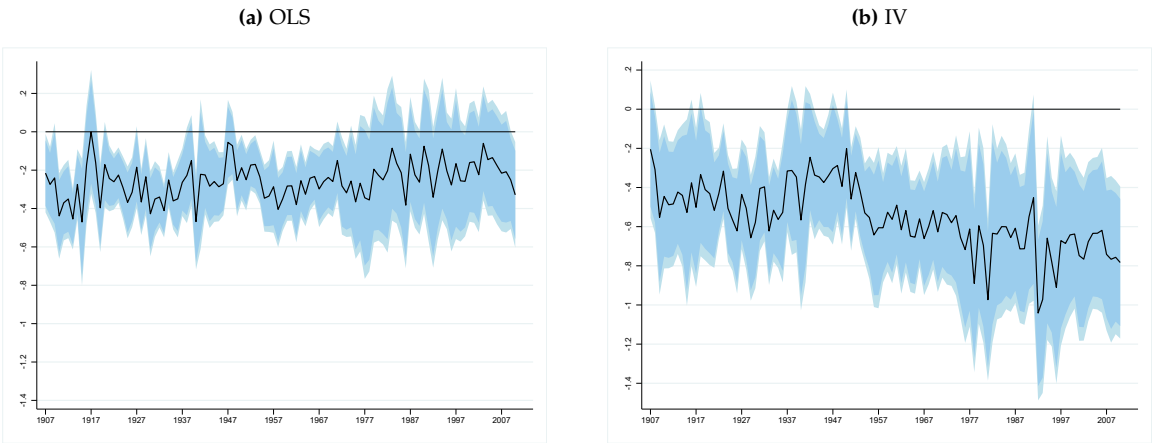
Section 4.3 argued that coal exposure did not only affect the amount of patenting activity, it also affected its quality, in the sense that it reduced breakthrough patenting activity, as measured using the data made available by Kelly et al. (2021). I provide complementary evidence of the negative and sustained impact of coal exposure on patent quality by studying the impact of the alternative coal exposure measure on patent quality; more specifically, Figures D5, D6, D7 and D8 respectively show the impact of a one standard deviation increase in coal exposure on innovation, when coal exposure is respectively measured as the exposure to Châtel (1931) coalfields, the (ln) surface of coalfields, the share of the territory covered with a coalfield or the (ln) average coal output in the onset of the XXth century.

FIGURE D5: Coal exposure measure and patenting quality, Châtel (1931) measure



Notes: Figure D5a the effect of exposure to coal as measured by the Châtel (1931) map on total innovation quality using the OLS specification while Figure D5b plots the same effect when instrumenting the coal exposure measure with distance to carboniferous rock strata.

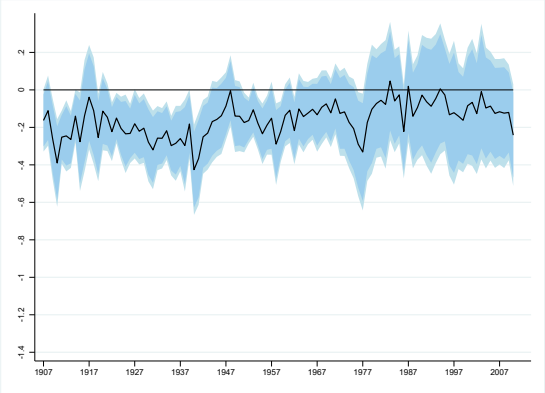
FIGURE D6: Coal exposure measure and patenting quality, (ln) surface measure



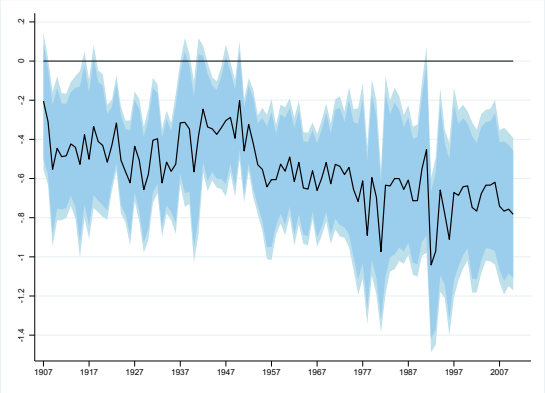
Notes: Figure D6a the effect of exposure to coal as measured by the (ln) surface of coalfields on total innovation quality using the OLS specification while Figure D6b plots the same effect when instrumenting the coal exposure measure with distance to carboniferous rock strata.

FIGURE D7: Coal exposure measure and patenting quality, share surface measure

(a) OLS



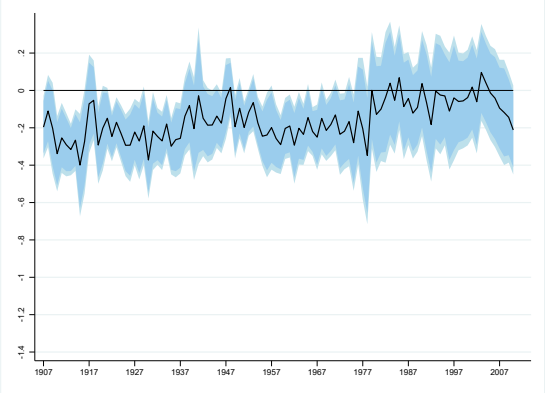
(b) IV



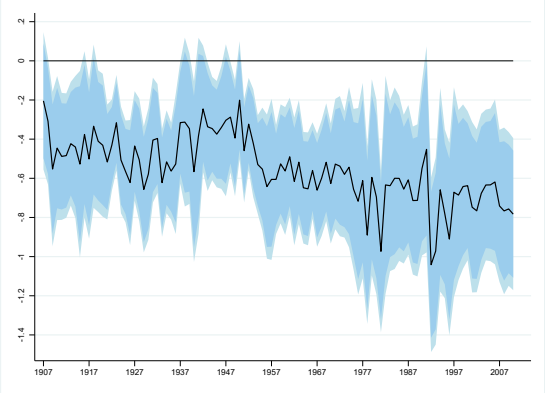
Notes: Figure D7a the effect of exposure to coal as measured by the share of the department surface covered with coalfields on total innovation quality using the OLS specification while Figure D7b plots the same effect when instrumenting the coal exposure measure with distance to carboniferous rock strata.

FIGURE D8: Coal exposure measure and patenting quality, output-based measure

(a) OLS



(b) IV



Notes: Figure D6a the effect of exposure to coal as measured by the average output of coal in the onset of the XXth century on total innovation quality using the OLS specification while Figure D6b plots the same effect when instrumenting the coal exposure measure with distance to carboniferous rock strata.

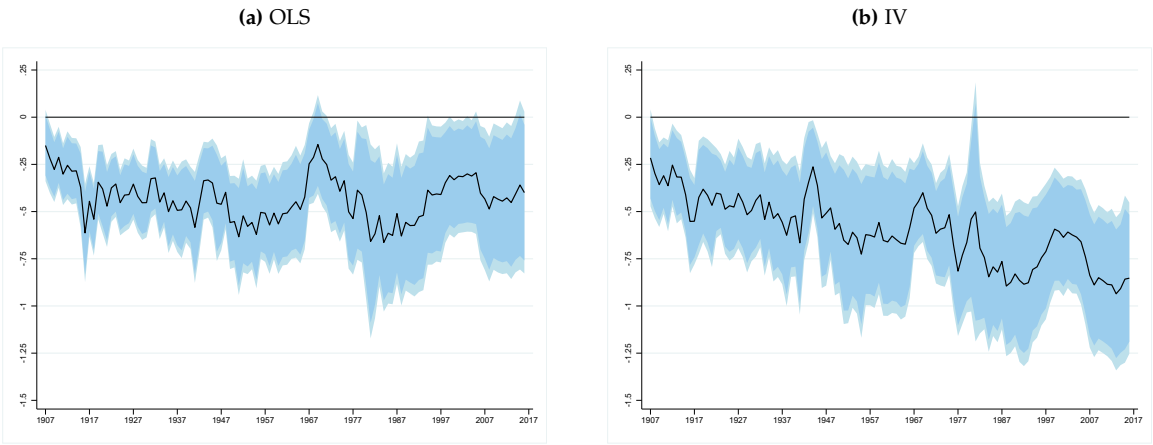
D.2 Flexible functional forms

It proves important as well to investigate whether results hold under alternative flexible functional forms. Indeed, the paper suggested a linear relation between coal exposure and patenting activity; it could however be that this flexible relation has some flaws, which pseudo poisson or hyperbolic sinusoidal forms may correct for.

D.2.1 Pseudo Poisson estimates

Figures D9, D10 and D11 show the impact of exposure to coal using the three main coal exposure measures when using a Pseudo Poisson functional form. More specifically, I plot for each measure the results obtained both under the Pseudo Poisson specification and using its instrumental counterpart. Again, results suggest a negative, significant and sustained impact of exposure to coal extraction on patenting activity.

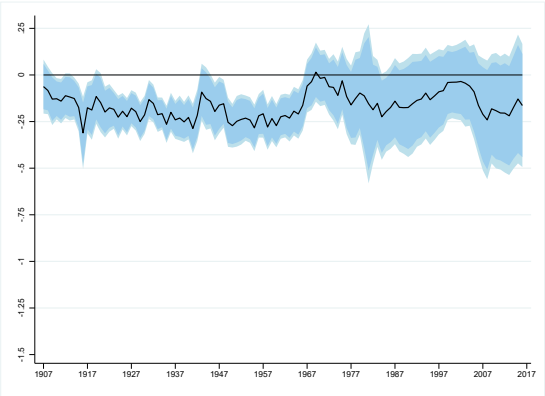
FIGURE D9: Coal exposure measure and patenting quantity, ERIH measure, Pseudo Poisson specification



Notes: Figure D9a shows the effect of exposure to coal as measured by the ERIH project on total innovation quality using the Poisson specification while Figure D9b plots the same effect when instrumenting the coal exposure measure with distance to carboniferous rock strata.

FIGURE D10: Coal exposure measure and patenting quantity, Châtel (1931) measure, Pseudo Poisson specification

(a) OLS



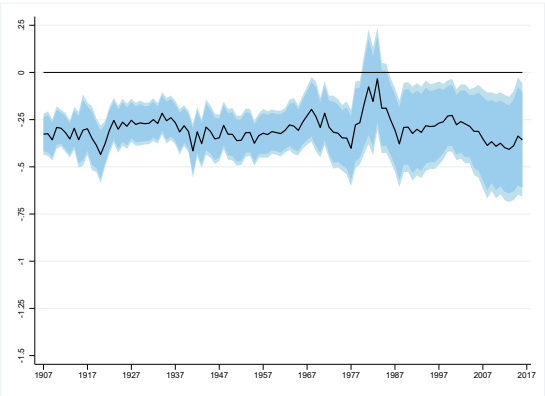
(b) IV



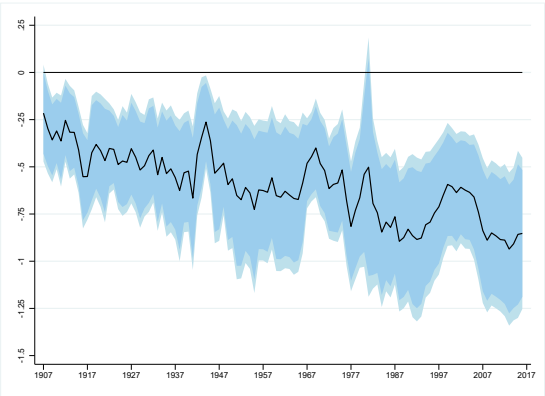
Notes: Figure D10a the effect of exposure to coal as measured by Châtel (1931) measure on total innovation quality using the Pseudo Poisson specification while Figure D10b plots the same effect when instrumenting the coal exposure measure with distance to carboniferous rock strata.

FIGURE D11: Coal exposure measure and patenting quantity, (ln) surface measure, Pseudo Poisson specification

(a) OLS



(b) IV

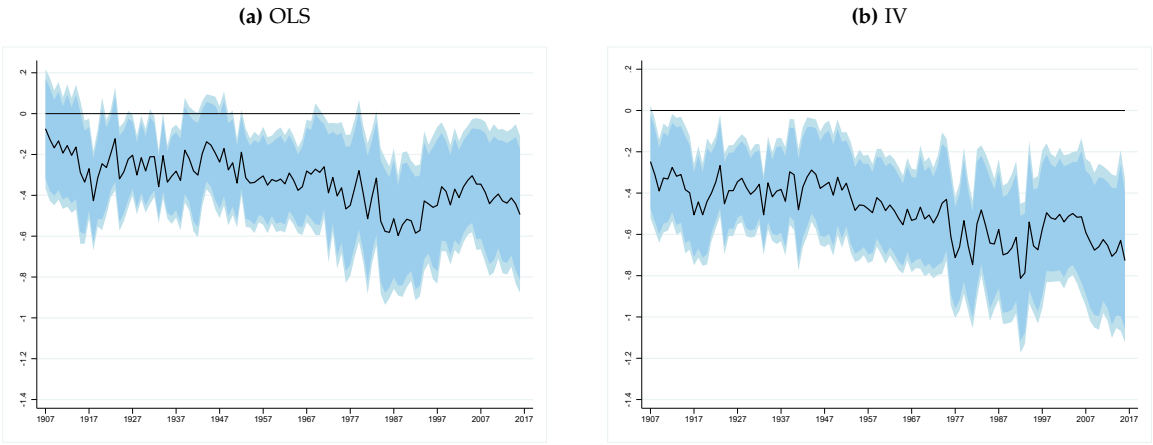


Notes: Figure D11a shows the effect of exposure to coal as measured by the (log) surface covered with coalfields on total innovation quality using the Pseudo POisson specification while Figure D11b plots the same effect when instrumenting the coal exposure measure with distance to carboniferous rock strata.

D.2.2 Inverse Hyperbolic Sine

Figures D12, D13 and D14 show the impact of exposure to coal using the three main coal exposure measures when using an Inverse Hyperbolic Sine functional form. More specifically, I plot for each measure the results obtained both under the Pseudo Poisson specification and using its instrumental counterpart. Again, results suggest a negative, significant and sustained impact of exposure to coal extraction on patenting activity. More importantly, the size of the coefficients is commensurate with that obtained under the general functional form, suggesting, again, that the estimates are very stable, not only across measures of coal but also across specifications, bringining further credence to the argument that coal exposure induced a negative and sustained impact over patenting activity in the long run.

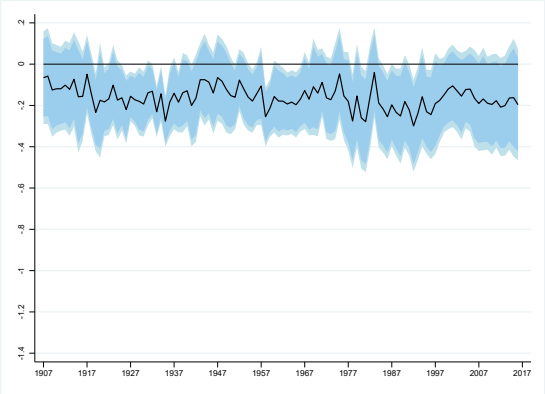
FIGURE D12: Coal exposure measure and patenting quantity, ERIH measure, Inverse Hyperbolic Sine specification



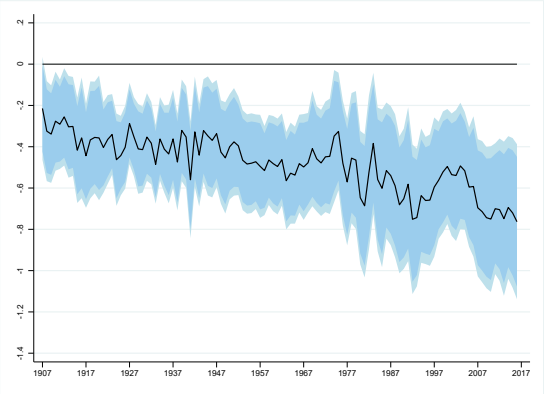
Notes: Figure D12a shows the effect of exposure to coal as measured by the ERIH project on total innovation quality using the inverse hyperbolic sine specification while Figure D12b plots the same effect when instrumenting the coal exposure measure with distance to carboniferous rock strata.

FIGURE D13: Coal exposure measure and patenting quantity, Châtel (1931) measure, Inverse Hyperbolic Sine specification

(a) OLS



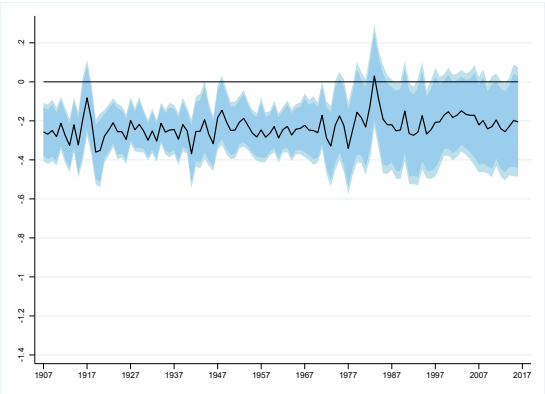
(b) IV



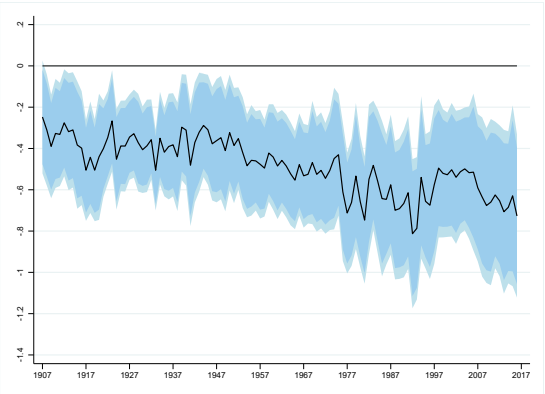
Notes: Figure D13a the effect of exposure to coal as measured by Châtel (1931) measure on total innovation quality using the Inverse Hyperbolic Sine specification while Figure D13b plots the same effect when instrumenting the coal exposure measure with distance to carboniferous rock strata.

FIGURE D14: Coal exposure measure and patenting quantity, (ln) surface measure, Inverse Hyperbolic Sine specification

(a) OLS



(b) IV



Notes: Figure D14a shows the effect of exposure to coal as measured by the (log) surface covered with coalfields on total innovation quality using the Inverse Hyperbolic Sine specification while Figure D14b plots the same effect when instrumenting the coal exposure measure with distance to carboniferous rock strata.

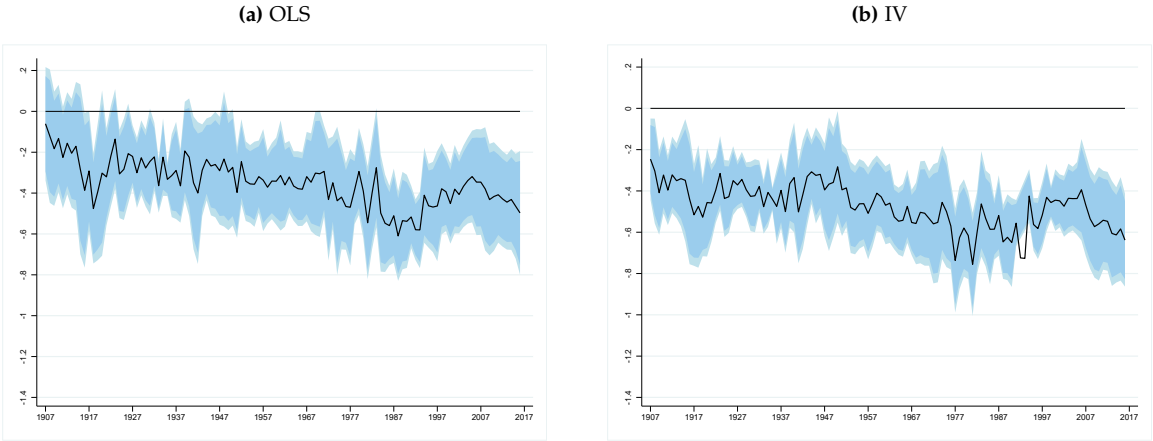
D.3 Alternative clustering methods

So far, all cross-sectional regressions were clustered using robust standard errors. Yet, it could be that errors terms are spatially correlated across departments, given that the main explanatory variable has a geographical structure which may imply some spillovers. In order to alleviate concerns that results may be biased because of such a correlation structure, I use the *acreg* command in Stata from Colella et al. (2020) which allows for spatial autocorrelation using flexible buffers; I let the main distance cutoff to be 175 kilometers (Section D.3.1), but check whether the estimates are unaltered when using alternative values such as 150 kilometers (Section D.3.2) and 200 kilometers (Section D.3.3). Under each assumption, I show results for the three main coal exposure measures.

D.3.1 Conley standard errors - cutoff of 175 kilometers

Figures D15, D16 and D17 show the effect of coal exposure on innovation, when coal exposure is measured respectively as the proximity to ERIH coalfields, to Châtel (1931) coalfields or as the (ln) surface of coalfields, while standard errors are clustered using Conley adjusted methods with a cutoff value of 175 kms. Clearly, results are robust to this alternative clustering method, and are still significant either at the 5% or even at the 1% level depending on the years considered.

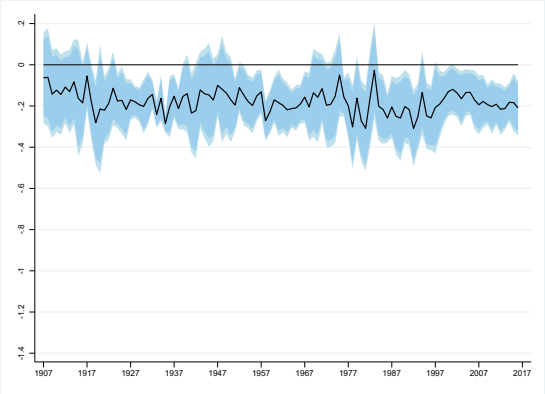
FIGURE D15: Coal exposure measure and patenting quantity, Conley s.e., 175 km buffer, main measure



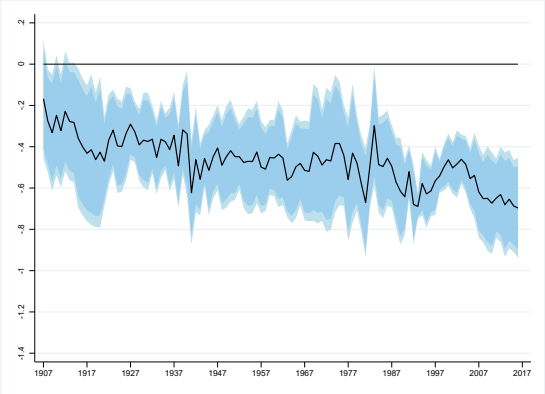
Notes: Figure D15a shows the effect of exposure to coal as measured by the ERIH project on total innovation quality using Conley adjusted standard errros with buffer 175 km while Figure D15b plots the same effect when instrumenting the coal exposure measure with distance to carboniferous rock strata.

FIGURE D16: Coal exposure measure and patenting quantity, Conley s.e., 175 km buffer, Châtel (1931) measure

(a) OLS



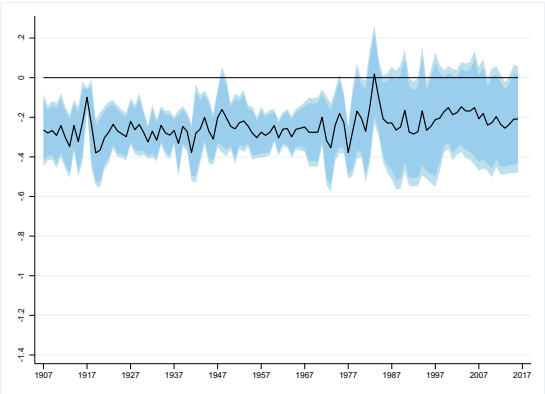
(b) IV



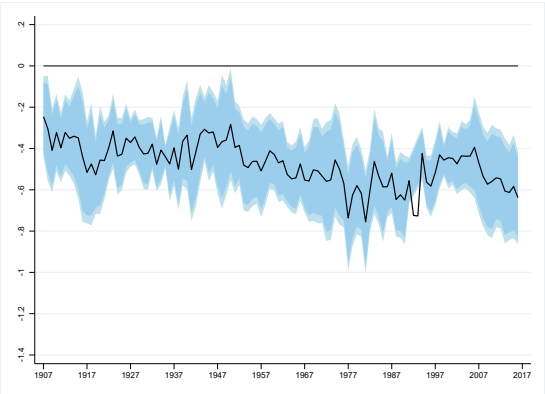
Notes: Figure D16a shows the effect of exposure to coal as measured by the ERIH project on total innovation quality using Conley adjusted standard erros with buffer 175 km while Figure D16b plots the same effect when instrumenting the coal exposure measure with distance to carboniferous rock strata.

FIGURE D17: Coal exposure measure and patenting quantity, Conley s.e., 175 km buffer, (ln) surface measure

(a) OLS



(b) IV



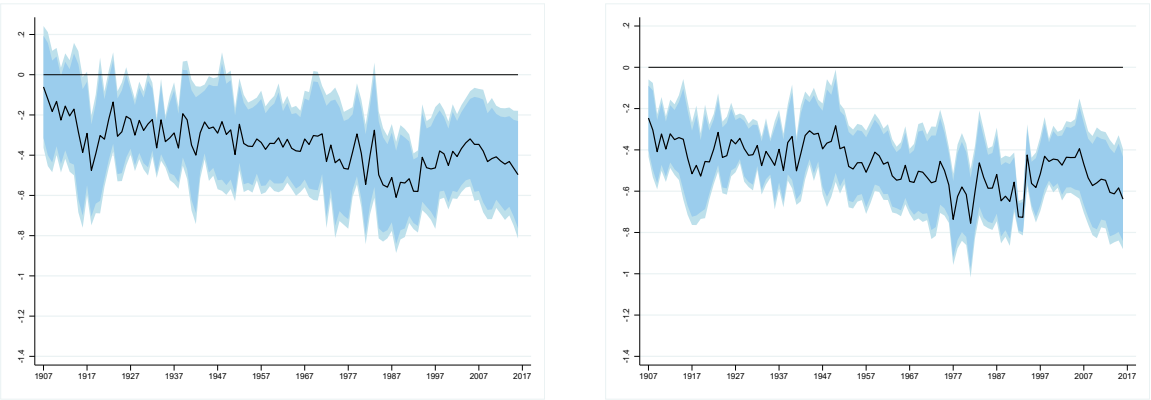
Notes: Figure D17a shows the effect of exposure to coal as measured by the ERIH project on total innovation quality using Conley adjusted standard erros with buffer 175 km while Figure D17b plots the same effect when instrumenting the coal exposure measure with distance to carboniferous rock strata.

D.3.2 Conley standard errors - cutoff of 150 kilometers

Figures D18, D19 and D20 show the effect of coal exposure on innovation, when coal exposure is measured respectively as the proximity to ERIH coalfields, to Châtel (1931) coalfields or as the (ln) surface of coalfields, while standard errors are clustered using Conley adjusted methods with a cutoff value of 150 kms. Clearly, results are robust to this alternative clustering method, and are still significant either at the 5% or even at the 1% level depending on the years considered.

FIGURE D18: Coal exposure measure and patenting quantity, Conley s.e., 150 km buffer, main measure

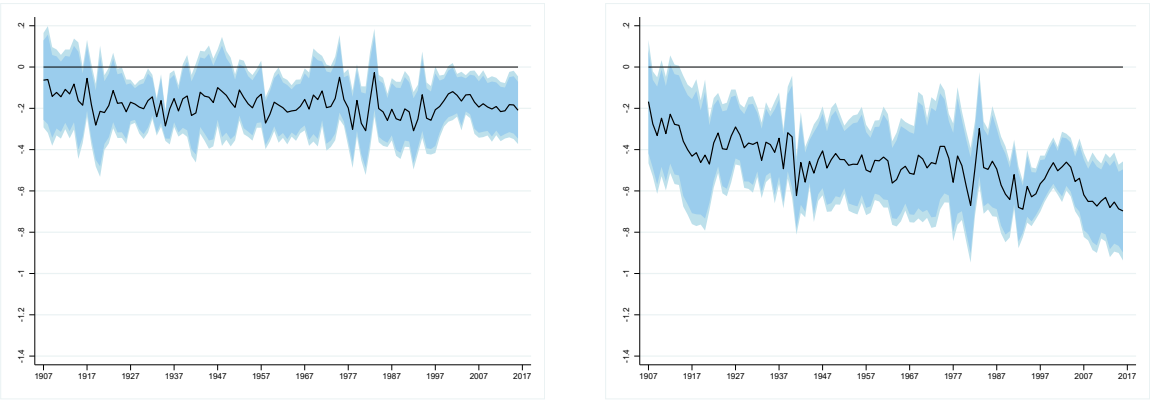
(a) OLS **(b)** IV



Notes: Figure D18a shows the effect of exposure to coal as measured by the ERIH project on total innovation quality using Conley adjusted standard errros with buffer 150 km while Figure D18b plots the same effect when instrumenting the coal exposure measure with distance to carboniferous rock strata.

FIGURE D19: Coal exposure measure and patenting quantity, Conley s.e., 150 km buffer, Châtel (1931) measure

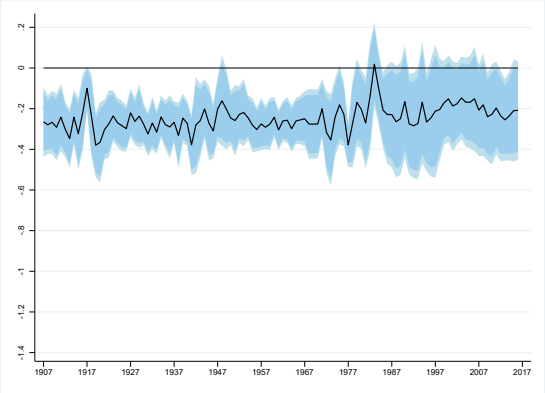
(a) OLS **(b)** IV



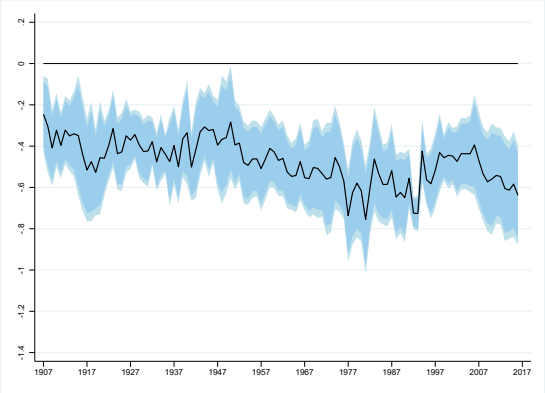
Notes: Figure D19a shows the effect of exposure to coal as measured by the ERIH project on total innovation quality using Conley adjusted standard errors with buffer 150 km while Figure D19b plots the same effect when instrumenting the coal exposure measure with distance to carboniferous rock strata.

FIGURE D20: Coal exposure measure and patenting quantity, Conley s.e., 150 km buffer, (ln) surface measure

(a) OLS



(b) IV

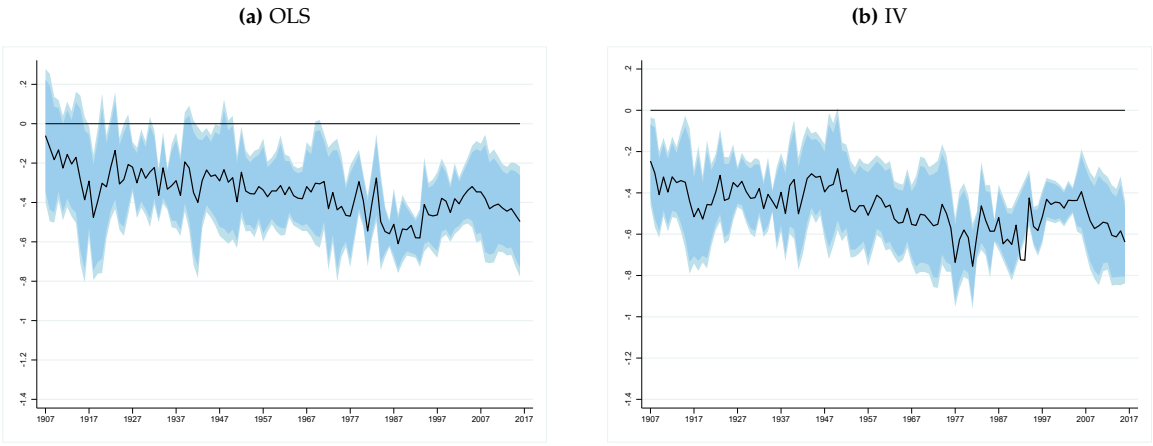


Notes: Figure D20a shows the effect of exposure to coal as measured by the ERIH project on total innovation quality using Conley adjusted standard erros with buffer 150 km while Figure D20b plots the same effect when instrumenting the coal exposure measure with distance to carboniferous rock strata.

D.3.3 Conley standard errors - cutoff of 200 kilometers

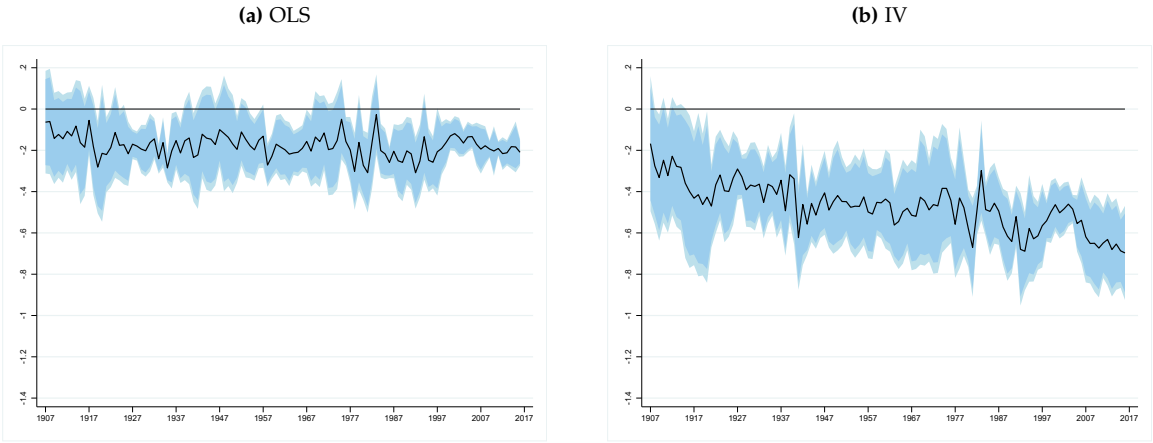
Figures D21, D22 and D23 show the effect of coal exposure on innovation, when coal exposure is measured respectively as the proximity to ERIH coalfields, to Châtel (1931) coalfields or as the (ln) surface of coalfields, while standard errors are clustered using Conley adjusted methods with a cutoff value of 200 kms. Clearly, results are robust to this alternative clustering method, and are still significant either at the 5% or even at the 1% level depending on the years considered.

FIGURE D21: Coal exposure measure and patenting quantity, Conley s.e., 200 km buffer, main measure



Notes: Figure D21a shows the effect of exposure to coal as measured by the ERIH project on total innovation quality using Conley adjusted standard errros with buffer 200 km while Figure D21b plots the same effect when instrumenting the coal exposure measure with distance to carboniferous rock strata.

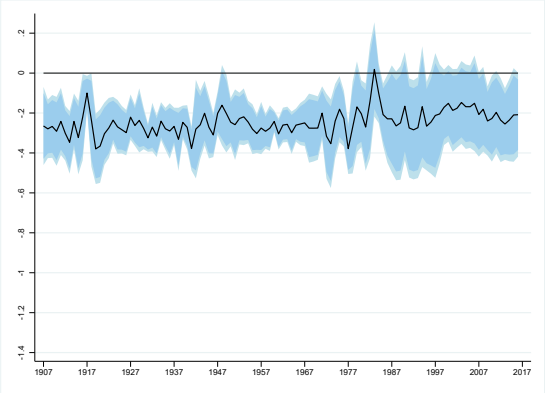
FIGURE D22: Coal exposure measure and patenting quantity, Conley s.e., 200 km buffer, Châtel (1931) measure



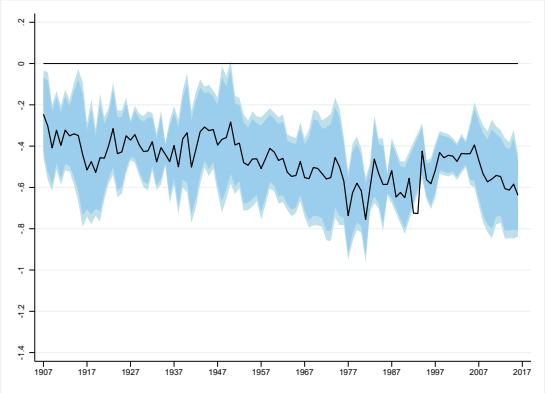
Notes: Figure D22a shows the effect of exposure to coal as measured by the ERIH project on total innovation quality using Conley adjusted standard errors with buffer 200 km while Figure D22b plots the same effect when instrumenting the coal exposure measure with distance to carboniferous rock strata.

FIGURE D23: Coal exposure measure and patenting quantity, Conley s.e., 200 km buffer, (ln) surface measure

(a) OLS



(b) IV



Notes: Figure D23a shows the effect of exposure to coal as measured by the ERIH project on total innovation quality using Conley adjusted standard erros with buffer 200 km while Figure D23b plots the same effect when instrumenting the coal exposure measure with distance to carboniferous rock strata.