INSTITUTE
OF ECONOMICS

Scuola Superiore
Sant'Anna

LEM | Laboratory of Economics and Management

Institute of Economics
Scuola Superiore Sant'Anna

Piazza Martiri della Libertà, 33 - 56127 Pisa, Italy
ph. +39 050 88.33.43
institute.economics@sssup.it

LEM
WORKING PAPER SERIES

Innovation, localized externalities, and the British Industrial Revolution, 1700–1850

Ugo M. Gragnolati [a,b,c]
Alessandro Nuvolari [d]

[a] Dipartimento di Scienze economiche ed aziendali, Università di Cagliari, Italy.
[b] Centre d'Économie de la Sorbonne, Université Paris 1 Panthéon-Sorbonne, France.
[c] CRENoS - Centro Ricerche Economiche Nord Sud, Italy.
[d] Institute of Economics, Scuola Superiore Sant'Anna, Pisa, Italy.

2023/26 **June 2023**
ISSN(ONLINE) 2284-0400

Innovation, localized externalities, and the British Industrial Revolution, 1700–1850

Ugo M. Gragnolati

Università di Cagliari, Dipartimento di Scienze economiche ed aziendali, Italy.
Université Paris 1 Panthéon-Sorbonne, Centre d'Économie de la Sorbonne, France.
CRENoS - Centro Ricerche Economiche Nord Sud, Italy.

Alessandro Nuvolari

Sant'Anna School of Advanced Studies, Institute of Economics, Italy.

Abstract

We study the determinants of the spatial distribution of patent inventors at the county level for Great Britain between 1700–1850. Our empirical analysis rests on the localization model by Bottazzi et al. (2007) and on the related estimation procedure by Bottazzi and Gragnolati (2015). Such an approach helps in particular to discriminate the role of localized externalities against other descriptors of county attractiveness. Our results show that, while the underlying geography of production remained a strong determinant of inventor location all throughout the industrial revolution, the effect of localized externalities among patent inventors went from being nearly absent in the early phases of industrialization to becoming a major driver of inventor location. In particular, local interactions among the "mass" of generic inventors turn out to be at least as important as interactions with "elite" inventors.

JEL codes: C31, N73, N93, O33, R12.

Keywords: Inventor location · Patents · Localized externalities · Industrial Revolution.

1 Introduction

Technological innovation is widely regarded as a major source of economic growth. Yet, the inventive activities that are responsible for innovation tend to be unevenly distributed in space, which possibly plays a role in the economic differentiation of cities, regions, and countries. In this sense, the determinants of inventor location may contribute to explain geographic heterogeneity in economic outcomes, while also shedding light on some of the mechanisms underlying technological innovation. Furthermore, the possibility of path-dependent dynamics imply that the present spatial configuration of economic activities may ultimately depend on early episodes of localization and development. For these reasons, the geography of inventive activities during key historical periods of early cluster formation may contain relevant information about the underlying enabling factors. In this respect, the British industrial revolution represents a key historical discontinuity characterized by an unprecedented rate of innovation, which ultimately allowed to escape Malthusian stagnation and to establish the modern growth regime (Kuznets, 1966). What then were the drivers of location for inventive activities during the British industrial revolution?

As a first approximation, the spatial distribution of inventive activities is commonly expected to reflect the location of production. When transportation costs are non negligible, for instance, the inventors of process innovations may have an incentive to site near productive activities as a way to minimize distance from demand. Incidentally, this was indeed the context that characterized the British industrial revolution. Yet, the spatial link between innovative and productive activities can also have a more profound underpinning. Specifically, as soon as technical change is ascribed to experience and learning by doing, it follows that "the very activity of production [...] gives rise to problems for which favorable responses are selected over time" (Arrow, 1962, p. 156). As a corollary, innovation would occur more frequently where production is larger (see the related discussion in Audretsch and Feldman, 1996). While taking for granted this basic tenet, however, a host of empirical studies point to some discrepancy between the spatial distributions of inventive and productive activities. For example, Lamoreaux and Sokoloff (2000) show that the geographic dispersal of patented inventions in glass manufacturing during the late 19th century presented non negligible deviations from the underlying location of production in the United States. Relatedly, Carlino and Kerr (2015) review a variety of studies that document the excess spatial concentration of innovation relative to other economic activities, so that some further agglomeration forces could act selectively on innovative activities. Even during the British industrial revolution, then, some specific forces could be potentially at play that affected the geography of innovation beyond its basic spatial linkage with production.

In this respect, knowledge spillovers have long been regarded as a prime suspect. Indeed, the manufacturing districts of 19th century England were famously described as an environment in which "if one man starts a new idea it is taken up by others and combined with suggestions of their own; and thus becomes the source of yet more new ideas" (Marshall, 1890, p. 332). If the possibility of such interactions among inventors are somehow bounded in space, the location of inventive activities becomes subject to a local self-reinforcing mechanism: namely, inventions tend to occur more frequently where other inventions have already occurred. This hypothesis has been tested using contemporary patent citations to conclude that indeed inventors tend to be more frequently linked to other spatially proximate inventors, even after controlling for the spatial distribution of production (see the discussion by Henderson et al., 2005, Jaffe et al., 1993, Thompson and Fox-Kean, 2005). A number of successive studies, however, have also pointed to several qualifications concerning the relation between knowledge spillovers, proximity, and innovation (see Agrawal et al., 2008, Breschi and Lissoni, 2009, Giuliani, 2007, Giuri and Mariani, 2013, Tubiana et al., 2020, among others). Nonetheless, despite these important caveats, the possible effect of local interactions among inventors on the resulting spatial distribution of innovative activities can be hardly ignored.[1]

Yet, if local interactions matter, inventors could find it especially attractive to site with peers of higher quality. Indeed, Mokyr (2016, 2018) argues that "elite inventors" had a pivotal role in the British industrial revolution, as they represented the portion of upper-tail human capital that was relevant for the production of new useful knowledge. Therefore, if local interactions among inventors had any effect at all on the siting of innovative activities, great inventors such as Abraham Darby, Richard Arkwright, or James Watt might have had an especially attractive role toward their peers. Indeed, when looking at the US economy after the mid-1980's, Kerr (2010) finds that breakthrough inventions played an empirically relevant role in driving the migration of innovation clusters. Relatedly, some studies point to the positive impact of contemporary superstar inventors on the innovative performance of their local peers or on the co-localization of a new industry (Almeida and Kogut, 1999, Zacchia, 2018, Zucker and Darby, 1996). By contrast, other studies do not find a comparable local effect for superstars in the specific context of academia (Azoulay et al., 2010, Oettl, 2012, Waldinger, 2012). In this perspective, the extent to which the great inventors of the industrial revolution influenced the spatial distribution of inventive activities may shed some light on the role of upper-tail human capital during the first major wave of innovation in modern economic history.

Under these premises, the present work aims at evaluating the determinants of the spatial distribution of patent inventors across British counties between 1700–1850. For this purpose, we consider all the patents awarded in Great Britain during the period of interest as collected by Woodcroft (1854, 1862) and the related industry classification and quality indicator introduced by Nuvolari et al. (2011) and Nuvolari et al. (2021). In particular, we test for the role of interdependencies in the siting of inventors relying on the localization model put forward by Bottazzi et al. (2007) and on the related maximum likelihood estimation described by Bottazzi and Gragnolati (2015). This approach allows to compare two different scenarios: one in which inventor location is driven only by the intrinsic features of counties, and an alternative scenario in which also local interdependencies in the siting of inventors come into play and lead to a specific agglomeration effect. The historical patent data on which we rely do not allow to define a citation network, as commonly the case for contemporary patent data. Therefore, our analysis cannot fully disentangle the different sources of local interactions among inventors, as done for instance by Agrawal et al. (2008) and Breschi and Lissoni (2009). Nonetheless, our empirical strategy and the specific type of inventors that characterized the British industrial revolution allow to provide at least some interpretation in this regard.

The present work differentiates its self from the related literature in economic history by focusing explicitly on the location of inventive activities, whereas most other contributions look at the geography of the British industrial revolution through the lens of production. For instance, Broadberry and Marrison (2002) and Leunig (2003) discuss the possible impact of lo-

[1] See Audretsch, 1998, Boschma, 2005, Feldman, 1999 and Carlino and Kerr, 2015 for more comprehensive surveys on the empirical literature assessing the role of knowledge spillovers on the geography of innovation. See also Breschi and Lissoni, 2001a,b for some further critical assessments.

calized externalities on productivity in the British cotton textiles industry looking in particular at Lancashire in 1900. On the other hand, Crafts and Wolf (2014) take a broader geographical view by assessing the county-level determinants of location for the textiles industry across the whole of the United Kingdom in 1838. Relatedly, Kelly et al. (2022) look in particular at English counties to analyze which pre-determined factors have a stronger impact on the local share of industry employment in 1831 and 1851. Finally, Hanlon and Miscio (2017) try to disentangle the various sources of agglomeration economies potentially at play in England between 1851–1911, but they focus on population as a dependent variable. Overall, then, the present work adds to the aforementioned literature by providing new county-level evidence regarding specifically the location of inventive activities in Great Britain between 1700–1850. In this sense, the historical focus adopted in this paper can also serve to establish a useful benchmark for the analysis of long-term developments in the geography of innovation. For example, Andrews and Whalley (2021) have recently adopted a long-term view to study how the spatial distribution of inventive activities has evolved within the US from 1865 to 2015. In a somewhat similar spirit, Crescenzi et al. (2020) discuss the patterns of spatial behavior that are specific to the more recent waves of innovation as compared to the earlier ones.

2 Data and descriptive statistics

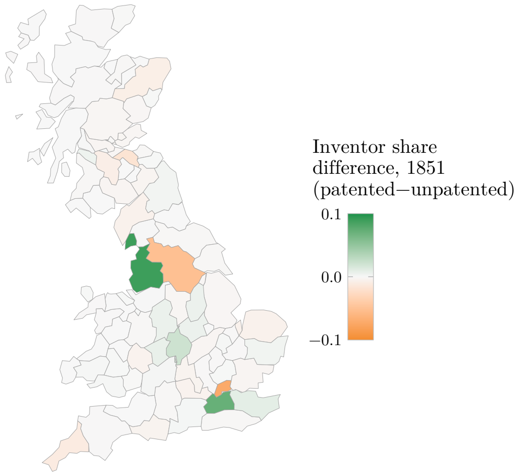
The location of inventive activities during the British industrial revolution is here identified by an inventor's county of residence at the moment in which they were awarded a patent in Great Britain between 1700–1850. Counties are defined according to their historic borders as reconstructed by the Cambridge Group (CAMPOPb, 2021, CAMPOPc, 2021), with the City of London merged into Middlesex. Our analysis then aims at explaining the cross-county distribution in the number of patented inventions observed over a given period of time. For this purpose, we link the patent data collected by Woodcroft (1862) to the related information on inventor residence reported by Woodcroft (1854). These data regard only inventions that obtained a patent from the administration of England and Wales, which remained separate from its counterpart in Scotland until the Patent Law Amendment Act in 1852. However, nearly all patents awarded in Scotland were granted also in England and Wales (see Bottomley, 2014a, p. 54). As a result, our analysis spans virtually all the inventions that were patented in Great Britain from the early developments in iron smelting and metallurgy, to the early introduction of steam engines, up to the mature phases of the cotton industry. In particular, each patent in our sample is associated with its inventors and their county of residence, so that the possible spatial mobility of recurring inventors is fully tracked.

The original patent data for the span 1700–1850 is filtered according to three criteria. First, we drop those observations for which the inventor's county of residence is entirely unknown (about 2% of the total). Second, one wants to ensure that each patentee coincides with the actual inventor and with the corresponding locus of innovation. For this reason, we drop the 399 patents that were registered in the name of patent agents on behalf of unknown actual inventors (about 3% of the total). Indeed, patent agents resided almost exclusively in London precisely to benefit from proximity to the only patent office in the country, whereas their clients

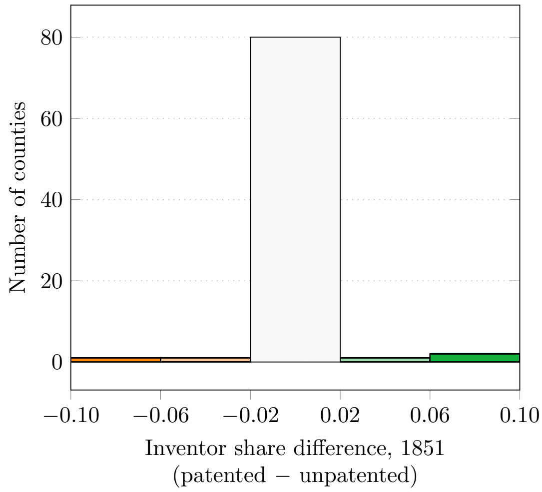
Table 1: Data coverage and sources.

VARIABLE BY COUNTY	YEARS	SOURCE	DEFINITION
Patent inventors	1699–1850	Woodcroft (1854) and Nuvolari and Tartari (2011).	Number of patentees residing in the county.
CONTROLS			
Population	1699–1850	HOC (1831, 1851), Webster (1952), Wrigley (2007, 2009).	Number of residents in the county.
Manufacturing dummy	1699–1850	Shaw-Taylor and Wrigley (2014), Trinder (2008), Wrigley (2007, 2009) and our own elaborations.	Binary variable with value 1 in counties that hosted historically known manufacturing centers while also experiencing an above-average population growth, and value 0 elsewhere.
City concentration (HHI)	1699–1850	Bairoch et al. (1988).	Herfindahl–Hirschman index (HHI) of population concentration in cities within each county.
Access to high-quality inventors	1699–1850	Woodcroft (1862) and Nuvolari et al. (2021).	Number of high-quality patent inventors by county summed to the number of high-quality inventors in other counties weighted by inverse-distance between counties. High-quality inventors are defined as those in the top 5% of the Bibliographic Reference Index (BRI) distribution.
Access to coal	1699–1850	Flinn (1984, p. 26) and Berg (2005, pp. 36–37).	Coal output by county summed to coal output in other counties weighted by inverse-distance between counties.
Distance from London	Fixed	Own elaboration.	Euclidean distance of each county centroid from London.
Employment by industry	1841	CAMPOPa (2021).	Number of employees by county in each industry.

COVERAGE: 12 021 patents corresponding to 92% of the total awarded between 1700–1850. The classification of patents by industry is taken from Nuvolari and Tartari (2011). Spatial units consist of 85 historic counties covering the whole of mainland Great Britain plus Bute and Anglesey, with the City of London merged to Middlesex.



(a) Map.



(b) Histogram.

Figure 1: Inventor share difference between patented and unpatented inventions.

The figure illustrates the difference between county shares of inventors with and without a patent. Patented inventions are taken from Woodcroft (1854, 1862), whereas unpatented inventions are taken from Moser's (2011) for the 1851 Crystal Palace world fair. *Left panel*: The map shows that the difference between county shares of patented and unpatented inventions is slightly positive in Lancashire (0.09) and Surrey (0.07), while being slightly negative in Middlesex (−0.07) and Yorkshire West Riding (−0.05). *Right panel*: The histogram shows that the difference between county shares of patented and unpatented inventions ha a symmetric empirical distribution closely centered around 0.

could well be based in other counties (see Bottomley 2014b, p. 65–72, Dutton 1984, p. 86–89, and Khan 2005, p. 58). Third, we consider only the contiguous counties that compose mainland Great Britain plus Bute and Anglesey, thus dropping all inventors that resided on other minor British isles, in Ireland, or abroad (about 3% of the total). This leaves us with 85 spatial units and 12 021 patents, which correspond to about 92% of the total between 1700–1850.

The typical inventor in our sample obtained only one patent during his lifetime and relied rather limitedly on co-patenting. For instance, about 72% of all inventor names in our sample occur only once throughout the period 1700–1850, while recurring names often correspond to widespread English names that are likely to identify homonyms. In this setting, the high cost of patenting can contribute to explain why it was a relatively rare event in the course of a lifetime. As expressed in 2010 constant prices, the overall cost of obtaining a patent in Great Britain during the period of interest is estimated in the range £107 000–£178 000 (see Bottomley, 2014b, p. 61). These high patenting costs did not seem to induce an especially high propensity to co-patent as a possible strategy for cost sharing. By means of comparison, in present-day United Kingdom "more than 80% of all patents are registered to more than one inventor [...] often in collaboration with universities, public agencies, and research centres" (Crescenzi et al., 2016, p. 177). On the other hand, less than 11% of all patents awarded in Great Britain over the time period 1700–1850 were associated to more than one patentee; moreover, when co-patenting occurred, it often involved father and son. Overall, then, we are looking at a sample of "independent inventors" that had to bear significant costs in order to obtain a patent.

Clearly, the use of patent data to capture the spatial distribution of innovative activities is

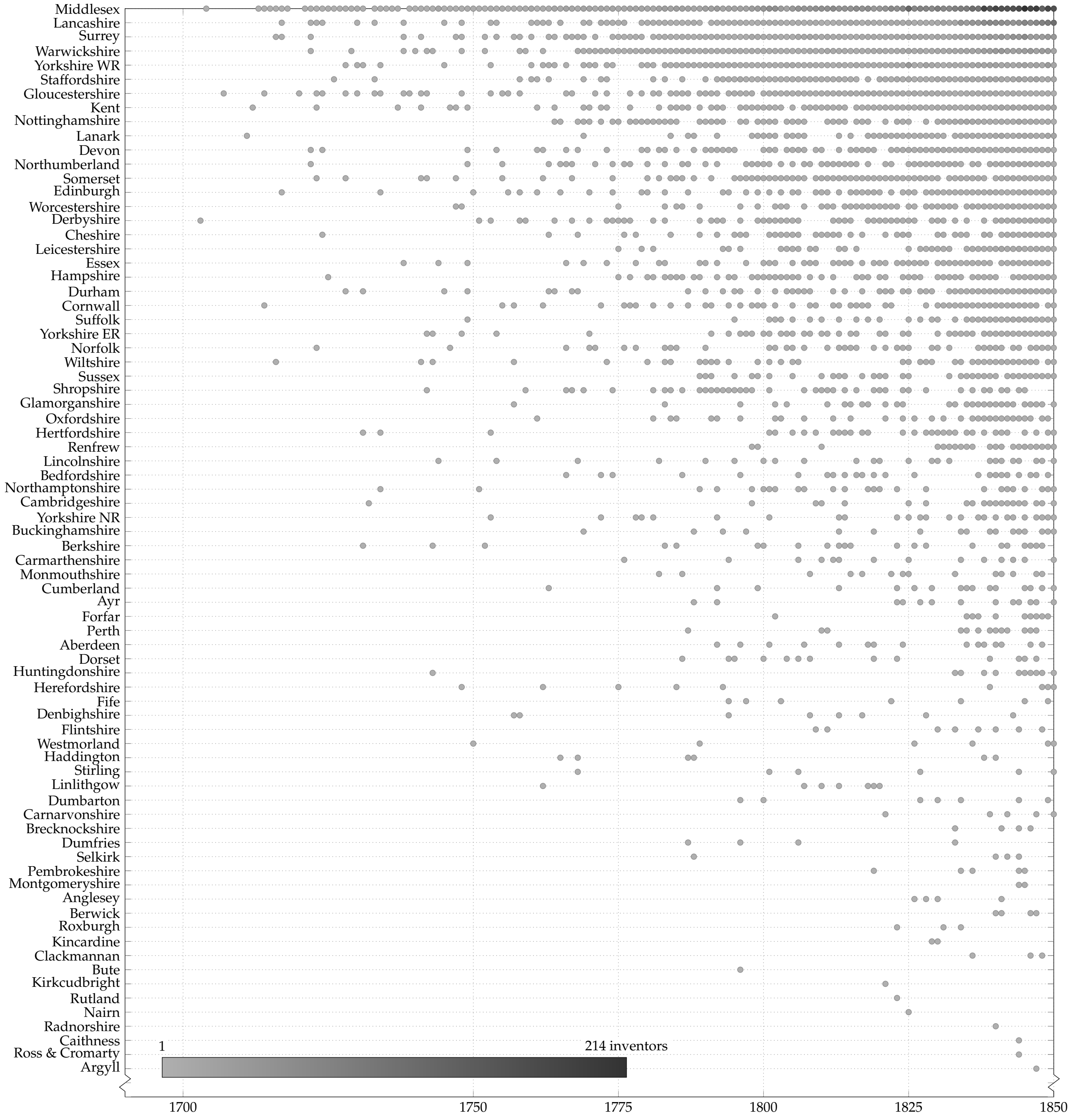


Figure 2: Arrival of patent inventors by county and year, 1700–1850.

potentially subject to some limitations. In particular, given the high cost of patenting, many inventions could remain unpatented and their distribution across counties could possibly diverge from that of patented inventions, for instance as a result of differential access to credit markets. In order to check whether our patent data suffer from this type of distortion, we compare the residence of patent inventors in our sample between 1826–1850 with the residence of other inventors that participated at the 1851 Crystal Palace World's Fair in London but did not obtain a patent (see Moser, 2011). As shown in Figure 1, these two types of inventors display a very limited discrepancy in their spatial distributions across counties. In particular, the map in Figure 1a shows that the difference between the county shares of inventors that patented their inventions and those that did not obtain a patent is nearly null in most counties, while being slightly higher in Lancashire, Surrey, Middlesex, and Yorkshire West Riding. In addition, the histogram in Figure 1b shows that the difference between county shares of inventors with and without patents has a symmetric distribution tightly centered around 0, so that the location of patent inventors does not seem to suffer from any systematic bias as compared to inventors that did not obtain a patent. These visual insights are furthered by a standard statistical test. Namely, the two-sample Kolmogorov-Smirnof test does not reject the null hypothesis that the county shares of patent inventors between 1826–1850 follow the same distribution as the county shares of 1851 Crystal Palace inventors that did not obtain a patent (p-value $= 0.45$). In fact, these two measures of inventor location have a correlation coefficient of 0.95. Overall, then, this evidence suggests that the patent data in our sample may well be representative of the spatial distribution of inventive activities by and large.

As shown in Figure 2, patent inventors in our sample are unevenly distributed across counties and become gradually more frequent over the course of British industrialization. Again, patenting remained a very expensive and complicated administrative process throughout the period 1700–1850, so that the increasing frequency of patents is unlikely to be associated with a systematic decline in their quality (see Bottomley 2014b, pp. 42–65, 74 and MacLeod 1988, p. 76). At any rate, each patent is linked to multiple quality indicators that serve here to identify high-quality inventions. Such indicators have been developed starting from Woodcroft's (1862) work for the English Patent Office, in which each patent is associated to the various references it receives in some of the main legal and technical publications of the time. Using this source, Nuvolari and Tartari (2011) construct a reference count known as the Woodcroft Reference Index (WRI), which can also be normalized by sub-period average (WRI*) to reduce the possible bias in favor of more recent patents. Then, to further correct for time and industry biases while also sensibly extending the set of publications considered in the reference count, Nuvolari et al. (2021) construct the so-called Bibliographic Composite Index (BCI). Being more accurate and sophisticated, this latter indicator represents our preferential way to identify high-quality inventions according to an upper-tail threshold. In our baseline specification, we set such threshold to the top 5% of the BCI distribution so as to identify the high-quality inventors shown in Figure 3. Yet, all of the estimates that will be discussed in Section 4.1 are robust to the use of any of the other patent quality indicators discussed above, as well as to sensible variations of the quality threshold.

Notably, our patent data display some industry specificities as soon as they are observed

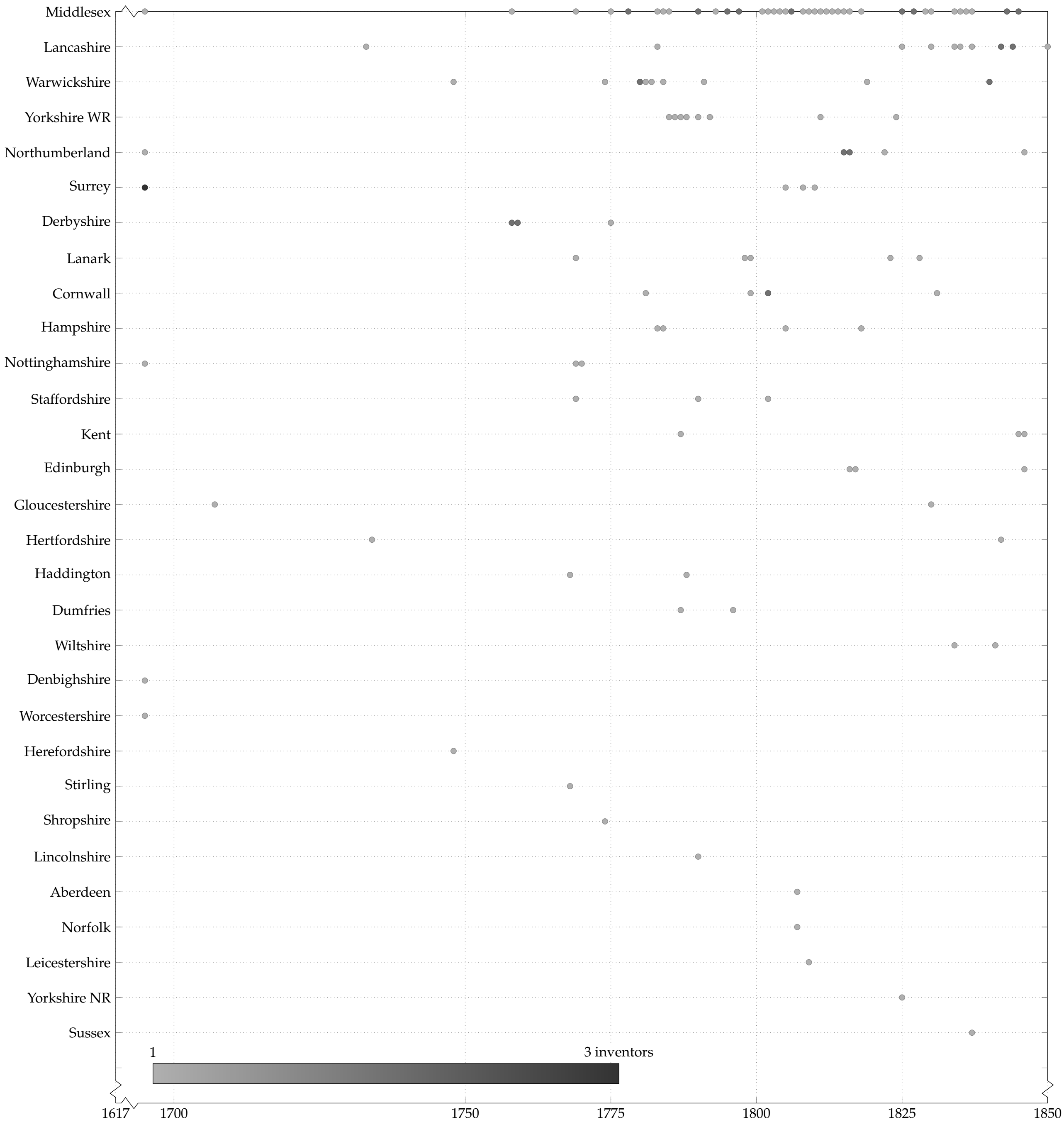


Figure 3: Arrival of high-quality patent inventors by county and year, 1617–1850 (top 5%).

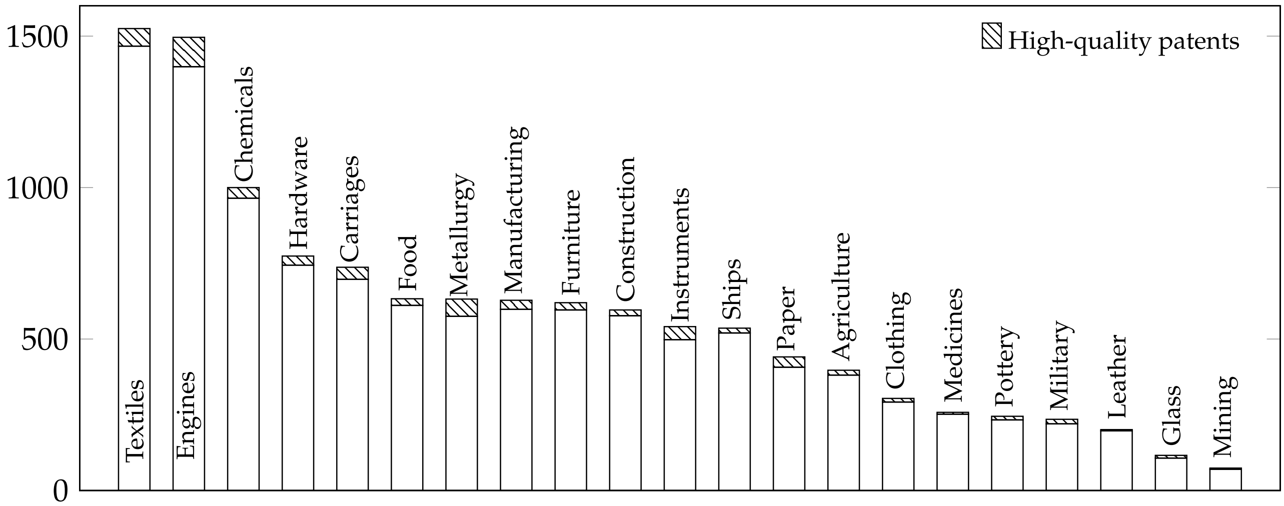


Figure 4: Number of patents by industry.

Note: The classification of patents by industry comes from Nuvolari and Tartari (2011). See text for further details.

from a spatial perspective. For each patent, we identify its industry of application according to Nuvolari and Tartari's (2011) classification, which rests on the synthetic description of invention provided in Woodcroft (1854).[2] As shown in Figure 4, almost all patents regard the secondary sector, with engines and textiles representing the most common industries of application for patents granted between 1700–1850. Specifically, both engines and textiles are associated to about 1500 patents each, accounting overall for nearly 25% of our sample. By contrast, the third most common industry of application (chemicals) follows from afar with about 1000 patents. More importantly for the purpose of this study, the spatial distributions of inventors in engines and textiles display significant specificities.

On the one hand, the typical county of residence differs substantively for inventors in engines and textiles as compared to their homologues in other industries. Figure 5 illustrates this type of evidence by associating each patent inventor to the centroid of his county of residence, so as to trace the resulting mean spatial center between 1700–1850. As a general rule, all industries tend to move from South-East to North-West during the industrial revolution: by and large, this evolution reflects the gradual growth of Lancashire, the West Riding of Yorkshire, and Warwickshire as possible alternatives to Middlesex and Surrey (see also Figure 2). Nonetheless, the mean spatial center of patent inventors differs across industries both in its starting point and subsequent spatial evolution. As shown in Figure 5, textiles inventors were not as driven toward London until the first half of the 19th century, and later on they moved even more markedly toward Lancashire. Inventors in engines, instead, were initially more comparable to their homologues in other industries from a spatial point of view, but during the mature phase of industrialization they took a more pronounced step toward Lancashire. For inventors in other industries, instead, the displacement over time is much less accentuated and London essentially continued to be their main hub all throughout the industrial revolution.[3]

[2]Following this approach, Nuvolari and Tartari (2011) classify all patents granted between 1617–1841. In this paper we extend their classification to 1850.

[3]To be sure, Figure 5 aggregates all industries other than engines and textiles precisely because the typical

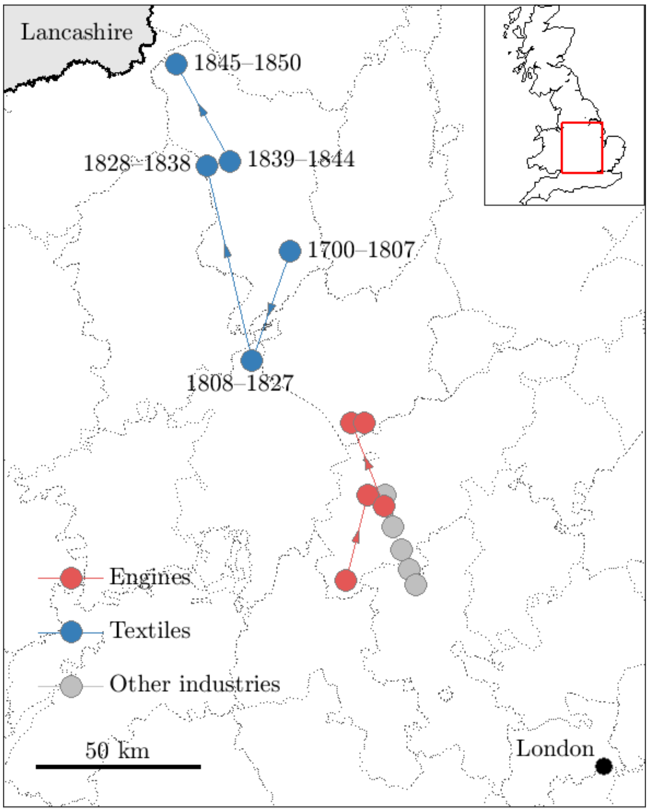


Figure 5: Evolution of the mean spatial coordinate of patent inventors, 1700–1850.

Note: This figure is constructed by associating each patent inventor to the centroid of his county of residence. Observations are then divided into five quantile groups according to the year of patent registration, so as define the time periods indicated on the map. This periodization ensures that the amount of spatial noise remains approximately constant across time periods, so that the evolution of the mean spatial center can be meaningfully compared over time and across industries with a sufficiently large number of observations.

On the other hand, industry specificities emerge also by looking at the dispersion of patent inventors across British counties. To illustrate this aspect, Figure 7 shows the cross-country distribution of patent inventors in each sub-period of the industrial revolution. For instance, between 1700–1750, about 90% of counties hosted less than 0.01% of inventors in textiles (see the leftmost white bar in Figure 7a), thus indicating that a vast majority of counties was essentially void of engines inventors while a few counties had large shares of them. A similar fact holds for inventors in the textiles industry during the same time period (see the leftmost white bar in Figure 7b). By contrast, just above 60% of counties were equally void of inventors operating in other industries between 1700–1750 (see the leftmost white bar in Figure 7c). In this sense, during the fist half of the 18th century, inventors in textiles and engines were less geographically common than their homologues in other industries. Then, over the course of industrialization, all patent inventors became more geographically common: that is, the leftmost bar in Figures 7a-7c systematically decreases as times elapses from 1700 to 1850. Nonetheless, the extent of dispersion—and thus of spatial concentration—is again rather heterogeneous across industries. Although a very few counties keep dominating the landscape by hosting a very large share of inventors all throughout the time period under scrutiny, the

county of residence for inventors in each of these other industries closely resembles the aggregate pattern. Hence, we chose to show only the aggregate pattern in order to facilitate the visualization of the figure.

Table 2: PST-industry matching.

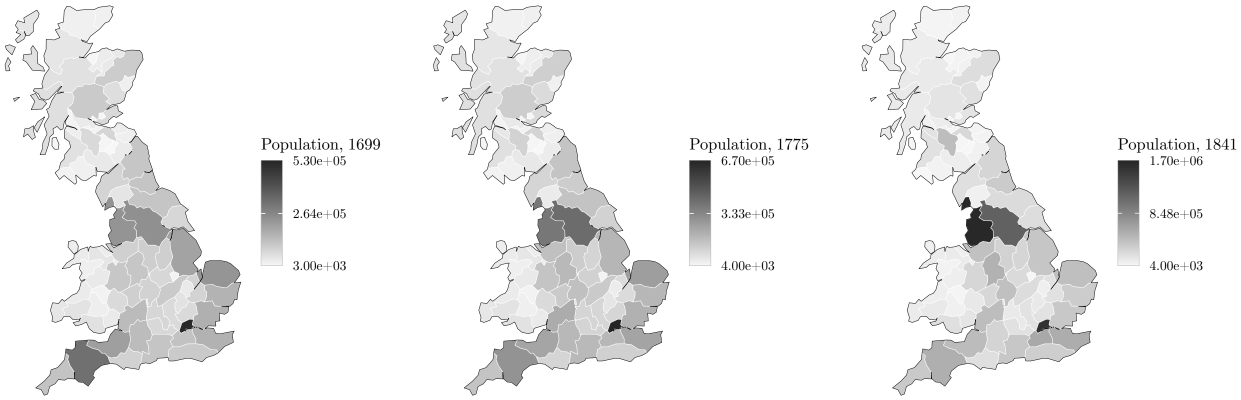
PST code	Group	Section	Occupation	Industry
2,65,1,1-2	Machines and tools	Machine making	Engine maker, pump maker, millwright, mechanic.	Engines
2,65,3,–		Engineering		
5,35,8,–	Professions	Engineering professions	Engineer	
5,36,8,–	Professional support			
2,20,–,–	Textiles	Manufactures of textiles	Spinner, weaver, dyer, printer, preparer, threader, twister, processor, maker, fuller, carder, scribbler, shearman, etc.	Textiles
2,[0-15],– 2,[25-63],– 2,65,1,[3-60] 2,65,2,– 2,65,4,– 2,65,99,60 2,[66-95],–	Other manufacturing occupations			Other manufacturing

Note: PST codes, groups, sections, and occupations are taken from the PST dictionary provided by CAMPOPa (2021), while the allocation of the various occupations to each industry is the authors' own elaboration.

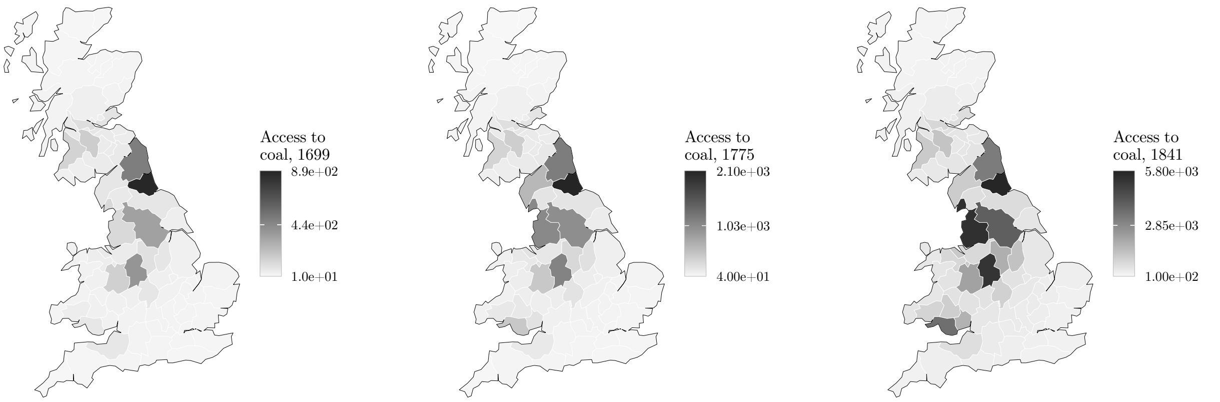
extent of such spatial dominance is heterogeneous across industries (see the rightmost bar in each sub-plot of Figure 7). For instance, the leading county may host up to 50% of industry inventors in engines (see Figure 7a), while its counterpart in textiles attracts at most 38% of industry inventors (see Figure 7b). Overall, then, these descriptive statistics point to a complex geography of innovation in which industry specificities are to be taken into account.

The present work rests also on a variety of other data that serve as descriptors of county attractiveness. Table 1 summarizes these data and their sources, while Figures 6 and 9 map some of the main control variables to provide a visual summary of their evolution over the course of British industrialization. Typically, only one or a few snapshots are available for each of the relevant variables, so that interpolation and extrapolation must be used to estimate missing observations when needed. Given the admittedly approximate nature of the data at stake, we resort simply to linear interpolation/extrapolation over time. Under these caveats, our control variables are broadly meant to trace the spatial distribution of productive activities as well as the different potential for the expansion of production that characterized each county.

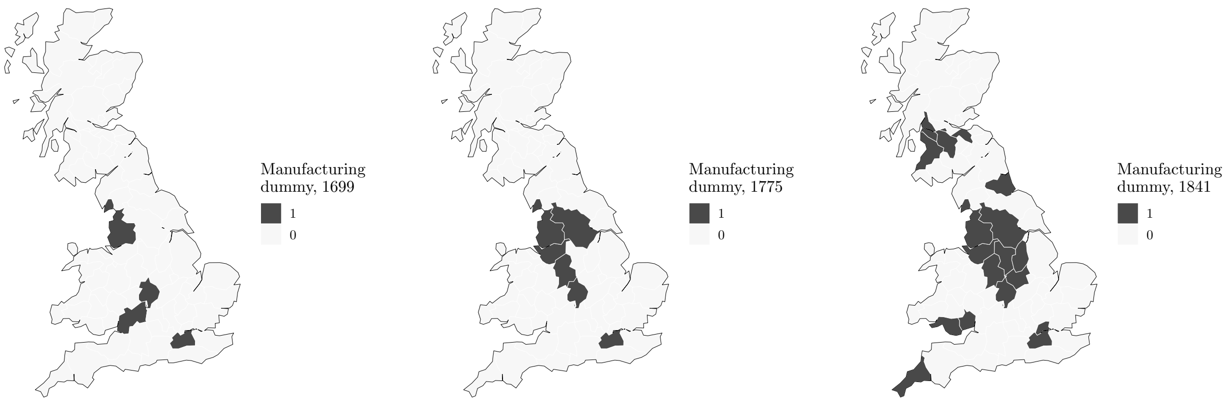
Controlling for the location of production poses some constraints to our empirical analysis because employment data by county become available only with the 1841 Census Report. As long as we run our analysis without any distinction on the industry in which inventors were active, we can overcome this data limitation by using population size by county as a proxy for the spatial distribution of aggregate production. Population size by county can be recovered also for the early phases of industrialization relying on the census data reported by the House



(a) Population.



(b) Access to coal.



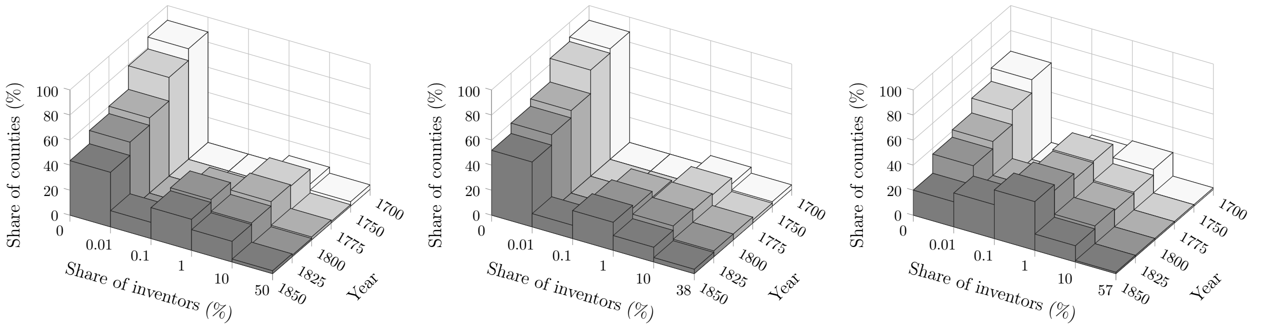
(c) Manufacturing dummy.

Figure 6: Some of the main control variables in 1699, 1775, and 1841.

of Commons Parliamentary Papers (1831, 1851) and by Webster (1952), as well as on the estimates by Wrigley (2007, 2009). The correlation between population size and total employment by county is indeed about 0.99 in 1841, so that little accuracy should be lost by using the former as proxy for the latter. At any rate, we also construct a dummy variable to identify counties that hosted fast-growing manufacturing centers (see in particular the identification of "textile or metalworking towns" made by Trinder, 2008 and the "industrial counties" defined in Shaw-Taylor and Wrigley, 2014, as well as the related account by Barry, 2008). Albeit admittedly crude, this dummy variable serves to highlight the distinct spatial distribution of manufacturing activities, so as to capture the specific linkages that patent inventors might have with this particular sector of production. For the post-1841 period, instead, we recover industry-level employment by exploiting occupational data from the Cambridge Group for the History of Population and Social Structure, using both male and female occupations for workers aged under and over twenty years (see CAMPOPa, 2021, for further details). Starting from occupational data classified into primary, secondary, and tertiary (PST) sectors and groups, we reconstruct industry-level employment by county according to the scheme reported in Table 2 (see Wrigley, 2010, for further details on the PST classification). Similarly to the manufacturing dummy discussed above, these various employment variables serve in particular to control for industry-specific linkages between inventive and productive activities, such as the ancillary relations discussed by Balderston (2010, p. 578).

The organization of production at the county level is further characterized in terms of its concentration across urban centers. When the population of a county concentrates in a few larger cities, the resulting local labor markets tend to be deeper than what they would be if the same number of individuals were dispersed across several smaller towns. In turn, deeper labor markets increase the potential for specialization, which may ultimately stimulate innovation. To account for this possible effect, we compute the Herfindahl–Hirschman index (HHI) of city concentration by county. Specifically, we use the data on city size by Bairoch et al. (1988) to obtain the county population share $s_{l,i}$ of city i in county l, so that the corresponding measure of city concentration by county is defined as $\mathrm{HHI}_l = \sum_i s_{l,i}^2$. Clearly, this measure can also be interpreted more broadly as an indicator of urbanization.

Another key aspect relating to the productive potential of counties during the British industrial revolution has to do with local access to coal. Authors such as Wrigley (1990) emphasize the key role of coal as the "new" source of energy that ultimately allowed for the large-scale growth of production during the industrial revolution. In parallel, Clark and Jacks (2007) show that the transportation cost of coal determined most of its final price, so that the access to this key resource generally decreased with the physical distance from the coal fields. To capture these aspects, we resort to the notion of potential accessibility from the transport geography literature (see Rodrigue, 2020, pp. 380–382). According to this approach, the potential access that a spatial unit has to a mobile resource depends on its local endowment and on the distance from other sources. Formally, the vector $\boldsymbol{e}_h = (e_{1,h}, \ldots, e_{L,h})'$ describes the endowment of resource h across L counties, while $\boldsymbol{W}_{L\times L}$ is the related spatial weight matrix. Given the distance measure $d_{i,j}$ between county i and j, matrix $\boldsymbol{W}$ is conventionally defined as having entries $w_{i,j} = 1/d_{i,j}$ for $i \neq j$ and 0 on the diagonal. To avoid dependence on the unit of measurement



(a) Engines. **(b)** Textiles. **(c)** Other industries.

Figure 7: Distribution of inventors across counties, 1700–1850.

Note: For each sub-period and industry, the leftmost bar includes all and only counties hosting exactly zero patent inventors, whereas the rightmost bar includes counties that host more than 10% of patent inventors in the relevant sub-period and industry.

while preserving symmetry, these entries can be transformed as $w^s_{i,j} = w_{i,j}L/\sum_i \sum_j w_{i,j}$, thus obtaining the globally standardized spatial weight matrix $\boldsymbol{W}^s$. A measure of potential access to resource h can then be defined as $\tilde{\boldsymbol{e}}_h = \boldsymbol{e}_h + \boldsymbol{W}^s \boldsymbol{e}_h$. In the present context, we use this definition to construct our measure of access to coal. In particular, we realy on Flinn's (1984) coal output data to construct the vector of endowments by county and we define $d_{i,j}$ as the Euclidean distance between counties. As shown in Table 1, we use the same approach also to quantify potential accessibility to high-quality and out-of-industry inventors. As a remark, the results presented here are qualitatively robust to alternative definitions of $\boldsymbol{W}$ relying on higher-order contiguity criteria, as well as to the use of row-normalization.

3 Localization model

We rely on the localization model introduced by Bottazzi et al. (2007) to assess the determinants of inventor location during the industrial revolution. In broad terms, such model studies how N agents distribute across L discrete alternatives when positing a stochastic process with Ehrenfest-like "disruptions" and Brillouin-like "creations". These can be invariably interpreted either as the genuine exit and entry of agents or as preference revision. The resulting stochastic process bears a close connection with the generalized Polya urn framework discussed in Arthur et al. (1987) and Dosi et al. (1994), particularly for the role played by positive feedback mechanisms. Here, however, the stochastic process entails an irreducible Markov chain characterized by a unique invariant distribution independent from the initial state of the system, thus being ergodic. Such distribution represents the stochastic equilibrium of the model and its governing parameters can be estimated thanks to ergodicity. In this section we provide only an essential summary of the model as introduced by Bottazzi et al. (2007) and the related working paper by Bottazzi and Secchi (2007). Instead, Garibaldi and Scalas (2010, esp. Ch. 7) provide a general characterization of models in the Ehrenfest-Brillouin class and further clarify their relation with the Polya urn framework.

Consider a population of N agents and L discrete locations. Due to agent heterogeneity and limited information about their preferences, the attractiveness of a location is defined in probabilistic terms rather than in a deterministic way. Following Arthur (1990), the probability p_l for site l to be selected is proportional to a fixed component plus a function of the number of agents in l. Namely,

$$p_l = \frac{a_l + bn_l}{A + bN} , \quad a > 0,\, b \geq 0 , \tag{1}$$

where n_l is the number of agents in l, $A = \sum_{j=1}^{L} a_j$, and $N = \sum_{j=1}^{L} n_j$. In equation (1), the fixed term a_l represents the "intrinsic attractiveness" of site l, as it captures those location-specific advantages that are independent from how many agents site in l. For example, a location that hosts more productive activities can prove intrinsically more attractive for inventors, as it may guarantee a higher local demand for technological innovation. On the other hand, equation (1) allows for p_l to depend also on a "social term", which grows linearly with the number of agents in l according to the magnitude of the global parameter b.[4] Specifically, the siting of agents is subject to a positive feedback mechanism for $b > 0$. In this case, having an extra agent that localizes in l leads other agents to select l with higher probability, thus generating an interdependence in the siting of agents. We thus refer to such an effect as a localized externality, given that it is entirely self-contained within site l without any spatial diffusion toward other sites.

What form of spatial order, if any, can emerge under this setup? To address this question, one can consider a dynamical system in which a state is represented by the occupancy vector

$$\boldsymbol{n} = (n_1, \ldots, n_l, \ldots, n_L) , \tag{2}$$

where $n_l \geq 0\ \forall l$ and $\sum_{l=1}^{L} n_l = N$. In such system, the transition from $\boldsymbol{n}$ to the subsequent state $\boldsymbol{n'}$ occurs via a unary move, which decreases the occupancy vector by one unit in m while increasing it by one unit in l. Formally,

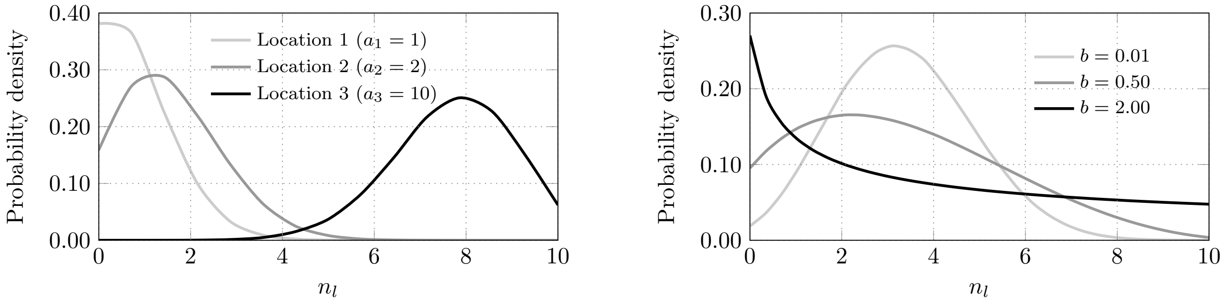
$$\boldsymbol{n'} = \boldsymbol{n} + \boldsymbol{\delta}_l - \boldsymbol{\delta}_m , \tag{3}$$

where $\boldsymbol{\delta}_j = (0, \ldots, 0, 1, 0, \ldots, 0)$ is a vector of length L with the j-th entry equal to 1 and all other entries equal to 0. This elementary event can be invariably interpreted either as the genuine exit of an incumbent followed by the entry of a new and generally different agent, or as choice revision on the side of an agent.

The evolution from state $\boldsymbol{n}$ to $\boldsymbol{n'}$ is specified by a transition probability. Assuming that none of the agents has a higher chance to revise its location, the probability of "disruption" in m is uniform. Hence, the so-called "Ehrenfest term" is $P(\boldsymbol{n}_m \,|\, \boldsymbol{n}) = n_m/N$, where $\boldsymbol{n}_m = \boldsymbol{n} - \boldsymbol{\delta}_m$. In turn, the probability of "creation" in l derives from conditioning equation (1) to the "disruption" in m, thus defining the so-called "Brillouin term" $P(\boldsymbol{n'} \,|\, \boldsymbol{n_m})$. It follows that the one-step transition probability from state $\boldsymbol{n}$ to $\boldsymbol{n'}$ is

$$P(\boldsymbol{n'} \,|\, \boldsymbol{n}) = P(\boldsymbol{n}_m \,|\, \boldsymbol{n}) \cdot P(\boldsymbol{n'} \,|\, \boldsymbol{n_m}) = \frac{n_m}{N} \cdot \frac{a_l + b(n_l - \delta_{l,m})}{A + b(N-1)} , \tag{4}$$

[4] See Bottazzi et al. (2017) for a version of the model in which p_l is a quadratic function of n_l.



(a) Marginal distribution of $\pi(\boldsymbol{n}; \boldsymbol{a}, b)$ for varying values of a_l and fixed $b = 0.01$, with $N = 10$ and $L = 3$.

(b) Marginal distribution of $\pi(\boldsymbol{n}; \boldsymbol{a}, b)$ for varying values of b and fixed $a_l = 1\ \forall l$, with $N = 10$ and $L = 3$.

Figure 8: Marginal distributions of $\pi(\boldsymbol{n}; \boldsymbol{a}, b)$

where $\delta_{l,m}$ is Kronecker's delta equal to 1 for $l = m$ and zero otherwise. Equation (4) determines a finite Markov chain having as state space the set of L-tuples $S_L^N = \{\boldsymbol{n}, n_l \geq 0, \sum_{l=1}^{L} n_l = N\}$. Given that $a_l > 0\ \forall l$, each state in S_L^N can be reached with positive probability, and a location that is empty once can be successively occupied again. Hence, all states in S_L^N are persistent and the related Markov chain is irreducible. Consequently, it admits a unique invariant distribution $\pi(\boldsymbol{n}; \boldsymbol{a}, b)$ independent from the initial state of the system. In this sense, the distribution $\pi(\boldsymbol{n}; \boldsymbol{a}, b)$ is ergodic (see Bottazzi and Secchi 2007, especially Propositions 3.3–3.4 and Garibaldi and Scalas 2010, especially Chs. 6–7).

The invariant distribution $\pi(\boldsymbol{n}; \boldsymbol{a}, b)$ is governed by the intrinsic attractiveness parameters $\boldsymbol{a} = (a_1, ..., a_L)$, and by the externality parameter b. Together, these $(L + 1)$ parameters fully determine the expected number of agents in each of the L locations in $\boldsymbol{n}$. In general, $\pi(\boldsymbol{n}; \boldsymbol{a}, b)$ has the Polya form

$$\pi(\boldsymbol{n}; \boldsymbol{a}, b) = \frac{N!\Gamma(A/b)}{\Gamma(A/b + N)} \prod_{l=1}^{L} \frac{1}{n_l!} \frac{\Gamma(a_l/b + n_l)}{\Gamma(a_l/b)} \,. \tag{5}$$

However, $\pi(\boldsymbol{n}; \boldsymbol{a}, b)$ reduces to a multinomial form in the specific case of null externalities (that is $b = 0$):

$$\pi(\boldsymbol{n}; \boldsymbol{a}, b = 0) = N! \prod_{l=1}^{L} \frac{1}{n_l!} \left(\frac{a_l}{A}\right)^{n_l} \,. \tag{6}$$

Therefore, the model predicts two qualitatively distinct scenarios depending on whether localized externalities have a role in the siting of agents.

The qualitative behavior of $\pi(\boldsymbol{n}; \boldsymbol{a}, b)$ is illustrated in Figure 8 by plotting the marginal distributions of $\pi(\boldsymbol{n}; \boldsymbol{a}, b)$ for varying values of a_l and b. In each sub-figure, only one of the underlying parameters is allowed to vary while all other parameters are kept constant, so as to highlight the effect of the varying parameter on the probability mass. On the one hand, locations associated to a higher value of a_l face a higher probability to host more agents (see Figure 8a). On the other hand, a higher value of b moves the probability mass from center to tails of the distribution other things being equal (see Figure 8b). Indeed, when localization is more interdependent, a site faces higher chances to end up being either highly populated or nearly void, *despite* being similar to all other locations in terms of its intrinsic attractiveness a_l. In this scenario, some degree of spatial concentration arises with certainty and its primary driver is the positive feedback underlying localization. By converse, the weakening of b drives the distribution to become more symmetrically centered around some "typical" value of n_l, which is ultimately determined by how location l fares relative to other locations in terms of its intrinsic attractiveness a_l. If b is sufficiently low and a_l is fairly uniform across locations, then the expected number of agents will tend to be similar across locations. In this case, spatial concentration would be low. But if locations are more heterogeneous in their intrinsic attractiveness a_l while b remains low, then those locations with a higher a_l will be expected to host more agents (see Figure 8a). Also in this case, therefore, spatial concentration would be relatively strong. In this sense, the observed degree of spatial concentration is an insufficient information to infer whether localized externalities are particularly strong. In fact, such concentration could emerge also from the combination of *low* externalities with highly heterogeneous intrinsic attractiveness (on a similar point, see also Ellison and Glaeser, 1997, pp. 896–897, Proposition 1). The present model and the related estimation procedure serve precisely to discriminate these different scenarios and disentangle the various underlying effects starting from the observed spatial distribution of patent inventors.

4 Empirical approach

The model described in Section 3 can be taken to the data presented in Section 2 with the methodology described by Bottazzi and Gragnolati (2015) for cross-sectional data. In broad terms, such an approach consists in implementing maximum likelihood estimation as well as a procedure of model selection, so as to ultimately derive the marginal elasticities associated to the various determinants of localization.

4.1 Estimation and marginal elasticities

The model presented in Section 3 has $L + 1$ unknown parameters, namely the intrinsic attractiveness vector $\boldsymbol{a} = (a_1, \ldots, a_L)$ and the externality coefficient b. Hence, the model cannot be estimated in such form on a single cross section with L observations. To overcome this limitation, one can instead interpret the equilibrium distribution $\pi(\boldsymbol{n})$ as being governed by a set of $H \ll L$ observed variables $\boldsymbol{x}_l = (x_{1,l}, \ldots, x_{h,l}, \ldots, x_{H,l})$ via the function $g(\boldsymbol{\beta}, \boldsymbol{x}_l)$, where $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_h, \ldots, \beta_H)$ is a vector of unknown parameters to be estimated. Notice in particular that $\pi(\boldsymbol{n})$ depends on the ratio a_l/b in equation (5), so that a generic functional specification of the Polya model can be obtained by setting $a_l/b = g(\boldsymbol{\beta}, \boldsymbol{x}_l)$. On the other hand, $\pi(\boldsymbol{n})$ depends only on a_l in equation (6), so that the corresponding functional specification for the Multinomial model is obtained by setting $a_l = g(\boldsymbol{\beta}, \boldsymbol{x}_l)$.

Substituting $g(\boldsymbol{\beta}, \boldsymbol{x}_l)$ for a_l/b in (5) and for a_l in (6), one can express the equilibrium distribution $\pi(\boldsymbol{n}; \boldsymbol{\beta})$ for the Polya and Multinomial model respectively. In each of these cases, the maximization of $\log \pi(\boldsymbol{n}; \boldsymbol{\beta})$ with respect to the unknown parameters $\boldsymbol{\beta}$ yields the maximum likelihood point estimates $\hat{\boldsymbol{\beta}}$. Specifically, the log-likelihood as function of the unknown parameters $\boldsymbol{\beta}$ for the Polya model reads

$$\log \pi(\boldsymbol{\beta} \mid \boldsymbol{n}, \boldsymbol{x}) = \log(N!) - \sum_{l=1}^{L} \log(n_l!) + \sum_{l=1}^{L} \sum_{k=0}^{n_l-1} \log(g(\boldsymbol{\beta}, \boldsymbol{x}_l) + k) - \sum_{k=0}^{N-1} \log(G + k) \ , \tag{7}$$

where $G = \sum_l a_l/b = \sum_l g(\boldsymbol{\beta}, \boldsymbol{x}_l)$. Maximizing (7) with respect to $\boldsymbol{\beta}$ yields the point estimates $\hat{\boldsymbol{\beta}}$ and the coefficients $\hat{g}_l = g(\hat{\boldsymbol{\beta}}, \boldsymbol{x}_l) = \widehat{a_l/b}$ and $\hat{G} = \sum_{l=1}^{L} \hat{g}_l$ for the Polya model. On the other hand, the log-likelihood for the Multinomial model reads

$$\log \pi(\boldsymbol{\beta} \mid \boldsymbol{n}, \boldsymbol{x}) = \log(N!) - \sum_{l=1}^{L} \log(n_l!) + \sum_{l=1}^{L} n_l \left(\log g(\boldsymbol{\beta}, \boldsymbol{x}_l) - \log G\right) \ , \tag{8}$$

where $G = \sum_l a_l = \sum_l g(\boldsymbol{\beta}, \boldsymbol{x}_l)$. Maximizing (8) with respect to $\boldsymbol{\beta}$ yields the point estimates $\hat{\boldsymbol{\beta}}$ and the coefficients $\hat{g}_l = \hat{a}_l = g(\hat{\boldsymbol{\beta}}, \boldsymbol{x}_l)$ and $\hat{G} = \sum_{l=1}^{L} \hat{a}_l$ for the Multinomial model.

The estimated coefficients $\hat{g}_l$ can then be straightforwardly plugged into equation (1) to estimate the probability p_l. By doing so, one can assess the impact of $x_{h,l}$ or n_l on the probability p_l and thus on the attractiveness of a location. In order to work on an unbounded measure of attractiveness, we consider the probability transformation

$$q_l = -\log(1 - p_l) \ . \tag{9}$$

Equation (9) is unbounded from above, so that enough room for variation is provided even when the overall attractiveness of a location is high. By measuring how much q_l varies in response to an infinitesimal change in $x_{h,l}$ and n_l, one can assess the impact of these variables on the attractiveness of a location.

As a last step for estimation, one needs to specify the function $g(\boldsymbol{\beta}, \boldsymbol{x}_l)$. Here we adopt the log-linear specification

$$g(\boldsymbol{\beta}, \boldsymbol{x}_l) = \exp\left(\sum_h \beta_h \log(x_{h,l}) + \beta_0\right) \ . \tag{10}$$

Equation (10) is equivalent to the standard Cobb-Douglas form $e^{\beta_0} \prod_h x_{h,l}^{\beta_h}$. Hence, if the different effects associated of the variables $\boldsymbol{x}_l = (x_{1,l}, \ldots, x_{H,l})$ were independent from each other, g_l could be interpreted from a probabilistic perspective as the accumulated multiplicative effect of these variables on the probability of localization. As a remark, the Multinomial log-likelihood function (8) is invariant for a rescaling factor, so that the transformation $a_l \to \lambda a_l$ applied to each a_l leaves the likelihood level unchanged. Hence, when estimating the Multinomial model, we set $\beta_0 = 0$ to avoid over specification.

Given specification (10) and considering that q_l is a logarithmic transformation of p_l, the proportionate impact of x_h an n on the overall attractiveness of a location can be estimated with the following marginal elasticities:

$$\frac{\partial q}{\partial \log x_h} = \sum_{l=1}^{L} \frac{\partial q_l}{\partial \log x_{h,l}} = \hat{\beta}_h \frac{\hat{G}}{N + \hat{G}} \ , \tag{11}$$

$$\frac{\partial q}{\partial \log n} = \sum_{l=1}^{L} \frac{\partial q_l}{\partial \log n_l} = \frac{N}{N + \hat{G}} \ , \tag{12}$$

for the Polya model and

$$\frac{\partial q}{\partial \log x_h} = \sum_{l=1}^{L} \frac{\partial q_l}{\partial \log x_{h,l}} = \hat{\beta}_h \ , \tag{13}$$

for the Multinomial model. Equations (11)–(13) measure of how much a proportionate increase in x_h or n affects total attractiveness q, and naturally these measures depend on the distribution of agents and features across locations. The statistical significance of the estimated marginal elasticities (11)–(13) is assessed by bootstrap resampling, which allows to associate a p-score to each estimate. To provide an intuitive illustration of estimate behavior under the Polya and Multinomial model, Appendix A discusses the result of estimation and model selection for three different toy examples representing easily recognizable scenarios.

4.2 Model selection

The invariant distributions (5) and (6) associated respectively to the Polya and Multinomial model represent two alternative scenarios. To evaluate which of these predictions is more compatible with the data, one can rely on a standard procedure of model selection. Specifically, we use the Akaike Information Criterion corrected by finite sample size (AICc) to compare the relative performance of alternative models contemplating a different number of parameters (see Akaike, 1974, Hurvich and Tsai, 1989, for a theoretical fundation). In the present case, for instance, the Polya model has the extra parameter β_0 as compared to the Multinomial. Formally,

$$\mathrm{AICc} = 2k - 2\ln(\mathcal{L}) + \frac{2k(k+1)}{L-k-1} \ , \tag{14}$$

where L is the sample size (i.e. the number of counties in our case), k is the number of parameters in the model, and $\mathcal{L}$ is the maximized value of the likelihood function. Between two alternative models, the one with a lower AICc is to be preferred, as it dissipates less information about the underlying data generating process. In this sense, the AICc prizes goodness of fit via $\mathcal{L}$ and penalizes the number of parameters through k.

As a remark, it is worth noting that the AICc will not generally lead to select the Multinomial over the Polya model only when localized externalities are nearly null. Also intermediate estimates of $\partial q/\partial \log n$ may be associated to Polya AICc values that exceed their Multinomial counterparts. In this sense, our criterion of model selection seems rather severe in admitting the superiority of a scenario that contemplates localized externalities. In Appendix A, we provide an example that further illustrates this point (see especially Figures A1e–A1f and the related discussion). Once a model is selected, one can also assess its absolute performance via standard pseudo-R^2 measures.

5 Estimating the drivers of inventor location

The H regressors that enter $g(\boldsymbol{\beta}, \boldsymbol{x}_l)$ ultimately define the intrinsic attractiveness of counties. Besides the obvious limitations regarding data availability for the time period under scrutiny, there are four considerations to be made about the various estimations presented in the following sections and the related variables summarized in Table 1.

First, all control variables are pre-determined relative to the time period at which the spatial distribution of patent inventors is observed. This serves to mitigate concerns of simultaneity between the location choices of inventors and those of other economic agents, which could otherwise bias the estimates.

Second, we insert all controls that manage to diminish the magnitude of $\partial q/\partial \log n$, so that the resulting estimates of localized externalities is as conservative as possible given the available data. In particular, we devote special attention to accounting for the spatial distribution of production. Additionally, we also take in consideration other key factors such as the specific role of high-quality inventors or the transportation costs involved in the patent application process.

Third, the choice of controls can also partly serve to characterize the type of externality captured by $\partial q/\partial \log n$. In particular, Scitovsky (1954) traces a distinction between pecuniary and technological externalities, as a way to capture the role of market exchanges in the unfolding of interdependicies among economic agents. In turn, Storper (1995) argues for an especially important role of "untraded interdependencies" in shaping the economic success of regions. The present analysis will then try to provide some interpretation in this regard based on the specific type of patent inventors that characterized the British industrial revolution as well as on the descriptors that shape the intrinsic attractiveness of counties.

Fourth, the relevant control variables necessarily differ between estimations operating at the aggregate level and at industry level. In particular, when looking at the aggregate pool of patent inventors, the county population size can be a sufficiently accurate proxy of aggregate employment. By contrast, industry-level estimation have to rely on industry-level employment, which in turn constrains our analysis of the engines and textiles industries to the time period 1842–1850.

Given these considerations, in the following sections we focus our presentation only on those specifications of $g(\boldsymbol{\beta}, \boldsymbol{x}_l)$ involving regressors whose estimated marginal elasticity reaches at least a 95% confidence level in at least some of the estimations at stake (such regressors are summarized in Table 1). We test also other specifications of $g(\boldsymbol{\beta}, \boldsymbol{x}_l)$ including more variables, but their marginal elasticities turn out to have a lower statistical significance and their removal does not significantly affect the results. For this reason, we streamline the exposition by presenting only the more minimal specifications of $g(\boldsymbol{\beta}, \boldsymbol{x}_l)$.

5.1 Aggregate estimations by sub-period

Our first application of the empirical approach presented in Section 4 consists in estimating the marginal elasticities associated to the various determinants of inventor location across the sub-periods in which the industrial revolution is traditionally subdivided. In doing so, one is essentially running an empirical exercise in comparative statics. Each sub-period is interpreted as a separate realization of the spatial equilibrium described in Section 3, so that different underlying parameters are generally expected to be governing such an equilibrium across sub-periods. The objective is then to trace how the resulting marginal elasticities have evolved over time.

In each estimation, the variable to be predicted at the county level is the number of resident patent inventors that obtained a patent during the sub-period under scrutiny. In this respect, it is worth noting that sub-periods must avoid being too short, considering that the minimum length of protection guaranteed by a patent was 14 years and inventions could well retain economic value beyond the length of protection. For this reason, we stick to the periodization that is traditionally used to study the industrial revolution. This facilitates also the comparability of our results with the existing literature. On the other hand, all control variables are pre-determined to minimize concerns of simultaneity: namely, if the sub-period under scrutiny runs between years $[t, t+k]$, the related control variables will date in year $t-1$. In the case of high-quality inventors, however, we consider all occurrences from the previous sub-period: for instance, if the sub-period of interest is 1776-1800, the high-quality patents that enter the controls are those that were granted between 1750–1775.

In our sub-period estimations, we control for the underlying geography of production by means of four separate variables, which are meant to capture different facets of the possible spatial relations between inventive and productive activities. First, we use population size by county to proxy the spatial distribution of aggregate production. While allowing to investigate also the early phases of industrialization, this control is especially important in virtually all of the aggregate estimations that we run. Second, we account specifically for the localization of manufacturing—as opposed to aggregate—production with a dummy variable that identifies counties which hosted historically known manufacturing centers. The main reason to control specifically for the spatial distribution of manufacturing is that a typical inventor of the industrial revolution possibly had to rely on a variety of intermediate inputs that were provided by auxiliary manufacturing activities. In this respect, the manufacturing dummy constitutes an admittedly imperfect variable, which will in fact undergo some refinement in the industry-level estimations of Section 5.2. Third, we further qualify the spatial distribution of aggregate production by measuring the extent to which population is spatially concentrated within each county, so as to characterize the typical depth of local labor markets. Deeper local labor markets indicate a higher potential for labor specialization, which in turn could favor the achievement of relevant innovations. To capture this possible effect, we include as a regressor the Herfindahl–Hirschman index of city concentration. Finally, we control for the energetic potential of counties as measured by local access to coal. While being tightly linked to the early development of specific energy-intensive activities such as iron smelting and metallurgy, the local availability of coal serves more broadly to characterize counties according to how much their energy base could support an expansion of production.

Besides production, we control for the access to high-quality patent inventors. Doing so helps to keep our empirical implementation aligned with the theoretical model on which estimation is based. Specifically, the term bn_l in equation (1) defines a role for the sheer *scale* of local interactions among inventors, while being silent about any effect related to the *quality* of such interactions. In this sense, controlling for the spatial distribution of high-quality inventors tends to remove from the estimate of $\partial q/\partial \log n$ the potentially confounding effect of patent quality. While facilitating the interpretation of our estimates according to the theoretical model in Section 3, this strategy can also serve to assess the specific influence of breakthrough innovations in shaping the spatial distribution of inventive activities. In principle, such an influence could materialize in different and possibly complementary ways. On the one hand, if knowledge spillovers were to have a role, superstars might contribute more valuable knowledge to the common local pool. On the other hand, the occurrence of a technological breakthrough may leave a legacy to its place of invention by fostering an upgrade in the skills and competences of the local labor force. For instance, Meisenzahl and Mokyr (2012) point to the importance of "tweakers" and "implementers" in favoring innovation during the industrial revolution. More broadly, then, the location of high-quality inventors is likely to closely reflect the spatial distribution of upper-tail human capital, which Mokyr (2016, 2018) regards as a key driver of innovation during the industrial revolution.

Finally, the distance from London controls for the transportation costs that inventors had to endure in order to carry out a patent application at the only office available in the country. Generally, these costs could be driven both by the sheer duration of travel to London and by the need to cover the same route multiple times, given that the application procedure normally lasted several months. For example, in reporting Samuel Taylor's diary details from when he was attending his patent petition, Bottomley (2014b, p. 44) notices how "it took Taylor five days to travel from Manchester to London on horseback" and "five months [...] until the patent was eventually sealed in February 1723".

5.1.1 Results by sub-period

With these caveats, the results of our sub-period estimations are reported in Tables 3–4 and can be summarized as follows.

First, the Multinomial model outperforms the Polya up to the mid-18th century, and then the reverse becomes true. More precisely, Tables 3–4 show that the AICc values associated to the Multinomial model are lower than their Polya counterpart for the sub-period 1700–1750, while Polya AICc values fall below Multinomial ones from the mid-18th century onward. Hence, localized externalities became too strong to go statistically unnoticed once the industrial revolution entered in its core phase. So, if "after about 1760 a wave of gadgets swept over England" (Ashton, 1955, p. 42), this seems to have co-occurred with the emergence of a qualitatively "new" determinant of inventor location. In this sense, the new era of sustained innovation and growth that characterized the industrial revolution appears to be tightly linked with a

Table 3: Marginal elasticities by sub-period with the Multinomial model.

Driver	1700–1750	1751–1775	1776–1800	1801–1825	1826–1850
Localized externalities	–	–	–	–	–
Pre-determined Covariates					
Population	$1.30e{+}00^{**}$	$7.06e{-}01^{**}$	$9.72e{-}01^{**}$	$9.43e{-}01^{**}$	$1.16e{+}00^{**}$
Manufacturing dummy	$3.39e{-}01^{**}$	$1.14e{-}01$	$2.64e{-}01$	$7.06e{-}01^{*}$	$9.26e{-}01^{**}$
City concentration (HHI)	$1.61e{-}01^{*}$	$2.20e{-}01^{*}$	$1.39e{-}01^{*}$	$9.18e{-}02^{**}$	$5.80e{-}02^{*}$
Access to coal	$3.12e{-}01^{*}$	$3.82e{-}01^{**}$	$-7.59e{-}02$	$-1.07e{-}01$	$4.32e{-}02$
Access to high-quality inventors	$-1.49e{-}01$	$1.32e{-}01$	$3.42e{-}01^{**}$	$1.96e{-}01$	$7.51e{-}02$
Distance from London	$-7.58e{-}01^{**}$	$-8.00e{-}01^{**}$	$-3.59e{-}01$	$-2.61e{-}01$	$-3.29e{-}01$
AICc	156.64	325.20	523.42	810.74	1669.25
McFadden adj. pseudo-R^2	0.85	0.87	0.91	0.93	0.95
Number of counties	85	85	85	85	85
Spatial auto-correlation of standardized residuals					
Moran's I	0.06	-0.03	0.04	-0.01	-0.12
(p-value)	(0.15)	(0.58)	(0.22)	(0.52)	(0.95)
Geary's C	0.92	1.02	0.96	1.10	1.19
(p-value)	(0.24)	(0.62)	(0.34)	(0.81)	(0.96)

Note: Marginal elasticities are computed with equation (13) and their statistical significance is assessed by bootstrap resampling. The symbol ** indicates a 99% confidence level, while * indicates a 95% confidence level. Spatial auto-correlation statistics are computed using the queen neighborhood criterion to define spatial weights and the related p-value is computed via Monte Carlo permutations.

structural change in the underlying geography of inventive activities.

Second, the size of local aggregate production as captured by county population size remains the primary component of intrinsic attractiveness throughout the whole time period 1700–1850. In particular, the magnitude of the marginal elasticity associated to population size generally exceeds those of other control variables according to both Multinomial and Polya estimates in Tables 3–4. Nonetheless, once localized externalities are accounted for with the Polya model, the resulting pull of population on the location of inventive activities is found to weaken over time according to the estimates reported in Table 4. A similar pattern emerges also for the other descriptors of aggregate production. Conversely, localized externalities grow stronger over the same period of time, which suggests an increasing relevance for spatially-bounded interactions among inventors.

Third, high-quality inventors turn out to have a punctual impact on the attractiveness of their county of residence. In particular, the marginal elasticity estimate associated to the local availability of high-quality inventors does not reach the 95% confidence threshold for most sub-periods under scrutiny, according to both the Multinomial and Polya estimates in Tables 3–4. The only clear exception to this rule is represented by sub-period 1776–1800. All else being equal, during these final years of the 18th century inventors were more likely to site in counties that had previously hosted high-quality inventors. With respect to this result, one should notice that some of the most famous inventors of the Industrial Revolution such as James Watt, John Wilkinson, James Hargraves, and Richard Arkwright indeed obtained their

Table 4: Marginal elasticities by sub-period with the Polya model.

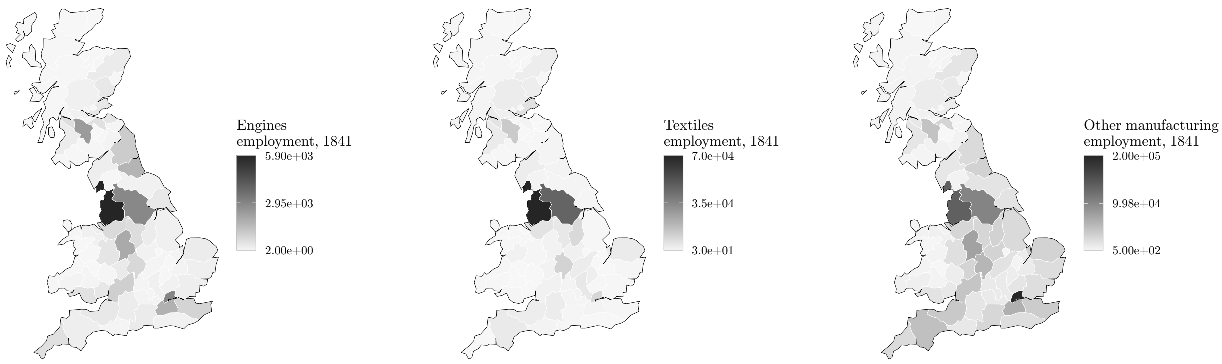
Driver	1700–1750	1751–1775	1776–1800	1801–1825	1826–1850
Localized externalities	1.60e−05**	7.19e−01**	8.09e−01**	8.97e−01**	9.60e−01**
Pre-determined Covariates					
Population	1.30e+00**	2.23e−01**	1.49e−01**	8.95e−02**	2.98e−02**
Manufacturing dummy	3.39e−01**	2.07e−02	4.05e−02	6.37e−02*	1.75e−02**
City concentration (HHI)	1.61e−01*	4.88e−02*	3.54e−02**	9.04e−03*	1.77e−03
Access to coal	3.12e−01*	1.23e−01**	7.49e−03	3.83e−03	6.99e−03
Access to high-quality inventors	−1.49e−01	4.96e−02*	6.19e−02**	1.97e−02	2.58e−02
Distance from London	−7.58e−01**	−2.24e−01**	−7.95e−02	−4.07e−02**	−7.16e−04
AICc	158.64	258.11	379.57	464.49	642.19
McFadden adj. pseudo-R^2	0.85	0.90	0.93	0.96	0.98
Number of counties	85	85	85	85	85
Spatial auto-correlation of standardized residuals					
Moran's I	0.06	−0.02	0.03	−0.01	−0.10
(p-value)	(0.13)	(0.54)	(0.24)	(0.44)	(0.96)
Geary's C	0.92	1.04	1.01	1.10	1.05
(p-value)	(0.21)	(0.65)	(0.56)	(0.80)	(0.69)

Note: Marginal elasticities are computed with equations (11) and (12) and their statistical significance is assessed by bootstrap resampling. The symbol ** indicates a 99% confidence level, while * indicates a 95% confidence level. Spatial auto-correlation statistics are computed using the queen neighborhood criterion to define spatial weights and the related p-value is computed via Monte Carlo permutations.

key patents in the "golden years" between 1769 and 1774, so that their impact is detected in our estimates precisely for sub-period 1776–1800. Hence, these superstar inventors have likely played some role in shaping the emerging geography of innovative activities during the British industrial revolution.

Fourth, Tables 3-4 show that the distance from London gradually became less of an impediment for a county to host patent inventors. Again, both the Multinomial and Polya models detect an overall decline in the magnitude of this effect. While suggesting a generic decrease in transportation costs, this result could be linked more specifically to the development of patent agency starting from the 1770s, which implied that "inventors were no longer required to travel to London to petition for the patent" (Bottomley, 2014b, p. 59).

Finally, Tables 3-4 report also some basic statistics of global spatial auto-correlation for the residuals of each estimation. These statistics serve to gauge the presence of spatial structure in the model errors. If residuals were to display global spatial auto-correlation, the estimated model would likely suffer from the omission of some spatially lagged terms; otherwise, interactions among neighboring spatial units should not be as concerning. For the sub-period estimations in Tables 3-4, residuals do not appear to have a clear spatial structure. Specifically, the null hypothesis Moran's $I = 0$ is never rejected for any level of statistical significance above 87%, thus suggesting that residuals are randomly distributed in space. Similarly, also the null hypothesis Geary's $C = 1$ is never rejected for any level of statistical significance above 80%, pointing again to the fact that neighboring residuals do not tend to be systematically simi-



(a) (b) (c)

Figure 9: Employment in engines, textiles, and other manufacturing industries, 1841.

lar nor dissimilar. In both cases, the lowest p-value is obtained for Multinomial estimates in sub-period 1700–1750. As a remark, p-values for Moran's I and Geary's C are here computed by Monte Carlo permutations, so as to avoid a priori hypotheses on their distribution. Furthermore, the values of I and C in Tables 3-4 are obtained using a queen contiguity criterion to determine spatial weights; however, qualitatively similar results obtain when using other contiguity criteria or distance-based weights.

5.2 Industry-level estimations between 1842–1850

The wave of innovation and productivity growth underlying the British industrial revolution has been alternatively interpreted as an economy-wide phenomenon or as succession of industry-specific shocks (see Crafts, 1985 and McCloskey, 2010 for rather opposing views). At any rate, the mechanization of textiles and the diffusion of steam engines are largely regarded as the two main processes that concretely made for the first wave of industrialization in Great Britain. Unsurprisingly, then, patents in textiles and engines are by far the most frequent in our data set (see Figure 4). Given these considerations, it becomes especially relevant to assess the determinants of localization for patent inventors in these two key industries.

Moving from aggregate to industry level in the analysis of inventor location requires that also the spatial distribution of production be controlled for at a finer disaggregation. As described in Section 2, this constrains our industry-level analysis to the final phase of British industrialization, given that employment in engines and textiles cannot be recovered at the county level before 1841. As a consequence, to avoid simultaneity between the observation to be explained and the related control variables, our industry-level analysis can only focus on inventors that obtained a patent between 1842–1850. In turn, the availability of occupational data allows to characterize the spatial distribution of productive activities more accurately. Instead of controlling for the spatial distribution of aggregate production as in Section 5.1, here our specification of $g(\boldsymbol{\beta}, \boldsymbol{x}_l)$ can separately account for various types of industry-level employment. For instance, when looking at the spatial distribution of engines inventors, we can control for county-level employment in engines, but also for employment in the other key industry (i.e. textiles), or for employment in other manufacturing activities. This allows to investigate the various types of local relations that inventors could have with productive activities within and beyond their own industry of application.

Relatedly, an industry-level assessment of inventor location can help to better characterize the local interactions among inventors that are responsible for external economies of scale. More precisely, one can estimate the Polya model at industry level controlling contemporaneously for the spatial distribution of production and for the location of inventors in other industries. Therefore, the resulting estimate of $\partial q/\partial \log n$ would primarily capture direct interactions accruing locally among inventors in the same industry. For this reason, here we include in the specification of $g(\boldsymbol{\beta}, \boldsymbol{x}_l)$ also the access to inventors who obtained a patent outside of the industry under scrutiny. Given that also these variables need to be pre-determined relative to the observation of interest, we construct them considering patents that were granted between 1816–1841, so as to maintain a consistent time lag with the specification used in Section 5.1.

5.2.1 Results for engines and textiles between 1842–1850

While further confirming some of the results discussed in the previous sections, moving from aggregate to industry level helps to have further insights on the determinants of inventor location during the final phase of the British industrial revolution.

In line with the aggregate sub-period estimates discussed in Section 5.1, the industry-level AICc values reported in Table 5 point to the superiority of the Polya model relative to the Multinomial during the time period 1842–1850. Indeed, such a conclusion holds even when the determinants of location for patent inventors are separately estimated for two key industries such as engines and textiles, rather than in the aggregate. Similarly, the impact of localized externalities at industry level proves to be sizable relative to other drivers, while also being comparable in magnitude to its aggregate counterpart in Table 4. In parallel, Table 5 confirms the close link between the location of inventive and productive activities. Both Multinomial and Polya estimates indicate that industry employment is the strongest determinant of a county's intrinsic attractiveness in the eye of patent inventors. In particular, under the Polya model, the impact of industry employment on the attractiveness of a county as measured in terms its marginal elasticity is in the same order of magnitude as localized externalities.

Notably, the estimates in Table 5 suggest that the spatial distribution of inventors at industry level is predominantly affected by within-industry determinants. On the one hand, the marginal elasticities associated to the location of employment in the other key industry or in other manufacturing activities remain systematically below the 95% confidence threshold, regardless of the particular model taken in consideration. We have also run some alternative estimations in which "Other manufacturing employment" was further fragmented, so as to possibly detect the role of more specific sectors such as metallurgy, precision instruments, or scientific professionals. However, none of these alternative specifications delivered qualitatively different results. On the other hand, also the marginal elasticity associated to the location of non high-quality inventors in other industries remains systematically below the 95% confidence threshold, irrespective of the model at stake. In parallel, we find some industry-specific role for high-quality inventors in textiles. Relatedly, also the effect of localized externalities captured

Table 5: Marginal elasticities at industry level, 1842–1850.

Driver	Multinomial Engines	Multinomial Textiles	Polya Engines	Polya Textiles
Localized externalities	–	–	$6.24e{-}01^{**}$	$7.57e{-}01^{**}$
Covariates in 1841				
Industry employment	$6.76e{-}01^{**}$	$6.85e{-}01^{**}$	$2.21e{-}01^{**}$	$1.81e{-}01^{**}$
Employment in other key industry	$-1.23e{-}02$	$2.25e{-}01$	$7.15e{-}03$	$9.67e{-}02$
Other manufacturing employment	$1.86e{-}01$	$4.52e{-}01$	$1.16e{-}01$	$-3.09e{-}02$
City concentration (HHI)	$2.48e{-}03$	$6.05e{-}02$	$-4.20e{-}03$	$2.18e{-}02$
Access to coal	$-2.37e{-}01$	$1.76e{-}01$	$-9.19e{-}02$	$4.72e{-}02$
Access to industry high-quality inventors	$2.18e{-}01$	$3.80e{-}01^{*}$	$9.77e{-}02$	$8.43e{-}02^{**}$
Access to other high-quality inventors	$4.31e{-}01$	$-1.24e{-}01$	$1.65e{-}01$	$-1.84e{-}02$
Access to other non high-quality inventors	$3.17e{-}01$	$-5.43e{-}01$	$9.32e{-}02$	$-7.68e{-}02$
Distance from London	$2.08e{-}01$	$-7.95e{-}01$	$8.19e{-}02$	$-1.80e{-}01$
AICc	260.96	306.27	236.20	230.58
McFadden adj. pseudo-R^2	0.89	0.89	0.90	0.92
Number of counties	85	85	85	85
Spatial auto-correlation of standardized residuals				
Moran's I	0.03	-0.03	-0.05	-0.07
(p-value)	(0.25)	(0.57)	(0.76)	(0.84)
Geary's C	1.07	1.04	1.10	1.15
(p-value)	(0.72)	(0.64)	(0.80)	(0.88)

Note: Marginal elasticities are computed with equations (11), (12), and (13) and their statistical significance is assessed by bootstrap resampling. The symbol ** indicates a 99% confidence level, while * indicates a 95% confidence level. Spatial auto-correlation statistics are computed using the queen neighborhood criterion to define spatial weights and the related p-value is computed via Monte Carlo permutations.

by the Polya estimates in Table 5 should be understood as having a clear within-industry connotation. That is, the controls in use imply that the estimates of $\partial q/\partial \log n$ reflect primarily the effect of local interactions among inventors operating within the same technological base.

Finally, the residuals for each of the estimations in Table 5 do not appear to be spatially auto-correlated at any meaningful level of statistical significance. Hence, our industry-level estimates do not seem to be worryingly affected by the omission of spatial lags. Again, the values of I and C in Table 5 are computed with a queen contiguity criterion to determine spatial weights, but other contiguity criteria or distance-based weights lead to the same conclusions about the spatial auto-correlation of residuals.

6 Conclusion

This paper has provided a quantitative assessment of the determinants of location for patent inventors during the British industrial revolution. As expected, the location of inventors turns out to be closely intertwined with the spatial distribution of productive activities throughout the time period under scrutiny. Starting from the mid-18th century, however, the geography of innovation underwent a structural change together with the whole economy. As new machines and products came to be invented with increasing frequency, the pull of productive activities lost some grip in driving inventor location. Conversely, another force gained ground. Namely, local interactions among inventors became key to the siting of inventive activities, ultimately making the location of inventors strongly interdependent. These localized externalities were possibly weak enough to go virtually unnoticed in the pre-industrial world, but they ended up becoming an indispensable ingredient in the new era of innovation and sustained growth.

The present paper has also characterized such externalities in at least two further respects. First, we systematically controlled for the local access to high-quality inventors, so that the resulting estimates of localized externalities tend to capture a pure scale effect. Second, we documented how local interactions among inventors had an industry-specific reach. In this sense, the spatial distribution of patent inventors during the British industrial revolution seems to be primarily shaped by "mass interaction" with other technologically similar peers, although also "select relations" with high-quality inventors had a non negligible role in particular toward the end of the 18th century. Overall, this evidence is consistent with a dynamic of discovery that is dense of incremental improvements resulting from trial and error procedures, in a context where productive knowledge is imperfectly transferable. As long as the details of an innovative attempt spread in space according to some decay function, inventors may well face an advantage to co-localize with their peers in order to learn more rapidly from their mistakes and to start from where the others left off. In this case, the need to learn from others brings inventors together, and in turn the probability to finally achieve a functioning invention is maximized where inventors are more numerous. Co-localization would then be a way to appropriate this type of externality.

Indeed, such narrative resonates well with the historical framework in which our analysis is centered. On the one hand, and in spite of the high application costs involved, the patents that were granted in England during the 18th and 19th century were often far from being fully functioning innovations. In this sense, their prototypical nature lends itself well to be interpreted as a step in a broader process of discovery based on trials, errors, and small incremental improvements. On the other hand, several historical accounts document how—even during the "Age of Enlightenment"—inventive activities were often based on tacit know-how and empirical rules of thumb, which are normally regarded as being more spatially sticky relative to codified information. For instance, Fox (2009) discusses several revealing examples in which detailed engineering drawings and documentation were insufficient to ensure the installation of machinery. Similarly, Macdonald (1979) shows that the diffusion of innovative agricultural techniques took place essentially by means of learning from neighbors, despite the growing availability of published information. Relatedly, MacLeod and Nuvolari (2009) highlight the important role of apprenticeships networks in transmitting skills within the mechanical engineering industry. Finally, a number of case studies illustrate how some communities of inventors engaged in deliberate forms of knowledge sharing (see Allen, 1983, Berg, 1993, Cookson, 1997, and Nuvolari, 2004). The present paper suggests that these various instances of technological externalities were the rule rather than the exception, and they had a key role in shaping the new geography of innovation that came to take form during the British industrial revolution.

Appendix A

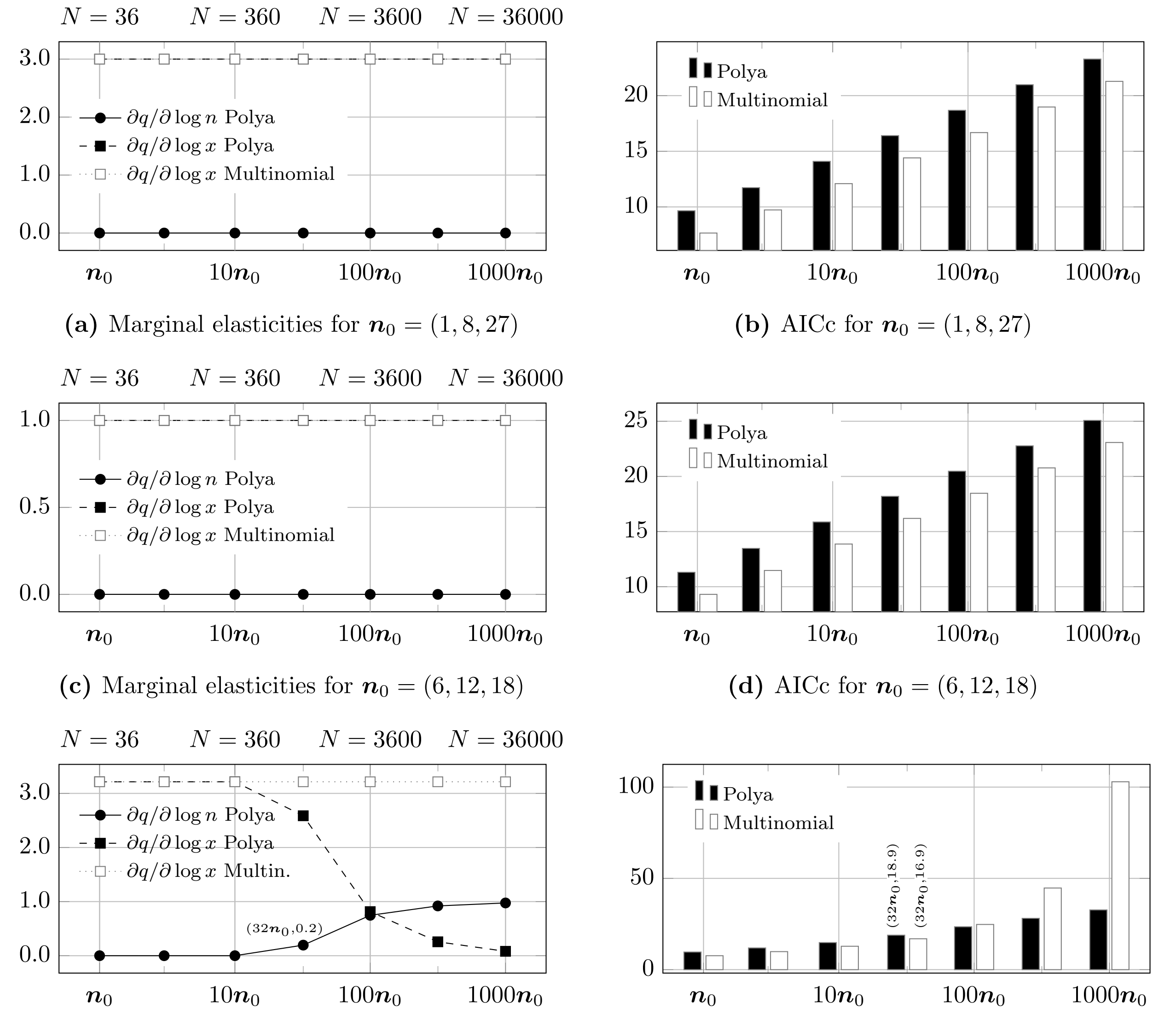
This section is meant to provide an illustration of estimate behavior under the Polya and Multinomial model, while also giving some insights about the related procedure of model selection. In particular, we estimate marginal elasticities under the Polya and Multinomial models for three different toy examples that represent qualitatively distinct and easily recognizable scenarios.

Consider a simplified setup with $L = 3$ alternatives, a single vector of intrinsic features $\boldsymbol{x} = (1, 2, 3)$, and some "initial" occupancy vector $\boldsymbol{n}_0 = (n_{1,0}, n_{2,0}, n_{3,0})$ with N_0 total agents. Given that the Polya marginal elasticities (11) and (12) are a function of the total number of agents N, we estimate $\log \pi(\boldsymbol{\beta} \mid \boldsymbol{x})$ for transformations of $\boldsymbol{n}_0$ that increase the total number of agents while preserving their distribution across alternatives. Specifically, we consider transformations of the type $\boldsymbol{n}_0 \to \lambda \boldsymbol{n}_0$ with $\lambda \in \mathbb{N}^*$, so that each resulting occupancy vector $\boldsymbol{n} = \lambda \boldsymbol{n}_0$ has $N = \lambda N_0$ agents following the same distribution across alternatives as $\boldsymbol{n}_0$. Figure A1 summarizes the outcomes of this exercise for three different specifications of the initial occupancy vector $\boldsymbol{n}_0$.

The results shown in Figures A1a–A1b are obtained using $\boldsymbol{n}_0 = (1, 8, 27)$ as an initial occupancy vector, so that in this case $n_l = \lambda x_l^3$ for all l. It follows that the marginal elasticity $\partial q/\partial \log x$ as estimated under the Multinomial model is centered around 3 for any value of N, precisely because external economies of scale are not supposed to have a role in the Multinomial case (see Figure A1a). Even when estimating the Polya model, however, the estimates of $\partial q/\partial \log x$ are almost perfectly overlapping with their Multinomial counterparts, while the estimates of $\partial q/\partial \log n$ are approximately at 0. This result occurs because x_l suffices to account entirely for n_l across all alternatives (via the function $n_l = x_l^3$), so that no significant room is left for localized externalities to have an impact. As further discussed in Section 4.2, this also induces to select the Multinomial over the Polya model, as shown by the Akaike Information Criterion corrected by finite sample size (AICc) in Figure A1b.

A qualitatively similar situation occurs if x_l entirely accounts for n_l across all locations but via a different function. For instance, the results shown in Figures A1c–A1d are obtained using $\boldsymbol{n}_0 = (6, 12, 18)$ as an initial occupancy vector, so that in this case $n_l = \lambda 6 x_l$ for all l. As a consequence, the resulting Multinomial estimates of $\partial q/\partial \log x$ now come to be centered around 1, and again their Polya counterparts are closely overlapping while $\partial q/\partial \log n$ is approximately at 0. Also in this case, the AIC values reported in Figure A1d indicate that the Multinomial model is to be selected over the Polya model.

Instead, the results shown in Figures A1e–A1f are based on the initial occupancy vector $\boldsymbol{n}_0 = (1, 7, 28)$, so that in this last case n_l cannot be expressed as a common function of x_l for all l. Moreover, $\boldsymbol{n}_0$ can also be regarded as an outcome of the transformation $(1, 8, 27) \to (1, 7, 28)$: that is, starting from the Multinomial configuration underlying Figures A1a–A1b, an agent moves from a less populated location to the most populated one. In this sense, localized externalities might be expected to gain weight, and indeed the results reported in Figures A1e–A1f confirm such an intuition. For sufficiently large values of N, the estimates of $\partial q/\partial \log n$ are now significantly greater than 0, while the Polya model outperforms the Multinomial according to their corresponding AICc values. Notably, however, localized externalities are irrelevant as long as the total number of agents N remains too small for interactions to matter. Moreover, even



(a) Marginal elasticities for $\boldsymbol{n}_0 = (1, 8, 27)$

(b) AICc for $\boldsymbol{n}_0 = (1, 8, 27)$

(c) Marginal elasticities for $\boldsymbol{n}_0 = (6, 12, 18)$

(d) AICc for $\boldsymbol{n}_0 = (6, 12, 18)$

(e) Marginal elasticities for $\boldsymbol{n}_0 = (1, 7, 28)$

(f) AICc for $\boldsymbol{n}_0 = (1, 7, 28)$

Figure A1: Estimation of a toy example with $L = 3$ and $\boldsymbol{x} = (1, 2, 3)$.

Left panels: Each dot corresponds to an estimated marginal elasticity obtained with $\boldsymbol{x} = (1, 2, 3)$ and some occupancy vector $\boldsymbol{n} = \lambda \boldsymbol{n}_0$, where $\lambda \in \{1, 3, 10, 32, 100, 316, 1000\}$. Each row of plots is associated to a different specifications of $\boldsymbol{n}_0$, as indicated in the sub-figure title. For each value of λ, the total number of agents $N = \lambda N_0$ is constant across the various specifications of $\boldsymbol{n}_0$, namely $N \in \{36, 108, 360, 1152, 3600, 11376, 36000\}$.

Right panels: Each group of bars shows the AICc values associated to the Polya and Multinomial model for the corresponding occupancy vector $\boldsymbol{n} = \lambda \boldsymbol{n}_0$, given the common intrinsic features $\boldsymbol{x} = (1, 2, 3)$. AICc values are to be compared only within a single group of bars, so as to select either the Polya or the Multinomial model for a specific $\boldsymbol{n}$ and $\boldsymbol{x}$. The model associated to a lower AICc value is to be preferred.

when the estimates of $\partial q / \partial \log n$ are non negligible in magnitude and statistically significant, the procedure of model selection may still favor the Multinomial over the Polya model. For instance, in In Figures A1e–A1f, we show that $\partial q / \partial \log n \approx 0.2$ for $\lambda = 32$ and $N = 1152$, yet the corresponding AICc values lead to select the Multinomial model over the Polya (with an AICc value of 16.9 and 18.9 respectively).

Among other things, the results illustrated in Figure A1 should also clarify that the link between Polya marginal elasticity estimates (11)–(12) and the total number of agents N is far from being mechanical. Namely, while an economy with more agents may potentially offer greater scope for localized externalities, having a more numerous economy will not mechanically increase the estimate of $\partial q/\partial \log n$. In fact, as shown in Figures A1a–A1d, the role of intrinsic attractiveness may well be so predominant as to annihilate entirely the effect of localized externalities even for large values of N. In this sense, the magnitude of the Polya marginal elasticities in equations (11) and (12) crucially depend on the *distribution* of agents and features across locations, so that the value of N per se is insufficient to reveal anything about the magnitudes of $\partial q/\partial \log x_h$ and $\partial q/\partial \log n$ under the Polya model.

References

Agrawal, A., Kapur, D., and McHale, J. (2008). How do spatial and social proximity influence knowledge flows? Evidence from patent data. *Journal of Urban Economics*, 64(2):258–269.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723.

Allen, R. (1983). Collective invention. *Journal of Economic Behavior & Organization*, 4(1):1–24.

Almeida, P. and Kogut, B. (1999). Localization of knowledge and the mobility of engineers in regional networks. *Management Science*, 45(7):905–917.

Andrews, M. and Whalley, A. (2021). 150 years of the geography of innovation. *Regional Science and Urban Economics*. In press.

Arrow, K. (1962). The economic implications of learning by doing. *Review of Economic Studies*, 29(3):155–173.

Arthur, W. (1990). 'Silicon Valley' locational clusters: When do increasing returns imply monopoly? *Mathematical Social Sciences*, 19(3):235–251.

Arthur, W., Ermoliev, Y., and Kaniovski, Y. (1987). Path-dependent processes and the emergence of macro-structure. *European Journal of Operational Research*, 30(3):294–303.

Ashton, T. (1955). *An economic history of England: The eighteenth century*. Methuen, London.

Audretsch, B. (1998). Agglomeration and the location of innovative activity. *Oxford Review of Economic Policy*, 14(2):18–29.

Audretsch, D. and Feldman, M. (1996). R&D Spillovers and the geography of innovation and Production. *American Economic Review*, 86(3):630–640.

Azoulay, P., Graff Zivin, J., and Wang, J. (2010). Superstar extinction. *Quarterly Journal of Economics*, 125(2):549–589.

Bairoch, P., Batou, J., and Chevre, P. (1988). *La population des villes européennes de 800 à 1850: banque de données et analyse sommaire des résultats*. Librairie Droz, Paris.

Balderston, T. (2010). The economics of abundance: Coal and cotton in Lancashire and the world. *Economic History Review*, 63(3):569–590.

Barry, J. (2008). South-West. In Clark, P., editor, *The Cambridge urban history of Britain*, volume 2. Cambridge University Press.

Berg, M. (1993). Small producer capitalism in eighteenth-century England. *Business History*, 35(1):17–39.

Berg, M. (2005). *The age of manufactures, 1700-1820: Industry, innovation and work in Britain*. Routledge.

Boschma, R. (2005). Proximity and innovation: A critical assessment. *Regional Studies*, 39(1):61–74.

Bottazzi, G., Dosi, G., Fagiolo, G., and Secchi, A. (2007). Modeling industrial evolution in geographical space. *Journal of Economic Geography*, 7(5):651–672.

Bottazzi, G. and Gragnolati, U. (2015). Cities and clusters: economy-wide and sector specific effects in corporate location. *Regional Studies*, 49(1):113–129.

Bottazzi, G., Gragnolati, U., and Vanni, F. (2017). Non-linear externalities in firm localization. *Regional Studies*, 51(8):1138–1150.

Bottazzi, G. and Secchi, A. (2007). Repeated choices under dynamic externalities. LEM Working Paper Series.

Bottomley, S. (2014a). Patenting in England, Scotland and Ireland during the industrial revolution, 1700–1852. *Explorations in Economic History*, 54:48–63.

Bottomley, S. (2014b). *The British patent system during the Industrial Revolution 1700–1852: From privilege to property*, volume 28. Cambridge University Press.

Breschi, S. and Lissoni, F. (2001a). Knowledge spillovers and local innovation systems: A critical survey. *Industrial and corporate change*, 10(4):975–1005.

Breschi, S. and Lissoni, F. (2001b). Localised knowledge spillovers vs. innovative milieux: Knowledge "tacitness" reconsidered. *Industrial and Corporate Change*, 10(4):975–1005.

Breschi, S. and Lissoni, F. (2009). Mobility of skilled workers and co-invention networks: An anatomy of localized knowledge flows. *Journal of Economic Geography*, 9(4):439–468.

Broadberry, S. and Marrison, A. (2002). External economies of scale in the Lancashire cotton industry, 1900–1950. *Economic History Review*, 55(1):51–77.

Cambridge Group for the History of Population and Social Structure (CAMPOPa) (2021). The occupational structure of Britain 1379–1911. Last consulted in April 2021.

Cambridge Group for the History of Population and Social Structure (CAMPOPb) (2021). 1851 England and Wales ancient counties. Last consulted in April 2021.

Cambridge Group for the History of Population and Social Structure (CAMPOPc) (2021). 1851 Scotland Ancient Counties GIS. Last consulted in April 2021.

Carlino, G. and Kerr, W. (2015). Agglomeration and innovation. In *Handbook of Regional and Urban Economics*, volume 5, pages 349–404. Elsevier.

Census Report (1841). Occupation abstract. London.

Clark, G. and Jacks, D. (2007). Coal and the industrial revolution, 1700–1869. *European Review of Economic History*, 11(1):39–72.

Cookson, G. (1997). Family firms and business networks: Textile engineering in Yorkshire, 1780–1830. *Business History*, 39(1):1–20.

Crafts, N. (1985). *British economic growth during the industrial revolution*. Oxford University Press.

Crafts, N. and Wolf, N. (2014). The location of the British cotton textiles industry in 1838: A quantitative analysis. *Journal of Economic History*, 74(04):1103–1139.

Crescenzi, R., Iammarino, S., Ioramashvili, C., Rodríguez-Pose, A., and Storper, M. (2020). The geography of innovation and development: Global spread and local hotspots. London School of Economics. Geography and Environment Discussion Paper Series, 4.

Crescenzi, R., Nathan, M., and Rodríguez-Pose, A. (2016). Do inventors talk to strangers? On proximity and collaborative knowledge creation. *Research Policy*, 45(1):177–194.

Dosi, G., Ermoliev, Y., and Kaniovski, Y. (1994). Generalized urn schemes and technological dynamics. *Journal of Mathematical Economics*, 23(1):1–19.

Dutton, H. (1984). *The patent system and inventive activity during the Industrial Revolution, 1750-1852*. Manchester University Press.

Ellison, G. and Glaeser, E. (1997). Geographic concentration in US manufacturing industries: A dartboard approach. *Journal of Political Economy*, 105(5):889–927.

Feldman, M. (1999). The new economics of innovation, spillovers and agglomeration: A review of empirical studies. *Economics of Innovation and New Technology*, 8(1):5–25.

Flinn, M. (1984). *The history of the British coal industry, 1700–1830: The Industrial Revolution*, volume 2. Clarendon Press, Oxford. With the assistance of Stoker, D.

Fox, C. (2009). *The arts of industry in the Age of Enlightenment*. Yale University Press for The Paul Mellon Centre for Studies in British Art.

Garibaldi, U. and Scalas, E. (2010). *Finitary probabilistic methods in Econophysics*. Cambridge University Press, Cambridge.

Giuliani, E. (2007). The selective nature of knowledge networks in clusters: Evidence from the wine industry. *Journal of Economic Geography*, 7(2):139–168.

Giuri, P. and Mariani, M. (2013). When distance disappears: Inventors, education, and the locus of knowledge spillovers. *Review of Economics and Statistics*, 95(2):449–463.

Hanlon, W. and Miscio, A. (2017). Agglomeration: A long-run panel data approach. *Journal of Urban Economics*, 99:1–14.

Henderson, R., Jaffe, A., and Trajtenberg, M. (2005). Patent citations and the geography of knowledge spillovers: A reassessment: Comment. *American Economic Review*, 95(1):461–464.

House of Commons Parliamentary Papers (HOC) (1831). Comparative account of the population of Great Britain in the years 1801, 1811, 1821, 1831. London.

House of Commons Parliamentary Papers (HOC) (1851). Return of the population of the several counties of England, Wales, and Scotland. Years 1831, 1841, 1851. London.

Hurvich, C. and Tsai, C. (1989). Regression and time series model selection in small samples. *Biometrika*, 76(2):297.

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

Kelly, M., Mokyr, J., and Ó Gráda, C. (2022). The mechanics of the industrial revolution. *Journal of Political Economy*. forthcoming.

Kerr, W. (2010). Breakthrough inventions and migrating clusters of innovation. *Journal of Urban Economics*, 67(1):46–60.

Khan, B. (2005). *The democratization of invention: patents and copyrights in American economic development, 1790-1920*. Cambridge University Press.

Kuznets, S. (1966). *Modern economic growth*. Yale University Press, New Haven. Conn.

Lamoreaux, N. and Sokoloff, K. (2000). The geography of invention in the American glass industry, 1870–1925. *Journal of Economic History*, 60(3):700–729.

Leunig, T. (2003). A British industrial success: Productivity in the Lancashire and New England cotton spinning industries a century ago. *Economic History Review*, pages 90–117.

Macdonald, S. (1979). The diffusion of knowledge among Northumberland farmers, 1780-1815. *Agricultural History Review*, 27(1):30–39.

MacLeod, C. (1988). *Inventing the industrial revolution: The English patent system, 1660-1800*. Cambridge University Press, Cambridge.

MacLeod, C. and Nuvolari, A. (2009). "Glorious Times": The emergence of mechanical engineering in early industrial Britain, c. 1700-1850. *Brussels Economic Review*, 52.

Marshall, A. (1890). *Principles of economics*. McMillan, London.

McCloskey, D. (2010). *Bourgeois dignity: Why economics can't explain the modern world*. University of Chicago Press.

Meisenzahl, R. and Mokyr, J. (2012). The rate and direction of invention in the British Industrial Revolution. In Lerner, J. and Stern, S., editors, *The rate and direction of inventive activity revisited*, pages 443–482. University of Chicago Press.

Mokyr, J. (2016). *A culture of growth*. Princeton University Press.

Mokyr, J. (2018). Bottom-up or top-down? The origins of the Industrial Revolution. *Journal of Institutional Economics*, 14(6):1003–1024.

Moser, P. (2011). Do patents weaken the localization of innovations? Evidence from world's fairs. *Journal of Economic History*, 71(2):363–382.

Nuvolari, A. (2004). Collective invention during the British Industrial Revolution: The case of the Cornish pumping engine. *Cambridge Journal of Economics*, 28(3):347–363.

Nuvolari, A. and Tartari, V. (2011). Bennet Woodcroft and the value of English patents, 1617–1841. *Explorations in Economic History*, 48(1):97–115.

Nuvolari, A., Tartari, V., and Tranchero, M. (2021). Patterns of innovation during the industrial revolution: A reappraisal using a composite indicator of patent quality. *Explorations in Economic History*. In press.

Nuvolari, A., Verspagen, B., and Von Tunzelmann, N. (2011). The early diffusion of the steam engine in Britain, 1700–1800: A reappraisal. *Cliometrica*, 5(3):291–321.

Oettl, A. (2012). Reconceptualizing stars: Scientist helpfulness and peer performance. *Management Science*, 58(6):1122–1140.

Rodrigue, J.-P. (2020). *The geography of transport systems*. Routledge, third edition.

Scitovsky, T. (1954). Two concepts of external economies. *Journal of Political Economy*, 62(2):143–151.

Shaw-Taylor, L. and Wrigley, E. (2014). Occupational structure and population change. In Floud, R., Humphries, J., and Johnson, P., editors, *The Cambridge economic history of modern Britain*, volume 1: 1700–1870.

Storper, M. (1995). The resurgence of regional economies, ten years later: The region as a nexus of untraded interdependencies. *European Urban and Regional Studies*, 2(3):191–221.

Thompson, P. and Fox-Kean, M. (2005). Patent citations and the geography of knowledge spillovers: A reassessment. *American Economic Review*, 95(1):450–460.

Trinder, B. (2008). Industrialising towns 1700–1840. In Clark, P., editor, *The Cambridge urban history of Britain*, volume 2. Cambridge University Press.

Tubiana, M., Miguelez, E., and Moreno Serrano, R. (2020). In knowledge we trust: Learning-by-interacting and the productivity of inventors. *AQR–Working Papers, 2020, AQR20/05.*

Waldinger, F. (2012). Peer effects in science: Evidence from the dismissal of scientists in Nazi Germany. *Review of Economic Studies*, 79(2):838–861.

Webster, A. (1952). An account of the number of people in Scotland in the year 1755. In Kyd, J., editor, *Scottish population statistics including Webster's analysis of population, 1755.* Constable, Edinburgh.

Woodcroft, B. (1854). *Titles of patents of invention chronologically arranged, 1617–1852.* GE Eyre & W. Spottiswoode, London.

Woodcroft, B. (1862). *Reference index of English patents of invention, 1617-1852.* GE Eyre & W. Spottiswoode, London.

Wrigley, A. (2010). The PST system of classifying occupations. Last consulted in April 2021.

Wrigley, E. (1990). *Continuity, chance and change: The character of the industrial revolution in England.* Cambridge University Press.

Wrigley, E. (2007). English county populations in the later eighteenth century. *Economic History Review*, 60(1):35–69.

Wrigley, E. (2009). Rickman revisited: The population growth rates of English counties in the early modern period. *Economic History Review*, 62(3):711–735.

Zacchia, P. (2018). Benefiting colleagues but not the city: Localized effects from the relocation of superstar inventors. *Research Policy*, 47(5):992–1005.

Zucker, L. and Darby, M. (1996). Star scientists and institutional transformation: Patterns of invention and innovation in the formation of the biotechnology industry. *Proceedings of the National Academy of Sciences*, 93(23):12709–12716.