Spinning the Web: Codifiability, Information Frictions and Trade

Réka Juhász[1] and Claudia Steinwender*[2]

[1]Columbia University & CEPR
[2]MIT Sloan, NBER, CEPR & CEP

October 5, 2018

Abstract

This paper uncovers a novel mechanism through which information frictions matter for trade in differentiated goods; the *product specification mechanism.* We estimate the effect of a reduction in communication time on imports of three product categories in 19th century cotton textile trade; yarn, plain cloth, and finished cloth. In order to identify causal effects, we use exogenous variation in the ruggedness of the submarine seafloor to predict in which year countries get connected to the global telegraph network. The telegraph dramatically reduced the time it took to exchange information expressed in words, but did not affect the exchange of physical objects such as product samples. Using evidence from cotton traders' communication, we show that the examined three products differed in their *codifiability*, that is, in the extent to which merchants specified product attributes in words. Empirically, we find that communication time reductions had the largest effect on imports of the most codifiable product; yarn, and the smallest effect on the non-codifiable product, finished cotton cloth. Our results suggest that the effect of ICT on trade and fragmentation of production depends on the technology-specific codifiability of product specifications.

JEL code: F14, F15, F61, N70
Keywords: Trade, Information Frictions, ICT, Codifiability

*We thank David Atkin, Esther Bøler, Rafael Dix-Carneiro, Kerem Coşar, Don Davis, Dave Donaldson, Joe Doyle, Federico Esposito, Ivan Fernandez-Val, Bob Gibbons, Walker Hanlon, Rick Hornbeck, Amit Khandelwal, Marta Martinez, Dávid Nagy, Suresh Naidu, Guillermo Noguera, Steve Pischke, Jim Rauch, Pietra Rivoli, João Santos Silva, Silvana Tenreyro, Jon Vogel, David Weinstein, Susan Wolcott and numerous conference and seminar participants for very helpful comments and suggestions. David Jacks, Markus Lampe, Luigi Pascali and Roland Wenzlhuemer kindly shared data. Yue Yu, Benjamin Wache, Lili Szabó, Laura Audet, Iain Bamford, Chelsea Carter, Dennis Egger, Benjamin Guillon, Guillermo Carranza Jordan, Claire Lagesse, William Little, Rachel Neumann, Caitlyn Schwartz, Jean Sohn, Pablo Torroella and Jean Philippe Wan provided excellent research assistance. The title of the paper is borrowed from www.spinningtheweb.org.uk. Contact: rj2446@columbia.edu and csteinwe@mit.edu.

1 Introduction

Trade at a distance is difficult. Firms need information about market conditions in faraway locations that are hard to acquire. While it is known that information about the prevailing price in a distant market affects trade flows for certain goods (Allen, 2014; Steinwender, 2018), for many other products, the ability of prices to act as demand signals is limited — especially when products are differentiated by many characteristics (Rauch, 1999). Consider the case of textile traders. To secure an order, information about product characteristics such as the type of yarn demanded, the weave of the cloth in style, or the prints in season are required. This paper goes beyond the price discovery mechanism and studies how information frictions related to the specification of *product characteristics* affect trade flows.

Understanding the effect of such information frictions on trade flows is difficult. It requires evidence about how traders communicate product specifications. In general, this type of correspondence, whether via emails, letters, or phone calls, is highly guarded proprietary information that firms are unwilling to share. Moreover, exogenous variation in communicating product specifications is needed. Improvements in information and communication technology (ICT) provide variation, but adoption is usually endogenous. Firms and governments invest in ICT where gains from trade are expected to be largest. Finally, we need products which vary in the degree to which ICT improvements affect the way that product specifications can be communicated.

In this paper, we overcome these empirical challenges by estimating the causal effect of communication time on imports of different cotton textile products in the 19th century. The setting is useful for the following reasons.

The global telegraph network in the 19th century dramatically reduced communication times and therefore information frictions. For the first time in history, information expressed in words was able to travel faster than physical goods. This aspect is key, as the telegraph greatly reduced information frictions for product characteristics that could be specified in words, but not for other types of information exchange. In this paper, we use the term *codifiability* to refer to the extent to which a product can be specified using only words as opposed to other forms of information exchange.

The 19th century cotton textile industry provides us with rich variation across products in the extent to which product specifications could be codified using only words. In particular, we examine trade in cotton yarn, plain cotton cloth and finished cotton cloth for the years 1845-1880, the time period that spans the roll-out of the global telegraph network. Yarn is a standardized product, meaning that it can be easily described using only one word that consists of the standardized variety. Plain cloth is not standardized, however, its handful of non-standardized characteristics can be described in words, such as the type of weave and the measurement of the cloth. Finished cotton cloth is much harder to describe in words as the intricate patterns inherent to the different finishing techniques cannot easily be described, rather they need to be seen.

Due to the economic importance of the cotton textile industry in the 19th century, records of the correspondence of cotton textile merchants have been preserved. This provides us with a rare opportunity to study how merchants exchanged information about product characteristics at a distance, both before and after they were able to use the telegraph. From business archives that contain the mail correspondence of a large number of textile merchants, we observe that before the telegraph, merchants described product specifications either in words in a letter (for yarn and plain cloth), or by sending a physical product sample (for finished cloth). Both letters and product samples were sent via the postal mail service, which used the fastest modes of transport available, usually a combination of railway and steamship. Pre-telegraph, the time it took to communicate product specifications was the same for all products, regardless of their codifiability.

We examine a telegraph codebook that was developed by cotton textile traders to economize on the use of words to understand how merchants used the telegraph to specify products. The use of the telegraph differed by type of product imported. We find codes that deal with securing an order for yarn or plain cloth, but no code for describing finished cloth. Instead, we find references that condition ordering on the receipt of a physical sample. The telegraph reduced information frictions for the communication of product specifications for codifiable products such as yarn and plain cloth, but not for non-codifiable products such as finished cotton cloth. We exploit this useful variation across products to understand how information frictions affected imports depending on the codifiability of the product.

The roll-out of the telegraph network provides a useful natural experiment to study the effect of a reduction in information frictions. Though countries with higher trade potential had a greater incentive to connect to the telegraph, we isolate exogenous variation in the timing of connection to the network by exploiting the fact that the laying of a successful cable depended on local conditions that were difficult for contemporaries to account for. We show evidence that these difficulties led to significant delays in establishing a working telegraph connection by documenting the number of years that passed between the initial attempt and the eventually successful cable for a number of lines. We use one exogenous local factor, ruggedness of the submarine terrain, to predict the timing of a successful connection. This was a factor that contemporaries were initially unaware of, and later was difficult to account for, given that the sea floor could not be mapped in detail at the time.

To conduct the empirical analysis, we assembled a number of novel datasets. We digitized annual product level imports from Britain for 75 countries for the period of 1845 to 1880 for the three products that form the basis of our analysis: yarn, plain cloth, and printed cloth. We relate trade to our second dataset that annually measures communication time and mail shipping time between London and each importing country by extracting all international shipping information from every May-July issue of the daily publication “Lloyd’s List”. This makes it possible for the first time to systematically examine reductions in communication time worldwide during the roll-out of the telegraph network. Third, for each country, we construct the date and route through which a successful telegraph connection to London was achieved. Finally, we construct an instrumental variable for the arguably endogenous laying of the telegraph network by using the predicted year of the telegraph connection based on the ruggedness of the sea floor along the cable route.

We estimate the effect of the reduction in communication time for each product of interest. We follow the literature and use pseudo-poisson maximum likelihood (PPML) estimation to address the bias in estimating trade elasticities that arise from heteroskedastic error terms in log-linearized models and the presence of zero trade observations, using our ruggedness measures as instruments for communication time, and including country and year fixed effects. We find that communication time improvements had heterogeneous effects on the imports of different products, with the effect being largest for the most codifiable product. According to our preferred specification, a 1% fall in communication time increased imports of yarn by 0.27%, by 0.13% for plain cloth, and by 0.03% for finished cloth, though this latter effect is generally not statistically significant.

We subject the results to a number of robustness checks. The results are robust to controlling for the time-varying effect of product specific import-tariffs, GDP and mail shipping time. Moreover, the results are not driven by the Civil War period, nor by British colonies. In addition, we examine whether other mechanisms besides codifiability could explain the estimated effect. Taking these results together, the paper provides causal empirical evidence for a novel mechanism through which information frictions affect trade flows. Our findings suggest that the effect of ICT on trade in more differentiated goods is heterogeneous across products and depends on the ability of the specific new technology to improve the communication of product specifications, that is, its codifiability with respect to the new technology.

The paper is related to several strands of the literature. The fact that trade at a distance is costly is widely recognized. However, the trade costs that are implied by the flow of goods across countries are much larger than what can be accounted for by directly observable costs such as transportation costs and tariffs (Anderson and van Wincoop, 2004). To explain why trade diminishes so drastically with distance, the literature has turned to examining other factors that may make trade costly, such as information frictions. Previous studies have shown that costly acquisition of price information in distant markets (the "*price discovery mechanism*") affect the spatial dispersion of prices (Jensen, 2007; Aker, 2010; Ejrnæs and Persson, 2010; Goyal, 2010) and trade flows (Allen, 2014; Steinwender, 2018), focusing on homogeneous goods.

However, a large number of products are differentiated by so many characteristics that information about the price alone does not convey the necessary information about market conditions (Rauch, 1999). This paper deepens our understanding about how information frictions matter for trade in these goods by showing the role that the *product specification mechanism* plays. As such, it contributes to a small literature that shows how communication such as telephone calls (Portes and Rey, 2005; Fink et al., 2005), web hosts (Freund and Weinhold, 2004), travel (Poole, 2010;

Cristea, 2011; Startz, 2017) or the presence of ethnic networks (Rauch and Trindade, 2002; Rauch, 2001, 1999) help to overcome information barriers in trade, particularly for differentiated products. Relative to these papers, the main advantage of our setting is that we are able to observe the type of communication that takes place between buyers and sellers. As such, we are able to deepen our understanding of the nature of information frictions that impede trade at a distance.

Codifiability of tasks or product specifications is one mechanism that has been argued to facilitate offshoring and fragmentation of production in the contemporary setting (Leamer and Storper, 2001; Autor et al., 2003). Empirical challenges have limited the extent to which this mechanism has been studied in the literature. The only empirical paper we are aware of in this literature is Fort (2016), who shows that adoption of communication technology that facilitates specifying product characteristics at a distance is systematically related to a firm's decision to outsource its manufacturing processes, in line with predictions by models such as Grossman and Rossi-Hansberg (2008) and Costinot et al. (2013).

Although our outcome variable is not offshoring, but rather trade in different products, we believe our findings have important insights for this literature. In particular, our results suggest that the effect of ICT is technology-dependent. While the telegraph improved the technology of communicating words, different technologies such as the telefax, the internet, or 3-D printing may improve the communication of a different set of product characteristics such as those that can be described by black and white patterns, high-resolution pictures, or physical prototypes. Our findings imply that these distinct technologies should affect trade, fragmentation and outsourcing differently depending on what product (or task) specifications become more easy to communicate. This insight is crucial when thinking about the effects of ICT. For example, a key assumption in Grossman and Rossi-Hansberg (2008) is that reduction to the costs of offshoring affect all tasks in the same way. Our results suggest the effect of ICT on the ability to communicate product specifications at a distance is more nuanced.

The paper is structured as follows. The next section presents evidence on cotton textile traders' information exchange before and after the telegraph. Section 3 introduces the data. Section 4 contains the empirical analysis. Section 5 concludes.

2 Communication of product specifications in the 19th century cotton textile trade

Timely information about distant markets was crucial in 19th century cotton textile trade. Much like the apparel industry today, cotton textiles back then were an industry characterized by rapidly changing product cycles (Llorca-Jana, 2012). As a merchant in Buenos Aires stated; "Fashion here is continually on her change & what is in vogue today may be out of date 2 or 3 months hence."[1] Novelty, however, was not sufficient, as demand was not only volatile, but also highly market-specific; "(...) goods must be manufactured entirely for these markets to make anything by them." (Llorca-Jana, 2012, p. 96).[2] These frequent changes meant that merchants needed to specify attributes of the specific type of product that was in fashion in a particular market at any given time.

Due to the importance of trade in cotton textiles in the 19th century,[3] records of the correspondence of cotton textile merchants have been preserved, both before and after they were able to use the telegraph. Before the telegraph was established, merchants used mail correspondence which is housed today in business archives and has been analyzed by a rich secondary literature. After the telegraph was established, telegraph codebooks were developed by cotton textile traders to economize on the use of words. By analyzing these sources, we are able to understand both how products were specified at a distance by international traders, and how this was affected by ICT.

The three main product categories within cotton textiles; cotton yarn, plain cotton cloth, and finished cotton cloth, differed in the way that product characteristics needed to be specified when securing an order for international trade.

The first category, yarn, includes cotton twist and yarn which can be untreated (plain), bleached, or dyed. Besides the color, the distinctive attribute for a yarn variety is its fineness (or density), for which a standardized, international product classification existed, labeled the "count".[4] The counts were numbers that ranged from 8 to about 140 (with an increment of 2). Around 50-70 product varieties were in use.[5] As a result of this standardization, it was sufficient to specify the color and count when importing yarn.

Plain cloth was a product category that included woven cloth which could be sold unbleached (gray) or bleached. Plain cloth was not standardized, meaning that there was no shorthand terminology that had the same meaning to all parties. Instead, a number of product characteristics had to be specified: the weaving pattern (e.g., how many "picks in the weft"), the length and width of the cloth, the measurements and weaving pattern of borders, whether it was supposed to be bleached or not, and the fineness or quality (which mainly depended on the count of the yarn used). All of these attributes could be specified using a combination of words or numbers. However, compared to yarn, a larger number of non-standardized attributes needed to be specified.

Finished cloth existed in countless varieties that differed in color, the specific printing patterns, or the embroidering pattern. Describing these attributes in words was very difficult, as the following quote illustrates; "Even if agents in the Southern Cone could establish with certainty the sort of

[1] GHR/5/1/3, Hodgson to Green (Liverpool). Buenos Aires, April 24, 1829, quoted in Llorca-Jana (2012).

[2] UGD/28/1/2, Wylie to Daglish (Glasgow). Buenos Aires, December 23, 1809, quoted in Llorca-Jana (2012).

[3] For example, in 1867 (Jan-Jul), cotton manufactures accounted for 40.1% of total exports in British and Irish produce and manufactures. Source: The Economist, July 27, 1867.

[4] The count is a measure of density and still used today. It is the number of hanks (840 yd or 770 m) of skein material that weigh 1 pound (0.45 kg). Under this system, the higher the number, the finer (and thinner) the yarn (Brooks, 1893).

[5] "Liddel's The 'Economic' Telegram Code for Piece Goods and General Business", p. 68

goods that would be in most demand, they still had great 'difficulty of specifying, in words, and at a distance of several thousand miles and several months, exactly what was wanted'."[6] Another merchant complains in a similar way; "it is another proof, out of great many which the writer has seen, of the great risk in sending fancy goods to a foreign market, when not selected by a person who knows exactly the article wanted, as it appears *the plainest written description* which can be given, is seldom correctly understood."[7] To overcome the inherent difficulty of communicating attributes of the finished cloth in words, the standard practice was to send product samples (Llorca-Jana, 2012). Chapman and Chassagne (1981) quote instructions given to the pattern drawer to illustrate how manufacturers used samples of what was in fashion to tailor their product to recent trends; "Enclosed you have a pattern of one of the Bury House's plate furnitures. Joseph Peel desires you will draw up and engrave two or three patterns similar, they should be showey and full of work."; "Try something in this colouring on a smaller scale and with more novelty in the figure. You must do your utmost to get novelty. You may as well do nothing as draw old figure." (Chapman and Chassagne, 1981, p. 84). The interesting aspect of these exchanges is how little instruction can be given in words on precisely what the pattern should look like.

Prior to the introduction of the telegraph, traders used two modes of communication to specify product characteristics for cotton textiles; either in a letter using only words, or by sending an accompanying product sample. Importantly for our setting, both letters and product samples were sent by postal mail. Postal mail included letters (up to a certain size), post cards, newspapers, books, patterns and samples (defined by specific size and weight limits, restricted to "bona fide trade patterns or samples of merchandise", see "British Postal Guide", 1st October 1876, London.), and reached its destination using the fastest method of transportation, usually a combination of railway and steamship. Before the telegraph, the time it took to communicate product specifications was therefore the same for all three cotton textile categories, independent of the medium of information used.

Given the need for timely information, it is unsurprising that the telegraph was rapidly adopted by textile traders. For the first time in history, information expressed in words could travel much faster than the fastest mode of pysical transportation. Llorca-Jana quotes the Argentine consul on the effect of the telegraph; "Goods are now ordered by telegraph as required, and brought out by steamer, thus diminishing market fluctuations and exposing traders to less risk. The absence of season demand is becoming noticeable, the extension of the telegraph and railways to supply their immediate want in lieu of, as formerly, laying in stock of twelve months." (Llorca-Jana, 2012, p. 98).

As a way of minimizing telegraph charges, which were a function of the number of words written, specific codes were developed for traders of cotton textiles and published in telegram codebooks.

[6] Smail (1999) p. 83, quoted in (Llorca-Jana, 2012, pp. 98-99).

[7] UGD/28/1/7, Wylie Cook to Daglish (Glasgow). San Luis de Potosi, June 22, 1834 quoted in (Llorca-Jana, 2012, pp. 98-99), own emphasis.

We examine the extent to which telegram communication differed by product type in “Liddel's The ‘Economic’ Telegram Code for Piece Goods and General Business”.

Figure 3 shows excerpts that are sorted in columns by product category. The first column shows the code that specifies the count of yarn. Consistent with the discussion above, as yarn was standardized and therefore a highly codifiable product, we see the code numbers corresponding to the count of yarn.

The second column shows examples of how the telegraph was used to order plain cloth. Several characteristics described above were in fact communicated using the telegraph. For example, “buy for us a trial lot one pick less in the weft” instructs the exporting merchant about the type of weave that is demanded, “send samples of this made of heavier yarns” instructs the exporting merchant about the type of yarn that should be used in the cloth, while the last photo shows a table that buyers can use to specify measurements. However, it is also clear that it was more costly to specify plain cloth compared to yarn, as more product attributes needed to be communicated in a non-standardized way.

Finally, the third column shows product specifications related to finished cloth. In stark contrast to plain cloth, we find no instructions to the exporting merchants about how the finishing should be done. Conversations about finished cloth were restricted to discussing product samples. For example; “We cannot give the following patterns – we are sending substitutes for approval”. Note how there are no details given about the substitute pattern. Instead, the telegraph was used to negotiate prices and quantities, and other terms and conditions of delivery – everything however, always subject to inspection of a product sample. (Farnie, 2004, p. 38) writes, “Mail advices became restricted to the dispatch of samples, general discussion, hypothetical inquiry, advice, admonition and complaint”.

We use the term *codifiability* to refer to the extent to which attributes of a product can be specified in words, and therefore coded in a telegram, as opposed to inspecting a sample of the product. This definition is key, as the introduction of the telegraph dramatically reduced communication time for exchanging words, but not for exchanging product samples which were still sent by postal mail.[8] The telegraph therefore affected communication times of cotton textile products depending on their codifiability, and in the empirical analysis we expect the telegraph to have had the largest effect on yarn, followed by plain, and then printed cloth.

3 Data

We estimate the effect of communication time on imports of different product categories based on a number of newly constructed datasets covering 75 countries for the years 1845 to 1880. This time period spans the initial roll-out of the global telegraph network. 1851 marks the year the

[8]In the analysis of our communication time data below we show that postal mail shipping time was indeed not affected by the telegraph.

first underwater telegraph was put in place (between Britain and France). By 1880, most countries were connected to the telegraph. The datasets include: 1) An annual product level trade dataset of imports from Great Britain for cotton textile product categories; 2) An annual dataset that contains measures of communication times and mail shipping times between Britain and each country; 3) A dataset that contains the year and route through which a country was first connected to the global telegraph network. We describe each dataset briefly below.[9]

Nineteenth century British trade provide a useful setting in which to estimate our effects of interest. Owing to a number of technological breakthroughs forged in Britain in the late 18th - early 19th century, the country dominated global trade in cotton textiles for much of the nineteenth century (Table 1). British cotton textiles penetrated destination markets worldwide (Figure 1) and trade was important in all three stages of production (Figure 2). British 19th century customs records report both quantities and values of trade, allowing us to disentangle quantity and unit value changes. Moreover, product level data are consistently reported at the level of the three categories that are useful for us to their variation with respect to codifiability: cotton yarn, plain cotton cloth, and finished cotton cloth. We digitized bilateral product level data for cotton textiles from handwritten customs records. The annual data were harmonized across years to yield a strongly balanced panel of 75 countries.

We collected annual data on communication and mail shipping time relative to Britain from "Lloyd's List", a daily London-based publication which printed the most up to date shipping information for ports worldwide. We digitized and extracted all international shipping information from every May, June, and July issue. We used optical character recognition (OCR) to convert images into text files and text matching tools to extract information on the dates of ship movements in foreign ports, and the date the information about this ship movement was published in London. We define the information lag as the difference between the date of the news and the date of publication.[10] "Communication time" (capturing the fastest way of communication, including telegraphic communication) is defined as the minimum across all observations in a port*year, and "mail shipping time" (capturing the transport time of postal mail) is defined as the median. We are able to track communication and mail shipping times separately, because even once the telegraph was adopted for a given port, both forms of technology were used to communicate information to London, depending on the importance of the news.[11] This data makes it possible to, for the first time, systematically analyze the change in communication times and mail shipping times worldwide in the second half of the nineteenth century.

In order to construct a dataset that provides the date and the route taken for the first telegraph connection for each destination market using the global telegraph network, we build on data set

[9] More details on the data collection and extraction are provided in Appendix A.3.

[10] Our method follows Kaukiainen (2001) who collected communication times for a number of ports at decadal intervals from the Lloyd's List between 1820 and 1870.

[11] More details about the construction of the dataset are provided in the online appendix.

covering submarine connections collected by Wenzlhuemer (2013) supplemented with a variety of other, country-specific, sources on international terrestrial connections to identify the date of telegraph connection for all countries. Timing of the first telegraph connection was found by calculating the shortest path between all nodes in the network in each year using Dijkstra's algorithm. We take the first time a country appears in the network as the date at which the country is connected to the global telegraph network.[12]

4 Empirical analysis

In this section we estimate the effect of a reduction in communication time on trade in yarn, plain and finished cloth. We first discuss how the introduction of the telegraph affected communication times around the globe. As the roll-out of the telegraph is potentially endogenous to trade potential of different countries, we discuss our identification strategy based on the ruggedness of the submarine terrain. We then present the estimation results.

4.1 The impact of the telegraph on communication times

It is hard to overstate the impact that telegraphy had on the transmission speed of information expressed in words (Standage, 2009). Prior to the development of this technology, information was only able to travel at the speed of the fastest mode of physical transportation. Even when frontier technology such as steamships and railways were used, the speed of communication was painfully slow, taking days, weeks, or even months.

Our newly collected data on communication time enables us to provide evidence for the astonishing reductions in communication time brought about by the telegraph. Figure 4 plots communication times (solid line) between three important foreign ports and London (Madras, New York and Constantinople). Communication time fell gradually in the lead-up to a telegraph connection, reflecting transportation technology improvements. However, there is a sharp drop in communication times at the vertical lines which mark the year the port was connected to London via the telegraph. For comparison, we are also able to examine mail shipping times (dashed line), which track the number of days it took to ship mail (i.e., letters, newspapers, and product samples) between London and the port in question using the fastest mode of transportation.[13] Communication and mail shipping time fell in parallel up to the year of connection to the telegraph. This reflects the fact that prior to the telegraph, information could only travel as fast as the quickest mode of transportation. Once a port was connected to the telegraph network, however, there was stark divergence between the two, as communication time dropped to within a couple of days, while mail

[12] While our sample period covers only the years 1845-1880, the data on the telegraph network runs until 1899, because we need to know the connection year for countries which are as yet unconnected in 1880 in order to construct our instruments below.

[13] This was not generally the mode of transportation taken for shipping merchandise.

shipping time continued on its previous trend.

Figure 5 shows heat-maps for communication times around the world during our sample period. Communication lags relative to Britain are shown for every decade between 1850 and 1880. In 1850, most of the world is colored in red or orange, meaning it takes several weeks to communicate between London and these countries. Communication time increased with distance, reflecting the fact that it was still determined by the fastest mode of physical transportation. By 1880, almost all ports of the world were within instantaneous communication with Britain. Distance now had very little role to play in determining communication times.

4.2 Identification strategy

The roll-out of the telegraph network provides a useful natural experiment to study the effect of a reduction in information frictions. Figure 6 shows the first telegraph links via which countries in our sample were connected to the network. We define a country as having a connection to the network if it has an active (terrestrial and/or submarine) connection to London.

As is apparent from the figure, it took decades for all countries to be connected to the network. One reason for the gradual roll-out of the telegraph is that the laying of a successful cable depended on local conditions that were difficult for contemporaries to account for. This is a crucial aspect of our setting, as it allows us to isolate exogenous variation in connection to the network. Table 2 shows the year of the first attempted connection and the year of successful connection for some submarine cables. A number of important points are worth noting. First, the time between the years passed between first and successful connection was large; in some cases spanning over a decade. This is key as it suggests that a potentially significant part of the variation in the timing of connection was driven by exogenous factors. Second, the difficulty in laying a cable was not systematically related to distance. Relatively short cables such as the one between Greece and Egypt could take far longer to succeed than longer ones such as the Ireland – Canada connection which was the key link for the first transatlantic cable. Third, the table supports the claim that the telegraph needed to be adapted to local conditions. Even in the 1870s, when much longer cables crossing oceans had been successfully laid, cables such as the Panama - Jamaica link took multiple attempts to succeed.

Our identification strategy relies on one geographic aspect of laying submarine cables that introduced quasi-random variation in the timing of a connection; the ruggedness of the underwater terrain. A rugged sea-floor made laying a telegraph cable particularly difficult as the cable would break if it became suspended between two peaks (Company, 1915). Laying a cable across a rugged sea bed also meant that the cable had to be longer, which made the transmitted signal weaker and noisier, because automatic signal re-transmission had not yet been invented. It also made the cable heavier, which put more strain on it, making it easier to break, e.g., when it chafed against sharp rocks as the cable moved with the currents. Variation in elevation also meant that the length of the cable needed to be longer and it made estimating the correct length more difficult. Given the state of contemporary technology, mapping the sea floor in any meaningful detail was impossible. Without detailed knowledge of the submarine terrain, progress was made through trial and error.

To illustrate the extent to which contemporaries were initially unaware of the ruggedness of the submarine terrain along a given connection, it is instructive to compare the elevation profile of the ocean floor along the Atlantic cable route (referred to as the Canada – Ireland connection in Table 2) in Figure 7 with the description the New York Times gave based on the soundings that had been taken; "[the telegraphic plateau] extends in a *continuous ledge* from Cape Clear, in Ireland, to Cape Race in Newfoundland; the greatest depths being in mid-ocean whence it *ascends imperceptibly* to the shore on either side."[14] Contemporaries were aware neither of the steep drop in elevation on both side of the Atlantic, nor the significant amount of ruggedness along what they called the telegraphic plateau, that modern data shows in Figure 7. This was problematic, as the amount of slack cable that should have been laid depends crucially on getting the elevation changes right. In the end, it took five attempts over the course of almost a decade to establish the first successful transatlantic telegraph connection.[15]

In fact, ruggedness of the submarine terrain is argued to have been a factor delaying a successful connection for all the connections in Table 2. Headrick (1991, p. 18) writes the following about the initial failure of the Egypt – India cable; "The cause of the disaster soon became apparent.(...) Soundings taken in early 1858 showed a soft bottom, but no proper sea-bed survey was ever made. The cable was laid in a straight line without slack; so it hung between underwater peaks and soon broke under the weight of the barnacles that fastened onto it.". Similarly, according to Ahvenainen (2011, p. 30-31); "The Mediterranean was a difficult area in which to operate (...). The sea charts were often incomplete, they were behindhand with the soundings and sudden variations in depth made it difficult to estimate the amount of cable needed."

One final example relates to the difficulty in laying submarine cables in the Caribbean, another location that witnessed significant delays according to Table 2; "The Caribbean sea floor did not favour submarine cables. The volcanic formation of the region was marked by great and sudden variations in the depth of the water, with a rocky, coralline and occasionally metalliferous sea floor on which the cables wore themselves out or were chemically affected until they broke." (Ahvenainen, 1996, p. 42).

Our identification strategy isolates exogenous variation in the year a country was connected to the telegraph by using variation in the ruggedness of the submarine terrain to predict the year of the first successful connection to a given country. To measure ruggedness of the submarine terrain, we use the Riley measure of local ruggedness as proposed by Riley et al. (1999) and first

[14] "The Atlantic Cable. Successful Completion of the Great Work", *The New York Times*, July 30, 1866 (own emphasis).

[15] Appendix **??** discusses additional historical evidence regarding the difficulty of laying submarine cables.

used in the economics literature by Nunn and Puga (2012).[16] As a robustness check, we also construct a second measure, *normalized cable length*, defined as the ratio between 3-dimensional and 2-dimensional distance along the shortest path. This measure naturally captures one way in which historical accounts claim ruggedness mattered; the amount of "slack" that is needed because of underwater peaks and troughs.

We operationalize both ruggedness measures as follows. The nodes that make up a given connection between a country c and London are taken from the global telegraph network that we constructed (Figure 6). This is the shortest path along which a country was first connected to London, using any combination of overland and submarine connections. To locate the actual path along which ruggedness will be measured, we take the two endpoints of all submarine connections and find the shortest sea-route between them. Note that the actual cable might have been laid on a different route. We prefer to use the shortest distance route, as ex ante this is the most economical (longer cables were costlier) and hence exogenous, while the path of the actual cable may be endogenous if routes were changed after difficulties were experienced. The Riley measure of ruggedness and the normalized cable length measure are then calculated along this route.

Calculation of the *Riley measure* proceeds as follows. To calculate ruggedness in a representative area around the shortest path, we take a 25km buffer along both sides of the shortest route. Within this area, we then calculate a measure of local ruggedness as proposed by Riley et al. (1999). This measure is designed to capture local changes in the terrain. This is well-suited to our purposes, as it was precisely local variation in elevation that was difficult to map for contemporaries, yet crucially important for the success of a given link. As knowledge about laying submarine cables advanced, depth was measured at a few points along a proposed route, giving an approximation of large changes in elevation. However, it was not possible to map the more local aspects of ruggedness, which is the type of terrain variation that our measure utilizes.

Riley ruggedness is defined at a point in space. The measure is the square root of the sum of the squared differences in elevation between a point and its eight neighbors (ie. the point to the north, north-east, east, south-east, south, south-west, west and north-west).[17] We calculate this measure of ruggedness for each cell within the 25km buffer. Following the methodology in Nunn and Puga (2012), we take the mean of the ruggedness measure (defined at the level of individual cells) along the 25km buffer of the given connection (edge), which we denote as R_i (mean ruggedness for edge i). In the majority of cases, a connection between London and a country was made up of multiple edges. We define the ruggedness of a connection in our dataset as the maximum across all edges that form a particular connection. This captures the idea that the edge with the largest ruggedness measure will be the binding constraint for establishing the full telegraphic connection to that place.

[16]Ruggedness measures are calculated using the General Bathymetric Chart of the Oceans (GEBCO). This contains elevation data at the fineness of a thirty arc-second grid.

[17]More formally, let e_{rc} denote elevation at the point located in row r and column c of a grid of elevation points. The terrain ruggedness index at point rc is calculated as $\sqrt{\sum_{i=r-1}^{r+1}\sum_{j=c-1}^{c+1}(e_{ij}-e_{rc})^2}$.

Our alternative measure of normalized cable length is calculated as the ratio between 3-dimensional and 2-dimensional distance along the shortest path (including only submarine edges).

We predict the year a country is first connected to London based on these two ruggedness measures using a linear probability model. This yields non-integer predictions, which we round to the nearest integer. We need integer predictions of the year of connection to the telegraph, as we will use this instrument in a panel regression where it will take the value of one in years including and after a predicted connection, and zero otherwise.

Figure 8 and Table 3 show the scatterplot and the estimates from the linear prediction model.[18] As the figure shows, ruggedness seems to be a better predictor of year of connection for early years, while the relationship seems weaker from about 1880. This is consistent with ruggedness of the submarine terrain mattering most at the earlier phases of the roll-out when the technology was still being developed.

While we have shown some suggestive evidence that short cables could be at least as difficult to lay as longer ones; there is a concern that longer connections may be more likely to feature areas of large elevation changes, meaning that distance to London may be somewhat correlated with these measures. For this reason, we use only the variation in ruggedness not explained by distance to London when predicting the year of the telegraph connection, as shown in Column (2) of Table 3. Another concern is that ruggedness may be correlated with shipping time, for example, because submarine ruggedness may be correlated with currents affecting ship speed. We show this is not the case in Column (3) – ruggedness does not predict mail shipping time. To build evidence in support of the underlying assumption that ruggedness is not correlated with country-time trends, we show that ruggedness is also uncorrelated with growth in yarn, plain and finished cloth imports (Columns (4)-(6)) prior to the introduction of the telegraph.

4.3 The effect of communication time improvements on textile imports

As previous sections have shown, the telegraph was rapidly adopted by cotton textile merchants to facilitate the exchange of information at a distance in an industry characterized by rapidly fluctuating demand. In this section, we ask whether communication time improvements differentially affected trade in products depending on their codifiability. We estimate the elasticity of imports to changes in communication time separately by product category using the following specification;

$$E[imp^i_{ct}|C_{ct}] = \exp\{\beta^i * \ln C_{ct} + \lambda_c + \gamma_t\} \tag{1}$$

imp^i_{ct} denotes quantity imported of product category i (yarn, plain cloth, finished cloth) in country c at time t. C_{ct} denotes communication time as defined in the previous section. λ_c and γ_t are country and time fixed effects, respectively. The parameters of interest are β^i, which

[18] Note that we use the full dataset of connections up to 1899 for this prediction, as we need to have a predicted year of connection for those countries that are not connected to the telegraph by 1880.

capture the elasticity of imports of product category i with respect to changes in communication time. We use Poisson pseudo-maximum likelihood (PPML) to estimate the equation following the extensive literature showing PPML to be a consistent and relatively efficient estimator for gravity equations featuring a large number of zeros and a multiplicative estimating equation (Santos Silva and Tenreyro, 2006; Head and Mayer, 2013).[19] We examine quantity imported as opposed to the value of imports so as to not confound price and volume effects. Standard errors are clustered at the level of countries to account for country level serial correlation in the error terms.

The endogeneity concern in equation (1) is that communication time between two locations is partially determined by ICT investments, which may be related to the trading potential between those locations. Countries that traded intensively with each other had larger incentives to invest in faster communication infrastructure. On the other hand, political considerations may have resulted in larger investments in infrastructure for more peripheral or conflict-ridden regions where trading potential was lower.[20]

As discussed in the previous section, we address these concerns by using the variation in telegraph connections that is predicted based on ruggedness. We transform the predicted telegraph year into a binary variable that takes the value of zero in years preceding a predicted connection, and one thereafter. However, the telegraph had a differential effect on the communication times of different countries: while communication time to France was reduced only by 1 day, communication time to Australia was reduced by several weeks. To exploit this cross-sectional variation in communication time reductions, we interact our predicted telegraph dummy with the log average reduction in the communication time of a country's neighbor. We use the communication time drop of the neighboring country rather than the country itself in order to avoid potential reverse causality: If a country was more important in terms of trade or trade potential, its communication times before the telegraph might have been influenced, e.g., by using faster mail ships instead of slower sail ships before the telegraph. Neighbors are defined according to sea-distance between the ports used for each country. We use the neighbor with the largest number of observations among the three closest neighbors of a country. The log drop in communication time is defined as the log of the difference in communication time across all years pre- and post-telegraph (excluding the year of the telegraph connection).

While there is no first stage defined in GMM similarly to that of the linear 2SLS case, it is still important to assess the strength of relationship between the endogenous variable, communication time, and each instrument, because weak instruments can invalidate the GMM estimates. In Table 4, we therefore report the equivalent first stages from an OLS regression with two-way fixed effects.

[19] Estimation of the relationship using ordinary least squares does not yield similar results. The reason in this particular case seems to be because of heteroscedasticity in the error term rather than differential shares of zeros, because non-linear least squares, which is a less efficient method to deal with heteroskedasticity yields similar results to PPML and diagnostic tests suggested by Head and Mayer (2013) also point to PPML being the preferred specification. Appendix A.1 contains a more in depth discussion of these issues.

[20] Headrick (1981) discusses in detail the extent to which the telegraph was a tool for empire building of the British.

It is reassuring to see that both instruments are highly statistically significant and have the expected sign.

Estimation of a Poisson regression model with endogenous regressors is implemented using generalized method of moments (GMM), following Windmeijer and Santos Silva (1997) and estimated using Stata's `ivpoisson` command. It should be noted that non-linear regressions with fixed effects may suffer from the incidental parameters problem (Neyman and Scott, 1948) . Fernández-Val and Weidner (2016) have shown that, for the case of exogenous or predetermined regressors, Poisson models do not suffer from the incidental parameters problem. However, it may well be a concern for the IV Poisson model. If both N and T are relatively large, as is the case in our data with 72 countries and 36 years, the bias arising from the incidental parameters problem is unlikely to be large. Nonetheless, to gain a better sense of the potential size of the bias, we implement a panel jackknife bias correction that has recently been proposed in the literature (Hahn and Newey, 2004; Fernández-Val and Weidner, 2016; Cruz-Gonzalez et al., 2017).

4.3.1 Baseline Results

Table 5 contains summary statistics for quantities and values imported for each group of products across the sample period. The table also splits the sample according to years before and after the telegraph was put in place in. This provides a first assessment of the extent to which the effect of the telegraph is apparent in the raw data without imposing additional structure. As can be seen, both in terms of quantities and in terms of import values, the percentage increase in imports after the telegraph was put in place is largest for yarn, followed by plain and finished cloth. We now turn to estimating the effect of interest; the elasticity of imports with respect to communication time.

Table 6 presents the baseline results. Panel A reports the coefficients for the effect of communication time on yarn, plain and finished cloth in the different columns, using country and year fixed effects.[21] We see a clear pattern: a 1% decrease in communication time increases the quantity of yarn imported by 0.183% (s.e. 0.041), plain cloth by 0.097% (s.e. 0.029), while there is no statistically significant effect on finished cloth (the point estimate is effectively zero; 0.006).

In order to explore which components of the time series and cross-sectional variation used are important for our results, Table 7 explores alternative regressors. Results using only (time series) variation from the telegraph dummy are similar, but less precise, which is to be expected as the cross-sectional variation in communication time drops is not exploited, which adds a lot of noise to the estimation. Results using the telegraph dummy interacted with the average communication time drop of a country yield similar results to those in Panel A (though the interpretation of the estimated coefficients are different). The importance of the cross-sectional variation motivates our inclusion of a plausibly exogenous component of it (the closest neighbor's drop in communication

[21] Results without fixed effects are shown in Table A.1.

time) in the instrument.

Panels B and C of Table 6 report the IV estimates based on the Riley measure and the normalized cable length, respectively. The estimated coefficients decrease in size as we move from yarn to plain and then to finished cloth, consistent with our previous results. The coefficients are very similar across the two IVs. Comparison of the coefficients between the Poisson specification and IV estimates shows that while the IV estimates are marginally larger, the two are fairly similar and statistically indistinguishable. This is suggestive of the fact that the timing of successful telegraph connections is predominantly determined by exogenous factors such as ruggedness of the submarine terrain, consistent with the historical evidence discussed previously.

In terms of magnitudes in panel B, a 1% drop in communication time is predicted to increase quantity imported by 0.272% (s.e. 0.076) for yarn, 0.128% (s.e. 0.065) for plain cloth, and by 0.030% (s.e. 0.066; not significant) for finished cloth. To gain a better sense of the economic effect of the telegraph, we scale the elasticities by the average drop in communication time caused by the telegraph. Based on this, introduction of the telegraph increased imports of yarn by 10.6%, of plain cloth by 5.0% and of finished cloth by 1.2%, though this latter effect is statistically indistinguishable from zero, consistent with the historical evidence that trading finished cloth required the physical shipping of samples, which could not be done via telegraph.

The difference between the estimated effect on yarn and the effect on plain cloth is statistically different (as is the estimated difference between yarn and finished cloth), as we show in Table A.2 of the online appendix, which presents the pooled baseline and IV specifications. The difference between the effect on plain and finished cloth is statistically different in the Poisson specification, but not in either IV specification.

4.3.2 Robustness

We explore the robustness of our results along a number of dimensions. First, Table 8 examines robustness of the results for various subsamples of the data.[22] Panel B reports results for the sample that excludes years during the American Civil War (1861-1865), when the supply of raw cotton to the British textile industry was severely disrupted because of the blockade of the raw cotton exporting Southern states (Hanlon, 2015). Comparison with the baseline IV results show virtually no change in the size of the coefficients and their statistical significance.

Panel C estimates the effects of interest by excluding British colonies from the sample. It is likely that trade with colonies was more regulated thus restricting the impact of communication time on trade. Indeed, the gradient of the effect across product categories is larger when British colonies are dropped, suggesting that our results hold (and are even stronger) in contexts that

[22]The robustness checks are based on the instrument that uses the Riley measure of ruggedness. We use this as our baseline instrument as this measure has been used previously in the literature. All results are similar when we use normalized cable length as reported in Table A.3, but it should be noted that the effect on plain cloth is not always statistically significant when normalized cable length is used as the instrument.

resemble competitive markets more closely.

Panel D includes time-varying product level ad-valorem tariffs from Tena-Junguito et al. (2012) for the subset of countries for which they are available. The coefficients of interest remain similar in both magnitude and significance, confirming that our identification strategy is robust. Tariffs enter with the expected negative sign, though they are only statistically significant for finished cotton cloth.

Panel E includes controls for GDP for the countries for which they are available using data from the Maddison Project Database (Bolt et al., 2018) and from Fouquin and Hugot (2016). The point estimates for the coefficients of interest retain their pattern, however, the estimated effect on yarn is somewhat larger and plain cotton cloth is no longer different statistically from zero. However, the latter is not due to the effect of including GDP, but rather the different (and substantially smaller) subsample on which this specification is estimated. The coefficients of interest are virtually identical when the sample is restricted to the same subsample and the control for GDP is omitted.

Panel F shows the estimated effect on import values. Interestingly, the magnitudes of the effect are smaller and insignificant, even though the same gradient of the effect across product categories is present. This suggests that unit values decreased as the quantity imported increased.

Table A.5 in the online appendix contains the panel jackknife bias correction to assess the size of bias arising from the incidental parameters problem. The effects on yarn and finished cloth are basically unchanged, but the effect on plain cloth is slightly reduced, turning it marginally insignificant. Reassuringly, the gradient of the effect across product categories is unchanged.

4.3.3 Alternative mechanisms

The results presented above show that communication time improvements driven by connection to the telegraph network had a larger effect on more codifiable product categories, that is, when buyers and sellers could specify and order at a distance using only words instead of needing to exchange product samples. Are there other mechanisms besides codifiability that could explain the empirical findings?

First, better communication technology may have reduced contracting frictions by increasing trust (Nunn, 2007). Contracting frictions are expected to be larger in product categories that are less standardized, as the hold-up problem is less severe for these products. If better communication increased contract enforcement, we would expect to see the pattern across products to be the opposite of what we find, so this is unlikely to explain our results.

It is also possible that the telegraph enabled buyers and sellers to sample more potential match partners in foreign countries. This could increase the match quality and thereby trade. Notice, however, that in this particular setting, the large majority of trade occurred between agents of multinational merchant houses, i.e., within the boundaries of the firm (Chapman, 1992). Trade within subsidiaries of the firm does not require search for trading partners, nor imply important contracting frictions. It is therefore unlikely that these explanations drive our findings. Nevertheless, we are able to test the robustness of our results to these two alternative mechanisms more formally, using data we collected on the number of international merchant houses active in a particular country in the cotton textile trade. In Panel A of Table 9, we explore whether our effect is indeed weaker or insignificant for those countries where international merchant trading houses were widespread, by adding an interaction term between the coefficient of interest and the number of international merchants. The effect on the interaction term is very small and not large enough to offset the pattern across product categories that we see, suggesting that contract enforcement or better match quality are unlikely to be the driving force behind the different estimated elasticities.[23]

Communication time could have a different effect on product categories if demand volatility varied across them. The larger volatility of demand, the more valuable is up to date information. Demand volatility would have to be largest for yarn and smaller for finished cloth in order to explain our empirical results. However, we believe that the opposite is more likely to be true, as there were a larger number of varieties, and it was therefore harder to predict fashion for finished cloth compared to yarn.

Another feature of our setting is that there is an input-output relationship between the product categories of our data. Yarn is an input to plain and finished cloth, and plain cloth can be printed and embroidered to make finished cloth. Could the same communication time shock have differential effects on imports of specific products because of their input-output linkages? Considering that one variety of plain cloth can be transformed into many varieties of finished cloth (by applying different printing patterns, for example), aggregation implies that the demand volatility for a specific variety of plain cloth is smaller than that for a specific variety of finished cloth. More timely information would therefore be expected to affect imports of finished cloth by more than plain cloth, the opposite of what we find. A similar argument can be made for yarn versus the other product categories.

Furthermore, assume demand elasticities are largest for yarn and smallest for finished cloth for an unspecified reason. In this case, the same 1% decrease in price will increase yarn imports by more than imports of finished cloth. However, note that this effect should hold for *any* reduction in trading frictions. To test for this, we can exploit the fact that data from the Lloyd's List also yields mail shipping time for letters and product samples, and explore whether the effects are similar when a different trading friction decreases. Panel B of Table 9 shows that while the estimated coefficients for mail shipping time are not statistically different from zero, the point estimates *increase* in magnitude as we move further downstream, which is precisely the opposite of what we find for communication time. This suggests that the main driver of the effect of communications times is not the generic effect of a reduction in trading frictions.

Interestingly, however, our preferred mechanism — codifiability — *is* highly consistent with mail shipping time having the opposite effect: reductions in mail shipping time, which are also

[23] Table 9 uses the Riley ruggedness measure as the instrument, the results with the cable length instrument are shown in Table A.4.

used for mailing samples, are most relevant for the non-codifiable product category; finished cloth.

5 Conclusion

This paper has shown what is, to the best of our knowledge, the first evidence for a novel mechanism through which information frictions affect trade; *the product specification mechanism.* We use the introduction of the global telegraph network as a natural experiment that affected the time it took to communicate product specifications expressed in words. We show that in the cotton textile industry, imports of more codifiable products – that is, products whose specifications can be formulated in words – increased more in response to the same decrease in communication time.

To what extent do the results from this historical episode inform our understanding of the role that information frictions play in impeding trade in the present day – a setting in which ICT is far more sophisticated? Our results suggest a nuanced view of how a specific ICT improvement will affect trade. In particular, when viewing the effect of a technology for products that are sufficiently differentiated such that the price mechanism in itself is not a sufficient signal of market conditions, in light of our findings, it is crucial to examine how the ability to communicate product specifications has changed. While the telegraph had an enormous effect on communication time for information expressed in words, it had no effect on the time it took to transmit an image, or a prototype. In contrast, 3D printing, a relatively new ICT, has no impact on the time it takes to exchange information expressed in words, but allows fast transmission of prototypes. In light of our results, these two ICTs would affect trade in particular products differently.

The focus on ICT's ability to enhance communication of product specifications is particularly important in the present day in which value chains are increasingly fragmented spatially. The literature to date has speculated that trade in intermediates and global value chains (GVCs) gained in importance only recently because communication costs did not meaningfully decrease until well into the 20th century (e.g., Baldwin and Martin (1999)). This view is in stark contrast with the revolutionary reduction in communication times that the telegraph and later the telephone made possible. Our paper suggests a way of resolving this tension. If the effects of ICT on differentiated products are technology dependent, then subsequent waves of ICT improvements such as the telegraph and the telephone that both facilitated only the communication of product specifications expressed in words may not have had much impact on the ability to communicate more complex product specifications that require transmitting high resolution images or prototypes. Transmission of the latter type of information are often what is required for (predominantly manufactured) inputs that form spatially fragmented value chains today. It also suggests that future improvements in ICT will have the biggest effect on the range of products traded at a distance if they facilitate the transmission of product characteristics that previous waves of ICT did not.

References

Ahvenainen, J. (1996). *The history of the Caribbean telegraphs before the First World War.* Acad. sci. Fennica.

Ahvenainen, J. (2011). *The history of the Near Eastern telegraphs: before the First World War.* Acad. Scientiarum Fennica.

Aker, J. C. (2010). Information from markets near and far: Mobile phones and agricultural markets in niger. *American Economic Journal: Applied Economics 2*(3), 46–59.

Allen, T. (2014). Information frictions in trade. *Econometrica 82*(6), 2041–2083.

Anderson, J. E. and E. van Wincoop (2004, September). Trade costs. *Journal of Economic Literature 42*(3), 691–751.

Autor, D. H., F. Levy, and R. J. Murnane (2003). The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics 118*(4), 1279 – 1333.

Baldwin, R. and P. Martin (1999). Two waves of globalisation: superficial similarities, fundamental differences. *National Bureau of Economic Research w6904*.

Bolt, J., R. Inklaar, H. de Jong, and J. L. van Zanden (2018). Rebasing 'Maddison': new income comparisons and the shape of long-run economic development. *Madison Project Working Paper 10*.

Brooks, C. P. (1893). *Weaving calculations: a guide to calculations relating to cotton yarn and cloth and all processes of cotton weaving* (3 ed.). E. & F. N. Son.

Chapman, S. (1992). *Merchant enterprise in Britain: From the industrial revolution to World War I.* Cambridge University Press.

Chapman, S. D. and S. Chassagne (1981). *European textile printers in the eighteenth century: A Study of Peel and Oberkampf.*

Company, T. C. C. (1915). *How submarine cables are made, laid, operated and repaired.* The Commerical Cable Company.

Costinot, A., J. Vogel, and S. Wang (2013). An elementary theory of global supply chains. *The Review of Economic Studies 80*(1), 109–144.

Cristea, A. D. (2011). Buyer-seller relationships in international trade: Evidence from us states' exports and business-class travel. *Journal of International Economics 84*(2), 207–220.

Cruz-Gonzalez, M., I. Fernández-Val, and M. Weidner (2017). probitfe and logitfe: Bias corrections for probit and logit models with two-way fixed effects. *Stata Journal fothcoming*.

Ejrnæs, M. and K. G. Persson (2010). The gains from improved market efficiency: Trade before and after the transatlantic telegraph. *European Review of Economic History 14*(03), 361–381.

Ellison, T. (1886). *The Cotton Trade of Great Britain.* Effingham Wilson.

Farnie, D. A. (2004). The role of merchants as prime movers in the expansion of the cotton industry, 1760-1990. In D. A. Farnie and D. J. Jeremy (Eds.), *The Fibre that Changed the World: The Cotton Industry in International Perspective, 1600-1990s*. Oxford University Press.

Fernández-Val, I. and M. Weidner (2016). Individual and time effects in nonlinear panel models with large N, T. *Journal of Econometrics 192*(1), 291–312.

Fink, C., A. Mattoo, and I. C. Neagu (2005). Assessing the impact of communication costs on international trade. *Journal of International Economics 67*(2), 428–445.

Fort, T. C. (2016). Technology and production fragmentation: Domestic versus foreign sourcing. *The Review of Economic Studies*, 1–38.

Fouquin, M. and J. Hugot (2016). Two centuries of bilateral trade and gravity data: 1827-2014.

Freund, C. L. and D. Weinhold (2004). The effect of the Internet on international trade. *Journal* *of International Economics 62*(1), 171–189.

Goyal, A. (2010). Information, direct access to farmers, and rural market performance in central india. *American Economic Journal: Applied Economics 2*(3), 22–45.

Grossman, G. M. and E. Rossi-Hansberg (2008). Trading tasks: A simple theory of offshoring. *The American Economic Review 98*(5), 1978–1997.

Hahn, J. and W. Newey (2004). Jackknife and analytical bias reduction for nonlinear panel models. *Econometrica 72*(4), 1295–1319.

Hanlon, W. W. (2015). Necessity is the mother of invention: Input supplies and directed technical change. *Econometrica 83*(1), 67–100.

Head, K. and T. Mayer (2013). Gravity equations: Workhorse, toolkit, and cookbook. In G. Gopinath, E. Helpman, and K. Rogoff (Eds.), *Handbook of International Economics*, Volume 4. Elsevier.

Headrick, D. R. (1981). *The tools of empire: Technology and European imperialism in the nineteenth century*. Oxford University Press.

Headrick, D. R. (1991). *The invisible weapon: Telecommunications and international politics, 1851-1945*. Oxford University Press.

Hugot, J. and C. U. Dajud (2016). Trade costs and the Suez and Panama Canals. Technical report, CEPII.

Jensen, R. (2007). The digital provide: Information (technology), market performance, and welfare in the south indian fisheries sector. *The Quarterly Journal of Economics 122*(3), 879–924.

Kaukiainen, Y. (2001). Shrinking the world: Improvements in the speed of information transmission, c. 1820–1870. *European Review of Economic History 5*(1), 1–28.

Leamer, E. E. and M. Storper (2001). The economic geography of the Internet age. *Journal of International Business Studies 32*(4), 641–655.

Llorca-Jana, M. (2012). *The British textile trade in South America in the nineteenth century*, Volume 97. Cambridge University Press.

Neyman, J. and E. L. Scott (1948). Consistent estimates based on partially consistent observations. *Econometrica 16*(1), 1–32.

Nunn, N. (2007). Relationship-specificity, incomplete contracts, and the pattern of trade. *The Quarterly Journal of Economics 122*(2), 569–600.

Nunn, N. and D. Puga (2012). Ruggedness: The blessing of bad geography in Africa. *Review of Economics and Statistics 94*(1), 20–36.

Poole, J. (2010). Business travel as an input to international trade.

Portes, R. and H. Rey (2005). The determinants of cross-border equity flows. *Journal of International Economics 65*(2), 269–296.

Rauch, J. E. (1999, June). Networks versus markets in international trade. *Journal of International Economics 48*(1), 7–35.

Rauch, J. E. (2001, December). Business and social networks in international trade. *Journal of Economic Literature 39*(4), 1177–1203.

Rauch, J. E. and V. Trindade (2002, February). Ethnic Chinese networks in international trade. *The Review of Economics and Statistics 84*(1), 116–130.

Riley, S. J., S. D. De Gloria, and R. Elliot (1999). A terrain ruggedness index that quantifies index that quantifies topographic heterogeneity. *Intermountain Journal of Sciences 5*(1-4), 23–27.

Santos Silva, J. and S. Tenreyro (2006). The log of gravity. *The Review of Economics and Statistics 88*(4), 641–658.

Standage, T. (2009). *The Victorian Internet: The Remarkable Story of the Telegraph and the Nineteenth Century's Online Pioneers*. Bloomsbury Publishing.

Startz, M. (2017). The value of face-to-face: Search and contracting frictions in Nigerian trade. *mimeo*.

Steinwender, C. (2018). Information frictions and the law of one price: "when the states and the kingdom became united". *American Economic Review 108*(3), 657–696.

Tena-Junguito, A., M. Lampe, and F. T. Fernandes (2012). How much trade liberalization was there in the world before and after Cobden-Chevalier? *The Journal of Economic History 72*(3), 708–740.

Wenzlhuemer, R. (2013). *Connecting the nineteenth-century world: The telegraph and globalization.* Cambridge University Press.

Windmeijer, F. and J. Santos Silva (1997). Endogeneity in count data models: An application to demand for health care. *Journal of Applied Econometrics 12*(3), 281–294.

A Tables

Table 1: Exports of cotton goods and yarn, by country

	Exports of cotton goods and yarn	
	Value exported 1876-1880	Ratio relative to British 1876-1880
Great Britain	68,457	
Holland	2,818	0.041
Germany[1]	2,727	0.040
France	2,616	0.038
United States	2,008	0.029
Belgium	770	0.011

Notes: British exports are declared value. There are no recorded exports of cotton goods and yarn from Russia, Austria, Spain or Italy. Source: Ellison (1886, p. 113).

[1] German exports are estimates.

Table 2: Date of first and successful telegraph attempt for selected cables

Cable link	First attempt	Successful attempt	Distance (km)
Ireland - Canada	1857	1866	3,070
Egypt - India	1858	1870	3,059
Algeria - France	1861	1870	740
Greece - Egypt	1856	1873	650
Panama - Jamaica	1870	1875	1,010

Table 3: Prediction of first successful telegraph connection based on Riley ruggedness and placebo checks

VARIABLES	(1) Telegraph year	(2) Telegraph year	(3) Mail time	(4) Growth yarn	(5) Growth plain cloth	(6) Growth finished cloth
Riley ruggedness	0.029***	0.026***	-0.011	-9.619	0.310	56.166
	(0.004)	(0.004)	(0.010)	(84.406)	(98.647)	(66.368)
ln(distance to London)		0.000***	0.007***	60,869.7**	5,705.4	-11,533.9
		(0.000)	(0.000)	(29,737.5)	(23,755.4)	(20,171.5)
Observations	73	73	63	56	63	62
R-squared	0.499	0.531	0.648	0.138	0.002	0.010

Notes: Estimation via linear probability model. Riley ruggedness defined as the maximum Riley measure across all submarine edges that make up a connection. Distance to London is the natural logarithm of great circle distance to London. Growth in columns 4-6 defined as the change in imports between 1846 and 1850 relative to average imports in 1846 and 1850. The denominator is the average of imports in the two periods because the large number of zeros would reduce the sample significantly if defined only for 1846. Robust standard errors in parentheses. Notation for statistical significance; *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 4: First stage equivalent

VARIABLES	(1) ln(comm time)	(2) ln(comm time)
(Pred tele Riley)*	-0.297***	
(comm time change neighbor)	(0.046)	
(Pred tele cable length)*		-0.296***
(comm time change neighbor)		(0.044)
Observations	2,150	2,150
R-squared	0.557	0.557
Number of countries	72	72

Notes: OLS specification for the pseudo-first stage. ln(commtime) defined as the natural logarithm of communication time (in days) to London. Regressors: (Pred tele Riley)* (comm time change neighbor) is a binary variable that takes the value of one including and after the predicted year of connection based on the Riley measure interacted with the closest neighbor's drop in communication time after connection to the telegraph. (Pred tele cable length)* (comm time change neighbor) is a binary variable that takes the value of one including and after the predicted year of connection based on the normalized cable length measure interacted with the closest neighbor's drop in communication time after connection to the telegraph. Year and country FEs included. Standard errors clustered by country in parentheses. Notation for statistical significance; *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Summary statistics

	All	**Before telegraph**	**After telegraph**	**Change**	**Change in %**
Num. observations (country*year)	2,684	1,884	800		
Quantities					
Yarn, million pounds[1]	2.33	0.78	5.99	5.21	668
Plain cloth, million yards	21.73	12.25	44.05	31.80	260
Finished cloth, million yards	10.72	7.45	18.41	10.95	147
Values (000's of £'s)					
Yarn	138.23	40.09	369.34	329.24	821
Plain cloth	300.73	172.97	601.60	428.62	248
Finished cloth	203.20	138.96	354.51	215.54	155

Notes: "Before telegraph" includes all country*year observations prior to achieving a connection to the telegraph network. "After telegraph" includes all country*year observations including and after achieving a connection to the telegraph network. On average, 1 pound of yarn yields 5 yards of cloth (Ellison, 1886).

Table 6: Baseline specifications

VARIABLES	(1) Yarn	(2) Plain cloth	(3) Finished cloth
Panel A. Baseline, no IV			
ln(comm time)	-0.183***	-0.097***	0.006
	(0.041)	(0.029)	(0.044)
Observations	2,150	2,150	2,150
Nr of countries	72	72	72
Panel B. Riley measure * comm time drop			
ln(comm time)	-0.272***	-0.128**	-0.030
	(0.076)	(0.065)	(0.066)
Observations	2,150	2,150	2,150
Nr of countries	72	72	72
Panel C. Normalized cable length * comm time drop			
ln(comm time)	-0.269***	-0.110*	0.007
	(0.077)	(0.065)	(0.064)
Observations	2,150	2,150	2,150
Nr of countries	72	72	72

Notes: ln(commtime) defined as the natural logarithm of communication time (in days) to London. The instrument in Panel B is a binary variable that takes the value of one including and after the predicted year of connection based on the Riley measure interacted with the closest neighbor's drop in communication time after connection to the telegraph.The instrument in Panel C is a binary variable that takes the value of one including and after the predicted year of connection based on the normalized cable length measure interacted with the closest neighbor's drop in communication time after connection to the telegraph. Year and country FEs included. Standard errors clustered by country in parentheses. Notation for statistical significance; *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Alternative regressors: Poisson

VARIABLES	(1) Yarn	(2) Plain cloth	(3) Finished cloth
Panel A.			
ln(comm time)	-0.183***	-0.097***	0.006
	(0.041)	(0.029)	(0.044)
Observations	2,150	2,150	2,150
Nr of countries	72	72	72
Panel B.			
Telegraph dummy	0.385**	0.092	-0.091
	(0.167)	(0.112)	(0.098)
Observations	2,540	2,540	2,540
Nr of countries	72	72	72
Panel C.			
(Tele dummy)*	0.190**	0.070*	-0.004
(change comm time)	(0.078)	(0.037)	(0.051)
Observations	2,540	2,540	2,540
Nr of countries	72	72	72

Notes: Regressors are ln(commtime) defined as the natural logarithm of communication time (in days) to London (Panel A), Telegraph dummy defined as a binary variable that takes the value of one in years including and after a telegraph connection is achieved (Panel B), (Tele dummy)*(change comm time) defined as a binary variable that takes the value of one in years including and after a telegraph connection is achieved interacted with the drop in communication time after connection to the telegraph. Year and country FEs included. Standard errors clustered by country in parentheses. Notation for statistical significance; *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Robustness

VARIABLES	(1) Yarn	(2) Plain cloth	(3) Finished cloth
Panel A. Baseline IV (Riley measure)			
ln(comm time)	-0.272***	-0.128**	-0.030
	(0.076)	(0.065)	(0.066)
Observations	2,150	2,150	2,150
Nr of countries	72	72	72
Panel B. Drop Civil War			
ln(comm time)	-0.265***	-0.128*	-0.036
	(0.076)	(0.069)	(0.080)
Observations	1,845	1,845	1,845
Nr of countries	72	72	72
Panel C. Drop British colonies			
ln(comm time)	-0.401***	-0.212***	0.067
	(0.122)	(0.073)	(0.107)
Observations	1,632	1,632	1,632
Nr of countries	55	55	55
Panel D. Control for tariffs			
ln(comm time)	-0.282***	-0.125**	-0.047
	(0.069)	(0.055)	(0.062)
tariff rate	-0.430	-0.041	-0.020***
	(0.276)	(0.293)	(0.001)
Observations	1,096	1,096	1,096
Nr of countries	36	36	36
Panel E. Control for GDP			
ln(comm time)	-0.428**	-0.076	0.153
	(0.195)	(0.093)	(0.094)
ln(GDP)	0.082	-0.680***	-0.757***
	(0.417)	(0.173)	(0.258)
Observations	920	920	920
Nr of countries	38	38	38
Panel F. Import values			
ln(comm time)	-0.158	-0.076	0.005
	(0.104)	(0.067)	(0.066)
Observations	2,150	2,150	2,150
Nr of countries	72	72	72

Notes: ln(commtime) defined as the natural logarithm of communication time (in days) to London. The instrument used across all specifications is a binary variable that takes the value of one including and after the predicted year of connection based on the Riley measure interacted with the closest neighbor's drop in communication time after connection to the telegraph. Controls: ad-valorem product specific tariffs from Tena-Junguito et al. (2012), current price annual (log) GDP values from Bolt et al. (2018) and Hugot and Dajud (2016). Appendix A.2 contains a detailed discussion of the construction of each variable. Year and country FEs included. Standard errors clustered by country in parentheses. Notation for statistical significance: *** p<0.01, ** p<0.05, * p<0.1.

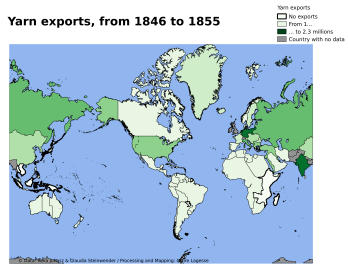
Table 9: Mechanism

VARIABLES	(1) Yarn	(2) Plain cloth	(3) Finished cloth
Panel A. Interaction with international merchants			
ln(comm time)	-0.335***	-0.230***	-0.094
	(0.095)	(0.059)	(0.065)
ln(comm time)*	-0.001	0.023**	0.020**
number int merchants	(0.020)	(0.010)	(0.009)
Observations	2,019	2,019	2,019
Nr of countries	68	68	68
Panel B. Control for mail shipping time			
ln(comm time)	-0.306***	-0.127*	-0.014
	(0.092)	(0.068)	(0.068)
ln(mail ship time)	0.201	-0.074	-0.212
	(0.148)	(0.123)	(0.142)
Observations	2,150	2,150	2,150
Nr of countries	72	72	72

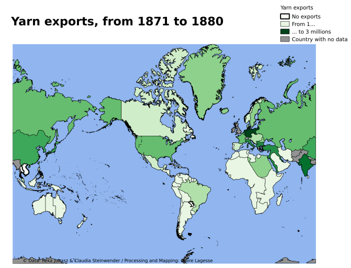
Notes: ln(commtime) defined as the natural logarithm of communication time (in days) to London. The instrument used across all specifications is a binary variable that takes the value of one including and after the predicted year of connection based on the Riley measure interacted with the closest neighbor's drop in communication time after connection to the telegraph. Controls: international merchants defined as the number of British merchant houses that have an affiliate merchant house in the destination market, mail shipping time defined as the natural logarithm of mail shipping time (in days) to London. Appendix A.2 contains a detailed discussion of the construction of each variable. Year and country FEs included. Standard errors clustered by country in parentheses. Notation for statistical significance; *** p<0.01, ** p<0.05, * p<0.1.

B Figures

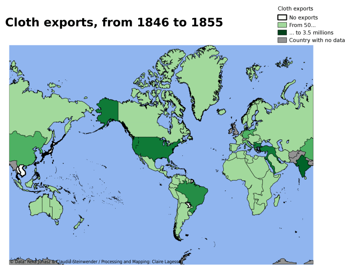
Figure 1: Destination markets for British cotton textiles



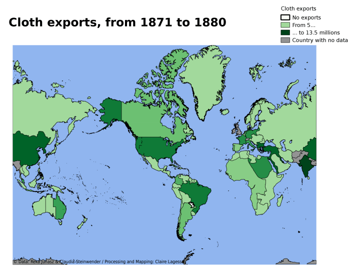
(a) Yarn, 1846-1855



(b) Yarn, 1871-1880

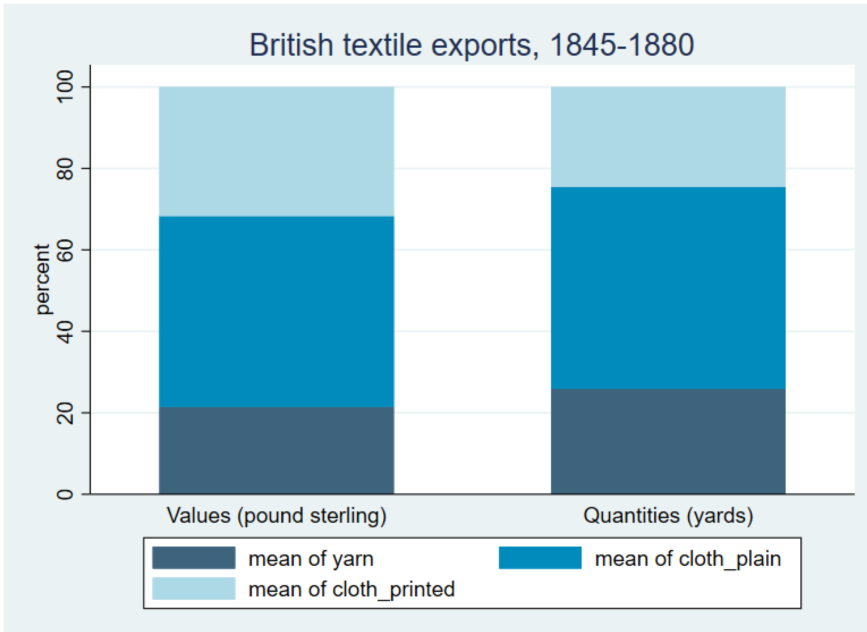


(c) Plain and finished cloth, 1846-1855



(d) Plain and finished cloth, 1871-1880

Figure 2: British cotton textile exports, by product category, (average 1845-1880)



Notes: For quantities chart, 1 pound of yarn yields on average 5 yards of cloth (Ellison, 1886)

Figure 3: Excerpts from *"Liddel's The 'Economic' Telegram Code for Piece Goods and General Business"*

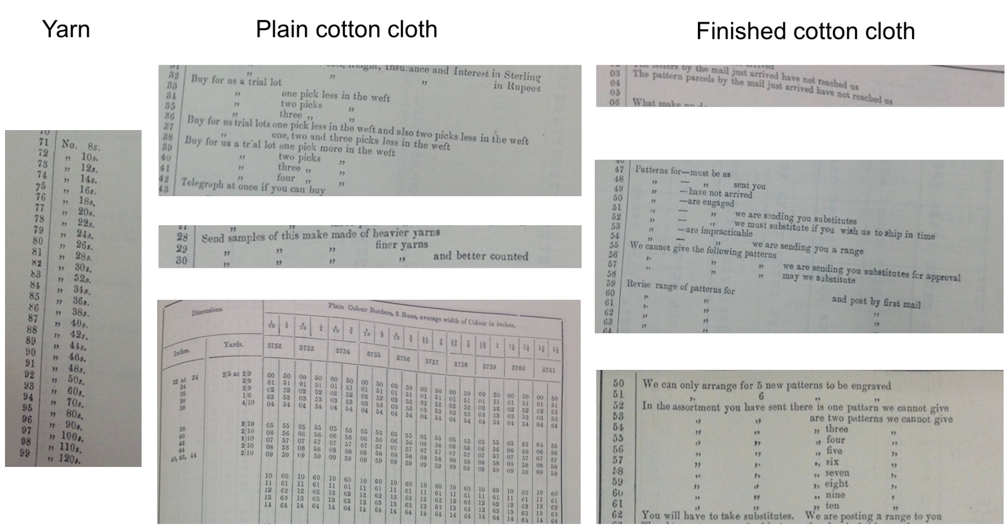
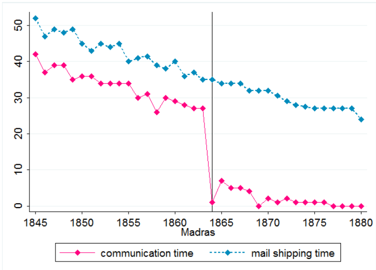
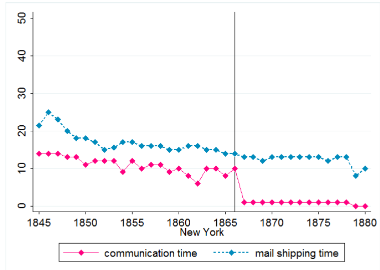


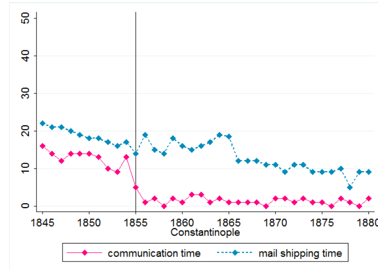
Figure 4: Communication and mail shipping times relative to London (in days)



(a) Madras



(b) New York



(c) Constantinople

Figure 5: Global communication times in days to London by decade, 1850-1880

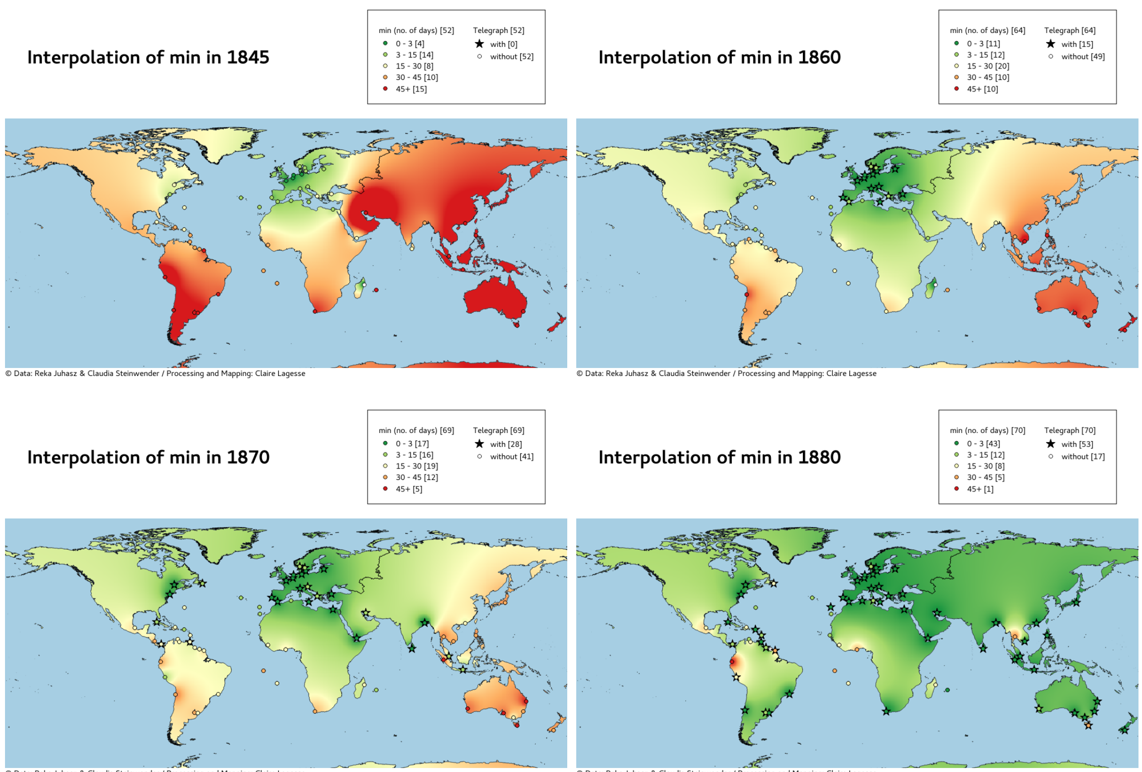


Figure 6: Year of first successful connection to London via telegraph, 1850-1899

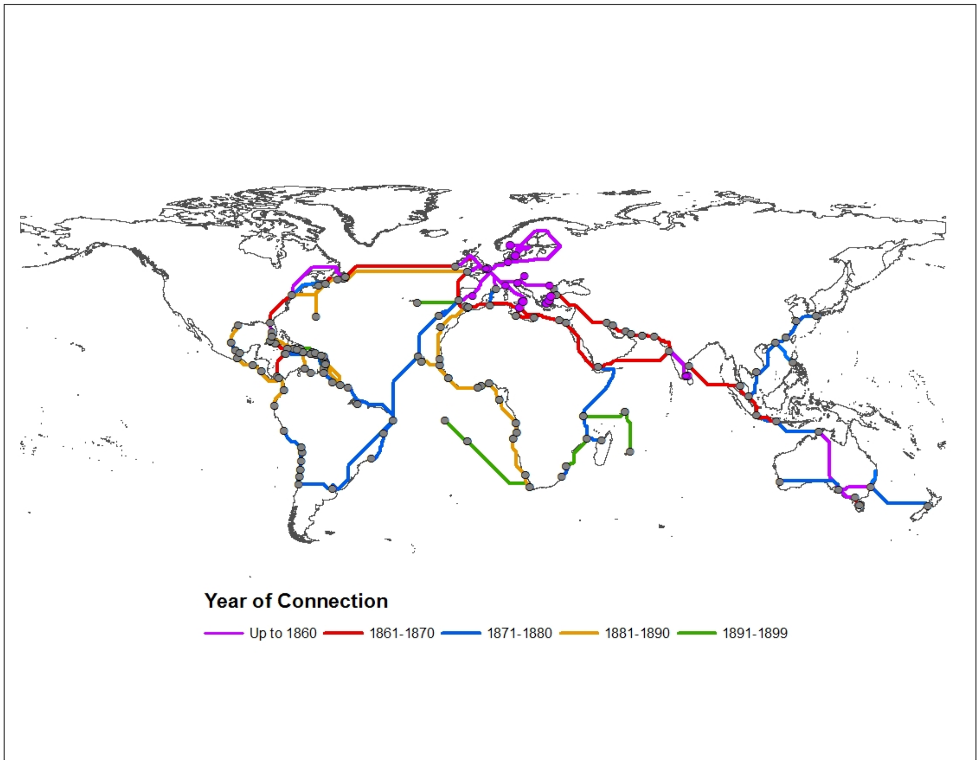


Figure 7: Ruggedness of "telegraphic plateau" (transatlantic telegraph)

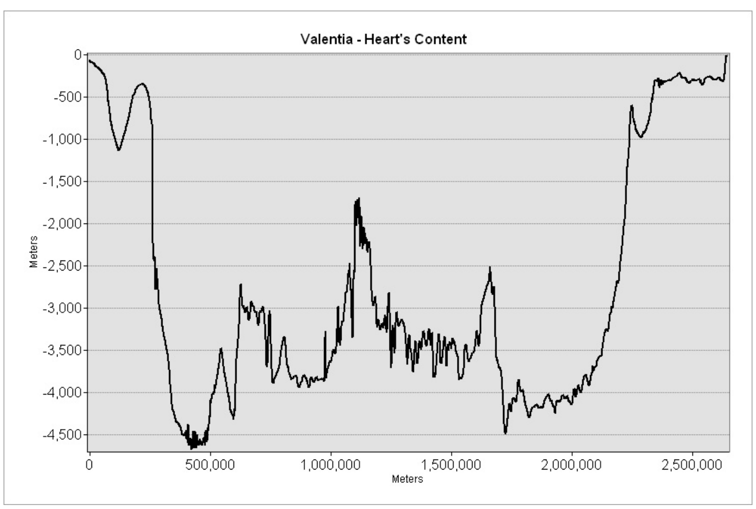
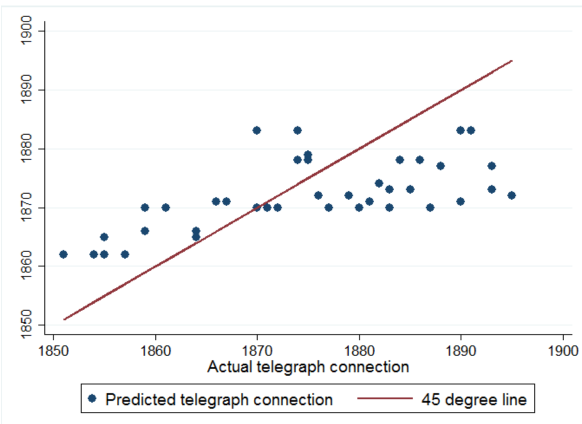
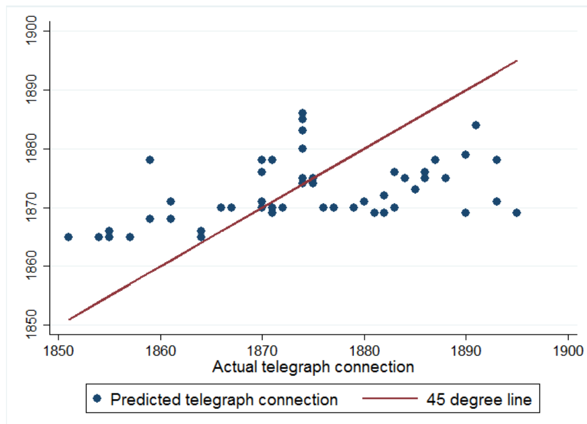


Figure 8: Predicted versus actual year of first successful telegraph connection



(a) Riley ruggedness



(b) Normalized Cable Length