



**Universitat
Pompeu Fabra**
Barcelona

Department
of Economics and Business

Economics Working Paper Series

Working Paper No. 1945

**The republic of letters and the rise of the
west***

**Andrea Melillo, Luigi Pascali, Mounu Prem, Francesca
Asja Trento**

May 2026

The Republic of Letters and the Rise of the West*

Andrea Melillo, Luigi Pascali, Mounu Prem, Francesca Asja Trento

May 28, 2026

Abstract

What role did the Republic of Letters play in Europe's transition to sustained innovation? We combine a corpus of digitized correspondence within the Republic of Letters with European aristocratic genealogies, historical postal routes, and a database of notable individuals to trace the diffusion and consequences of Enlightenment correspondence between 1600 and 1850. We first show that the Republic spread, in part, through aristocratic kinship networks: aristocrats connected to already participating peers entered earlier, and their probability of entry declined sharply with network distance. To isolate a causal channel, we exploit changes in postal distances along pre-existing kinship paths to already-inoculated aristocrats, while controlling directly for local postal access. We then aggregate this variation to European grid cells and estimate the effect of exposure to the Republic on the rise of applied science, innovation, and economic activity. Cells instrumented into the Republic experienced a near-doubling in applied scientists and inventors roughly three decades after first contact, with no pre-trends and effects concentrated in scientific and technical correspondence rather than religion or philosophy. These findings suggest that the Republic of Letters helped reshape the geography of innovation before industrialization.

Keywords: Enlightenment, Republic of Letters, The Great Divergence, useful knowledge, innovation, social networks, economic growth

JEL Codes: N13, O31, O33, Z13

*Melillo: LSE; Email: A.Melillo@lse.ac.uk. Pascali: LUISS Guido Carli, EIEF, Barcelona School of Economics (BSE) and Pompeu Fabra University, and affiliated with IPEG and CEPR; Email: luigi.pascali@upf.edu. Prem: EIEF, and affiliated with IZA and CEPR; Email: francisco.munoz@eief.it. Trento: EIEF; Email: ftrento@redphd.it. Luigi Pascali acknowledges financial support from the Spanish Agencia Estatal de Investigación (AEI), through the Severo Ochoa Programme for Centres of Excellence in RD (Barcelona School of Economics CEX2024-001476-S).

1 Introduction

Between the late seventeenth and the early nineteenth century, a narrow strip of western Europe began to produce engineers, instrument makers, agronomists, hydrographers, and inventors at a rate without precedent elsewhere in the world. Newton's *Principia* appeared in 1687; Diderot and d'Alembert's *Encyclopédie* began circulating in 1751; James Watt patented his separate condenser in 1769; the *Exposition des produits de l'industrie française* opened in Paris in 1798. Behind these landmarks lay a quieter transformation: the emergence of a transnational community of scholars who corresponded across confessional and political borders and gradually came to share a common intellectual framework associated with the Enlightenment. Contemporaries called this community the *Republic of Letters*. This paper asks whether, and how, the Republic of Letters reshaped the human capital of the places it reached—whether the arrival of an Enlightenment correspondent in a town in Lower Saxony or Languedoc increased the local production of the applied scientists and inventors who would eventually carry Europe across the threshold of sustained growth.

The question matters because one of the leading cultural accounts of the Industrial Revolution rests on it. Mokyr (2017) argues that the transnational Republic of Letters transformed Europe's emerging "culture of growth" into a functioning pan-European market for useful knowledge, and provided a key coordinating mechanism behind what he calls the *Great Enrichment*. In this account, the cultural breakthrough preceded the technological one: a new belief that the natural world was intelligible, and that useful knowledge could be accumulated, scrutinized, and shared, made eighteenth-century macroinventions thinkable in the first place and initiated a self-reinforcing interaction between macro- and microinvention. The view is contested. Allen (2009) emphasizes high wages and cheap coal; McCloskey (2010) the rehabilitation of bourgeois values; Landes (1998), Clark (2007), Jacob (2014), and Acemoglu et al. (2005) alternative cultural, institutional, and demographic configurations. The historical claim at stake is that a small

transnational elite—roughly twelve hundred active scholars in 1680, perhaps ten times that number a century later (Dibon, 1978)—rewired the production of useful knowledge across an entire continent. To the best of our knowledge, the empirical counterpart of that claim, traced through the interpersonal correspondence network itself and at the scale of Europe, is so far missing.

This paper makes two empirical claims. The first is historical: the Republic of Letters spread, in part, through the European aristocratic kinship network. The second is economic: exposure to the Republic of Letters contributed to the local rise of applied scientists, inventors, and commercial entrepreneurs.

To study these questions, we combine four new sources of data. The first is the *Electronic Enlightenment* corpus (Bodleian Libraries, Oxford University Press), a scholarly digital edition of 81,765 letters exchanged among 10,497 correspondents—including Voltaire, Locke, Rousseau, Newton, Hume, Franklin, and Bentham—between 1600 and 1850. The second is a European aristocratic kinship graph reconstructed from *The Peerage* and the *Europäische Stammtafeln*, covering almost two hundred thousand individuals between 1550 and 1850. The third is a new digitization of eleven pan-European postal-network snapshots assembled from seventeenth- to nineteenth-century postal atlases and itineraries. The fourth is the cross-verified database of notable individuals assembled by Laouénan et al. (2022), which reconciles biographical records across Wikipedia and Wikidata. We restrict the sample to 323,556 individuals born or deceased in Europe between 1550 and 1850 and use occupational classifications to trace the evolution of applied scientists, inventors, and commercial entrepreneurs across European grid cells over time.

The European aristocratic kinship network contributed to the spread of the Republic of Letters. The Republic did not float free of European elite society. Aristocrats were heavily over-represented among its correspondents, patrons, salon hosts, and scientific amateurs, and the kinship ties binding the European nobility together also carried the introductions, recommendations, and personal acquaintances through which Enlightenment correspon-

dence first spread.

To study this process, we construct an aristocrat-decade panel and model the spread of Enlightenment ideas across the kinship network as an epidemiological process. The “inoculation event” for an aristocrat is the receipt of a first letter from a member of the Republic of Letters—an event we observe directly in the *Electronic Enlightenment* data. Using an event-study framework, we show that aristocrats whose network neighbors had already been inoculated entered the Republic earlier themselves: exposure to Enlightenment correspondents through the kinship network predicts the timing at which a given aristocrat first joins the Republic of Letters. Consistent with this pattern, panel regressions with aristocrat and decade fixed effects show that the probability of inoculation declines sharply with network distance to the nearest already-inoculated aristocrat.

The central identification challenge is that aristocratic connections are not randomly assigned. To isolate a causal channel, we exploit the interaction between the pre-existing topology of the aristocratic network and the gradual expansion of the European postal network. For each aristocrat, we hold fixed at the year of her birth both the structure of the kinship graph and the identity of the nearest already-inoculated aristocrat, and construct a time-varying measure of exposure equal to the postal distance separating the two along the kinship network. Changes in postal routes generate plausibly exogenous variation in this distance over time. The resulting variation does not simply capture generic postal accessibility: it arises from the interaction between postal expansion and the pre-existing topology of the aristocratic network and infection status of connected nodes. Aristocrat fixed effects absorb features of the kinship network and the identity of the nearest already-inoculated aristocrat at birth, while flexible controls for local postal accessibility account for the endogenous expansion of the postal system.

Our results suggest that the aristocratic network played a meaningful role in the spread of the Republic of Letters. At the individual level, an uninfected aristocrat who moves from a kinship-and-postal distance of one thousand kilometres to one hundred kilome-

tres from the nearest already-corresponding aristocrat experiences roughly a twenty-fold increase in the probability of entering the Republic in a decade, relative to a baseline infection rate of about one in a thousand per decade. The instrumented effect is nearly identical to the descriptive one. Moreover, we show that this effect operates through proximity along the aristocratic network itself, rather than through generic postal accessibility or overall network centrality.

We do not claim that aristocratic networks were the only conduit through which Enlightenment ideas spread. Scholarly travel, printed books, learned societies, and systems of local patronage also transmitted useful knowledge across Europe, and aristocrats were by no means the only participants in the Republic of Letters: in our data, no more than one fifth of correspondents belong to the aristocracy. Our claim is narrower. Decade by decade, the aristocratic kinship network constituted one important channel through which the Republic expanded, and one that leaves a measurable footprint in the data. In the next step of the analysis, we use this footprint to construct plausibly exogenous variation in local exposure to Enlightenment ideas and the Republic of Letters.

Exposure to the Republic of Letters and the rise of applied science, innovation, and economic activity.

We next ask whether the arrival of Enlightenment correspondence in a cell increased the local production of *applied scientists*—engineers, cartographers, agronomists, hydrographers, and other measurement-driven professions that Mokyr (2002) identifies as carriers of useful knowledge (Section 4). To do so, we construct a cell-by-decade panel covering 1,453 European grid cells between 1600 and 1850. An event study estimated using the Borusyak et al. (2024) imputation estimator shows monotonically increasing post-treatment effects in the decades after a cell receives its first letter from the Republic of Letters.

Establishing causality requires variation in a cell’s exposure to the Republic that is unrelated to its underlying intellectual trajectory. Motivated by our first set of results,

we aggregate to the cell-decade level the same plausibly exogenous variation in exposure that we construct at the aristocrat-decade level. A cell becomes more exposed when postal infrastructure reaches deeper into the kinship paths connecting its local aristocrats to those already corresponding with the Republic. Our 2SLS estimates suggest that cells instrumented into the Republic of Letters subsequently experienced a near-doubling in the local stock of applied scientists roughly three decades after first contact. The effect emerges with a two-to-three-decade lag, precisely as a slow human-capital channel would predict: we observe no effect in the decade of first contact and no anticipatory increase beforehand. The results remain robust to controlling for local postal accessibility over time.

We further show descriptively that the applied-scientist effect is concentrated in correspondence on technical and scientific topics—the subjects most directly related to useful knowledge—with no detectable effect from correspondence on religion or philosophy. This result is a within-cell association rather than a separate causal estimate, and we interpret it as suggestive evidence on the underlying mechanism.

In the final part of the paper, we examine the effects of the Republic of Letters on broader economic outcomes. Our 2SLS estimates suggest that cells instrumented into the Republic subsequently experienced a near-doubling in the number of *inventors* three decades later, with effects that are precisely estimated and already detectable at earlier horizons. *Merchants, entrepreneurs, and bankers* move in the same direction and display broadly similar long-run magnitudes, although the entrepreneurial response is noisier and becomes statistically distinguishable only after three decades.

We interpret these findings as the empirical micro-counterpart of the macro-historical mechanism proposed by Mokyr (2017). In this view, the Republic of Letters functioned as the institutional carrier of Europe’s emerging culture of growth—a pan-European market for useful knowledge that supplied a continental audience, a form of peer review by correspondence, and a transnational reputational structure for individuals producing

applicable knowledge about the natural world. Our evidence is consistent with this interpretation: where the network spread, the local stock of applied scientists and inventors rose persistently in the decades following first contact. The asymmetry between sharply identified effects on inventors and noisier effects on entrepreneurs is itself consistent with the distinction between institutions that supported the production of useful knowledge and those that translated such knowledge into firms and markets—property rights, credit arrangements, and commercial organization—which the Republic of Letters did not directly provide.

Contribution to the literature.

This paper contributes to several literatures.

Most directly, it provides, to our knowledge, the first pan-European empirical test of the thesis, advanced most forcefully by Mokyr (2002; 2005; 2017), that interpersonal exchange within the Republic of Letters helped generate the useful knowledge underlying Western economic growth. The closest existing papers are Squicciarini and Voigtländer (2015) and Cinnirella et al. (2025), which study the role of upper-tail human capital in different European contexts during the First and Second Industrial Revolutions. Squicciarini and Voigtländer (2015) show that upper-tail human capital—proxied by subscription density to Diderot’s *Encyclopédie* (1751–72)—predicted city-level industrialization in France during the First Industrial Revolution. Cinnirella et al. (2025) show that German economic societies—institutional bodies founded from the mid-eighteenth century to promote useful knowledge—played an important role in regional technological progress during the Second Industrial Revolution. Our paper complements this literature by studying the earlier, informal, and transnational correspondence network that preceded, and in many cases inspired, both the *Encyclopédie* and the German economic societies.¹

We also relate to work on institutions of knowledge diffusion and upper-tail human capital. Cantoni and Yuchtman (2014) show that medieval universities shaped long-run

¹See also Hornung (2014), who studies the long-run productivity effects of high-skilled immigration using the Huguenot migration into Prussia as a natural experiment.

economic outcomes through the diffusion of human capital within Europe. Our paper studies a distinct but complementary mechanism: the spread of useful knowledge through a decentralized interpersonal correspondence network operating largely outside the formal academic establishment.

Our work also contributes to the literature on the economic consequences of communication technologies, from the printing press (Dittmar, 2011; ?; Rubin, 2014) to the telegraph (Steinwender, 2018). Most closely, we relate to a recent literature on postal networks, communication costs, and economic development. Sasaki (2024) studies the effects of the state-sponsored postal service on city growth in France between 1500 and 1850. Hanlon et al. (2022) exploit the introduction of the Uniform Penny Post in Britain in 1840 to estimate the effect of lower communication costs on scientific knowledge flows, applied research, and technological development. In a recent working paper, Matteo Cervellati (2026) study the long-run effects of a 1635 postal-system reform in England on the content of correspondence within the Republic of Letters and innovation. Our paper is complementary but distinct. Existing work uses exogenous changes in postal infrastructure or postal prices to estimate the effects of communication costs on correspondence and economic activity. We use the postal network differently: not as the object of interest, but as one component to construct a time-varying instrument for local exposure to the Republic of Letters along with the aristocratic kinship network and the infected status of connected nodes. In all main specifications, we control directly for local access to the postal system: so we are not able to estimate the direct impact of the postal services. We do construct, however, plausibly exogenous variation in exposure to the Republic of Letters. This allows us to isolate the specific mechanism emphasized by Mokyr (2017)—the diffusion of useful knowledge through a transnational intellectual network—from other channels through which postal access may affect development, including state capacity (Acemoglu et al., 2016) and financial integration (Heblich and Trew, 2019).

The rest of the paper is organized as follows. Section 2 describes the data. Section 3

documents how the Republic of Letters spread along the aristocratic-and-postal network. Section 4 estimates the effect of Enlightenment correspondence on the local production of applied scientists. Section 5 extends the analysis to inventors and to merchants, entrepreneurs, and bankers. Some concluding remarks close the paper.

2 Data

2.1 Enlightenment correspondence

Our main correspondence data come from the *Electronic Enlightenment* (henceforth EE), a scholarly digital edition assembled by the Bodleian Libraries and Oxford University Press.² The current release contains 81,765 letters exchanged among 10,497 correspondents between 1600 and 1850.³ The corpus combines scholarly editions of individual letter collections—among them those of Voltaire, Locke, Rousseau, Newton, Hume, Franklin, and Bentham—with smaller collections of less prominent correspondents. For each letter, we observe the full text, the sender, the recipient, the date, and the place of writing and the place of receipt. We geocode sender and recipient locations to the city level and retain the 51,519 letters in which the sender’s or recipient’s location falls within continental Europe or the British Isles.⁴

For robustness, we repeat the analysis using a second corpus, *Early Modern Letters Online* (EMLO), maintained by the Cultures of Knowledge project at the University of Oxford.⁵ EMLO records approximately 180,000 letters exchanged between 1500 and 1800 and is broader than EE in two respects: it covers correspondence on commercial, family,

²See <https://www.e-enlightenment.com>.

³We drop an additional 21 letters dated 1490–1510: they predate the period of interest and are too sparse to be informative.

⁴Starting from the raw “written from” and “written to” fields, we extract the set of unique location strings and query the Google Maps Geocoding API to obtain latitude and longitude coordinates. We then assign each geocoded point to its corresponding European LAU (Local Administrative Unit) polygon via a spatial join.

⁵See <https://emlo.bodleian.ox.ac.uk>.

and political topics in addition to scholarly exchange, and it extends back to the early sixteenth century, predating the conventional onset of the Enlightenment. These two features motivate two placebo tests: the broader topical coverage allows us to verify that our effects are driven by correspondence on science and technology rather than by other types of exchange, and the earlier temporal coverage allows us to verify that the estimated effects are absent prior to the Enlightenment.

We translate each letter into English and classify it into one of seven topical categories: Religion, Politics, Science, Technology, Philosophy, Art, and Others. We adopt an LLM-as-a-judge framework in which a secondary model independently audits the primary annotator’s classifications, flagging potentially erroneous entries with structured justifications; a third model then re-examines the flagged cases in light of these justifications to produce a final classification.⁶ Figure A.1 reports the word clouds generated from these classifications. For instance, the Technology cloud features terms such as *mechanical*, *invention*, and *steam engine*—a finding consistent with the centrality of that technology to eighteenth-century industrialization.

The correspondence data are subject to an important selection process. The EE editorial team chose which surviving correspondence to digitize, with a stated focus on figures central to the long eighteenth century’s intellectual exchange. This curatorial selection is precisely what makes EE a useful proxy for the Republic of Letters and, at the same time, what limits its representativeness as a sample of surviving early modern correspondence more broadly. We therefore interpret our estimates as informative about the causal effect of exposure to the canonical Enlightenment network, rather than about early modern correspondence at large.

⁶See Appendix A.1 for details on the translation procedure, the models employed, and the prompts used.

2.2 The aristocratic network

We combine two data sources to construct a dataset of European aristocratic networks spanning 1550 to 1850. The first is the *Peerage* project (Lundy, 2020), which compiles data on more than 720,000 European nobles.⁷ The *Peerage* database records nobles' dates of birth, death, and marriage.⁸ In addition, *Peerage* contains information that allows us to link each noble to his or her parents, siblings, spouses, and children. By way of illustration, Figure A.5 shows the *Peerage* entry for Wolfgang, Duke of Braunschweig-Grubenhagen. On the basis of this information, we reconstruct the European nobility network at any point in time. The resulting panel, at the aristocrat-decade level, comprises 175,377 individuals born between 1550 and 1850.

To supplement these data, we digitize information from the *Europäische Stammtafeln* (European Family Trees, Schwennicke (1998)), which covers 1,379 aristocratic family trees. Figure A.6 presents a sample family tree for the Dukes of Braunschweig-Grubenhagen, and Figure A.7 shows a detailed individual entry. This entry corresponds to Wolfgang, the same individual featured in the *Peerage* example above. The *Stammtafeln* provide valuable additional information on the *locations* of births, deaths, and marriages, which allows us to link individual nobles to cities.⁹ Where available, we further supplement these data with location information from *Peerage* itself.

2.3 Notable individuals

Our primary outcome data are drawn from the cross-verified database of notable people compiled by Laouénan et al. (2022). The database records 2.29 million individuals from across the world between 3500BC and 2018AD, identified through a cross-verification procedure that reconciles biographical entries across Wikipedia (in twenty-seven lan-

⁷Accessed 20 January 2020.

⁸In some instances, only the date of birth *or* death is available. In such cases, we impute the missing date using the median age of all nobles known to be alive at the same time as the noble in question; see Appendix A.2 for details.

⁹Appendix A.2 describes the geocoding procedure.

guage editions) and Wikidata. For each individual, we observe the name, birth year and place, death year and place, and a primary occupation reported as a text string in the language of the underlying Wikipedia entry.¹⁰ To the best of our knowledge, this database constitutes the largest harmonized prosopography currently available for the early modern period, and it has rapidly become a benchmark source in the economics of innovation (?).

We restrict the sample to individuals whose birthplace or place of death falls within continental Europe or the British Isles and whose birth year and primary occupation are both observed. These restrictions yield 323,556 individuals, whom we treat as the universe of *later-remembered* Europeans active between the late Middle Ages and the mid-nineteenth century.¹¹

The raw occupation field contains 3,982 distinct strings, reported in the language of each Wikipedia edition (e.g. *mathematician*, *Mathematiker*, *matemático*, *ingénieur*). We classify each string into one of eight occupational categories spanning the arts, humanities, and sciences, three categories of disruptive professions, and a residual category.

Figures A.2 and A.3 display the time profile of the resulting groups, with the conventional Enlightenment window (1650–1789) shaded for reference. The series align closely with established historical chronology, which we read as suggestive evidence that the underlying source captures real shifts in the composition of educated activity. Figure A.2 plots the scholarly professions. Panel A shows that the share of *visual artists* peaks around 1500, the conventional end of the Italian Renaissance, while the performing and literary

¹⁰Using the latitude and longitude coordinates of birth and death locations reported in the raw data, we assign each location to its corresponding LAU polygon via a spatial join.

¹¹Admittedly, the label “notable” identifies people who became famous enough for a posthumous Wikipedia entry, not people who were innovative in their own time, and the two need not coincide: an obscure provincial inventor of a textile loom may matter for the industrial revolution without ever entering the corpus, while a sufficiently colorful court poet may be in the corpus without ever advancing the technological frontier. Coverage also varies across centuries—the twelfth century is recorded more sparsely than the eighteenth—because later biographical sources devote more attention to recent figures. Our outcome should therefore be read as a composition measure of the educated stratum that posterity remembered, not as a count of innovative activity. The cell-by-decade fixed effects we introduce in Section 4 absorb any secular drift in coverage at the cell level; identification rests on *within-cell*, *between-decade* movements in composition, not on the raw cross-sectional level.

arts rise steadily into the early nineteenth century. Panel B shows a sharp spike in *religious studies* during the sixteenth century, coinciding with the Reformation, and a subsequent resurgence of letters, history, and philology after 1600. Panel C shows that the hard sciences (formal and natural sciences) accelerate from the fifteenth century onward: formal sciences follow a relatively flat trajectory from the mid-sixteenth century, while natural sciences grow steadily, consistent with the onset of the Scientific Revolution. Panel D, on applied sciences, anchors our main empirical exercise: unlike any other scholarly profession, the share of applied scientists only begins to accelerate with the onset of the Enlightenment, following centuries of stagnation and a preceding century of pronounced decline. Figure A.4 shows that the applied-science category is dominated by terms we associate with practical, measurement-driven knowledge—*engineer, cartographer, horticulturist, agronomist, meteorologist, topographer, hydrographer*. These are precisely the kinds of occupations that Mokyr (2002) identifies as the carriers of useful knowledge during the long eighteenth century.

Figure A.3 displays three categories that we read as proxies for disruptive activity: inventors, businesspeople and entrepreneurs, and political activists and revolutionaries. All three grow as a share of notable people during the Enlightenment, and three patterns merit attention. First, the share of inventors and entrepreneurs peaks around 1850, but the rise begins clearly in the early eighteenth century—well before the conventional onset of the Industrial Revolution. We read this as consistent with the view that the human-capital shifts underlying the Industrial Revolution were already accumulating during the Enlightenment, rather than emerging suddenly after 1760. Second, the share of political activists and revolutionaries is essentially flat before 1700 and then jumps in the final decades of the eighteenth century, peaking in the 1780s and falling thereafter—a pattern consistent with the Atlantic Revolutionary period and, in particular, with the years immediately preceding 1789. Third, the share of businesspeople and entrepreneurs moves in tandem with that of inventors from the onset of the Enlightenment onward.

Using these data, we construct a cell-by-decade panel by assigning each individual to each decade in which he or she was alive and at least twenty years of age, and to the cell in which he or she was born.¹²

2.4 Postal network

We draw on multiple sources to construct snapshots of the European postal network from 1630 to 1820. Most are historical atlases that document the postal network from the eighteenth and nineteenth centuries and that we digitize manually. For earlier decades, we supplement these atlases with a corpus of nine printed postal itineraries digitized by the EMDigIt project (Midura (2025), Midura (2021)).¹³ The Online Appendix A lists all itineraries and atlas maps used in the analysis.

From each itinerary or map, we construct an undirected graph $G_t = (V_t, E_t)$. Vertices V_t are post relays: the stations at which horses were changed along official postal routes, each identified by a place name and a geocoded coordinate. We call these the *observed* nodes. Edges E_t connect adjacent relays along documented route segments; because postal communication is assumed equally costly in both directions, edges are undirected. Edge weights are haversine (great-circle) distances in kilometres between segment endpoints.¹⁴ By way of illustration, Figure 1 shows a raw map from 1814 alongside the corresponding digitized graph.

Once we have digitized graphs with sufficient coverage of Europe in a given year,

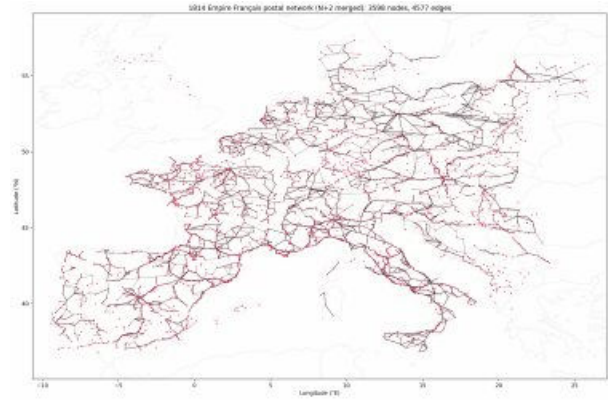
¹²We drop the first two decades of each individual’s life, since few people make notable contributions to science, business, or politics before age twenty. In robustness checks, results are unchanged when individuals are instead assigned to the cell in which they died.

¹³We thank Rachel Midura for kindly sharing the data with us.

¹⁴Some relays are visible on the map but cannot be matched to a gazetteer entry, and some route segments pass through stretches in which intermediate stations are illegible. To preserve route geometry and connectivity, we insert *synthetic* vertices in three cases. (a) A relay marker is present on the map but its label is illegible: we recover coordinates by fitting a linear transformation to known anchor points on the same map sheet. (b) A label is legible but cannot be matched to a gazetteer: we geocode it using the same linear transformation. (c) The relay’s position is interpolated linearly along the corridor between two known neighbours. Synthetic vertices participate fully in the graph: they are incident to edges and lie on shortest paths.



(a) Raw map.



(b) Digitized network.

Figure 1: Example of a digitized postal map: raw map and digitized version. *Notes:* Raw map: Tardieu (1814), *Carte des Routes de Postes de l'Empire Français, du Royaume d'Italie et de la Confédération du Rhin*.

we compute all-pairs shortest paths on G_t using Dijkstra's algorithm, with edge weights equal to haversine distances in kilometres. The spatial unit used in the analysis is a 0.25×0.25 latitude–longitude grid restricted to European land (1,453 cells). The bilateral postal distance between cells A and B is the sum of three components: (i) the distance from the cell- A centroid to its snap relay (the nearest relay node to the cell centroid); (ii) the shortest path through the network between the two snap relays; (iii) the distance from the cell- B snap relay to the cell- B centroid.

3 The Spread of the Enlightenment

3.1 Descriptive evidence

We begin by examining whether proximity to the Republic of Letters within the aristocratic network predicts participation in it. We construct an individual-decade panel from

The Peerage, covering 175,377 aristocrats born between 1550 and 1850.¹⁵ For each aristocrat i in decade t , we define Y_{it} as an indicator equal to one if i sends or receives at least one letter from a member of the Republic of Letters. In total, 7,851 aristocrats have an infected relative within twelve hops (steps in the kinship network) at some point during their lifetime, and 668 aristocrats are ever infected.

Before turning to a regression framework, we document three descriptive patterns. First, the aggregate share of infected aristocrats follows an S-shaped trajectory over time (Figure A.8). Infection rates remain close to zero before 1700, accelerate through the mid-eighteenth century, and peak around the 1790s—a pattern consistent with a diffusion process that reaches saturation as the pool of susceptible aristocrats shrinks. Second, the timing of infection varies systematically with initial network distance. Aristocrats initially close to the Enlightenment network (1–3 hops) are infected earlier and at higher rates than those farther away (Figure A.9). This ordering is consistent with a contagion process that spreads outward from the initial hubs. Third, the transition probabilities by distance bin reveal a steep monotonic gradient (Figure A.10): among aristocrats uninfected at t , the per-decade transition rate is largest at distance 1 and falls steadily to a near-zero baseline beyond twelve hops. For the regression analysis, we truncate the network distance at twelve hops, beyond which the contagion gradient is essentially flat and statistical power becomes negligible.

These patterns motivate a regression framework that relates infection to network proximity while absorbing individual and time heterogeneity. We define $\text{AristDist}_{i,t}$ as the number of hops separating aristocrat i from the nearest infected aristocrat in the kinship network in decade t . We estimate:

$$Y_{it} = \alpha_i + \gamma_t + \beta \text{InvAristDist}_{i,t-1} + \mathbf{X}'_{it} \boldsymbol{\delta} + \varepsilon_{it}, \quad (1)$$

¹⁵The panel is restricted to aristocrats for whom both birth and death decades are available (whether observed or imputed). Each aristocrat enters the panel two decades after birth—the earliest age at which one could plausibly contribute to scholarly correspondence—and exits in the decade of death. We discuss the implications of this restriction below.

where α_i and γ_t are aristocrat and decade fixed effects, $\text{InvAristDist}_{i,t-1} = 1/\text{AristDist}_{i,t-1}$ is the reciprocal network distance lagged by one decade,¹⁶ and \mathbf{X}_{it} is a vector of time-varying controls including degree centrality and eigenvector centrality in the aristocratic kinship network. Standard errors are clustered at the aristocrat level throughout. The aristocrat fixed effects absorb all time-invariant individual characteristics—dynasty, birth order, country of birth—so β is identified from within-aristocrat variation in network proximity to infected nodes over time. A positive $\hat{\beta}$ indicates that an aristocrat is more likely to become infected in decades in which he or she is closer to an already-infected member of the network.

Table 1 presents the results. Columns (1)–(3) restrict the sample to reachable aristocrats—those within twelve hops of an infected node—while columns (4)–(6) use the full sample, coding unreachable aristocrats as $\text{InvAristDist} = 0$. In column (1), without controls, the coefficient on reciprocal distance is positive and statistically significant at the 5 percent level. To interpret the magnitude of this coefficient: moving from distance 10 ($\text{InvAristDist} = 0.1$) to distance 1 ($\text{InvAristDist} = 1.0$) is associated with a 1.5 percentage point increase in the probability of infection—more than an order of magnitude larger than the baseline infection rate of approximately 0.07 percent per decade in the reachable subsample. In columns (2) and (3), we add degree centrality and eigenvector centrality as controls. Reassuringly, the estimated coefficient on InvAristDist is essentially unchanged, indicating that the distance gradient is not driven by the general connectedness of aristocrats in the kinship network.¹⁷ Columns (4)–(6) expand the sample to include unreachable aristocrats: estimated coefficients are similar to those in columns (1)–(3), but are now esti-

¹⁶The reciprocal transformation avoids dropping the large number of unreachable aristocrats: they enter with $\text{InvAristDist} = 0$. The transformation also captures the nonlinearity visible in the transition probabilities (Figure A.10): the difference between 1 and 2 hops is far more consequential than the difference between 9 and 10 hops. Accordingly, in our benchmark specification, we set $\text{InvAristDist}_{i,t-1} = 0$ for aristocrats with no infected neighbor within 12 hops.

¹⁷Within the reachable subsample, neither degree centrality (0.0013, s.e. 0.0013) nor eigenvector centrality (0.0854, s.e. 0.1256) enters significantly once aristocrat fixed effects are absorbed. In the full sample (columns (4)–(6)), degree centrality is positive and precisely estimated (0.0002, $p < 0.01$), consistent with better-connected aristocrats being more exposed at the extensive margin. Controlling for either measure leaves the distance coefficient unaffected.

mated with considerably greater precision and are statistically significant at the 1 percent level.

Table 1: Network Distance and Probability of Infection

Outcome: Receiving or sending a letter (dummy)						
	(1)	(2)	(3)	(4)	(5)	(6)
	Reachable only			Full sample (unreachable=0)		
[1em] InvAristDist	0.0163** (0.00773)	0.0165** (0.00773)	0.0162** (0.00773)	0.0124*** (0.00197)	0.0124*** (0.00197)	0.0124*** (0.00197)
Degree Centrality		0.00129 (0.00127)			0.000176*** (0.0000561)	
Eigenvector Centrality			0.0854 (0.126)			0.00184 (0.00538)
Aristocrat FE	yes	yes	yes	yes	yes	yes
Decade FE	yes	yes	yes	yes	yes	yes
Observations	10,505	10,505	10,505	972,434	972,434	972,434
Aristocrats	3,580	3,580	3,580	155,724	155,724	155,724
Outcome mean	0.007	0.007	0.007	0.001	0.001	0.001
R2	0.405	0.405	0.405	0.197	0.197	0.197

Note: Dependent variable: infected (=1 if sent/received Enlightenment letter). Cols 1-3: sample restricted to reachable aristocrats (distance 1-12). Cols 4-6: full sample; unreachable aristocrats coded as `inv_lagged_dist = 0`. All specifications include aristocrat and decade fixed effects. Robust standard errors are clustered at aristocrat level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

To document the dynamics of contagion, we estimate a second specification that traces how the probability of infection evolves over the decades following a neighbor’s infection. Specifically, we define $\mathbb{1}[\text{NearbyInf}_{i,t}]$ as an indicator equal to one if aristocrat i has at least one infected aristocrat within 3 hops in decade t .¹⁸ We then estimate the following specification separately for each horizon $h = 1, 2, 3, 4$:

$$Y_{i,t+h} = \alpha_i + \gamma_t + \beta_h \mathbb{1}[\text{NearbyInf}_{i,t}] + \varepsilon_{it}, \quad (2)$$

on the sample of aristocrats who are uninfected in decade t . This is a local projection that asks: conditional on being uninfected in decade t , does having an infected neighbor predict infection h decades later?

¹⁸The 3-hop threshold is motivated by the transition probabilities in Figure A.10, which show that the infection gradient is steepest at distances 1 through 3 and flattens substantially beyond that point.

Figure 2 presents the estimated coefficients $\hat{\beta}_h$. All four are positive and statistically distinguishable from zero at the 1 percent level, and the profile trends upward across horizons. An aristocrat with an infected neighbour within 3 hops is 0.70 percentage points more likely to become infected one decade later ($h = 1$), rising to 1.07 percentage points at $h = 2$, 1.00 percentage points at $h = 3$, and 1.33 percentage points at $h = 4$. To put these magnitudes in context, the raw infection rate among aristocrats *without* a nearby infected neighbour is on the order of 0.07 percent per decade; among those *with* a nearby infected neighbour, the implied rate at $h = 4$ is roughly twenty times higher (Figure A.11). The upward trend in coefficients at longer horizons is consistent with cumulative network exposure: once a neighbour becomes infected, additional aristocrats in the same kinship neighbourhood are also drawn into the Enlightenment network, opening new transmission paths and reinforcing the local contagion process.¹⁹

Taken together, these results document a clear epidemiological pattern: proximity to infected aristocrats in the kinship network predicts the timing and probability of infection, the association is robust to controls for network centrality, and the effect operates dynamically—exposure to a nearby infected neighbour is followed by a sustained increase in the probability of own infection over the subsequent four decades. These patterns, however, document correlation rather than causation. Proximity in the aristocratic network is endogenous: aristocrats who formed more connections to intellectually active peers may have been predisposed to engage with the Enlightenment regardless of network diffusion. The within-aristocrat estimates absorb time-invariant confounders, but cannot address time-varying selection into intellectual circles. We turn to this identification challenge in the next subsection.

¹⁹Because infection is absorbing—once an aristocrat sends or receives an Enlightenment letter, he or she remains “infected” in all subsequent decades—there are no pre-treatment periods to examine. The dynamics in equation (2) are therefore purely forward-looking. The absorbing-state structure mechanically implies that the cumulative probability of infection by horizon h is weakly increasing in h ; the empirical finding is that the *marginal* probability mass accruing at each horizon remains positive throughout the four decades we examine, consistent with sustained rather than one-off intellectual transmission.

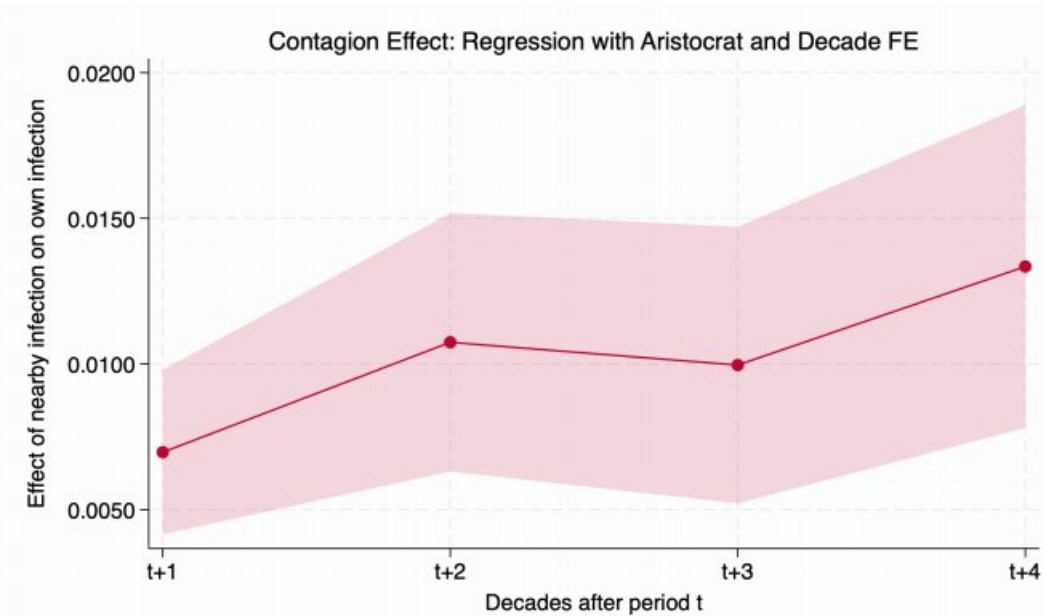


Figure 2: Contagion dynamics: the probability of infection after a neighbor becomes infected

Note: The figure plots coefficients $\hat{\beta}_h$ from equation (2), estimated separately for horizons $h = 1$ through $h = 4$. The dependent variable is an indicator for infection at $t + h$; the regressor is an indicator for having an infected aristocrat within 3 hops at t . Sample restricted to aristocrats uninfected at t . All regressions include aristocrat and decade fixed effects. Standard errors clustered at the aristocrat level; shaded area shows 95% confidence intervals.

3.2 2SLS estimates

To isolate a causal channel, we exploit the expansion of the postal network. We assign to each edge (j, k) in the aristocratic network a weight equal to the postal distance between the cities where j and k reside, measured using the postal network as of decade t . For each aristocrat i , we fix the topology of the aristocratic network as of her year of birth. We then compute the shortest weighted path from aristocrat i to the nearest already-infected aristocrat at the time of her birth, along the edges of this fixed network with postal distances as weights, and denote it $\text{AristPostalDist}_{it}$. This variable changes over time not because the aristocratic network changes—we hold its topology fixed at the aristocrat’s year of birth—but because postal routes expand, reducing the communication cost along existing aristocratic ties. A postal expansion that reduces the communication cost along the

specific kinship ties connecting i to infected aristocrats lowers $\text{AristPostalDist}_{it}$, bringing i closer to infected nodes in terms of postal-weighted network distance.

We use the reciprocal of this variable, $\text{InvAristPostalDist}_{i,t-1} = 1/\text{AristPostalDist}_{i,t-1}$ as an instrument for the reciprocal network distance $\text{InvDist}_{i,t-1}$ in a two-stage least squares (2SLS) framework. The second stage is:

$$Y_{it} = \alpha_i + \gamma_t + \beta \text{InvDist}_{i,t-1} + \delta \text{InvPostalDist}_{c(i),t-1} + \mathbf{X}'_{it} \boldsymbol{\phi} + \varepsilon_{it}, \quad (3)$$

where $\text{InvPostalDist}_{c(i),t-1} = 1/\text{PostalDist}_{c(i),t-1}$ and $\text{PostalDist}_{c(i),t-1}$ is the direct postal distance from city $c(i)$ to the nearest city containing an infected aristocrat in decade $t - 1$, computed on the postal network alone.

The first stage relates network proximity to postal proximity along aristocratic ties:

$$\text{InvDist}_{i,t-1} = \alpha_i + \gamma_t + \pi \text{InvAristPostalDist}_{i,t-1} + \lambda \text{InvPostalDist}_{c(i),t-1} + \mathbf{X}'_{it} \boldsymbol{\psi} + \eta_{it}, \quad (4)$$

The inclusion of $\text{InvPostalDist}_{c(i),t-1}$ as a control is crucial. Controlling for it ensures that the instrument does not simply capture the direct effect of postal access on Enlightenment exposure. What identifies β is the residual variation: a postal expansion that reduces the weighted path along the specific edges of the *pre-existing aristocratic network* connecting i to *pre-existing infected aristocrats*, over and above any direct effect of a city's postal proximity to centers of Enlightenment activity. The impact of the time invariant network and proximity to infection at birth is picked up by the aristocrat fixed effects.

The exclusion restriction requires that changes in $\text{InvAristPostalDist}_{i,t-1}$ —conditional on the aristocratic network structure, the direct postal distance, and aristocrat and decade fixed effects—affect infection only through their impact on postal-weighted network proximity.

Table 2 reports the OLS, first-stage, and 2SLS estimates.²⁰ Column (1) replicates the

²⁰As in Section 3.1, $\text{InvAristDist}_{i,t-1}$ and $\text{InvAristPostalDist}_{i,t-1}$ are set to zero if aristocrat i has no in-

full-sample OLS specification from Table 1, now conditioning on direct postal distance to the nearest infected city; the estimated coefficient on reciprocal network distance is essentially unchanged. The direct postal-distance control $\text{InvPostalDist}_{c(i),t-1}$ enters with a small negative coefficient that is statistically indistinguishable from zero, consistent with the view that what drives individual infection is proximity along the kinship network rather than a city’s general postal connectivity. Column (2) reports the first stage. The instrument $\text{InvAristPostalDist}_{i,t-1}$ enters with a large, positive, and precisely estimated coefficient; the control $\text{InvPostalDist}_{c(i),t-1}$ also enters positively, as expected given that a city’s direct postal distance to the nearest infected city is mechanically related to the postal-weighted distance along its residents’ aristocratic ties. The Kleibergen–Paap first-stage F -statistic is 365.22, well above the Stock–Yogo critical value of 16.38 for 10 percent maximal IV size. Figure A.12 plots the conditional first-stage relationship and confirms that, after partialing out aristocrat and decade fixed effects and direct postal proximity, the residual covariation between postal-weighted aristocratic distance and raw network distance is tight, positive, and approximately linear. Column (3) reports the 2SLS estimate. The coefficient on $\text{InvAristDist}_{i,t-1}$ is essentially indistinguishable from the OLS estimate. Columns (4) to (6) replicate the analysis of columns (1) to (3), adding controls for degree centrality and eigenvector centrality. The inclusion of these controls leaves the OLS and 2SLS estimates unchanged.

These 2SLS results thus confirm the descriptive conclusion of Section 3.1: aristocrats become more likely to engage with the Republic of Letters precisely in the decades when a postal expansion reduces the weighted path along their own kinship ties to an already-infected relative, over and above any direct effect of the city’s postal access. Taken together with the descriptive evidence of Section 3.1, these 2SLS estimates provide our individual-level causal estimate of network-mediated diffusion. We turn next to city-level effects of the Republic of Letters’ expansion.

infected relative within 12 hops.

Table 2: The spread of the Enlightenment: OLS and 2SLS estimates

	Outcome: Receiving or sending a letter (dummy)					
	(1) OLS	(2) First stage	(3) 2SLS	(4) OLS	(5) First stage	(6) 2SLS
InvAristDist	0.013*** (0.003)		0.011* (0.005)	0.013*** (0.003)		0.011* (0.005)
InvAristPostalDist		16.039*** (0.839)			16.039*** (0.839)	
InvPostalDist	-0.036 (0.077)	4.437*** (0.667)	-0.003 (0.086)	-0.037 (0.077)	4.435*** (0.667)	-0.003 (0.086)
Degree Centrality				0.000** (0.000)	0.000*** (0.000)	0.000** (0.000)
Eig. Centrality				0.001 (0.005)	0.019** (0.007)	0.001 (0.005)
Aristocrat FE	yes	yes	yes	yes	yes	yes
Decade FE	yes	yes	yes	yes	yes	yes
Sample	Europe	Europe	Europe	Europe	Europe	Europe
Observations	972,434	972,434	972,434	972,434	972,434	972,434
Cells	155,724	155,724	155,724	155,724	155,724	155,724
Outcome mean	0.001	0.001	0.001	0.001	0.001	0.001
K-P F-Stat	.	.	365.22	.	.	365.20

Note: This table presents the OLS, first stage, and 2SLS of the probability of receiving or sending a letter based on the distance to the nearest infected aristocrat in the kinship network. Standard errors are clustered at the 1×1 degree cell level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

4 The Republic of Letters and the rise of Applied Sciences

To study how the Republic of Letters stimulated the development of applied sciences, we partition Europe into 1×1 degree cells. We merge the letter and notable-individuals datasets at the cell-decade level, restricting the sample to the decades 1550–1850 to allow for sufficient pre-treatment periods. The panel is balanced: each cell enters in 1550 and is observed through 1850, yielding 31 observations per cell. We first present descriptive event-study evidence on the relationship between Enlightenment correspondence and the emergence of applied scientists, exploiting the staggered timing with which different cells entered the Republic of Letters. We then turn to the instrumental-variable strategy.

4.1 Event-study framework

We adopt the imputation estimator proposed by Borusyak et al. (2024), which is robust to treatment effect heterogeneity under staggered adoption. Define E_c as the *cohort* of cell c —the first decade in which a resident of that cell who is a member of the Republic of Letters sends or receives a letter. Cells that never receive or send such a letter have $E_c = \infty$ and serve as never-treated controls. The estimator proceeds in two steps. First, it estimates counterfactual outcomes for treated cell-decade observations using only untreated observations—never-treated cells and pre-treatment observations of eventually-treated cells:

$$Y_{c,t} = \alpha_c + \gamma_t + \varepsilon_{ct} \quad \text{for all } (c, t) \text{ such that } t < E_c, \quad (5)$$

where Y_{ct} is the log of one plus the number of applied scientists who died in cell c and had reached adulthood—defined as being born at least two decades before decade t —during decade t , and α_c and γ_t are cell and decade fixed effects.²¹ Second, the estimator imputes the treatment effect for each treated observation as $\hat{\tau}_{ct} = Y_{ct} - \hat{Y}_{ct}(0)$, where $\hat{Y}_{ct}(0)$ is the predicted counterfactual from equation (5). The event-study coefficients are then obtained by averaging these imputed effects at each horizon k relative to the treatment onset:

$$\tau_k = \mathbb{E}[\hat{\tau}_{ct} \mid t - E_c = k]. \quad (6)$$

We estimate τ_k for horizons $k \in \{0, 1, \dots, 5\}$ (i.e., from the decade of first letter receipt to five decades after treatment onset) and report five pre-treatment coefficients ($k \in \{-5, \dots, -1\}$) to assess the plausibility of the parallel trends assumption. We restrict the analysis to cohorts entering treatment between 1650 and 1830, the core period of the Republic of Letters.

Figure 3 presents the event-study estimates. The post-treatment coefficients display a

²¹Results are robust to using other transformations of the count of applied scientists, such as the inverse hyperbolic sine transformation.

clear and increasing pattern: cells that receive their first Enlightenment letter produce progressively more applied scientists in the decades that follow, pointing to an acceleration in the emergence of scientists after the first letter arrives. The pre-treatment coefficients are all statistically indistinguishable from zero, supporting the parallel trends assumption. However, there seem to be a spike in the $k = -1$ coefficient. There could be a couple of potential explanations for this pattern. The first is mechanical. Scientists enter the panel upon reaching their third decade of life; in most cases, their scientific contributions come later, but we record an increase in applied-science activity as soon as they enter the panel. If the arrival of a letter inspires individuals who are already in the panel to pursue applied science, this would generate a spike before the treatment date. The second explanation is that letters are a noisy marker of a cell’s initial engagement with the Republic of Letters: if some of the intellectual ferment that preceded letter exchange was already producing scientists, the $k = -1$ coefficient would capture this anticipation effect.

4.2 OLS and 2SLS

We now turn to the OLS and 2SLS estimates. Specifically, we estimate equations of the following form:

$$Y_{c,t+s} = \alpha \text{Letters}_{c,t} + \beta \text{InvPostalDist}_{c,t-1} + \mathbf{X}'_{ct} \boldsymbol{\phi} + \alpha_c + \gamma_t + \varepsilon_{it}. \quad (7)$$

The unit of observation is a cell-decade. The outcome $Y_{c,t+s}$ is the log of one plus the number of notable individuals classified as applied scientists who died in cell c and were at least twenty years old during decade $t + s$, with $s \in \{-20, 0, +10, +20, +30\}$ (decades). The endogenous regressor $\text{Letters}_{c,t}$ is the log of one plus the count of letters sent or received in cell c during decade t . The control $\text{InvPostalDist}_{c,t-1}$ is the inverse postal distance from c to the nearest already-treated cell, as defined in Section 3.2; it absorbs the direct effect of general postal accessibility to the Republic of Letters on the

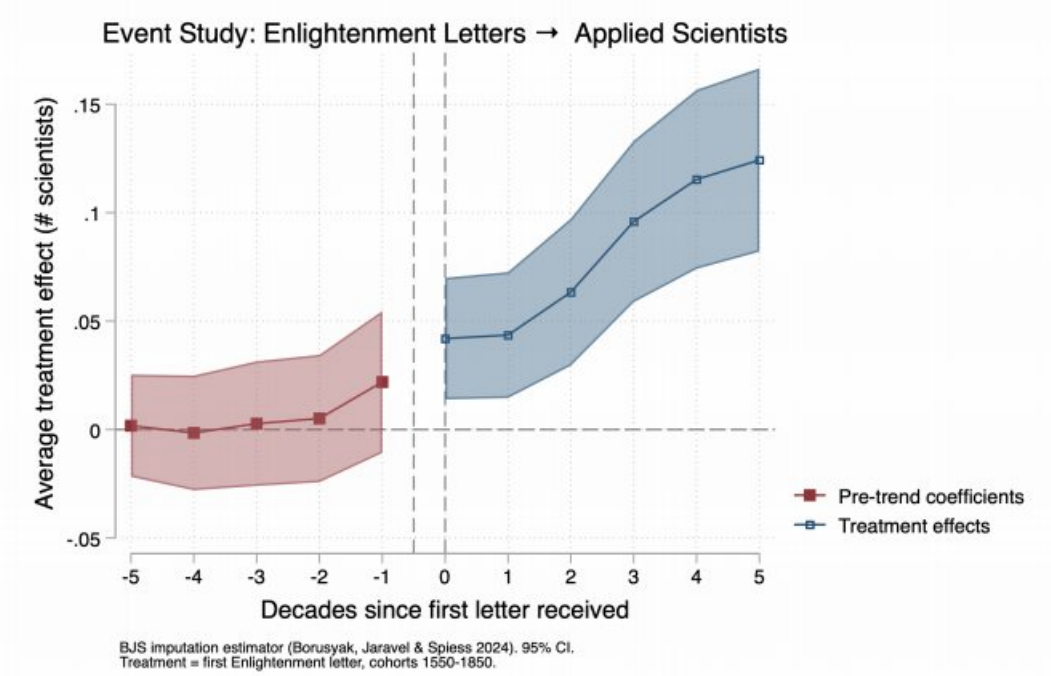


Figure 3: Event study: Enlightenment letters and Applied Science

Note: The figure plots event-study coefficients estimated using the imputation estimator of Borusyak et al. (2024). The outcome is the log of one plus the number of notable individuals classified as applied scientists active in a given cell-decade. Treatment onset is defined as the first decade in which a cell receives a letter from a member of the Republic of Letters. Red markers show pre-treatment coefficients ($k = -5$ to $k = -1$); blue markers show post-treatment coefficients ($k = 0$ to $k = 5$). Shaded areas indicate 95% confidence intervals. Treatment cohorts: 1650–1830.

emergence of applied scientists. \mathbf{X}_{ct} additionally controls for the count of all notable individuals in cell-decade (c, t) , ensuring that the estimated coefficient captures shifts in the composition of the educated stratum towards applied sciences rather than changes in coverage. The fixed effects α_c and γ_t absorb time-invariant cell heterogeneity and Europe-wide trends. Standard errors are clustered at the cell level. The sample covers 1,453 cells over the decades 1600–1850.

In the 2SLS specification, we instrument $\text{Letters}_{c,t}$ with $\text{InvAristPostalDist}_{c,t-1}$, defined as the maximum value of $\text{InvAristPostalDist}_{i,t-1}$ among all aristocrats i associated with cell c in decade $t - 1$. That is, the instrument captures the shortest postal-weighted kinship-network path from any aristocrat residing in cell c to an already-infected aristocrat, as described in Section 3.2. The exclusion restriction requires that, conditional on

the direct postal-accessibility control and the fixed effects included in the specification, $\text{InvAristPostalDist}_{c,t-1}$ affects applied-science output only through its effect on the volume of Enlightenment correspondence reaching the cell. Two features of the instrument's construction support this assumption. First, within-cell variation in $\text{InvAristPostalDist}_{c,t-1}$ is driven entirely by changes in postal routes over time. Second, the direct postal-accessibility control $\text{InvPostalDist}_{c,t-1}$ absorbs any effect of a cell's general postal connectivity on applied-science output, so identification comes exclusively from postal expansions that shortened the weighted path along the *specific kinship ties* connecting the cell's aristocrats to already-infected ones.

Table 3 reports the results, with columns 1–5 OLS and columns 6–10 2SLS at horizons $t - 20, t, t + 10, t + 20, t + 30$. We discuss OLS first. At $t - 20$ the estimated coefficient on log letters is -0.004 (s.e. 0.008); at t it is -0.001 (s.e. 0.010). Both are economically and statistically indistinguishable from zero. We read this as a pre-trend check that the specification passes: more correspondence today is not associated with more applied scientists already active two decades earlier, nor with more applied scientists active in the same decade. The coefficient turns positive and significant from $t + 10$ onward: 0.022 (s.e. 0.010) at one decade, 0.029 (s.e. 0.011) at two decades, and 0.042 (s.e. 0.011) at three decades. The monotone increase across leads is what a slow human-capital channel predicts: cells that receive letters today produce more applied scientists thirty years later than ten years later, because the relevant cohorts have had time to reach productive age.

Columns 6–10 turn to 2SLS. The pattern in the leads mirrors the OLS one. At $t - 20$ and t , the 2SLS coefficients are again statistically indistinguishable from zero. The 2SLS coefficient is positive and significant at $t + 20$ (0.389, s.e. 0.163) and at $t + 30$ (0.453, s.e. 0.173); at $t + 10$ it is 0.332 (s.e. 0.176) and borderline. To interpret these magnitudes, we anchor them to a dependent-variable mean of approximately 0.055 at the lead horizons. A 10% increase in letters maps to an increase of about 0.045 applied scientists in levels after 3 decades. On a baseline level of 0.055, this is an implied LATE of about 83% of the

baseline. Thus, for the cell-decades whose correspondence is moved by postal-induced changes in aristocratic-network proximity, a modest proportional shift in letter volume nearly doubles the local stock of applied scientists three decades later.

The 2SLS point estimates are roughly an order of magnitude larger than their OLS counterparts. The natural interpretation has two parts. The first is classical attenuation bias: $\text{Letters}_{c,t}$ is a noisy proxy for a cell's true exposure to Enlightenment correspondence, since the EE corpus is a selected subset of surviving letters digitized by the Bodleian Libraries, and measurement error in the regressor attenuates the OLS estimate toward zero. The second is the LATE interpretation: the 2SLS recovers the effect for the subset of cell-decades whose letter volume is responsive to postal-induced changes in network proximity. These are cells that were not already in the canonical core of the Republic of Letters—the compliers in the language of the IV literature (??)—and where each marginal letter arriving over a newly shortened postal route plausibly carries more informational content than an additional letter in already-dense correspondence hubs such as Paris or London.

Figure 4 examines which categories of correspondence drive the result. We estimate equation (7) at the $t + 20$ horizon, replacing the total number of letters with the count of letters classified into each of the following topical categories: technology, science, politics, art, philosophy, and religion. We report OLS estimates only: while we have an instrument for total correspondence, we do not have one for the topical composition of correspondence. These estimates should therefore be interpreted as descriptive within-cell associations, conditional on the same fixed effects and controls as Table 3.

The pattern in Figure 4 is consistent with the mechanism we propose. Technological Progress has the largest point estimate (0.072, s.e. 0.030), followed by Science (0.055, s.e. 0.020). Both are significant at the 5% level and their point estimates lie above the coefficient on total correspondence (0.029) from column (4) of Table 3. Politics (0.030, s.e. 0.014) and Art (0.029, s.e. 0.017) fall between the two, with Politics significant at the 5% level and

Art at the 10% level. Finally, Religion (0.010, s.e. 0.014) and Philosophy (0.022, s.e. 0.017) yield smaller estimates that are statistically indistinguishable from zero. The finding that the two categories most directly associated with useful knowledge carry the largest estimates is consistent with the contention of Mokyr (2002) that the Industrial Enlightenment's central dynamic was a feedback loop between practical mechanics and science, in which advances on each side raised the returns to advances on the other. We read the figure as suggestive evidence that the channel running from the Republic of Letters to applied science operates through the topical categories one would expect—and not, for example, primarily through religious correspondence.

Taken together, Table 3 and Figure 4 establish that Enlightenment correspondence is associated with the local emergence of applied scientists with a two-to-three-decade lag, that the association survives instrumenting letter volume with postal-induced variation in aristocratic-network proximity, and that it is concentrated in the topical categories most plausibly linked to useful knowledge. We turn next to whether the same source of variation explains the local emergence of inventors and entrepreneurs.

Table 3: Enlighted Letters and Applied Scientists

Outcome: ln (1 + # notable individuals working in the applied sciences)										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	t-20	t	t+10	t+20	t+30	t-20	t	t+10	t+20	t+30
ln (1+ # enlightened letters)	-0.004 (0.008)	-0.001 (0.010)	0.022* (0.010)	0.029** (0.011)	0.042*** (0.011)	0.225 (0.152)	0.304 (0.182)	0.332 (0.176)	0.389* (0.163)	0.453** (0.173)
# Notable Individuals	0.001** (0.000)	0.001* (0.000)	0.001* (0.001)	0.001* (0.001)	0.001* (0.001)	0.001** (0.000)	0.001* (0.000)	0.001* (0.001)	0.001* (0.001)	0.001 (0.001)
Inverse Postal Distance	2.651* (1.185)	3.577** (1.332)	4.318* (1.816)	4.767*** (1.275)	5.296*** (1.387)	-0.383 (2.611)	-0.455 (2.977)	0.300 (2.583)	0.059 (2.316)	0.142 (2.532)
City FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Decade FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Sample	Europe	Europe	Europe	Europe	Europe	Europe	Europe	Europe	Europe	Europe
Estimator	OLS	OLS	OLS	OLS	OLS	2SLS	2SLS	2SLS	2SLS	2SLS
Observations	37,778	37,778	36,325	34,872	33,419	37,778	37,778	36,325	34,872	33,419
Cells	1,453	1,453	1,453	1,453	1,453	1,453	1,453	1,453	1,453	1,453
Outcome mean	0.041	0.085	0.054	0.055	0.057	0.041	0.052	0.054	0.055	0.057
K-P F-Stat	8.39	8.39	10.43	14.11	13.42

Note: Standard errors are clustered at the 1×1 degree cell level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

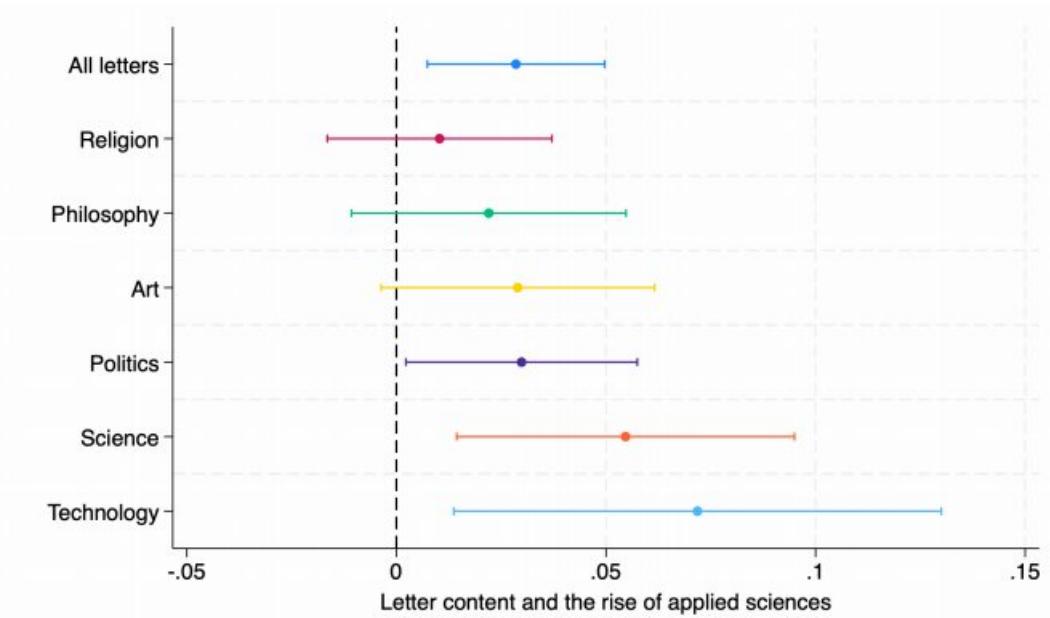


Figure 4: Number of letters and applied scientists after two decades

Note: The figure plots the coefficients of equation (7) with the dependent variable captured at $t+2$ and the number of letters distinguished by their content. The lines indicate 95% confidence intervals.

5 Enlightenment Exposure and Innovation

Section 4.2 suggests that postal-induced variation in exposure to the transnational Republic of Letters raised the local stock of applied scientists with a two-to-three-decade lag. This is in line with the theory of Mokyr (2017), according to which the Republic of Letters was a market for ideas in which useful knowledge was produced, certified, and rewarded. The natural empirical follow-up is whether the same source of variation also produced the complementary economic agents who translated that knowledge into output. We ask the question by estimating equation 4.2 with two new outcomes: inventors (Table 4) and the broader commercial category of merchants, entrepreneurs, and bankers (Table 5).

Table 4 reports the inventor results, with columns 1–5 OLS and columns 6–10 2SLS at horizons $t - 20$, t , $t + 10$, $t + 20$, $t + 30$. In both the OLS and the 2SLS estimates, the leads are insignificant in the pre-period and become significant from $t + 10$. In the 2SLS estimates, the leads from $t + 10$ have also a monotonic a monotone-increasing profile

across the three lead horizons. To translate the estimates into substantive magnitudes, a 10% increase in letters maps to approximately 0.029 added inventors in levels. On a baseline level of 0.028, this is roughly a doubling of the inventor stock three decades later. As in Section 4.2, the 2SLS point estimate exceeds its OLS counterpart by roughly an order of magnitude.

Table 5 turns to the broader commercial category. The pattern in the pre-period columns repeats: the coefficients at $t - 20$ and t are economically and statistically null under both estimators. Focusing on the 2SLS estimates, we find that the coefficient on letters are statistically indistinguishable from zero at every horizon except $t + 30$. Translated into levels, the 2SLS estimate at the longest lead implies roughly a fourteen percent rise in the local stock of merchants, entrepreneurs, and bankers three decades later for compliers. The asymmetry between the two tables— inventor effect after 1 decade, entrepreneur effect after 3 decades—is informative on its own terms, and it reads cleanly through the lens of Mokyr (2017): they are what one should expect if correspondence intensity loads sharply on the supply of useful knowledge and only translated later into new commercial organisations.

Table 4: Enlighted Letters and Inventors

	Outcome: ln (1 + # notable individuals working as inventors)									
	t-20	t	t+10	t+20	t+30	t-20	t	t+10	t+20	t+30
ln (1+ # enlightened letters)	-0.010 (0.006)	-0.006 (0.006)	0.022* (0.010)	0.014* (0.007)	0.018** (0.007)	0.144 (0.088)	0.176 (0.101)	0.223* (0.094)	0.252** (0.088)	0.285** (0.106)
# Notable Individuals	0.001*** (0.000)	0.001*** (0.000)	0.001* (0.001)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001** (0.000)
Inverse Postal Distance	0.210 (0.555)	1.295* (0.614)	4.318* (1.816)	2.083*** (0.600)	2.481*** (0.569)	-1.836 (1.569)	-1.113 (1.734)	-1.280 (1.516)	-1.032 (1.337)	-0.864 (1.514)
City FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Decade FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Sample	Europe	Europe	Europe	Europe	Europe	Europe	Europe	Europe	Europe	Europe
Estimator	OLS	OLS	OLS	OLS	OLS	2SLS	2SLS	2SLS	2SLS	2SLS
Observations	37,778	37,778	36,325	34,872	33,419	37,778	37,778	36,325	34,872	33,419
Cells	1,453	1,453	1,453	1,453	1,453	1,453	1,453	1,453	1,453	1,453
Outcome mean	0.010	0.085	0.014	0.014	0.015	0.010	0.013	0.014	0.014	0.015
K-P F-Stat	8.39	8.39	10.43	14.11	13.42

Note: Standard errors are clustered at the 1×1 degree cell level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Enlighted Letters and entrepreneurship

	Outcome: ln (1 + # notable individuals working as inventors)									
	t-20	t	t+10	t+20	t+30	t-20	t	t+10	t+20	t+30
ln (1+ # enlightened letters)	-0.000 (0.009)	0.008 (0.010)	0.022* (0.010)	0.033*** (0.010)	0.040*** (0.010)	0.142 (0.191)	0.157 (0.213)	0.196 (0.192)	0.308 (0.171)	0.395* (0.181)
# Notable Individuals	0.001* (0.001)	0.001* (0.001)	0.001* (0.001)	0.001* (0.001)	0.001 (0.001)	0.001* (0.001)	0.001* (0.001)	0.001* (0.001)	0.001 (0.001)	0.001 (0.001)
Inverse Postal Distance	2.333 (1.454)	3.730* (1.760)	4.318* (1.816)	4.813* (1.887)	5.650** (1.870)	0.457 (1.878)	1.762 (2.159)	2.077 (2.016)	1.224 (2.163)	1.197 (2.485)
City FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Decade FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Sample	Europe	Europe	Europe	Europe	Europe	Europe	Europe	Europe	Europe	Europe
Estimator	OLS	OLS	OLS	OLS	OLS	2SLS	2SLS	2SLS	2SLS	2SLS
Observations	37,778	37,778	36,325	34,872	33,419	37,778	37,778	36,325	34,872	33,419
Cells	1,453	1,453	1,453	1,453	1,453	1,453	1,453	1,453	1,453	1,453
Outcome mean	0.068	0.085	0.088	0.090	0.093	0.068	0.085	0.088	0.090	0.093
K-P F-Stat	8.39	8.39	10.43	14.11	13.42

Note: Standard errors are clustered at the 1×1 degree cell level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6 Conclusion

This paper has asked whether the Republic of Letters—the transnational correspondence network of Enlightenment scholars—helped reshape the human capital of the places it reached in the centuries preceding industrialization, and how this affected innovation and economic activity.

Two findings stand out. First, the Republic of Letters spread, in part, through the European aristocratic kinship network. Second, where the Republic arrived, the local production of useful knowledge, innovation, and entrepreneurship rose persistently in the decades that followed. Cells instrumented into the Republic experienced a near-doubling in the local stock of applied scientists—our proxy for the production of useful knowledge in the sense of Mokyr (2017)—and a near-doubling in the local stock of inventors roughly three decades after first contact. Merchants, entrepreneurs, and bankers move in the same direction and reach broadly similar long-run magnitudes, although the entrepreneurial response is noisier and becomes statistically distinguishable only at longer horizons.

We interpret these findings as the empirical micro-counterpart of the macro-historical mechanism proposed by Mokyr (2017). The asymmetry between the sharply identified effects on applied scientists and inventors and the noisier effects on entrepreneurs is itself consistent with the distinction between institutions that supported the production of useful knowledge and the property-rights regimes, credit arrangements, and commercial structures that translated such knowledge into firms and markets—institutions that the Republic of Letters did not directly provide. At the same time, our estimates capture only one of several channels through which Enlightenment ideas circulated, and the cell-level analysis cannot recover individual-level returns to correspondence.

The Republic of Letters did not produce the Industrial Revolution on its own. Yet where its correspondents arrived, the local stock of applied scientists and inventors rose persistently in the decades that followed—consistent with the view that a pan-European

market for useful knowledge helped reshape the geography of innovation before industrialization.

We believe that these findings speak to a broader contemporary debate about the relationship between the circulation of ideas and economic progress. In recent years, a movement often referred to as the “Dark Enlightenment” has questioned whether the values associated with the Enlightenment—openness, free inquiry, tolerance of dissent, and representative institutions—are conducive to economic dynamism or instead in tension with it (Land, 2013). Our results suggest that the diffusion of Enlightenment ideas was not a hindrance to growth but, on the contrary, contributed to the emergence of sustained innovation in early modern Europe.

References

- Acemoglu, Daron, Jacob Moscona, and James A Robinson**, “State capacity and American technology: evidence from the nineteenth century,” *American Economic Review*, 2016, 106 (5), 61–67.
- , **Simon Johnson, and James A. Robinson**, “Institutions as a Fundamental Cause of Long-Run Growth,” *Handbook of Economic Growth*, 2005, 1A, 385–472.
- Allen, Robert C.**, *The British Industrial Revolution in Global Perspective*, Cambridge: Cambridge University Press, 2009.
- Borusyak, Kirill, Xavier Jaravel, and Jann Spiess**, “Revisiting Event-Study Designs: Robust and Efficient Estimation,” *Review of Economic Studies*, 2024, 91 (6), 3253–3285.
- Brown, Tom B, Ben Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse,**

- Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei**, “Language models are few-shot learners,” *Advances in Neural Information Processing Systems*, 2020, 33, 1877–1901.
- Cantoni, Davide and Noam Yuchtman**, “Medieval Universities, Legal Institutions, and the Commercial Revolution,” *Quarterly Journal of Economics*, 2014, 129 (2), 823–887.
- Cinnirella, Francesco, Erik Hornung, and Julius Koschnick**, “Flow of Ideas: Economic Societies and the Rise of Useful Knowledge,” *Economic Journal*, 2025, 135 (669), 1496–1536.
- Clark, Gregory**, *A Farewell to Alms: A Brief Economic History of the World*, Princeton, NJ: Princeton University Press, 2007.
- Dibon, Paul**, “Communication in the Respublica Litteraria of the 17th Century,” *Res Publica Litterarum*, 1978, 1, 43–55.
- Dittmar, Jeremiah E.**, “Information Technology and Economic Change: The Impact of the Printing Press,” *Quarterly Journal of Economics*, 2011, 126 (3), 1133–1172.
- Hanlon, W Walker, Stephan Hebllich, Ferdinando Monte, and Martin B Schmitz**, “A penny for your thoughts,” Technical Report, National Bureau of Economic Research 2022.
- Hebllich, Stephan and Alex Trew**, “Banking and industrialization,” *Journal of the European Economic Association*, 2019, 17 (6), 1753–1796.
- Hornung, Erik**, “Immigration and the diffusion of technology: The Huguenot diaspora in Prussia,” *American Economic Review*, 2014, 104 (1), 84–122.
- Jacob, Margaret C.**, *The First Knowledge Economy: Human Capital and the European Economy, 1750–1850*, Cambridge: Cambridge University Press, 2014.

Land, Nick, *The Dark Enlightenment* 2013. Originally published as a series of essays at <http://www.thedarkenlightenment.com/the-dark-enlightenment-by-nick-land/>.

Landes, David S., *The Wealth and Poverty of Nations: Why Some Are So Rich and Some So Poor*, New York: W. W. Norton, 1998.

Laouénan, Morgan, Daniela Mantovani, Etienne Mercenier, Nicolas Zuber, and Roberto Galbiati, “A Cross-Verified Database of Notable People, 3500 BC–2018 AD,” *Scientific Data*, 2022, 9, 290.

Lazzaroni, Gianni Marciante Paolo Masella Matteo Cervellati Sara, “The Rise of the Knowledge Economy,” *NBER Conference Proceedings, Economics of Culture and Institutions, Spring 2026*, 2026.

Lundy, Darryl, *The Peerage*, <http://www.thepeerage.com>, 2020.

McCloskey, Deirdre N., *Bourgeois Dignity: Why Economics Can't Explain the Modern World*, Chicago: University of Chicago Press, 2010.

Midura, Rachel, “Itinerating Europe: early modern spatial networks in printed itineraries, 1545–1700,” *Journal of Social History*, 2021, 54 (4), 1023–1063.

– , “Early Modern Digital Itineraries: The Italian Core Set,” *DigItalia*, 2025, 20 (2), 97–110.

Mokyr, Joel, *The Gifts of Athena: Historical Origins of the Knowledge Economy*, Princeton, NJ: Princeton University Press, 2002.

– , “The Intellectual Origins of Modern Economic Growth,” *Journal of Economic History*, 2005, 65 (2), 285–351.

– , *A Culture of Growth: The Origins of the Modern Economy* The Graz Schumpeter Lectures, Princeton, NJ: Princeton University Press, 2017.

- Rubin, Jared**, "Printing and Protestants: An Empirical Test of the Role of Printing in the Reformation," *Review of Economics and Statistics*, 2014, 96 (2), 270–286.
- Sasaki, Yu**, "Postal Growth: How the State-Sponsored Post Affected Growth in Preindustrial France, 1500–1850," *Available at SSRN 4774412*, 2024.
- Schwennicke, Detlev**, *Europäische Stammtafeln, Neue Folge, Vol. I.1-I.3*, Vittorio Klostermann, 1998.
- Serafinelli, Michel and Guido Tabellini**, "Creativity over time and space," *Available at SSRN 3070203*, 2017.
- Squicciarini, Mara P. and Nico Voigtländer**, "Human Capital and Industrialization: Evidence from the Age of Enlightenment," *Quarterly Journal of Economics*, 2015, 130 (4), 1825–1883.
- Steinwender, Claudia**, "Real Effects of Information Frictions: When the States and the Kingdom Became United," *American Economic Review*, 2018, 108 (3), 657–696.

A Appendix: Data

A.1 Enlightenment correspondence

A substantial share of the letters in our corpus is not originally written in English: 47,959 letters are in other languages, predominantly French and Italian. To obtain a homogeneous English-language corpus suitable for subsequent analysis, we translate these letters into English using Google’s GEMINI 2.5-FLASH together with a few-shot prompting strategy.²² The prompt supplied to the model is reproduced verbatim below.

Translation prompt

Role: Act as an expert professional translator, with a perfect mastery of the English language and its stylistic and cultural nuances.

Task: Translate the provided text exclusively into the English language.

Operational Constraints:

- Fidelity and Register: Rigorously maintain the tone, style, and thematic nuances of the original text. Do not add explanations, opinions, or translator’s notes.

Idiomatic Localization: Adapt idiomatic expressions with natural equivalents in English, ensuring that the meaning and emotional impact remain unchanged.

- Terminological Precision: Translate every reference, including names of entities or specific concepts, so that the entire output is in harmonious English.
- Formatting Integrity: Exactly preserve the original structure (paragraphs, lists, bold, italics, and spacing).

²²Few-shot prompting is a technique in which a small number of input–output examples (typically fewer than ten) are included in the prompt before the main query. These demonstrations allow the model to infer the desired output format, style, or reasoning pattern without updating its parameters—a form of in-context learning (Brown et al., 2020).

- Prohibition of Metalanguage: The output must contain solely and exclusively the translation of the text. Do not include introductions (e.g., "Here is the translation:") or final comments.

Final Instruction: Respond exclusively in English. If the original text contains passages in other languages, these must be translated into English, unless they are untranslatable proper nouns.

Combining the letters originally written in English with the translated letters, we obtain a homogeneous corpus on which we perform topic classification. Each letter is assigned a binary indicator for each of seven thematic categories: *Religion, Family, Technological Progress, Science, Philosophy, Art, and Politics*. We classify each letter using GEMINI 2.5-FLASH, again with few-shot prompting.

To reduce the risk that a single-pass LLM classifier overattributes themes—labeling a category as present even when it is only tangentially mentioned in the text—we adopt an LLM-as-a-judge design. A second model, CLAUDEHAIKU4.5, independently audits the primary annotator’s classifications, flagging entries it deems potentially erroneous and providing a structured justification for each. Finally, GEMINI 2.5-FLASH re-examines the flagged letters in light of these justifications and produces a revised classification. The prompt used in the classification step is reproduced verbatim below.

Topic-classification prompt

You are a text classifier. Determine whether specific themes are present in a given text.

Rules:

- Base your judgment ONLY on what the text explicitly states.
- Do not infer themes from external historical knowledge.

- A theme should be marked "true" only if it is a main and central focus of the text. If you cannot classify because no valid text is provided simply, respond with "false". Do not give any explanations.

Category definitions:

- Religion: theology, faith, worship, church affairs, God, salvation, scripture, religious institutions as religious bodies
- Family: marriage, children, kinship, inheritance, household matters
- Technological Progress: inventions, machines, tools, manufacturing techniques, agricultural improvement, practical crafts
- Science: investigation of natural phenomena — experiments, theories, observations, mathematics, natural philosophy
- Philosophy: logic, ethics, metaphysics, epistemology, moral reasoning, aesthetics as abstract inquiry
- Art: painting, sculpture, music, architecture, literature, poetry, theater, aesthetic discussion
- Politics: governance, legislation, war, diplomacy, elections, political institutions, taxation, sovereignty

Return ONLY valid JSON. Schema: "Religion": true/false, "Family": true/false, "Technological_Progress": true/false, "Science": true/false, "Philosophy": true/false, "Art": true/false, "Politics": true/false

A.2 The aristocratic network

We aim to construct a panel at the aristocrat-decade level comprising individuals born between 1550 and 1850. In some cases, the year of birth and/or death is missing, so we implement a multi-stage imputation procedure to recover the decade of birth and death. Of the more than 720,000 aristocrats in the original dataset, we begin by dropping those for whom the year of death is known and falls before 1400 or after 1900, leaving 582,610 individuals. Within this set, only 14.0% have both the year of birth and death observed, 31.5% have only the year of birth, 7.8% have only the year of death, and 46.7% have neither. In the first stage, for the 31.5% of individuals with a known birth decade but missing death decade, we assign the average death decade of same-gender individuals born in the same decade; this cohort average is available for 9.3% of individuals, while for the remaining 22.2% whose gender-decade cell contains insufficient observations, we use the gender-level average lifespan in our 1550–1850 sample. For the 7.8% with a known death decade but missing birth decade, we apply the same procedure in reverse, subtracting the cohort or gender-level average lifespan; cohort averages cover nearly all of this group, with only 12 individuals requiring the gender-level lifespan. In the second stage, for the 46.7% of individuals missing both dates, we exploit kinship links from the genealogical network. For those with a recorded marriage, we use the marriage decade to infer their birth decade from the average birth decade of same-gender individuals who married in the same decade, and then derive the death decade from the resulting birth cohort (5.5% of aristocrats). For those without a recorded marriage who appear as parents, we apply the same logic using the decade of their first child (1.5%). All imputed values are rounded to the nearest decade. This procedure raises the share of individuals with both birth and death decades observed from 14.0% to 60.3%, with the remaining 39.7% lacking complete information—predominantly individuals with no observed dates and no kinship links in the network. This leaves us with 175,377 aristocrats born between 1550 and 1850.

To assign geographic locations to individuals in the *Peerage* dataset, we extract place names from the narrative field of each entry using regular expressions. We identify six types of locations: birth, death, burial, education, marriage, and residence. Each narrative is parsed for patterns such as “born at [place],” “died at [place],” “educated at [place],” “married [person] on [date] at [place],” “buried at [place],” and “lived at [place],” with multiple locations recorded where applicable (e.g., multiple marriages or educational institutions). We clean the extracted place strings to remove parenthetical notes and circumstantial descriptions (e.g., “killed in action”). We then geocode the unique European place strings using the Google Maps Geocoding API, obtaining latitude and longitude coordinates for each location. Finally, we spatially join the geocoded points to European LAU (Local Administrative Unit) polygons, assigning each location a standardised administrative code.

B Appendix: Additional Figures

Evolution of the share of notable people working in different professions

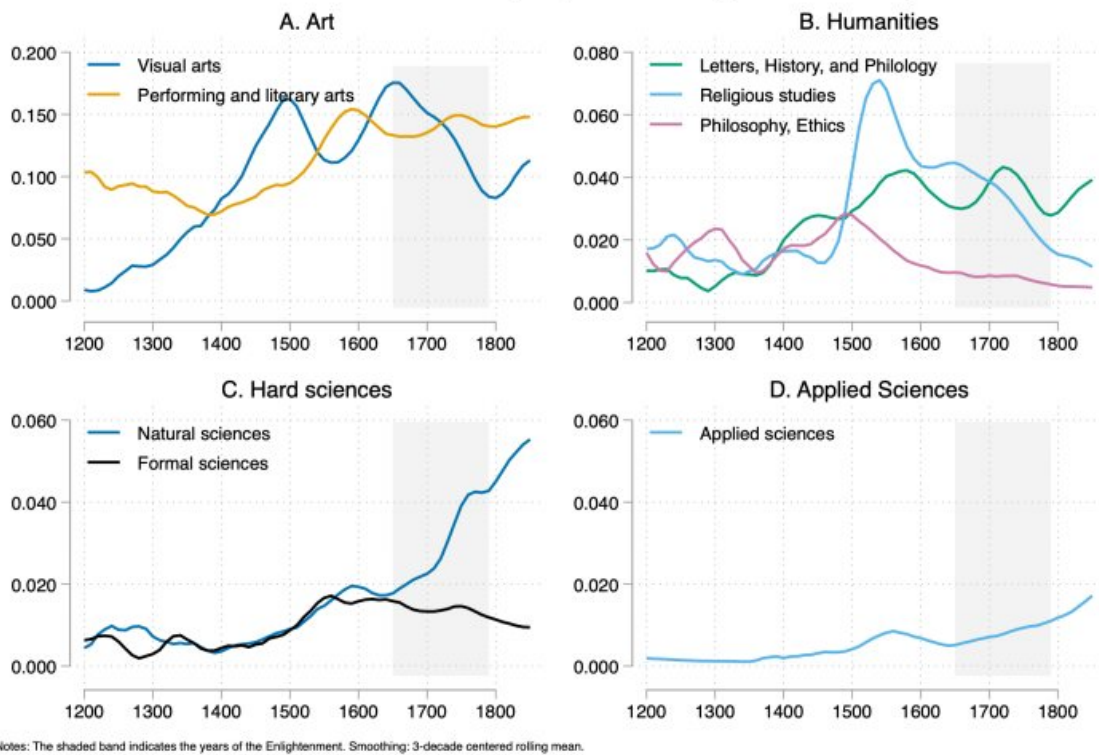


Figure A.2: Evolution of the share of notable people working in different professions, 1200–1850.

Evolution of the share of disruptive professions, 1200-1850

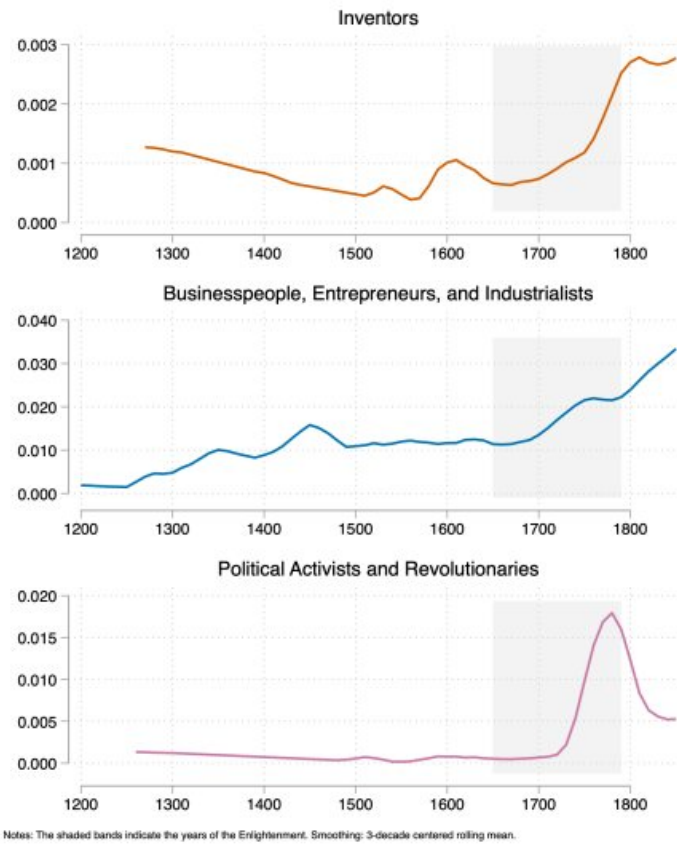


Figure A.3: Evolution of the share of disruptive professions, 1200–1850.

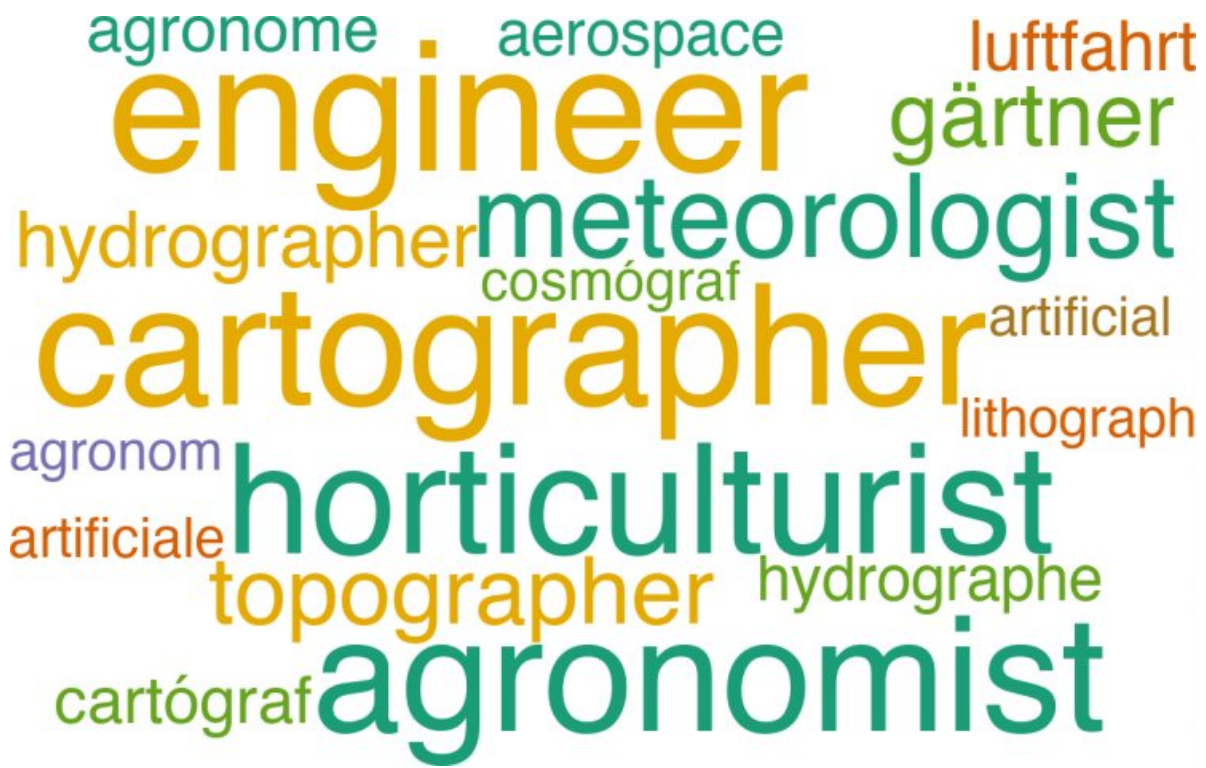



Figure A.4: Word cloud of sectors classified into Applied Sciences.

Figure A.5: The Peerage Example

Wolfgang Herzog von Braunschweig-Grubenhagen

M, #7388, b. 6 April 1531, d. 14 March 1595



Last Edited=10 May 2003

Wolfgang Herzog von Braunschweig-Grubenhagen was born on 6 April 1531. He was the son of [Philip Herzog von Braunschweig-Grubenhagen](#) and [Catherine von Mansfeld](#). He married [Dorothea von Sachsen-Lauenburg](#), daughter of [Franz I Herzog von Sachsen-Lauenburg](#) and [Sybille von Sachsen-Freiberg](#), on 10 December 1570 at [Osterode, Germany](#)^g. He died on 14 March 1595 at age 63.

He gained the title of *Herzog von Braunschweig-Grubenhagen*.

Note: Example from Darryl Lundy's genealogical website *The Peerage* (<http://www.thepeerage.com/>, accessed 04/11/2017) for Wolfgang, Duke of Braunschweig-Grubenhagen. Wolfgang is one of the over 680,000 nobles we use to reconstruct the European nobility network each year (see text for details).

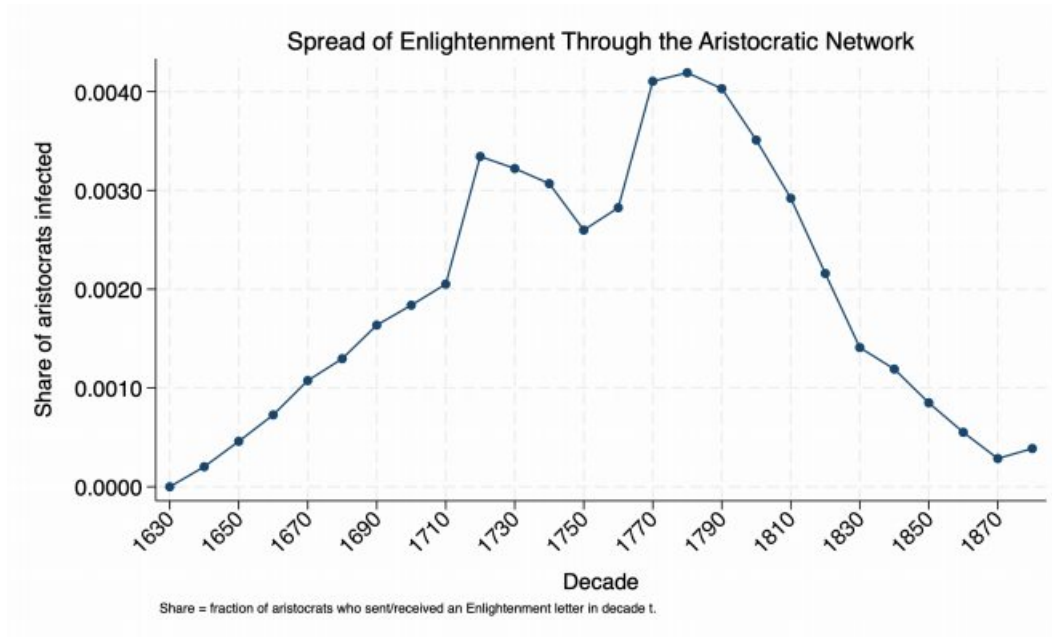


Figure A.8: The spread of the Enlightenment through the aristocratic network. The figure shows the share of aristocrats who are infected (i.e., have sent or received at least one Enlightenment letter) in each decade. The S-shaped trajectory is consistent with a diffusion process, with infection rates rising from near zero before 1700, accelerating through the mid-eighteenth century, and peaking around the 1790s.

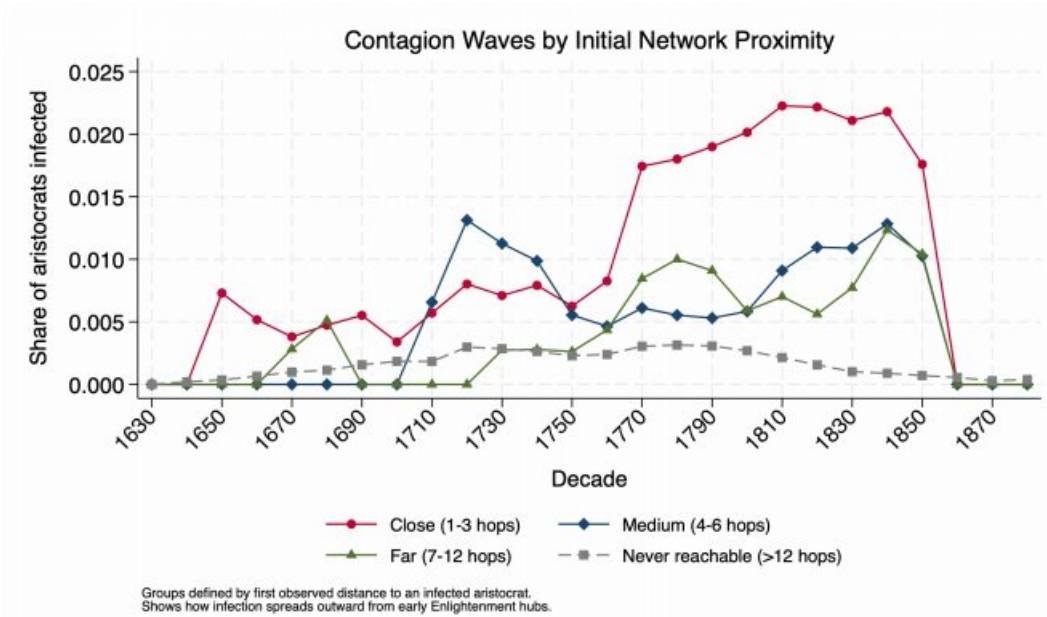


Figure A.9: Contagion waves by initial network proximity. Aristocrats are grouped by their first observed network distance to an infected node: close (1–3 hops), medium (4–6), far (7–10), and never reachable. Close aristocrats are infected earlier and at higher rates, consistent with a diffusion process that spreads outward from early Enlightenment hubs.

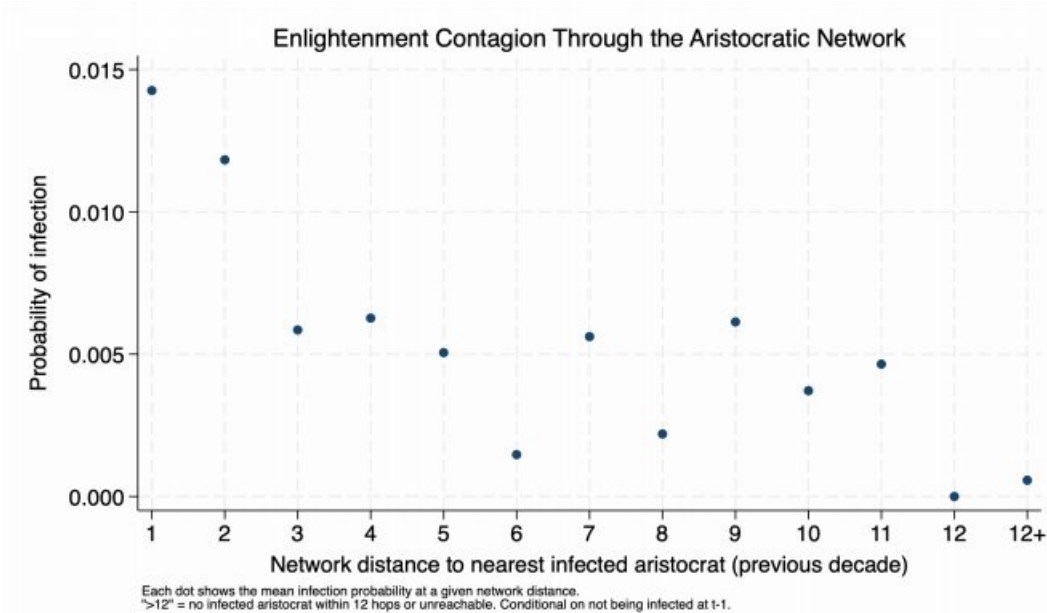


Figure A.10: Transition probabilities by network distance. For each distance bin, the figure shows the fraction of aristocrats who are uninfected at t and become infected at $t + 1$. Labels indicate the number of observations in each bin. The transition probability declines monotonically from 1.3% at distance 1 to 0.06% at distance >10 , confirming that proximity to infected nodes predicts the timing of infection.

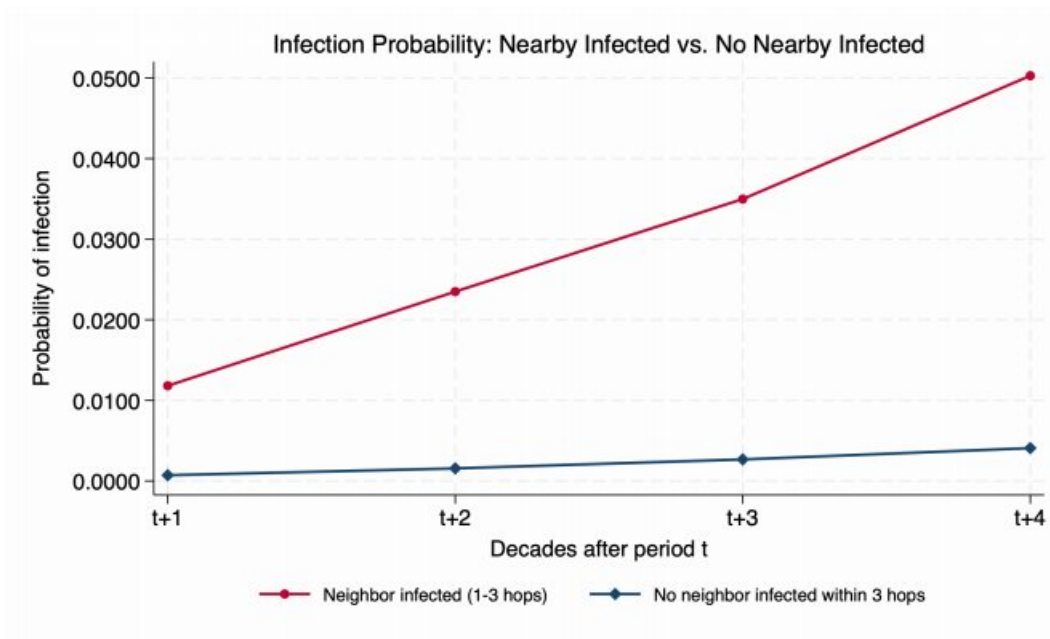


Figure A.11: Infection probability: aristocrats with a nearby infected neighbor vs. those without. The figure compares the raw probability of infection at horizons $t + 1$ through $t + 4$ for two groups of aristocrats, both uninfected at t : those with at least one infected aristocrat within 3 hops (red) and those with no infected neighbor within 3 hops (blue). At $t + 1$, the infection rate is 0.96% among the treated group and 0.06% among the control group.

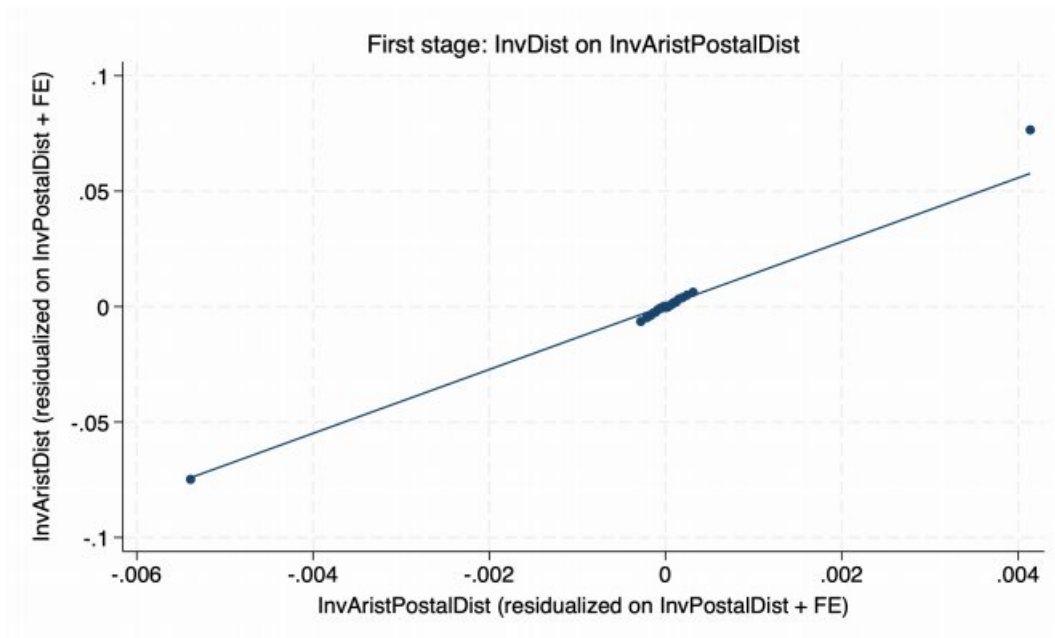


Figure A.12: First-stage relationship between reciprocal network distance and reciprocal postal-weighted aristocratic distance. The figure plots a binned scatter of $\text{InvAristDist}_{i,t-1}$ against $\text{InvAristPostalDist}_{i,t-1}$ after partialling out aristocrat and decade fixed effects and the reciprocal direct postal distance to the nearest infected city. The overlaid line is the first-stage best-fit slope (coefficient 16.04, s.e. 0.84; Kleibergen-Paap F = 365.22, well above the Stock-Yogo critical value of 16.38 for 10 percent maximal IV size). Sample: 972,434 aristocrat-decade observations covering 155,724 aristocrats.