NBER WORKING PAPER SERIES

WOMEN, RAILS AND TELEGRAPHS:
AN EMPIRICAL STUDY OF INFORMATION DIFFUSION AND COLLECTIVE ACTION

Camilo García-Jimeno
Angel Iglesias
Pinar Yildirim

Working Paper 24495
http://www.nber.org/papers/w24495

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2018

We acknowledge the financial support of the Dean's Research Fund at the Wharton School, and the Hal Varian Visiting Assistant Professor research fund at MIT. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Women, Rails and Telegraphs: An Empirical Study of Information Diffusion and Collective Action
Camilo García-Jimeno, Angel Iglesias, and Pinar Yildirim
NBER Working Paper No. 24495
April 2018
JEL No. D71,D83,N11,N31,N71,O18,Z12

ABSTRACT

How do social interactions shape collective action, and how are they mediated by the availability of networked information technologies? To answer these questions, we study the Temperance Crusade, one of the earliest instances of organized political mobilization by women in the U.S. This wave of protest activity against liquor dealers spread between the winter of 1873 and the summer of 1874, covering more than 800 towns in 29 states. We first provide causal evidence of social interactions driving the diffusion of the protest wave, and estimate the roles played by information traveling along railroad and telegraph networks. We do this by relying on exogenous variation in the rail network links generated by railroad worker strikes and railroad accidents. We also develop an event-study methodology to estimate the complementarity between rail and telegraph networks in driving the spread of the Crusade. We find that railroad and telegraph-mediated information about neighboring protest activity were main drivers of the diffusion of the protest movement. We also find strong complementarities between both networks. Using variation in the types of protest activities of neighboring towns and in the aggregate patterns of the diffusion process, we also find suggestive evidence of social learning as a key mechanism behind the effect of information on protest adoption.

Camilo García-Jimeno
Department of Economics
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19104
and NBER
gcamilo@sas.upenn.edu

Angel Iglesias
University of Pennsylvania
3718 Locust Walk
McNeil Building
Philadelphia, PA 19147
angelig@sas.upenn.edu

Pinar Yildirim
Marketing Department
The Wharton School
University of Pennsylvania
Philadelphia, PA 19104
pyild@wharton.upenn.edu

1 Introduction

The political power of disenfranchised groups relies on their ability to organize and exercise collective action. Effective collective action requires coordination, and effective coordination requires information. But how does information impact the ability of groups to solve the collective action problem? Two features of the informational environment appear to be key: the types of information technologies available, and how these groups use the information they receive.

Internet-based social media platforms, for example, played a key role in fostering the Arab Spring (Acemoglu et al. (2014); Hassanpour (2014); Tufekci and Wilson (2012)). Similarly, television and radio were key instruments for the organization of the Civil Rights Movement in the U.S (Andrews and Biggs (2006); Morris (1984)). Governments that seek to prevent or undermine collective action are well aware of the threat posed by readily available information technologies (Dagaev et al. (2013)). As attested by Trotsky in his *History of the Russian Revolution*, one of the priorities of the Revolutionary government was to control all forms of communication technologies:

> The soviet seized all the post and telegraph bureaus, the wireless, all the Petrograd railroad stations, all the printing establishments, so that without its permission it was impossible to send a telegram, to leave Petrograd, or to print an appeal. (Trotsky, 1932, p. 120)

Communication technologies matter not only for the quantity, but also for the nature of the information they carry. For example, internet-based news may become ineffective if fake news become prevalent. Closer to this paper, and reminiscent of Trotsky's account, telegraphs during the second half of the 19th Century transmitted words at the speed of light, while trains moved newspapers and people, albeit at a much slower pace.

Information flows can have different effects depending on how a group aggregates and uses them. In threshold models of social interactions, for example, people care about how many others have already adopted a behavior when deciding whether to adopt it too (Granovetter (1978)). In contrast, in models of social learning, people care about others' behavior only insomuch as it leads them to conclude it may be profitable to change their behavior as well.

In this paper, we contribute to the literatures on social interactions and collective action in several ways. We study social interactions by estimating how women's decisions to participate in the Temperance Crusade depended upon the collective action decisions of women in neighboring towns. We do this by tracing how information about protest activity was mediated through the railroad and telegraph networks, the two key information transmission technologies in the second half of the 19th Century. The Temperance Crusade is an ideal historical setting to investigate these issues for a variety of reasons. Despite the very different technological characteristics of rails and telegraphs, their geographic distribution appears to be closely correlated with the spread of the Crusade. Although it was the first movement of organized collective action by women in the U.S. (with the Seneca Falls Convention of 1848 as its only antecedent of significance), the protests in each town were quite parochial in their aims. This had two implications. First, besides a few leaders involved in spreading the word, protest decisions at the local level were driven only by local objectives, and not by strategic interests such as affecting the subsequent diffusion of the movement. As a result, we can safely abstract from issues of strategic learning or strategic information transmission in our empirical and theoretical approaches. Second, the protests' local nature implied that resistance to them was also purely local. Moreover, in 1874 women in the U.S. were still disenfranchised in all states except for Wyoming and Utah. Women's ability to exercise collective action was their only direct source of political power, allowing us to abstract away from alternative channels of political influence as potential omitted variables. The historical setting also restricts the number of potential communication technologies we must consider, justifying our emphasis on railroad and telegraph networks. Finally, we have access to detailed, daily variation in the occurrence and type of Crusade-related events, and daily variation in railroad worker strikes and railroad accidents. Crucially for our empirical strategy, the time-series variation in these disruptions to the information transmission infrastructure is likely unrelated to other shocks driving women's collective action decisions. Moreover, our detailed knowledge of the rollout of the Crusade in each town allows us to provide some evidence distinguishing between pure contagion and social learning as drivers of the protest diffusion process.

A vast literature in Economics has studied social interactions in the adoption of behaviors and activities at the individual level. A similarly large body of work in Political Science, Sociology, and Economics has examined the determinants of collective action. Our main contribution is to bring these areas of inquiry together by providing causal evidence of social interactions in a collective action setting, and tracing the roles that alternative communication technologies play in allowing groups to aggregate and use information. We do this using a variety of methodological approaches. Estimating whether Crusade-related protests in a community had a causal effect on the subsequent crusading decisions of neighboring communities presents an array of empirical challenges first emphasized by Manski (1993). The potential for unobserved correlated effects is particularly serious in our context because towns are embedded in several communication networks.[1] We tackle this challenge with a panel instrumental variables strategy relying on exogenous variation in network links caused by railroad worker strikes and railroad accidents during the months of the Crusade. To the best of our knowledge, this is the first observational study to exploit exogenous variation in network links to identify social interaction effects.[2] The only other comparable exercise of which we are aware of is the work by Koudjis (2016), who uses weather disruptions in the English Channel to identify the effect of information on stock prices in 18th Century Amsterdam. His strategy, however, abstracts away from any network considerations.

Our instrumental variables strategy allows us to separately identify the effects of information

[1] In fact, in a setting with simultaneous networks in place, the information flows traveling along a network will appear as correlated effects from the point of view of other networks.

[2] Discussing the challenges of identification in social network settings, Breza (Forthcoming) argues: "Because the social network encodes patterns of interactions of individuals in real life, it is often extremely hard, if not impossible to find sources of exogenous variation in network structure... The possibility of using exogenous network change to better understand causal links between network shape and other real outcomes is exciting. However, such an exercise would require that the underlying change to the network not be directly correlated with the outcome of interest..." (p. 22).

transmission along railroad and telegraph networks on crusading activity. It also allows us to establish the central role played by railroad and telegraph-mediated information flows in the diffusion of the Crusade. To further estimate the complementarity between these communication networks, we rely on an event-study methodology that exploits cross-sectional variation in access to railroad and telegraph networks across towns. Studying short time windows after women in a town have undertaken a protest, we compare the relative likelihood of subsequent Temperance Crusade events between neighboring towns that vary in their rail and telegraph access within narrow spatial clusters. The time-series variation in protest activity allows us to control for spatially correlated and town-specific unobservables, making the comparison of towns with varying types of network access quite reasonable.

We find positive and precisely estimated average social interaction effects mediated through the railroad network. During the phase of fastest spread of the Crusade, one additional Crusade event among neighboring towns linked by rail led to a six-fold increase in the probability of holding a Crusade event in the following 10 days. Our estimates for the average effect of information transmission through the telegraph network are larger but less precisely estimated. We show that this is partly explained by a complementarity between railroad and telegraph networks: in the absence of rail connections, telegraph connections were not an effective means of protest diffusion, though they boosted the responsiveness of neighboring towns when railroad links were present. Conditional on a neighboring town experiencing a Crusade event, the average probability of holding a Crusade event in the following 2 weeks was 10 percentage points larger for towns with both a rail link and telegraph access compared to towns with only one of the technologies. These results are very precisely estimated, and highlight the importance of network complementarities in social learning settings. Moreover, as would be expected in a network setting of information transmission, we find a clear pattern of decay in the effectiveness of signals over increasing distances. Consistent with the qualitatively different nature of both communication technologies, our findings indicate that information transmitted through the railroad network had a delayed effect on neighboring towns' collective action relative to information transmitted through the telegraph.

We also provide an array of robustness exercises and tests of the validity of our empirical strategy, including specification tests and placebo exercises. Our results are very similar when we vary the way in which we construct our railroad link instruments, when we vary the number of days within periods in our panel, when we vary the lag structure of our models, and when we vary the definition of a link in the railroad network. Additionally, we show the first stage relationship between railroad accidents and strikes and neighboring protest activity is very strong. Results are similar (and more precise) when we use GMM instead of IV.

Our results contribute to the literature on social interactions in adoption settings. Most of this literature has focused on *individual* social learning. Its origins are commonly traced to Bikhchandani et al. (1992) and Banerjee (1992), who model social learning as a sequential process. These models allow for informational cascades and inefficient herding despite Bayesian behavior.[3] More recent theoretical contributions on social learning focus on agents interacting in networks, establishing the relationship between network topology and long-run learning under various behavioral assumptions (Acemoglu et al. (2011); Bala and Goyal (1998); Golub and Jackson (2010); Mossel et al. (2015)).[4]

Parallel to these theoretical contributions, there is a vast empirical literature interested in identifying social interactions in the adoption of behaviors, from hybrid corn adoption (Griliches (1957)) to bank panics (Kelly and OGrada (2000)). Economists have been keen on finding evidence of social learning when social interactions are present. At its heart is the challenge of distinguishing it from other forms of social interactions, such as simple contagion or imitation (Young (2009)). Observing both choices and outcomes is necessary to isolate patterns of social learning from other channels. For example, several papers using observational data have studied technology adoption decisions of farmers learning from their neighbors (Bandiera and Rasul (2006); Conley and Udry (2010); Foster and Rosenzweig (1995)).[5] Experimental studies such as Kremer and Miguel (2007) and Duflo and Saez (2003) have also studied social learning in the adoption of deworming medicine uptake or retirement savings decisions.[6] Moretti (2011) uses data on movie attendance to show that over or under-performance of a movie at the box office in its first week leads to even larger over or under-performance in subsequent weeks. In a similar vein, Knight and Schiff (2010) estimate a model of learning in primaries, and find an important role played by momentum in electoral success: candidates benefit in late-voting states from performing unexpectedly well in early states. These effects are consistent with Bayesian updating by consumers and voters.

Our paper also attempts to distinguish between social learning and other sources of correlation in behavior such as contagion. In contrast to the previous literature, however, our emphasis is on identifying *collective* social learning, where a group of people aggregates and uses information generated by other groups. We do not observe the outcomes of protest activity at the local level –for example, whether the women managed to close the saloons in their towns–. We do observe, however, different types of collective action events. Under social learning, different kinds of events should generate signals with different informational content. Observing heterogeneity in social interaction

[3]For a recent generalization delineating the conditions required for learning to obtain, see Smith and Sorenson (2000). Another noteworthy strand of the literature on individual social learning models sequential learning under bounded rationality, where players learn based on rules of thumb or where their ability to interpret information is limited (See Ellison and Fudenberg (1993, 1995); Guarino and Jehiel (2013)).

[4]These models extend the popular DeGroot-type models where agents are embedded in networks but aggregate information in simple, non-Bayesian ways (DeGroot (1974)). For a recent empirical experimental application using a DeGroot-type model to distinguish social learning from endorsement, see Banerjee et al. (2013). For a field experiment showing agents in a network appear to be more sophisticated than the standard DeGroot model assumes, see Mobius et al. (2015).

[5]Foster and Rosenzweig (1995) rely on their ability to observe both the adoption decisions of Indian farmers (choices) and their profits (outcomes), to provide evidence that the changes in adoption over time are driven by learning about the profitability of the new crop variety and not simply by a preference for imitation of neighbors.

[6]Other papers providing evidence of social learning in a variety of contexts include: Munshi and Myaux (2006), who look at the fertility transition in Bangladesh; Burke et al. (2007), who analyze physicians' stent adoption decisions; Henkel and Maurer (2010), who document R&D patterns in stem cell research that are consistent with social learning; Oster and Thornton (2012), who find evidence of peer effects in menstrual cup use in Nepal; Dupas (2014), who shows that a subsidy for bed nets increases the adoption rate of neighbors in Kenya.

effects across types of events can provide evidence of a learning mechanism. We find this to be the case; both for railroad and telegraph information flows, crusading women were more responsive to meetings in neighboring towns than to petitions or rallies. We show this is not driven by differential media reporting of different types of Crusade events, and interpret it as suggestive of social learning.

Our paper also relates to a growing literature studying the diffusion of behaviors in online social networks, much of it by computer scientists (Aral et al. (2009); Aral and Walker (2012); Bakshy et al. (2012); Gruhl et al. (2004); Lerman and Ghosh (2010)). These papers document contagion in a variety of online activities such as news consumption or app adoption relying on observational and experimental data. This is made possible by their ability to both map the social networks and trace the information flows in detail. In a different and historical setting, we are able to undertake a similar exercise by mapping the railroad and telegraph networks, and by observing each instance of information generation. Conveniently, the relatively short time span of the Temperance Crusade allows us to take these communication networks as fixed and abstract from endogenous network formation considerations. This is a major empirical difficulty in online social network studies because correlations in behavior across agents can be driven by selection into friendships.

Our paper is also directly related to the literature on collective action and political mobilization. Beginning with Olson (1965), most early work by political scientists emphasized characteristics such as group size and group heterogeneity as important determinants of successful collective action. In his classic study on collective action, Tilly (1978) highlights four dimensions of the problem: interests, organization, mobilization, and opportunity. Beginning with Granovetter (1973, 1978), sociologists in turn have emphasized the importance of group identity, social ties, and preferences for conformity in galvanizing collective action. These ideas have been applied to various settings such as worker strikes, the diffusion of trade unions, and political protest and unrest (Biggs (2003); Gould (1991); Hedstrom (1994); Opp and Gern (1993)). Along similar lines, the recent theoretical work by Passarelli and Tabellini (2017) suggests that emotions may drive protest activity when citizens organize in response to perceptions of unfair policies.

Economists also are increasingly interested in understanding collective action. González (2016) studies how high school classmate networks were key drivers of a recent protest movement in Chile. Aidt et al. (2017) study the English Swing riots of 1830-31, and emphasize the importance of communication constraints and economic fundamentals as drivers of their diffusion. We are unaware, however, of other studies besides ours focused on the dynamics of collective action in a setting with competing networks, and on the complementary roles of alternative communication networks in driving protest diffusion. Ours is also the only study exploiting exogenous variation in network links. Other recent empirical studies of protest activity and social networks estimate how the spatial roll-out of new information technologies such as internet-based social media platforms and cell phone coverage impact the likelihood of collective action (Christensen and Garfias (2015); Enikolopov et al. (2016); Pierskalla and Hollenbach (2013)).[7] In contrast, we directly trace how information trans-

[7] Also related is the recent literature studying the role of social networks such as Twitter in shaping political

mitted along established communication networks leads to social learning and subsequent collective action. Other recent work studying the nature and consequences of collective action and protest activity from a political economy perspective includes Cantoni et al. (2017), Madestam et al. (2013), and Yanagizawa-Drott (2014).

The rest of the paper proceeds as follows. In Section 2 we provide a historical overview of the Temperance Crusade, and a discussion of the role of railroads and telegraphs as the main communication technologies of the time. In section 3 we introduce and describe the data collected and used in the paper. Section 4 then discusses our empirical strategy to identify the effects of network-mediated information flows on protest diffusion, presents our main estimates, and describes our findings. After having established the importance of network-mediated information flows in for the diffusion of the Crusade, in section 5 we analyze the aggregate dynamics of the protest movement, testing a variety of models of social interactions. Section 6 then turns to the estimation of technological complementarities between railroads and telegraphs, describing our event-study approach and main results. Section 7 concludes. Finally, online Appendices A, B, and C contain additional results and a more detailed description of our sources and data.

2 Historical Overview of the Temperance Crusade

The Temperance Crusade was striking in the speed and scope of its spread. In less than a year, groups of disenfranchised women mobilized and took to the streets in hundreds of towns across 29 states around a single cause: to demand the closure of saloons and liquor dealers. Effectively, it was the largest social movement involving political action in the nineteenth century, counting almost 150,000 women (Blocker (1985)).[8] Other major social movements of this period such as abolition and temperance societies had reached larger enrollments, but few had engaged in active militant action (Bordin (1981); Degler (1981); Tyrrell (1979)). Perhaps more strikingly, and in contrast with other reform movements, the Crusade was truly grass-roots, with no planning or central organization, and taking place two years before Graham Bell invented the telephone, five years before Edison invented the lightbulb, and when much of the West was still frontier territory.

Notwithstanding these limitations, historical accounts of the Crusade agree that communication technologies were key to its diffusion. By the early 1870s, both the railroad and the telegraph networks had expanded considerably across the U.S. The First Transcontinental Railroad, linking California to the eastern states, had already opened, and close to 45,000 miles of track had been laid (Stover (1999)). Trains were by far the main mode of transportation of travelers and freight. Why was the railroad so important for the diffusion of the Crusade? It allowed for the movement of local

communication (Halberstam and Knight (2016)).

[8]The Temperance Crusade also preceded all other Progressive-era female organizations except for the Women's Suffrage Movement begun at Seneca Falls in 1848. The WCTU was founded in 1874 as a direct result of the Crusade, while the General Federation of Women's Clubs was founded in 1890, the National Congress of Mothers in 1897, the Women's Trade Union League in 1903, and the National Birth Control League in 1910 only (Cooney (2005); Schneider and Schneider (1993)).

leaders across towns, and for the flow of newspapers reporting on Crusade activities. Contemporary accounts all agree on the importance of 'visitors, emissaries, missionaries, and delegates' spreading the word, particularly during the early phase of the Crusade. For example, after Dr. Dio Lewis gave the speech on temperance in Fredonia, NY, that led to the first Crusade, he traveled to three other towns in New York and Ohio, giving speeches that had the same effect. According to (Blocker, 1985, pp. 11-12), "...the four actions initiated by Lewis became the forerunners of a national women's movement... Lewis provided the initial impetus for the Crusade, but other agencies produced its growth from a local incident to a national movement."

Possibly more important than the role of leaders was the role of local newspapers. As studied recently by Gentzkow et al. (2011) and noted early on by De Tocqueville, nineteenth-century printed newspapers were widespread and central to the political and civic culture of the U.S.: "... the number of periodicals and occasional publications in the United States exceeds all belief.... scarcely any hamlet lacks its newspaper" (DeTocqueville, 1835, p. 215). Less than a month after the first Crusade in upstate New York, newspapers in Columbus, Cleveland, Detroit, Minneapolis, New York City, Baltimore, and Newark were already reporting on it. Newspaper reports of protest activity were read out loud and shared during the organizational meetings where women discussed whether to undertake protests themselves (Blocker (1985)). As we will discuss in Section 3, our own newspaper search recovered more than four thousand articles on Temperance Crusade activity, many of them reporting on events taking place in distant locations.[9]

Of similar importance for the diffusion of the Crusade was the telegraph network, which by then had reached California as well. To a large extent, this network operated in lines running parallel to the railways. Rails and telegraph cables did not, however, completely overlap, as we illustrate in Table 1. The table presents the joint distribution of rail and telegraph access across all 15,971 towns in the 1870 U.S. Census, the analogous joint distribution for the 802 towns which experienced Crusade activity, and the respective conditional probabilities of collective action. The table makes two points. First, the railroad network had much wider coverage: while two thirds of all towns had rail access, only six percent of towns had telegraph access; second, towns with telegraph were very likely to have rail access as well: 87 percent of towns with a telegraph were also in the rail network.

Despite its much smaller geographic scope, the telegraph was significantly more efficient at information transmission.[10] As a result, it became central to the operation of the newspaper industry. The telegraph industry had been rapidly expanding starting in the 1840s. It also had experienced intense competition. However, after the Civil War, Western Union managed to consolidate an effective monopoly of the telegraph cables, controlling 37,000 miles of routes and 2,250 offices (Swindler

[9]Recent studies have similarly shown the importance of the railroad for economic growth and market integration in the U.S. during the nineteenth century (Donaldson and Hornbeck (2016); Feigenbaum and Rotemberg (2015)).

[10]This became especially true after the invention in the late 1860s of the automatic repeater, which retransmitted incoming telegraph messages onto the next circuit without the intervention of a human operator, and the invention of the duplex cable, which permitted messages to be sent simultaneously over the same wire from opposite ends (Schwarzlose (1990)).

Joint Distribution of Railroad and Telegraph Access across U.S. Towns c. 1870, and Temperance Crusade Activity

		Distribution of Towns				Distribution of Crusade Activity				$\mathbb{P}$(Crusade|R, T)	
		T				T				T	
		0	1			0	1			0	1
R	0	5538	122	R	0	126	3	R	0	0.022	0.024
R	1	9514	797	R	1	609	64	R	1	0.064	0.080

Table 1: Joint Distribution of Rail and Telegraph Access across U.S. Towns, c. 1870 and Temperance Crusade Activity. The left-most table presents the joint distribution of railroad and telegraph access across all 15,971 towns in the 1870 U.S. Census. The central table reports the distribution of the 802 crusading towns across the support of the joint distribution of rail and telegraph access. The right-most table presents the corresponding conditional probabilities of Crusade activity across the support of the joint distribution of rail and telegraph access.

(1946), Thompson (1947)).[11] This gave Western Union a strong bargaining position in relation to the newspaper industry, forcing the Associated Press to enter a collusive agreement with it. While Western Union promised to transmit reports of AP member dailies only, the Associated Press promised to use Western Union lines exclusively.[12]

Besides the importance of the telegraph for the newspaper industry, leaders of the Crusade also relied on direct telegraph communication to coordinate and spread information. Telegrams appear to have been key in generating enthusiasm.[13] In her memories, one of the most prominent Crusade leaders known as Mother Stewart provides a fascinating account of her role in the movement, illustrating the strong complementarities between the telegraph and railroad networks:

> The Crusaders in Bucyrus were having a peculiarly hard time with the liquor men and their allies, which were not only the low drunkards, but the city Mayor and his officials also. So they wrote me to come to them for the evening of the 21st, but my friends at home insisted that I must go with them to Cincinnati. I telegraphed I could not go at that time. Rev. Baltzly telegraphed back: It will be very disastrous to us if you do not come now. It was now twelve o'clock, and the train left at one. I ran to Rev. Mr. Hamma for advice. He said go, and rising from the dinner-table, bade me sit down and eat my dinner, while he ran to the telegraph office to notify them that I was coming... I sprang in and was driven a half mile to my home... in time for my train. The sisters still insisted that I must return in time to accompany them to Cincinnati next morning. (Stewart, 1890, p. 316)

While the communications infrastructure may have been necessary for the spread of the protest

[11] The postal system, which played a key role in the distribution of newspapers, increasingly relied on the rail and telegraph networks as these expanded. In fact, the Pony Express was ended in 1861 when the Transcontinental telegraph line was completed, and beginning in 1864 railway post offices started to open, effectively linking the postal system to the railroad network (USPS (2012)). For this reason we do not consider the Postal Service as an independent communication network in our analysis.

[12] Schwarzlose (1990) estimates that in 1872, 8 percent of all Western Union revenue came directly from transmitting press reports.

[13] A local newspaper from Ottawa, IL, for example, reported: "The success at Washington Court House encouraged like efforts in other quarters, and now the prayer enthusiasm has broke out all over the state. The Ohio papers are full of it..." (*The Ottawa Free Trader*, Feb. 7, 1874)

movement, in the absence of local grievances women in crusading towns would have lacked the motivation to engage in the costly and risky collective action that meetings, petitions, and marches entailed.

As part of the broader temperance movement, Temperance crusaders –mostly affiliated to Protestant churches– were religiously motivated. As precursors of the Progressive-era reform movement, many also believed that state and community should be involved in promoting moral and social values (Gusfield (1955)). And although there was debate around the issue among crusaders, many supported the women's suffrage movement (Blocker (1989)). Historians, however, disagree on their motivations. Epstein (1981) argues that most crusaders were middle-class women reacting against working-class immigrants and their increasing social influence. For Blocker (1985), in contrast, crusaders' main motivations were the private costs they faced from their male relatives' drinking. The rapid growth of the liquor industry in the decade prior to the Crusade is consistent with this view. Between 1864 and 1873, the number of liquor dealers registered as federal taxpayers grew from 80 to 200 thousand, a 17 percent annual growth rate well above the 2.6 percent annual population growth rate of the decade. The geographic distribution of the Crusade also appears to agree with this view. Both the number of liquor dealers and the levels of alcohol consumption were highest in the Midwest, where more than 75 percent of the protest activity took place. The alcohol markets were smaller in New England and in the South, where liquor restrictions were more common and had been enforced more strongly (Cherrington (1920)). The changing political economy of the Midwest may also have motivated women in that region. Starting in the 1860s, politicians became less willing to enforce Prohibition measures. Several states adopted civil-damage legislation – such as the Adair law in Ohio– allowing victims to sue alcohol dealers for damages. These statutes were intended as substitutes for prohibitory measures and may have, therefore, increased drinking.

Crusaders faced strong local opposition –especially in reaction to their rallies–, from liquor dealers, consumers, and in some cases from preachers and local authorities. Historians also believe resistance from public officials was strongest in larger cities. The opposition turned violent in some instances. Some women were blasted with force-pumped water; others were pelted with rotten eggs and even bricks. (Blocker (1985); Bolton (1944)). In a few cases, saloon owners requested court injunctions against crusaders picketing their businesses.[14]

Saloon visits and sit-ins, referred to as marches, were the most radical form of collective action. They were not, however, the only type of activity. Crusaders also held organizational meetings, often in churches, and sometimes addressed by traveling Crusade leaders. Meetings were well documented

[14] A daily from Rutland, VT, reported on events taking place in Hillsboro, OH: "In Hillsboro, the women who have been laboring for many days with a Mr. Dunn, a druggist, who refuses to accede to their demands, have been astonished at seeing the object of their persistent attentions assume the aggressive and invoke the majesty of the law to sustain him in his defiant attitude. Mr. Dunn has entered suit against the ladies who have been engaged in the Crusade against him, claiming 10,000 damages for trespass and defamation of character. He has also procured from Judge Safford an injunction restraining them from further interference with his business... A correspondent of the Cincinnati Enquirer had an interview with Judge Safford, and asked him what he should do if any of the citizens disobeyed the injunction... 'Do, Sir? Why, I shall have them arrested, and will put them in jail on bread and water for contempt', said the Judge." (*The Rutland Daily Globe*, Feb. 11, 1874).

in the press, and had the purpose of motivating participants, sharing information, and coordinating further action. There is variation across towns in whether meetings took place before militant action was undertaken. There is also variation across towns in whether meetings led to subsequent petitions or rallies; in some towns, the Crusade stopped at the meeting stage. Where further action did happen, there is cross-town variation in the time it took the crusaders to move from a meeting to a formal petition or a march. Similarly, we observe variation in whether and how long it took crusading women to move from petitions to marches. Exploiting this information will allow us to partially distinguish between social learning and other mechanisms driving social interactions in the diffusion of the Crusade. If towns were learning from each other, different crusading paths may have led to different inferences about the expected costs and benefits of collective action. Blocker (1985), for example, suggests that towns experiencing petitions but not subsequently holding marches were mostly places where opposition was strong, and the women concluded marching would have been very risky. Similarly, he argues that isolation from communication networks explains why many towns holding meetings did not move onto rallies.

Despite the rapid spread of the Crusade and its significant geographic reach, its effectiveness in permanently closing saloons was short-lived. The protests themselves may even have generated a backlash from men at the ballot box. In Ohio, the state with the most Crusade activity, the Democratic party –by then the anti-temperance party–, made large gains in the 1874 elections. The Crusade's most direct consequence was the creation of the Women's Christian Temperance Union at a Convention in Cleveland in 1874. The WCTU would become a key player in the movements leading to both constitutional Prohibition (García-Jimeno (2016)) and women's suffrage forty years later (Gusfield (1955)).

3 Data Description

In this section, we describe our data collection effort and our main sources of information, and provide summary statistics describing the evolution of the Crusade.

Temperance Crusade Activity

Jack Blocker's research, described in his book *Give to the Winds Thy Fears: The Women's Temperance Crusade, 1873-1874* (Blocker, 1985), is our source for Temperance Crusade activity. Using his files, we recovered the name of every town where an event related to the Crusade took place, as well as the nature of these events, classified as *meetings*, *petitions*, or *marches*.[15] Based on Blocker (1985) and on contemporary newspaper sources, meetings were town hall gatherings, often held in

[15] In personal communication with us, Professor Blocker claimed his archive includes the complete record of all towns experiencing Crusade-related events. Besides his own newspaper and archival research, he used the record of Crusades compiled by Annie Wittenmyer, the first President of the Womans National Christian Temperance Union, and by historian Susan Dye Lee (Lee (1980); Wittenmyer (1878)). Our own newspaper search described below did not find any crusading towns not already in his archive.

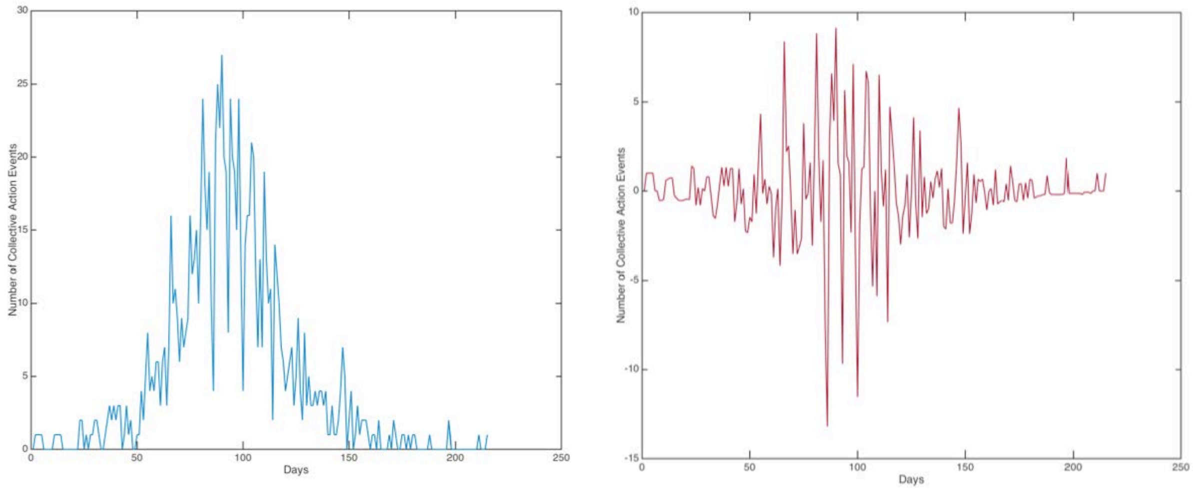

Figure 1: Number of Temperance Crusade Events, Dec. 1873 - July 1874. Evolution of the Temperance Crusade protest activity. The figure reports the total number of events per day, including meetings, petitions, and marches based on the data from Blocker (1985). The left-hand side picture reports the raw count of events. The right-hand side figure reports deviations from a 14-day moving average.

churches, where attendees discussed potential Crusade activity. Petitions were written requests to either the local authorities or the saloon owners directly, demanding closure of the stores.[16] Most strikingly, marches were public demonstrations, often organized in front and inside targeted saloons, involving prayer, singing, and sit-ins. These were effectively wars of attrition between crusaders and store owners, extending over several days in many documented cases.

We observe the type and beginning date of each event. Moreover, we observe that meetings never took place *after* petitions or marches, and petitions never occurred after marches. Petitions and marches, however, were not always preceded by meetings. There is significant variation in the observed histories across towns: some experienced meetings only, some experienced marches only, some experienced meetings, petitions, and marches. The first recorded Temperance Crusade event took place on December 14, 1873, and corresponds to the meeting in Fredonia, NY, mentioned in Section 2. The last recorded march in our dataset took place on July 15, 1874. During this 214-day period, 483 towns held a meeting, 264 towns circulated a petition, and 464 towns staged a march. Figure 1 describes the number of events at a daily frequency.[17] The spread was very slow early on, picking up speed only around 50 days after the first event. The number of incidents peaked after a hundred days into the protest wave.

[16]This is an example of a petition written by the women of Fredonia, NY, in the form of a direct pledge to saloon-owners: *"In the name of God and humanity we make our appeal: Knowing, as we do, that the sale of liquor is the parent of every misery, prolific in all woes in this life and the next, potent alone in evil, blighting every fair hope, desolating families, the chief incentive to crime, these, mothers, wives and daughters, representing the moral and religious sentiment of our town, to save the loved members of our households from the strong temptation of drink,... do earnestly request that you will pledge yourself to cease the traffic here in those drinks forthwith and forever. We will also add the hope that you will abolish your gambling tables."* (Stewart, 1890, p. 87)

[17]We were unable to match six events from Blocker's archive to any town in the 1870 U.S. Census: Stanislaus, CA; Richmond, NY; Lake, IL; Montgomery, OH; Ross, OH; and Highland, OH.

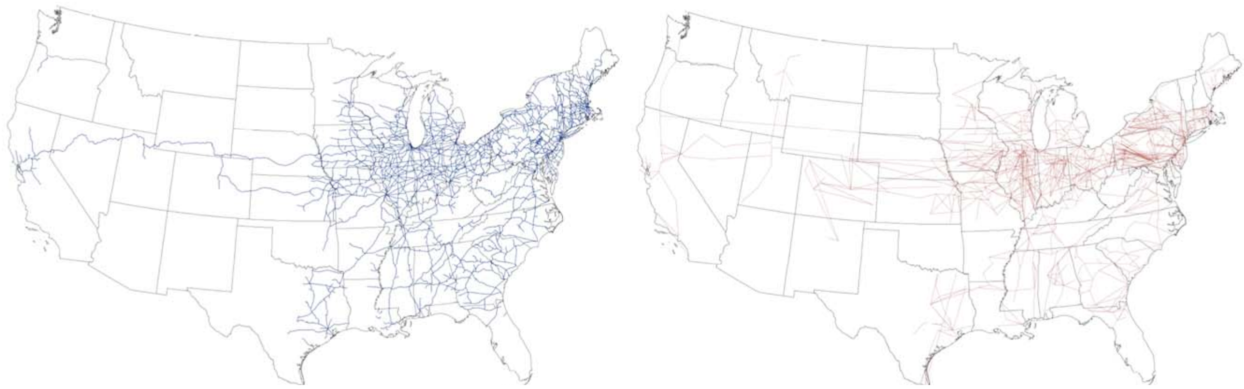

Figure 2: The Railroad and Telegraph Networks, c.1870. The figure on the left depicts the railroad network in 1870, based on Atack (2013). The figure on the right depicts the telegraph network in 1874 based on our own geo-referencing of the maps in WesternUnion (1874).

Railroad and Telegraph Networks

We constructed our railroad and telegraph networks based on the universe of towns in the 1870 U.S. Census, by first geo-referencing each 1870 town using the 2000 Census Tiger/Line shape-files. We matched each town by GPS coordinates, county, and state. We manually verified the cases for which a county had changed its name, or the town became part of a larger settlement.[18] Our final dataset contains 15,971 towns.

Our railroad network data comes from Jeremy Atack's archive at Vanderbilt University[19]. We use his 1870 ArcGIS shape-file, which covers all rail lines in the continental U.S. as of 1870. Using the train routes marked on the maps, we represent the railroad network, denoted by $\mathbf{R}$, as an undirected graph, where railroad lines form the edges, and town centroids from the geo-referenced 1870 U.S. Census are the vertices. The geographical distances between towns within the railroad line serve as weights. We classify a town to be on the railroad network if its centroid is within 10 Kms. of the rail line. Our benchmark rail network classifies a pair of towns as directly linked if they are adjacent along a rail line, and no other towns lie in between.

We next geo-referenced the telegraph network using the 1874 Western Union Telegraph Directory (WesternUnion, 1874).[20] The directory contains maps of U.S. states and territories, depicting the location of telegraph offices and towns with telegraph connections between them.[21] As an illustration, Figure C.1 reproduces the telegraph map for Connecticut and Rhode Island from the directory. Western Union offices are represented by dots, and telegraph cables appear as solid lines.

We geocoded the information in these maps by merging it with the 1870 town-level boundary shape-file. We successfully located 92 percent of telegraph offices from the directory. Using the office

[18]To verify our matching procedure, we used two types of shape-files: sub-county division and place division. For the observations we were unable to match, we manually verified that the town was not part of a larger metropolitan area, or whether the town had changed counties from 1870 to 2000.

[19]The collection can be accessed at `https://my.vanderbilt.edu/jeremyatack/data-downloads/`.

[20]The Western Union directories are accessible through the Hathi Trust Archives at `www.hathitrust.org`.

[21]During the period of study, Western Union controlled more than 90% of the market share of telegraph communications, making our telegraph network almost completely comprehensive.

coordinates, we constructed the telegraph network as an undirected graph, where each town is a vertex and each telegraph line is an edge. In our empirical application, we consider the telegraph network as complete among towns in the network, based on the very nature of the information transmission technology.[22] Throughout we denote this network by $\mathbf{\Gamma}$.

Figure 2 illustrates our resulting railroad and telegraph networks. The 1870 railroad network from Atack (2013) is on the left, and the 1874 telegraph network from our calculations based on WesternUnion (1874) is on the right. As expected, the networks are densest in the most populated regions. At a birds-eye view the two appear highly correlated, but the figures conceal a great deal of variation across towns at the local level. In our sample, 4.8 percent of towns had access to both the railroad and telegraph, 59.7 percent had access to the railroad only, and 0.7 percent had access to the telegraph only. To our knowledge, this paper is the first to collect and use data from the historical telegraph directories.

Railroad Strikes and Accidents

As described in the Introduction, our identification strategy relies on using plausibly exogenous variation in network connectivity across towns over time, induced by railroad worker strikes and train accidents. We obtained data on railroad strikes from the historical *Railroad Gazette* (Wright and Forney, 1873-4), a weekly newspaper publishing railroad news about the whole industry.[23] Ideally for our purposes, it also included reports on service disruptions by rail line and cause. We used the annual compilation of the Gazette for the years 1873 and 1874, and manually extracted information about which railway companies were affected by railroad labor strikes, together with the start and end dates of each strike. We complemented these data with information on strikes from Gutman (1961), and then matched the strike data with information from the *Travelers' Official Railway Guide for the United States and Canada*, from 1870 (Vernon, 1870). The guide contains detailed information for all railroad companies, including the towns along each railroad company's line, and the train schedules for each route. We manually matched each railway company's route in the Guide to the railroad company suffering a strike in the Gazette, to identify which segments of the network were affected. We then coded the affected links as broken during the days when the corresponding strike took place.

The *Railroad Gazette* also contains a monthly compilation of all railroad accidents in the U.S., providing details of each accident (explosion, derailment, collision, people involved, its date, and its location). We manually matched the location of the accidents to our universe of towns to determine which of them suffered from reduced railroad access over the period of an accident. The Gazette is silent about the geographic and time extent of the disruption induced by the accidents, forcing us to make some assumptions about which links in the rail network were affected, and for how long. We

[22] That is, we consider all towns with telegraph access as linked to each other in the telegraph network.

[23] The full collection of *Railroad Gazette* volumes is accessible through the Hathi Trust Archives at www.hathitrust.org.

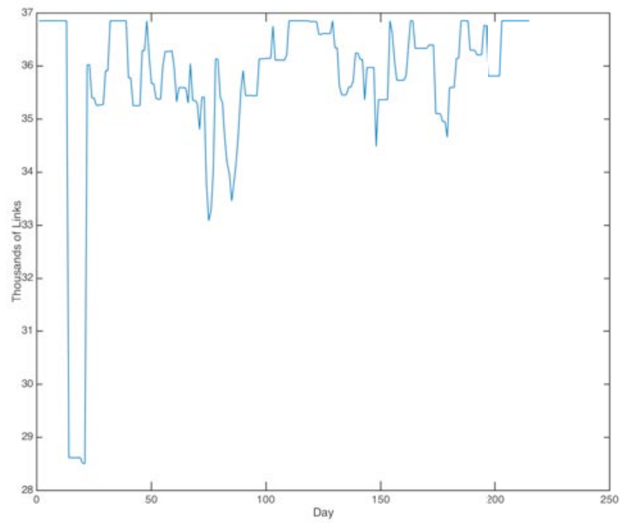

Figure 3: Number of Active Links in the Railroad Network, Dec. 1873 - July 1874. Variation in the number of active links comes from railroad worker strikes and railroad accidents. In the figure, railroad accidents are defined as breaking all links in a 50 kms. radius of the accident's location, for a window of seven days. Accident reports are taken from Vernon (1870). Railroad worker strikes are defined as breaking all links in the line under strike, between the exact beginning and end dates of the respective strike. Strike reports are taken from Gutman (1961) and Wright and Forney (1873-4).

assume throughout that following an accident, the affected edges remained broken for seven days, and compute alternative measures allowing the affected area to include all edges inside either a 50, 80, or 120 kms. radius from the accident location.

Using the disruptions caused by railroad worker strikes and accidents, we compute a time-varying railroad network $\mathbf{R}_t$. Figure 3 plots the "active" number of links in the railroad network at the day level, using the 50 kms. radius definition for the accidents. It illustrates the substantial time series variation in the network structure induced by these disruptions.

Town Characteristics

We also collected information on town characteristics from an array of sources. From the 1870 Census –Volume I (Table III)– we obtain town-level total population, foreign-born population, and black population information. From the University of Minnesota *National Historical Geographic Information System* database[24], we collected additional demographic characteristics at the county level which we then matched to our universe of 1870 towns. These include the female to male ratio and the number of religious sittings (the total seat capacity for each religion in each county). Based on the religious sittings information, we created a Herfindahl index to capture religious heterogeneity within counties.[25]

[24] The dataset can be accessed at www.nhgis.org.

[25] The specification for the Herfindahl index in county j is:

$$HHI_j = \sum_{i=1}^{N_j} \left(\frac{\text{Sittings of Religion i in county j}}{\text{Total Religious Sittings in county j}} \right)^2,$$

We also collected town-level information on the number of local alcohol vendors (saloons, distillers, wine retailers, wine wholesalers, and breweries) from 46 state business directories covering the years 1860-1885. We manually matched each town's vendor data with the 1870 town-level boundary shapefile, based on name and county of each town. Based on the *N.W. Ayer & Son's American Newspaper* directory, we also collected data on the number of newspapers circulating in each town in 1880. Finally, we collected data on the existence of a United States Post Office in each town.[26] The data contain information about each USPO office, geocoded location, and years of operation. During our period of study, 9,130 towns had a post office.

Table 2 presents survivor descriptive statistics for towns that had not yet experienced a Crusade-related event at a given point in time, and towns that had. The table illustrates the patterns of selection on observables induced over time as the protest wave took place. By the end of the Crusade, crusading towns appear to be significantly less religiously heterogeneous (Herfindahl index of 0.23 compared to 0.27 for non-crusading towns, with a t-statistic for a difference of means of 11.3). Crusading towns, disproportionately located in the mid-west, also had much smaller shares of black population (t-statistic for a difference of means of 14.1). The final sample of crusading towns also appears to be much better connected within the railroad network as measured by the betweenness and the degree centrality statistics (t-statistics of -2.5 and -8.3). End-of-Crusade differences for the remaining covariates in the table are not statistically significant, although during the first half of the Crusade, protesting towns had significantly fewer alcohol vendors per capita.

Newspaper Coverage of the Crusade

We also collected data on newspaper coverage of the Crusade. This information comes from the *Chronicling America* online newspaper archive of the Library of Congress.[27] Based on a battery of keywords related to the Temperance movement, we collected a sample of relevant articles in the years around the beginning and end of the Crusade.[28] We used the news article texts to collect data on mentions of specific event types ("meeting", "march", "petition"), by searching for the occurrence of these keywords, as well as the mentions of the towns where the events took place. These searches yielded 4,006 articles in 156 newspaper titles within the relevant time frame.[29] Figure 4 plots the daily number of articles mentioning a crusading event during the relevant period. Comparing it to

where N is the total number of religions in the county.

[26] Miklos Coren from the Central European University created the data we use, merging US post office data from www.postalhistory.com with data from https://geonames.usgs.gov. It can be accessed at https://github.com/ceumicrodata/us-postal-history.

[27] This newspaper repository includes images of publications starting in 1690, and allows for keyword searches of the text in the images themselves (http://chroniclingamerica.loc.gov/). We thank Jesse Shapiro for pointing us to this source.

[28] We used the following keywords: Crusade, Dio Lewis, temperance, war on whisky, whisky war, women protest, women's war, ladies league, women movement, and saloon pledge.

[29] As an illustration, an article in our dataset in a newspaper from Wayne County, IN, reported on Crusade meetings taking place in the towns of Shelbyville, Marietta, Waldron, and Fairland: "These meetings and these lectures have done much toward awakening and strengthening a healthy feeling on this important question of Temperance." (*The Richmond Palladium*, Jan. 3, 1874).

Empirical Survivor Town Characteristics

Days since beginning	Newspapers pc: No Event Yet	Newspapers pc: Some Event	Post Office: No Event Yet	Post Office: Some Event	Alcohol Vendors pc: No Event Yet	Alcohol Vendors pc: Some Event	Religious Herfindhal Index: No Event Yet	Religious Herfindhal Index: Some Event
0	0.27	–	0.57	–	0.40	–	0.27	–
	(1.31)	–	(0.49)	–	(2.92)	–	(0.14)	–
50	0.27	0.26	0.57	0.51	0.40	0.07	0.27	0.21
	(1.31)	(0.75)	(0.49)	(0.51)	(2.93)	(0.44)	(0.14)	(0.07)
100	0.27	0.28	0.57	0.56	0.41	0.24	0.27	0.23
	(1.30)	(1.52)	(0.49)	(0.50)	(2.95)	(1.67)	(0.14)	(0.10)
150	0.26	0.30	0.57	0.58	0.41	0.30	0.27	0.23
	(1.30)	(1.59)	(0.49)	(0.49)	(2.96)	(1.95)	(0.14)	(0.10)
215	0.26	0.31	0.57	0.59	0.40	0.40	0.27	0.23
	(1.30)	(1.57)	(0.49)	(0.49)	(2.95)	(2.26)	(0.14)	(0.10)

Days since beginning	Black Pop. Share: No Event Yet	Black Pop. Share: Some Event	Rail Betweenness Centrality: No Event Yet	Rail Betweenness Centrality: Some Event	Rail Degree Centrality: No Event Yet	Rail Degree Centrality: Some Event	Number of Towns: No Event Yet	Number of Towns: Some Event
0	0.17	–	29	–	2.31	–	15971	0
	(1.05)	–	(106.3)	–	(3.74)	–		
50	0.18	0.04	29	14	2.31	1.71	15936	35
	(1.06)	(0.09)	(106.4)	(26.9)	(3.75)	(1.36)		
100	0.18	0.03	29	33	2.26	3.72	15470	501
	(1.07)	(0.11)	(106.2)	(107.6)	(3.69)	(4.85)		
150	0.18	0.04	28	40	2.24	3.64	15197	774
	(1.08)	(0.15)	(105.6)	(118.1)	(3.68)	(4.68)		
215	0.18	0.04	28	39	2.24	3.63	15169	802
	(1.08)	(0.15)	(105.7)	(116.2)	(3.67)	(4.71)		

Table 2: Survivor Table for Town Characteristics During the Temperance Crusade. The table presents means and standard deviations for a set of town characteristics, after the number of days since the beginning of the Crusade indicated in the leftmost column. For each covariate, the columns to the left report summary statistics for towns not yet having experienced any Crusade event. The columns to the right report summary statistics for towns already having experienced any Crusade event. Newspapers and Alcohol vendors per capita are multiplied by 100. The betweenness centrality statistic is expressed per 1000.

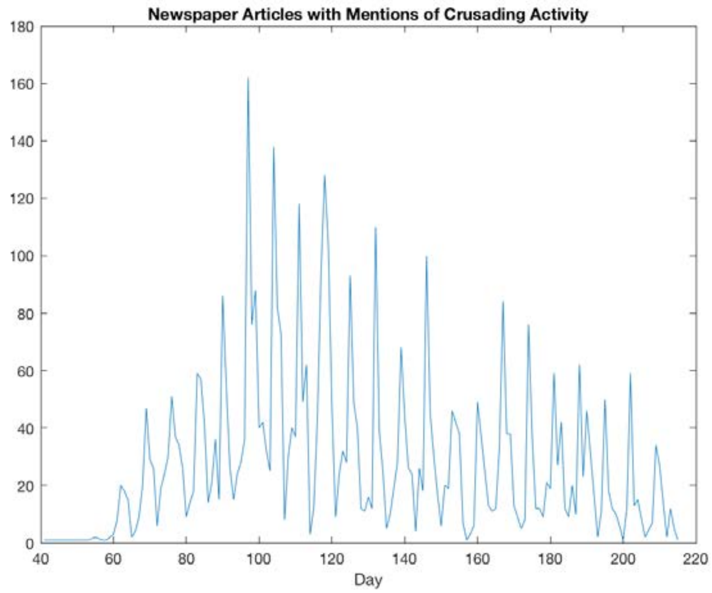

Figure 4: Newspaper Coverage of Temperance Crusade Activity, Jan. 1873 - July 1874. The figure presents the number of articles about events related to Temperance Crusade activity from all newspapers in the *Chronicling America* online newspaper repository of the Library of Congress, based on our text analysis search described in Appendix C, during the period of protest activity.

Figure 1 illustrates the close correlation between protest activity and its newspaper coverage.

We organized these data in a network format, by coding for each town with a local newspaper in our sample, the mentions of other towns' Crusade events reported in its newspaper. This allows us to measure both how often a crusading town was mentioned elsewhere, and how much Crusade-related information a given town received. Appendix C contains additional details about the newspaper article search.

4 Information Technologies and Social Interactions: Empirical Strategy and Results

Our objective in this study is to establish the role played by the main communication technologies of the 1870s, namely railroads and telegraphs, in mediating the information flows related to the Crusade, and leading to its geographic diffusion. To do so, we employ several complementary empirical strategies. In this section, we rely on a linear model of social interactions, and the exogenous time-series variation in network links induced by railroad worker strikes and accidents, to estimate the average effect of information about neighboring crusade events on the likelihood of crusade activity. We find large and precise effects from railroad information flows, and even larger albeit less precisely estimated effects from telegraph information flows. We also find that meetings induced the largest responses in neighboring towns. In section 5 we study the aggregate patterns of the Crusade movement to distinguish between alternative mechanisms underlying the information-mediated social interactions we find in section 4. Our findings here are suggestive of social learning. In section 6 we then use an event study methodology to estimate the complementarities between railroad and telegraph access in fostering the spread of the movement.

4.1 The Impact of Rail and Telegraph-mediated Information Flows

Consider a set of towns $i = 1, 2, ..., n$ embedded in several communication networks. At time t each town i is connected by rail to a set $R_t(i)$ of other towns. For towns without railroad access, $R_t(i)$ is empty. Because railroad worker strikes and accidents disrupt the network, the set of connected towns changes over time. Similarly, each town is connected by the telegraph to a set $\Gamma(i)$ of other towns. Throughout we will assume that the telegraph network is complete among towns with telegraph access: $\Gamma(i) = \Gamma$ for all towns with telegraph access, and $\Gamma(i) = \varnothing$ for all towns without it. We will, however, allow for the strength of a link between two towns in the telegraph network to depend on distance. Similarly, the strength of railroad links depends on distance. We denote by $\mathbf{R}_t$ and $\boldsymbol{\Gamma}$ the rail and telegraph network matrices. Both of these are symmetric matrices and their diagonals are zeroes. $r_{ij,t} \in [0, 1]$ is a typical element of $\mathbf{R}_t$, and $\gamma_{ij} \in [0, 1]$ is a typical element of $\boldsymbol{\Gamma}$. Finally, denote by $\mathbf{r}_{i,t}$ the ith row of the rail network matrix, and by $\boldsymbol{\gamma}_i$ the ith row of the telegraph network matrix. Information may travel through alternative means, such as roads and waterways. These constitute latent networks, through which the same information may flow. We capture these latent networks using the distance matrix of all U.S. towns, and call it $\mathbf{D}$, with typical entry d_{ij}. We denote by $\mathbf{d}_i$ the ith row of the distance network.

Information may travel at different speeds in different networks because of their distinct nature. As a result, the same signal may have effects at different frequencies. The lag structure of the effectiveness of information flows in inducing protest activity elsewhere may differ across different networks. Part of our empirical strategy will entail estimating the relevant lag structure. There is also a potential for interaction effects between networks if, for example, the rail-mediated information is useful for protesters especially when additional telegraph-mediated information arrives. To allow for collective action in some towns to generate informative signals about the prospects for collective action in neighboring towns, we consider the following linear probability specification,

$$a_{i,t} = \sum_{\ell=0}^{L_r} \beta_r^\ell \mathbf{r}_{i,t-\ell}\mathbf{a}_{t-\ell} + \sum_{\ell=0}^{L_\gamma} \beta_\gamma^\ell \boldsymbol{\gamma}_i \mathbf{a}_{t-\ell} + \sum_{\ell=0}^{L_{r\gamma}} \beta_{r\gamma}^\ell \boldsymbol{\gamma}_i \cdot \mathbf{r}_{i,t-\ell}\mathbf{a}_{t-\ell} + \sum_{\ell=0}^{L_d} \beta_d^\ell \mathbf{d}_i \mathbf{a}_{t-\ell} + \mu_i + \xi_t + \varepsilon_{i,t} \quad (1)$$

where $a_{i,t}$ is an indicator of collective action in town i at time t, and $\mathbf{a}_{t-\ell}$ denotes the column vector of these indicators for all towns at time $t-\ell$. Equation (1) allows for up to L_r lags of rail signals, L_γ lags of telegraph signals, $L_{r\gamma}$ lags of their interaction ($\cdot$ denotes element by element multiplication), and L_d lags of other latent networks, to induce Crusade activity in town i. The μ_i are town fixed effects, capturing all time-invariant unobservables that may make women in a given town more or less prone to collective action. The ξ_t are time fixed effects, capturing time-varying shocks affecting

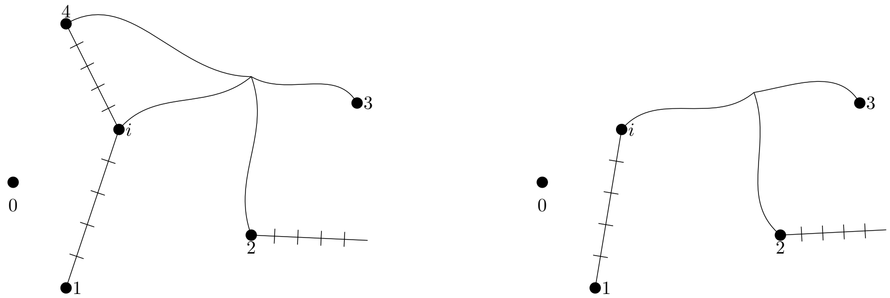

Figure 5: Types of Connections between Neighboring Towns and Identification. The figure to the left illustrates all the potential types of connections between towns: town i and town 0 are not connected by an observed network; town i and town 1 are connected by a direct rail link; town i and town 2 are connected through the telegraph, and town 2 also has railroad access; town i and town 3 are connected through the telegraph and town 3 does not have railroad access; town i and town 4 are connected both by a direct rail link and by the telegraph. The figure to the right shows that effectively in our sample there are no pairs of towns like $(i, 4)$, making the identification of the interaction effect from equation (1) not possible.

all towns in a given period. In practice, these will capture the aggregate time-path of the Crusade we illustrated in Figure 1. Finally, the $\varepsilon_{i,t}$ capture all time-varying unobservables relevant for the collective action decisions of women, possibly including a lagged dependent variable.

Unfortunately, recovering the technological complementarity effects $\{\beta^{\ell}_{r\gamma}\}_{\ell=0}^{L_{r\gamma}}$ in equation (1) is infeasible given the structure of the telegraph network. The reason is that the spatial distribution of telegraph stations across U.S. towns at the time was highly negatively correlated. Neighboring towns of a town with a telegraph were very unlikely to have a telegraph station themselves. Telegraph companies explicitly followed a strategy that located telegraph stations far apart from each other. As a result, there are very few pairs of towns directly linked by the railroad *and* with access to the telegraph network. Figure 5 illustrates why our observed network structure does not allow for the identification of the interaction effects in equation (1).

As such, we delay our discussion of technological complementarities to Section 6, and focus here on an econometric specification of the form

$$a_{i,t} = \sum_{\ell=0}^{L_r} \beta^{\ell}_{r} \mathbf{r}_{i,t-\ell} \mathbf{a}_{t-\ell-1} + \sum_{\ell=0}^{L_\gamma} \beta^{\ell}_{\gamma} \boldsymbol{\gamma}_i \mathbf{a}_{t-\ell-1} + \sum_{\ell=0}^{L_d} \beta^{\ell}_{d} \mathbf{d}_i \mathbf{a}_{t-\ell-1} + \mu_i + \xi_t + \varepsilon_{i,t} \tag{2}$$

Thanks to the panel structure of our data, the usual reflection problem that arises in the estimation of social interactions is not a concern in this setting (Manski (1993)). Another recurrent empirical concern in the network peer effects literature is the endogeneity of the network structure itself. When links in a network are created based on characteristics that are also correlated with the behavior under study, it is hard to assess whether a correlation in behavior across linked agents is the result of a social interaction effect, or simply of the selection into the friendship. In our setting, unless rails and telegraph cables were laid as a function of the similarity of neighboring towns along characteristics relevant for collective action, this concern will be minor. Population size is important for collective action, and the geographic distribution of both networks is strongly correlated with population density. Our ability to include town-level fixed effects thanks to the panel structure of our data, however, would require that the effect of population size on protest activity be time-varying for this to be a concern. As a result, we treat all networks as predetermined. The short duration of the Crusade makes this assumption quite reasonable.

Estimation of the coefficients in equation (2), however, is fraught with other econometric challenges. The possibility of persistent unobserved characteristics correlated with the collective action choices of neighboring towns, and relevant for the collective action decisions of women in town i, is particularly serious. This is most obviously the case if we consider the existence of a latent network (roads, waterways, etc.) through which information about *the same* neighboring actions is also transmitted. Were we to leave the term $\mathbf{d}_i \mathbf{a}_{t-\ell-1}$ inside the error, even an instrument that generates exogenous variation in information flows $\mathbf{a}_{t-\ell-1}$ will be invalid in equation (2). In a setting with multiple networks transmitting correlated information, an instrumental variables strategy will not be useful if a subset of the networks is left as latent. To the best of our knowledge, this econometric challenge has not been highlighted by previous literature, and is the reason why we explicitly include the 'distance' network in our econometric specification, as a way of capturing alternative channels of information transmission. Even if we can control for all relevant communication networks, residual sources of correlation across neighboring towns remain a concern that make network information flow variables (rail, telegraph, and distance) endogenous at all relevant lags in equation (2).

Exclusion Restrictions and Identification

Our strategy to deal with these issues relies on disruptions of the railroad network caused by rail worker strikes and railroad accidents happening during the months of the Crusade. Using this information, we can consider the rail network as time varying, with each link being switched on or off depending on these events.[30] We code $r_{ij,t} = 0$ if despite there being a rail connection between towns i and j, at time t either a strike or an accident affecting towns i or j took place. Our key identifying assumption will be that $\text{Cov}(r_{ik,s}, \varepsilon_{i,t} | \mu_i) = 0$ for all neighboring towns (i, k) and for adjacent time period pairs (s, t). We believe this exclusion restriction is reasonable in our context; disruptions in the railroad caused by strikes or accidents should be unlikely to predict time-varying unobservables relevant to the crusaders' protest decisions. The identifying assumption is especially plausible because for a large fraction of our sample, strikes and accidents affecting a given pair of towns took place relatively far from the pair.[31] We rely on this assumption to construct valid instruments for all the endogenous variables in equation (2).[32] These instruments directly affect the flows of information,

[30]See Section 3 for a description of our classification criteria for link breaks.

[31]For example, because railroad strikes are themselves a form of collective action, a concern would be that information about them motivated women to engage in their own collective action. These strikes do not appear to have been widely reported by newspapers, and we found no mention of a relationship between the strikers and the crusaders in our research.

[32]An additional concern with the use of instruments in network settings is spatial correlation of the instrument. When this is the case, it is possible that the cross sectional variation in the instrument picks up some of the variation in the spatially correlated unobservables (see Acemoglu et al. (2015) for a discussion of this issue). In our setting, strikes and railroad accidents naturally affected neighboring towns. Notice, however, that our empirical strategy does not use the cross-sectional variation in the railroad disruptions. We exploit the within-town time-series variation they

the main channel of social interactions we are exploring. As such, we expect them to be strong predictors of protest activity in neighboring towns. Below we will document this to be the case.

Consider first $\mathbf{r}_{i,t}\mathbf{a}_{t-1}$, the weighted sum of Crusade events of town i's railroad neighbors one period earlier. It varies both because the set of effective rail neighbors of town i, $R_t(i)$, varies *exogenously* over time as accidents and railroad-worker strikes take place, and because $\mathbf{a}_{t-1}$ varies *endogenously* over time and across i's railroad neighbors $j \in R_t(i)$. If equation (2) applies for any town, then $a_{j,t-1}$ varies exogenously because the set of effective rail neighbors of town j, $R_{t-1}(j)$, is varying over time. This provides us with a number of valid instruments for $\mathbf{r}_{i,t}\mathbf{a}_{t-1}$: i) the sum of town i's active railroad links themselves: $\mathbf{r}_{i,t}\boldsymbol{\iota} = \sum_{j\in R_t(i)} r_{ij,t}$;[33] ii) the sum across i's neighbors, of each of their active railroad links in the previous period: $\mathbf{r}_{it}\mathbf{R}_{t-1}\boldsymbol{\iota} = \sum_{j\in R_t(i)} r_{ij,t} \sum_{k\in R_{t-1}(j)} r_{jk,t-1}$; iii) following the same idea one neighbor away, the sum across i's active railroad neighbors, of the sum across each of their active railroad links in the previous period, of the sum across each of their active rail links in the period before that: $\mathbf{r}_{it}\mathbf{R}_{t-1}\mathbf{R}_{t-2}\boldsymbol{\iota} = \sum_{j\in R_t(i)} r_{ij,t} \sum_{k\in R_{t-1}(j)} r_{jk,t-1} \sum_{q\in R_{t-2}(k)} r_{kq,t-2}$.

We can construct instruments for the telegraph network and for the distance network information flows following the same idea. For telegraph information flows we use the rail link variation of telegraph neighbors, and the rail link variation of rail neighbors of own telegraph neighbors. Our instruments for $\boldsymbol{\gamma}_i\mathbf{a}_{t-1}$ are: i) $\boldsymbol{\gamma}_i\mathbf{R}_{t-1}\boldsymbol{\iota} = \sum_{j\in\Gamma(i)} \gamma_{ij} \sum_{k\in R_{t-1}(j)} r_{jk,t-1}$, and ii) $\boldsymbol{\gamma}_i\mathbf{R}_{t-1}\mathbf{R}_{t-2}\boldsymbol{\iota} = \sum_{j\in\Gamma(i)} \gamma_{ij} \sum_{k\in R_{t-1}(j)} r_{jk,t-1} \sum_{q\in R_{t-2}(k)} r_{kq,t-2}$. For latent network information flows we use the rail link variation of distance neighbors, and the rail link variation of rail neighbors of own distance neighbors. Our instruments for $\mathbf{d}_i\mathbf{a}_{t-1}$ are: i) $\mathbf{d}_i\mathbf{R}_{t-1}\boldsymbol{\iota} = \sum_{j\in D(i)} d_{ij} \sum_{k\in R_{t-1}(j)} r_{jk,t-1}$, and ii) $\mathbf{d}_i\mathbf{R}_{t-1}\mathbf{R}_{t-2}\boldsymbol{\iota} = \sum_{j\in D(i)} d_{ij} \sum_{k\in R_{t-1}(j)} r_{jk,t-1} \sum_{q\in R_{t-2}(k)} r_{kq,t-2}$. Lags of each of these instruments will be valid instruments for the corresponding lags of the endogenous regressors in equation (2).

Model Selection and Inference

In our context, an important question relates to the relevant lag structure of equation (2). Information travels at different speeds along different communication networks. Naturally, we expect telegraph information flows to have effects at higher frequencies, and thus for shorter lags to be most relevant for that technology. The distance network, in turn, is intended to capture communication taking place along roads and rivers, so we expect information traveling through these alternative communication networks to be the slowest. Higher lags should be most relevant in this case.

Our empirical strategy will begin with a formal model selection statistical test to find the lag structure that most closely approximates the relevant frequencies at which information affected protest diffusion. We rely on Andrews and Lu (2001), who propose a model selection test for panel data and GMM estimation, ideally suited to our setting. The test is based on the J statistic for over-identifying restrictions, and incorporates a degrees-of-freedom adjustment that takes into account varying degrees of over-identification across models being compared. When we estimate a model

generate.

[33] $\boldsymbol{\iota}$ represents a column vector of ones.

with the wrong lag structure, the true lags (or a subset of them) are left in the error term. As a result, the instruments should be correlated with the residual from such a model, leading to a large J statistic. In contrast, if we estimate a model with the correct lag structure, valid instruments will be uncorrelated with the residual, leading to a small value for the J statistic. The test simply selects the model with the smallest test statistic. Naturally, we are unable to compare the vast number of possible combinations of lag structures. As we will discuss below, however, our exercise very quickly pointed us to a lag structure that both dominates all other models based on the Andrews and Lu (2001) test, and is the most parsimonious.

Before presenting our main results, we briefly discuss inference, which is particularly challenging in network settings. As pointed out by Chandrasekhar and Lewis (2011), estimation of network models based on sampled nodes –even if at random- leads to biased estimates and incorrect inference. In our setting, our estimation sample includes all towns in the U.S. based on the 1870 Census, effectively allowing us to observe the population of nodes in all of our networks. Spatial correlation in unobservables is another challenge for inference in network settings. To address this issue, we compute standard errors that allow for contemporaneous correlation in the residuals across railroad neighboring towns in the spirit of Conley (1999), and allow for arbitrary inter-temporal correlation in the errors within a town.[34]

Main Estimates

Table 3 presents our benchmark 2SLS estimation results of equation (2) for an array of competing lag structures. We build the estimation sample panel as follows. First, because the responsiveness to information was likely heterogeneous over time as the Crusade spread, we restrict attention to the period between day 50 and day 150 after the first Crusade event took place. As we illustrated in Figure 1, this is the interval where most activity happened. We then created synthetic time periods of five contiguous days, within which we aggregated all our variables. In a series of robustness checks discussed below, we alter the definition of a period to include either three days or seven days. In the benchmark specification, we do not distinguish between types of Crusade events. The dependent variable is a dummy for whether any type of crusading event took place (meetings, petitions, or marches) within the five-day period. The regressors include any type of neighboring event as well. We instrument each regressor with the corresponding lag of the instruments described above. Finally, based on our knowledge of the structure of the Crusade, we eliminate all time periods after a town

[34] Defining $\mathbf{X}$ to be the matrix of regressors, and $\mathbf{Z}$ the matrix of instruments, the robust network-correlation corrected variance matrix of the 2SLS fixed effects estimator takes the form:

$$(\mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{W}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X}(\mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X})^{-1}$$

where

$$\mathbf{W} = \sum_{i=1}^{n} \mathbf{Z}_i'\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i'\mathbf{Z}_i + \sum_t \left(\sum_{i=1}^{n} \sum_{j \in R_t(i)} \mathbf{z}_{it}'\varepsilon_{it}\varepsilon_{jt}\mathbf{z}_{jt} \right).$$

experienced a march, leading to a slightly unbalanced panel. After such an event had taken place, no further collective action could occur.

In column (1) we begin with the simplest specification, including only the first lag of neighboring Crusade activity signals along rail, telegraph, and distance networks. Column (2) presents results for a specification that includes lags of order two instead. Neither first or second lags of distance network neighboring signals have any effect on the likelihood of collective action, with coefficients of zero. The effect of railroad network signals, in contrast, is highly significant and positive, and very similar in magnitude for the first or second lag models. The coefficient for the telegraph network signals is larger and also similar across lags, but has a large standard error. In column (3) we include instead the third lag of the distance network signals, which appears highly significant and positive. This specification includes the second lag for the network signals and the first lag for the telegraph signals. The last three rows of the table present the J statistic with its associated p-value for testing the null hypothesis of the joint validity of our instruments, along with Andrews and Lu (2001) model selection criterion. We are unable to reject the null hypothesis that our instrument set is valid across all model specifications. However, the model in column (3) has the smallest model selection test statistic (-49.69), not only across all models we report in Table 3 but also among alternative lag structures we do not show to conserve space.

Our preferred model has a parsimonious structure, and coincides with our initial hypothesis about the nature of information flows across the different networks. Signals along the telegraph appear to have the fastest effect (with the first lag consistently being the largest and most precisely estimated), followed by signals along the railroad (for which the second lag is the most relevant), followed by signals along the distance network (for which only the third lag is significant). The remaining columns in the table explore alternative lag structures, including models with several lags of the same network simultaneously. None of these, however, outperform the model in column (3). The quantitative effects implied by the coefficients of this model are very large. On average, 50 towns experienced Crusade activity every 5-day period. In a population of more than 15 thousand towns, this implies a mean for the dependent variable of 0.003. A coefficient of 0.02 (s.e. = 0.007) on $\mathbf{r}_{i,t}\mathbf{a}_{t-2}$ means that on average, a Crusade in a rail-neighboring town happening 5 to 10 days before, multiplies the likelihood of undertaking collective action by 6.6 ($= 0.02/0.003$).

For completeness, Appendix Table A.1 presents the R^2 and F statistics for the corresponding first stages of each of the models in Table 3. These statistics are presented, from top to bottom, in the same order as their corresponding endogenous regressor appears in Table 3. Across all specifications, they show we have very strong first stages for all lags of our endogenous regressors. All F statistics in the table are above 14. In the first three columns of Table A.2, we narrow our attention to our preferred model, and present the estimates of each of the three first stages corresponding to the regressors in column (3) of Table 3. As column (1) reports, variation in own railroad link disruptions, $\mathbf{r}_{i,t-1}\boldsymbol{\iota}$, appears to be strongly correlated with neighboring railroad signals generated a period before, $\mathbf{r}_{i,t}\mathbf{a}_{t-2}$. The effect is positive, implying that in periods with a higher than average number of active

Causal Effects of Crusade Signals along the Railroad and Telegraph Networks: Lag Specification Model Selection

	Dependent Variable: Any Crusade Activity a_{it} -Meetings, Petitions, Marches-									
Second stages:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\mathbf{r}_{i,t}\mathbf{a}_{t-1}$	0.0277			-0.0163		-0.0169	-0.0094			0.0069
	[0.0074]			[0.0162]		[0.0162]	[0.0174]			[0.0136]
	(0.0068)			(0.0146)		(0.0146)	(0.0154)			(0.0120)
$\mathbf{r}_{i,t-1}\mathbf{a}_{t-2}$		0.0287	0.0201	0.0243	0.0198	0.0247	0.0277		0.0208	-0.004
		[0.0080]	[0.0078]	[0.0136]	[0.0078]	[0.0137]	[0.0149]		[0.0077]	[0.0105]
		(0.0069)	(0.0072)	(0.0124)	(0.007)	(0.0124)	(0.0131)		(0.0072)	(0.0100)
$\mathbf{r}_{i,t-2}\mathbf{a}_{t-3}$								0.0132		0.0103
								[0.0081]		[0.0104]
								(0.0071)		(0.0089)
$\boldsymbol{\gamma}_i\mathbf{a}_{t-1}$	0.0714		0.0919	0.1031	0.1354	0.1291	0.0563	0.0978		0.1272
	[0.0704]		[0.0633]	[0.0637]	[0.0854]	[0.0844]	[0.0795]	[0.0634]		[0.0751]
	(0.0610)		(0.0572)	(0.0575)	(0.0716)	(0.0709)	(0.0653)	(0.0571)		(0.0649)
$\boldsymbol{\gamma}_i\mathbf{a}_{t-2}$		0.0852			-0.0237	-0.0155	0.0228			0.0033
		[0.0620]			[0.0659]	[0.065]	[0.0654]			[0.067]
		(0.0517)			(0.0582)	(0.0576)	(0.0566)			(0.0593)
$\boldsymbol{\gamma}_i\mathbf{a}_{t-3}$									0.019	0.004
									[0.059]	[0.0364]
									(0.0521)	(0.0316)
$\mathbf{d}_i\mathbf{a}_{t-1}$	0.0001						0.0006			-0.0022
	[0.0005]						[0.0015]			[0.0015]
	(0.0005)						0.0013			(0.0014)
$\mathbf{d}_i\mathbf{a}_{t-2}$		-0.0006					-0.0006			0.0003
		[0.001]					[0.0020]			[0.0013]
		(0.0008)					(0.0017)			(0.0011)
$\mathbf{d}_i\mathbf{a}_{t-3}$			0.0027	0.0025	0.0027	0.0025		0.0023	0.0019	0.0045
			[0.0011]	[0.0011]	[0.0012]	[0.0011]		[0.0011]	[0.0011]	[0.0019]
			(0.0009)	(0.0009)	(0.0009)	(0.0009)		(0.0009)	(0.0009)	(0.0017)
No. of towns	15960	15947	15934	15934	15934	15934	15947	15934	15934	15934
Max. no. of periods	18	17	16	16	16	16	17	16	16	16
Observations	299154	283194	267247	267247	267247	267247	283194	267247	267247	267247
J-test statistic	3.36	1.09	0.295	3.22	0.32	3.37	6.27	0.328	2.4	17.9
J-test p-value	0.498	0.895	0.990	0.781	0.997	0.848	0.616	0.988	0.662	0.116
Andrews-Lu (2001) stat.	-47.07	-49.12	-49.69	-46.76	-49.66	-46.61	-43.94	-49.66	-47.58	-32.01

Table 3: The Effect of Information along the Rail and Telegraph Networks: Lag Specification Model Selection. The table presents panel 2SLS estimates of competing lag structure specifications of equation (2) on the universe of U.S. 1870 Census towns. In all models a time period is defined as a 5-day interval. The dependent variable is an indicator of crusading activity -meetings, petitions, or marches-. All models include period fixed effects and town fixed effects. Standard errors in square brackets are robust and allow for spatial correlation between neighboring towns along the railroad network. Standard errors in parentheses are clustered at the town level. The last row of the table reports the model selection test statistic of Andrews and Lu (2001). Appendix table A.1 reports the first stage R squared and F-statistics corresponding to each endogenous regressor in the corresponding column, from top to bottom. The first three columns of Appendix table A.2 report the first stage coefficients of the model in column (3). Instruments in all specifications are based on a 50 Km. rail accident radius.

links, a town was more likely to receive neighboring information through the railroad. Our second and third instruments for $\mathbf{r}_{i,t-1}\mathbf{a}_{t-2}$ are $\mathbf{r}_{i,t-1}\mathbf{R}_{t-2}\boldsymbol{\iota}$ and $\mathbf{r}_{i,t-1}\mathbf{R}_{t-2}\mathbf{R}_{t-3}\boldsymbol{\iota}$. Railroad disruptions of neighboring towns and neighbors of neighboring towns are significantly negatively correlated with railroad information flows.

Column (2) presents the first stage for $\boldsymbol{\gamma}_i\mathbf{a}_{t-1}$, telegraph information flows. In this case, rail signal disruptions of neighbors in the telegraph network, $\boldsymbol{\gamma}_i\mathbf{R}_{t-1}\boldsymbol{\iota}$, are very strong predictors of less information being generated by those neighbors. In column (3), we report the first stage for the third lag of distance-network information flows, $\mathbf{d}_i\mathbf{a}_{t-3}$. Once again, variation in railroad links of neighbors and neighbors of neighbors correlate strongly with this endogenous regressor.[35]

Heterogeneous Effects across Type of Crusade Events

We now report results that disaggregate the average effect of different types of signals. Recall that we can distinguish between meetings, petitions, and marches. While meetings were organizational events where women in crusading communities debated whether to engage in militant action, petitions and marches entailed additional risks associated with protest activity. Meetings may also have generated information about the success of such activity (which we do not observe). If the diffusion of the Crusade involved social learning, we would expect these different types of events to have induced differential enthusiasm in neighboring towns. As mentioned in section 3, differential inferences based on the type of event would be less likely in a contagion or social influence setting.

In Table 4 we report 2SLS estimates of models using the lag structure found as optimal by our model selection test (lag 2 for the railroad network, lag 1 for the telegraph network, lag 3 for the distance network). The dependent variable still includes all types of Crusade events, while the endogenous regressors include only meetings (in columns (1) to (4)), or petitions and marches (in columns (5) to (8)). The top panel of the table reports the second stage estimates, while the bottom panel presents the R^2 and F statistics of the corresponding first stages. To probe the robustness of these estimates, we present results using two alternative classification criteria for the impact of railroad accidents. In columns (1), (2), (5), and (6), we use our benchmark 50 km. radius definition. In columns (3), (4), (7), and (8), we use a 120 km. radius definition instead. Similarly, we present results for two alternative panel period definitions. In odd-numbered columns we use our benchmark 5-day periods, while in even-numbered columns we use a 3-day period definition.

The first four columns allow us to assess the responsiveness of collective action to information about meetings in neighboring towns. Across all specifications, the magnitude of the coefficients is very similar, and the pattern of significance for the railroad signals and the telegraph signals is the same as that for the models that do not distinguish between types of events. We find this remarkable,

[35] Recall that our baseline definition of a rail link break is based on turning off all links within a 50 km. radius of a railroad accident. As a robustness exercise, the remainder of Table A.1 presents additional first stage results for our benchmark model, varying the definition of a rail link break. In Columns (4)-(6), we define rail links to be broken around an 80 km. radius of each railroad accident. In columns (7)-(9), we define rail links to be broken around a 120 km. radius of each railroad accident. In both cases, the patterns and magnitude of effects are very similar to those of the benchmark 50 km. case.

particularly as we compare the 5-day period models with the 3-day period models. By construction, 5-day panels have up to 16 time periods, while 3-day panels have up to 30 time periods. The magnitude of the coefficients for the telegraph signals is around twice as large as the magnitude of the coefficients for the railroad signals. Once again, the railroad signals are very precisely estimated while standard errors for the telegraph signals are large. However, telegraph signals always have positive point estimates. All point estimates appear to be slightly larger when using the 120 Km. definition for the instruments compared to the 50 Km. definition. In contrast to our previous results, however, when we disaggregate protest activity by type of event, the distance network effects do not show up as significant.

In the last four columns we assess the responsiveness of collective action to information about petitions or marches in neighboring towns. The qualitative results are very similar to those for meetings. A key difference appears; we find the effects of petitions and marches to be smaller in magnitude than the effects of meetings. The differences are statistically significant for 6 out of the 8 comparisons. Take, for example, the coefficient on rail meeting signals in column (4) of 0.07 (s.e. = 0.018), and the coefficient on rail petitions or marches signals in column (8) of 0.039 (s.e = 0.015). Both are highly significant, but the effect for meetings is 80 percent larger. We observe a similar pattern for the effects of information traveling along the telegraph.

A possible explanation for this result is differential media coverage of meetings compared to petitions and marches. Our newspaper article search from the Library of Congress online repository (described in Section 3), however, suggests this was not the case. Meetings were not reported at a higher rate, being mentioned in 3379 out of the 4006 Crusade-related articles in our sample. In fact, most articles mentioning meetings also mention sit-ins, marches, etc. Social learning would suggest differential responses to different types of events in neighboring towns. A simple model of learning where meetings are especially informative about the value of holding a meeting, and rallies are especially informative about the value of holding a rally, but where the cost of organizing a meeting is on average lower than the cost of organizing a rally, for example, could rationalize larger responses to neighboring meetings.

Specification Tests and Robustness

We now discuss a battery of specification tests, placebo exercises, and additional robustness checks to probe the sensitivity of our findings. We begin with Appendix Table A.3. This table presents estimates of our benchmark lag structure specification without distinguishing different types of Crusade events, but using alternative estimators. In columns (1) and (2), we present results using a 5-day period panel, and in columns (3) and (4) results use a 3-day period panel. Odd-numbered columns present OLS estimates. Regression coefficients in these specifications are positive and significant for all network information flows. They are, however, smaller than their 2SLS counterparts, and point

Heterogeneity in Type of Crusade Event Signal

Endogenous Regressors $\mathbf{a}_t$:	Meetings				Petitions and Marches			
Instrument Variation:	50km accident radius		120km accident radius		50km accident radius		120km accident radius	
Period Definition:	5 days	3 days	5 days	3 days	5 days	3 days	5 days	3 days
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathbf{r}_{i,t-1}\mathbf{a}_{t-2}$	0.052	0.050	0.076	0.070	0.030	0.035	0.062	0.039
	[0.019]	[0.015]	[0.024]	[0.018]	[0.013]	[0.010]	[0.018]	[0.015]
	(0.018)	(0.014)	(0.022)	(0.016)	(0.012)	(0.010)	(0.016)	(0.013)
$\boldsymbol{\gamma}_i\mathbf{a}_{t-1}$	0.026	0.129	0.143	0.133	0.086	0.119	0.091	0.090
	[0.075]	[0.082]	[0.179]	[0.079]	[0.069]	[0.084]	[0.061]	[0.065]
	(0.065)	(0.072)	(0.175)	(0.070)	(0.062)	(0.077)	(0.056)	(0.060)
$\mathbf{d}_i\mathbf{a}_{t-3}$	0.003	0.000	0.001	0.001	0.004	0.000	0.004	0.001
	[0.002]	[0.001]	[0.001]	[0.001]	[0.002]	[0.001]	[0.002]	[0.001]
	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
No. of towns	15934	15950	15934	15950	15934	15950	15934	15950
Max. no. of periods	16	30	16	30	16	30	16	30
Observations	267247	487548	267247	487548	267247	487548	267247	487548
J-test statistic	3.673	1.782	5.885	0.695	3.159	3.85	1.635	8.08
J-test p-value	0.452	0.776	0.208	0.952	0.532	0.427	0.802	0.089
Panel B:	First Stages							
F-statistic	33.59	36.62	28.56	28.54	22.44	20.34	18.36	24.87
R squared	0.011	0.010	0.009	0.008	0.014	0.012	0.009	0.258
F-statistic	15.59	18.89	10.59	12.01	30.39	21.06	26.62	27.7
R squared	0.024	0.012	0.008	0.012	0.022	0.008	0.027	0.012
F-statistic	461.5	993.3	955.7	1648.2	211.3	770.0	430.9	197.4
R squared	0.239	0.259	0.245	0.264	0.247	0.255	0.250	0.008

Table 4: The Effect of Information along the Rail and Telegraph Networks: Heterogeneity. The table presents panel 2SLS estimates of equation (2) across alternative specifications. The dependent variable is an indicator of crusading activity -meetings, petitions, or marches-. All models include period fixed effects and town fixed effects. Standard errors in square brackets are robust and allow for spatial correlation between neighboring towns along the railroad network. Standard errors in parentheses are clustered at the town level. All columns use the lag structure identified as optimal by the Andrews and Lu (2001) test in Table 3 (second order lag for the rail signals, first order lag for the telegraph signals, third order lag for the geographic signals). Columns (1)-(4) restrict $\mathbf{a}_t$ in the definition of the endogenous regressors to include only meetings. Columns (5)-(8) restrict $\mathbf{a}_t$ in the definition of the endogenous regressors to include only petitions or marches. Columns (1)-(2) and (5)-(6) use the benchmark 50 Km. radius definition of rail accidents for the instruments. Columns (3)-(4) and (7)-(8) use the 120 Km. radius definition of rail accidents for the instruments. Odd-numbered columns use the benchmark 5-day interval period definition. Even-numbered columns use the 3-day interval period definition. Panel B reports the first stage R squared and F-statistics corresponding to each endogenous regressor in the corresponding column, from top to bottom.

out the importance of instrumenting neighboring protest activity.[36] Columns (2) and (4) then present GMM results based on moments constructed using the same set of instruments we employ for our 2SLS empirical strategy. The magnitude of the GMM estimates is remarkably close to that of our benchmark 2SLS estimates for the three communication networks. Moreover, the standard errors for the telegraph information flows are now smaller, making these coefficients statistically significant at standard levels.

Appendix Table A.4 reproduces the model selection exercise from Table 3, but using a panel based on 3-day periods instead of our benchmark 5-day period definition. Quantitatively and qualitatively, the results point to the same conclusions we derived in Table 3, and suggest that neither our model selection exercise nor the magnitude and significance of our results are driven by our choice of time period definition.[37]

In Table 5, we move on to a more exhaustive set of robustness checks. The first five columns use the optimal lag structure from Table 3 and our benchmark definition of a railroad link. In columns (1)-(4) we use the full sample of towns, and vary the radius used for defining railroad accidents when building our instrument set, and the number of days per period in the panel. Columns (1) and (2) use the 5-days period definition, but use 80 Km. and 120 Km. radii for the instrument construction. Results are unchanged. In columns (3) and (4) we fix the accident radius at 50 Kms., but present results for panels based on 3-day or 7-day periods. Despite the very different number of effective periods, coefficients are once again very similar, although the standard error for the railroad network signals is larger in the 7-day specification. In column (5) we change the sample of towns for estimation, by excluding all towns for which there is no within-town variation in any of the instruments. Although this reduces the sample size considerably, results are unchanged. The over-identification test in this case similarly cannot reject the validity of the instrument set. We conclude this table by changing the definition of a link in the railroad network. Notice this changes the construction of the railroad network endogenous regressor and the construction of all of the instruments. We re-define this network by classifying not only neighbors, but also neighbors of neighbors, as directly linked in the railroad network. In this case, the lag structure model selection test (which we omit to save space)

[36] A downward bias of OLS is precisely what we would expect in our setting: if the error term in equation (2) contains a lagged dependent variable and crusading activity is *negatively* autocorrelated, then a positive correlation between own and neighboring protest activity will lead to a downward-biased OLS estimator. To illustrate this point, consider a simplified model where only the first lag of railroad network information has an effect, but where a lagged dependent variable is present and left in the error term:

$$a_{i,t} = \beta_r \mathbf{r}_{i,t}\mathbf{a}_{t-1} + (\rho a_{i,t-1} + \varepsilon_{i,t}) + \mu_i$$

For simplicity, suppose the lagged dependent variable is the only source of endogeneity of $\mathbf{r}_{i,t}\mathbf{a}_{t-1}$. Then the probability limit of the OLS estimator of this model will be:

$$\beta_r^{OLS} = \beta_r + \rho \frac{\text{Cov}(\mathbf{r}_{i,t}\mathbf{a}_{t-1}, a_{i,t-1})}{\text{Var}(\mathbf{r}_{i,t}\mathbf{a}_{t-1})}$$

which is smaller than β_r if $\rho < 0$ and the covariance term is positive. In our setting, the within-town auto-correlation in crusading activity is negative, because periods immediately following an event are very unlikely to exhibit an event as well. The average autocorrelation in $a_{i,t}$ across towns is -0.25.

[37] Appendix Table A.5 presents the corresponding first stage results for the models in Appendix Table A.4.

chooses a model that includes lags 2 and 3 of the railroad network, lags 1, 2, 3 and 4 of the telegraph network, and lag 3 of the distance network. In column (6) we use the full sample of towns, while in column (7) we use the restricted sample of towns with within-town instrument variation. In both cases, and despite the different lag structure, for both rail and telegraph information flows the net effect across lags is very similar in magnitude to our benchmark results.[38]

In Appendix table A.7, we present a placebo exercise to support the validity of our instruments. We build false instruments, by taking the base railroad network and simulating railroad link breaks at random every day, at the same rate we observe them break in the data. Using our benchmark lag structure identified as optimal, columns (1) and (2) present results for instruments based on a false 50 Km. accident radius. Columns (3) and (4) present results for instruments based on a false 120 Km. accident radius. In odd-numbered columns the panel is based on 5-day periods, while in even-numbered columns the panel is based on 3-day periods. No coefficients appear to be statistically significant. Moreover, the railroad network coefficients are negative in some of the specifications, and always very small.

Next, we present results of a specification test of our model and the exclusion restrictions it is based on, following Acemoglu et al. (2015). We recover the 2SLS residuals from our benchmark specification, and regress them on three different network centrality statistics for the railroad network: degree, betweenness, and eigenvector centrality.[39] These are commonly used statistics in the network literature that capture different dimensions of connectivity. If our model is close to correctly specified, these residuals should be uncorrelated with towns' centrality characteristics. In these models we do not include town-level fixed effects because our interest is to assess the relationship between residuals and centrality measures *across* towns. The models include the level of the centrality statistic, and a full set of interactions between the centrality statistic and time fixed effects. We present these results graphically in Figure 6.[40] The figure plots the point estimates and associated confidence intervals for the interaction terms in each of the models. The left-most figure plots the results for the model based on degree centrality. The central figure plots the results for the model based on betweenness centrality. The right-most figure plots the results for the model based on eigenvector centrality. We find no discernible pattern over time, and only one of the 16 interactions is significant at the 95 percent level in the degree centrality model. In the betweenness centrality model, only two of the 16 interactions is statistically different from zero. For the eigenvector centrality model, four of the 16 coefficients are statistically significant. However, some of these are positive and some are negative. Moreover, neither of the baseline centrality measures is significantly correlated with the 2SLS residuals. Taken together, these results suggest very little correlation between network centrality and the unexplained variation in protest adoption.

[38] Appendix Table A.6 presents the first stage results corresponding to the models in Table 5.

[39] Notice that based on our definition of the telegraph network –where all towns with telegraph access are assumed to be linked to each other–, these centrality statistics do not vary within towns in the telegraph network. Thus, this exercise is only meaningful for the railroad network.

[40] For completeness, we report the full set of estimates in Appendix Table A.8.

Causal Effects of Crusade Signals along the Railroad and Telegraph Networks: Robustness

Rail Neighbors:	First order links					First and second order links	
Sub-sample:	All towns				Instruments vary	All towns	Instruments vary
Instrument Variation: (accident radius)	80km	120km	50km	50km	50km	50km	50km
Period Definition:	5 days	5 days	3 days	7 days	5 days	5 days	5 days
Second stages:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\mathbf{r}_{i,t-1}\mathbf{a}_{t-2}$	0.029	0.036	0.023	0.010	0.019	-0.017	-0.017
	[0.0082]	[0.0105]	[0.0064]	[0.0077]	[0.0093]	[0.0114]	[0.0138]
	(0.0076)	(0.0093)	(0.0061)	(0.0079)	(0.0082)	(0.010)	(0.0118)
$\mathbf{r}_{i,t-2}\mathbf{a}_{t-3}$						0.021	0.019
						[0.0097]	[0.0120]
						(0.0083)	(0.0096)
$\boldsymbol{\gamma}_i\mathbf{a}_{t-1}$	0.094	0.073	0.130	0.108	0.089	0.146	0.145
	[0.070]	[0.0510]	[0.0575]	[0.056]	[0.0730]	[0.0285]	[0.0331]
	(0.0670)	(0.0471)	(0.0516)	(0.0450)	(0.0655)	(0.0245)	(0.0280)
$\boldsymbol{\gamma}_i\mathbf{a}_{t-2}$						-0.133	-0.123
						[0.0488]	[0.0532]
						(0.0413)	(0.0448)
$\boldsymbol{\gamma}_i\mathbf{a}_{t-3}$						-0.022	-0.023
						[0.0349]	[0.0414]
						(0.0302)	(0.0342)
$\boldsymbol{\gamma}_i\mathbf{a}_{t-4}$						0.141	0.134
						[0.0451]	[0.0493]
						(0.0384)	(0.0420)
$\mathbf{d}_i\mathbf{a}_{t-3}$	0.0013	0.0015	-0.0002	0.0028	0.0025	-0.0028	-0.0008
	[0.0008]	[0.0007]	[0.0006]	[0.001]	[0.0012]	[0.0024]	[0.0022]
	(0.0007)	(0.0007)	(0.0005)	(0.0009)	(0.0010)	(0.0020)	(0.0019)
No. of towns	15934	15934	15950	15906	5095	15909	4628
Max. no. of periods	16	16	30	11	16	15	15
Observations	267247	267247	487548	188384	85533	251313	73035
J-test statistic	2.428	1.857	0.704	1.463	0.379	9.035	9.062
J-test p-value	0.658	0.762	0.951	0.833	0.984	0.434	0.432

Table 5: The Effect of Information along the Rail and Telegraph Networks: Robustness The table presents panel 2SLS estimates of equation (2) across alternative specifications. The dependent variable is an indicator of crusading activity -meetings, petitions, or marches-. All models include period fixed effects and town fixed effects. Standard errors in square brackets are robust and allow for spatial correlation between neighboring towns along the railroad network. Standard errors in parentheses are clustered at the town level. Columns (1)-(5) use the benchmark railroad edge definition of a first order connection, and use the lag structure identified as optimal by the Andrews and Lu (2001) test in Table 3 (second order lag for the railroad neighbors' Crusade events, first order lag for the telegraph neighbors' Crusade events, and third order lag for the geographic neighbors' Crusade events). Columns (6)-(7) use an alternative railroad edge definition of first or second order connections. Columns (1)-(4) and (6) use the full universe of 1870 U.S. Census towns. Columns (5) and (7) restrict the sample to those towns where at least one instrument varies over time. Column (1) uses the 80 Km. radius definition of rail accidents for the instruments. Column (2) uses the 120 Km. radius definition of rail accidents for the instruments. Columns (3)-(7) use the benchmark 50 Km. radius definition of rail accident for the instruments. Columns (1), (2), (5)-(7) use the benchmark 5-day interval period definition. Column (3) uses an alternative 3-day interval period definition. Column (4) uses an alternative 7-day interval period definition. Appendix Table A.6 reports the first stage R-squared and F-statistics corresponding to each endogenous regressor in the corresponding column, from top to bottom.

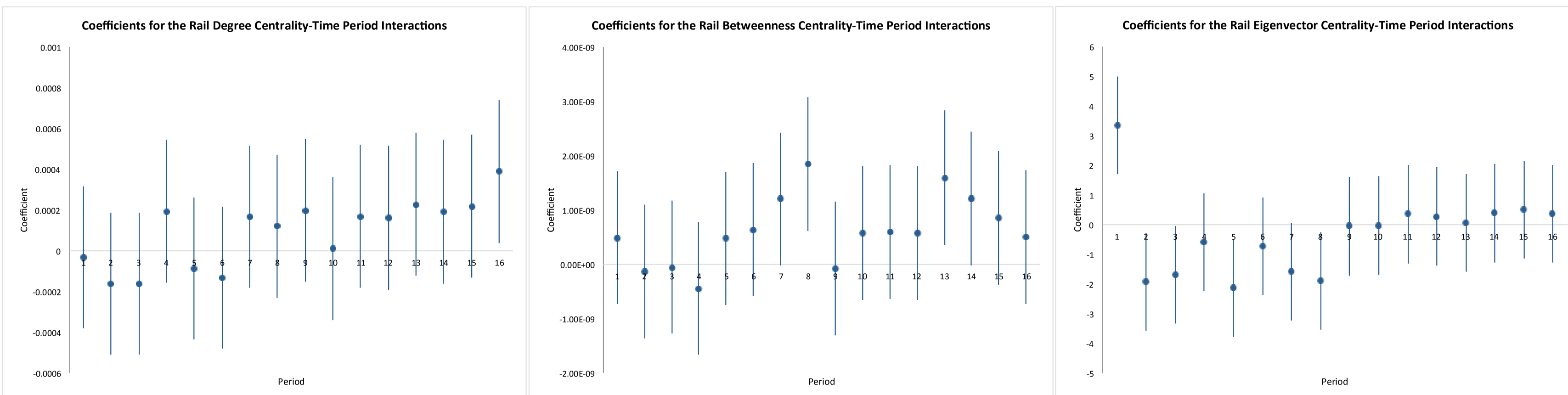

Figure 6: Specification Tests: Coefficients on the Railroad Centrality Statistics - Time Period Dummies Interactions. The figures plot the coefficients and corresponding confidence intervals on the interactions between time period dummy variables and alternative railroad network centrality statistics, from the OLS estimates reported in Appendix Table A.8. The dependent variable in each case corresponds to the 2SLS residual from the optimally selected lag specification in column (3) of Table 3. The left-most figure uses degree as the centrality statistic. The central figure uses betweenness as the centrality statistic. The right-most figure uses eigenvector as the centrality statistic. All models include period fixed effects and the corresponding centrality statistic.

Newspaper Coverage Responsiveness to Temperance Crusade Events					
	Dependent variable: Mentions of town i in Crusade-related articles of other town newspapers				
Between days:	[t,t+10)	[t+10,t+20)	[t+20,t+30)	[t+30,t+40)	[t+40,t+50)
	(1)	(2)	(3)	(4)	(5)
$a_{i,t}$	0.162	0.306	0.178	0.134	0.024
	(0.081)	(0.100)	(0.058)	(0.057)	(0.040)
R squared	0.43	0.45	0.47	0.49	0.53
No. of towns	802	802	802	802	802
No. of periods	21	20	19	18	17
Observations	16842	16040	15238	14436	13634

Table 6: Newspaper Coverage of Temperance Crusade Events. The table presents panel regression estimates for the number of articles mentioning town i in a given ten-days interval, across *all* newspapers in the *Chronicling America* online newspaper repository of the Library of Congress, excluding town i newspapers. The explanatory variable measures the number of Temperance Crusade events taking place in town i during time period t. The sample includes all crusading towns. All specifications include town fixed effects and period fixed effects. The dependent variable in column (1) is the contemporaneous number of article mentions. The dependent variable in column (2) is the first lead of the number of article mentions. The dependent variable in column (3) is the second lead of the number of article mentions. The dependent variable in column (4) is the third lead of the number of article mentions. The dependent variable in column (5) is the fourth lead of the number of article mentions. Standard errors and robust and clustered at the town level.

The Local Newspaper Channel

We conclude this section presenting some complementary evidence of the importance of the newspaper as a channel of information diffusion of the Temperance Crusade. We do so relying on our newspaper text analysis described in Section 3 and in more detail in Appendix C. We recorded the number of articles in newspapers from any other towns, reporting a Crusade-related event happening in town i at the town-day level.[41] Using this variable we perform two predictive exercises. First, on the panel of crusading towns, we explore whether a collective action event in town i at time t is predictive of news reports about it in other towns at future dates. We explore the effects at between 10 and 50 days ahead by using different leads of the dependent variable, after aggregating the 215 days of the Crusade into twenty-one 10-day periods for the panel. We report the estimates from these exercises in Table 6. These specifications include town and period fixed effects, and standard errors are clustered at the town level.[42]

The table illustrates that a given town is mentioned 0.16 times more in the 10 days following its Crusade event compared to days before the event. It is mentioned 0.3 times more between 10 and 20 days, 0.17 times more between 20 and 30 days, 0.13 times more between 30 and 40 days, and 0.02 times more between 40 and 50 days after its collective action event has occurred. The effects between 10 and 40 days are statistically significant, and overall they reveal an inverted U-shaped pattern that

[41]The *Chronicling America* online newspaper repository from the Library of Congress reports the town to which each newspaper was registered, and this is the information we use. Local newspapers had additional circulation in other towns, but we do not have detailed data on the geographic circulation of local newspapers.

[42]The econometric specification is:

$$\text{Crusade-related article mentions about town } i_{[t,t+\tau]} = \alpha_i + \beta a_{i,t} + \xi_t + \varepsilon_{i,t}.$$

peaks at between 10 and 20 days after the event has taken place.

In a second exercise, we look at how the likelihood of a newspaper report about Crusade-related events varies with the network path-length between the newspaper's home town and the town experiencing protest activity.[43] We do this on a panel of all newspaper home town-crusading town pairs, controling for network centrality characteristics of both towns, and for the physical distance between them. Because each town is a member of several pairs, we can alternatively include newspaper town and crusading town fixed effects. We report these results on Table A.9. Columns 1 and 2 report results for the models looking at railroad path lengths, and columns 3 and 4 report results for the models looking at telegraph path lengths. In both cases, distance along the networks makes the likelihood of a newspaper report lower, conditional on the physical distance between the pair of towns. Results are very precisely estimated in the fixed effects specifications. A one standard deviation higher number of links along the railroad (42 links) reduces the likelihood of a newspaper report by 0.79 percentage points ($= 0.00019 \times 42 \times 100$), which is close to half the baseline probability of a newspaper report in the sample. Quantitatively, the effect is similar along the telegraph network.

Given the completely different sources of our protests and newspaper text analysis datasets, we find these results consistent with the historical literature highlighting the vibrancy of the newspaper industry and of local newspaper outlets as sources of information for the Crusade. Results are also strongly suggestive of the major role that newspapers played as a channel through which the rail and telegraph networks had the effects we identified above.

5 Aggregate Dynamics: Testing Models of Social Interactions

Having identified the social interaction effects of railroad and telegraph-mediated information flows, in this section we empirically evaluate whether the patterns of spread of the Temperance Crusade across towns are consistent with the aggregate implications of any of the basic diffusion mechanisms suggested by Young (2009). In that important article, he discusses how to distinguish between alternative mechanisms of diffusion in a population –inertia, contagion, social influence, and social learning–. Each of these, under very general conditions, leaves distinguishing signatures on the aggregate path of the diffusion process. Albeit only suggestive, and similar to his analysis of the adoption of hybrid corn in the 1930s, we find evidence favoring social learning over alternative mechanisms.

Let $p(t)$ be the adoption curve: the fraction of the population who has adopted the behavior under study by time t. An adoption process driven by inertia is one where at any given time, players who have not yet adopted do so at some exogenous rate. As a result, any such process

[43] The path length between towns i and j along the corresponding network is computed as the shortest number of links between both towns (intermediate towns along the rail line for the railroad network, and intermediate stations along the telegraph network).

must be characterized by a *concave* adoption curve.[44] The top left panel in Figure 7 presents the diffusion curves of the Temperance Crusade. Eventually, 5 percent of all U.S. towns experienced some Crusade-related event, as the blue line illustrates. The figure also depicts the adoption curves separately for meetings (red line), petitions (green line), and marches (purple line). Petitions were the least frequent type of event, eventually occurring in 1.5 percent of all towns, while meetings and marches eventually took place in around 3 percent of towns. Either aggregated or separately, all adoption curves are clearly S-shaped, suggesting that inertia alone cannot explain the diffusion of the Crusade.

Contagion is a popular alternative type of adoption process, frequently used in the epidemiology literature. Under contagion dynamics, players adopt when others they are in touch with have adopted.[45] In contrast to an inertial model, models of contagion have S-shaped adoption curves. Because agents adopt when more agents have adopted, there must be a period where diffusion is fast, generating the steep region of the adoption curve. While other models of diffusion also generate S-shaped adoption curves, in any process driven only by contagion, however, the relative hazard rate, $\dot{p}(t)/p(t)(1-p(t))$, must be non-increasing (see Young (2009)). As a way to indirectly probe this aggregate implication, in the first column of Table 7 we report the estimates of an OLS regression of the relative hazard rate of the adoption curve for all types of events, on a fifth-order polynomial in time.[46] Similar to the result of an exercise by Young (2009) on hybrid corn adoption, we find a non-monotonic relative hazard rate. Indeed, the top-right panel of Figure 7 depicts both the relative hazard rate (in blue), and the fitted values based on the estimates from the model in column 1 of Table 7. This curve is initially decreasing but subsequently increases reaching a local maximum before starting to decrease again, easily ruling out a non-increasing relative hazard rate.[47]

Young (2009) also considers models of social influence and social learning. In a social influence model, such as the classic threshold model of Granovetter (1978), agents are heterogeneous in the threshold fraction of other agents that must have adopted before they are willing to adopt. As a result, the dynamics of these models depend closely on the distribution F of thresholds in the population. The simplest model of social influence is described by the differential equation $\dot{p}(t) = \lambda[F(p(t)) - p(t)]$. Models of social learning are varied, depending on the specific assumptions made about the informational environment and the information-processing abilities of agents. The simplest such model, where risk-neutral and myopic agents observe others' outcomes –besides others' choices–, turns out to have a structure similar to that of a social influence model. However, in this case the

[44]The simplest such inertial process is characterized by the differential equation $\dot{p}(t) = \lambda(1-p(t))$, where each instant a fraction λ of the population that has not yet adopted does so. Young (2009) demonstrates that the adoption curve will not be S-shaped even if there is heterogeneity in the λs across the population.

[45]Contagion of behaviors can be micro-founded with preferences for conformity. Young (2009) shows that a simple such model is given by the differential equation $\dot{p}(t) = (ap(t) + \lambda)(1-p(t))$. The fraction of non-adopters adopting at a given instant has an inertial component but also a component that is proportional to the fraction who already have adopted.

[46]Because the adoption curve is almost flat after around 125 days into the Crusade, we estimate this regression for the first 125 days of the Crusade only.

[47]As Young (2009) points out, this finding does not imply the absence of contagion dynamics, but it strongly suggests that contagion by itself cannot explain the diffusion of the Crusade.

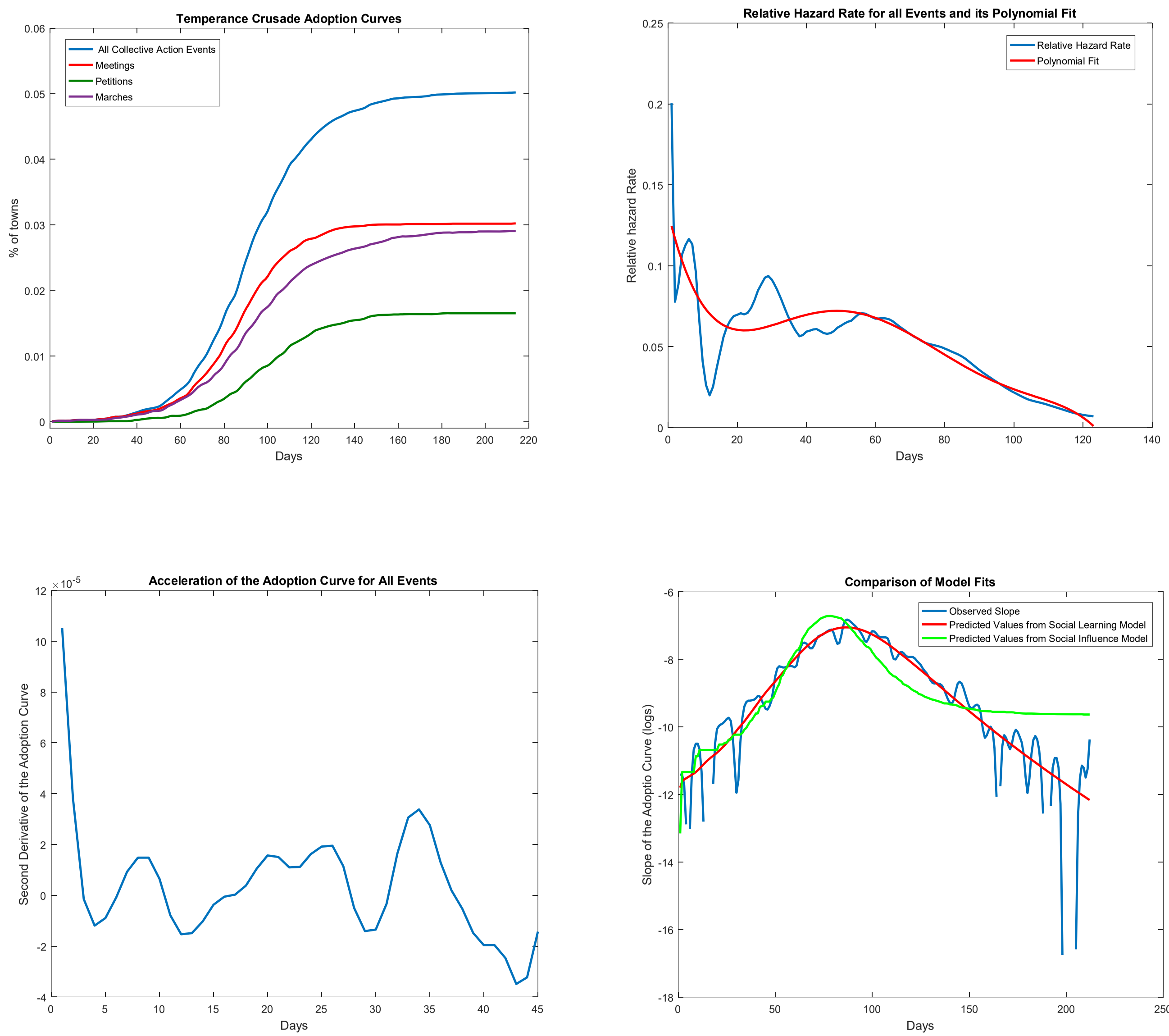

Figure 7: Tests of Alternative Protest Diffusion Signatures, based on Young (2009). The figure in the top left presents the adoption curves $p(t)$ of the Crusade. The blue line includes all types of Temperance Crusade events –meetings, petitions, and marches–. The red line includes only meetings, the green line includes only petitions, and the purple line includes only marches. The figure in the top right presents the relative hazard rate of the adoption curve of all events $\dot{p}(t)/p(t)(1-p(t))$ (in blue), and the fitted values for the relative hazard rate based on the coefficients in column (1) of Table 7 (in red). The bottom left figure presents the second derivative of the adoption curve $\ddot{p}(t)$ for the first 45 days of the Crusade. The bottom right figure presents the log slope of the adoption curve $\ln[\dot{p}(t)]$ (in blue), and the fitted values of the log slope of the adoption curve based on the coefficients in column (2) of Table 7 (in green) and on the coefficients in column (3) of Table 7 (in red).

Evaluating Alternative Protest Diffusion Signatures, based on Young (2009)

	Contagion: Relative Hazard Rate Monotonically Decreasing		Social Influence: Slope of the Adoption Curve Proportional to its Level		Social Learning: Slope of the Adoption Curve Proportional to its Integral	Schennach-Wilhelm (2016) Model Selection Test
	$\dot{p}(t)/[p(t)(1-p(t))]$ (1)		$\ln[\dot{p}(t)]$ (2)		$\ln[\dot{p}(t)]$ (3)	Ho: (2) = (3) Ha: (3) $\succ$ (2)
t	-0.0095	$\ln[p(t)]$	-174.7	$\ln[\int_0^t p(s)ds]$	-0.933	t-statistic: 6.834
	(0.0016)		(18.3)		(0.045)	p-value: 0.000
t^2	0.00039	$(\ln[p(t)])^2$	-57.9	$(\ln[\int_0^t p(s)ds])^2$	-0.744	
	(0.00007)		(6.64)		(0.034)	
t^3	-6.62E-06	$(\ln[p(t)])^3$	-9.27	$(\ln[\int_0^t p(s)ds])^3$	-0.139	
	(1.49E-06)		(1.16)		(0.020)	
t^4	4.92E-08	$(\ln[p(t)])^4$	-0.72	$(\ln[\int_0^t p(s)ds])^4$	-0.012	
	(1.30E-08)		(0.10)		(0.004)	
t^5	-1.34E-10	$(\ln[p(t)])^5$	-0.022	$(\ln[\int_0^t p(s)ds])^5$	-0.00044	
	(4.09E-11)		(0.003)		(0.00022)	
R squared	0.74		0.79		0.90	
Observations	123		177		177	

Table 7: Alternative Protest Diffusion Signatures. Column (1) presents OLS results from a regression of the relative hazard rate of the adoption curve on a fifth-order polynomial in time, between the beginning of the Crusade and day 124. Column (2) presents OLS results from a regression of the log slope of the adoption curve on a fifth-order polynomial in the log of the level of the adoption curve. Column (3) presents OLS results from a regression of the log slope of the adoption curve on a fifth-order polynomial in the log of the integral of the adoption curve. In all models, the adoption curve is based on all types of Temperance Crusade events –meetings, petitions, and marches–. Standard errors are robust to arbitrary heteroskedasticity. The last column presents the test statistic and associated p-value of the model selection test from Schennach and Wilhelm (Forthcoming), comparing the models from columns (2) and (3).

individual thresholds depend not on how many others have adopted, but on how much information has been generated by the adoption decisions of others. Young (2009) shows that the differential equation characterizing a social-learning diffusion process is given by $\dot{p}(t) = \lambda\left[F\left(\int_0^t p(s)ds\right) - p(t)\right]$.

The area under the adoption curve captures the amount of information that has been generated up to time t. It is much harder to distinguish between social influence and social learning based on the aggregate patterns of the adoption curve alone. Its shape will depend on the distribution of thresholds and on subtle features of the informational environment. When social learning is present, however, two key signatures should be observed: first, because information is scarce early on, most social learning processes should exhibit a rocky beginning with slow growth. In fact, they should exhibit *deceleration* in their early phase.[48] In Figure 1 we already illustrated the slow and bumpy start of the Crusade. In the bottom left panel of Figure 7 we reiterate this point by graphing the second derivative of the adoption curve for all events during the first 45 days of the movement. Overall, the rate of change of the slope of the adoption curve decreases in this period, and moreover, the acceleration is *negative* for around half the time span under consideration.

The second distinguishing signature of social learning emphasized by Young (2009) follows directly from the equations describing social influence and social learning: under social influence, the slope

[48] The reason for this, in Young (2009)'s words is that "... the initial block of optimists... exerts a decelerative drag on the process: they contribute at a decreasing rate as their numbers diminish, while the information generated by the new adopters gathers steam slowly because there are so few of them to begin with" (p. 1913)

of the adoption curve should be proportional to its level. Under social learning, in contrast, the slope of the adoption curve should be proportional to its integral. Taking logs of both equations, we approximate the right-hand side functions as fifth-order polynomials of either the adoption curve or its integral, and estimate them by OLS. We report the results in columns (2) and (3) of Table 7. Naturally, both polynomials fit the log slope of the adoption curve quite well, but the model based on the integrals under the adoption curve has an R squared of 0.9 compared to an R squared of only 0.79 for the model based on the levels.

We go further in the last column of the table, by performing a model selection test based on Schennach and Wilhelm (Forthcoming). This parametric test compares the fit of the models by building a t-statistic that has a normal limiting distribution centered at zero under the null hypothesis that both models are equally good at fitting the data. We easily reject the null in favor of the social learning model, with a t-statistic of 6.83 and an associated p-value of 0 to twelve decimal places.[49] The much better fit of the model in column (3) of the table can also be seen graphically. In the bottom right panel of Figure 7 we plot the log slope of the adoption curve (blue curve), together with the predicted values from the social influence model (green curve) and the social learning model (red line), using the estimated coefficients from Table 7. The picture shows the much better fit of the social learning model, despite both being polynomials of the same order. The social influence model under-predicts the slope of the adoption curve between days 100 and 150 into the Crusade, and over-predicts it after that. In contrast, the flexible polynomial in $\ln(\int_0^t p(s)ds)$ easily follows the observed rate of change of the adoption curve. Taken together we see these pieces of evidence to strongly suggest that social learning across towns was at the heart of the spread of the Temperance Crusade.[50]

6 Estimating Technological Complementarities

In this section, we describe our empirical strategy to identify the complementary roles that railroads and telegraphs played in fostering the diffusion of the Temperance Crusade. Dimensions of information such as the speed or range with which communication networks allow information transmission can have different implications over the resulting patterns of social interaction. Moreover, when several information transmission networks are in place, a natural question that arises is whether they operate as complements or substitutes for information transmission and social learning. Our findings here show that the effects of information can depend crucially on the technological features of the

[49] The Schennach and Wilhelm (Forthcoming) test requires providing a tuning parameter ε_n. We follow their advice and compute ε_n based on their suggested optimal choice.

[50] We should emphasize that the adoption models in Young (2009) are all based on the assumption that agents are matched randomly in the population. He points out that when interaction in the population is mediated by a network, the signature patterns on the aggregate adoption curve may be different because the network constrains how agents can interact. Although in our setting towns were embedded in several networks –rail and telegraph foremost–, we find it encouraging that all of the footprints from the adoption curve analysis point strongly to social learning as a key driver of protest diffusion.

different communication technologies available, and on their interaction. Our results indicate that access to the telegraph boosted the effectiveness of railroad connections.

A Cluster Event Study Approach

Were railroads and telegraphs complementary or substitute technologies for the diffusion of the Temperance Crusade? Answering this question is empirically challenging. Towns with rail or telegraph access were likely different from towns without access to these technologies, particularly along dimensions which might have made them more responsive to information or more prone to collective action. To address this difficulty, we propose a methodology resembling an event study for each of the collective action events during the Temperance Crusade, relying on the variation across towns in their rail and telegraph network connections.

For each collective action event during the Crusade (meeting, petition, or march), we take all towns falling within a given geographic radius of the town experiencing the event, and observe their collective action responses within a window of time following the event. We then compare the responsiveness of towns with different network characteristics within this geographic cluster, and average across all event studies. We can control for all unobserved time-invariant characteristics of the town because given town will fall within several event studies. We can also control for all unobserved features common to all towns in a given event study cluster because we average across many events This allows us to compare the response of towns with and without a direct rail link to the signal-generating town, and how this response varies with additional access to the telegraph network. Thus, we exclude towns without railroad access from the analysis.[51]

We construct our clusters for the event study regressions as follows: for every town $i = 1, ..., 802$ with a Crusade event –the signal-generating town–, we draw a circle of radius d from the town's centroid. Using ArcGIS, we calculate the geodesic distance between town i's centroid and all the town centroids in our census dataset. We keep all towns with centroids at a distance d or less from town i –the signal-recipient towns–.

For every signal-generating town i experiencing a Crusade event at time t, we define $G_d(i, t)$ to be the set of all signal-recipient towns j within distance d to it. We also define $F(t)$ to be the set of towns which, by time t, have not yet experienced a march. This is the subset of towns which can still hold collective action events at time t. We denote by $r_{ij} \in \{0, 1\}$ a dummy variable equal to one if signal-generating town i and signal-recipient town j have a *direct* railroad connection. We denote by $\gamma_j \in \{0, 1\}$ a dummy variable equal to one if town j has access to the telegraph network. Finally, $a_{j[t,t+\tau]}$ denotes a dummy variable equal to one if signal-recipient town j had any collective action event within the time window $[t, t + \tau]$.

For event study window size τ, and pooling across all event studies (i, t), consider the following

[51] Towns not in the 1870 railroad network were very different along most observable characteristics to towns with access to at least one communication network.

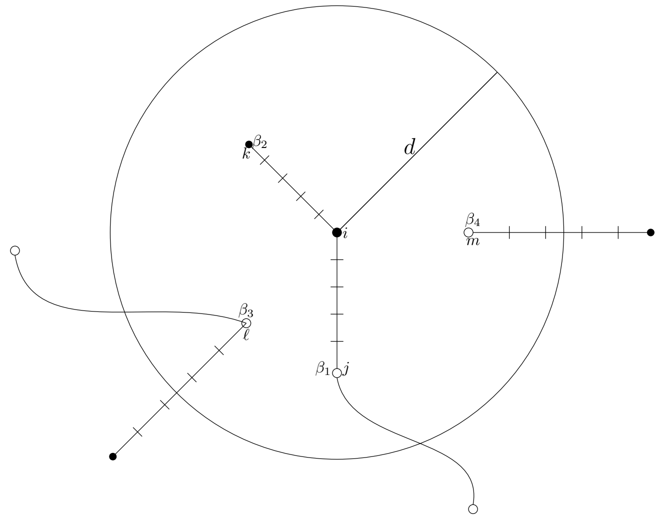

Figure 8: Identification of Technological Complementarities under the Cluster Event Study Approach: Illustration. The figure illustrates the sources of variation we exploit to identify the railroad-telegraph complementarities, based on the cluster event study approach. Within a given cluster radius d around a town i experiencing a Crusade event at time t, there exist towns j directly linked to i through the railroad and with telegraph access, towns k directly linked to i through the railroad but without telegraph access, towns ℓ with both railroad and telegraph access but not directly linked to i by rail, and towns m with railroad access but without a direct link to i, and with no telegraph access.

econometric specification:

$$a_{j[t,t+\tau]} = \beta_1 r_{ij}\gamma_j + \beta_2 r_{ij}(1-\gamma_j) + \beta_3(1-r_{ij})\gamma_j + \beta_4(1-r_{ij})(1-\gamma_j) + \rho d_{ij} + \varepsilon_{ij,t}, \tag{3}$$

for all $j \in G_d(i,t) \cap F(t)$, where d_{ij} is the geographic distance between towns i and j. Based on this specification, one could compute the following quantities of interest: i) The average effect of telegraph access among towns with rail connection: $\beta_1 - \beta_2$; ii) The average effect of telegraph access among towns without rail connection: $\beta_3 - \beta_4$; iii) The average effect of a rail connection among towns with telegraph access: $\beta_1 - \beta_3$; iv) The average effect of a rail connection among towns without telegraph access: $\beta_2 - \beta_4$; v) The differential effect of a rail connection between towns with and without telegraph access: $(\beta_1 - \beta_3) - (\beta_2 - \beta_4)$. This last effect is what we define as the technological complementarity between the rail and telegraph networks. Figure 8 illustrates graphically our empirical strategy. In practice, however, each recipient town either has a rail link to the generating town or does not, and either has telegraph access or does not: the first four regressors in equation (3) are perfectly collinear, and $(1-r_{ij})(1-\gamma_j)$ must be dropped. In this case, the network complementarity can be recovered as $\beta_1 - \beta_2 - \beta_3$.

In the context of network effects, a key confounder is the possibility of an unobserved shock that makes both towns $j, k \in G_d(i,t)$ experience collective action. We can, however, include event-study fixed effects $\delta_{(i,t)}$, comparing signal-recipient towns that vary in their network characteristics. Event-study fixed effects subsume any common shocks to all towns in $G_d(i,t)$. Furthermore, this empirical strategy is also immune to unobservables that affect the likelihood of collective action at the signal-generating town i and at the signal-recipient town j because signal-generating towns are not included in the event study defined by their collective action event.

Perhaps more importantly, heterogeneity across towns in their proclivity to collective action may also be correlated with their network characteristics. We can partially address this concern controlling for an array of town characteristics potentially relevant for collective action such as their religious heterogeneity, access to newspapers or post offices, gender ratio, or the number of liquor dealers. Even after controlling for these characteristics, other unobservables remain a concern. However, signal-recipient towns j are members of several different event study clusters $G_d(i,t)$, so we can go further and include town fixed effects ξ_j. In this way, we can control for all time-invariant town unobservables, and all time-varying cluster unobservables. Naturally, in the models where we include town fixed effects, an additional network interaction term must also be dropped. We drop $(1-r_{ij})\gamma_j$, so we recover the network complementarity as $\beta_1 - \beta_2$. Finally, a fraction of event studies straddle state boundaries, so we are also able to include state fixed effects in all our specifications.

Main Results

Table 8 presents our main results, where we do not distinguish between types of Crusade events. We fix the cluster radius at $d = 30$ Kms., but allow for three different time windows following the signal-generating event: $\tau \in \{2 \text{ weeks}, 3 \text{ weeks}, 4 \text{ weeks}\}$. We report standard errors clustered two-ways: at the event study and at the recipient-town level. The row labeled "Complementarity" reports our estimate for the network complementarities, computed as $\beta_1 - \beta_2 - \beta_3$ for the models without town fixed effects (in columns 1-6), and as $\beta_1 - \beta_2$ for the models including town fixed effects (in columns 7-9).[52]

The first three columns present results for models without recipient-town fixed effects, for 2-week, 3-week, and 4-week windows. We do not include any additional covariates besides the distance between signal-generating and signal-recipient towns. The coefficients on $r_{ij}\gamma_j$ and on $(1-r_{ij})\gamma_j$ are positive and very precisely estimated, while the effect on $r_{ij}(1-\gamma_j)$ is small and statistically insignificant. The resulting network complementarity effect is 0.137 (s.e. = 0.06) for the 3-week window model.

This estimate, however, may capture the effect of town-level characteristics correlated with network access and important for collective action. The increase in the magnitude of the complementarity coefficient as we increase the window size from 2 to 4 weeks in columns 1-3 is symptomatic of the presence of such confounders. In columns (4)-(6) we include the following town-level covariates

[52] Its standard error is computed using the full variance-covariance matrix of the vector of estimated coefficients:

$$\text{Var}(\beta_1 - \beta_2 - \beta_3) = \sum_{i=1}^{3} \text{Var}(\beta_i) - 2\text{Cov}(\beta_1, \beta_2) - 2\text{Cov}(\beta_1, \beta_3) + 2\text{Cov}(\beta_2, \beta_3)$$

or

$$\text{Var}(\beta_1 - \beta_2) = \sum_{i=1}^{2} \text{Var}(\beta_i) - 2\text{Cov}(\beta_1, \beta_2).$$

Rail and Telegraph Technological Complementarities: 30Kms Cluster Event Studies

	2 week (1)	3 week (2)	4 week (3)	2 week (4)	3 week (5)	4 week (6)	2 week (7)	3 week (8)	4 week (9)
$r_{ij}\gamma_j$	0.293	0.358	0.405	0.253	0.308	0.347	0.106	0.109	0.100
	(0.059)	(0.062)	(0.063)	(0.056)	(0.058)	(0.061)	(0.047)	(0.032)	(0.033)
$r_{ij}(1-\gamma_j)$	0.006	0.011	0.013	0.007	0.013	0.015	-0.002	0.003	0.007
	(0.007)	(0.008)	(0.009)	(0.007)	(0.008)	(0.009)	(0.005)	(0.005)	(0.004)
$(1-r_{ij})\gamma_j$	0.177	0.211	0.233	0.142	0.166	0.183			
	(0.029)	(0.032)	(0.035)	(0.028)	(0.031)	(0.034)			
Complementarity	0.110	0.137	0.159	0.104	0.130	0.149	0.108	0.106	0.093
	(0.059)	(0.060)	(0.064)	(0.058)	(0.058)	(0.062)	(0.046)	(0.032)	(0.033)
Signal-recipient distance	-0.0008	0.0021	0.0038	-0.0002	0.0026	0.0040	-0.0045	0.0004	0.0044
	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.0028)	(0.0027)	(0.0028)
Cluster FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
State FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Recipient town FE	N	N	N	N	N	N	Y	Y	Y
Mean of dep. var.	0.048	0.062	0.073	0.048	0.062	0.073	0.048	0.062	0.073
R squared	0.111	0.125	0.138	0.138	0.163	0.181	0.066	0.064	0.06
Observations	29592	29592	29592	29497	29497	29497	29592	29592	29592

Table 8: Rail and Telegraph Technological Complementarities: Cluster Event Studies. The table presents estimation results of the cluster event study approach based on equation (3), using the benchmark 30 Km. radius clusters. The dependent variable is a dummy variable for whether a town within the cluster radius experienced a Crusade event within the time window in each column header following the cluster-defining town experiencing its Crusade event. All models include event-cluster fixed effects, state fixed effects, and the distance between generating and recipient towns. Columns (4)-(6) include the following set of controls: native share, black share, female to male ratio, newspapers per capita, post office dummy, religious ascriptions Herfindahl index, Presbyterian share, and log population. Columns (7)-(9) include recipient-town fixed effects. In columns (1)-(6) the complementarity interaction effects are computed as the difference between the coefficients on $r_{ij}\gamma_j$, $r_{ij}(1-\gamma_j)$, and $(1-r_{ij})\gamma_j$. In columns (7)-(9) the complementarity interaction effects are computed as the difference between the coefficients on $r_{ij}\gamma_j$ and $r_{ij}(1-\gamma_j)$. Standard errors are robust and clustered two-ways, at the event-cluster and at the recipient town levels.

to address this concern: the native-born share of the population, the black share of the population, and the female to male gender ratio, the number of newspapers per capita in circulation, a post office dummy, the religious heterogeneity Herfindahl index, the share of Presbyterians, and log population. The inclusion of these controls reduces the magnitude of the interaction effects, but the estimated complementarity effect is almost unaffected in its magnitude and precision. These estimates, however, still grow in magnitude as we look at event studies with longer time windows. In the last three columns, we move on to models including recipient-town fixed effects. We now find the coefficient on the $r_{ij}\gamma_j$ interaction shrinks considerably, from around 0.3 in the models with covariates to around 0.1, suggesting the importance of omitted unobservables. This coefficient is, however, very precisely estimated, and leads to a similarly precisely estimated network complementarity effect of 0.1 irrespective of the time window we use. In the models from columns 1-6, the coefficients on the complementarity effect become larger as we study longer time windows. This is no longer the case for the models in columns 7-9. We see this as strong evidence that the simultaneous inclusion of cluster and town-level fixed effects is sufficient for identifying the rail-telegraph complementarity. The combination of a direct rail link and telegraph access increases the likelihood of collective action by 10 percentage points relative to having access to just the rail link or to telegraph access. This is 1.6 times the mean of the dependent variable (0.062), a quantitatively large effect.

Robustness

We conclude this section with some additional robustness exercises that strengthen the validity of these results. In Appendix Table B.1 we present models similar to those in Table 8, for alternative cluster radii. All these models include recipient-town fixed effects. We continue to find positive and significant network complementarities. As we would expect in a network setting (where distance imposes frictions on information flows), the magnitude of the effect decreases as we increase the radius of the clusters: from around 0.8 for the 50Kms. specifications to around 0.5 for the 120 Kms. specifications. Similar to the baseline 30 Kms. results, different time windows for the event studies make no difference to the estimated magnitudes.

In Appendix Table B.2, we then test for evidence of heterogeneity in the effects of these network complementarities. The table reports estimates for different cluster radii (30 and 50 Kms.), but fixing a 2-week event study time window. We include interactions between each of the network interaction variables and the number of newspapers per capita (columns (1) and (2)), the post office dummy (columns (3)-(4)), the religious heterogeneity index (columns (5) and (6)), and the gender ratio (columns (7) and (8)). Across seven of the eight specifications, we find the network complementarity to remain stable around 0.1, and no evidence of any significant heterogeneity.

In Table 9 we present results of a placebo test on the event study methodology, to address the possibility of residual time persistent unobservables. Instead of a dummy for a Crusade event in the signal-recipient town in the weeks following the Crusade event in the signal-generating town, in this exercise the dependent variable is a dummy for a Crusade event in the signal-recipient town

Rail and Telegraph Technological Complementarities: Placebo Event Studies using Previous Weeks Responses

	30 KMS				50 KMS		80 KMS		120 KMS	
	2 weeks (1)	4 weeks (2)	2 weeks (3)	4 weeks (4)	2 weeks (5)	4 weeks (6)	2 weeks (7)	4 weeks (8)	2 weeks (9)	4 weeks (10)
$r_{ij}\gamma_j$	0.076	0.150	-0.052	-0.049	-0.001	0.017	-0.004	0.017	-0.001	0.021
	(0.023)	(0.038)	(0.031)	(0.032)	(0.022)	(0.024)	(0.018)	(0.021)	(0.017)	(0.019)
$r_{ij}(1-\gamma_j)$	0.0016	0.0063	0.0032	0.0063	0.0007	0.0045	-0.0014	0.0042	-0.0003	0.0048
	(0.005)	(0.007)	(0.004)	(0.004)	(0.003)	(0.003)	(0.0021)	(0.0026)	(0.0019)	(0.0024)
$(1-r_{ij})\gamma_j$	0.106	0.162								
	(0.018)	(0.024)								
Complementarity	-0.032	-0.018	-0.055	-0.055	-0.001	0.013	-0.002	0.012	-0.0005	0.016
	(0.026)	(0.041)	(0.031)	(0.032)	(0.022)	(0.024)	(0.018)	(0.020)	(0.017)	(0.019)
Signal-recipient distance	0.0005	0.0035	-0.0012	0.0008	-0.0038	-0.0030	-0.0029	-0.0021	-0.0021	-0.0014
	(0.003)	(0.004)	(0.002)	(0.003)	(0.001)	(0.001)	(0.0007)	(0.0008)	(0.0005)	(0.0005)
Cluster FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
State FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Recipient town FE	N	N	Y	Y	Y	Y	Y	Y	Y	Y
Mean of dep. var.	0.023	0.036	0.023	0.036	0.020	0.033	0.01	0.032	0.018	0.030
R squared	0.075	0.101	0.042	0.062	0.017	0.026	0.019	0.017	0.007	0.014
Observations	30557	30557	30557	30557	81858	81858	200603	200603	421712	421712

Table 9: Rail and Telegraph Technological Complementarities: Placebo Event Studies using Previous Weeks Responses. The table presents estimation results of the cluster event study approach based on equation (3), where the dependent variable is a dummy variable for whether a town within the cluster radius experienced a Crusade event within the time window in each column header *prior* to the cluster-defining town experiencing its Crusade event. Columns (1)-(4) use 30 Km. radius clusters. Columns (5)-(6) use 50 Km. radius clusters. Columns (7)-(8) use 80 Km. radius clusters. Columns (9)-(10) use 120 Km. radius clusters. All models include event-cluster fixed effects, state fixed effects, recipient-town fixed effects, and the distance between generating and recipient towns. In columns (1)-(2) the complementarity interaction effects are computed as the difference between the coefficients on $r_{ij}\gamma_j$, $r_{ij}(1-\gamma_j)$, and $(1-r_{ij})\gamma_j$. In columns (3)-(10) the complementarity interaction effects are computed as the difference between the coefficients on $r_{ij}\gamma_j$ and $r_{ij}(1-\gamma_j)$. Standard errors are robust and clustered two-ways, at the event-cluster and at the recipient town levels.

in the weeks *prior* to the Crusade event in the signal-generating town. The table presents results for different cluster radii definitions and different time windows, with and without town-level fixed effects. We find no statistically significant network complementarity estimates, with both negative and positive point estimates across different specifications.

Finally, in a second placebo exercise we address the possibility of unobserved similarities between the signal-generating town and the signal-recipient towns in its cluster. We create false clusters by replacing each true signal-generating town i for its closest match k, using a matching algorithm based on covariate similarity between towns.[53] We then estimate the response of town i's signal-recipient towns $j \in G_d(i,t)$, to the Crusade event of town k, which generically took place on a different date than i's. We report these results in Appendix Table B.3, for different cluster radii and event study time windows. Once again, we find no systematic pattern of signs for the estimated network complementarities, and all but one of the coefficients across specifications is statistically significant at the 5 percent level. Taken together, these results indicate that complementarities between the railroad and telegraph communication networks were an important channel of social interactions in the diffusion of the Temperance Crusade.

7 Concluding Remarks

We study how communication networks mediate social interactions leading to the geographic spread of protest activity. We do so in the context of the Temperance Crusade, the first movement of mass collective action by women in the 19th Century, and focus on the two main communication networks of the time: railroads and telegraphs. Using exogenous variation in rail network links induced by railroad worker strikes and railroad accidents, we use a linear model of social interactions to estimate the causal effect of railroad-mediated and telegraph-mediated information flows about Crusade activity on the Crusade activity of neighboring towns. We find evidence of large effects, which can account for the spatial diffusion of the protest movement. We also provide evidence consistent with social interactions based on these information flows, and on the importance of newspapers for the diffusion of the Crusade. We then propose an event study methodology allowing us to identify large complementarities between the railroad and telegraph networks in the responsiveness of crusading women to information about neighboring protest activity. We also provide evidence of social learning driving the social interaction effects we estimate, both by studying the aggregate patterns of the diffusion process and the heterogeneous rates of protest adoption in response to varying types of neighboring Crusade activities. Our findings confirm the importance of group heterogeneity as a limiting factor in successful collective action, the importance of communication networks as drivers of protest diffusion when social interactions are important, and the key role that organizational stages can have in fostering protest movements.

[53] We use the Mahalanobis distance metric to find the closest matches, using the native-born population share, the black share, the gender ratio, the number of newspapers per capita, the number of alcohol vendors per capita, the religious Herfindahl index, the number of Presbyterian sittings per capita, and log population.

Our paper is the first observational study that relies on exogenous variation in network links to identify social interaction effects. It is also a first attempt at establishing the complementary roles of competing communication networks. Taken together, our results highlight that collective action, specifically in the context of protest activity, is shaped strongly by network effects. It also highlights that the type of information technologies available, their network structure, and their interaction, are first-order mediators of social interactions. Our results also suggest that future research on collective action should study the role of organizational meetings and how they lead to information aggregation and coordination, as these appear to have been of key importance for the spread of the Crusade. We hope these results also encourage further research on the role of competing networks in shaping the quantity and quality of information relevant for political mobilization, public good provision, and other forms of collective action, particularly in contemporary settings where online networks co-exist with more traditional communication technologies.

References

Acemoglu, D., M. Dahleh, I. Lobel, and A. Ozdaglar (2011): "Bayesian Learning in Social Networks," *The Review of Economic Studies*, 78, 1201–1236.

Acemoglu, D., C. García-Jimeno, and J. A. Robinson (2015): "State Capacity and Economic Development: A Network Approach," *The American Economic Review*, 105, 2364–2409.

Acemoglu, D., T. Hassan, and A. Tahoun (2014): "The Power of the Street: Evidence from Egypt's Arab Spring," NBER Working Paper no. 20665.

Aidt, T., G. Leon, and M. Satchell (2017): "The Social Dynamics of Collective Action: Evidence from the Captain Swing Riots, 1830-31," Unpublished.

Andrews, D. and B. Lu (2001): "Consistent Model and Moment Selection Procedures for GMM estimation with Application to Dynamic Panel Data Models," *Journal of Econometrics*, 101, 123–164.

Andrews, K. and M. Biggs (2006): "The Dynamics of Protest Diffusion: Movement Organizations, Social Networks, and News Media in the 1960 Sit-Ins," *American Sociological Review*, 71, 752–77.

Angrist, J. and J.-S. Pischke (2008): *Mostly Harmless Econometrics*, Princeton, NJ: Princeton University Press.

Aral, S., L. Muchnik, and A. Sundararajan (2009): "Distinguishing Influence-based Contagion from Homophily-driven Diffusion in Dynamic Networks," *Proceedings of the National Academy of Sciences*, 106, 21544–21549.

Aral, S. and D. Walker (2012): "Identifying Influential and Susceptible Members of Social Networks," *Science*, 337, 337–341.

Atack, J. (2013): "On the Use of Geographic Information Systems in Economic History: The American Transportation Revolution Revisited," *The Journal of Economic History*, 73, 313–338.

Bakshy, E., I. Rosenn, C. Marlow, and L. Adamic (2012): "The Role of Social Networks in Information Diffusion," *Proceedings of the 21st International World Wide Web Conference*, 519–528.

Bala, V. and S. Goyal (1998): "Learning from Neighbors," *The Review of Economic Studies*, 65, 595–621.

Bandiera, O. and I. Rasul (2006): "Social Networks and Technology Adoption in Northern Mozambique," *The Economic Journal*, 116, 869–902.

Banerjee, A. (1992): "A Simple Model of Herd Behavior," *The Quarterly Journal of Economics*, 107, 797–817.

Banerjee, A., A. Chandrasekhar, E. Duflo, and M. Jackson (2013): "The Diffusion of Microfinance," *Science*, 341, 363–370.

Biggs, M. (2003): "Positive Feedback in Collective Mobilization: The American Strike Wave of 1886," *Theory and Society*, 32, 217–254.

Bikhchandani, S., D. Hirshleifer, and I. Welch (1992): "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *The Journal of Political Economy*, 100, 992–1026.

Blocker, J. (1985): *Give to the Winds Thy Fears*, Santa Barbara, CA: Praeger.

——— (1989): *American Temperance Movements. Cycles of Reform*, Boston, MA: Twayne Publishers.

Bolton, S. K. (1944): "Sarah Knowles Bolton Papers, 1881-1944. Autobiography," Schlesinger Library, Radcliffe Institute, Harvard University.

Bordin, R. (1981): *Woman and Temperance: The Quest for Power and Liberty, 1873-1900*, Philadelphia, PA: Temple University Press.

Breza, E. (Forthcoming): *Field Experiments, Social Networks, and Development*, Oxford University Press, vol. Oxford Handbook on the Economics of Networks, chap. TBD.

Burke, M., G. Fournier, and K. Prasad (2007): "The Diffusion of a Medical Innovation: Is Success in the Stars?" *Southern Economic Journal*, 73, 558–603.

Cantoni, D., D. Yang, N. Yuchtman, and Y. J. Zhang (2017): "Are Protests Games of Strategic Complements or Substitutes? Experimental Evidence from Hing Kong's Democracy Movement," NBER Working Paper no. 23110.

Chandrasekhar, A. and R. Lewis (2011): "Econometrics of Sampled Networks," MIT Working Paper.

Cherrington, E. H. (1920): *The Evolution of Prohibition in the United States of America*, Westerville, OH: The American Issue Press.

Christensen, D. and F. Garfias (2015): "Can You Hear me Now? How Communication Technology Affects Protest and Repression," Stanford University.

Conley, T. (1999): "GMM estimation with Cross-Sectional Dependence," *Journal of Econometrics*, 92, 1–45.

Conley, T. and C. Udry (2010): "Learning about a New Technology: Pineapple in Ghana," *American Economic Review*, 100, 35–69.

Cooney, R. (2005): *Winning the Vote: The Triumph of the American Woman Suffrage Movement*, Half Moon Bay, CA: American Graphic Press.

Dagaev, D., N. Lamberova, A. Sobolev, and K. Sonin (2013): "Technological Foundations of Political Instability," CEPR Working paper no. 9787.

Degler, C. (1981): *At Odds: Women and the Family in America from the Revolution to the Present*, Oxford, UK: Oxford University Press.

DeGroot, M. H. (1974): "Reaching a Consensus," *Journal of the American Statistical Association*, 69, 118–121.

DeTocqueville, A. (1835): *Democracy in America*, Suffolk, England: Penguin Books, 2003 ed.

Donaldson, D. and R. Hornbeck (2016): "Railroads and American Economic Growth: A Market Access Approach," *The Quarterly Journal of Economics*, 131, 799–858.

Duflo, E. and E. Saez (2003): "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment," *Quarterly Journal of Economics*, 118, 815–842.

Dupas, P. (2014): "Short-run Subsidies and Long-run Adoption of New Health Products: Evidence from a Field Experiment," *Econometrica*, 82, 197–228.

Ellison, G. and D. Fudenberg (1993): "Rules of Thumb for Social Learning," *The Journal of Political Economy*, 101, 612–643.

——— (1995): "Word-of-Mouth Communication and Social Learning," *The Quarterly Journal of Economics*, 110, 93–125.

Enikolopov, R., A. Makarin, and M. Petrova (2016): "Social Media and Protest Participation: Evidence from Russia," .

Epstein, B. (1981): *The Politics of Domesticity: Women, Evangelism, and Temperance in Nineteenth-Century America*, Middletown, CT: Wesleyan University Press.

Feigenbaum, J. and M. Rotemberg (2015): "Communication and Manufacturing: Evidence from the Expansion of Postal Services," Harvard University.

Foster, A. and M. Rosenzweig (1995): "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture," *Journal of Political Economy*, 103, 1176–1209.

García-Jimeno, C. (2016): "The Political Economy of Moral Conflict: An Empirical Study of Learning and Law Enforcement Under Prohibition," *Econometrica*, 84, 511–570.

Gentzkow, M., J. Shapiro, and M. Sinkinson (2011): "The Effect of Newspaper Entry and Exit on Electoral Politics," *The American Economic Review*, 101, 2980–3018.

Golub, B. and M. Jackson (2010): "Naive Learning in Social Networks and the Wisdom of Crowds," *American Economic Journal: Microeconomics*, 2, 112–149.

González, F. (2016): "Collective Action in Network: Evidence from the Chilean Student Movement," Unpublished.

Gould, R. (1991): "Multiple Networks and Mobilization in the Paris Commune, 1871," *American Sociological Review*, 56, 716–729.

Granovetter, M. (1973): "The Strength of Weak Ties," *American Journal of Sociology*, 78, 1360–1380.

——— (1978): "Threshold Models of Collective Behavior," *American Journal of Sociology*, 83, 1420–1443.

Griliches, Z. (1957): "Hybrid Corn: An Exploration in the Economics of Technological Change," *Econometrica*, 25, 501–522.

Gruhl, D., R. Guha, D. Liben-Nowell, and A. Tomkins (2004): "Information Diffusion Through Blogspace," *Proceedings of the 13th International World Wide Web Conference*, 491–501.

Guarino, A. and P. Jehiel (2013): "Social Learning with Coarse Information," *American Economic Journal: Microeconomics*, 5, 147–174.

Gusfield, J. (1955): "Social Structure and Moral Reform: A Study of the Women's Christian Temperance Union," *American Journal of Sociology*, 61, 221–232.

Gutman, H. (1961): "Trouble on the Railroads in 1873-1874: Prelude to the 1877 Crisis?" *Labor History*, 2, 215–235.

Halberstam, Y. and B. Knight (2016): "Homopholy, Group Size, and the Diffusion of Political Information in Social Networks: Evidence from Twitter," *Journal of Public Economics*, 143, 73–88.

Hassanpour, N. (2014): "Media Disruption and Revolutionary Unrest: Evidence from Mubarak's Quasi-Experiment," *Political Communication*, 31, 1–24.

Hedstrom, P. (1994): "Contagious Collectivities: On the Spatial Diffusion of Swedish Trade Unions, 1890-1940," *American Journal of Sociology*, 99, 1157–1179.

Henkel, J. and S. Maurer (2010): "Network Effects in Biology," *American Economic Review: Papers and Proceedings*, 100.

Kelly, M. and C. OGrada (2000): "Market Contagion: Evidence from the Panics of 1854 and 1857," *The American Economic Review*, 90, 1110–1124.

Knight, B. and N. Schiff (2010): "Momentum and Social Learning in Presidential Primaries," *Journal of Political Economy*, 118, 1110–1150.

Koudjis, P. (2016): "The Boats that Did Not Sail: Asset Price Volatility in a Natural Experiment," *The Journal of Finance*, 71, 1185–1226.

Kremer, M. and T. Miguel (2007): "The Illusion of Sustainability," *Quarterly Journal of Economics*, 122, 1007–1065.

Lee, S. E. D. (1980): *Evangelical Domesticity: The Origins of the Woman's National Christian Temperance Union Under Frances E. Wilard*, Chicago, IL: Northwestern University Press.

Lerman, K. and R. Ghosh (2010): "Social Contagion: An Empirical Study of Information Spread on Digg and Twitter Follower Graphs," ArXiv.

Madestam, A., D. Shoag, S. Veuger, and D. Yanagizawa-Drott (2013): "Do Political Protests Matter? Evidence from the Tea Party Movement," *The Quarterly Journal of Economics*, 128, 1633–1685.

Manski, C. (1993): "Identification of Endogenous Social Effects: The Reflection Problem," *The Review of Economic Studies*, 60, 531–542.

Mobius, M., T. Phan, and A. Szeidl (2015): "Treasure Hunt: Social Learning in the Field," Unpublished.

Moretti, E. (2011): "Social Learning and Peer Effects in Consumption: Evidence from Movie Sales," *The Review of Economic Studies*, 78, 356–393.

Morris, A. (1984): *The Origins of the Civil Rights Movement*, New York: Free Press.

Mossel, E., A. Sly, and O. Tamuz (2015): "Strategic Learning and the Topology of Social Networks," ArXiv:1209.5527v2.

Munshi, K. and J. Myaux (2006): "Social Norms and the Fertility Transition," *Journal of Development Economics*, 80, 1–38.

Olson, M. (1965): *The Logic of Collective Action*, Cambridge, MA: Harvard University Press.

Opp, K.-D. and C. Gern (1993): "Dissident Groups, Personal Networks, and Spontaneous Cooperation: The East German Revolution of 1989," *American Sociological Review*, 58, 659–680.

Oster, E. and R. Thornton (2012): "Determinants of Technology Adoption: Peer Effects in Menstrual Cup Take-up," *Journal of the European Economic Association*, 10, 1263–1293.

Passarelli, F. and G. Tabellini (2017): "Emotions and Political Unrest," *Journal of Political Economy*, 125, 903–946.

Pierskalla, J. and F. Hollenbach (2013): "Technology and Collective Action: The Effect of Cell Phone Coverage on Political Violence in Africa," *American Political Science Review*, 107, 207–224.

Schennach, S. and D. Wilhelm (Forthcoming): "A Simple Parametric Model Selection Test," *Journal of the American Statistical Association*.

Schneider, D. and C. Schneider (1993): *American Women in the Progressive Era*, New York, NY: Facts on File, Inc.

Schwarzlose, R. (1990): *The Nation's Newsbrokers*, vol. 2, Evanston, IL: Northwestern University Press.

Smith, L. and P. Sorenson (2000): "Pathological Outcomes of Observational Learning," *Econometrica*, 68, 371–398.

Stewart, M. (1890): *Memories of the Crusade. A Thrilling Account of the Great Uprising of the Women of Ohio in 1873, Against the Liquor Crime*, Chicago, IL.: H. J. Smith and Co.

Stover, J. F. (1999): *The Routledge Historical Atlas of the American Railroads*, New York, NY: Routledge.

Swindler, W. (1946): "The AP Antitrust Case in Historical Perspective," *Journalism Quarterly*, 23, 40–57.

Thompson, R. L. (1947): *Wiring a Continent: The History of the Telegraph Industry in the United States, 1832-1866*, Princeton, NJ: Princeton University Press.

Tilly, C. (1978): *From Mobilization to Revolution*, New York: Random House.

Trotsky, L. (1932): *History of the Russian Revolution*, vol. 1, Chicago: Haymarket Books, 2008 ed.

Tufekci, Z. and C. Wilson (2012): "Social Media and the Decision to Participate in Political Protest: Observations from Tahrir Square," *Journal of Communication*, 62, 363–379.

Tyrrell, I. (1979): *Sobering Up: From Temperance to Prohibition in Antebellum America, 1800-1860*, Westport, CT: Grenwood Press.

USPS (2012): *The United States Postal Service. An American History, 1775-2006*, Washington, DC: Government Relations, United States Postal Service.

Vernon, E. (1870): *Travelers' Official Railway Guide for the United States and Canada*, Philadelphia, PA: The Leisenring Printing House.

WesternUnion (1874): *Western Union Telegraph Directory*, New York: The Western Union Telegraph Company.

Wittenmyer, A. (1878): *History of the Woman's Temperance Crusade*, Boston, MA: W. F. Brown and Co.

Wright, S. and M. Forney (1873-4): "Railroad Gazette," in *Railroad Gazette*, vol. 5-6.

Yanagizawa-Drott, D. (2014): "Propaganda and Conflict: Evidence from the Rwandan Genocide," Harvard University.

Young, P. (2009): "Innovation Diffusion in Heterogeneous Populations: Contagion, Social Influence, and Social Learning," *The American Economic Review*, 99, 1899–1924.

A Online Appendix A: Reduced Form Additional Results

Causal Effects of Crusade Signals along the Railroad and Telegraph Networks:
Lag Specification Model Selection First Stages

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
F-statistic	62.3	53.51	48.11	40.21	39.64	33.67	34.01	54.89	47.91	26.33
R squared	0.021	0.021	0.019	0.020	0.019	0.020	0.021	0.020	0.019	0.022
F-statistic	29.89	26.56	18.93	43.45	20.02	37.71	35.67	19.42	27.2	28.44
R squared	0.013	0.016	0.014	0.025	0.023	0.025	0.026	0.014	0.016	0.027
F-statistic	1054.2	464.6	307.37	14.28	23.4	15.86	18.92	330.09	462.66	28.27
R squared	0.292	0.265	0.254	0.014	0.030	0.023	0.024	0.255	0.255	0.028
F-statistic				227.51	311.42	17.77	19.22			13.64
R squared				0.255	0.255	0.030	0.030			0.031
F-statistic						239.56	742.14			14.01
R squared						0.255	0.295			0.038
F-statistic							467.89			14.75
R squared							0.275			0.047
F-statistic										884.92
R squared										0.301
F-statistic										500.19
R squared										0.274
F-statistic										346.18
R squared										0.265

Table A.1: The Effect of Information along the Rail and Telegraph Networks: Lag Specification Model Selection First Stages. The table presents the first stage R-squared and F-statistics corresponding to each column of the 2SLS models reported in Table 3. The statistics for each first stage, from top to bottom, are reported in the same order as the endogenous regressors appear in Table 3. Following Angrist and Pischke (2008), the F-statistics are corrected for the presence of multiple endogenous regressors.

First Stages for Optimally Selected Lag Structure Model

Instrument variation:	50km accident radius			80km accident radius			120km accident radius		
Dependent Variable:	$\mathbf{r}_{i,t}\mathbf{a}_{t-2}$	$\gamma_i\mathbf{a}_{t-1}$	$\mathbf{d}_i\mathbf{a}_{t-3}$	$\mathbf{r}_{i,t}\mathbf{a}_{t-2}$	$\gamma_i\mathbf{a}_{t-1}$	$\mathbf{d}_i\mathbf{a}_{t-3}$	$\mathbf{r}_{i,t}\mathbf{a}_{t-2}$	$\gamma_i\mathbf{a}_{t-1}$	$\mathbf{d}_i\mathbf{a}_{t-3}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\mathbf{r}_{i,t-1}\boldsymbol{\iota}$	0.029	-2.28E-5	-0.113	0.0159	-1.45E-4	-0.136	0.011	1.24E-5	-0.114
	(0.004)	(4.42E-4)	(0.015)	(0.0016)	(2.70E-4)	(0.019)	(0.001)	(2.19E-4)	(0.016)
$\mathbf{r}_{i,t-1}\mathbf{R}_{t-2}\boldsymbol{\iota}$	-0.0024	-9.12E-5	0.025	-0.0002	-6.53E-5	0.031	-0.0001	0.0000	0.023
	(0.0007)	(6.10E-5)	(0.002)	(0.0002)	(3.45E-5)	(0.003)	(0.0002)	(2.65E-5)	(0.002)
$\mathbf{r}_{i,t-1}\mathbf{R}_{t-2}\mathbf{R}_{t-3}\boldsymbol{\iota}$	-8.99E-5	2.08E-6	-6.50E-4	-0.0001	7.64E-7	-0.0008	-0.0001	0.0000	-0.0006
	(1.67E-5)	(1.58E-6)	(5.49E-5)	(6.08E-6)	(1.00E-6)	(0.0001)	(4.98E-6)	(8.25E-7)	(5.95E-5)
$\gamma_i\mathbf{R}_{t-1}\boldsymbol{\iota}$	-0.0014	7.59E-3	-0.154	-0.0023	0.0015	-0.137	-0.0017	-4.85E-4	-0.082
	(0.0015)	(1.62E-3)	(0.011)	(0.0011)	(0.0002)	(0.013)	(0.0008)	(0.0001)	(0.010)
$\gamma_i\mathbf{R}_{t-1}\mathbf{R}_{t-2}\boldsymbol{\iota}$	1.99E-4	-2.18E-4	0.017	0.0002	1.24E-4	0.0153	1.74E-4	3.47E-4	0.010
	(1.57E-4)	(1.81E-4)	(0.001)	(0.0001)	(1.55E-5)	(0.0011)	(7.69E-5)	(1.27E-5)	(0.0009)
$\mathbf{d}_i\mathbf{R}_{t-3}\boldsymbol{\iota}$	7.35E-5	3.53E-6	-0.0087	6.34E-6	-2.82E-6	-0.0101	-2.93E-5	-6.09E-6	-0.008
	(2.95E-5)	(3.81E-6)	(0.0003)	(2.07E-5)	(3.42E-6)	(0.0002)	(1.51E-5)	(2.51E-6)	(0.0002)
$\mathbf{d}_i\mathbf{R}_{t-3}\mathbf{R}_{t-4}\boldsymbol{\iota}$	-3.95E-6	-3.75E-7	1.84E-4	-8.06E-7	-2.35E-7	0.0003	1.50E-6	-3.11E-8	2.05E-4
	(1.76E-6)	(2.17E-7)	(1.42E-5)	(1.18E-6)	(1.95E-7)	(1.40E-5)	(9.02E-7)	(1.50E-7)	(1.08E-5)
F statistic	48.1	18.9	307.4	46.46	11.9	683.6	33.07	14.27	680.0
R squared	0.019	0.014	0.254	0.018	0.012	0.258	0.014	0.019	0.259
No. of towns	15934	15934	15934	15934	15934	15934	15934	15934	15934
Max. no. of periods	16	16	16	16	16	16	16	16	16
Observations	267247	267247	267247	267247	267247	267247	267247	267247	267247

Table A.2: First Stages for Optimal Lag Structure Models. The table presents the first stage coefficient estimates and standard errors for the preferred lag specification. All models include town fixed effects and period fixed effects, and use a 5-day period definition. Columns (1)-(3) present the first stages corresponding to the three endogenous regressors of column (3) in Table 3, with instruments based on a 50 Km. radius for the railroad accidents. Columns (4)-(6) present the first stages corresponding to the three endogenous regressors of column (1) in Table 5, with instruments based on an 80 Km. radius for the railroad accidents. Columns (7)-(9) present the first stages corresponding to the three endogenous regressors of column (2) in Table 5, with instruments based on a 120 Km. radius for the railroad accidents. The dependent variable in columns (1), (4), and (7) is the second order lag of railroad neighbors' Crusade events. The dependent variable in columns (2), (5), and (8) is the first order lag of telegraph neighbors' Crusade events. The dependent variable in columns (3), (6), and (9) is the zeroth order lag of geographic neighbors' Crusade events. Following Angrist and Pischke (2008), the F-statistics are corrected for the presence of multiple endogenous regressors.

Causal Effects of Crusade Signals along the
Railroad and Telegraph Networks: OLS and GMM

	Dependent Variable: Any Crusade Activity a_{it} -Meetings, Petitions, Marches-			
Period definition:	5 days		3 days	
	OLS	GMM	OLS	GMM
	(1)	(2)	(3)	(4)
$\mathbf{r}_{i,t-1}\mathbf{a}_{t-2}$	0.0053	0.019	0.0024	0.023
	(0.0010)	(0.0107)	(0.0007)	(0.012)
$\boldsymbol{\gamma}_i\mathbf{a}_{t-1}$	0.014	0.089	0.009	0.127
	(0.0059)	(0.0430)	(0.0045)	(0.069)
$\mathbf{d}_i\mathbf{a}_{t-3}$	0.001	0.0026	0.0010	-0.0003
	(0.0001)	(0.0011)	(0.0001)	(0.0006)
No. of towns	15934	15934	15950	15950
Max. no. of periods	16	16	30	30
Observations	267247	267247	487548	487548

Table A.3: The Effect of Information along the Rail and Telegraph Networks: OLS and GMM Estimates. The table presents panel estimates of equation 2 for the optimally chosen lag structure model (second order lag for the railroad neighbors' Crusade events, first order lag for the telegraph neighbors' Crusade events, and third order lag for the geographic neighbors' Crusade events). The dependent variable is an indicator of crusading activity -meetings, petitions, or marches-. All models include period fixed effects and town fixed effects. Standard errors are robust and clustered at the town level. In columns (1)-(2) a period is defined as a 5 day interval. In columns (3)-(4) a period is defined as a 3 day interval. Columns (1) and (3) report OLS estimates. Columns (2) and (4) report GMM estimates based on moment conditions using the same set of instruments of column (3) in Table 3. Instruments are based on a 50 Km. radius for the railroad accidents.

Causal Effects of Crusade Signals along the Railroad and Telegraph Networks: Lag Specification Model Selection

	Dependent Variable: Any Crusade Activity a_{it} -Meetings, Petitions, Marches-									
Second stages:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\mathbf{r}_{i,t}\mathbf{a}_{t-1}$	0.0253 [0.0065] (0.0059)			-0.0006 [0.0132] (0.0124)		0.0022 [0.0132] (0.0123)	0.0041 [0.0134] (0.0125)			0.0168 [0.0119] (0.0109)
$\mathbf{r}_{i,t-1}\mathbf{a}_{t-2}$		0.0261 [0.0064] (0.0059)	0.0232 [0.0064] (0.0061)	0.024 [0.0137] (0.0129)	0.0225 [0.0063] (0.0061)	0.0206 [0.0136] (0.0127)	0.0223 [0.0138] (0.0128)		0.0235 [0.0064] (0.0061)	0.0045 [0.0159] (0.0141)
$\mathbf{r}_{i,t-2}\mathbf{a}_{t-3}$								0.0229 [0.0066] (0.0060)		0.0003 [0.01] (0.0091)
$\boldsymbol{\gamma}_i\mathbf{a}_{t-1}$	0.0787 [0.0726] (0.0625)		0.130 [0.0575] (0.0516)	0.1349 [0.0580] (0.0521)	0.082 [0.043] (0.0386)	0.0833 [0.0432] (0.0388)	0.0813 [0.048] (0.0425)	0.12 [0.0573] (0.0514)		-0.0726 [0.0647] (0.0602)
$\boldsymbol{\gamma}_i\mathbf{a}_{t-2}$		0.0772 [0.0712] (0.0637)			0.0029 [0.0407] (0.0368)	0.0034 [0.0408] (0.0369)	-0.0065 [0.0399] (0.0361)			-0.0989 [0.0567] (0.0479)
$\boldsymbol{\gamma}_i\mathbf{a}_{t-3}$									0.0815 [0.0626] (0.0581)	0.1672 [0.0732] (0.0655)
$\mathbf{d}_i\mathbf{a}_{t-1}$	0.0011 [0.0008] (0.0007)						0.0013 [0.0012] (0.0011)			0.0000 [0.0012] (0.0010)
$\mathbf{d}_i\mathbf{a}_{t-2}$		0.0004 [0.0005] (0.0005)					-0.0009 [0.0012] (0.0010)			0.0005 [0.0017] (0.0015)
$\mathbf{d}_i\mathbf{a}_{t-3}$			-0.0002 [0.0006] (0.0005)	-0.0002 [0.0006] (0.0005)	-0.0002 [0.0006] (0.0005)	-0.0002 [0.0006] (0.0005)		0.0000 [0.0006] (0.0005)	-0.0002 [0.0006] (0.0005)	-0.0002 [0.001] (0.0009)
No. of towns	15969	15958	15950	15950	15950	15950	15958	15950	15950	15950
Max. no. of periods	32	31	30	30	30	30	31	30	30	30
Observations	519475	503506	487548	487548	487548	487548	503506	487548	487548	487548
J-test statistic	0.647	2.15	0.704	1.88	2.43	3.85	2.93	3.37	0.951	7.18
J-test p-value	0.95	0.71	0.95	0.93	0.79	0.79	0.94	0.49	0.92	0.85
Andrews-Lu (2001) stat.	-52.00	-50.36	-51.7	-50.51	-49.96	-48.54	-49.58	-49.01	-51.44	-45.21

Table A.4: The Effect of Information along the Rail and Telegraph Networks: 3-day Periods. The table presents panel 2SLS estimates of competing lag structure specifications of equation (2) on the universe of U.S. 1870 Census towns. In all models a time period is defined as a 3-day interval. The dependent variable is an indicator of crusading activity -meetings, petitions, or marches-. All models include period fixed effects and town fixed effects. Standard errors in square brackets are robust and allow for spatial correlation between neighboring towns along the railroad network. Standard errors in parentheses are clustered at the town level. The last row of the table reports the model selection test statistic of Andrews and Lu (2001). Appendix Table A.5 reports the first stage R squared and F-statistics corresponding to each endogenous regressor in the corresponding column, from top to bottom. Instruments in all specifications are based on a 50 Km. rail accident radius.

Causal Effects of Crusade Signals along the Railroad and Telegraph Networks:
Lag Specification Model Selection First Stages

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
F-statistic	38.46	40.42	40.3	30.5	34.46	26.97	29.48	42.16	42.62	21.8
R squared	0.017	0.017	0.017	0.018	0.017	0.0175	0.018	0.017	0.017	0.018
F-statistic	15.68	13.79	20.09	33.35	20.07	32.34	32.08	19.9	15.67	22.93
R squared	0.007	0.007	0.009	0.017	0.014	0.0173	0.018	0.009	0.008	0.018
F-statistic	888.1	1333.0	1038.4	14.65	26.28	15.38	14.35	1137.4	1184.3	22.58
R squared	0.286	0.290	0.284	0.009	0.016	0.014	0.012	0.285	0.285	0.020
F-statistic				774.7	892.0	20.47	20.9			11.16
R squared				0.285	0.285	0.016	0.016			0.015
F-statistic						678.2	1090.2			14.71
R squared						0.285	0.302			0.018
F-statistic							1266.1			14.5
R squared							0.301			0.020
F-statistic										783.5
R squared										0.301
F-statistic										862.7
R squared										0.303
F-statistic										838.2
R squared										0.297

Table A.5: The Effect of Information along the Rail and Telegraph Networks: 3-day Period First Stages. The table presents the first stage R-squared and F-statistics corresponding to each column of the 2SLS models reported in Table A.4. The statistics for each first stage, from top to bottom, are reported in the same order as the endogenous regressors appear in Table A.4. Following Angrist and Pischke (2008), the F-statistics are corrected for the presence of multiple endogenous regressors.

Causal Effects of Crusade Signals along the Railroad and Telegraph Networks: Robustness First Stages							
Rail Neighbors:	First order links					First and second order links	
Sub-sample:	All towns				Instruments vary	All towns	Instruments vary
Instrument Variation: (accident radius)	80km	120km	50km	50km	50km	50km	50km
Period Definition:	5 days	5 days	3 days	7 days	5 days	5 days	5 days
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
F statistic	46.46	33.07	40.3	54.95	47.28	28.87	28.0
R squared	0.018	0.014	0.017	0.021	0.0346	0.019	0.036
F statistic	11.9	14.27	20.09	18.75	24.15	33.07	33.94
R squared	0.012	0.019	0.009	0.021	0.0273	0.024	0.047
F statistic	683.6	680.0	1038.4	324.0	452.8	25.92	26.7
R squared	0.258	0.259	0.284	0.246	0.374	0.125	0.133
F statistic						15.38	19.0
R squared						0.057	0.068
F statistic						25.05	26.0
R squared						0.073	0.082
F statistic						25.04	25.52
R squared						0.066	0.077
F statistic						124.1	148.8
R squared						0.239	0.366

Table A.6: The Effect of Information along the Rail and Telegraph Networks: Robustness First Stages. The table presents the first stage R-squared and F-statistics corresponding to each column of the 2SLS models reported in Table 5. The statistics for each first stage, from top to bottom, are reported in the same order as the endogenous regressors appear in Table 5. Following Angrist and Pischke (2008), the F-statistics are corrected for the presence of multiple endogenous regressors.

Causal Effects of Crusade Signals along
the Railroad and Telegraph Networks:
Placebo Exercise Using Random Variation in Rail Link Breaks

Dependent Variable:	Any Crusade Activity a_{it} -Meetings, Petitions, Marches-			
Instrument variation:	50km accident radius		120km accident radius	
Period Definition:	5 days	3 days	5 days	3 days
	(1)	(2)	(3)	(4)
$\mathbf{r}_{i,t-1}\mathbf{a}_{t-2}$	-0.008	-0.009	0.003	0.0015
	(0.011)	(0.013)	(0.009)	(0.010)
$\boldsymbol{\gamma}_i\mathbf{a}_{t-1}$	0.198	0.330	0.082	0.075
	(0.12)	(0.19)	(0.194)	(0.05)
$\mathbf{d}_i\mathbf{a}_{t-3}$	0.001	0.002	0.0015	0.0013
	(0.0003)	(0.0004)	(0.0002)	(0.0002)
No. of towns	15934	15950	15934	15950
Max. no. of periods	16	30	16	30
Observations	267247	487548	267247	487548

Table A.7: Random Variation in Rail Link Breaks: Placebo Instruments. The table presents panel 2SLS estimates of equation (2) across alternative specifications, using the lag structure identified as optimal by the Andrews and Lu (2001) test in Table 3 (second order lag for the railroad neighbors' Crusade events, first order lag for the telegraph neighbors' Crusade events, and third order lag for the geographic neighbors' Crusade events). The dependent variable is an indicator of crusading activity -meetings, petitions, or marches-. All models include period fixed effects and town fixed effects. The instruments for the endogenous regressors are built by simulating random link breaks in the railroad network for every day, respecting the true aggregate rate of link breaks in the data. Standard errors are clustered at the town level.

Specification Test: Correlation between Residuals from the Benchmark Specification and Railroad Network Centrality Statistics

Dependent Variable:	Residuals $\hat{\varepsilon}_{i,t}$		
Centrality statistic:	Degree (1)	Betweenness (2)	Eigenvector (3)
Degree	-0.004		
	(0.013)		
Betweenness		-0.561	
		(0.446)	
Eigenvector			0.295
			(0.597)
Centrality statistic ×			
Period 2 dummy	-0.003	0.485	3.342
	(0.018)	(0.625)	(0.838)
Period 3 dummy	-0.016	-0.132	-1.929
	(0.018)	(0.625)	(0.838)
Period 4 dummy	-0.016	-0.058	-1.696
	(0.018)	(0.625)	(0.838)
Period 5 dummy	0.019	-0.444	-0.590
	(0.018)	(0.625)	(0.838)
Period 6 dummy	-0.009	0.474	-2.127
	(0.018)	(0.625)	(0.838)
Period 7 dummy	-0.013	0.641	-0.730
	(0.018)	(0.627)	(0.838)
Period 8 dummy	0.017	1.210	-1.578
	(0.018)	(0.627)	(0.844)
Period 9 dummy	0.012	1.850	-1.880
	(0.018)	(0.627)	(0.844)
Period 10 dummy	0.020	-0.075	-0.050
	(0.018)	(0.629)	(0.844)
Period 11 dummy	0.001	0.570	-0.027
	(0.018)	(0.629)	(0.844)
Period 12 dummy	0.017	0.593	0.361
	(0.018)	(0.629)	(0.844)
Period 13 dummy	0.016	0.575	0.284
	(0.018)	(0.629)	(0.844)
Period 14 dummy	0.023	1.590	0.064
	(0.018)	(0.629)	(0.844)
Period 15 dummy	0.019	1.210	0.398
	(0.018)	(0.629)	(0.844)
Period 16 dummy	0.022	0.858	0.510
	(0.018)	(0.630)	(0.844)
Period 17 dummy	0.039	0.505	0.379
	(0.018)	(0.630)	(0.844)
No. of towns	19534	19534	19534
Max. no. of periods	17	17	17
Observations	267247	267247	267247

Table A.8: Specification Test: Benchmark Model Residuals and Railroad Network Structure. The table presents panel OLS regression estimates. The dependent variable in all columns corresponds to the 2SLS residual from the optimally selected lag specification in column (3) of Table 3. All models include period fixed effects. Column (1) includes the railroad network degree centrality, and a full set of interactions between the degree centrality and period dummy variables as additional regressors. Column (2) includes the railroad network betweenness centrality, and a full set of interactions between the betweenness centrality and period dummy variables as additional regressors. Column (3) includes the railroad network eigenvector centrality, and a full set of interactions between the betweenness centrality and period dummy variables as additional regressors. The coefficients in column (1) are multiplied by 100. The coefficients of column (2) are multiplied by 10^8.

Newspaper Coverage along the Railroad and Telegraph Networks				
Dependent variable:	Dummy for town i newspaper report about crusading town j			
	(1)	(2)	(3)	(4)
Railroad network path length $i \to j$	-0.117	-0.189		
	(0.050)	(0.060)		
Telegraph network path length $i \to j$			-2.202	-5.180
			(2.191)	(0.814)
Geographic distance between towns i and j	-0.152	0.539	-0.232	-2.120
	(0.102)	(0.214)	(0.488)	(1.530)
Newspaper town covariates				
Railroad network betweenness centrality	0.0009		13.5	
	(0.0009)		(9.81)	
Telegraph network dummy	0.008			
	(0.009)			
Crusading town covariates				
Railroad network betweenness centrality	-0.001		-0.329	
	(0.0002)		(0.144)	
Telegraph network dummy	-0.013			
	(0.0016)			
Newspaper town fixed effects	No	Yes	No	Yes
Crusading town fixed effects	No	Yes	No	Yes
R squared	0.004	0.32	0.05	0.62
No. of observations	50076	50076	402	402

Table A.9: Newspaper Coverage along the Railroad and Telegraph Networks: Path Lengths
The table presents OLS regression estimates on a panel of pairs of newspaper home towns-x-crusading towns. The dependent variable in all columns is a dummy variable taking the value of one if the newspaper in town i reported on any Crusade activity of town j. Standard errors are robust and clustered at the newspaper home town level. The coefficients on the railroad and telegraph network path length variables are multiplied by 1000. The coefficients on the geographic distance between towns are in kms. and multiplied by 10^5. The coefficients on the betweenness centrality statistic are multiplied by 10^6.

B Online Appendix B: Technological Complementarities: Additional Results

Rail and Telegraph Technological Complementarities: Alternative Cluster Radii Event Studies

	50 KMS			80 KMS			120 KMS		
	2 week (1)	3 week (2)	4 week (3)	2 week (4)	3 week (5)	4 week (6)	2 week (7)	3 week (8)	4 week (9)
$r_{ij}\gamma_j$	0.081	0.081	0.086	0.060	0.072	0.074	0.047	0.059	0.069
	(0.034)	(0.026)	(0.027)	(0.027)	(0.021)	(0.021)	(0.025)	(0.020)	(0.021)
$r_{ij}(1-\gamma_j)$	0.0016	0.0045	0.0047	0.0021	0.0035	0.0038	0.0026	0.0036	0.0039
	(0.004)	(0.004)	(0.003)	(0.0034)	(0.0032)	(0.0028)	(0.0031)	(0.0028)	(0.0025)
Complementarity	0.080	0.076	0.081	0.058	0.068	0.070	0.045	0.056	0.065
	(0.034)	(0.025)	(0.026)	(0.026)	(0.020)	(0.021)	(0.025)	(0.020)	(0.021)
Signal-recipient distance	-0.0042	-0.0024	-0.0014	-0.0019	-0.0006	-0.0003	-0.0013	-0.0001	0.0001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.0005)	(0.0005)	(0.0005)
Cluster FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
State FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Recipient town FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Mean of dep. var.	0.043	0.057	0.067	0.041	0.053	0.063	0.037	0.050	0.059
R squared	0.032	0.033	0.029	0.021	0.021	0.018	0.018	0.017	0.013
Observations	79134	79134	79134	193745	193745	193745	407618	407618	407618

Table B.1: Rail and Telegraph Technological Complementarities: Alternative Cluster Radii Event Studies. The table presents estimation results of the cluster event study approach based on equation (3) for alternative cluster radii definitions. Columns (1)-(3) use 50 Km. radius clusters. Columns (4)-(6) use 80 Km. radius clusters. Columns (7)-(9) use 120 Km. radius clusters. The dependent variable is a dummy variable for whether a town within the cluster radius experienced a Crusade event within the time window in each column header following the cluster-defining town experiencing its Crusade event. All models include event-cluster fixed effects, state fixed effects, recipient-town fixed effects, and the distance between generating and recipient towns. The complementarity interaction effects are computed as the difference between the coefficients on $r_{ij}\gamma_j$ and $r_{ij}(1-\gamma_j)$. Standard errors are robust and clustered two-ways, at the event-cluster and at the recipient town levels.

Rail and Telegraph Technological Complementarities: 2-Week Window Cluster Event Studies–Heterogeneity

Interaction variable:	Newspapers per capita		Post Office Dummy		Religious Ascriptions Hefindahl Index		Gender Ratio	
	30KM (1)	50KM (2)	30KM (3)	50KM (4)	30KM (5)	50KM (6)	30KM (7)	50KM (8)
$r_j^i \gamma_j$	0.122	0.119	0.144	0.084	0.081	0.007	0.108	0.079
	(0.042)	(0.034)	(0.057)	(0.061)	(0.156)	(0.121)	(0.052)	(0.045)
$r_j^i(1-\gamma_j)$	-0.0006	0.0029	-0.0091	-0.0075	0.0002	0.0045	-0.0014	0.0017
	(0.005)	(0.004)	(0.005)	(0.004)	(0.0117)	(0.0095)	(0.0057)	(0.0066)
$r_j^i \gamma_j \times$ Interaction	-58.16	-86.56	-0.044	-0.003	0.119	0.339	-0.004	0.005
	(101.7)	(36.2)	(0.078)	(0.071)	(0.580)	(0.474)	(0.040)	(0.057)
$r_j^i(1-\gamma_j) \times$ Interaction	-21.69	-20.35	0.013	0.016	-0.010	-0.014	-0.0008	-0.0001
	(20.2)	(17.8)	(0.007)	(0.005)	(0.047)	(0.035)	(0.0062)	(0.0097)
Complementarity	0.123	0.116	0.153	0.092	0.081	0.003	0.109	0.077
	(0.042)	(0.033)	(0.057)	(0.061)	(0.156)	(0.121)	(0.050)	(0.043)
Complementarity interaction	-36.47	-66.21	-0.057	-0.019	0.129	0.353	-0.004	0.005
	(104.2)	(38.9)	(0.078)	(0.070)	(0.581)	(0.475)	(0.040)	(0.055)
Signal-recipient distance	-0.0045	-0.0042	-0.0043	-0.0042	-0.0045	-0.0042	-0.0045	-0.0042
	(0.0028)	(0.0015)	(0.0028)	(0.0015)	(0.0028)	(0.0015)	(0.0028)	(0.0015)
R squared	0.066	0.033	0.066	0.033	0.066	0.032	0.066	0.032
Observations	29592	79133	29592	79134	29590	79129	29592	79133

Table B.2: Testing for Heterogeneity in the Rail-Telegraph Complementarity. The table presents estimation results of the cluster event-study approach based on equation (3), allowing for interaction terms between the railroad and telegraph characteristics with either the number of newspapers per capita, a Post Office dummy, the Herfindahl index of religious ascriptions, and the gender ratio (females/males). All models are estimated for the 2-week window responses, and include event-cluster fixed effects, state fixed effects, recipient town fixed effects, and the distance between generating and recipient towns. Odd-numbered columns present models based on 30 Km. radius clusters. Even-numbered columns present models based on 50 Km. radius clusters. The complementarity effects are computed as the difference between the coefficients on $r_{ij}\gamma_j$ and $r_{ij}(1-\gamma_j)$. The complementarity interaction effects are computed as the difference between the coefficients on $r_{ij}\gamma_j \times$ Interaction and $r_{ij}(1-\gamma_j) \times$ Interaction. Standard errors are robust and clustered two-ways, at the event-cluster and at the recipient town levels.

Rail and Telegraph Technological Complementarities: Placebo Event Studies using Close Match Signal-Generating Towns

	30 KMS				50 KMS		80 KMS		120 KMS	
	2 weeks (1)	4 weeks (2)	2 weeks (3)	4 weeks (4)	2 weeks (5)	4 weeks (6)	2 weeks (7)	4 weeks (8)	2 weeks (9)	4 weeks (10)
$r^i_j \gamma_j$	0.358	-0.021	0.022	-0.064	0.196	-0.018	-0.236	0.322	-0.214	-0.014
	(0.139)	(0.019)	(0.133)	(0.054)	(0.220)	(0.010)	(3.760)	(0.160)	(0.098)	(0.071)
$r^i_j (1-\gamma_j)$	0.0013	-0.0044	0.0078	-0.0163	0.0706	-0.0206	-0.0170	0.0131	-0.0080	0.0291
	(0.046)	(0.005)	(0.041)	(0.009)	(0.039)	(0.008)	(0.0188)	(0.0046)	(0.0068)	(0.0190)
$(1-r^i_j)\gamma_j$	0.154	0.0191								
	(0.022)	(0.007)								
Complementarity	0.203	-0.036	0.014	-0.048	0.125	0.003	-0.219	0.309	-0.206	-0.043
	(0.150)	(0.018)	(0.136)	(0.056)	(0.236)	(0.011)	(3.760)	(0.159)	(0.098)	(0.057)
Signal-recipient distance	-0.0072	0.0039	-0.0031	0.0033	0.0012	-0.0019	-0.0012	0.0012	-0.0029	-0.0027
	(0.006)	(0.003)	(0.004)	(0.002)	(0.002)	(0.001)	(0.0017)	(0.0010)	(0.0013)	(0.0007)
Cluster FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
State FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Recipient town FE	N	N	Y	Y	Y	Y	Y	Y	Y	Y
Mean of dep. var.	0.046	0.011	0.046	0.011	0.046	0.012	0.044	0.011	0.041	0.011
R squared	0.096	0.056	0.056	0.058	0.037	0.032	0.025	0.019	0.022	0.013
Observations	20399	20399	20399	20399	52399	52399	123140	123140	238097	238097

Table B.3: Rail-Telegraph Complementarities: Placebo Event Studies using Close Match Signal-Generating Towns. The table presents estimation results of the cluster event-study approach based on equation (3), where the signal generating town i for each event study is replaced by its closest match within the set of Crusading towns, along the following observable characteristics: native share, black share, newspapers per capita, female to male ratio, alcohol vendors per capita, religious ascriptions Herfindahl index, Presbyterian sittings per capita, and log population. The dependent variable is a dummy variable for whether a town within the cluster radius experienced a Crusade event within the time window in each column header following the placebo town experiencing its Crusade event. All models include event-cluster fixed effects, state fixed effects, and the distance between generating and recipient towns. Models in columns (3)-(10) include recipient town fixed effects. Columns (1)-(4) use 30 Km. radius clusters. Columns (5)-(6) use 50 Km. radius clusters. Columns (7)-(8) use 80 Km. radius clusters. Columns (9)-(10) use 120 Km. radius clusters. In columns (1)-(2) the complementarity interaction effects are computed as the difference between the coefficients on $r_{ij}\gamma_j$, $r_{ij}(1-\gamma_j)$, and $(1-r_{ij})\gamma_j$. In columns (3)-(10) the complementarity interaction effects are computed as the difference between the coefficients on $r_{ij}\gamma_j$ and $r_{ij}(1-\gamma_j)$. Standard errors are robust and clustered two-ways, at the event-cluster and at the recipient town levels.

C Online Appendix C: Supplementary Data Description

Newspaper Articles Data Construction

We collected newspaper data from the "Chronicling of America" Newspaper database of the Library of Congress. The archive contains images of historic newspapers from 1690 to present. Its online interface allows the researcher to carry out keyword searches.

We searched for the following keywords (or combination of keywords when one of the keywords is likely to generate a large number of false positives) to identify mentions of events related to the Temperance Crusade: Crusade; Dio Lewis; Saloon pledge; Temperance; Temperance & Women; War & Whisky; Women & Protest; Women & War.

We scraped the website to download the texts that contain any of the keywords or keyword combinations (together) in their body. We also downloaded key information about the newspaper - such as its name and location.

The output from these searches resulted in several thousand articles which carried at least one of these keywords, some of which may be duplicates. The output is an image of the newspaper page. The data set also allows the output to be downloaded as text derived from the processing of this image.

To reduce image-to-text processing issues, we first implemented the following steps:

1. The text was turned into all lowercase to reduce differences due to lower and upper case.
2. We removed punctuation and signs that were likely to be included in the output due to imperfect image processing, such as or |.
3. We searched for words which may have been unintentionally divided with a space to create two unintelligible terms. For example, if the image processing software resulted in the word "development" to separate into two consecutive words like "deve" and "lopment," we would combine the two words since this would result in a meaningful new word. Unfortunately, while these steps reduce errors, they can also generate combination words which were not in the original text. For example, if two words 'up" and "date" were consecutively available in the text, we would form the word "update". This is an unavoidable trade-off in our search algorithm.

We ended up with 5,749 mentions of events from 194 newspapers titles.

Beginning and End of Articles

Our goal is to find town events are mentioned in articles related to the Temperance Crusade. To reduce the number of false positive mentions of a town, we would like to identify the text of an article which mentions the Women's Crusade event, rather than the text of an entire newspaper page. So we aim to carry out the search of town names only in the text of a Women's Crusade event related article.

Unfortunately, aside from the messiness of the text due to imagine processing, working with the historical newspaper data is challenging because in historic newspapers, there were very little indicators of where an article begins and where it ends. Modern day newspapers, for example, contain indicators such as the name of the reporter, or the first letter being a slightly larger font than the rest, which could be used to separate the beginning of a text from the ending of it. This is not the case for older newspapers. Although many newspapers were smaller in physical size for the duration of interest, inclusion of other articles may still result in picking up the names of towns that are not relevant to the women's protests.

We followed the following approach to determine the body of a text. We first suppose that if there are multiple sequential pages turning in the search, these are likely coming from the same article spread across multiple pages. So we combine texts from sequential pages (if there are hits from the same newspaper pages but the pages are not sequential, we suppose there may be multiple articles).

In this combined body, we then take the first keyword hit and the last keyword hit as indicators of where an article may likely be located. It is unlikely that these words coincide with the first and the last words of an article, so we take all the text between these words and add combine a length of text before and after these words in order to cover the earlier and later texts. We add 100, 200 and 500 words in three different versions of the processing as these "padding texts" before and after the body, and do the town name search in this combined text.

Searching for the Names of Towns

We searched for town name mentions within the combined text capturing an article body. We then matched the list of recovered names in all articles to our list of Crusading towns. This procedure can give rise to false positives. For example, "Union" is a town in NY. However, the word itself is also meaningful and quite commonly used in this period. To reduce such errors, we searched for the names of towns by looking for the words starting with a capital letter.

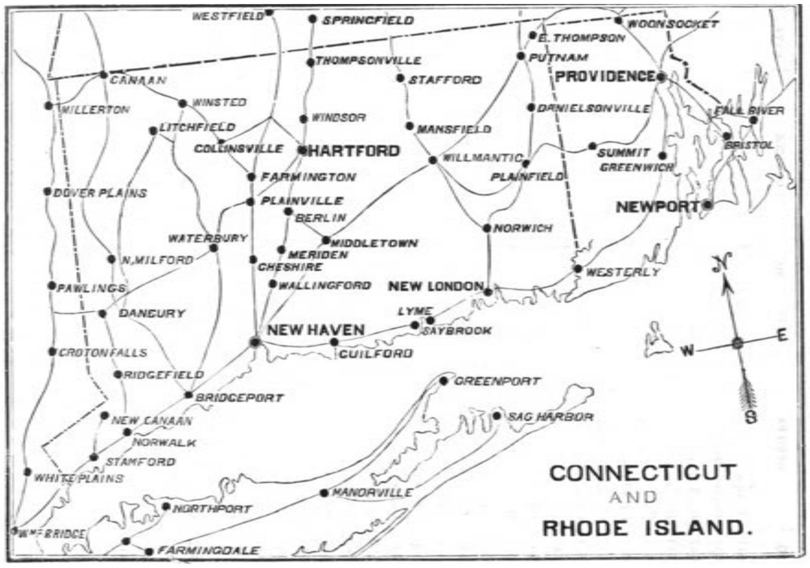

Figure C.1: The Telegraph Network in Connecticut and Rhode Island, 1874. The figure reproduces the Western Union telegraph lines map in WesternUnion (1874) for the states of Connecticut and Rhode Island.

For any town in our list, it was possible for it to be the only town with its name. It was also possible for some towns to have identical names in different states. While for some towns we could be sure which town was mentioned exactly, in other cases we were less confident about the identity of the mentioned town. To deal with the latter case, we took two measures: i) we checked whether the state of the town was also mentioned. If only one state was mentioned, we coded the town in the corresponding state as mentioned. If there were multiple states with a possible match mentioned, we assign a probability equal to 1/number of mentioned towns. If there were, for example, five towns with identical names in the 800+ towns in our search, and three of them whose states were mentioned, we assigned a 1/3 probability.

For each article, each town resulted coded as either a 0 (no mention), a 1 (town + state name mentioned, or there is only one town, or town+state combination) or a number between 0 and 1 (when there is only partial information and are multiple towns or town+state names matching). Unfortunately after these steps, there were still some town names which had high levels of false positives. Whenever we ran into similar cases with likely high number of false positives, a research assistant manually checked the names of the towns mentioned in the corresponding article. We then took this output and created a town to town mention matrix for each day in our study. In the rows we report the newspapers town location and in the rows the Crusade town mention. Thus we mark whether a town - through its own newspaper - hears about the events in another town.

Illustration of a Telegraph Map from WesternUnion (1874)

Figure C.1 reproduces the Western Union telegraph lines map in WesternUnion (1874) for the states of Connecticut and Rhode Island as an illustration. We geo-referenced these maps for all states using GIS software to create the telegraph network data.