THE RISE OF THE KNOWLEDGE ECONOMY

Matteo Cervellati Sara Lazzaroni Gianni Marciante Paolo Masella *

April 22, 2026

Abstract

The roots of sustained growth in England are increasingly located in the period from the Scientific Revolution to the Enlightenment. Productivity increases have been suggestively linked to the development of communication infrastructures and the diffusion of *useful knowledge*. In this paper, we isolate the empirical impact of the evolution of the postal system in England and Wales over the period 1570-1769 on interpersonal communication exchanges in the context of the *Republic of Letters*. We exploit a quasi-natural experiment due to a postal system reform leading to the opening of the pre-existing network to the public in 1635. We build a novel geo-referenced database involving about 60,000 letters by roughly 9,000 correspondents. The reform leads to a substantial increase in interpersonal communications. The results of DiD/IV and event studies confirm the baseline findings, bolstering confidence in a causal interpretation of the estimates. Looking at the patterns of interactions, we detect a sizable intensification of communications between scholars and professionals. We collect the content of around 30,000 letters, classify their topics exploiting unsupervised text analysis, and detect a sizable impact of the reform on the evolution of communications related to useful knowledge. Finally, drawing on biographical data, we document the effects on innovation activities. Taken together, the findings provide the first systematic evidence of the role of communication infrastructure for the rise of Knowledge Economies in the early modern period.

*We thank Joel Mokyr, Jeremiah Dittmar, and participants at the Louvain FRESH Workshop, ASREC Europe in Thessaloniki, University of Bologna, the Deep-Rooted Factors in Comparative Development Conference at Brown University, University of Tor Vergata, the CesIFO Summer Forum 2025, and the PSD-GLP Seminars at the University of Paris 1 - Panthéeon Sorbonne for useful comments. We thank Arno Bosse and Mark Rogerson for their help with the data at EMLO and EE, respectively. We are grateful to Luca De Angelis and Giacomo Plevani for their technical help. We thank Luca Tancetti for his important contribution to an early development of the project, and Devis Decet and Elena Sala for excellent research assistance. The project has been developed at the Department of Economics University of Bologna and financial support from the PRIN Grant # PRIN 2020 is gratefully acknowledged. Cervellati: m.cervellati@unibo.it (Dep. Economics University of Bologna, CEPR-London); Lazzaroni: sara.lazzaroni@unibo.it (Dep. Economics University of Bologna); Marciante: gianni.marciante@unibo.it (Dep. Economics University of Bologna); Masella: p.masella@unibo.it (Dep. Economics University of Bologna, IZA-Bonn).

The short answer as to why the West is so much richer today than it was two centuries ago is that collectively, these societies "Know" more. (Mokyr, 2005, p.287)

(...) production can draw on accumulated useful knowledge, there has to be social contact between "knowers" and "doers." (Mokyr, 2002, p.63)

1 Introduction

The emergence of sustained economic growth in England was associated with unprecedented rises in productivity that gained momentum from the early modern period. In this age of the Enlightenment, knowledge produced after the Scientific Revolution was increasingly converted into the productive know-how that paved the way for the Industrial Revolution (Landes, 1969; Pomeranz, 2000; Jones, 2003; Mokyr, 2010). Innovations like the printing press greatly favored the spread of codified knowledge. Nonetheless, direct interactions between *knowers* and *doers* have been considered key for the creation and diffusion of *useful knowledge* in terms of productive—technical, organizational, and institutional—innovations (Mokyr, 2016).

Person-to-person information flows across locations largely rested on epistolary exchanges. Historians underscore a peculiar social institution known as the *Republic of Letters*. The term refers to a pan-European network of individuals sharing the goal of cooperating for the advancement of knowledge. Besides intellectual elites like academics and thinkers, it involved individuals in apical positions like statesmen, and segments of the rising productive elites, including practitioners, entrepreneurs, and innovators. England was a forerunner in the development of a functioning national postal system, which was open to the general public already in the first half of the seventeenth century and which might have bolstered its lead in the rise of its modern *Knowledge Economy* (Merton, 1938; Mokyr, 2002). In line with this perspective, historical narratives portray an intensification of epistolary exchanges before the Industrial Revolution.

In spite of the eminent arguments, limited formal evidence exists on the role of communication infrastructures during the Enlightenment and several unsettled questions stand out. Did the early access to communication infrastructures in England impact the intensity of information exchange across locations? How did it affect the patterns of person-to-person communications and, in particular, interactions between the knowers (i.e., scholars) and the doers (i.e., professionals)? What was the content of these communications, and can it be traced to the production and diffusion of useful knowledge? Did it matter for innovation activity? In this paper, we address these questions by tracking the evolution of the postal network across time and space in England and Wales, and by isolating its impacts on person-to-person epistolary exchanges in the context of the Republic of Letters in the period 1570-1769.

The context under investigation offers an ideal setting both for the general relevance and for the validity of the empirical identification. The role of the Industrial Enlightenment in the economic rise of England is a crucial and long-lasting topic in the social sciences. More generally, the emergence of knowledge-based economies before the Industrial Revolution is central to debates on the rise of Europe (the "Great Divergence"). Methodological challenges and data availability have hitherto limited the scope of formal investigations of the rise of knowledge economies and innovation activities before industrialization.

Important advancements have recently been achieved by looking at the spread of knowledge codified in published books and manuals (as discussed below), but empirical investigations of the role of interpersonal communications for knowledge diffusion are still missing. Finally, the role of England's early development of a national postal system, while extensively discussed by historians, has not been formally investigated.

In this paper, we contribute towards filling these gaps by building a novel geo-referenced panel database from a set of primary and secondary sources. The peculiar features of the creation of the national postal system imply a quasi-natural experiment that can be leveraged for empirical identification. See the historical background in Section 2. A communication infrastructure at the service of the Crown was created before the seventeenth century during war times as straightaway connections between London and the battlefronts (Robinson, 1948). Based on a costly network of horse posts, the infrastructure faced a serious shortage of resources with the death of Queen Elizabeth I in 1603. A few years later, Mr. Thomas Witherings, a reputed merchant, proposed a reform aimed at making the system a self-financing state monopoly. In 1635, the King approved the proposal, and the Royal Post was born. The opening to the public of the pre-existing postal routes implied that within the subset of locations with a postal stage, all individuals could send and receive letters at low prices and fixed scheduled days of delivery. Using difference-in-differences, we estimate the impact of the reform in locations that got access to the pre-existing network compared to locations without a postal stage before the reform. Event studies allow us to track the impact of the reform over time.

We produce a set of geo-referenced maps that track the evolution of access to the postal system across locations in England and Wales over the period 1570-1690. Information on epistolary communications is extracted from the universe of letters made available in the catalogues of Electronic Enlightenment (EE) and Early Modern Letters Online (EMLO), which have been assembled by large scale projects in digital humanities over the last three decades. These data cover communications across all of Europe (but also the Americas and Asia). A process of geo-time referencing is applied to locate the origin and destination of each letter. We restrict our focus to around 56,000 letters sent and received within England and Wales. See Section 3 for a discussion of the main sources and the empirical strategies.

The analysis proceeds in several steps. First, in Section 4, we estimate the impacts of the reform on the intensity of correspondence across treated and non treated locations. The results of the DID analysis show that the reform led to a differential increase in epistolary exchanges in the treated locations. The patterns are robust to a variety of robustness and sensitivity checks.[1] Information on the location of historical battlefronts allows to predict the least-cost paths connecting London to the main battlefronts which can be used to instrument the location of pre-reform postal stages. The main results are robust to this alternative identification strategy. The use of event studies allow checking the validity of the identification assumptions and tracking the effects over time. The results shows that the reform significantly increased the intensity of epistolary exchanges immediately after the opening of the postal system. Taken together, these checks bolster confidence on a causal interpretation of the results.

As a second step, in Section 5, we look at communications across different types of correspondents.

[1] These include alternative measurement of the dependent variable and empirical specifications. We also check the sensitivity of the results to outliers and historically peculiar locations, the coding of treated locations, and alternative coding of controls, among others.

Exploiting information on names, dates of birth/death, and any other relevant biographic information made available in the sources, we disambiguate each sender and recipient across catalogues. This process allowed isolating around 9,000 correspondents. Our interest is, in particular, in unfolding the impacts on the patterns of personal interactions among the conceptual categories of "knowers" and the "doers". We interpret knowers as the individuals involved in the production and dissemination of scientific knowledge across different academic disciplines. These individuals, that are labeled as *scholars*, include natural scientists, humanists but also non-academic thinkers. "Doers" are classified looking at a broad category, labeled *professionals*, that includes practitioners, businessman, technicians, and statesmen, as well as individuals involved in the information industry.[2] The results document that, on top of an increase in correspondence intensity, the opening of the postal system involved changing patterns of communication across different correspondents' categories. Communications between scholars became particularly significant after the end of the Civil War. Interestingly, and in line with perspective that postal systems represented a main instrument for accessing useful knowledge by part of non scholars, we find that the reform disproportionately increased the intensity, and even the share, of communication exchanges between scholars and professionals.

The content of epistolary communications is studied in Section 6. The analysis relies on information from around 30,000 communications for which abstracts or the full text is available. To perform the text analysis, the content of the letters (written in multiple languages including archaic English, French, German, and Latin) has been translated into English. Using a Latent Dirichlet Allocation (LDA) algorithm the content of the letters is classified into macro topics. The relative majority of communications relate to what we label *ordinary life* matters and include, for instance, communications with relatives and friends, running the household or family estates, scheduling visits, etc. Our main interest, is on the macro-topic labeled *knowledge*. Further macro topics are on politics and religion.[3] The results document that compared to letters referring to ordinary life matters, access to the postal system had a more gradual effect on knowledge-related correspondence, peaking in the second half of the seventeenth century.

The analysis in Section 7 focuses on the changing content of communications by types of correspondents, and in particular on knowledge-related correspondences involving scholars and professionals. The sub-topics identified by unsupervised text analysis align well with the conceptualization of useful knowledge discussed by historians and cover discussions on advances in theoretical and practical knowledge through experiments, cataloguing, the process of publishing, as well as more practical applications and discussions on how to make knowledge universally accessible. We find that knowledge-related epistolary exchanges increased among scholars, but also between scholars and professionals. In fact, the effect of the postal reform are largest for the intensity of knowledge exchanges

[2] We interpret doers across professional categories involved in the application of knowledge for productive purposes from a broad perspective (see, e.g., Mokyr, 2002). These involve, e.g. physicians, manufacturers, businessmen, builders, craftsmen, engineers (jointly labeled as practitioners). The category of statesmen involves personnel of government administration, diplomats, colonial agents, as well as MPs, who were increasingly involved in the regulation of economic activities. Given their focus on knowledge transfer, we also track individuals involved in the information industry, which include editors, librarians, typographers, antiquarians, encyclopedists, etc. Additional categories are clergymen (e.g., members of the church, Jesuits, missionaries, etc.) and people that could not be classified in any of these aforementioned groups of scholars and professionals.

[3] The topic *politics* spams discussions of royal and court affairs, the management of colonies and wars, local politics, etc. Letters with a *religion* content revolve around teleological discussions and organization of the Church. A final topic, which is excluded from the macro category of knowledge, is labeled *academic gossip*.

with professionals. Event study analysis document, however, a delayed effect on these communications compared, most prominently, to letters on ordinary life matters.

Finally, in Section 8 we evaluate the effects of the reform on innovation activities. To explore the impacts beyond cross-sectional correlations, we assemble a novel dataset of around 600 innovators active in England and Wales between 1570 and 1769, by extracting and classifying individuals included in the Oxford Dictionary of National Biography (ODNB).[4] For each innovator, we geo-reference life events (e.g., visits, apprenticeships) and innovations, to obtain proxies of innovation activities over time and space. The results show that the access to the postal system in treated locations increased innovation activities in subsequent decades only for technical innovators (i.e., in natural sciences), while no significant effects for soft innovators (involved in arts and humanities) can be detected. To interpret the findings, we also validate the analysis crossing information from the database of correspondents and of innovators. The results confirm that the postal reform significantly increased communications between scholars and technical innovators.

This paper contributes to several strands of the literature in long-term development and economic history. The background is the tradition in the social sciences aiming at understanding why sustained growth first occurred in Europe—the so-called "Great Divergence"—and more specifically why in England—the so-called "Little Divergence". Initially portrayed as a sudden leap into industrial production out of hitherto stagnation, the process of long-term development is increasingly interpreted as resulting from economic, technological, social, and institutional changes that progressively gained momentum in pre-industrial times.

In this paper, we explore the empirical validity of arguments tracing the roots of industrial development to the rise of knowledge economies during the Age of the Enlightenment. Recent empirical estimates indirectly support these perspectives by locating the onset of the increases in living conditions in England earlier than initially thought, i.e., already starting from the second half of the seventeenth century (see, e.g., Humphries and Weisdorf, 2019; Broadberry, 2021; Bouscasse et al., 2025).

The analysis and results contribute to a rising area of research on the legacy of history for social interactions, cultural values, and long-run economic development (e.g., Spolaore and Wacziarg, 2013; Giuliano and Nunn, 2021, for surveys). Within this broader topic, we focus on the intensification of person-to-person direct interactions related to the production and diffusion of useful knowledge during the Enlightenment (Mokyr, 2016). Conceptually, the analysis is also indirectly informative on the role of the postal system in making locations more densely connected and unfolding scale effects.[5]

We contribute a first systematic measurement and empirical identifications of the impacts of communication infrastructure on the intensity, patterns, and content of person-to-person interactions across different locations in pre-industrial times. More specifically, the analysis contributes to a small but rising body of formal works that have pushed the investigations of the role of science, knowledge diffusion, and long-term development by performing econometric analysis for the pre-industrial period.

[4] As preliminary exploration, we use existing data and document that locations exposed to the opening of postal stages in the early seventeenth century produce more patents in the longer run (until mid nineteenth century).

[5] The diffusion of postal stages and the rising territorial state capacity (and the associated increased density of economic activities) has been investigated in the context of nineteenth century US and UK (Acemoglu et al., 2016; Rogowski et al., 2022; Hanlon et al., 2022; Aneja and Xu, 2024). We are not aware of any attempt to estimate the impacts of postal infrastructure for the patterns of person-to-person communications and the diffusion of useful knowledge before the Industrial Revolution.

The focus of our analysis is on the empirical role of infrastructures that favor person-to-person interactions. In this respect, our findings conceptually complement the analysis that studies the role of historical innovations that favor the diffusion of information before the industrial revolution. Dittmar (2011) shows that the adoption of the printing press across locations fostered city growth over the period 1400-1600. Related to the question of useful knowledge, Dittmar and Seabold (2019) show that the printing of books with a more applied content, like business manuals often printed in vernacular languages, disproportionately affect city growth while religious books with a protestant content impact institutional reforms.

The findings on the role of communication exchanges on practical knowledge between scholars and professional complement the existing evidence on the role of diffusion of technical knowledge among individuals in prominent societal positions. Most notably, Squicciarini and Voigtländer (2015) track the diffusion of the *Encyclopédie* in mid-eighteenth century and isolate a key role of upper tail human capital for the onset of industrial development in France. Our focus on the content of person-to-person communications exchanges complements the analysis by Almelhem et al. (2026) on the content of books in England showing that the language of science became more independent from religion and more progress-oriented from the early eighteenth century.

The analysis conceptually relates, and complements, the investigations of the role of universities and academic scholars. Cantoni and Yuchtman (2014) show that the foundation of universities in Germany in the middle ages favored commerce and economic activities. De la Croix et al. (2024) builds a large novel database on the universe of scholars in pre-industrial Europe and tracks the evolution of the academic market over the period 1000-1800. De la Croix et al. (2025) study the effects of affiliations on the diffusion of main scientific ideas looking at interpersonal links by means of academic affiliations.

Finally, our analysis adds to a growing body of formal investigations of the determinants of long-term development during pre-industrial times across different locations within England. Recent works have pushed the analysis by exploiting spatially disaggregated data. Most notably, Heldring et al. (2021) document the role of the dissolution of English monasteries in the sixteenth century in shaping the process of commercialization and paving the way to the industrial development at the onset of the Industrial Revolution. Angelucci et al. (2022) isolate the impact of trade across towns for the patterns of political representation in the Middle Ages, participation to the Civil War in the seventeenth century, and longer term outcomes.

We contribute to these streams of literature with a first systematic evidence on the impact of communication infrastructures in shaping person-to-person interactions for the rise of knowledge economies in seventeenth-century England and Wales. Complementing existing evidence based on codified knowledge published in books by academic scholars, the results uncover spatial patterns of intensified interactions between knowledge elites and professionals in the context of the diffusion of useful knowledge. The results also complement the evidence on the role of local trade and market history by adding evidence on the role of local exposure to communication infrastructure for innovation activities.

2 Historical Background

2.1 Postal Communications and the 1635 Reform

During the early sixteenth century, the so-called *King's Posts* was an unrefined postal system lacking a structured network of postal routes (Brayshay, 1991, p. 378). Its aim was to "provide beleaguered English monarchs as quickly as possible with news from distant places" (Campbell-Smith, 2011, p. 7). The system, in place since the days of King John, was based on royal messengers, called *cursores*, who were paid according to the length of their travels (Hemmeon, 1912, p. 3). During the Wars of the Roses (1455-1485) and the English invasion of Scotland of 1482, Edward IV, in need of keeping a close track of the conflict events, mandated to place temporary stables with fresh horses at intervals of roughly a dozen miles along the ways to the battlefronts. The idea of 'laying the posts', or more simply arranging 'post stations' to be used during war times, outlived Edward IV and became the cornerstone of future developments of the postal system, especially after the appointment of Brian Tuke, a court official, as the first 'Master of the Posts' by King Henry VIII in 1517. Tuke arranged permanent stables run by *deputy postmasters*. The first two main post roads to be equipped with permanent staging posts were the North Road, connecting London with the Scottish border, and the Dover Road, linking London with the port of Dover on the English Channel. In the last decades of the sixteenth century, three new post roads with staging posts were set up in response to a rebellion against the Crown happening in Southern Ireland: the West Road, linking London to the Cornwall county, in the far southwest of England; the Bristol Road, stretching from London to Bristol; and the Irish Road, extending from London to the city of Holyhead, in northwestern Wales.

Despite this expansion of the postal network, "little was done to improve the general administration of the Posts" in that period (Campbell-Smith, 2011, p. 19). As the deputy postmasters were remunerated with a basic state salary, their financial conditions critically deteriorated in the second half of the sixteenth century, when a combination of high inflation rates and a sizable wage cut ordered by Queen Elizabeth I pushed them close to their financial limits. The general postmaster Tuke and his successors could not be of much help, as they were not allowed to decree any increase of the deputy postmasters' wages without the royal consent (Hemmeon, 1912, p. 6). In 1583 Elizabeth I made a few concessions to 'uncommissioned' private users of the royal postal service to help supporting the running of the system, but post-riders had to always prioritize the delivery of papers and letters for Her Majesty's affairs over the consignment of private letters. For this reason, by 1600 posting was still a "service of the court [... and] the real beginning of a public service only [came] in the next century" (Robinson, 1948, p. 21). Queen Elizabeth's death in 1603 did not bring about a general improvement in the national postal system either. As James I took power, the Crown's concerns over the financial stability of the King Posts became recurring and the lack of resources became so severe that the government ended up falling into arrears with the payment of deputy postmasters' state wages (Robinson, 1948, p. 35). By the early 1630s the system was thus in need of a deep overhaul, as also demonstrated by the poor quality of most post roads and the slow conveyance of messages.

In 1632 the King's Principal Secretary of State, John Coke, appointed Thomas Witherings, a respected merchant from London, as general postmaster of England for the Foreign Posts, the official and private postal business along the Dover Road. Before the end of that year, Witherings handed to Coke a detailed plan to

reorganize the branch of the postal network under his supervision, through the use of mounted couriers (*staffeti*) who were to travel day and night from stage to stage. Witherings' plan was an evident success, as it soon reduced the journey time between London and the port of Antwerp in the Flanders region from eight to four days (Sharpe, 1984, p. 160). More importantly, the plan kick-started a much more ambitious reform that Witherings himself carefully crafted in that period, whose final goal was to make the Posts a self-supporting business and "no longer a charge to the state" (Hemmeon, 1912, p. 13).

In his "Proposition for settling staffets or packet posts betwixt London and all parts of His Majesty's dominions" dated June 1635, Witherings set up a fixed scale of charges for the carriage of private letters on all the main post roads (now amounting to six, after the addition of a new postal route from London to Norwich via Colchester).[6] Deputy postmasters would then be not only forbidden from pocketing the fees paid for private mail, but would also be required to remit the money to a newly-established 'Letter Office' in London, which was responsible for paying out their salaries. The "Proposition" also prohibited the deputies from hiring out their horses on the days on which the post was due for delivery. Finally, the day and hour of the arrival and departure of *staffeti* to and from London had to be "always certain" (Robinson, 1948, p. 30) and "mail brought along one road to London and addressed to a town on another road" was to be "transferred at the London office" (Beale, 1998, p. 227). On July 31 of that year, Coke approved Witherings' reform proposal and issued a royal proclamation that authorized "the settling of the Letter Office of England and Scotland". On that day the Royal Mail was born.

The reform of 1635 marked the beginning of a radically new period of English communication history. Crucially for our purposes, the change to a revenue-paying basis was made possible both by making the postal system openly available to the public (for the first time ever) and by establishing a statutory monopoly over the carriage of official and unofficial letters.[7] According to Hemmeon (1912), Witherings' work was so effective that "by 1636 his reform was in full and profitable operation" (cf. p. 15). As a result, the whole kingdom experienced a remarkable improvement in postal arrangements: the postal service became much more rapid, reliable, and financially sustainable, and the volume of letters passing along the kingdom's post roads gradually rose (Campbell-Smith, 2011, p. 27). In the following decades, the postal network expanded to serve the population off the six main routes.[8] Still, the administration of the Letter Office kept following

[6]Details on the payment can be found in Beale (1998): "[The proposition] gave postage rates of 2 pence per single letter [i.e. a letter taking only one sheet of paper] for journeys of under 80 miles, of 4 pence for 80 to 140 miles and of 6 pence for journeys above that. Letters to or from Scotland were 8 pence. The same payments were to be applied to the road to Ireland" (cf. p. 227). Also, "the postmasters in the country were not to take any charge for a letter, except one penny for carriage to the next market town. [...] Letters for any of the towns off the great roads [...] were to be left at the post-towns nearest the country towns to which they were directed. They were then to be carried to their destination by foot-posts to a distance of six or eight miles and for each letter these foot-posts were to charge 2 pence, the same price that was charged by the country carriers" (Hemmeon, 1912, pp. 13-14). Based on data from Allen (2001), a London building laborer earned about 7.1 grams of silver per day in 1600-1649, while a building craftsman (a professional) earned approximately 11.3 grams. Converting 2 pence into silver at the post-1601 mint standard, the cost of sending a one-sheet letter within 80 miles amounted to roughly 14% of a laborer's daily wage and about 9% of a craftsman's daily wage, or approximately two loaves of bread (Allen, 2009).

[7]Private carriers kept operating in competition with the Letter Office after 1635, but were increasingly overloaded with legal constraints slowing down their activity and therefore lost relevance within a few decades (Campbell-Smith, 2011, pp. 27-28).

[8]During the 1660s and 1670s, deputy postmasters on the main routes were allowed to establish new branches, called *by-posts*, between the main post roads and the kingdom's largest towns. In the 1690s the number of by-posts increased significantly and some postmasters were granted the permission to establish *cross-posts*, i.e. links between the six main postal routes.

closely the practices laid down in the "Proposition" of 1635 for almost two centuries (Robinson, 1948).[9]

2.2 Epistolary Communications and the Diffusion of Useful Knowledge

Between the XVI and XVIII century and before the advent of scientific journals, letters were by far the most direct and important means of communication between intellectuals, noblemen, bureaucrats and merchants over short to long distances. As mentioned above, in the context of the Enlightenment, historians highlight that the *Republic of Letters*—a pan-European epistolary network—allowed people from all over Europe to be in touch with persons they had never physically met. Communications over the network aimed at exchanging information, discussing new theories, and critiquing ideas, to advance knowledge and technological progress (Mokyr, 2005, 2016).

Crucial to interactions across different locations was the possibility to communicate at moderate costs. As Mokyr (2016) argues, in fact, "[c]ultural innovation post-1500, from the Reformation to Copernican astronomy to calculus, depended on the printing press, better postal services, and the improved capability of ships and land-based modes of transport to spread information around the world" (cf. p. 42). Concerning postal services, from the 1530s the Taxis family—initially responsible for operating the Imperial post across Europe—extended postal services to private individuals for a fee. While the service was relatively reliable, rates varied substantially across routes and jurisdictions due to custom barriers, the availability of postal stages, and local privileges. Limited trust in these arrangements led monarchs in emerging nation-states to establish their own postal systems for official communication (Headrick, 2015). With the 1635 reform, England was the first to implement a public national postal system under state monopoly at relatively affordable rates (Brix, 2026).

According to Merton (1938, p. 580), the postal reform of 1635 was key for the development of science: "*Although a postal service for the conveyance of state letters had been extant since the beginning of the sixteenth century, it was not until the middle of the next century, that one general post office of England for private mail was established. And what was perhaps even more important, international postal relations were continually improved so that by the end of the century, communication with the Continent was constant and regular. These extensive developments in communication and transport [...] also promoted intercourse in the domain of thought. This increase of social interaction facilitated the development of science in no small measure.*" In this sense, the postal infrastructure formed part of the broader innovation ecosystem that made Britain "a society that provided both the incentives and the opportunities to apply existing useful knowledge to technology" (Mokyr, 2002, p. 55).

Along similar lines, Mokyr (2002, p. 58) points out that, although postal rates remained high and their structure complex prior to Rowland Hill's postal reform of 1840—most notably introducing the Uniform Penny Post—"postal services in England long before [1840] were providing easy and reliable access to knowledge generated elsewhere." An excerpt from a correspondence between the London-based botanist and physician Nehemiah Grew and the Italian scholar and physician Marcello Malpighi, today referred as

[9] According to Robinson (1948), "the 2-pence charge for a single letter was to remain the basic rate for two centuries. The increasing cost, according to the distance a letter traveled, was also to a be a long-lived device of the Post Office. The use of the footpost was to remain an important part of the system both in town and country" (cf. p. 31).

the founders of (plant and) microscopical anatomy—a prerequisite for advances in the fields of physiology, embryology, and practical medicine (Riva and Toffoletto, 2025)—illustrates this point: "*Most illustrious man, since you have indeed given me many occasions to write to you, and the most illustrious Lord Oldenburg has shown me the way to send letters, I could not (without prejudice to the decorum of my manners) hold my pen any longer. I see that almost all of your observations agree with mine; however, I have chosen to pass over some of them concerning the flower, fruit, and seed. [...] These anatomical observations on roots, which I have collected in the last year, together with figures, both corresponding to the size of the roots and microscopic ones, I am going to give in a short printer [...] I have also decided to preface the idea of Phytological Science, as outlined in our mind [...] Continue, most learned man, in your excellent undertakings, and strive to bind us and posterity more and more daily to your glorious memory.*" (see Grew to Malpighi, March 5, 1673; in Adelmann, 1975, pp. 646-47).[10]

According to scholars, the rise of an intense correspondence-based network made possible by the postal system facilitated the creation, circulation, evaluation, and refinement of what Mokyr (2002) terms *useful knowledge*: knowledge with practical applications, encompassing scientific principles, technical skills, artisanal expertise, and the institutional arrangements that support innovation. Furthermore, technical advances rest on incremental improvements to existing instruments and are not based only on entirely new inventions. The development of techniques that work was often the result of sustained interactions among scholars and skilled practitioners, including physicians, instrument makers, engineers, and artisans, as well as individuals involved in crafting norms and regulations of socio-economic activities.

The correspondence between the English educational, social, and agricultural reformer Samuel Hartlib and the English natural philosopher Robert Boyle illustrates the importance of extensive discussions about new knowledge, experiments, and technical innovations, but also about the individuals with useful knowledge and skills: "*I thank you for your new discoveries in anatomy, and enquiries of other useful and ingenious knowledges. I shewed Mr. Hobbes your letter, who liked it so well, that he desired me to lend it him, which I did. Your worthy friend and mine Mr. Gassend is reasonable well, and hath printed a book of the life and manners of Epicurus [...] in which I believe we shall have much of his own philosophy, which doubtless will be an excellent work. There is an experiment how to shew, as they suppose, that there is or may be vacuum. It were too long to write all the particulars, but in brief thus. They prepare a long tube like a weather-glass [... description of experiment and instruments used continued...].*" (see Hartlib to Boyle, May 19, 1648; Hunter et al., 2001, pp. 65-67).[11] The letter also illustrates the intensity of intellectual exchange sustained through correspondence: it references experimental findings, recently published works, philosophical debates, and technical descriptions of instruments. Such exchanges were not merely informational, as they facilitated the replication, adaptation, and improvement of instruments in England, where skilled mechanics were available to implement and refine new designs (see Hartlib to Boyle, July 10, 1657; Hunter et al., 2001, p. 220-22).[12]

The cumulative effects of these interactions were tangible. In 1659, Boyle and Robert Hooke—an

[10] An abstract summarizing the content of this letter is available at the following link: `https://emlo.bodleian.ox.ac.uk/profile/work/f41e567d-6def-4af0-8fe8-0f6df88ea0cb`.

[11] Letter available at the following link: `https://www.e-enlightenment.com/item/boylroPC0010065e1c`.

[12] Letter available at the following link: `https://www.e-enlightenment.com/item/boylroPC0010220a1c`.

exceptionally versatile natural philosopher active in physics, astronomy, geology, and meteorology—eventually constructed the air pump that would later bear Boyle's name. In fact, Hooke had earlier been recommended to Boyle by Hartlib, following a letter from the mathematician John Pell (see Pell to Hartlib, February 14, 1658; British Library, Manuscript Collections).[13] The Boleyan air pump became a crucial instrument in the experimental study of pneumatics. Together with related work by Otto von Guericke in Germany and Denis Papin in France, it contributed to advances in the understanding of air pressure and vacuum, developments that ultimately fed into the theory of steam power and the later application for the design of Newcomen's steam engine, subsequently improved by James Watt.

Besides these relevant and highly suggestive examples, whether and how the development of the postal system favored an intensification of communications in the context of the British Enlightenment, significantly boosted the connections and the interactions between knowers and doers across the affected locations, facilitated the production and spread of useful knowledge, and ultimately had a detectable impact on innovation activities and socio-economic outcomes in the longer term, are all open empirical questions, to which we turn next.

3 Data and Empirical Strategies

We build a geo-referenced panel dataset drawing on multiple primary and secondary sources. This section presents the main data sources, the procedures for the construction and treatment of the baseline data and the main empirical specifications. The collection and treatment of further data are presented in the corresponding sections below.

3.1 Data

We exploit spatial variation across grid cells measuring 0.09° in latitude by 0.09° in longitude (approximately 10×10 km at the equator), covering England and Wales. All the data are aggregated into a panel at the grid cell level spanning the decades 1570-1769.[14] The resulting dataset is a balanced panel of 2,764 grid cells observed over 20 decades, yielding a total of 55,280 grid cell-decade observations.[15] Data Appendix A provides additional details on the construction of all variables.

The Postal System. We extract information from several historical sources describing the evolution of the English postal system over time and reconstruct the distribution of the network's components from a map series (Carr, 1668; Greene, 1690; Edwards, 1987; Brayshay, 1991, 2011; Campbell-Smith, 2011). We geo-reference the location of every postal stage and build a digitized version of the network represented in each map. Pre-reform information covers the years 1570, 1580, 1595, 1603, and 1628. For the post-reform

[13] An abstract of the content of this letter is available at the following link: `https://emlo.bodleian.ox.ac.uk/profile/work/ad164c11-9815-4e1b-b1eb-06e4bf8f7a97`.

[14] Each decade begins in a year ending in zero: for example, the period 1620-1629 corresponds to the decade 1620.

[15] The smallest administrative unit in early modern England and Wales was the parish. By 1850, there were 17,898 parishes (see Kain and Oliver, 2001, for detailed information). This implies that, on average, each grid cell in our dataset encompasses approximately six parishes.

period, information is available for the years 1668, 1677, and 1690. The final grid cell-level panel dataset is obtained after first mapping all geolocated postal stages onto our $0.09° \times 0.09°$ grid map and then allocating digitized postal maps to each decade from 1570 to 1760.[16] We then define the indicator variable *Stage Dummy*$_{it}$, equal to one if the grid cell i hosts at least one postal stage in decade t, and zero otherwise. In addition, we construct the time-invariant indicator *Stage Dummy 1620*$_i$, which equals one for all decades if the grid cell i hosts at least one postal stage in 1628—the last observed network configuration in the pre-reform period—and zero otherwise.

Epistolary Communications. To obtain information on early modern correspondence in England and Wales, we draw on a major investment in the digital humanities made over the past decade (and still ongoing), aimed at collecting information on epistolary exchanges within the Republic of Letters. These collection efforts have so far produced two online platforms: Early Modern Letters Online (EMLO) and Electronic Enlightenment (EE). EMLO is a union of catalogues that collect person-to-person correspondence for the sixteenth-, seventeenth-, and eighteenth-century. The project, developed starting in 2009 by the Cultures of Knowledge Project at the University of Oxford, aims at tracking the epistolary exchanges in the context of the Enlightenment in the period 1550-1750.[17] Conceived in 1995 by the Bodleian Libraries at the University of Oxford, the EE project has assembled a wide-ranging digital collection of interlinked correspondence of central actors (not necessarily scholars), as well as more 'marginal' figures, to represent a cross-section of early modern and Enlightenment society.[18]

For each letter, we extract information on the date, place of origin, destination, sender (including bibliographical details), recipient (including bibliographical details), and either an abstract or the full text. We retrieve all letters sent from or received in England and Wales between 1570 and 1769. We retain only letters for which at least one location (either origin or destination) is identified.[19] The data is then enriched with geo-time information to make it suitable for spatial panel analysis. First, we geocode the origin and destination of each letter. When geographic coordinates are not provided by the source databases, we manually geocode locations using the platform OpenCage.[20] Next, we remove EE letters that also appear in EMLO, as well as duplicate entries in EE corresponding to original and translated versions of the same letter. This procedure yields 55,539 unique letters. For letters involving multiple senders and/or recipients, we treat each sender-recipient pair as a separate observation. This approach results in 55,778 letter-sender-recipient combinations (hereafter "letters" for simplicity). After excluding observations with either an unidentified sender or recipient (182 letters, corresponding to 185 sender-recipient combinations), the final dataset comprises 55,593 unique letter-sender-recipient combinations. In Sections 5.1 and 6.1 we describe in detail our classification of correspondents by occupation and the text analysis of the letters'

[16] More information on the construction of the postal dataset is available in Data Appendix A.1.

[17] For further details, see the Cultures of Knowledge Project website: `https://www.culturesofknowledge.org/`.

[18] More details on EE are available at the following websites: `https://www.e-enlightenment.com/index.html` and `https://digital.humanities.ox.ac.uk/project/electronic-enlightenment-project`.

[19] Data from EMLO were extracted via web-scraping procedures in Python, for which we obtained permission from the platform manager. The full EE dataset was obtained under a license agreement between the University of Bologna and the University of Oxford. We use the EMLO version updated in 2020 and the EE version updated in 2018.

[20] Link: `https://opencagedata.com/`.

content, respectively.

To construct the final panel dataset, we map the origin and destination of each letter to the $0.09° \times 0.09°$ grid, which is defined for each different decade. We then aggregate letters at the grid cell-decade level. For each grid cell-decade, we compute the number of outgoing letters (sent from the cell) and incoming letters (received in the cell). We sum these two quantities to obtain a measure of epistolary intensity for each grid cell-decade.[21]

Other Variables. Data on Roman roads is extracted from McCormick et al. (2013), while major medieval routes come from Edwards (1987). These sources are used to compute the variable *Distance to Historical Routes*, defined as the distance from each grid cell's centroid to the nearest segment of either network.[22] Proximity to medieval urban centers is tracked using data from Angelucci et al. (2022). For each grid cell, we compute the variable *Distance to Medieval Boroughs*, measuring the distance from the cell centroid to the nearest medieval mesne or royal borough.

3.2 Empirical Strategies

Preliminaries: Correlational Analysis. The relationship between the presence of a postal stage and correspondence intensity can be explored by estimating two-way fixed effects regressions of the following form using OLS:

$$Y_{it} = \beta StageDummy_{it} + \phi_1'(\mathbf{X}_i \times \eta_t) + \phi_2(StageDummy_{it} \times CivilWar_t) + \theta_i + \eta_t + \varepsilon_{it} \tag{1}$$

where Y_{it} represents the inverse hyperbolic sine of the number of incoming and outgoing letters in a grid cell-decade. Covariates are denoted by $\mathbf{X}_i$. The main coefficient of interest is β, the coefficient associated with $StageDummy_{it}$, which is an indicator taking the value one if grid cell i hosted at least one postal stage in decade t, and equal to zero otherwise. The specification includes grid cell fixed effects θ_i and decade fixed effects η_t. Grid cell fixed effects absorb time-invariant local characteristics, including historical, socio-economic, and political factors that may affect both the presence of a postal stage and our outcomes of interest. Decade fixed effects control for common shocks and aggregate trends that affect all locations simultaneously.

The estimates of the effect of time-varying exposure to postal stages do not necessarily lend themselves to a causal interpretation. For instance, concerning the analysis on the determinants of correspondence intensity, one cannot rule out that the establishment of a postal stage in a given location after the creation of the Royal Post for the public is endogenous. Moreover, time-varying confounders may simultaneously drive, e.g., changes over time in the intensity of correspondence recorded in the data and the establishment of a postal stage in a given location. Evaluating the likelihood of such biases is not straightforward. Possibly, this problem is less likely to arise when dealing with other outcomes of interest, such as letter flows across

[21] Double counting problems do not arise as long as origin and destination grid cells differ. However, 924 letters originate from and are delivered within the same grid cell, with approximately 86% of them being associated with the grid cell containing the City of London. Results are robust to excluding these letters.

[22] More details on how we generate this variable can be found in Data Appendix A.3.

different types of correspondents or measures of variation in the letters' content. Ultimately, while useful to explore the empirical patterns within locations over time, the use of time-varying information on the presence of a postal stage in each location and period is a strategy that is prone to such biases.

The Impact of the 1635 Reform: Difference-in-Differences Analysis. The 1635 reform of the postal system, which opened the pre-existing network to the public, provides a useful natural experiment. Prior to the reform, the network of postal stages had been developed primarily for military purposes, as a service of the Crown aimed at connecting London to the battlefronts as quickly as possible. Consequently, postal stages were not systematically located in large, commercial, or rapidly growing cities. The details of the reform implementation imply that the opening of the pre-existing stages to the public can be considered plausibly exogenous to local trends in epistolary activity. As a result, the location of these stages is, by construction, unlikely to be affected by potential selection issues in the letters written after the reform, such as differences in the types of correspondents or in the content of the letters.

We take advantage of the exogenous variation driven by the 1635 reform and adopt a simple difference-in-differences (DID) strategy, where we compare locations with differential exposure to the pre-reform postal infrastructure before and after the reform. We use this identification method to estimate the effect of the reform on a range of outcomes (e.g., intensity of epistolary communications, their contents, the level of innovation activity, etc.). More precisely, we follow a conservative approach and restrict the treatment group to locations where a postal stage was already in place right before the reform. Specifically, a location is considered treated in our setting if it hosted at least one postal stage according to the 1628 postal map (the latest map before the reform). Locations without a pre-existing stage as per the 1628 map act all as untreated units.[23] Following narratives on the rapid implementation of the reform (see, e.g., Hemmeon, 1912), we interpret the decade 1630 as the first post-treatment period (although the reform de facto affected outcomes only for the second half of that decade). This amounts to estimating through OLS different specifications of a DID regression model of the form:

$$Y_{it} = \delta(Stage\,Dummy\,1620_i \times Post_t) + \phi_1'(\mathbf{X}_i \times \eta_t) + \phi_2(Stage\,Dummy\,1620_i \times CivilWar_t) + \theta_i + \eta_t + \varepsilon_{it} \quad (2)$$

where the time-invariant indicator $Stage\,Dummy\,1620_i$ equals one if grid cell i hosted at least one postal stage in the decade 1620 (based on the postal map of 1628), and zero otherwise. $Post_t$ represents instead a time-specific dummy variable equal to one for all decades from 1630 onward, and zero otherwise. The baseline coefficient of interest is δ, which captures the impact of the postal reform. It measures the differential change in a given outcome after 1630 between locations that hosted a postal stage prior to the reform and those that did not.

The vector $\mathbf{X}_i$ includes a set of time-invariant geographic and historical controls (expressed in logarithms): minimum distance to major rivers, Roman roads, and medieval roads; distance to London; and distance to royal and mesne boroughs. These variables proxy for long-run connectivity and economic

[23] We expect this approach to generate a downward bias in the estimated effect of the reform, given the expansion of the network after 1635 that largely affected locations in the control group.

importance since the medieval period. Because these covariates do not vary over time, they are interacted with decade fixed effects, allowing for their influence on the outcomes to vary flexibly across decades. Finally, to account for potential disruptions to the postal infrastructure during the English Civil War, we additionally include an interaction between the indicator variable $Stage\,Dummy\,1620_i$ and a time-specific dummy equal to one only for the decades 1640 and 1650 ($CivilWar_t$). This term captures specific shocks to our outcomes during that period of conflicts, based on the postal network that was available in those decades.

Dynamic Impacts: Event Study. To assess the validity of the parallel trends assumption and to study the dynamic effects of the postal reform, we also estimate event-study specifications. Specifically, this is done by interacting the indicator $Stage\,Dummy\,1620_i$ with the full set of decade dummies. The decade 1620 is omitted and serves as the reference period. This specification allows us both to test whether treated and control locations followed similar trends prior to the reform and to trace the evolution of treatment effects over time. Specifically, we use OLS to estimate the following equation:

$$Y_{it} = \sum_{\tau \neq 1620} \delta_\tau (Stage\,Dummy\,1620_i \times \mathbb{1}\{t = \tau\}) + \phi_1'(\mathbf{X}_i \times \eta_t) + \phi_2(Stage\,Dummy\,1620_i \times CivilWar_t) + \theta_i + \eta_t + \varepsilon_{it} \quad (3)$$

Under the parallel trends assumption, the coefficients δ_τ corresponding to pre-reform decades ($\tau < 1630$) should be statistically indistinguishable from zero. Post-reform coefficients trace the dynamic response of correspondence and innovation outcomes to the opening of the postal system.

Clustering. Across specifications, standard errors are clustered at the grid cell level to account for serial correlation within locations. Given the fine spatial resolution of the data, we also report standard errors corrected for spatial correlation, following Conley (1999). Our baseline specification assumes a spatial correlation radius of 50 km, corresponding to approximately 100 neighboring cells (roughly half the area of the historical county of Yorkshire). We use a Bartlett kernel with linear distance decay, thereby allowing correlation to decline smoothly with geographic distance even among observations within the same cluster.

4 The Postal System and Intensity of Epistolary Communications

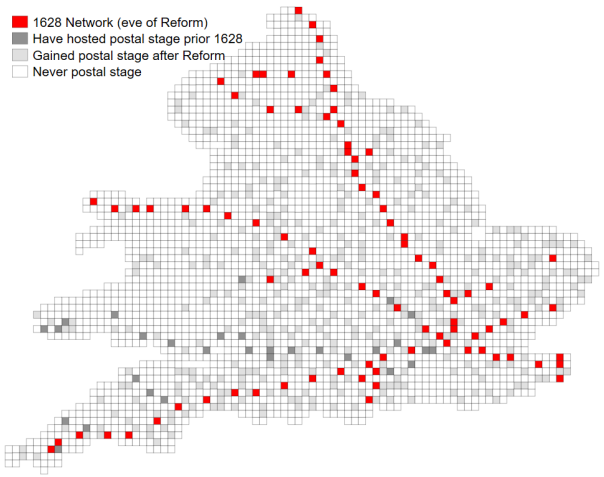
We start by investigating the impact of access to the postal network on spatial variation in patterns of interpersonal exchange through private correspondence across England and Wales. While the question may appear intuitive, its empirical answer is far from straightforward and has broader implications for understanding the role of communication infrastructures in enabling scale effects and intensifying connectivity across locations. Historical narratives suggest a significant effect of the postal system on the intensity of epistolary communications, but whether, how, and when the local access to this communication infrastructure impacts the intensity of communications, remain all important open questions.

We first study correlational patterns (Section 4.1) and then exploit the quasi-natural experiment associated with the postal reform of 1635 to estimate the effect of the reform, using a DID framework (Section 4.2). The section continues with a discussion of a set of sensitivity checks and an event study analysis to investigate the timing of the reform's impact (Section 4.3).

4.1 Descriptive Evidence

Figure 1 depicts the raw data about the spatial distribution of the postal stages. Grid cells marked in red indicate locations along the main routes of the King's Posts, where postal stages were active at the time of the 1635 reform. The population in these locations gained immediate access to the postal network following its opening to the public. Locations outside the early post roads with at least one postal stage established after the reform are depicted in light gray.

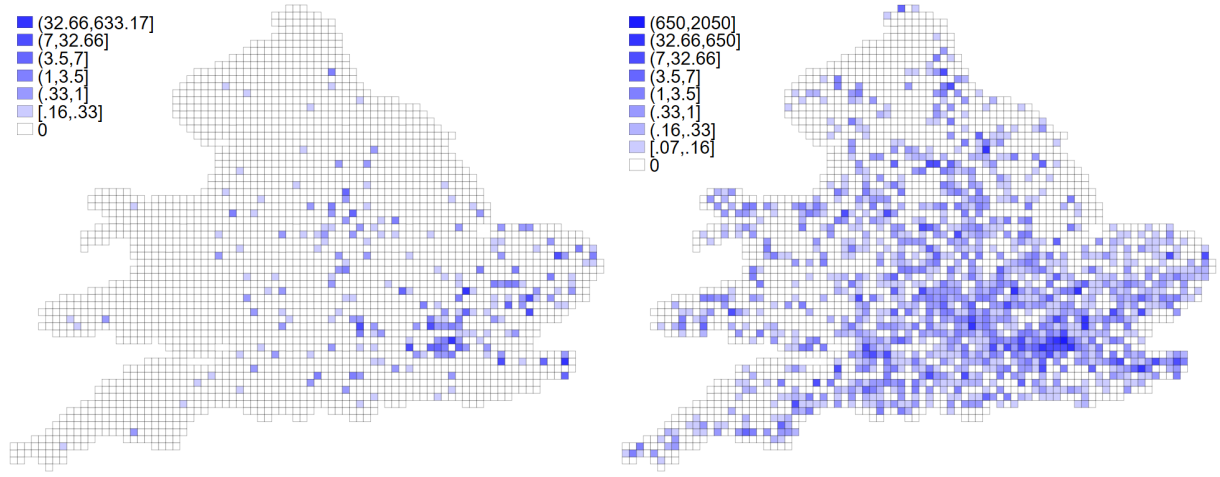
Figure 1: Postal Stages in England and Wales, Decades 1570-1760



Notes: Grid cells in red hosted at least one postal stage in 1628 that was opened to the public with the 1635 reform. Grid cells in dark gray hosted at least one postal stage in the decades before 1628 but were no longer active at the time of the reform. Grid cells in light gray correspond to postal stages established after the reform. Grid cells in white never hosted a postal stage during the entire observation period. *Sources*: own elaborations based on Carr (1668), Greene (1690), Edwards (1987), Brayshay (1991), Brayshay (2011) and Campbell-Smith (2011).

Figure 2 provides a summary of the intensity of epistolary communications across different locations, in the decades before the postal reform (Panel a) and after the reform (Panel b). Several patterns stand out. First, the intensity of letter exchanges tends to increase after 1635 across all areas of England and Wales, with a larger concentration of letters along the pre-reform routes. Second, the intensification of communications occurs both at the extensive margin—with many locations that remain inactive for correspondence—and at the intensive margin. Third, there is substantial local heterogeneity in the intensity of epistolary exchanges, even across nearby locations. Fourth, a subset of locations, among which most notably London, exhibits a disproportionately large number of letters already before the reform.

Figure 2: Epistolary Communications Before and After the 1635 Postal Reform



(a) Avg. Letters per Decade, 1570-1620

(b) Avg. Letters per Decade, 1630-1760

Notes: The figure shows the average total number of incoming and outgoing letters in each grid cell per decade, over the pre-reform period (Panel a) and the post-reform period (Panel b). *Sources*: Early Modern Letters Online and Electronic Enlightenment. See text and Appendix for details.

As a preliminary step, we assess the relative importance of a postal stage in a correlational exercise, where we regress the intensity of correspondence over time and space on the time-varying variable *Stage Dummy*$_{it}$. As illustrated in Figure 2, the distribution of the total number of letters across periods and locations contains many zeros and is highly right-skewed, reflecting the presence of observations with large values and excess variance. Therefore, we use *IHS(All Letters*$_{it}$*)* as our baseline outcome variable, defined as the inverse hyperbolic sine transformation of the total number of incoming and outgoing letters in grid cell i during decade t. This allows us to perform estimations in the full sample, mitigate the influence of grid cells showing disproportionately larger correspondence flows, and stabilize variance. Given the characteristics of this variable's distribution, we check the sensitivity of our findings to alternative transformations of the dependent variable (see Section 4.3).

The results, obtained by estimating different versions of Equation (1) through OLS, are reported in Table 1. Column (1) reports the results of unconditional specifications that exploit within-cell variation over time by including two-way fixed effects. Cell fixed effects account for time-invariant location-specific characteristics, which may affect communication intensity both before and after the introduction of a postal stage, while time fixed effects flexibly absorb nonlinear trends in the total number of letters across different decades. Given the peculiar role of the capital city also in the organization of the postal system, Column (2) controls also for distance to London, while Column (3) further conditions on distance to historical transportation routes. Given their role for long-term development documented in the literature, Column (4) also controls for distance to medieval boroughs. All distance measures are expressed in logarithms and interacted with decade fixed effects. Finally, Column (5) augments the specification with an interaction term between the variable *Stage Dummy*$_{it}$ and a dummy equal to one only in the decades affected by the English Civil War. This accounts for the presence of heterogeneous variation in the functioning of the postal network

during this specific period. Standard errors clustered at the grid cell level are reported in parentheses, while Conley (1999)'s spatially corrected standard errors are reported in squared brackets. The results document a positive and statistically significant coefficient on the postal stage indicator. Although the magnitude of the estimated effect tends to decline by increasingly conditioning on covariates, it remains remarkably stable across all specifications.

Table 1: Presence of a Postal Stage and Intensity of Correspondence – Correlational Estimates

Dependent Variable	IHS(All Letters)				
	(1)	(2)	(3)	(4)	(5)
Stage Dummy	0.111	0.099	0.095	0.096	0.091
	(0.024)***	(0.024)***	(0.023)***	(0.024)***	(0.024)***
	[0.015]***	[0.015]***	[0.015]***	[0.015]***	[0.015]***
Decade FE	✓	✓	✓	✓	✓
Cell FE	✓	✓	✓	✓	✓
Log Dist. to London × Decade FE		✓	✓	✓	✓
Log Dist. to Historical Routes × Decade FE			✓	✓	✓
Log Dist. to Medieval Boroughs × Decade FE				✓	✓
Stage Dummy × Civil War Dummy					✓
N. Clusters	2,764	2,764	2,764	2,764	2,764
N. Observations	55,280	55,280	55,280	55,280	55,280
Adj. R-Squared	0.398	0.403	0.404	0.406	0.407
Mean of DV	0.119	0.119	0.119	0.119	0.119

Notes: Observations are grid cells observed in each decade from 1570 to 1769. In each specification the dependent variable is the inverse hyperbolic sine (IHS) transformation of the total number of incoming and outgoing letters in each grid cell-decade. The explanatory variable *Stage Dummy*$_{it}$ equals one if the grid cell i includes at least one postal stage in decade t, and zero otherwise. OLS estimates. Standard errors clustered by grid cell are reported in parentheses, while Conley's standard errors accounting for spatial correlation up to 50 km with Bartlett kernel decay in space are reported in squared brackets. * = p-value $<10\%$, ** = p-value $<5\%$, *** = p-value $<1\%$.

4.2 The Impact of the 1635 Reform: Difference-in-Differences Analysis

To estimate the causal effect of the 1635 postal reform on correspondence intensity, we use the difference-in-differences strategy outlined in Equation (2). We focus on the coefficient of the interaction between the pre-reform postal stage indicator (*Stage Dummy 1620*$_i$, our treatment variable) and the post-reform dummy (*Post-1635*$_t$). OLS estimates of Equation (2) are presented in Table 2, where we gradually expand the set of covariates in the estimating equation as done in Table 1.

Table 2: Impact of the 1635 Postal Reform on Intensity of Correspondence – DID Estimates

Dependent Variable	IHS(All Letters)				
	(1)	(2)	(3)	(4)	(5)
Stage Dummy 1620 × Post-1635	0.309	0.284	0.262	0.255	0.229
	(0.077)***	(0.079)***	(0.077)***	(0.077)***	(0.080)***
	[0.039]***	[0.041]***	[0.040]***	[0.040]***	[0.041]***
Decade FE	✓	✓	✓	✓	✓
Cell FE	✓	✓	✓	✓	✓
Log Dist. to London × Decade FE		✓	✓	✓	✓
Log Dist. to Historical Routes × Decade FE			✓	✓	✓
Log Dist. to Medieval Boroughs × Decade FE				✓	✓
Stage Dummy 1620 × Civil War Dummy					✓
N. Clusters	2,764	2,764	2,764	2,764	2,764
N. Observations	55,280	55,280	55,280	55,280	55,280
Adj. R-Squared	0.399	0.404	0.405	0.407	0.407
Pre-1630 Mean of DV	0.039	0.039	0.039	0.039	0.039

Notes: Observations are grid cells observed in each decade from 1570 to 1769. In each specification the dependent variable is the inverse hyperbolic sine (IHS) transformation of the total number of incoming and outgoing letters in each grid cell-decade. The interaction variable *Stage Dummy 1620 × Post-1635* equals one in all grid cells containing at least one postal stage in the decade 1620 for all decades from 1630 onwards, and zero otherwise. OLS estimates. Standard errors clustered by grid cell are reported in parentheses, while Conley's standard errors accounting for spatial correlation up to 50 km with Bartlett kernel decay in space are reported in squared brackets. * = p-value $<10\%$, ** = p-value $<5\%$, *** = p-value $<1\%$.

A positive and statistically significant effect of the reform emerges consistently across specifications. To interpret this result, recall that, by construction, the control group includes locations that did not host a postal stage in 1628, but may have acquired one following the opening of the postal system in 1635. To the extent that the establishment of new postal stages after the reform increased correspondence activity in these locations (see Table 1), their inclusion in the control group of the DID setting biases the estimated treatment effect toward zero. From this perspective, the DID estimates reported in Table 2 can be interpreted as a lower bound of the true effect of the reform.

4.3 Sensitivity Analysis and Dynamic Effects of the 1635 Reform

We perform a battery of sensitivity checks to aid the interpretation of the results, assess their validity and robustness, and examine the dynamic effects of the postal reform of 1635. We briefly present the main findings below.

4.3.1 Measurement of the Dependent Variable

As discussed previously, the dataset features a large number of grid cells with zero values, a common characteristic in spatial settings. The distribution of letters across space and time is also characterized by overdispersion, due to some cells with a disproportionately large number of letters across periods (primarily the grid cell including the City of London).

As a baseline, we follow the literature and deal with excess variance by using an inverse hyperbolic

sine transformation of the total number of letters.[24] This transformation is similar to a logarithm and allows to retain zero-valued observations. A caveat of this strategy, which generally applies to any log-like transformation, is that the interpretation of the magnitude of the results is scale sensitive and it changes depending on the value of the constant term added to the variable of interest. Conceptually, increasing the constant reduces the relative contribution of the extensive margin (grid cells moving from zero to positive correspondence), thereby shifting weight toward variation in higher levels of correspondence. On the contrary, reducing the constant yields the opposite effect. Given the questions under investigation, we are mainly interested in unfolding the main empirical patterns rather than pinning down the exact magnitude of treatment effects.

Nonetheless, the literature suggests a range of robustness checks to assess the sensitivity of the results to the coding of the dependent variable and also to aid interpretation when both intensive and extensive margins are involved (see, e.g., Chen and Roth, 2024). In particular, all results are robust to varying the constant term between 0.001 and 8.[25] This is done for both the inverse hyperbolic sine transformation and the natural logarithm transformation.[26] The point estimates, reported in Appendix Figure B1, are larger for smaller values of the constant term and remain essentially unchanged for larger values of the constant. This pattern suggests that our baseline outcome variable provides a consistent estimate of the treatment effect. Moreover, Appendix Table B1 reports additional robustness checks. First, the results are confirmed when focusing on the mere presence of letters in a grid cell, thereby isolating the extensive margin (Columns 1 and 3). Second, our findings are unaffected when using the total number of letters as the outcome variable, addressing excess variance by winsorizing the distribution of these data at the 99th percentile (Columns 2 and 4).

Taken together, these results suggest that our findings are robust to alternative transformations of the outcome variable and are not driven by the relative weight assigned to the intensive or extensive margins. The magnitude of the effect is economically large: the specification in Column 4 of Appendix Table B1 shows that, relative to the control group, treated cells display an increase of 0.5 letters, corresponding to more than seven times the pre-reform mean of the outcome (0.070).

4.3.2 Implementation of the DID Analysis

The open access to postal stages that was enabled by the 1635 reform is conceptually orthogonal to the process generating the observed correspondence, including the collection of letters, the identity of correspondents, and the content of exchanges across macro-topics before and after the reform. Nevertheless, to assess the validity of our empirical strategy and the robustness of the results, we conduct a series of checks. These robustness exercises examine the sensitivity of the baseline estimates to: (i) the influence of specific locations on the selection of correspondence, (ii) the definition of treated locations, (iii) the composition of the treatment and control groups, and (iv) the use of optimal paths to historical battlefronts as an instrument for the treatment variable.

[24] The inverse hyperbolic sine function is defined as: $IHS(x) = \ln(x + \sqrt{x^2+1})$.

[25] The upper bound of 8 corresponds to the 98th percentile of the distribution of the total number of letters.

[26] These sensitivity checks are informative irrespective of the econometric specification being considered, and are performed focusing on both the impact of the presence of a postal stage and the DID estimates of the effect of the 1635 reform.

Selection of Correspondence in Relevant Locations. As described in Section 3.1, our analysis focuses on small grid cells within England and Wales, drawing on the universe of epistolary exchanges provided by two large-scale digital collections, i.e., EMLO and EE. These sources are not restricted to specific geographic areas and include thousands of correspondents across Europe, the Americas, and Asia from the early sixteenth to the mid-nineteenth century. Since these large-scale collections of historical correspondence span a wide range of locations and correspondents, the process by which letters are collected is unlikely to be systematically correlated with treatment status, reducing concerns about differential selection between treated and control locations. However, the presence of grid cells with disproportionately large numbers of letters, as well as the historical specificity of certain locations, warrants an assessment of their potential influence on the results. Appendix Table B2 reports the findings.

First and foremost, a disproportionate share of letters is concentrated in London, reflecting its role as both the capital city and the central hub of the postal network. Since the grid cell including London is classified as treated, it is important to verify that it does not drive the results. Dropping the grid cell hosting the City of London from the sample, or alternatively excluding all letters to or from London, marginally reduces the point estimate, but leaves the main result unchanged (Columns 1 and 2, respectively). Second, to address concerns about concentration among prolific correspondents, we exclude letters from the top 15 correspondents. The results remain similar (Column 3). Third, we consider locations hosting major universities and libraries, such as Oxford and Cambridge, where scholarly correspondence may be overrepresented in the digitized sources. In the baseline, these locations are classified as untreated and included in the control group, as they were not located along military routes and did not host postal stages prior to the reform. Reclassifying Oxford and Cambridge as treated makes the estimates slightly larger, suggesting that the baseline results are on the conservative side (Column 4).

Definition of Treated Locations. We next assess the robustness of the difference-in-differences estimates to alternative definitions of treated locations, i.e., the set of grid cells hosting postal stages that were opened to the public in 1635 and are therefore considered treated in the DID setting. As discussed in Section 3.1, identifying these locations requires reconstructing the evolution of the postal network over time, which necessarily involves some measurement choices. We evaluate their implications through a series of checks, reported in Appendix Table B3. First, regarding timing, the baseline specification defines the first post-treatment period as the decade 1630-1639, even though the reform was implemented in 1635. Redefining the first treated period as 1640–1649 yields similar results (Column 1). Second, some locations along the Bristol road (extending to Dale) hosted postal stages in 1603, but are not documented as active immediately before the reform, according to the postal map of 1628. In the baseline, we classify as treated only locations with confirmed active stages in 1628. Expanding the definition to include locations with postal stages active before 1628 strengthens the results and reduces estimation noise (Column 2).

Comparing Treated and Control Locations. We then examine the sensitivity of the results to alternative definitions of the control group. Results are reported in Appendix Figure B2, which also displays the baseline estimate for comparison, as per Column (5) in Table 2. First, we restrict the control group to

'never-treated' grid cells by excluding locations where postal stages were established after 1635. Although this reduces the sample size by approximately 16%, the estimated effect increases in magnitude. Second, we retain the full sample and apply entropy balancing weights (Hainmueller and Xu, 2013) to match the distribution (mean and variance) of key geographic covariates—distance to London, distance to medieval boroughs, and distance to historical routes—between treated and control units. The estimated coefficient decreases only slightly. Third, we further improve comparability by progressively restricting the sample to grid cells in higher quintiles of the entropy-weight distribution, thereby focusing on increasingly comparable subsets. The results remain quantitatively similar across all sub samples. Finally, we implement a synthetic difference-in-differences estimator that constructs counterfactual outcomes by weighting control units to match pre-treatment trends. The estimates are again virtually unchanged.

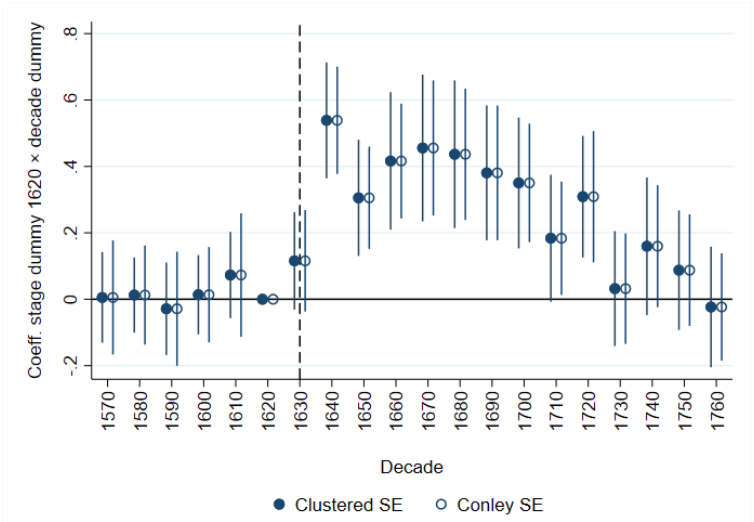
Optimal Routes to the Battlefronts. As a final check, we employ an instrumented difference-in-differences (DID-IV) design, which provides additional evidence on the determinants of treatment assignment and potential measurement error. The instrument exploits the arguably exogenous location of sixteenth-century battlefronts at the edges of the kingdom of England, which led the Crown to establish postal stages along the routes connecting London to these areas for military communication (see Section 2.1). Specifically, we construct least-cost paths from London to historical battlefronts and use a measure of proximity to these routes as an instrument for the presence of postal stages across England and Wales right before the reform.[27] This approach helps address potential non-classical measurement error in the geo-referencing of postal stages, due to possible inaccuracies in the postal map of 1628, and it provides an additional test of the identification strategy. The results, shown in Appendix Table B4, confirm the baseline findings and yield slightly larger but comparable estimates in magnitude.

4.3.3 Dynamic Effects: Event Study

To shed light on the dynamic effect of the opening of the postal system to the public over time, we perform an event study analysis. Specifically, we use a more flexible estimating equation relying on the interaction between the treatment variable (*Stage Dummy 1620*$_i$) and each decade dummy, using 1620 as the baseline period. Figure 3 presents the OLS results of this exercise, based on Equation (3). The results show no evidence of differential trends in correspondence intensity before 1635, when the postal system was not available for public use. The intensity of epistolary exchanges starts differentially increasing after 1635, with all the coefficients remaining positive and statistically significant up to 1720. In subsequent decades, the estimated treatment effects then gradually decline, consistent with the progressive expansion of postal stages into control areas, which induces attenuation bias.

[27] For more details on how we constructed the predicted network of least-cost paths, we refer the interested reader to Data Appendix A.1.

Figure 3: Dynamic Effects of the 1635 Postal Reform on Intensity of Correspondence – Event Study



Notes: Each dot represents the estimated coefficient of the interaction *Stage Dummy 1620* × decade fixed effect from Equation (3), using the specification in Table 2, Column 5. The dependent variable is the inverse hyperbolic sine transformation of the total number of incoming and outgoing letters in each grid cell-decade. The omitted interaction term is *Stage Dummy 1620* × decade 1620 fixed effect. OLS estimates. 90% confidence intervals based on standard errors clustered by grid cell (filled marker) or Conley's spatially-corrected standard errors (accounting for spatial correlation up to 50 km with Bartlett kernel decay in space, empty marker) are reported. The dashed vertical line indicates the decade of the postal reform.

All things considered, the findings shown in this section provide robust evidence on the sizable impact of the opening of the postal system on the intensity of epistolary communications in the affected locations of England and Wales. Taken together, the stability of our results to all the above robustness checks bolsters our confidence in the validity of the main DID strategy.

5 Person-to-Person Interactions: Scholars and Professionals

In this section, we study whether and how the reform had an impact on the patterns of communications of the conceptual categories of *knowers* (scholars) and *doers* (professionals).

5.1 Classification of Correspondents

We classify letters according to the occupations of their senders and recipients, using biographical information from EMLO and EE. The raw information on individuals' biographies and occupations varies across sources. Biographies in EE consist of structured lists of single-word occupational descriptors. We first compile the full list of distinct keywords appearing in EE biographies and group them into subcategories (e.g., "management," "state bureaucracy," "manufacturing").[28] Each correspondent is assigned to one or more subcategories based on the keywords listed in her biography. For EMLO correspondents, whose biographies are written in narrative form, we segment each biography into short text fragments and

[28]The Online Appendix reports the full list of keywords and their corresponding subcategories.

extract multi-word phrases that capture occupational information. These phrases are then submitted to the generative AI model Google Gemini for a preliminary classification, which we subsequently review manually to cross-check accuracy and ensure a coherent classification. For the 691 correspondents appearing in both EMLO and EE, we cross-check classifications to ensure consistency.[29]

Correspondents are then classified in macro-categories. For the knowers we focus on *scholars* (e.g., natural scientists, humanists, academics). For the doers we instead consider a set of categories that include: *practitioners*, encompassing occupations that apply theoretical knowledge to practical problems, including medicine and law (e.g., physicians, lawyers), business (e.g., managers), and technical fields (e.g., engineers, craftsmen); *statesmen* (e.g., public administrators, diplomats, judges, colonial agents, military officers, members of parliament); and individuals in the *information industry* (e.g., editors, publishers, booksellers, translators), given our focus on the diffusion of knowledge. Taken together, these categories are labeled as *professionals*. Finally, we identify a subset of correspondents belonging to the macro-category *clergy* (e.g., churchmen, Jesuits, missionaries). Importantly, the correspondents' classification is not mutually exclusive, as an individual may belong to more than one category. For example, Francis Bacon is classified as both a scholar and a statesman, as reflected from his recorded biography "lord chancellor and politician; philosopher; first baron verulam of verulam; viscount st alban". Individuals who cannot be assigned to any of the above macro-categories are placed in a residual category labeled *other*.[30]

For each individual, we construct a set of indicator variables. The dummy corresponding to a category s equals one if the correspondent is associated with category s, and zero otherwise. For each letter we construct a vector of dummy variables $d_{s_1 s_2}$, where s_1 and s_2 denote the occupational category of the sender and the recipient, respectively.[31] Appendix Table A1 reports the distribution of correspondents across macro-categories, overall and by source, as well as the share of letters involving at least a correspondent in a given macro-category.

Occupation-pair indicators are aggregated at the grid cell-decade level following the same procedure used to aggregate letters. These variables provide occupation-specific measures of local cross-occupational epistolary exchanges.[32]

[29]The dataset contains 9,765 distinct correspondents, of whom 8,205 are drawn from EMLO and 1,560 from EE. We identify duplicate correspondents by matching first name, surname, year of birth, year of death, and honorific titles (e.g., "Sir," "Lord," "Reverend"), when available. See Appendix Table A2 for the list of the 30 most active correspondents in our dataset and their classification.

[30]For example, the "natural philosopher and mathematician" Isaac Newton (classified as a scholar) used to correspond with his mother, Hannah Ayscough, whose biography reads: "Mother of Isaac Newton." Since no clear occupation appears in this biography, we assign Hannah Ayscough to the *other* category.

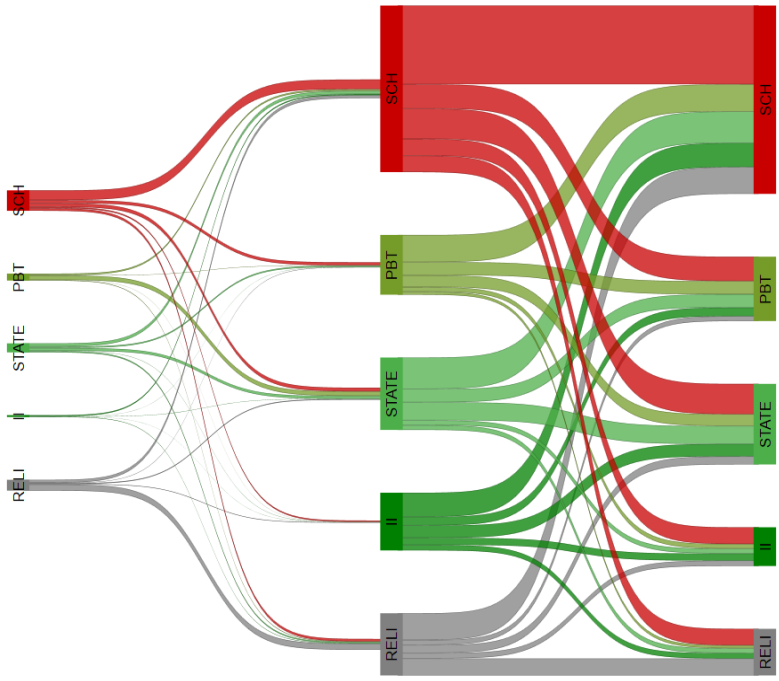
[31]More specifically, these indicators allow us to distinguish among three types of letter flows: (i) letters involving at least one correspondent belonging to macro-category s (i.e., $s_1 = s$ or $s_2 = s$) ; (ii) letters exchanged between correspondents belonging to the same macro-category (i.e., $s_1 = s_2$), and (iii) letters exchanged between correspondents belonging to different macro-categories (i.e., $s_1 \neq s_2$). In addition, for robustness checks, to specifically account for individuals exclusively assigned to a single macro-category, we construct occupation-pair indicators involving "pure" occupations. For instance, a physician who is neither engaged in academic activities nor affiliated with the clergy, state bureaucracy, or the information industry is classified as a "pure practitioner" (and therefore as a "pure professional").

[32]Letters sent from occupation s_1 to s_2 and letters sent from occupation s_2 to s_1 are combined into a single measure of interaction between occupations s_1 and s_2. When a correspondent holds multiple occupational classifications, the corresponding letter is included in all occupation-pair counts associated with those classifications.

5.2 Descriptive Evidence

Figure 4 illustrates the evolution of the average letter flows by occupational category of correspondents in the decades before and after the 1635 reform.

Figure 4: Letters by Occupation of Correspondents, Decades 1570-1760



Notes: The figure shows the average number of letters by occupational macro-category of correspondents per decade, before the 1635 postal reform (left half) and after the reform (right half). The width of bands indicates the average size of letter flows. In each half of the figure, the macro-categories of letter senders are reported on the left vertical axis, while those of letter recipients are reported on the right vertical axis. Correspondence involving *scholars* is reported in red. Epistolary communications involving the macro-category *professionals* are reported in different shades of green across the following sub-categories: practitioners, businessmen and tradesmen; statesmen; and individuals employed in the information industry. Finally, correspondence involving *clergy* is reported in gray. Figure made with Stata command `sankey`, version 1.81 (Naqvi, 2022a).

The raw data presented in the figure highlight several patterns. First, there is a sharp increase in the average number of epistolary exchanges per decade in the post-reform period, across all categories of correspondents. Two additional facts stand out, although they are less visible due to the relatively small number of exchanges before the postal reform. On the one hand, communications involving professionals (across all macro-categories) become relatively more frequent compared to, for example, those involving scholars. On the other hand, following the postal reform, correspondence between scholars and professionals increases both in absolute terms and as a share of total exchanges. While purely suggestive, these patterns in the raw data—averaged across all locations—point to a rise in the intensity of communications between knowers and doers after the reform.

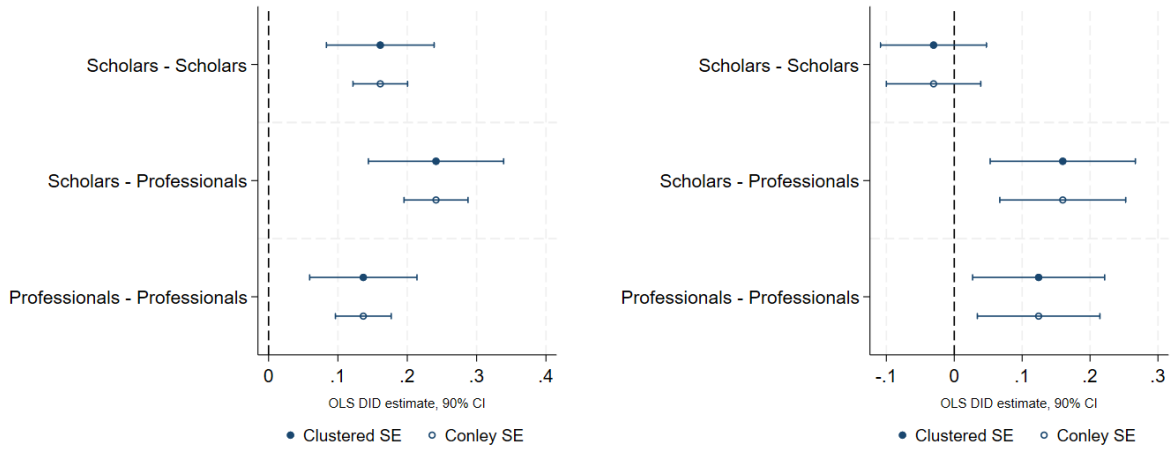
5.3 The Impact of the 1635 Reform on Interactions between Scholars and Professionals

Difference-in-Differences estimates. Figure 5 presents the results of the difference-in-differences analysis for the subset of epistolary communications involving scholars and professionals. Correspondence in which both the sender and the recipient are scholars is labeled "Scholars-Scholars". To examine interactions between the conceptual categories of knowers and doers, we also consider exchanges between a scholar and a professional (either as sender or recipient), labeled "Scholars-Professionals". Finally, we analyze communication flows between professionals, labeled "Professionals-Professionals". Each DID coefficient is estimated using our preferred specification, as per Column (5) in Table 2. The figure reports point estimates along with 90% confidence intervals.

Panel (a) shows OLS estimates using the inverse hyperbolic sine transformation of the total number of letters exchanged within each occupation-pair (indicated on the y-axis). The results suggest a significant increase in correspondence involving both scholars and professionals following the reform. This effect is particularly pronounced for exchanges between scholars and professionals, a pattern that remains robust to the exclusion of letters associated with the most prolific correspondents from the sample (see Appendix Table B7). To better understand this pattern, Panel (b) considers as the dependent variable the share of letters exchanged within each occupation pair relative to the total number of letters in each grid cell-decade. This specification is more demanding, as it restricts the analysis to grid cell-decades with positive correspondence, thereby reducing the sample size. In particular, observations with no recorded letters are excluded, since the share variable is undefined in those cases.[33] With this caveat in mind, the results show that the reform primarily increased the share of exchanges involving professionals, and in particular those between scholars and professionals. Taken together, these findings indicate that access to the postal system had a disproportionate effect on both the absolute intensity and the relative frequency of communication between scholars and professionals. These results are consistent with the broader patterns illustrated in Figure 4.

[33] An extensive sets of estimates of the impact of the postal reform on the intensity of correspondence and the shares of letters across different types of correspondents are reported in Appendix Tables B5 and B6, respectively.

Figure 5: Impact of the Reform on Correspondence between Knowers and Doers – DID Estimates



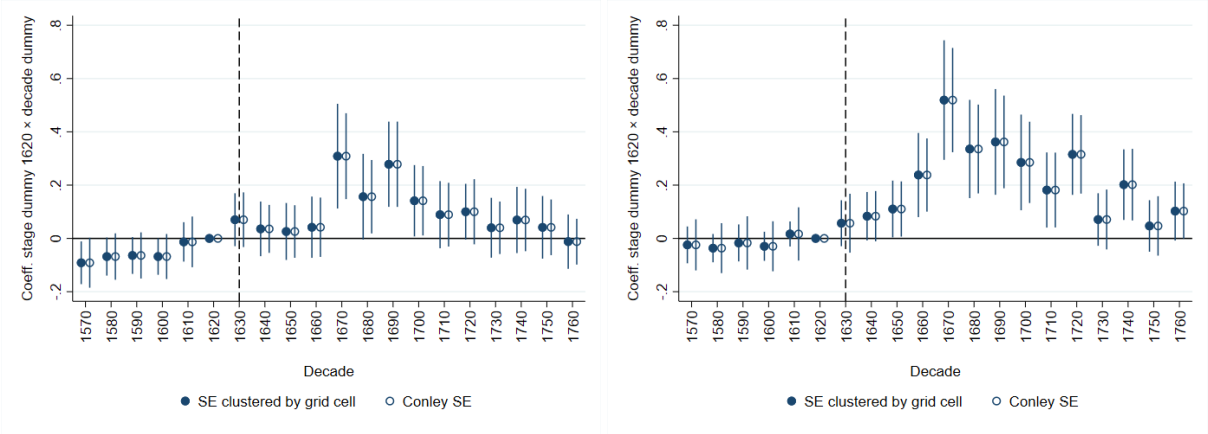
(a) INTENSITY OF COMMUNICATIONS – IHS(Letters)

(b) RELATIVE PATTERNS – Share of Letters

Notes: The figure shows the estimated coefficient of the interaction *Stage Dummy 1620* $\times$ *Post-1635* from Equation (2), using the specification in Table 2, Column 5. In Panel (a) the dependent variable is the inverse hyperbolic sine transformation of the total number of incoming and outgoing letters of a given occupation pair of correspondents in each grid cell-decade. In Panel (b) the dependent variable is the share of letters of a given occupation pair of correspondents over the total number of incoming and outgoing letters in each grid cell-decade. The labels on the y-axis indicate the specific subset of letters being considered to define the dependent variable, according to their senders and recipients. The macro-category *Professionals* indicates a single group of people including practitioners, statesmen, and people employed in the information industry. OLS estimates. 90% confidence intervals based on standard errors clustered by grid cell (filled marker) or Conley's spatially-corrected standard errors (accounting for spatial correlation up to 50 km with Bartlett kernel decay in space, empty marker) are reported.

Dynamic Effects: Event Study. Figure 6 plots the dynamic treatment effects of the postal reform on the intensity of letter flows involving either only scholars or scholars and professionals. We do not find significant differences in pre-trends of correspondence intensity across places with and without pre-existing postal stages, regardless of the specific type of interlocutor with whom scholars interacted. More interestingly, Panel (a) shows that exchanges between scholars increase gradually in the decades following the reform. By contrast, Panel (b) indicates that correspondence between scholars and professionals rises soon after the opening of the postal system, with larger and more persistent effects over time.

Figure 6: Correspondence involving Knowers and Doers (on any topic) – Event Study



(a) SCHOLAR → SCHOLAR

(b) SCHOLAR ↔ PROFESSIONAL

Notes: Each dot in a given subfigure represents the estimated coefficient of the interaction *Stage Dummy 1620* × decade fixed effect from Equation (3), using the specification in Table 2, Column 5. The dependent variable, constructed using letters from and to different categories of correspondents, is the inverse hyperbolic sine transformation of the total number of incoming and outgoing letters in each grid cell-decade. The omitted interaction term is *Stage Dummy 1620* × decade 1620 fixed effect. OLS estimates. 90% confidence intervals based on standard errors clustered by grid cell (filled marker) or Conley's spatially-corrected standard errors (accounting for spatial correlation up to 50 km with Bartlett kernel decay in space, empty marker) are reported. The dashed vertical line indicates the decade of the postal reform.

6 The Content of Epistolary Communications

In this section, we turn our attention to the analysis of the text of the letters and we estimate the impacts of the 1635 reform on the content of the epistolary communications.

6.1 Data and Text Analysis

We retrieved an abstract or the full text for 31,090 letters. All letters in EE include either an abstract or the full text, while EMLO reports abstracts for 18,462 letters. For an additional 2,514 EMLO letters without textual information, we complement the dataset using the Hartlib Papers project, a digital archive of correspondence involving Samuel Hartlib curated by the Digital Humanities Institute at the University of Sheffield.[34] Textual data are preprocessed and standardized following common practices in the text-analysis literature. Even restricting attention to communications with senders and/or recipients in England and Wales, the letters are written in multiple languages, such as English and Latin (but also Italian, French, German, Dutch, and Spanish). As a first step, we translate all text into English using Google Translate to ensure comparability across observations.[35]

To characterize the content of the correspondence, we implement a Latent Dirichlet Allocation (LDA) model (Blei et al., 2003). The algorithm identifies 25 latent topics based on word co-occurrence patterns.

[34] Link: `https://www.dhi.ac.uk/hartlib/`.

[35] Appendix A.2.2 provides details on translation and text cleaning procedures.

For each letter, the model returns the parameter γ_j, representing the share of topic j in the document.[36] To reduce dimensionality and facilitate interpretation, we aggregate these topics into five broader macro-topics that are labelled: (i) *Ordinary Life*, covering the topics business, child-rearing, romantic relationships, real estate, patrons, finances, friendship, and travels; (ii) *Knowledge*, including the topics practical knowledge, concepts and methods, experimenting, cataloguing, publishing, and pansophia; (iii) *Politics*, featuring the topics court affairs, military matters, civil war, and colonial activities; (iv) *Religion*, comprising the topics doctrines, Protestant unity (2 topics), and religious controversies; and (v) *Academic Gossip* (2 topics). We assess the coherence of this aggregation by examining both the text of randomly chosen letters within each topic, and the 30 letters with the highest γ_j for each topic. For this last subsample of letters we use the Large Language Model (ChatGPT 5.1) to summarize the content. Figure 8 presents the word clouds of the knowledge-related topics. Their summaries are reported in Appendix A.2.2. The word clouds and summaries of all remaining topics are reported in the Online Appendix.

To make the content suitable for the empirical analysis, for each letter we compute the score associated to each macro-topic (which is given by the sum of γ_j across topics belonging to the same macro-topic).[37] Each letter is assigned to the macro-topic with the highest aggregate score. We subsequently aggregate information at the grid cell-decade level. Specifically, for each grid cell-decade, we compute the inverse hyperbolic sine of the number of letters assigned to each macro-topic. We also construct occupation-pair-specific counts of letters by macro-topic, based on the same occupational classifications of senders and recipients presented above. This allows, in particular, to study the content of the communications involving scholars and professionals.

6.2 Descriptive Evidence

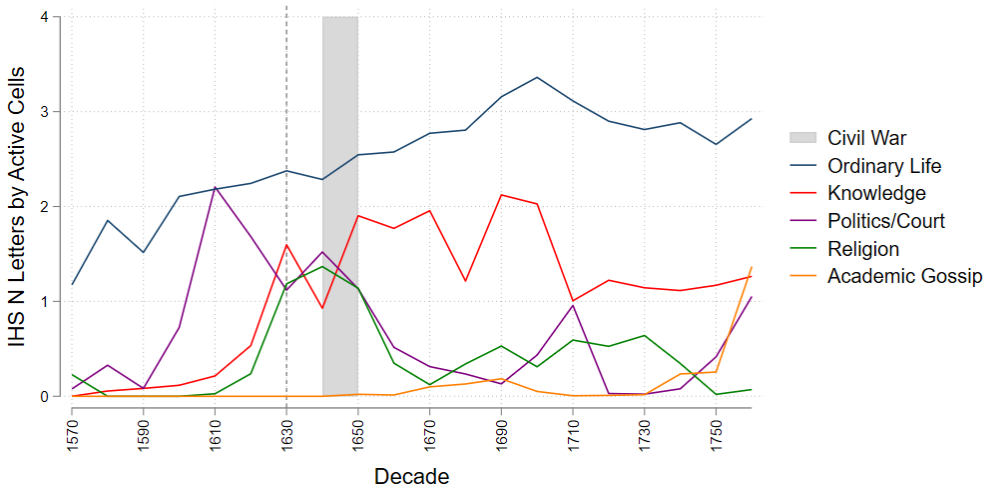
As a preliminary illustration of the raw data, Figure 7 reports the the inverse hyperbolic sine of the number of incoming and outgoing letters per active grid cell for each macro-topic across time. The gray vertical line indicates the decade of the postal reform, while the shaded gray area highlights the decades of the English Civil War. Four main broad patterns emerge. First, in relative terms, the main macro topic involves correspondences about ordinary life matters, which tend to increase steadily over time. Second, knowledge-related correspondence sharply increases from the 1630s, eventually replacing communications about politics and the court as the main topic (besides private ordinary life matters). Interestingly, and different from ordinary life communications, the relative number of letters with a predominant knowledge content tends to drop in the eighteenth century. Finally, the remaining macro-topics do not display specific trends and tend to present spikes associated with relevant historical events. For instance, political correspondence peaks during periods of political turmoil like the death of the Queen in 1603, the Civil War and the restoration, and at the time of major conflicts in the eighteenth century. Similarly, religion-related correspondence spikes during the decades surrounding the Civil War, consistent with the confessional disputes of this period. Finally, academic gossip reaches its maximum in the latest decades (e.g., the disputes

[36] Appendix Figure A3 reports the top ten words associated with each topic.

[37] Because we are particularly interested in interactions involving useful knowledge, our sum of γ_j for the macro-topic *Knowledge* discards the topic which specifically focuses on literature and arts. However, results (available upon request) are robust and similar in magnitude if we also include this topic in the *Knowledge* macro-topic.

surrounding Rousseau and Hume).

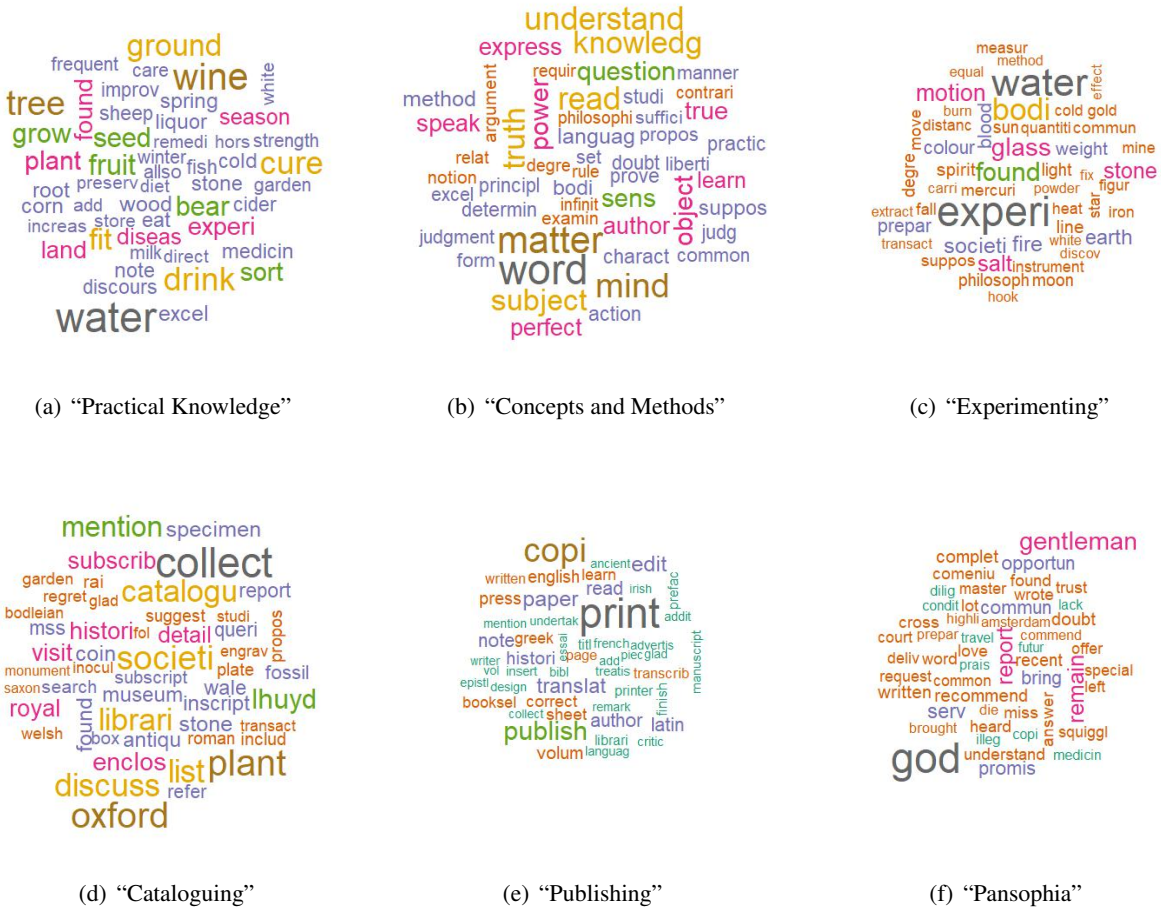
Figure 7: Correspondence Intensity by Macro-Topic



Notes: Each line represent the evolution of the inverse hyperbolic sine of the number of incoming and outgoing letters per active grid cell in each decade over the period 1570-1760 for the following five macro-topics: ordinary life (blue), knowledge (red), politics/court (purple), religion (green), and academic gossip (orange). The gray vertical line indicates the decade of the postal reform, whereas the gray shaded area highlights the decades of the English Civil War.

The *Knowledge* macro-topic includes six sub-topics. First, *Practical Knowledge* discusses agricultural, horticultural, botanical, and medical practices aimed at, among other activities, selecting crops, preparing drugs, and increasing health and wealth through the circulation of practical knowledge. Second, *Concepts and Methods* reports debates about general concepts such as the nature of mind, will, liberty, and language, and methodological guidance on how to reason properly and communicate knowledge effectively to facilitate its application. Third, *Experimenting* provides detailed descriptions of tests and scientific experiments, e.g., on gravity, the existence of vacuum, and the possibility of blood transfusion between animals. Fourth, *Cataloguing* reports discussions how to record and classify new specimens of, e.g., minerals, plants, fossils. Fifth, *Publishing* mentions the management of the publication process, including discussions on new manuscripts, appropriate translations, annotations, and the resolution of printing issues. Finally, *Pansophia* centers on the importance of disseminating scientific advances universally and educating all individuals to the importance of knowledge.

Figure 8: Wordclouds of Macro-Topic Knowledge



(a) "Practical Knowledge" (b) "Concepts and Methods" (c) "Experimenting"

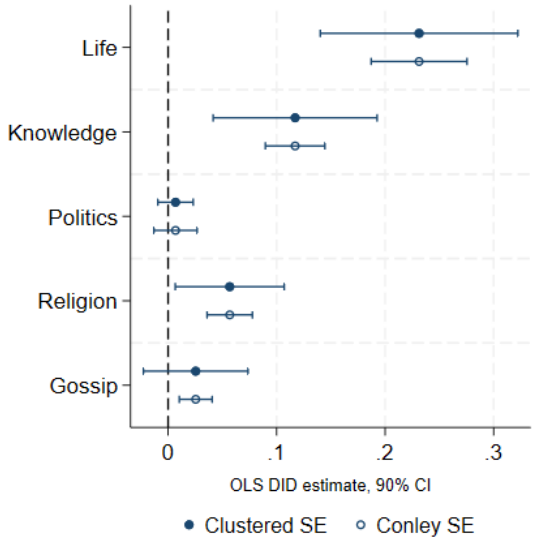
(d) "Cataloguing" (e) "Publishing" (f) "Pansophia"

Notes: Wordclouds of the knowledge-related sub-topics identified by applying the LDA algorithm to the corpus of letters with text over the period 1570-1769. In each panel, larger words are more recurrent.

6.3 Impact of the 1635 Reform on Macro-Topics

For each macro-topic we estimate the impact of the reform on the intensity of communications on that macro-topic. Figure 9 presents the results using the difference-in-differences specification outlined in Equation 2, including the full set of controls (as in Column 5 of Table 2). The dependent variable, indicated on the y-axis, is the inverse hyperbolic sine transformation of the number of letters on each macro-topic.

Figure 9: Impact of the Reform on Intensity of Correspondence by Macro-topic – DID Estimates



Notes: The figure shows the estimated coefficient of the interaction *Stage Dummy 1620* × *Post-1635* from Equation (2), using the specification in Table 2, Column 5. The dependent variable is the inverse hyperbolic sine transformation of the total number of incoming and outgoing letters assigned to a given macro-topic in each grid cell-decade. The labels on the y-axis indicate the specific macro-topic being considered to define the dependent variable. OLS estimates. 90% confidence intervals based on standard errors clustered by grid cell (filled marker) or Conley's spatially-corrected standard errors (accounting for spatial correlation up to 50 km with Bartlett kernel decay in space, empty marker) are reported.

The results document that the opening to the public of the pre-existing postal stages significantly increases private communications on ordinary life matters. Among the remaining macro-topics knowledge-related exchanges display the strongest response.[38] The impact on the reform on the remaining macro-topics tends to be lower and less significant. In the following Section we focus on these two prominent macro-topics and examine the interactions between scholars and professionals.

7 Knowledge-related Correspondence between "Knowers" and "Doers"

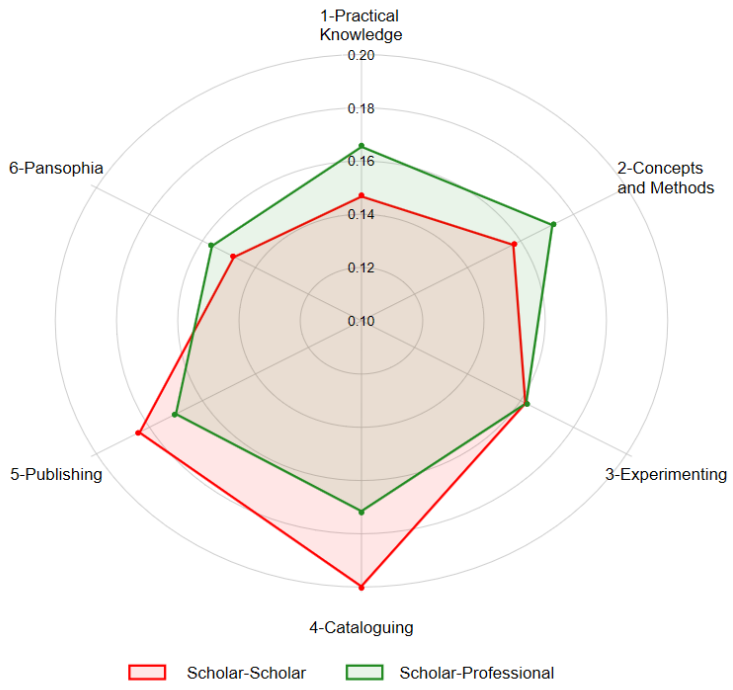
The historical literature emphasizes the role of communication infrastructures for the creation and diffusion of useful knowledge, as well as the importance of interactions between knowers and doers (Merton, 1938; Mokyr, 2002). Useful knowledge is broadly understood as the application of insights derived from the study of natural phenomena using experimental methods. For instance, the use of thermodynamic principles to build functioning vacuum pumps suggests how scientific understanding can be translated into practical innovations. Another example is systematic documentation, classification, and diffusion of information on the "facts that work", such as the development of botanical catalogs that contributed to improve agricultural productivity and drug production. More generally, person-to-person exchanges of information in the context of the Industrial Enlightenment have been portrayed as fundamental to the rise of cultural values that foster cooperation to enhance the production of useful knowledge, and make it universally accessible to

[38] Appendix Table B8 reports the full set of estimates underlying the results.

everybody. In this section, we push the econometric analysis toward its limits by focusing on the subset of knowledge-related communications and isolating the impacts of the postal reform on communications involving scholars and professionals.

Descriptive Evidence. Data availability limits the number of communications when disaggregated at the grid cell-decade-subtopic level. This prevents the possibility of relying on the econometric analysis exploiting within-cell and time variation to estimate the impacts of the reform on the content of communications between different types of correspondents at the subtopic level. Nonetheless, examining the raw data on subtopics is useful to understand the nature of knowledge-related communications involving scholars and professionals. Figure 10 presents a spider graph summarizing the content of knowledge-related correspondence across subtopics. In particular, the figure distinguishes between types of correspondents, reporting separately the composition of communications among scholars (red) and between scholars and professionals (green). Knowledge-related communications tend to span all sub-topics (see the overlapping area); however, correspondence among scholars is relatively more concentrated on publishing and cataloguing, while exchanges between scholars and professionals discuss more intensively practical knowledge, understanding concepts and methods, and how to disseminate knowledge to a wider audience.

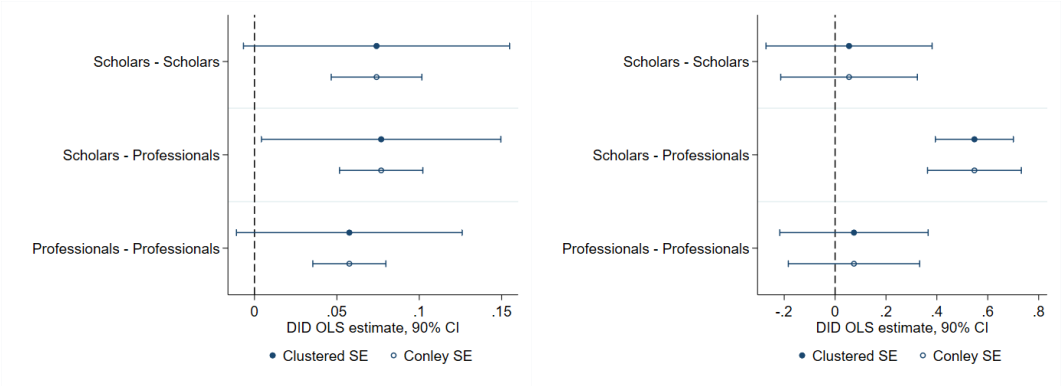
Figure 10: Knowledge-related Sub-Topics involving Scholars and Professionals



Notes: Average share of text devoted to discuss 6 knowledge-related topics in the correspondence between scholars (red) and scholars and professionals (green) over the period 1570-1769. The 6 topics are identified by applying the LDA algorithm to the corpus of letters with text and displayed on each spike of the spider plot. Figure made with Stata command `spider`, version 1.54 (Naqvi, 2022b).

Difference-in-Differences estimates of the effects of the reform. Figure 11 reports the estimated effects of the reform on the intensity of knowledge-related correspondence involving scholars and professionals, based on our preferred specification. In panel (a), the dependent variable is the log number of incoming and outgoing knowledge-related letters in each grid cell-decade, for the unordered occupation pairs shown on the y-axis. The estimated effect for exchanges between scholars and professionals is positive and statistically significant, while the estimates for scholars-scholars and professionals-professionals exchanges are less precisely estimated. Panel (b) replicates the analysis using, for each occupation-pair, the share of knowledge-related exchanges relative to the total knowledge-related communications. Although these specifications entail a substantial reduction in sample size, the results indicate that the reform disproportionately increased knowledge-related exchanges between scholars and professionals. Together with the descriptive evidence above, these findings suggest that the opening of the postal system facilitated discussions related to the creation and diffusion of useful knowledge in England and Wales.

Figure 11: Impact of the Reform on Knowledge-related Communications involving Scholars and Professionals – DID Estimates



(a) COMMUNICATION INTENSITY – IHS(Letters) (b) RELATIVE PATTERNS – Share of Letters

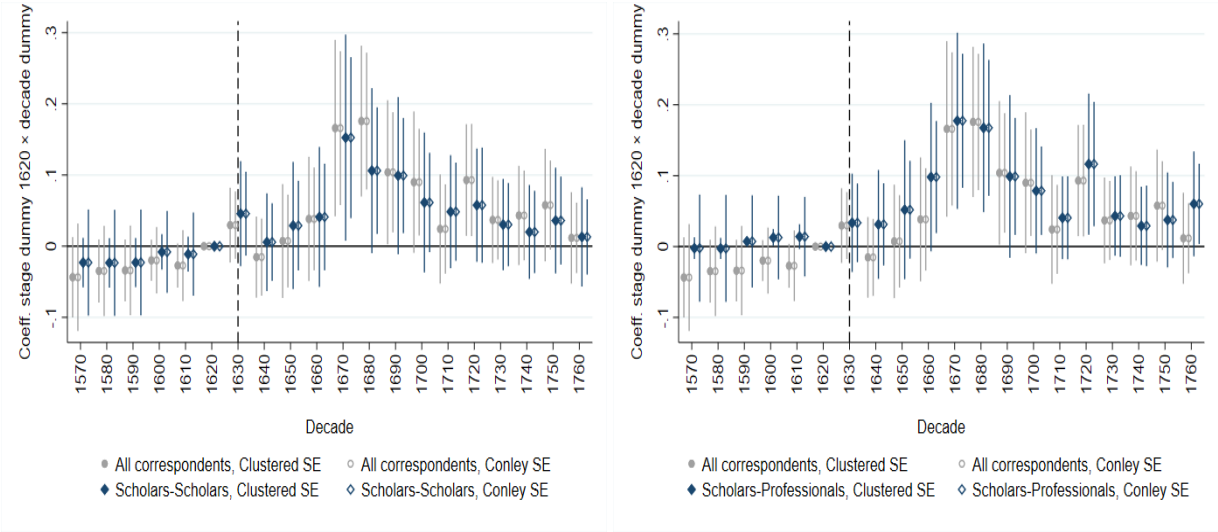
Notes: The figure shows the estimated coefficient of the interaction *Stage Dummy 1620* × *Post-1635* from Equation (2), using the specification in Table 2, Column 5. In Panel (a) the dependent variable is the inverse hyperbolic sine (IHS) transformation of the total number of incoming and outgoing letters assigned to the topic *Knowledge* in each grid cell-decade. In Panel (b) the dependent variable is the share of letters assigned to the topic *Knowledge* over the total number of incoming and outgoing letters in each grid cell-decade. The labels on the y-axis indicate the specific subset of letters being considered to define the dependent variable, according to their senders and recipients. The macro-category *Professionals* indicates a single group of people including practitioners, statesmen, and people employed in the information industry. OLS estimates. 90% confidence intervals based on standard errors clustered by grid cell (filled marker) or Conley's spatially-corrected standard errors (accounting for spatial correlation up to 50 km with Bartlett kernel decay in space, empty marker) are reported.

Dynamic effects of the reform. Figure 12 replicates the event-study analysis focusing on knowledge-related communications. In Panel (a), the estimates in blue refer to exchanges among scholars, while in Panel (b) they refer to exchanges between scholars and professionals. In both panels, the gray estimates capture exchanges involving all correspondents. Two main patterns emerge.

First, the dynamics documented in Figure 6 on all topics consistently emerge when restricting attention to knowledge-related communications. Second, compared to the baseline dynamic analysis for all

correspondents and topics reported in Figure 2, which shows positive and significant effects immediately after the opening of the postal system, knowledge-related communications involving scholars display a more gradual response. Understanding the reasons of this delay is not straightforward. One possible explanation is that, besides communications about ordinary life matters, topics related to politics and religion became particularly prominent during the Civil War, thereby reducing the likelihood that a letter is classified as mainly related to knowledge. Consistent with this interpretation, the event studies focusing on political and religious exchanges peak in the decades surrounding the war (see Figure B3).

Figure 12: SCHOLARS and PROFESSIONALS: Communications on KNOWLEDGE – Event Study



(a) KNOWLEDGE: Scholar-Scholar

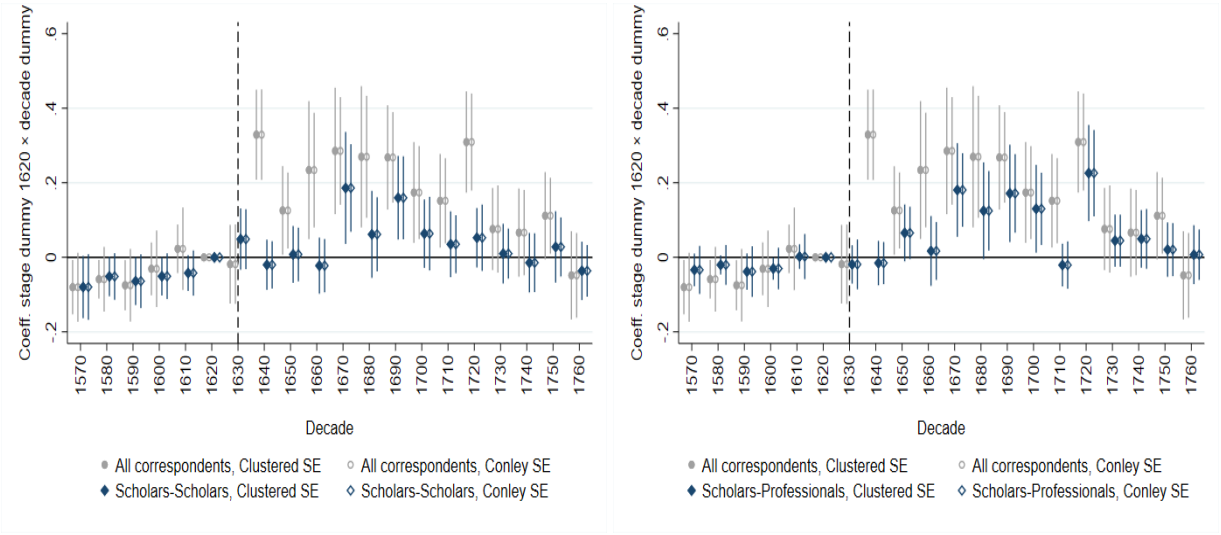
(b) KNOWLEDGE: Scholar-Professional

Notes: Each dot in a given subfigure represents the estimated coefficient of the interaction *Stage Dummy 1620* × decade fixed effect from Equation (3), using the specification in Table 2, Column 5. In both panels, estimates reported in gray use as dependent variable the inverse hyperbolic sine transformation of the total number of incoming and outgoing letters assigned to the macro-topic *Knowledge* in each grid cell-decade. In Panel (a), estimates reported in blue focus only on correspondence between scholars, whereas in Panel (b) they focus only on correspondence between scholars and professionals. For each set of estimates the omitted interaction term is *Stage Dummy 1620* × decade 1620 fixed effect. OLS estimates. 90% confidence intervals based on standard errors clustered by grid cell (filled marker) or Conley's spatially-corrected standard errors (accounting for spatial correlation up to 50 km with Bartlett kernel decay in space, empty marker) are reported. The dashed vertical line indicates the decade of the postal reform.

A second, non mutually exclusive explanation for the delayed response is that communications between individuals without pre-existing personal ties may take time to develop. In this respect, communications on ordinary life matters provide a useful conceptual counterfactual. Besides involving a relative majority of exchanges, these letters mostly concern household matters or interactions between relatives and friends, and thus are more likely to occur between individuals with established personal ties. The results, depicted in Figure 13, provide support for this interpretation. Communications related to ordinary life by all correspondents, shown in gray, increase immediately following the opening of the postal system, consistent with a rapid expansion of exchanges along existing social ties. By contrast, the response of ordinary life

communications involving scholars appears more gradual. Taken together with the delayed increase in knowledge-related correspondence, these patterns are consistent with the possibility of a more gradual formation of new connections among scholars and between scholars and professionals, facilitated by the reform.

Figure 13: SCHOLARS and PROFESSIONALS: Communications on ORDINARY LIFE – Event Study



(a) ORDINARY LIFE: Scholars-Scholars

(b) ORDINARY LIFE: Scholar-Professional

Notes: Each dot in a given subfigure represents the estimated coefficient of the interaction *Stage Dummy 1620* × decade fixed effect from Equation (3), using the specification in Table 2, Column 5. In both panels, estimates reported in gray use as dependent variable the inverse hyperbolic sine transformation of the total number of incoming and outgoing letters assigned to the macro-topic *Ordinary Life* in each grid cell-decade. In Panel (a), estimates reported in blue focus only on correspondence between scholars, whereas in Panel (b) they focus only on correspondence between scholars and professionals. For each set of estimates the omitted interaction term is *Stage Dummy 1620* × decade 1620 fixed effect. OLS estimates. 90% confidence intervals based on standard errors clustered by grid cell (filled marker) or Conley's spatially-corrected standard errors (accounting for spatial correlation up to 50 km with Bartlett kernel decay in space, empty marker) are reported. The dashed vertical line indicates the decade of the postal reform.

8 Innovation Activity

The results thus far have documented the impact of the 1635 reform of the postal system on the intensity of interactions between scholars and professionals on knowledge-oriented topics. As a final step, we examine whether these changes translated into differences in innovation activity across locations with varying access to the communication infrastructure. Empirical research on historical innovation typically relies on patent data, which provide an objective measure of inventions formally protected by intellectual property rights, particularly during the Industrial Revolution. In our setting, patent records offer a useful indicator of innovation over the long run, covering the two centuries following the reform. However, patenting activity was relatively rare before the eighteenth century. To assess the shorter-term effects of the reform—especially during the period of the Industrial Enlightenment—we therefore complement patent data with a newly

assembled dataset of innovation events drawn from the biographical records of individual innovators.

Patents in Period 1672-1852. To provide a first assessment of the long-run effects of the postal reform on innovation, we look at the total number of agricultural patents registered in a parish between 1672 and 1852, as compiled by Dowey (2013). We take these data from Heldring et al. (2021). From the same source, we also retrieve the average tax revenue per taxpayer and an indicator for whether a monastery owned a manor, both measured at the parish level. These variables are mapped onto our $0.09^\circ \times 0.09^\circ$ grid by assigning parish centroids to the corresponding grid cells. We estimate the cross-sectional relationship between the presence of postal stages right before the 1635 reform—the treatment variable from the DID setting outlined in Equation (2)—and the number of patents in the following period across grid cells. The results in Table 3 document a positive relationship between long-run patenting activity and the open access to the postal system brought about by the 1635 reform, regardless of the specification being considered. These patterns provide first suggestive evidence that locations treated by the 1635 postal reform are associated with higher levels of innovation over time.

Table 3: Postal Stages that Opened to the Public in 1635 and Number of Patents in the Long Run

Dependent Variable	IHS(No. Patents, 1672-1852)			
	(1)	(2)	(3)	(4)
Stage Dummy 1620	0.284	0.270	0.261	0.229
	(0.093)***	(0.094)***	(0.092)***	(0.093)**
	[0.078]***	[0.079]***	[0.079]***	[0.077]***
County FE	✓	✓	✓	✓
Dissolution Monastery	✓	✓	✓	✓
Lay Subsidy Revenue	✓	✓	✓	✓
Log Dist. to London		✓	✓	✓
Log Dist. to Historical Routes			✓	✓
Log Dist. to Medieval Boroughs				✓
N. Clusters	47	47	47	47
N. Observations	2,273	2,273	2,273	2,273
Adj. R-Squared	0.023	0.037	0.038	0.059
Mean of DV	0.107	0.107	0.107	0.107

Notes: Observations are grid cells. The dependent variable is the inverse hyperbolic sine (IHS) transformation of the total number of agricultural patents that were registered to people living in a parish between 1672 and 1852. The indicator *Stage Dummy 1620* equals one if the grid cell hosted at least one postal stage in the decade 1620, and zero otherwise. OLS estimates. Standard errors clustered by county are reported in parentheses, while Conley's standard errors accounting for spatial correlation up to 50 km with Bartlett kernel decay in space are reported in squared brackets. * = p-value $<10\%$, ** = p-value $<5\%$, *** = p-value $<1\%$.

Innovations Events in Period 1570-1769. To move beyond cross-sectional variation and examine dynamic effects of the postal reform, we build a geo-referenced dataset of innovators active in England and Wales during the early modern period. We extract biographical information from the *Oxford Dictionary of National Biography* (henceforth ODNB), the most comprehensive compendium of notable individuals. The dataset is assembled in four steps. First, we perform a full-text search of all biographies of individuals born between 1520 and 1799 using keywords such as "create", "construct", "invent", "innov", "improv",

"patent", and "devise". These terms serve as markers of narratives potentially describing innovative activity. This initial screening yields 6,441 biographies. Second, for each biography, we extract short text excerpts surrounding the keywords and manually review them to determine whether the individual can reasonably be classified as an innovator.[39] We distinguish between two categories: (i) *technical innovators*, defined as individuals engaged in innovative activities in science, technology, engineering, or mathematics; and (ii) *soft innovators*, defined as individuals contributing to innovations in the arts and humanities. Third, for individuals in both categories, we extract information on key life events and activities, including dates and locations of birth and death, residence, education, temporary visits, inventions, collaborations, memberships in societies and associations, apprenticeships, publications, and honors (e.g., knighthood). Finally, we geocode each recorded location and map it onto our grid at decadal frequency, thereby linking information on innovation activity to our dataset on letters and postal stages.[40]

This procedure allows us to identify 363 technical innovators and 216 soft innovators.[41] Leveraging this information, we construct three outcome variables at the grid cell-decade level: (i) the number of innovators present in a given cell-decade (irrespective of birthplace); (ii) the number of innovation-related events involving these individuals; and (iii) the number of recorded inventions. The first measure simply captures the presence of innovators in a given grid cell-decade and may exceed one if multiple individuals are active in the same grid cell-decade. The second variable accounts for the number of different activities performed by innovators in each grid cell-decade. We apply the inverse hyperbolic sine transformation to all outcomes to account for zero-valued observations and compute them for the full sample of innovators, as well as separately for technical and soft innovators. Appendix Figure B4 shows the average number of events associated with all innovators in each grid cell before and after the 1635 postal reform.

Table 4 presents estimates obtained by replicating the difference-in-differences specification in Equation (2). In the baseline, we focus on innovation-related events.[42] Columns (1) and (2) report estimates for overall innovation activity. Columns (3) to (6) show that the positive and statistically significant effect of the postal reform is driven entirely by technical innovators, with no detectable impact on soft innovators.

[39] Excerpts consist of the 40 characters preceding and following each keyword occurrence.

[40] Geolocation is conducted using the website Geonames (https://www.geonames.org/). For locations not available through Geonames, we manually identify coordinates using additional historical and online sources.

[41] Examples include Sir Thomas Aylesbury (1576–1657), who participated in trials of new machinery designed to improve coinage production through stamping rather than hammering, and Robert Boyle (1627–1691), one of the founders of modern chemistry, best known for formulating Boyle's Law, which describes the inverse relationship between pressure and volume in a gas at constant temperature within a closed system. Examples of soft innovators include Sir Nathaniel Bacon (1585–1627), an English painter who experimented with varnishes and pigments and developed the yellow pigment known as 'Pinke', widely adopted by contemporaries such as Peter Oliver and John Hoskins; and Richard Bentley (1662–1742), a classical scholar, critic, philologist, and theologian, credited with laying the foundations of the English school of Hellenism.

[42] Results using alternative outcome definitions are reported in Appendix Table B9 and convey a similar message.

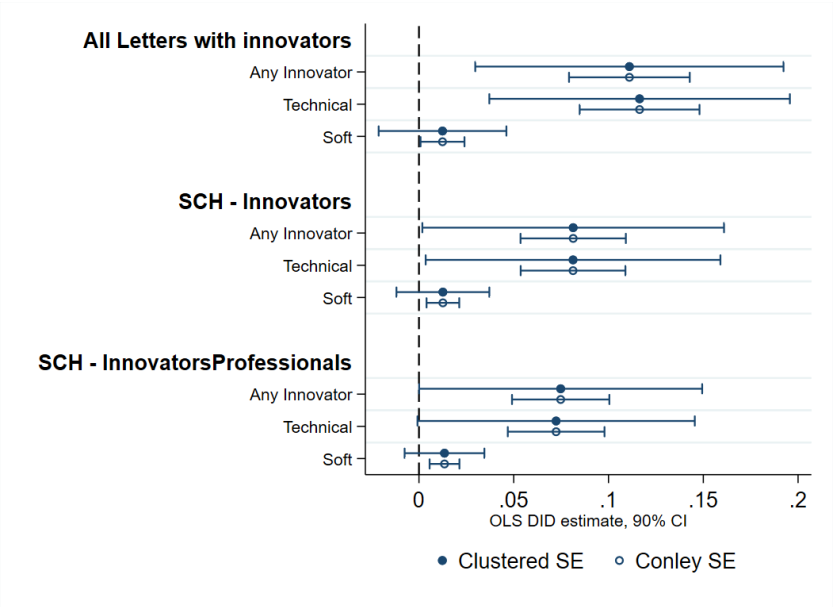
Table 4: Impact of the 1635 Postal Reform on Innovation Events – DID Estimates

Dependent Variable	IHS(Events Involving Innovators)					
	All		Technical		Soft	
	(1)	(2)	(3)	(4)	(5)	(6)
Stage Dummy 1620 × Post-1635	0.047	0.044	0.058	0.058	-0.006	-0.009
	(0.025)*	(0.026)*	(0.019)***	(0.020)***	(0.014)	(0.014)
	[0.016]***	[0.017]***	[0.014]***	[0.014]***	[0.010]	[0.010]
Decade FE	✓	✓	✓	✓	✓	✓
Cell FE	✓	✓	✓	✓	✓	✓
Log Dist. to London × Decade FE		✓		✓		✓
Log Dist. to Historical Routes × Decade FE		✓		✓		✓
Log Dist. to Medieval Boroughs × Decade FE		✓		✓		✓
Stage Dummy 1620 × Civil War Dummy		✓		✓		✓
N. Clusters	2,764	2,764	2,764	2,764	2,764	2,764
N. Observations	55,280	55,280	55,280	55,280	55,280	55,280
Adj. R-Squared	0.463	0.465	0.459	0.461	0.566	0.567
Pre-1630 Mean of DV	0.016	0.016	0.013	0.013	0.006	0.006

Notes: Observations are grid cells in each decade from 1570 to 1769. The dependent variables are: the inverse hyperbolic sine (IHS) transformation of the number of events involving all innovators (Columns 1 and 2), only technical innovators (Columns 3 and 4), and only soft innovators (Columns 5 and 6). In all columns, events exclude the innovator's birth in a given grid cell-decade. The interaction variable *Stage dummy 1620 × Post-1635* equals one in all grid cells containing at least one postal stage in the decade 1620 from 1630 onwards, and zero otherwise. OLS estimates. Standard errors clustered by grid cell are reported in parentheses, while Conley's standard errors accounting for spatial correlation up to 50 km with Bartlett kernel decay in space are reported in squared brackets. * = p-value $<10\%$, ** = p-value $<5\%$, *** = p-value $<1\%$.

As a final exercise to aid interpretation, we link the epistolary correspondence data with the innovation dataset by identifying individuals appearing in both sources. We match individuals based on full names and, for each category of innovator (any, technical, and soft), we compute: (i) the number of letters involving at least one innovator; (ii) the number of letters exchanged between a scholar and an innovator; and (iii) the number of letters exchanged between a scholar and an innovator who is also a practitioner, a statesman, or an individual working in the information industry. Although based on a relatively small sample, the results in Figure 14 further qualify our findings by showing that the opening of the postal system led to a significant increase in interactions between scholars and technical innovators.

Figure 14: Impact of the 1635 Postal Reform on Communications involving Innovators – DID Estimates



Notes: The figure shows the estimated coefficient of the interaction *Stage dummy 1620* × *Post-1635* from Equation (2), using the specification in Table 2, Column 5. The dependent variable is the inverse hyperbolic sine transformation of the number of incoming and outgoing letters in each grid cell-decade. The labels on the y-axis indicate the specific subset of letters being considered to define the dependent variable, according to their senders and recipients. The macro-category *Professionals* indicates a single group of people including practitioners, statesmen, and people employed in the information industry. OLS estimates. 90% confidence intervals based on standard errors clustered by grid cell (filled marker) and Conley's spatially-corrected standard errors (accounting for spatial correlation up to 50 km with Bartlett kernel decay in space, empty marker) are reported.

9 Concluding Remarks

The empirical analysis in this paper identifies the impact of postal infrastructure on patterns of interpersonal communications between "knowers" and "doers" in early modern England and Wales. The econometric analysis relies on a novel panel dataset with geo-temporal information on both the intensity and the content of epistolary communications between different types of correspondents over the period 1570-1769. We leverage a quasi-natural experiment whereby different locations gained differential access to the the postal network following the 1635 reform that opened the system to the public. The results document, in particular, significant increases in the intensity of knowledge-related communications. These findings support the view that the early development of a national postal system open to the public contributed to the emergence of a knowledge-based economy during the British Industrial Enlightenment.

The analysis of epistolary communications across topics and occupational categories of correspondents unfolds several interesting patterns. In particular, we find that access to the postal system disproportionately increased both the intensity and the relative frequency of communications between scholars and professionals. These results support the argument put forward by historians on the key role of person-to-person information exchanges between scholars and professionals for the creation and diffusion of applied—or *useful*—knowledge. The dynamic analysis further suggests that access to epistolary communication was particularly valuable for professionals. The results also document that, compared to

communications regarding ordinary life matters, knowledge-related exchanges exhibit a delayed response to the reform but more persistent effects over time, particularly for interactions involving professionals. These patterns are corroborated by evidence from innovation data and are further supported by linking correspondence to renowned technical innovators.

The data and results also point to several promising avenues for future research, although pursuing them would require substantial additional data collection and processing. While our analysis focuses on interactions between scholars and professionals, the raw data and the unsupervised text analysis reveal extensive communication involving other social groups, such as the clergy, as well as topics beyond knowledge, including politics and religion. Investigating these dimensions—potentially using supervised text analysis—falls outside the scope of this paper but represents a natural extension. Our findings also complement a growing body of evidence on the role of access to codified knowledge (e.g., books and, later, scientific journals) and on the importance of universities and academic markets in facilitating direct interactions among scholars. Exploring how these different channels of knowledge diffusion interact across the various actors involved in the production of useful knowledge remains a fruitful direction for future research. Finally, while this paper focuses on England and Wales, our results raise broader questions about the role of communication infrastructures and knowledge exchange in shaping economic development across pre-industrial Europe.

References

Acemoglu, D., J. Moscona, and J. A. Robinson (2016). "State capacity and American technology: Evidence from the nineteenth century". *American Economic Review 106*(5), 61–67.

Adelmann, E. B. (1975). The correspondence of marcello malpighi. Available at: `https://bibdig.museogalileo.it/tecanew/opera?bid=333274_2&seq=222`, accessed February 11, 2026.

Allen, R. C. (2001). The great divergence in european wages and prices from the middle ages to the first world war. *Explorations in economic history 38*(4), 411–447.

Allen, R. C. (2009). *The British industrial revolution in global perspective*. Cambridge University Press.

Almelhem, A., M. Iyigun, A. Kennedy, and J. Rubin (2026). Enlightenment ideals and belief in progress in the run-up to the industrial revolution: A textual analysis. *The Quarterly Journal of Economics, 141*(1), 263–314.

Aneja, A. and G. Xu (2024). "strengthening state capacity: Civil service reform and public sector performance during the gilded age". *American Economic Review 114*(8), 2352–2387.

Angelucci, C., S. Meraglia, and N. Voigtländer (2022). How merchant towns shaped parliaments: From the norman conquest of england to the great reform act. *American Economic Review 112*(10), 3441–87.

Beale, P. (1998). *A history of the post in England from the Romans to the Stuarts*. Routledge.

Blei, D. M., A. Y. Ng, and M. I. Jordan (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research 3*(Jan), 993–1022.

Bouscasse, P., E. Nakamura, and J. Steinsson (2025). When did growth begin? new estimates of productivity growth in england from 1250 to 1870. *The Quarterly Journal of Economics, 140*(2), 835–888.

Brayshay, M. (1991). "Royal post-horse routes in England and Wales: The evolution of the network in the later-sixteenth and early-seventeenth century". *Journal of Historical Geography 17*(4), 373–389.

Brayshay, M. (2011). Tudor and Stuart provincial posting and the development of England's Exchequer-funded postal network. Technical report, Second international conference of Cultures of Knowledge: An Intellectual Geography of the Seventeenth-Century Republic of Letters.

Brix, A. C. (2026). Enciclopedia britannica: postal system. Available at: https://www.britannica.com/topic/postal-system. Accessed 7 April 2026.

Broadberry, S. (2021). Chapter 24 - the industrial revolution and the great divergence: recent findings from historical national accounting. In A. Bisin and G. Federico (Eds.), *The Handbook of Historical Economics*, pp. 749–771. Academic Press.

Campbell-Smith, D. (2011). *"Masters of the post: The authorized history of the Royal Mail"*. Allen Lane.

Cantoni, D. and N. Yuchtman (2014). "medieval universities, legal institutions, and the commercial revolution". *The Quarterly Journal of Economics, 129*(2), 823–887.

Carr, R. (1668). *A description of al the postroads in England, from London to Edenborough in Scotland*. British Library London.

Chen, J. and J. Roth (2024). Logs with zeros? some problems and solutions. *The Quarterly Journal of Economics 139*(2), 891–936.

Conley, T. G. (1999). Gmm estimation with cross sectional dependence. *Journal of Econometrics 92*(1), 1–45.

De la Croix, D., F. Docquier, A. Fabre, and R. Stelter (2024). "the academic market and the rise of universities in medieval and early modern europe (1000–1800)". *Journal of the European Economic Association, 22*(4), 1541–1589.

De la Croix, D., R. Scebba, and C. Zanardello (2025). Flora, cosmos, salvatio: Pre-modern academic institutions and the spread of ideas. Technical report, Centre for Economic Policy Research.

Dittmar, J. E. (2011). "Information technology and economic change: the impact of the printing press". *The Quarterly Journal of Economics, 126*(3), 1133–1172.

Dittmar, J. E. and S. Seabold (2019). "new media and competition: Printing and Europe's transformation after Gutenberg". Centre for Economic Performance Discussion Paper No. 1600.

Dowey, J. (2013). *Mind over Matter: Access to Knowledge and the British Industrial Revolution*. Ph. D. thesis, London School of Economics.

Edwards, J. F. (1987). *The transport system of medieval England and Wales: a geographical synthesis*. University of Salford (United Kingdom).

Giuliano, P. and N. Nunn (2021). Understanding cultural persistence and change. *The Review of Economic Studies 88*(4), 1541–1581.

Greene, R. (1690). *A New Map of England; Containing the Adjacent parts of Scotland, Ireland, France, Flanders, and Holland. Shewing the true Situation and Distance of London from Edinburg, Dublin, Paris, Maastricht, Antwerp, Amsterdam, & c. With a Description of the Post Roads, and their several branches, from Town to Town, never before extent*. At ye Rose and Crowne in ye Middle of Budge.row., London.

Hainmueller, J. and Y. Xu (2013). Ebalance: A stata package for entropy balancing. *Journal of Statistical Software 54*, 1–18.

Hanlon, W. W., S. Heblich, F. Monte, and M. B. Schmitz (2022). "a penny for your thoughts". NBER Working Paper No. 30076.

Headrick, D. (2015). International communication: History. In J. D. Wright (Ed.), *International Encyclopedia of the Social & Behavioral Sciences (Second Edition)* (Second Edition ed.)., pp. 490–494. Oxford: Elsevier.

Heldring, L., J. A. Robinson, and S. Vollmer (2021). "The Long-Run Impact of the Dissolution of the English Monasteries". *The Quarterly Journal of Economics 136*(4), 2093–2145.

Hemmeon, J. C. (1912). *The history of the British Post Office*. Harvard University Press.

Humphries, J. and J. Weisdorf (2019). Unreal wages? real income and economic growth in england, 1260–1850. *The Economic Journal, 129*(623), 2867–2887.

Hunter, M., A. Clericuzio, and L. M. Principe (2001). *The Correspondence of Robert Boyle*, Volume 1 (1636–61). Routledge.

Jones, E. (2003). *The European miracle: environments, economies and geopolitics in the history of Europe and Asia*. Cambridge University Press.

Kain, R. J. and R. R. Oliver (2001). *The Historic Parishes of England and Wales: An Electronic Map of Boundaries Before 1850 with a Gazetteer and Metadata*. Colchester: History Data Service.

Landes, D. S. (1969). *The unbound Prometheus: technological change and industrial development in Western Europe from 1750 to the present*. Cambridge University Press.

McCormick, M., G. Huang, G. Zambotti, and J. Lavash (2013). Roman Road Network (version 2008). Available at: `https://doi.org/10.7910/DVN/TI0KAU`.

Merton, R. K. (1938). Science, technology and society in seventeenth century england. *Osiris 4*, 360–632.

Mokyr, J. (2002). *The Gifts of Athena*. Princeton, NJ: Princeton University Press.

Mokyr, J. (2005). Long-term economic growth and the history of technology. In *Handbook of economic growth*, Volume 1, pp. 1113–1180. Elsevier.

Mokyr, J. (2010). *The Enlightened economy an economic history of Britain 1700-1850*. Yale University Press.

Mokyr, J. (2016). *"A Culture of Growth: the Origins of the Modern Economy"*. Princeton University Press.

Naqvi, A. (2022a). Sankey: Stata module for sankey diagrams. Statistical Software Components S459154, Boston College Department of Economics, revised 28 Jul 2025.

Naqvi, A. (2022b). Spider: Stata module for spider plots. Statistical Software Components S459136, Boston College Department of Economics, revised 28 Jan 2026.

Pape, D. (2004). CIA World DataBank II. Available at: `https://www.evl.uic.edu/pape/data/WDB/`.

Pomeranz, K. (2000). *The great divergence: China, Europe, and the making of the modern world economy*. Princeton, NJ: Princeton University Press.

Riva, A. and E. Toffoletto (2025). Enciclopedia britannica: Marcello malpighi. Available at: `https://www.britannica.com/biography/Marcello-Malpighi`, accessed November 26, 2025.

Robinson, H. (1948). *The British post office: A history*. Princeton University Press.

Rogowski, J. C., J. Gerring, M. Maguire, and L. Cojocaru (2022). Public infrastructure and economic development: Evidence from postal systems. *American Journal of Political Science 66*(4), 885–901.

Sharpe, K. (1984). Thomas witherings and the reform of the foreign posts, 1632–40. *Historical Research 57*(136), 149–164.

Spolaore, E. and R. Wacziarg (2013). How deep are the roots of economic development? *Journal of economic literature 51*(2), 325–69.

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

A Data Appendix

A.1 Evolution of the Postal System

In this section we illustrate the procedure used to reconstruct the evolution of the postal network across England and Wales. We also explain how we build an artificial network to predict the actual postal network of 1628.

Tracking postal stages in time. To reconstruct the evolution of the postal system during the period 1570-1769 we follow four steps.

The first step consists of conducting a systematic search for historical maps documenting the development of the postal system over time. We collect a set of maps depicting the historical postal system in place in the following years: 1570, 1580, 1595, 1603, 1628, 1668, 1677, and 1690. As an example, Panel (a) of Figure A1 presents the 1668 map digitized from Carr (1668).

Figure A1: Tracking the Postal Stages in 1668



(a) (b)

Notes: Panel (a) displays the map of the postal system in England and Wales in 1668. Panel (b) shows a magnified view of the region enclosed by blue box in Panel (a), with the existing postal stages marked in red.

In the second step, each map is digitized, subdivided into manageable sections, and georeferenced using ArcGIS 10.8. This procedure facilitates the identification of towns and villages by enabling a direct comparison between the historical map and contemporary geographic coordinates (whenever information on the locations name is hard to extract from a given map).

In the third step, we examine the map legend to determine which symbols denote postal stages. For example, the legend shown in Panel (b) of Figure A1 — which zooms in on the area highlighted in blue in Panel (a) — indicates that asterisks represent postal stages. In Panel (b), we mark each asterisk with a red circle for clarity. Because each asterisk is accompanied by a place name, we can identify the corresponding locality and record its geographic coordinates.

In the fourth step, we match the geographic coordinates of localities identified as postal stages to the corresponding grid cells in a fishnet layer covering England and Wales (using a spatial join GIS operation) and repeat this process for all our maps.

The resulting dataset is a grid cell-level panel covering the entire study area, allowing us to observe both the presence and absence of postal stages across cells in each map year. Because historical maps are available for a limited number of years, we use the decade as unit of temporal aggregation. This choice is motivated by the fact that changes in the postal network were gradual rather than abrupt, making it unlikely that substantial reconfigurations occurred within a few years. To construct a balanced panel covering all decades from 1570 to 1760, we assign to each decade the postal stage configuration depicted in the map produced within that decade. For decades without a corresponding map, we assume that the postal network remained unchanged, and so we carry forward the most recently observed configuration. This implies that the 1603 map is assigned in the final dataset to the decades 1600 and 1610, the 1628 map to all decades from 1620 to 1650, and so on.

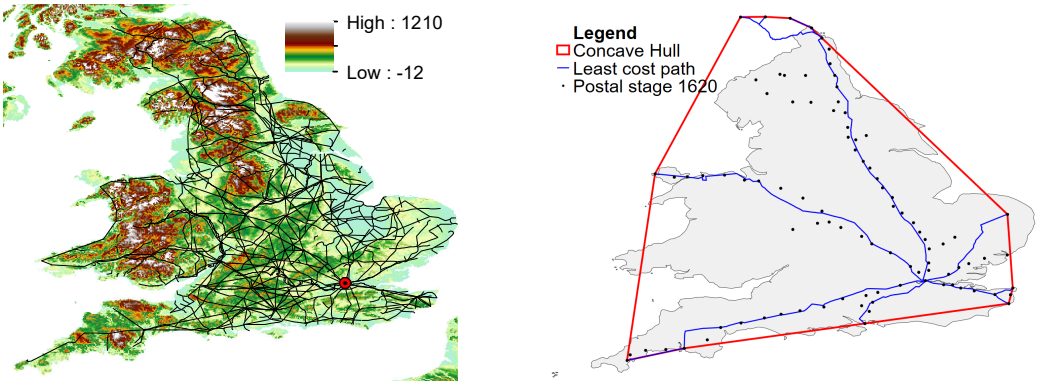
Constructing the predicted network. Estimating the effect of having a postal stage using the actual postal network over time would imply the assumption that postal stages randomly appeared within England and Wales throughout the decades. To mitigate the concern of non-random allocation of the postal stages after the reform, we focus on the network of postal stages of 1628, that is, the latest year in the pre-reform period in which we observe the full postal network configuration. To further address the threats to identification coming from potential non-random allocation of postal stages in 1628, we construct a hypothetical least cost path network and use it as instrumental variable for the actual route placement in that year (see Appendix Figure A2). We refer to this as the predicted network. Given that the purpose of the postal system prior to the reform was to connect the Queen/King with distant locations (due to local turmoil or territorial control) or to continental Europe (see Section 2.1), the proposed instrument addresses the question of which route the royal messengers would have taken to reach the extremes of the English reign, subject to cost minimization.

The least cost path network displayed in Panel (b) of Appendix Figure A2 (blue line) is built following five steps. First, we geolocate all postal stages included in the map of the 1628 postal network (the black dots in the panel). Second, we construct the concave hull containing the 1628 network (the red polygon in the panel) to have a measure of the extent of the reign that the King's messengers were originally able to reach. Third, we identify unique least cost paths from London to each vertex of the concave hull, based on the Dijkstra's optimal route algorithm. To compute transportation costs, we rely on a simple construction cost function from the transport engineering literature and assigned higher transportation costs to land parcels without transportation routes and with steeper slope of the terrain. We build historical routes by assembling the shapefiles of historical Roman roads (McCormick et al., 2013), medieval transport routes (Edwards, 1987), and rivers (Pape, 2004), while slope of the terrain was computed based on the European Digital Elevation Model (DEM).[43] Fourth, we assemble all the separate least cost paths to form a unique predicted network connecting London to the extremes of the 1628 postal network. Finally, for each grid cell we calculate the geodesic distance (in kilometers) from the centroid of the cell to the full predicted least cost

[43]Link: https://land.copernicus.eu/imagery-in-situ/eu-dem/eu-dem-v1.1?tab=metadata. Last access: 30th December 2018.

path network and construct a dummy variable equal to one if such distance is within 10 kilometers, and equal to zero otherwise. As this distance-based variable is time-invariant, we interact it with the post-reform indicator to define our instrumental variable in the final panel dataset.

Figure A2: Pre-Reform Postal Stages and Predicted Network



(a) Historical Routes and Terrain

(b) 1628 Postal Stages and Least Cost Path

Notes: Panel (a) displays the historical routes of England and Wales (black lines), altitude of the terrain in meters, and the location of London (red circle). Panel (b) presents the postal stages in 1628 (black dots), the concave hull including all the vertices of the 1628 network (red polygon), and the least cost path network joining London with each vertex of the concave hull (blue line). *Sources*: Historical routes of England and Wales combine Roman roads from McCormick et al. (2013), medieval routes from Edwards (1987), and rivers from Pape (2004), and own elaborations.

A.2 Details on Data about the Republic of Letters

A.2.1 Descriptive statistics on Correspondents

Table A1: Letter Correspondents by Occupation Macro-Categories

Category	Tot. correspondents (%)	Tot. EMLO correspondents (%)	Tot. EE correspondents (%)	Tot. letters "belonging" to the category (%)
Scholars	19.6	18.5	25.7	78.5
Practitioners	14.5	11.8	28.5	36.0
Statesmen	20.8	17.9	36.0	43.6
Information industry	5.4	4.1	12.0	28.3
Clergy	19.6	20.5	14.6	32.5
Other	39.6	43.5	19.2	20.4

Notes: The table shows the shares of correspondents in each occupation macro-category and the share of letters "belonging" to each macro-category, i.e., letters whose sender or recipient is associated with that macro-category. Our dataset includes 9,765 correspondents in total, of which 8,205 are from EMLO and 1,560 from EE. The percentages in each column are not meant to add up to 100%, as the correspondents' different occupations could fall into more than one macro-category and therefore a letter involving a given correspondent (either as sender or recipient) could "belong" to more than one macro-category.

Table A2: Top 30 Letter Correspondents

#	Letters			Correspondent full name	Macro-Category				
	Total	Sent	Received		Scholar	Practitioner	Statesman	Information industry	Clergy
1	6027	1719	4308	Locke, John, 1632-1704	1	1	0	0	0
2	3320	289	3031	Hartlib, Samuel, 1600-1662	1	0	1	0	0
3	3263	1530	1733	Oldenburg, Henry, 1619-1677	1	0	1	1	0
4	2967	617	2350	Lhwyd, Edward, 1659-1709	1	0	0	0	0
5	2366	602	1764	Ministers, elders and consistory of the Dutch Church in London	0	0	0	0	1
6	1771	476	1295	Lister, Martin, 1639-1712	1	1	0	0	0
7	1613	926	687	Smith, Thomas, 1638-1710	1	0	0	1	1
8	1440	1099	341	Wallis, John, 1616-1703	1	0	0	1	0
9	1276	762	514	Flamsteed, John, 1646-1719	1	0	1	0	0
10	1259	241	1018	Boyle, Robert, 1627-1691	1	0	0	0	0
11	1166	344	822	Brett, Thomas, 1667-1744	1	0	0	0	1
12	1154	510	644	Aubrey, John, 1626-1697	1	0	0	1	0
13	1104	102	1002	Charlett, Arthur, 1655-1722	1	1	0	0	0
14	1084	0	1084	Witt, Johan de, 1625-1672	1	0	1	0	0
15	1010	825	185	Pope, Alexander, 1688-1744	0	0	0	1	0
16	951	28	923	Wood, Anthony, 1632-1695	1	0	0	0	0
17	942	810	132	Gray, Thomas, 1716-1771	1	0	0	0	0
18	933	888	45	Evelyn, John, 1620-1706	1	0	0	1	0
19	841	454	387	Newton, Isaac, 1642-1727	1	0	0	0	0
20	840	728	112	Dury, John, 1596-1680	1	0	0	0	1
21	807	385	422	Casaubon, Isaac, 1559-1614	1	0	0	0	0
22	696	34	662	Sancroft, William, 1617-1693	0	0	0	0	1
23	681	237	444	Clarke, Edward, 1649-1710	0	0	1	0	0
24	655	544	111	Hume, David, 1711-1776	1	0	0	0	0
25	644	592	52	Addison, Joseph, 1672-1719	0	0	1	1	0
26	611	405	206	Le Clerc, Jean, 1657-1736	1	0	0	1	0
27	604	178	426	King, Peter, 1669-1734	0	0	1	0	0
28	589	85	504	Rawlinson, Richard, 1690-1755	1	0	0	0	1
29	563	230	333	Jurin, James, 1684-1750	1	1	0	0	0
30	555	93	462	Richardson, Richard, 1663-1741	1	1	0	0	0

A.2.2 Topic Analysis

To explore the content of the letters, we perform an unsupervised text analysis using the Latent Dirichlet Allocation algorithm (Blei et al., 2003). Based on this algorithm, each letter is a document and its text can be represented as a mixture of latent topics, that have a distribution over words. The algorithm assigns each word to a temporary topic semi-randomly (following a Dirichlet distribution) and then iteratively checks and updates topic assignments, looping over words in documents. For each word, the topic assignment is updated based on how prevalent that word is across topics and how prevalent are topics in the document, so that for each document it provides the share of topic j discussed in the document (gamma). Before running the algorithm, we follow the literature and pre-process the text following steps:

1. We remove html text that was left in the abstract field, together with some characters such as "nbsp" and "&amp", that stand for "non-breaking space" and "and", respectively;
2. Remove the Latin keywords: "recto" and "folio" indicating the front of a page or a page in general;
3. We translate the text of all letters in English.
4. We replace archaic forms of English with the contemporary ones.
5. We remove appellations such as Sir and Lady and correct the spelling and remove some appellations that were reported with spaces. For instance, the keyword "Mister" is often reported as "M r".
6. We remove names of months and weekdays (and their abbreviations). For instance, the month of October is sometimes spelled as "Oct ";
7. We put everything to lower case and remove accents;
8. We clean the text from the sender and recipient surnames (whenever they appear). We homogenize the surnames of Harlib, Comenius, and Durey because we realize that they were reported differently in time;
9. We remove the keywords "Editorial comment", "Extract of a letter", "Extract of letter" that appear whenever the text of a letter is not fully reported;
10. We clean parenthesis in unicode characters that are reported whenever a word has been reported as reconstructed from the original text of the letter;
11. We clean the text from the place of origin and destination of the letter (whenever available);
12. We remove all appellations such as "Sir", "Mister", "Lord", "Miss", "Lady", etc.;
13. We remove numbers, punctuation, name of months;
14. We further substitute some archaic forms of words. For instance we substitute "soe" with "so", "doe" with "do", etc.

15. We clean starting greetings. For example, most letters begin with the keyword "Dear" or "Honorable", "Illustrious", "Noble" or "Friend", which are clearly standard forms of addressing a correspondent rather than a proper letter content. To clean the text from these words we focus on the first 150 characters and remove them;

16. We clean closing greetings. For instance, most letters end with the keywords "Adieu", "Farewell", and wordings such as "I am your most humble and obedient servant" or "I am your most affectionate friend". To clean the text from these words we focus on the last 150 characters and remove them;

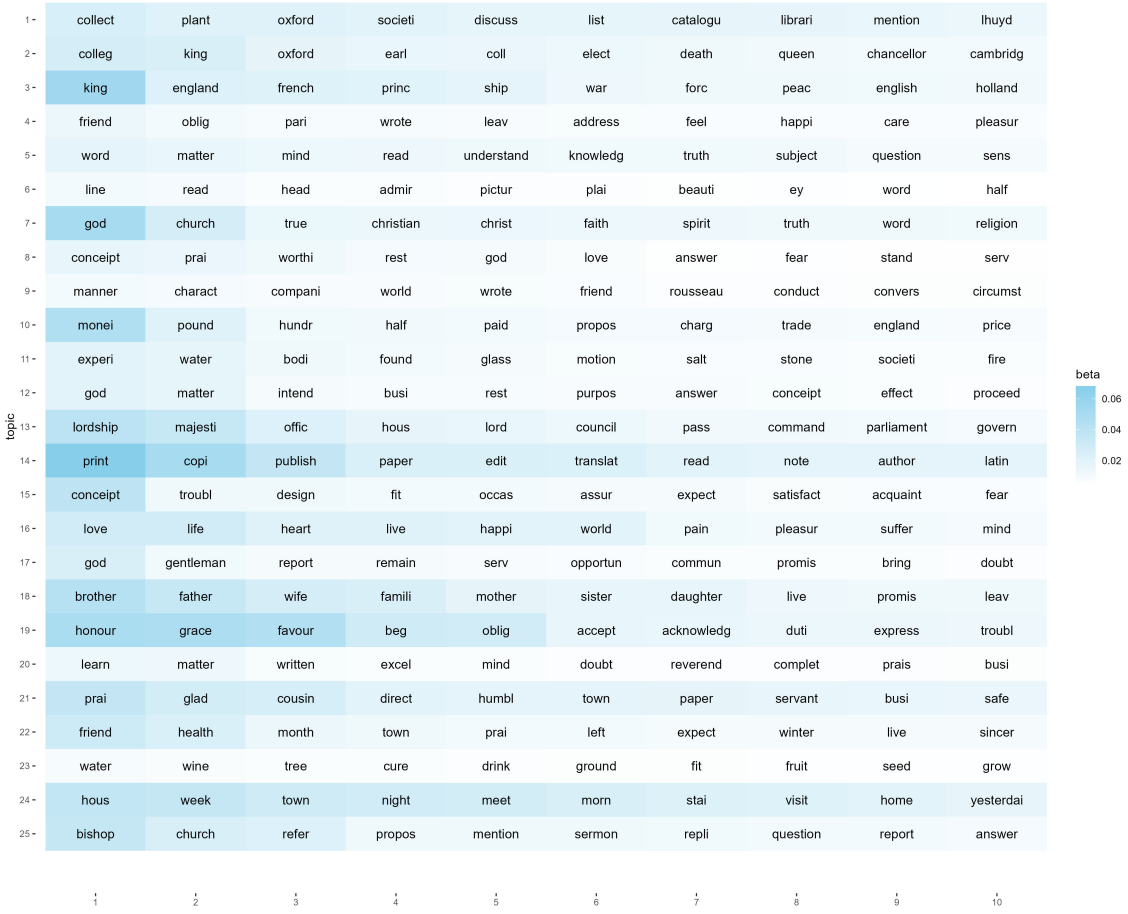
17. We remove English stopwords and replace words with their stems using the Porter stemmer;

18. We compute the term frequency-inverse document frequency (tf-idf) index and drop words whose rank is above 24,000 (thus excluding words that are too common or too rare to meaningfully qualify the content of a document).

Then, we ask the LDA algorithm to classify the content of the letters into 25 topics.[44] Figure A3 reports the 15 topics identified by the algorithm, showing the top-10 terms that mostly represent their content.

[44] As a standard in this literature, we set the parameter α to 50/2 and the parameter η to 0.1. These parameters are meant to control the mixture of topics for any given document and to control the distribution of words per topic, respectively.

Figure A3: Topic Matrix



Notes: The matrix displays the 10 keywords with highest weight in defining the 25 topics of the LDA. A lighter color is assigned to keywords with lower weight. Source: own elaborations over the text of 31,069 letters from EMLO and EE.

Finally, we group together topics that pertain to the same macro-topic. To do this, in the first place we use both our own understanding of the top-10 keywords and the wordclouds of each topic (these are reported in Section 6.2 and in the Online Appendix). In addition, we exploit the understanding of a Large Language Model. In particular, we submitted to Chat GPT 5.1 the following prompt *"I am giving you the text of a series of letters. Abstracting from authors/recipients, please summarize the content of the text in one paragraph. Then, explain in a second paragraph whether the content pertains more to one of the following 5 macro-topics: (1) knowledge; (2) religion; (3) politics; (4) ordinary life matters (i.e.: health, finances, work); or (5) academic gossip."*, followed by the text of the 30 documents with the highest gamma in that topic.

Below we present the short summary of the content of the topics we finally included in the *Knowledge* macro-topic, based on the 30 documents with the highest gamma and the LLM summaries and classification.

“Practical Knowledge”: This extraordinarily long series of letters is a vast compendium of practical agricultural, horticultural, botanical, medical, and household knowledge, written in an instructive and experimental spirit. The writer records detailed observations on producing high-quality cider and perry; managing orchards, grafting, soils, and fertilizers; raising goats, sheep, and cattle; preventing livestock diseases; cultivating grains, beans, and New World crops; malting maize; brewing beer; and improving land with irrigation, tree-planting, and new crops. Interwoven are long digressions on the medical uses of herbs, dietary practices, cures for fevers, wounds, ulcers, gout, scurvy, and other ailments; reflections on climate, air, early modern science, and natural philosophy; notes on beekeeping, fishponds, pearls, and shellfish; and occasional moral or religious commentary. Overall, the letters form an encyclopedic manual of rural improvement, experimental husbandry, and vernacular science, aimed at increasing the wealth, health, and self-sufficiency of England through empirical trial, observation, and the circulation of practical knowledge.

“Concepts and Methods”: The letters you provided form a vast, intellectually dense correspondence focused both on abstract reasoning about concepts and guidance on how to think, argue, and act. The writers debate at length the nature of mind, will, liberty, language, and knowledge, emphasizing that clarity of ideas and precise definitions are essential to avoid confusion and error. It also advances metaphysical and theological claims, such as the distinction between will and desire, the dependence of human action on prior judgment, and the unity of God derived from the notion of a perfectly infinite being. Alongside this, the text advises how to conduct inquiry (by carefully analyzing ideas and avoiding ambiguous language), how to reason properly (by grounding arguments in clear definitions rather than metaphor), and how to pursue knowledge and communication effectively. Practical guidance appears as well, for instance in discussions of teaching language to a deaf person, structuring thought and meditation, and engaging critically yet constructively with other authors. Overall, the text intertwines theoretical reflection with methodological instruction, presenting both what knowledge is and how it should be developed and applied.

“Experiments”: The letters you provide form an immense, wide-ranging body of seventeenth- and eighteenth-century correspondence concerned almost entirely with scientific, philosophical, alchemical, medical, and natural-historical investigations. Across hundreds of pages, the writers describe speculative theories of matter and ether; optical, mechanical, and astronomical problems; the tides, gravity, pendulum isochronism, magnetism, and the vacuum; chemical theories about salts, solvents, “alcahest,” alchemical mercury, metallurgical transmutation, and the generation of gems; anatomical dissections (including brains, ears, conjoined twins), physiology, and experimental transfusion of blood between animals; electrical effects, combustion, air chemistry, and maritime meteorology; mining, mineralogy, and natural resources; optical instruments, telescopes, microscopes, burning mirrors, and barometers; and numerous reports of experiments performed for the Royal Society. The letters are overwhelmingly technical, experimental, speculative, and conceptual—an active exchange of methods, observations, and theories among early modern natural philosophers.

“Cataloguing”: This collection of letters is an extensive exchange among naturalists, antiquaries, philologists, and members of the Royal Society in the late 17th and early 18th centuries. The writers report

fossil discoveries, mineral specimens, plant collections, shell catalogues, and observations gathered during field travels across England, Wales, and beyond; they discuss proposed publications on fossils, natural histories, plant anatomy, and regional floras; they describe museums, collections, rare books, and newly printed scientific works; and they request specimens, manuscripts, linguistic materials, and antiquarian information. Many letters include very detailed descriptions of local geology, botany, zoology, coins, inscriptions, manuscript sources, and etymologies. At the same time, they comment on scholarly disputes, printing delays, library fires, and the activities of various learned societies.

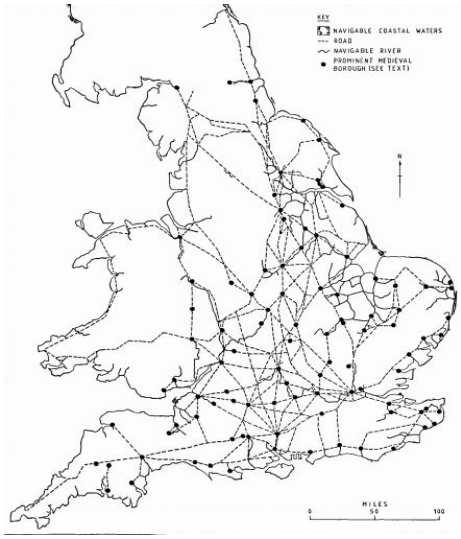
"Publishing": The text is an enormous compendium of early-modern scholarly correspondence—thousands of names, titles, and excerpts—centering on the exchange of manuscripts, critical editions, philological notes, emendations, translations, and textual scholarship across Europe. The letters discuss editing Greek and Latin authors (Ignatius, Clement, Theocritus, Menander, Phalaris, Josephus, Callimachus, Theocritus), translating the Bible into numerous languages (Irish, Turkish, Lithuanian, Indigenous American languages), publishing treatises on chronology, ecclesiastical history, chemistry, medicine, astronomy, and antiquities, acquiring manuscripts from libraries across the Continent and the East, and improving scholarly editions through new readings and annotations. They include details about printing problems, amanuenses, booksellers' disputes, manuscript transcription, academic controversies, and highly technical philological debates. Altogether, the collection represents the entire machinery of seventeenth- and eighteenth-century learned communication.

"Pansophia": This vast body of correspondence is an early-modern intellectual network in motion: long, dense letters exchanging news, manuscripts, draft treatises, critiques, and practical arrangements surrounding the work of Comenius and other scholars engaged in projects of pansophia, universal language, universal knowledge, pedagogy, mathematics, natural philosophy, alchemy, chemistry, medicine, mechanics, printing, and book distribution. The writers report on the progress of books in press, circulate excerpts, seek funding for collaborators, describe experiments (pyrotechnic, metallurgical, botanical, astronomical), debate the possibility of perpetual motion or the philosopher's stone, evaluate mathematical methods (analysis, algebra, optics, memory arts), discuss proposed reforms of learning and universal systems of knowledge, and report on scholarly reputations and patronage. Scattered through the letters are references to political events (peace treaties, wars), practical travel and health difficulties, and interpersonal frictions, but their center of gravity is the cooperative—and sometimes contentious—intellectual labor of producing, evaluating, and disseminating ambitious scholarly and scientific works.

A.3 Details on Other Variables

Medieval Routes. We georeference the map of the transport system of medieval England and Wales from Edwards (1987) (see Figure A4). Geo-referentiation is done using the "Thin Plate Spline (TPS)" transformation and the "nearest neighbour" resampling method. Lines connecting points have been created through a MultiLineString by writing a CSV file where we listed each segment connecting starting and ending points.

Figure A4: Transport System of Medieval England and Wales



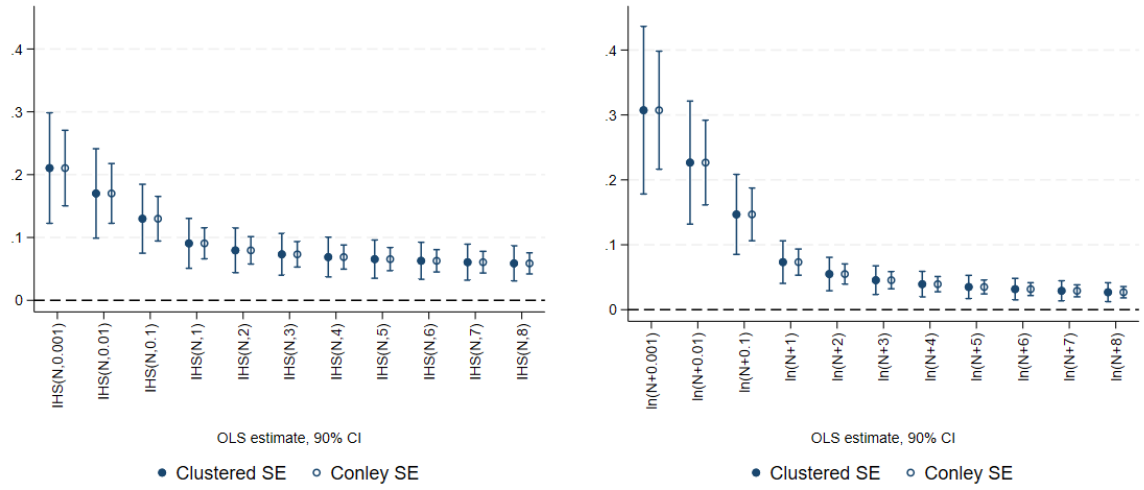
(a) Original Map

Notes: The source is Edwards (1987).

B Robustness Checks

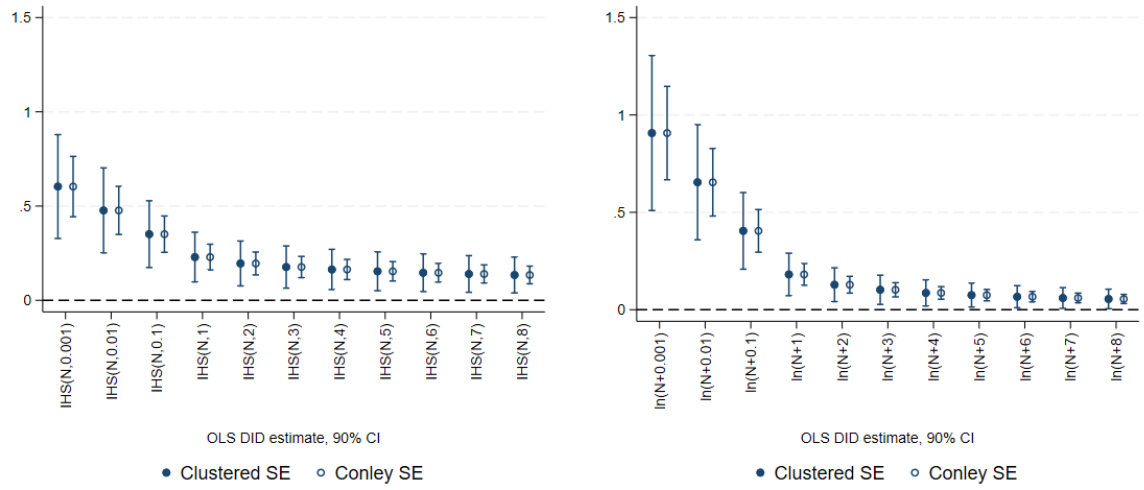
B.1 Sensitivity Analysis

Figure B1: Robustness: Different Transformations of the Dependent Variable



(a) Correlational estimates: IHS(All Letters, c)

(b) Correlational estimates: Ln(All Letters + c)



(c) DID estimates: IHS(All Letters, c)

(d) DID estimates: Ln(All Letters + c)

Notes: In Panels (a) and (b), each dot represents the estimated coefficient of the indicator variable *Stage Dummy*$_{it}$ from Equation (1), using the specification in Column 5 of Table 1. In Panels (c) and (d), each dot represents the estimated coefficient of the interaction term *Stage Dummy 1620*$_i$ $\times$ *Post-1635*$_t$ from Equation (2), using the specification in Column 5 of Table 2. In all panels, N indicates the total number of incoming and outgoing letters in each grid cell-decade. The dependent variable in Panels (a) and (c) is IHS$(N,c) = ln(N + \sqrt{N^2 + c})$, while the dependent variable in Panels (b) and (d) is $ln(N + c)$, for different values of c. OLS estimates. 90% confidence intervals based on standard errors clustered by grid cell (filled marker) or Conley's spatially-corrected standard errors (accounting for spatial correlation up to 50 km with Bartlett kernel decay in space, empty marker) are reported.

Table B1: Robustness: Extensive Margin and Winsorized Dependent Variable

Empirical Strategy	Correlational Evidence		DID Evidence	
Dependent Variable	Dummy Letter	All Letters Winsorized 99%	Dummy Letter	All Letters Winsorized 99%
	(1)	(2)	(3)	(4)
Stage Dummy	0.035	0.198		
	(0.009)***	(0.053)***		
	[0.007]***	[0.033]***		
Stage Dummy 1620 × Post-1635			0.110	0.502
			(0.028)***	(0.180)***
			[0.018]***	[0.096]***
Decade FE	✓	✓	✓	✓
Cell FE	✓	✓	✓	✓
Log Dist. to London × Decade FE	✓	✓	✓	✓
Log Dist. to Historical Routes × Decade FE	✓	✓	✓	✓
Log Dist. to Medieval Boroughs × Decade FE	✓	✓	✓	✓
Stage Dummy × Civil War Dummy	✓	✓	✓	✓
N. Clusters	2,764	2,764	2,764	2,764
N. Observations	55,280	55,280	55,280	55,280
Adj. R-Squared	0.248	0.344	0.249	0.344
Mean of DV	0.066	0.220		
Pre-1630 Mean of DV			0.023	0.070

Notes: Observations are grid cells observed in each decade from 1570 to 1769. In Columns 1 and 3, the dependent variable is a dummy equal to one in each grid cell-decade ever receiving or sending a letter, and zero otherwise. In Columns 2 and 4 the dependent variable is the total number of incoming and outgoing letters in each grid cell-decade. This number is winsorized at the 99th percentile to reduce excess variance. OLS estimates. Standard errors clustered by grid cell are reported in parentheses, while Conley's standard errors accounting for spatial correlation up to 50 km with Bartlett kernel decay in space are reported in squared brackets. * = p-value $<$10%, ** = p-value $<$5%, *** = p-value $<$1%.

B.2 Main Difference-in-Differences Analysis

Table B2: Robustness: Effect of the 1635 Postal Reform on Intensity of Correspondence, Dealing with Relevant Locations and Correspondents

Dependent Variable	IHS(All Letters)			
	No London grid cell	No letters to/from London	No top-15 correspondents	Oxford/Cambridge with postal stage
	(1)	(2)	(3)	(4)
Stage Dummy 1620 × Post-1635	0.226	0.218	0.190	0.284
	(0.081)***	(0.080)***	(0.061)***	(0.086)***
	[0.042]***	[0.042]***	[0.035]***	[0.043]***
Decade FE	✓	✓	✓	✓
Cell FE	✓	✓	✓	✓
Log Dist. to London × Decade FE	✓	✓	✓	✓
Log Dist. to Historical Routes × Decade FE	✓	✓	✓	✓
Log Dist. to Medieval Boroughs × Decade FE	✓	✓	✓	✓
Stage Dummy 1620 × Civil War Dummy	✓	✓	✓	✓
N. Clusters	2,763	2,764	2,764	2,764
N. Observations	55,260	55,280	55,280	55,280
Adj. R-Squared	0.361	0.361	0.425	0.408
Pre-1630 Mean of DV	0.036	0.036	0.035	0.039

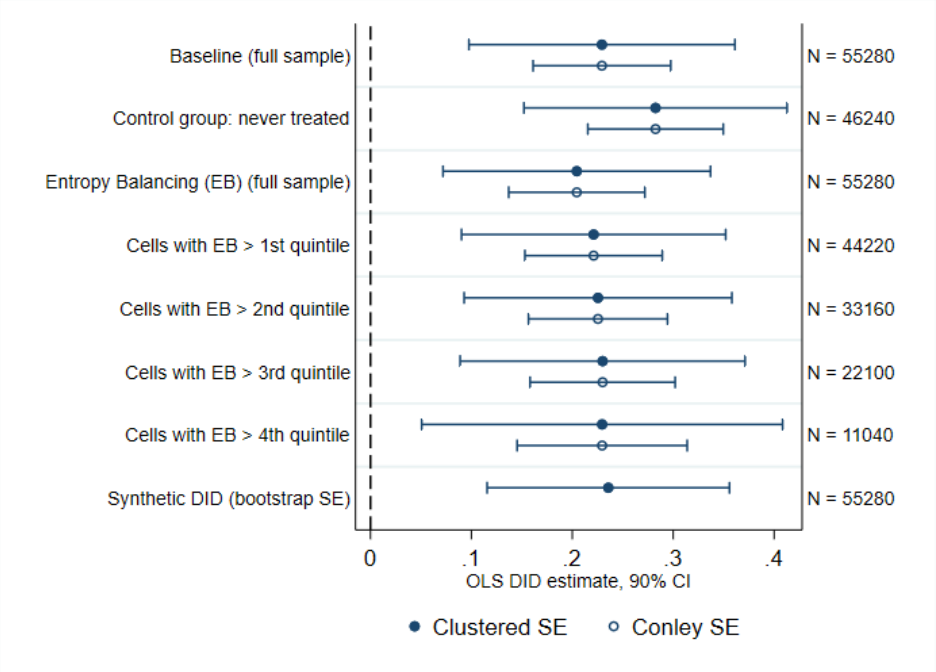
Notes: Observations are grid cells observed in each decade from 1570 to 1769 in all columns except Column 1 (where the sample excludes the cell hosting the London General Post Office). The dependent variable is the inverse hyperbolic sine (IHS) transformation of the total number of incoming and outgoing letters in each grid cell-decade, except Column 2 (where the dependent variable is computed excluding all letters sent to and received from London) and Column 3 (in which the dependent variable is computed excluding all letters sent to or received by a top-15 correspondent). The explanatory variable *Stage Dummy 1620 × Post-1635* equals one in all the grid cells containing at least one postal stage in the decade 1620 from 1630 onward, and zero otherwise. In Column 4 the grid cells including Oxford and Cambridge are recoded as having a postal stage before 1635. OLS estimates. Standard errors clustered by grid cell are reported in parentheses, while Conley's standard errors accounting for spatial correlation up to 50 km with Bartlett kernel decay in space are reported in squared brackets. * = p-value $<$10%, ** = p-value $<$5%, *** = p-value $<$1%.

Table B3: Robustness: Effect of the 1635 Postal Reform on Intensity of Correspondence, Alternative Treatment Definitions

Dependent Variable	IHS(All Letters)	
	Post-1635 = 1 from 1640 onward	Treatment = all postal stages pre-1635
	(1)	(2)
Stage Dummy 1620 × Post-1635	0.226	
	(0.079)***	
	[0.041]***	
Stage Dummy Pre-1635 × Post-1635		0.296
		(0.071)***
		[0.035]***
Decade FE	✓	✓
Cell FE	✓	✓
Log Dist. to London × Decade FE	✓	✓
Log Dist. to Historical Routes × Decade FE	✓	✓
Log Dist. to Medieval Boroughs × Decade FE	✓	✓
Stage Dummy 1620 × Civil War Dummy	✓	✓
N. Clusters	2,764	2,764
N. Observations	55,280	55,280
Adj. R-Squared	0.407	0.409
Pre-1630 Mean of DV	0.047	0.039

Notes: Observations are grid cells observed in each decade from 1570 to 1769. The dependent variable is the inverse hyperbolic sine (IHS) transformation of the total number of incoming and outgoing letters in each grid cell-decade. The main explanatory variable in Column 1 is the interaction term *Stage Dummy 1620 × Post-1635*, which equals one in all the grid cells containing at least one postal stage in the decade 1620 from 1640 onward, and zero otherwise. The main explanatory variable in Column 2 is the interaction term *Stage Dummy Pre-1635 × Post-1635*, which equals one in all the grid cells ever hosting a postal stage before 1635 in each decade from 1630 onward, and zero otherwise. OLS estimates. Standard errors clustered by grid cell are reported in parentheses, while Conley's standard errors accounting for spatial correlation up to 50 km with Bartlett kernel decay in space are reported in squared brackets. * = p-value $<10\%$, ** = p-value $<5\%$, *** = p-value $<1\%$.

Figure B2: Robustness: Effect of the 1635 Postal Reform on Intensity of Correspondence, Different Samples



Notes: Each dot represents the estimated coefficient of the interaction *Stage Dummy 1620* × *Post-1635* from Equation (2), using the specification in Table 2, Column 5. The dependent variable is the inverse hyperbolic sine transformation of the total number of incoming and outgoing letters in each grid cell-decade. The left y-axis indicates the estimation sample and method, whereas the right y-axis indicates the related number of observations. OLS estimates. 90% confidence intervals based on standard errors clustered by grid cell (filled marker) or Conley's spatially-corrected standard errors (accounting for spatial correlation up to 50 km with Bartlett kernel decay in space, empty marker) are reported. For the synthetic DID estimation (last row in the figure), bootstrapped standard errors are reported (bootstrap replications = 100).

Table B4: Effect of the 1635 Postal Reform on Intensity of Correspondence, IV Results

Dependent Variable	IHS(All Letters) (ITT)		Stage Dummy 1620 × Post-1635 (first stage)		IHS(All Letters) (second stage)	
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}\{\text{Distance Least-Cost Path} < 10\text{km}\} \times$ Post-1635	0.087	0.029	0.115	0.104		
	(0.019)***	(0.017)*	(0.014)***	(0.014)***		
	[0.013]***	[0.011]***	[0.003]***	[0.003]***		
Stage Dummy 1620 × Post-1635					0.756	0.282
					(0.169)***	(0.163)*
					[0.116]***	[0.106]***
Decade FE	✓	✓	✓	✓	✓	✓
Cell FE	✓	✓	✓	✓	✓	✓
Log Dist. to London × Decade FE		✓		✓		✓
Log Dist. to Historical Routes × Decade FE		✓		✓		✓
Log Dist. to Medieval Boroughs × Decade FE		✓		✓		✓
Stage Dummy 1620 × Civil War Dummy		✓		✓		✓
N. Clusters	2,764	2,764	2,764	2,764	2,764	2,764
N. Observations	55,280	55,280	55,280	55,280	55,280	55,280
Adj. R-Squared	0.398	0.406	0.709	0.724	-0.004	0.016
Kleibergen-Paap Wald F-Statistic			67.165	60.988		

Notes: OLS (Columns 1-4) and 2SLS (Columns 5-6) estimates. Observations are grid cells in each decade from 1570 to 1769. In Columns 1, 2, 5 and 6 the dependent variable is the inverse hyperbolic sine (IHS) transformation of the total number of incoming and outgoing letters in each grid cell-decade. In Columns 3 and 4 the dependent variable is the interaction term *Stage dummy 1620* × *Post-1635*, equal to one in all the grid cells containing at least one postal stage in the decade 1620 from 1630 onwards, and zero otherwise. The indicator *Stage dummy 1620* is instrumented with a dummy variable equal to one if the centroid of the grid cell lies within 10 kilometers from a predicted least-cost-path network, and equal to zero otherwise (see Data Appendix A.1 for more details on the IV). Standard errors clustered by grid cell are reported in parentheses, while Conley's standard errors accounting for spatial correlation up to 50 km with Bartlett kernel decay in space are reported in squared brackets. * = p-value $<10\%$, ** = p-value $<5\%$, *** = p-value $<1\%$.

B.3 Presence of Postal Stage and Effect of the 1635 Reform on Intensity of Correspondence by Occupation of Correspondents

Table B5: Effect of the 1635 Postal Reform on Intensity of Correspondence by Occupation of Correspondents – DID Estimates

Dependent Variable	IHS(Letters)					
	Scholars → Scholars	Scholars ↔ Non Scholars	Scholars ↔ Professionals	Scholars ↔ Clergy	Scholars ↔ Other	Professionals → Professionals
	(1)	(2)	(3)	(4)	(5)	(6)
Stage Dummy 1620 × Post-1635	0.161	0.219	0.241	0.133	0.108	0.137
	(0.047)***	(0.057)***	(0.059)***	(0.051)***	(0.033)***	(0.047)***
	[0.024]***	[0.029]***	[0.028]***	[0.025]***	[0.019]***	[0.024]***
Decade FE	✓	✓	✓	✓	✓	✓
Cell FE	✓	✓	✓	✓	✓	✓
Log Dist. to London × Decade FE	✓	✓	✓	✓	✓	✓
Log Dist. to Historical Routes × Decade FE	✓	✓	✓	✓	✓	✓
Log Dist. to Medieval Boroughs × Decade FE	✓	✓	✓	✓	✓	✓
Stage Dummy 1620 × Civil War Dummy	✓	✓	✓	✓	✓	✓
N. Clusters	2,764	2,764	2,764	2,764	2,764	2,764
N. Observations	55,280	55,280	55,280	55,280	55,280	55,280
Adj. R-Squared	0.381	0.349	0.365	0.369	0.327	0.373
Pre-1630 Mean of DV	0.008	0.012	0.009	0.008	0.006	0.012

Notes: Observations are grid cells in each decade from 1570 to 1769. In each specification the dependent variable is the inverse hyperbolic sine (IHS) transformation of the number of incoming and outgoing letters in each grid cell-decade, of which both the sender and the recipient are scholars (Column 1). In Columns 2 to 6 the dependent variable is the inverse hyperbolic sine (IHS) transformation of the number of incoming and outgoing letters in each grid cell-decade, respectively, from scholars to non-scholars (and vice versa), from scholars to professionals (and vice versa), from scholars to clergy (and vice versa), from scholars to uncategorised people (and vice versa), and from professionals to professionals. The macro-category *Professionals* indicates a single group of people including practitioners, statesmen, and people employed in the information industry. The explanatory variable *Stage dummy 1620 × Post-1635* equals one in all the grid cells containing at least one postal stage in the decade 1620 from 1630 onwards, and zero otherwise. OLS estimates. Standard errors clustered by grid cell are reported in parentheses, while Conley's standard errors accounting for spatial correlation up to 50 km with Bartlett kernel decay in space are reported in squared brackets. * = p-value $<10\%$, ** = p-value $<5\%$, *** = p-value $<1\%$.

Table B6: Effect of the 1635 Postal Reform on Intensity of Correspondence by Occupation of Correspondents, Share of Letters as Dependent Variable – DID Estimates

Dependent Variable	Share Letters					
	Scholars → Scholars	Scholars ↔ Non Scholars	Scholars ↔ Professionals	Scholars ↔ Clergy	Scholars ↔ Other	Professionals → Professionals
	(1)	(2)	(3)	(4)	(5)	(6)
Stage Dummy 1620 × Post-1635	-0.030	0.021	0.160	-0.027	-0.065	0.124
	(0.047)	(0.065)	(0.065)**	(0.056)	(0.051)	(0.059)**
	[0.042]	[0.059]	[0.056]***	[0.052]	[0.045]	[0.055]**
Decade FE	✓	✓	✓	✓	✓	✓
Cell FE	✓	✓	✓	✓	✓	✓
Log Dist. to London × Decade FE	✓	✓	✓	✓	✓	✓
Log Dist. to Historical Routes × Decade FE	✓	✓	✓	✓	✓	✓
Log Dist. to Medieval Boroughs × Decade FE	✓	✓	✓	✓	✓	✓
Stage Dummy 1620 × Civil War Dummy	✓	✓	✓	✓	✓	✓
N. Clusters	1,220	1,220	1,220	1,220	1,220	1,220
N. Observations	3,639	3,639	3,639	3,639	3,639	3,639
Adj. R-Squared	0.244	0.206	0.236	0.220	0.180	0.278
Pre-1630 Mean of DV	0.134	0.249	0.162	0.149	0.095	0.248

Notes: Observations are grid cells in each decade from 1570 to 1769. In each specification the dependent variable is the share of letters of a given occupation pair of correspondents over the total number of incoming and outgoing letters in each grid cell-decade. In Columns 1 to 6 the dependent variable is constructed using letters, respectively, from scholars to scholars, from scholars to non-scholars (and vice versa), from scholars to professionals (and vice versa), from scholars to clergy (and vice versa), from scholars to uncategorised people (and vice versa), and from professionals to professionals. The macro-category *Professionals* indicates a single group of people including practitioners, statesmen, and people employed in the information industry. The explanatory variable *Stage dummy 1620 × Post-1635* equals one in all the grid cells containing at least one postal stage in the decade 1620 from 1630 onwards, and zero otherwise. OLS estimates. Standard errors clustered by grid cell are reported in parentheses, while Conley's standard errors accounting for spatial correlation up to 50 km with Bartlett kernel decay in space are reported in squared brackets. * = p-value $<10\%$, ** = p-value $<5\%$, *** = p-value $<1\%$.

Table B7: Effect of the 1635 Postal Reform on Intensity of Correspondence by Occupation of Correspondents, Excluding Letters Involving Top 15 Correspondents – DID Estimates

Dependent Variable	IHS(Letters)					
	Scholars → Scholars	Scholars ↔ Non Scholars	Scholars ↔ Professionals	Scholars ↔ Clergy	Scholars ↔ Other	Professionals → Professionals
	(1)	(2)	(3)	(4)	(5)	(6)
Stage Dummy 1620 × Post-1635	0.088	0.142	0.151	0.088	0.052	0.075
	(0.032)***	(0.041)***	(0.040)***	(0.038)**	(0.021)**	(0.031)**
	[0.019]***	[0.024]***	[0.021]***	[0.022]***	[0.015]***	[0.019]***
Decade FE	✓	✓	✓	✓	✓	✓
Cell FE	✓	✓	✓	✓	✓	✓
Log Dist. to London × Decade FE	✓	✓	✓	✓	✓	✓
Log Dist. to Historical Routes × Decade FE	✓	✓	✓	✓	✓	✓
Log Dist. to Medieval Boroughs × Decade FE	✓	✓	✓	✓	✓	✓
Stage Dummy 1620 × Civil War Dummy	✓	✓	✓	✓	✓	✓
N. Clusters	2,764	2,764	2,764	2,764	2,764	2,764
N. Observations	55,280	55,280	55,280	55,280	55,280	55,280
Adj. R-Squared	0.438	0.384	0.412	0.402	0.371	0.410
Pre-1630 Mean of DV	0.008	0.012	0.009	0.008	0.006	0.012

Notes: Observations are grid cells in each decade from 1570 to 1769. In each specification the dependent variable is the inverse hyperbolic sine (IHS) transformation of the number of incoming and outgoing letters in each grid cell-decade, of which both the sender and the recipient are scholars (Column 1). In Columns 2 to 6 the dependent variable is the inverse hyperbolic sine transformation of the number of incoming and outgoing letters in each grid cell-decade, respectively, from scholars to non-scholars (and vice versa), from scholars to professionals (and vice versa), from scholars to clergy (and vice versa), from scholars to uncategorised people (and vice versa), and from professionals to professionals. Letters involving the top 15 correspondents either as sender or recipient are excluded from the sample. The macro-category *Professionals* indicates a single group of people including practitioners, statesmen, and people employed in the information industry. The explanatory variable *Stage dummy 1620 × Post-1635* equals one in all the grid cells containing at least one postal stage in the decade 1620 from 1630 onwards, and zero otherwise. OLS estimates. Standard errors clustered by grid cell are reported in parentheses, while Conley's standard errors accounting for spatial correlation up to 50 km with Bartlett kernel decay in space are reported in squared brackets. * = p-value $<10\%$, ** = p-value $<5\%$, *** = p-value $<1\%$.

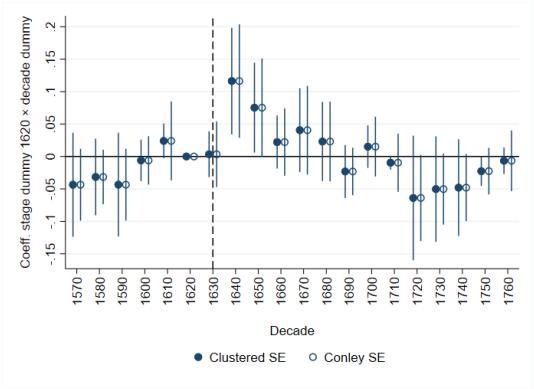
B.4 Presence of Postal Stage and Effect of the 1635 Reform on the Letters' Content

Table B8: Effect of the 1635 Postal Reform on Intensity of Correspondence by Macro-topic – DID Estimates

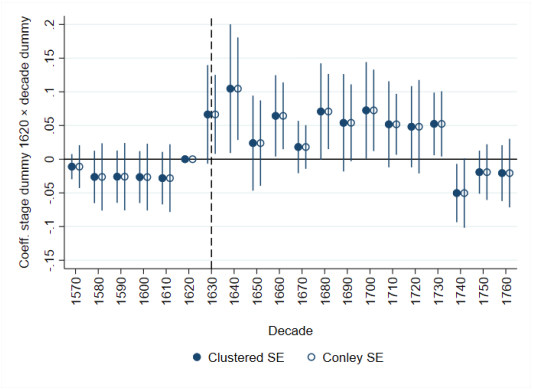
Dependent Variable	IHS(Letters)				
	Life	Knowledge	Politics	Church	Gossip
	(1)	(2)	(3)	(4)	(5)
Stage Dummy 1620 × Post-1635	0.159	0.148	0.006	0.053	0.025
	(0.051)***	(0.048)***	(0.010)	(0.031)*	(0.028)
	[0.026]***	[0.019]***	[0.012]	[0.013]***	[0.009]***
Decade FE	✓	✓	✓	✓	✓
Cell FE	✓	✓	✓	✓	✓
Log Dist. to London × Decade FE	✓	✓	✓	✓	✓
Log Dist. to Historical Routes × Decade FE	✓	✓	✓	✓	✓
Log Dist. to Medieval Boroughs × Decade FE	✓	✓	✓	✓	✓
Stage Dummy 1620 × Civil War Dummy	✓	✓	✓	✓	✓
Avg No. Words and (Avg No. Words)^2	✓	✓	✓	✓	✓
N. Clusters	2,764	2,764	2,764	2,764	2,764
N. Observations	55,280	55,280	55,280	55,280	55,280
Adj. R-Squared	0.516	0.412	0.417	0.321	0.261
Pre-1630 Mean of DV	0.012	0.003	0.002	0.001	0.000

Notes: Observations are grid cells in each decade from 1570 to 1769. In each specification the dependent variable is the inverse hyperbolic sine (IHS) transformation of the number of incoming and outgoing letters assigned to the macro-topic indicated in the column title. The interaction variable *Stage dummy 1620 × Post-1635* equals one in all grid cells containing at least one postal stage in the decade 1620 from 1630 onwards, and zero otherwise. OLS estimates. Standard errors clustered by grid cell are reported in parentheses, while Conley's standard errors accounting for spatial correlation up to 50 km with Bartlett kernel decay in space are reported in squared brackets. * = p-value $<10\%$, ** = p-value $<5\%$, *** = p-value $<1\%$.

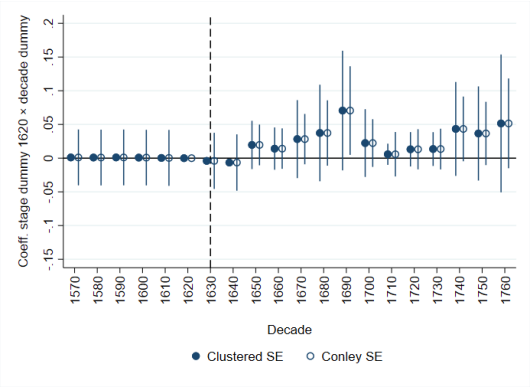
Figure B3: Dynamic Effect of the 1635 Postal Reform on Intensity of Correspondence for the Macro-topics Politics, Religion, and Academic Gossip – Event Study



(a) Topic Politics



(b) Topic Religion

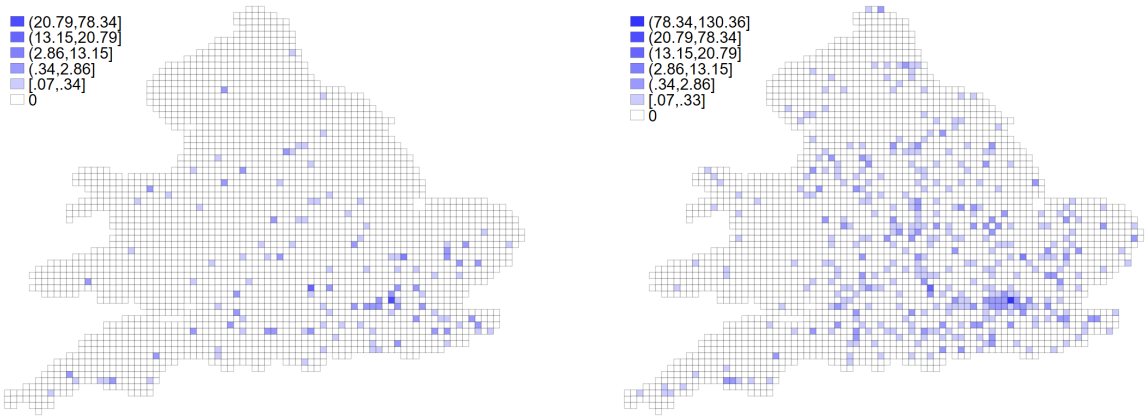


(c) Topic Academic Gossip

Notes: Each dot in a given subfigure represents the estimated coefficient of the interaction *Stage Dummy 1620* $\times$ decade fixed effect from Equation (3), using the specification in Table 2, Column 5. The dependent variable is the inverse hyperbolic sine transformation of the total number of incoming and outgoing letters assigned to a given macro-topic in each grid cell-decade. The omitted interaction term is *Stage Dummy 1620* $\times$ decade 1620 fixed effect. OLS estimates. 90% confidence intervals based on standard errors clustered by grid cell (filled marker) or Conley's spatially-corrected standard errors (accounting for spatial correlation up to 50 km with Bartlett kernel decay in space, empty marker) are reported. The dashed vertical line indicates the decade of the postal reform.

B.5 Presence of Postal Stage and Effect of the 1635 Reform on Innovation Activity

Figure B4: Innovation Events Before and After the 1635 Postal Reform



(a) Avg. Events involving Innovators per Decade, 1570-1620

(b) Avg. Events involving Innovators per Decade, 1630-1760

Notes: The figure shows the average number of events involving innovators in each grid cell per decade, over the pre-reform period (Panel a) and the post-reform period (Panel b). *Sources*: own elaborations based on Oxford Dictionary of National Biography.

Table B9: Impact of the 1635 Postal Reform on Innovation Activity – DID Estimates

Dependent Variable	All Innovators			Tech Innovators			Soft Innovators		
	IHS(Innovators)	IHS(Events)	IHS(Inventions)	IHS(Innovators)	IHS(Events)	IHS(Inventions)	IHS(Innovators)	IHS(Events)	IHS(Inventions)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Stage Dummy 1620 × Post-1635	0.041	0.044	0.013	0.047	0.058	0.015	0.004	-0.009	0.007
	(0.021)**	(0.026)*	(0.012)	(0.018)***	(0.020)***	(0.013)	(0.011)	(0.014)	(0.008)
	[0.013]***	[0.017]***	[0.005]**	[0.010]***	[0.014]***	[0.006]***	[0.008]	[0.010]	[0.005]
Decade FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Cell FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Log Dist. to London × Decade FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Log Dist. to Historical Routes × Decade FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Log Dist. to Medieval Boroughs × Decade FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Stage Dummy 1620 × Civil War Dummy	✓	✓	✓	✓	✓	✓	✓	✓	✓
N. Clusters	2,764	2,764	2,764	2,764	2,764	2,764	2,764	2,764	2,764
N. Observations	55,280	55,280	55,280	55,280	55,280	55,280	55,280	55,280	55,280
Adj. R-Squared	0.516	0.465	0.572	0.526	0.461	0.537	0.537	0.567	0.540
Pre-1630 Mean of DV	0.011	0.016	0.002	0.008	0.013	0.001	0.005	0.006	0.001

Notes: Observations are grid cells in each decade from 1570 to 1769. The dependent variables are: the number of innovators (not considering if they were born in that grid cell-decade); the number of events involving innovators (not considering their birth in that grid cell-decade); the number of inventions made by innovators. All these variables are IHS transformed. In Columns 4 to 6 and Columns 7 to 9 the above variables are measured using, respectively, only the subset of "technical" innovators and only the subset of "soft" innovators. In Columns 1 to 3 the above variables do not distinguish between different types of innovators. The interaction variable *Stage dummy 1620 × Post-1635* equals one in all grid cells containing at least one postal stage in the decade 1620 from 1630 onwards, and zero otherwise. OLS estimates. Standard errors clustered by grid cell are reported in parentheses, while Conley's standard errors accounting for spatial correlation up to 50 km with Bartlett kernel decay in space are reported in squared brackets. * = p-value <10%, ** = p-value <5%, *** = p-value <1%.