Online Appendices for ‘Information Networks and Collective Action: Evidence from the Women’s Temperance Crusade, by Camilo Garcia-Jimeno, Angel Iglesias, and Pinar Yildirim’

A Online Appendix: Social Interactions Figures and Tables

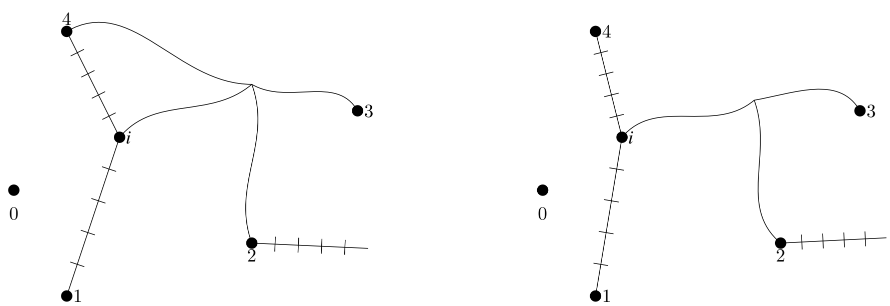
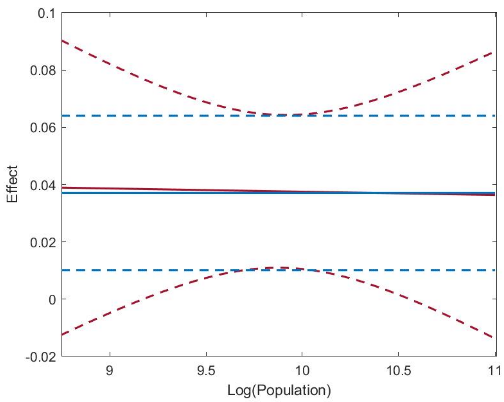
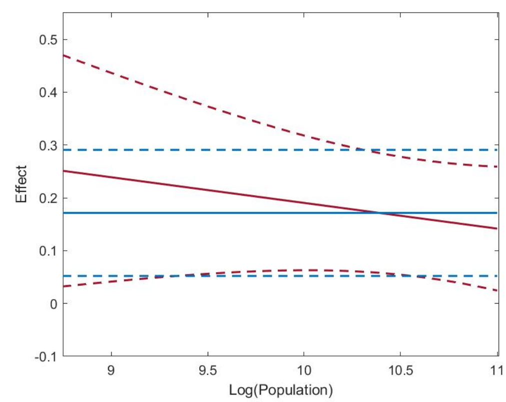


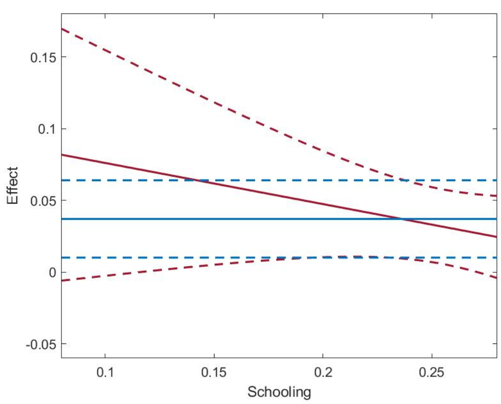
Figure A.1: Types of Connections between Neighboring Towns and Identification. The figure on the left illustrates all the potential types of connections between towns: town i and town 0 are not connected by an observed network; town i and town 1 are connected by a direct rail link; town i and town 2 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 the pair $(i, 4)$ from the left-hand side figure, making the identification of interaction effects between rail and network-mediated information flows not possible using time series variation in rail link activity.



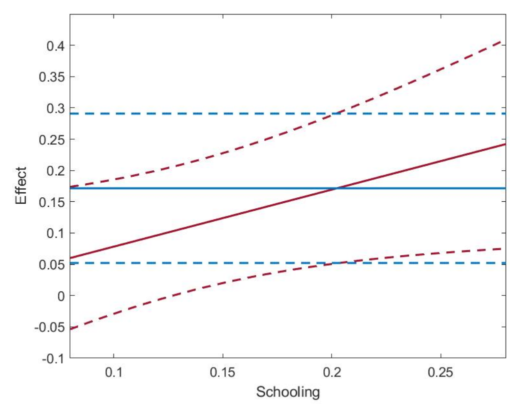
(a) Population Heterogeneity: Rail



(b) Population Heterogeneity: Telegraph

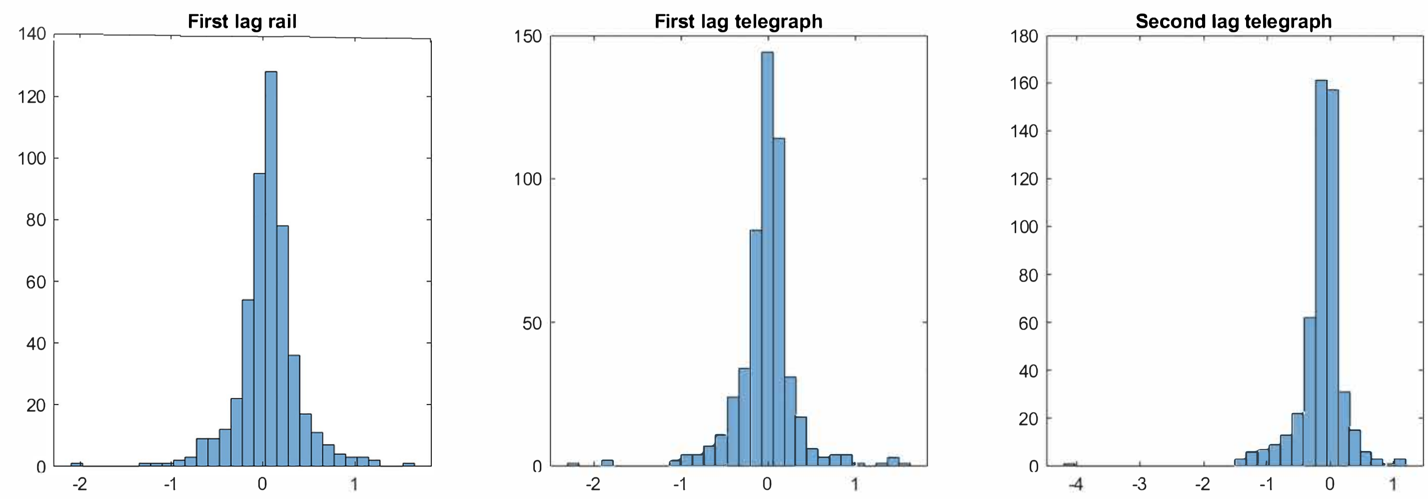


(c) Schooling Heterogeneity: Rail

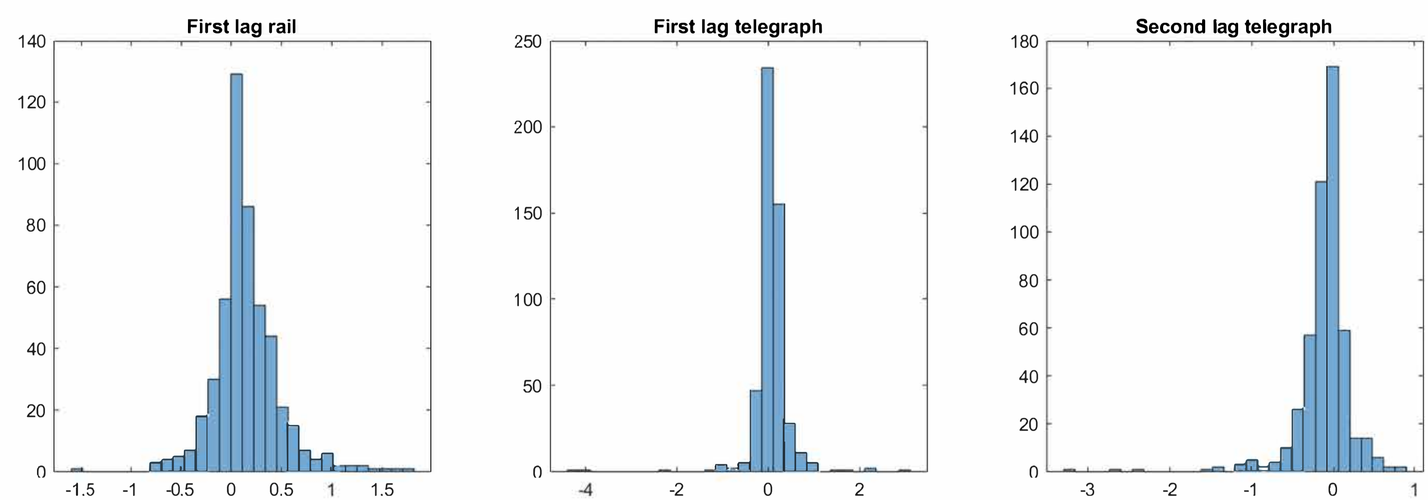


(d) Schooling Heterogeneity: Telegraph

Figure A.2: The Effect of Information along the Rail and Telegraph Networks: Allowing for heterogeneity along the Schooling and Population Gradients: The figure presents estimated heterogeneous effects from panel IV models based on equation (1), using our benchmark lag-structure specification (first order lag for the railroad neighbors' Crusade events, first and second order lags for the telegraph neighbors' Crusade events, and third order lag for the geographic neighbors' Crusade events), including interactions terms. The dependent variable is an indicator of crusading activity -meetings, petitions, or marches-. Sub-figures (a) and (b) correspond to a model that includes interactions between log population and the first lags of rail and telegraph-mediated information. Sub-figures (c) and (d) correspond to a model that includes interactions between average schooling and the first lags of rail and telegraph-mediated information. The implied heterogeneous effects are represented in red. For ease of comparison, the corresponding homogeneous effect (from our benchmark estimates) are represented in blue. All models include period fixed effects and town fixed effects, use the benchmark 50 km. radius definition of rail accidents for the instruments, and use the 5-day interval period definition. Dashed curves represent 95 percent confidence intervals.



(a) 5-day periods



(b) 3-day periods

Figure A.3: Random Variation in Rail Link Breaks: Placebo Instruments The figure presents empirical distributions of IV estimates from equation (1) across 500 simulations, using the lag structure identified as optimal by the Andrews and Lu (2001) test in Table A.3 (first order lag for the railroad neighbors' Crusade events, first and second 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. In each simulation, the instruments for the endogenous regressors are built by generating random link breaks in the railroad network day by day, to match the marginal distribution of rail link breaks in the data. The top panel reports results using the 5-day period definition. The bottom panel reports results using the 3-day period definition.

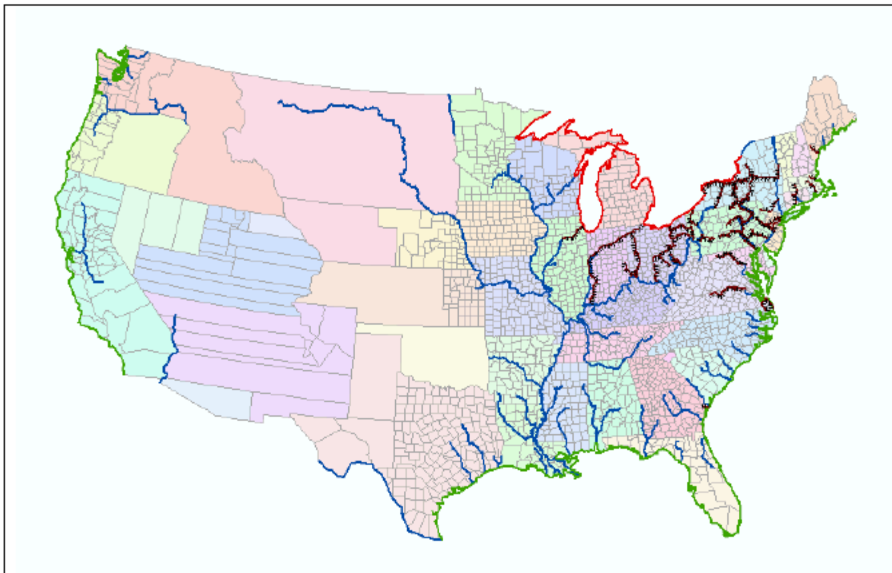


Figure 3: America's System of Navigable Waterways Mapped Against 1860 County Boundaries

Figure A.4: The figure illustrates the network of navigable rivers, canals, and waterways as of 1860 in the United States, which we borrow from Atack, Bateman and Margo (2007).

Empirical Survivor Town Characteristics

Days since	Newspapers pc		Post Office		Alcohol Vendors pc		Religious Herfindhal Index	
beginning	No Event Yet	Some Event	No Event Yet	Some Event	No Event Yet	Some Event	No Event Yet	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)
t-stat for equality of means	-0.805		-0.921		0.004		11.342	
	0.056		0.018		0.083		0.004	

	Black Pop. Share		Rail Betweenness Centrality		Rail Degree Centrality		Number of Towns	
	No Event Yet	Some Event	No Event Yet	Some Event	No Event Yet	Some Event	No Event Yet	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)		
t-stat for equality of means	14.138		-2.547		-8.253			
	0.010		4.192		0.169			

Table A.1: Survivor Table for Town Characteristics During the Temperance Crusade. The table presents means and standard deviations for a set of town characteristics, at different points in time (indicated in the leftmost column) after the beginning of the Crusade. 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 a Crusade event. The last row of each panel reports t-statistics for the equality of means between both groups of towns at day 215. Newspapers and Alcohol vendors per capita are multiplied by 100. The betweenness centrality statistic is multiplied by 10^3.

Town Characteristics by Rail Accidents Exposure

	Newspapers pc		Post Office		Alcohol Vendors pc		Religious Herfindhal Index	
	Below	Above	Below	Above	Below	Above	Below	Above
Mean	0.22	0.24	0.57	0.60	0.36	0.34	0.26	0.25
Std. Dev.	(0.90)	(1.37)	(0.50)	(0.49)	(2.56)	(2.19)	(0.12)	(0.13)

	Black Pop. Share		Rail Betweenness Centrality		Rail Degree Centrality		Number of Towns	
	Below	Above	Below	Above	Below	Above	Below	Above
Mean	0.14	0.13	29	60	2.14	5.01	5155	5156
Std. Dev.	(0.58)	(0.75)	(91.67)	(157.13)	(0.61)	(5.46)		

Table A.2: Town Characteristics by Railroad Accident Exposure. The table presents means and standard deviations for a set of town characteristics, comparing towns below and above the median of the distribution of rail link breaks (accidents) experienced during the Temperance Crusade period, using the 50 km./ radius definition. For each covariate, the columns to the left report summary statistics for towns below median. The columns to the right report summary statistics for towns above median. Newspapers and Alcohol vendors per capita are multiplied by 100. The betweenness centrality statistic is multiplied by 10^3.

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)	(11)	(12)
First lag rail	0.049			0.049		0.037	0.037	0.051	0.029			0.022
$(\mathbf{r}_{i,t}\mathbf{a}_{t-1})$	[0.012]			[0.018]		[0.013]	[0.013]	[0.018]	[0.019]			[0.016]
	(0.020)			(0.022)		(0.013)	(0.013)	(0.022)	(0.021)			(0.017)
Second lag rail		0.039	0.020	-0.013	0.021			-0.014	0.010		0.023	-0.006
$(\mathbf{r}_{i,t-1}\mathbf{a}_{t-2})$		[0.011]	[0.012]	[0.014]	[0.012]			[0.014]	[0.014]		[0.011]	[0.015]
		(0.017)	(0.013)	(0.021)	(0.013)			(0.021)	(0.021)		(0.014)	(0.025)
Third lag rail										0.018		0.017
$(\mathbf{r}_{i,t-2}\mathbf{a}_{t-3})$										[0.011]		[0.012]
										(0.012)		(0.018)
First lag telegraph	0.108		0.134	0.131	0.165	0.132	0.172	0.168	0.114	0.137		0.1426
$(\boldsymbol{\gamma}_i\mathbf{a}_{t-1})$	[0.025]		[0.025]	[0.025]	[0.032]	[0.025]	[0.033]	[0.033]	[0.030]	[0.025]		[0.047]
	(0.043)		(0.043)	(0.043)	(0.060)	(0.043)	(0.061)	(0.060)	(0.055)	(0.043)		(0.069)
Second lag telegraph		0.087			-0.065		-0.068	-0.065	-0.035			-0.072
$(\boldsymbol{\gamma}_i\mathbf{a}_{t-2})$		[0.024]			[0.030]		[0.031]	[0.031]	[0.029]			[0.067]
		(0.042)			(0.076)		(0.076)	(0.076)	(0.072)			(0.106)
Third lag telegraph											0.049	0.034
$(\boldsymbol{\gamma}_i\mathbf{a}_{t-3})$											[0.023]	[0.047]
											(0.034)	(0.049)
First lag distance	0.000								0.001			0.001
$(\mathbf{d}_i\mathbf{a}_{t-1})$	[0.002]								[0.001]			[0.001]
	(0.003)								(0.001)			(0.001)
Second lag distance		-0.002							-0.001			-0.001
$(\mathbf{d}_i\mathbf{a}_{t-2})$		[0.002]							[0.001]			[0.001]
		(0.002)							(0.001)			(0.001)
Third lag distance			0.004	0.005	0.005	0.005	0.006	0.006		0.005	0.000	0.001
$(\mathbf{d}_i\mathbf{a}_{t-3})$			[0.001]	[0.001]	[0.002]	[0.002]	[0.002]	[0.002]		[0.002]	[0.001]	[0.001]
			(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)		(0.002)	(0.002)	(0.001)
No. of towns	15,960	15,947	15,934	15,934	15,934	15,934	15,934	15,934	15,947	15,934	15,934	15,934
Max. no. of periods	18	17	16	16	16	16	16	16	17	16	16	16
Observations	299,154	283,194	267,247	267,247	267,247	267,247	267,247	267,247	283,194	267,247	267,247	267,247
Kleibergen-Paap Wald	141.4	237.8	178.0	18.9	112.3	185.8	112.3	17.6	8.4	165.2	359.9	4.9
J-test statistic	2.14	11.76	7.58	2.93	6.91	2.97	2.16	2.20	9.50	4.32	17.41	17.16
J-test p-value	0.711	0.019	0.110	0.818	0.228	0.563	0.827	0.948	0.302	0.365	0.002	0.144
Andrews-Lu (2001) stat.	-12.16	-2.54	-11.49	-16.14	-16.92	-16.10	**-21.68**	-21.63	-0.03	-5.22	3.11	-1.91

Table A.3: The Effect of Information along the Rail and Telegraph Networks: Lag Specification Model Selection. The table presents panel IV estimates of competing lag-structure specifications of equation (1) 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.4 reports the first stage F-statistics and p-values corresponding to each endogenous regressor in the corresponding column, from top to bottom. The first four columns of Appendix Table A.5 report the first stage coefficients of the model in column (7). 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 Stage F-statistics

First stages:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
First lag rail	52.95			35.24		48.28	38.40	30.08	27.30			19.93
($\mathbf{r}_{i,t}\mathbf{a}_{t-1}$)	0.000			0.000		0.000	0.000	0.000	0.000			0.000
Second lag rail		51.59	47.84	39.22	37.54			32.79	31.28		46.94	20.16
($\mathbf{r}_{i,t-1}\mathbf{a}_{t-2}$)		0.000	0.000	0.000	0.000			0.000	0.000		0.000	0.000
Third lag rail										51.26		25.30
($\mathbf{r}_{i,t-2}\mathbf{a}_{t-3}$)										0.000		0.000
First lag telegraph	84.82		67.75	54.69	74.17	75.00	76.43	58.92	58.08	65.21		41.78
($\boldsymbol{\gamma}_i\mathbf{a}_{t-1}$)	0.000		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000
Second lag telegraph		88.29			64.64		65.50	50.56	50.08			55.58
($\boldsymbol{\gamma}_i\mathbf{a}_{t-2}$)		0.000			0.000		0.000	0.000	0.000			0.000
Third lag telegraph											86.34	50.43
($\boldsymbol{\gamma}_i\mathbf{a}_{t-3}$)											0.000	0.000
First lag distance	116.21								1320.86			1330.32
($\mathbf{d}_i\mathbf{a}_{t-1}$)	0.000								0.000			0.000
Second lag distance		182.13							1433.60			1272.00
($\mathbf{d}_i\mathbf{a}_{t-2}$)		0.000							0.000			0.000
Third lag distance			158.45	133.10	192.97	135.60	174.94	162.11		183.90	253.78	1133.02
($\mathbf{d}_i\mathbf{a}_{t-3}$)			0.000	0.000	0.000	0.000	0.000	0.000		0.000	0.000	0.000

Table A.4: The Effect of Information along the Rail and Telegraph Networks: Lag Specification Model Selection First Stages. The table presents the first-stage F-statistics and p-values corresponding to each column of the IV models reported in Table A.3. The statistics for each first stage, from top to bottom, are reported in the same order as the endogenous regressors appear in Table A.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			
Dependent Variable:	First Lag Rail $\mathbf{r}_{i,t}\mathbf{a}_{t-1}$	First Lag Telegraph $\boldsymbol{\gamma}_i\mathbf{a}_{t-1}$	Second Lag Telegraph $\boldsymbol{\gamma}_i\mathbf{a}_{t-2}$	Third Lag Distance $\mathbf{d}_i\mathbf{a}_{t-3}$	First Lag Rail $\mathbf{r}_{i,t}\mathbf{a}_{t-1}$	First Lag Telegraph $\boldsymbol{\gamma}_i\mathbf{a}_{t-1}$	Second Lag Telegraph $\boldsymbol{\gamma}_i\mathbf{a}_{t-2}$	Third Lag Distance $\mathbf{d}_i\mathbf{a}_{t-3}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Rail links	0.728	0.091	0.024	5.480	0.931	0.111	0.021	2.466
$(\mathbf{r}_{i,t}\boldsymbol{\iota})$	(0.117)	(0.019)	(0.019)	(1.433)	(0.106)	(0.017)	(0.018)	(1.323)
Rail links of rail neighbors	0.180	-0.012	0.001	-0.112	0.154	-0.015	0.004	0.263
$(\mathbf{r}_{i,t}\mathbf{R}_{t-1}\boldsymbol{\iota})$	(0.016)	(0.003)	(0.003)	(0.191)	(0.016)	(0.003)	(0.003)	(0.197)
Rail links of rail neigh. of rail neigh.	-0.015	0.000	0.000	0.018	-0.017	0.000	0.000	0.008
$(\mathbf{r}_{i,t}\mathbf{R}_{t-1}\mathbf{R}_{t-2}\boldsymbol{\iota})$	(0.001)	(0.000)	(0.000)	(0.007)	(0.001)	(0.000)	(0.000)	(0.009)
Rail links of telegraph neighbors	0.017	0.568	-0.349	7.840	-0.123	0.385	-0.068	-0.248
$(\boldsymbol{\gamma}_i\mathbf{R}_{t-1}\boldsymbol{\iota})$	(0.153)	(0.025)	(0.025)	(1.870)	(0.111)	(0.018)	(0.019)	(1.390)
Rail links of rail neigh. of telegraph neigh.	0.004	0.001	0.125	-0.602	0.014	-0.011	0.088	-0.008
$(\boldsymbol{\gamma}_i\mathbf{R}_{t-1}\mathbf{R}_{t-2}\boldsymbol{\iota})$	(0.014)	(0.002)	(0.002)	(0.172)	(0.010)	(0.002)	(0.002)	(0.123)
Lag of rail links of telegraph neigh.	-0.041	0.791	0.921	-1.594	0.132	0.997	0.766	0.782
$(\boldsymbol{\gamma}_i\mathbf{R}_{t-2}\boldsymbol{\iota})$	(0.156)	(0.025)	(0.026)	(1.909)	(0.112)	(0.018)	(0.019)	(1.396)
Lag of rail links of rail neigh. of teleg. neigh.	0.001	-0.154	-0.166	1.149	-0.015	-0.132	-0.140	0.560
$(\boldsymbol{\gamma}_{it}\mathbf{R}_{t-2}\mathbf{R}_{t-3}\boldsymbol{\iota})$	(0.010)	(0.002)	(0.002)	(0.119)	(0.008)	(0.001)	(0.001)	(0.094)
Rail links of distance neighbors	0.002	0.001	0.001	0.201	-0.005	0.000	0.001	0.359
$(\mathbf{d}_i\mathbf{R}_{t-3}\boldsymbol{\iota})$	(0.002)	(0.000)	(0.000)	(0.021)	(0.001)	(0.000)	(0.000)	(0.018)
Rail links of rail neigh. of distance neigh.	0.000	0.000	0.000	-0.023	0.000	0.000	0.000	-0.052
$(\mathbf{d}_i\mathbf{R}_{t-3}\mathbf{R}_{t-4}\boldsymbol{\iota})$	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)

Table A.5: First Stages for Optimal Lag Structure Models. The table presents the first-stage coefficient estimates and standard errors for our preferred lag specification. All models include town fixed effects and period fixed effects, and use a 5-day period definition. Columns (1)-(4) present the first stages corresponding to the four endogenous regressors of column (7) in Table A.3, with instruments based on a 50 km. radius for the railroad accidents. Columns (5)-(8) present the first stages corresponding to the four endogenous regressors of column (1) in Table 3, with instruments based on an 80 km. radius for the railroad accidents. The dependent variable in columns (1) and (5) is the first order lag of railroad neighbors' Crusade events. The dependent variable in columns (2) and (6) is the first order lag of telegraph neighbors' Crusade events. The dependent variable in columns (3) and (7) is the second order lag of telegraph neighbors' Crusade events. The dependent variable in columns (4) and (8) is the third order lag of geographic neighbors' Crusade events. All coefficients and standard errors are multiplied by 100.

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)	(11)	(12)
First lag rail	0.050			0.022		0.033	0.032	0.022	0.038			0.032
$(\mathbf{r}_{i,t}\mathbf{a}_{t-1})$	[0.013]			[0.021]		[0.014]	[0.014]	[0.021]	[0.022]			[0.013]
	(0.018)			(0.028)		(0.015)	(0.015)	(0.027)	(0.030)			(0.018)
Second lag rail		0.052	0.039	0.015	0.037			0.013	0.003		0.048	0.000
$(\mathbf{r}_{i,t-1}\mathbf{a}_{t-2})$		[0.013]	[0.013]	[0.022]	[0.013]			[0.022]	[0.022]		[0.013]	[0.020]
		(0.021)	(0.018)	(0.034)	(0.018)			(0.034)	(0.036)		(0.019)	(0.029)
Third lag rail										0.019		0.005
$(\mathbf{r}_{i,t-2}\mathbf{a}_{t-3})$										[0.012]		[0.017]
										(0.018)		(0.021)
First lag telegraph	0.036		0.059	0.057	0.072	0.064	0.072	0.072	0.074	0.081		0.079
$(\boldsymbol{\gamma}_i\mathbf{a}_{t-1})$	[0.068]		[0.048]	[0.047]	[0.015]	[0.048]	[0.015]	[0.015]	[0.015]	[0.047]		[0.015]
	(0.108)		(0.088)	(0.087)	(0.040)	(0.089)	(0.041)	(0.040)	(0.040)	(0.086)		(0.038)
Second lag telegraph		-0.100			0.017		0.020	0.018	0.018			0.017
$(\boldsymbol{\gamma}_i\mathbf{a}_{t-2})$		[0.058]			[0.024]		[0.024]	[0.024]	[0.024]			[0.015]
		(0.090)			(0.036)		(0.036)	(0.036)	(0.035)			(0.017)
Third lag telegraph											-0.192	-0.014
$(\boldsymbol{\gamma}_i\mathbf{a}_{t-3})$											[0.070]	[0.023]
											(0.126)	(0.041)
First lag distance	0.000								0.002			0.000
$(\mathbf{d}_i\mathbf{a}_{t-1})$	[0.001]								[0.001]			[0.001]
	(0.001)								(0.002)			(0.001)
Second lag distance		-0.001							-0.002			0.002
$(\mathbf{d}_i\mathbf{a}_{t-2})$		[0.001]							[0.002]			[0.002]
		(0.001)							(0.002)			(0.002)
Third lag distance			-0.001	-0.001	-0.001	-0.001	-0.001	-0.001		-0.002	-0.001	-0.002
$(\mathbf{d}_i\mathbf{a}_{t-3})$			[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]		[0.001]	[0.001]	[0.002]
			(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)		(0.001)	(0.001)	(0.002)
No. of towns	15,969	15,958	15,950	15,950	15,950	15,950	15,950	15,950	15,958	15,950	15,950	15,950
Max. no. of periods	32	31	30	30	30	30	30	30	31	30	30	30
Observations	519,475	503,506	487,548	487,548	487,548	487,548	487,548	487,548	503,506	487,548	487,548	487,548
Kleibergen-Paap Wald	26.5	27.3	35.4	5.0	26.5	35.5	25.2	4.1	3.0	53.7	31.5	2.2
J-test statistic	1.26	4.01	0.94	1.47	1.25	1.03	1.13	1.79	2.90	2.14	1.31	2.62
J-test p-value	0.867	0.405	0.918	0.962	0.940	0.905	0.951	0.971	0.941	0.710	0.860	0.998
Andrews-Lu (2001) criterion	-13.04	-10.29	-18.13	-17.60	-22.58	-18.04	-22.70	-22.04	-6.63	-7.39	-12.99	-16.45

Table A.6: The Effect of Information along the Rail and Telegraph Networks: 3-day Periods. The table presents panel IV estimates of competing lag-structure specifications of equation (1) 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.7 reports the first-stage F-statistics and p-values 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 Stage F-statistics

First stages:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
First lag rail	58.42			41.56		46.98	36.64	34.68	34.42			21.61
($\mathbf{r}_{i,t}\mathbf{a}_{t-1}$)	0.000			0.000		0.000	0.000	0.000	0.000			0.000
Second lag rail		61.88	54.18	39.03	42.31			32.70	32.05		53.82	24.61
($\mathbf{r}_{i,t-1}\mathbf{a}_{t-2}$)		0.000	0.000	0.000	0.000			0.000	0.000		0.000	0.000
Third lag rail										62.46		24.16
($\mathbf{r}_{i,t-2}\mathbf{a}_{t-3}$)										0.000		0.000
First lag telegraph	14.91		24.13	18.46	111.90	23.21	113.21	86.09	73.59	25.29		62.74
($\boldsymbol{\gamma}_i\mathbf{a}_{t-1}$)	0.000		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000
Second lag telegraph		16.72			57.03		54.02	45.36	41.34			40.31
($\boldsymbol{\gamma}_i\mathbf{a}_{t-2}$)		0.000			0.000		0.000	0.000	0.000			0.000
Third lag telegraph											15.17	38.14
($\boldsymbol{\gamma}_i\mathbf{a}_{t-3}$)											0.000	0.000
First lag distance	107.14								1193.35			887.12
($\mathbf{d}_i\mathbf{a}_{t-1}$)	0.000								0.000			0.000
Second lag distance		113.79							443.61			1161.43
($\mathbf{d}_i\mathbf{a}_{t-2}$)		0.000							0.000			0.000
Third lag distance			66.74	76.24	86.44	77.60	93.77	82.60		85.11	84.75	1230.24
($\mathbf{d}_i\mathbf{a}_{t-3}$)			0.000	0.000	0.000	0.000	0.000	0.000		0.000	0.000	0.000

Table A.7: The Effect of Information along the Rail and Telegraph Networks: 3-day Period First Stages. The table presents the first-stage F-statistics and p-values corresponding to each column of the IV models reported in Table A.6. The statistics for each first stage, from top to bottom, are reported in the same order as the endogenous regressors appear in Table A.6. 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
Fully connected Rail Network under Alternative Lag Specifications

Dependent Variable:	Any Crusade Activity a_{it} -Meetings, Petitions, Marches-					
Second stages:	(1)	(2)	(3)	(4)	(5)	(6)
First lag rail	0.0029					-0.0007
($\mathbf{r}_{i,t}\mathbf{a}_{t-1}$)	(0.0005)					(0.0006)
Second lag rail		0.0037		0.0023	0.0024	0.0007
($\mathbf{r}_{i,t-1}\mathbf{a}_{t-2}$)		(0.0006)		(0.0007)	(0.0007)	(0.0011)
Third lag rail			0.0031			0.0025
($\mathbf{r}_{i,t-2}\mathbf{a}_{t-3}$)			(0.0006)			(0.0011)
First lag telegraph	0.0985			0.0408		0.2185
($\boldsymbol{\gamma}_i\mathbf{a}_{t-1}$)	(0.0514)			(0.0774)		(0.1051)
Second lag telegraph		0.1383		0.1007		-0.1253
($\boldsymbol{\gamma}_i\mathbf{a}_{t-2}$)		(0.0613)		(0.0591)		(0.1152)
Third lag telegraph			0.1393		0.1563	0.1335
($\boldsymbol{\gamma}_i\mathbf{a}_{t-3}$)			(0.0533)		(0.0533)	(0.0660)
First lag distance	-0.0003					0.0001
($\mathbf{d}_i\mathbf{a}_{t-1}$)	(0.0004)					(0.0006)
Second lag distance		0.0001				-0.0009
($\mathbf{d}_i\mathbf{a}_{t-2}$)		(0.0004)				(0.0010)
Third lag distance			0.0013	0.0019	0.0018	0.0016
($\mathbf{d}_i\mathbf{a}_{t-3}$)			(0.0004)	(0.0005)	(0.0005)	(0.0007)
No. of towns	15,934	15,934	15,934	15,934	15,934	15,934
Max. no. of periods	16	16	16	16	16	16
Observations	267,247	267,247	267,247	267,247	267,247	267,247
Kleibergen-Paap Wald	30.2	26.7	20.7	8.8	51.4	5.3
J-test statistic	20.62	37.10	12.01	21.11	22.70	25.97
J-test p-value	0.000	0.000	0.017	0.001	0.000	0.011

Table A.8: The Effect of Information along the Rail and Telegraph Networks: Fully Connected Rail Network. The table presents panel IV estimates of equation (1) under the alternative fully-connected rail network described in page 24, for a variety of lag 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 are clustered at the town level. All columns use the benchmark 5-day interval period definition. Appendix Table A.9 reports the first stage F-statistics and p-values corresponding to each endogenous regressor in the corresponding column, from top to bottom.

Causal Effects of Crusade Signals along the Railroad and Telegraph Networks
Fully connected Rail Network under Alternative Lag Specifications, First Stage F-statistics

First stages:	(1)	(2)	(3)	(4)	(5)	(6)
First lag rail	1299.06					1137.07
($\mathbf{r}_{i,t}\mathbf{a}_{t-1}$)	0.000					0.000
Second lag rail		1192.32		927.92	1165.66	945.58
($\mathbf{r}_{i,t-1}\mathbf{a}_{t-2}$)		0.000		0.000	0.000	0.000
Third lag rail			1159.83			967.09
($\mathbf{r}_{i,t-2}\mathbf{a}_{t-3}$)			0.000			0.000
First lag telegraph	55.15			56.09		32.86
($\boldsymbol{\gamma}_i\mathbf{a}_{t-1}$)	0.000			0.000		0.000
Second lag telegraph		55.33		41.21		29.5
($\boldsymbol{\gamma}_i\mathbf{a}_{t-2}$)		0.000		0.000		0.000
Third lag telegraph			54.58		54.25	27.83
($\boldsymbol{\gamma}_i\mathbf{a}_{t-3}$)			0.000		0.000	0.000
First lag distance	937.23					1208.91
($\mathbf{d}_i\mathbf{a}_{t-1}$)	0.000					0.000
Second lag distance		925.55				1128.84
($\mathbf{d}_i\mathbf{a}_{t-2}$)		0.000				0.000
Third lag distance			1088.08	798.95	1018.22	604.71
($\mathbf{d}_i\mathbf{a}_{t-3}$)			0.000	0.000	0.000	0.000

Table A.9: The Effect of Information along the Rail and Telegraph Networks: Fully Connected Rail Network First Stages. The table presents the first-stage F-statistics and p-values corresponding to each column of the IV models reported in Table A.8. The statistics for each first stage, from top to bottom, are reported in the same order as the endogenous regressors appear in Table A.8. Following Angrist and Pischke (2008), the F-statistics are corrected for the presence of multiple endogenous regressors.

The Effect of Information along the Rail and Telegraph Networks
Robustness Exercise: Using the Entire Crusade Period

Dependent Variable:	Any Crusade Activity a_{it} -Meetings, Petitions, Marches-			
Instrument Variation:	50km accident radius		80km accident radius	
Period Definition:	5 days	3 days	5 days	3 days
	(1)	(2)	(3)	(4)
First lag rail	0.113	0.078	0.142	0.114
$(\mathbf{r}_{i,t}\mathbf{a}_{t-1})$	[0.035]	[0.034]	[0.023]	[0.028]
	(0.045)	(0.056)	(0.043)	(0.058)
First lag telegraph	0.174	0.124	0.136	0.123
$(\boldsymbol{\gamma}_i\mathbf{a}_{t-1})$	[0.037]	[0.044]	[0.033]	[0.037]
	(0.073)	(0.092)	(0.057)	(0.068)
Second lag telegraph	-0.061	-0.021	-0.029	-0.012
$(\boldsymbol{\gamma}_i\mathbf{a}_{t-2})$	[0.031]	[0.064]	[0.029]	[0.040]
	(0.063)	(0.116)	(0.054)	(0.064)
Third lag distance	0.000	-0.005	0.002	0.000
$(\mathbf{d}_i\mathbf{a}_{t-3})$	[0.001]	[0.001]	[0.001]	[0.002]
	(0.001)	(0.002)	(0.001)	(0.002)
No. of towns	15,967	15,969	15,967	15,969
Max. no. of periods	38	66	38	66
Observations	612,539	1,052,681	612,539	1,052,681
Kleibergen-Paap Wald	9.6	14.9	10.8	118.1
Panel B:	First Stages (F-statistics)			
First lag rail	14.95	14.14	12.79	9.31
$(\mathbf{r}_{i,t}\mathbf{a}_{t-1})$	0.000	0.000	0.000	0.000
First lag telegraph	46.86	57.89	75.93	68.35
$(\boldsymbol{\gamma}_i\mathbf{a}_{t-1})$	0.000	0.000	0.000	0.000
Second lag telegraph	85.65	36.94	69.44	61.10
$(\boldsymbol{\gamma}_i\mathbf{a}_{t-2})$	0.000	0.000	0.000	0.000
Third lag distance	327.07	61.83	323.82	96.90
$(\mathbf{d}_i\mathbf{a}_{t-3})$	0.000	0.000	0.000	0.000

Table A.10: The Effect of Information along the Rail and Telegraph Networks: Entire Crusade Period. The table presents IV estimates of equation (1) in a panel covering the full time period of Crusade activity (from day 1 to day 215). 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 benchmark railroad link definition, and use the lag structure identified as optimal by the Andrews and Lu (2001) test in Table A.3 (first order lag for the railroad neighbors' Crusade events, first and second order lags for the telegraph neighbors' Crusade events, and third order lag for the geographic neighbors' Crusade events). Columns (1)-(2) use the benchmark 50 km. radius definition of rail accidents for the instruments. Columns (3)-(4) use an alternative 80 km. radius definition of rail accidents for the instruments. Columns (1) and (3) use the benchmark 5-day interval period definition. Columns (2) and (4) use an alternative 3-day interval period definition. Panel B reports the first-stage F-statistics and p-values corresponding to each endogenous regressor in the corresponding column, from top to bottom.

The Effect of Information along the Rail and Telegraph Networks
Additional Robustness to Alternative Information Networks

Dependent Variable:	Any Crusade Activity a_{it} -Meetings, Petitions, Marches-						
Network:	Benchmark	Watercanals Network			Hybrid Network		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
First lag rail ($\mathbf{r}_{i,t}\mathbf{a}_{t-1}$)	0.037 (0.013)	0.036 (0.013)	0.039 (0.013)	0.039 (0.013)	0.037 (0.013)	0.037 (0.013)	0.037 (0.013)
First lag telegraph ($\boldsymbol{\gamma}_i\mathbf{a}_{t-1}$)	0.171 (0.061)	0.178 (0.061)	0.156 (0.058)	0.131 (0.055)	0.171 (0.061)	0.171 (0.061)	0.171 (0.061)
Second lag telegraph ($\boldsymbol{\gamma}_i\mathbf{a}_{t-2}$)	-0.068 (0.076)	-0.070 (0.076)	-0.051 (0.073)	-0.039 (0.072)	-0.068 (0.076)	-0.068 (0.076)	-0.068 (0.076)
Third lag distance ($\mathbf{d}_i\mathbf{a}_{t-3}$)	0.006 (0.002)	0.008 (0.002)	0.006 (0.002)	0.003 (0.001)	0.006 (0.002)	0.006 (0.002)	0.006 (0.002)
First lag watercanals ($\mathbf{w}_i\mathbf{a}_{t-1}$)		-0.0003 (0.0001)	-0.0006 (0.0001)	-0.0004 (0.0001)			
Second lag watercanals ($\mathbf{w}_i\mathbf{a}_{t-2}$)			0.0004 (0.0001)	0.0005 (0.0002)			
Third lag watercanals ($\mathbf{w}_i\mathbf{a}_{t-3}$)				-0.0001 (0.0001)			
First lag hybrid ($\mathbf{h}_{i,t}\mathbf{a}_{t-1}$)					-0.038 (0.015)	-0.037 (0.015)	-0.038 (0.015)
Second lag hybrid ($\mathbf{h}_{i,t}\mathbf{a}_{t-2}$)						-0.008 (0.003)	-0.006 (0.003)
Third lag hybrid ($\mathbf{h}_{i,t}\mathbf{a}_{t-3}$)							-0.012 (0.004)
Observations	267,247	267,247	267,247	267,247	267,247	267,247	267,247
Kleibergen-Paap Wald	112.3	68.9	68.8	92.6	112.2	112.2	112.2
J-test statistic	2.16	5.62	7.15	15.57	2.16	2.16	2.16
J-test p-value	0.827	0.467	0.413	0.049	0.827	0.827	0.826

Table A.11: Additional Robustness to Alternative Information Networks. The table presents panel IV estimates of equation (1) using alternative information networks. 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 parentheses are clustered at the town level. All columns use the benchmark railroad link definition, and the lag structure identified as optimal by the Andrews and Lu (2001) test in Table A.3 (first order lag for the railroad neighbors' Crusade events, first and second order lags for the telegraph neighbors' Crusade events, and third order lag for the geographic neighbors' Crusade events). All columns use the benchmark 50 km. radius definition of rail accidents for the instruments, and the benchmark 5-day interval period definition. Column (1) reports the benchmark estimates from column (2) in Table 2. Columns (2)-(4) progressively include higher lags of waterway-mediated Crusade events, instrumenting them with the corresponding rail-link variation of neighbors. Columns (5)-(7) progressively include higher lags of hybrid network-mediated Crusade events as exogenous control variables.

The Effect of Information along the Rail and Telegraph Networks
Robustness to Alternative Subsets of Instruments

Dependent Variable:	Any Crusade Activity a_{it} -Meetings, Petitions, Marches-						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
First lag rail	0.053	0.051	0.036	0.042	0.048	0.051	0.037
($\mathbf{r}_{i,t}\mathbf{a}_{t-1}$)	[0.018]	[0.018]	[0.013]	[0.014]	[0.077]	[0.018]	[0.013]
	(0.020)	(0.021)	(0.013)	(0.014)	(0.081)	(0.021)	(0.013)
First lag telegraph	0.177	0.164	0.179	0.171	0.166	0.165	0.171
($\boldsymbol{\gamma}_i\mathbf{a}_{t-1}$)	[0.046]	[0.049]	[0.039]	[0.049]	[0.044]	[0.033]	[0.033]
	(0.059)	(0.066)	(0.061)	(0.064)	(0.070)	(0.062)	(0.061)
Second lag telegraph	-0.086	-0.074	-0.076	-0.079	-0.065	-0.065	-0.068
($\boldsymbol{\gamma}_i\mathbf{a}_{t-2}$)	[0.038]	[0.048]	[0.035]	[0.048]	[0.034]	[0.031]	[0.031]
	(0.091)	(0.084)	(0.080)	(0.084)	(0.078)	(0.076)	(0.076)
Third lag distance	0.007	0.006	0.007	0.007	0.006	0.006	0.006
($\mathbf{d}_i\mathbf{a}_{t-3}$)	[0.002]	[0.003]	[0.002]	[0.003]	[0.003]	[0.002]	[0.002]
	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)
Observations	267,247	267,247	267,247	267,247	267,247	267,247	267,247
Kleibergen-Paap Wald	130.8	59.9	125.4	66.0	3.2	71.4	112.3
J-test statistic	–	1.10	1.28	1.08	1.38	1.44	2.16
J-test p-value	–	0.295	0.257	0.299	0.710	0.838	0.827
Panel B:	First Stages (F-statistics)						
First lag rail	40.53	39.19	56.55	62.25	8.89	27.44	38.40
($\mathbf{r}_{i,t}\mathbf{a}_{t-1}$)	0.000	0.000	0.000	0.000	0.000	0.000	0.000
First lag telegraph	39.01	75.94	119.55	77.14	87.03	76.27	76.43
($\boldsymbol{\gamma}_i\mathbf{a}_{t-1}$)	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Second lag telegraph	90.70	105.54	110.55	106.27	79.35	72.81	65.50
($\boldsymbol{\gamma}_i\mathbf{a}_{t-2}$)	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Third lag distance	258.05	190.48	194.57	189.39	217.71	196.67	174.94
($\mathbf{d}_i\mathbf{a}_{t-3}$)	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table A.12: The Effect of Information along the Rail and Telegraph Networks: Robustness to alternative subsets of instruments. The table presents panel IV estimates of equation (1) using alternative subsets of instruments. 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)-(4) use the benchmark railroad link definition, and use the lag structure identified as optimal by the Andrews and Lu (2001) test in Table A.3 (first order lag for the railroad neighbors' Crusade events, first and second order lags for the telegraph neighbors' Crusade events, and third order lag for the geographic neighbors' Crusade events). All columns use the benchmark 50 km. radius definition of rail accidents for the instruments, and the benchmark 5-day interval period definition. Panel B reports the first stage F-statistics and p-values corresponding to each endogenous regressor in the corresponding column, from top to bottom. Column (1) excludes $\mathbf{r}_{i,t}\boldsymbol{\iota}$, $\mathbf{r}_{i,t}\mathbf{R}_{t-1}\boldsymbol{\iota}$, $\boldsymbol{\gamma}_i\mathbf{R}_{t-1}\boldsymbol{\iota}$, $\boldsymbol{\gamma}_i\mathbf{R}_{t-2}\mathbf{R}_{t-3}\boldsymbol{\iota}$ and $\mathbf{d}_i\mathbf{R}_{t-3}\boldsymbol{\iota}$. Column (2) excludes $\mathbf{r}_{i,t}\mathbf{R}_{t-1}\mathbf{R}_{t-2}\boldsymbol{\iota}$, $\boldsymbol{\gamma}_i\mathbf{R}_{t-1}\mathbf{R}_{t-2}\boldsymbol{\iota}$, $\boldsymbol{\gamma}_i\mathbf{R}_{t-2}\mathbf{R}_{t-3}\boldsymbol{\iota}$, and $\mathbf{d}_i\mathbf{R}_{t-3}\mathbf{R}_{t-4}\boldsymbol{\iota}$ from the instrument set. Column (3) excludes $\mathbf{r}_{i,t}\boldsymbol{\iota}$, $\boldsymbol{\gamma}_i\mathbf{R}_{t-1}\boldsymbol{\iota}$, $\boldsymbol{\gamma}_i\mathbf{R}_{t-2}\boldsymbol{\iota}$, and $\mathbf{d}_i\mathbf{R}_{t-3}\boldsymbol{\iota}$ from the instrument set. Column (4) excludes $\mathbf{r}_{i,t}\mathbf{R}_{t-1}\boldsymbol{\iota}$, $\boldsymbol{\gamma}_i\mathbf{R}_{t-1}\mathbf{R}_{t-2}\boldsymbol{\iota}$, $\boldsymbol{\gamma}_i\mathbf{R}_{t-2}\mathbf{R}_{t-3}\boldsymbol{\iota}$, and $\mathbf{d}_i\mathbf{R}_{t-3}\mathbf{R}_{t-4}\boldsymbol{\iota}$ from the instrument set. Column (5) excludes $\mathbf{r}_{i,t}\mathbf{R}_{t-1}\boldsymbol{\iota}$ and $\mathbf{r}_{i,t}\mathbf{R}_{t-1}\mathbf{R}_{t-2}\boldsymbol{\iota}$ from the instrument set. Column (6) excludes $\mathbf{r}_{i,t}\mathbf{R}_{t-1}\mathbf{R}_{t-2}\boldsymbol{\iota}$ from the instrument set. Column (7) includes all nine instruments for comparison.

Weak Instruments Diagnosis: Exactly Identified Models

Endogenous Regressor Instrumented				
(1) First Lag of Rail $\mathbf{r}_{i,t}\mathbf{a}_{t-1}$	(2) First Lag of Telegraph $\gamma_i\mathbf{a}_{t-1}$	(3) Second Lag of Telegraph $\gamma_i\mathbf{a}_{t-2}$	(4) First Lag of Distance $\mathbf{d}_i\mathbf{a}_{t-3}$	(5) Kleibergen-Paap statistic
$\mathbf{r}_{i,t}\boldsymbol{\iota}$	$\boldsymbol{\gamma}_i\mathbf{R}_{t-1}\boldsymbol{\iota}$	$\boldsymbol{\gamma}_i\mathbf{R}_{t-2}\boldsymbol{\iota}$	$\mathbf{d}_i\mathbf{R}_{t-3}\boldsymbol{\iota}$	0.1
			$\mathbf{d}_i\mathbf{R}_{t-3}\mathbf{R}_{t-4}\boldsymbol{\iota}$	0.7
		$\boldsymbol{\gamma}_i\mathbf{R}_{t-2}\mathbf{R}_{t-3}\boldsymbol{\iota}$	$\mathbf{d}_i\mathbf{R}_{t-3}\boldsymbol{\iota}$	0.1
			$\mathbf{d}_i\mathbf{R}_{t-3}\mathbf{R}_{t-4}\boldsymbol{\iota}$	0.7
	$\boldsymbol{\gamma}_i\mathbf{R}_{t-1}\mathbf{R}_{t-2}\boldsymbol{\iota}$	$\boldsymbol{\gamma}_i\mathbf{R}_{t-2}\boldsymbol{\iota}$	$\mathbf{d}_i\mathbf{R}_{t-3}\boldsymbol{\iota}$	0.1
			$\mathbf{d}_i\mathbf{R}_{t-3}\mathbf{R}_{t-4}\boldsymbol{\iota}$	0.6
		$\boldsymbol{\gamma}_i\mathbf{R}_{t-2}\mathbf{R}_{t-3}\boldsymbol{\iota}$	$\mathbf{d}_i\mathbf{R}_{t-3}\boldsymbol{\iota}$	0.1
			$\mathbf{d}_i\mathbf{R}_{t-3}\mathbf{R}_{t-4}\boldsymbol{\iota}$	0.6
$\mathbf{r}_{i,t}\mathbf{R}_{t-1}\boldsymbol{\iota}$	$\boldsymbol{\gamma}_i\mathbf{R}_{t-1}\boldsymbol{\iota}$	$\boldsymbol{\gamma}_i\mathbf{R}_{t-2}\boldsymbol{\iota}$	$\mathbf{d}_i\mathbf{R}_{t-3}\boldsymbol{\iota}$	46.4
			$\mathbf{d}_i\mathbf{R}_{t-3}\mathbf{R}_{t-4}\boldsymbol{\iota}$	10.7
		$\boldsymbol{\gamma}_i\mathbf{R}_{t-2}\mathbf{R}_{t-3}\boldsymbol{\iota}$	$\mathbf{d}_i\mathbf{R}_{t-3}\boldsymbol{\iota}$	46.6
			$\mathbf{d}_i\mathbf{R}_{t-3}\mathbf{R}_{t-4}\boldsymbol{\iota}$	10.7
	$\boldsymbol{\gamma}_i\mathbf{R}_{t-1}\mathbf{R}_{t-2}\boldsymbol{\iota}$	$\boldsymbol{\gamma}_i\mathbf{R}_{t-2}\boldsymbol{\iota}$	$\mathbf{d}_i\mathbf{R}_{t-3}\boldsymbol{\iota}$	43.4
			$\mathbf{d}_i\mathbf{R}_{t-3}\mathbf{R}_{t-4}\boldsymbol{\iota}$	10.8
		$\boldsymbol{\gamma}_i\mathbf{R}_{t-2}\mathbf{R}_{t-3}\boldsymbol{\iota}$	$\mathbf{d}_i\mathbf{R}_{t-3}\boldsymbol{\iota}$	44.7
			$\mathbf{d}_i\mathbf{R}_{t-3}\mathbf{R}_{t-4}\boldsymbol{\iota}$	10.8
$\mathbf{r}_{i,t}\mathbf{R}_{t-1}\mathbf{R}_{t-2}\boldsymbol{\iota}$	$\boldsymbol{\gamma}_i\mathbf{R}_{t-1}\boldsymbol{\iota}$	$\boldsymbol{\gamma}_i\mathbf{R}_{t-2}\boldsymbol{\iota}$	$\mathbf{d}_i\mathbf{R}_{t-3}\boldsymbol{\iota}$	49.9
			$\mathbf{d}_i\mathbf{R}_{t-3}\mathbf{R}_{t-4}\boldsymbol{\iota}$	123.5
		$\boldsymbol{\gamma}_i\mathbf{R}_{t-2}\mathbf{R}_{t-3}\boldsymbol{\iota}$	$\mathbf{d}_i\mathbf{R}_{t-3}\boldsymbol{\iota}$	49.9
			$\mathbf{d}_i\mathbf{R}_{t-3}\mathbf{R}_{t-4}\boldsymbol{\iota}$	121.6
	$\boldsymbol{\gamma}_i\mathbf{R}_{t-1}\mathbf{R}_{t-2}\boldsymbol{\iota}$	$\boldsymbol{\gamma}_i\mathbf{R}_{t-2}\boldsymbol{\iota}$	$\mathbf{d}_i\mathbf{R}_{t-3}\boldsymbol{\iota}$	53.5
			$\mathbf{d}_i\mathbf{R}_{t-3}\mathbf{R}_{t-4}\boldsymbol{\iota}$	130.8
		$\boldsymbol{\gamma}_i\mathbf{R}_{t-2}\mathbf{R}_{t-3}\boldsymbol{\iota}$	$\mathbf{d}_i\mathbf{R}_{t-3}\boldsymbol{\iota}$	53.2
			$\mathbf{d}_i\mathbf{R}_{t-3}\mathbf{R}_{t-4}\boldsymbol{\iota}$	134.7

Table A.13: Weak Instrument Diagnosis across Exactly Identified Models: Kleibergen-Paap Wald rk F statistics. The table presents the Kleibergen-Paap Wald rk F statistics corresponding to the twenty-four exactly identified models on the benchmark specification with four endogenous regressors including the first lag of rail, first and second lags of telegraph, and third lag of distance.

Weak Instruments-Robust Inference (Andrews, 2018)

	Allowing the coefficients on all four endogenous regressors to be weakly identified				Assuming the coefficients on $\boldsymbol{\gamma}_i \mathbf{a}_{t-2}$ and $\mathbf{d}_i \mathbf{a}_{t-3}$ are strongly identified			
	Exactly identified model Col. (1) in Table A.12		Over-identified model Col (7) in Table A.12		Exactly identified model Col. (1) in Table A.12		Over-identified model Col (7) in Table A.12	
	(1) Non-robust CS	(2) Robust CS	(3) Non-robust CS	(4) Robust CS	(5) Non-robust CS	(6) Robust CS	(7) Non-robust CS	(8) Robust CS
First lag rail ($\mathbf{r}_{i,t} \mathbf{a}_{t-1}$)	[0.017, 0.089]	[0.022, 0.084]	[0.011, 0.064]	[0.014, 0.060]	[0.017, 0.089]	[0.018, 0.091]	[0.011, 0.064]	[0.012, 0.065]
First lag telegraph ($\boldsymbol{\gamma}_i \mathbf{a}_{t-1}$)	[0.059, 0.295]	[0.075, 0.319]	[0.054, 0.288]	[0.071, 0.313]	[0.059, 0.295]	[0.088, 0.278]	[0.054, 0.288]	[0.106, 0.260]
Second lag telegraph ($\boldsymbol{\gamma}_i \mathbf{a}_{t-2}$)	[-0.257, 0.086]	[-0.293, 0.062]	[-0.209, 0.073]	[-0.238, 0.053]	-	-	-	-
Third lag distance ($\mathbf{d}_i \mathbf{a}_{t-3}$)	[0.003, 0.012]	[0.004, 0.011]	[0.003, 0.010]	[0.003, 0.010]	-	-	-	-

Table A.14: Weak Instruments-Robust Inference. The table presents weak instrument-robust confidence sets for a number of exactly identified model specifications reported in Table A.12, based on the methodology proposed by (Andrews, 2018). The distortion level γ is 5% in all columns. In columns (1)-(4) the confidence sets are based on a $(30 \times 30 \times 30 \times 30)$-size grid, and all four endogenous regressors are allowed to be weakly identified. In columns (5)-(8) the confidence sets are based on a (100×100)-size grid, and only the first lag of rail and the first lag of telegraph are allowed to be weakly identified. Columns (1)-(2) and (5)-(6) report confidence sets for the exactly identified model in column (1) of Table A.12. Columns (3)-(4) and (7)-(8) report confidence sets for the over-identified model in column (7) of Table A.12.

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 $i \rightarrow j$	-0.117	-0.189		
	(0.050)	(0.060)		
Telegraph network path length $i \rightarrow j$			-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	50,076	50,076	402	402

Table A.15: 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-times-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 and standard errors on the railroad and telegraph network path-length variables are multiplied by 1000. The coefficients and standard errors on the geographic distance between towns are in km. and multiplied by 10^5. The coefficients and standard errors on the betweenness centrality statistic are multiplied by 10^6.

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)
Rail and Telegraph	0.081	0.081	0.086	0.060	0.072	0.074	0.047	0.059	0.069
$r_{ij}\gamma_j$	(0.034)	(0.026)	(0.027)	(0.027)	(0.021)	(0.021)	(0.025)	(0.020)	(0.021)
Rail and No Telegraph	0.0016	0.0045	0.0047	0.0021	0.0035	0.0038	0.0026	0.0036	0.0039
$r_{ij}(1-\gamma_j)$	(0.004)	(0.004)	(0.003)	(0.0034)	(0.0032)	(0.0028)	(0.0031)	(0.0028)	(0.0025)
Network Interaction	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	79,134	79,134	79,134	193,745	193,745	193,745	407,618	407,618	407,618

Table A.16: Rail and Telegraph Technological Interaction Effects: Alternative Cluster Radii Event Studies. The table presents estimation results of the cluster event study approach based on equation (2) 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 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 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)
Rail and Telegraph	0.122	0.119	0.144	0.084	0.081	0.007	0.108	0.079
$r_{ij}\gamma_j$	(0.042)	(0.034)	(0.057)	(0.061)	(0.156)	(0.121)	(0.052)	(0.045)
Rail and No Telegraph	-0.0006	0.0029	-0.0091	-0.0075	0.0002	0.0045	-0.0014	0.0017
$r_{ij}(1-\gamma_j)$	(0.005)	(0.004)	(0.005)	(0.004)	(0.0117)	(0.0095)	(0.0057)	(0.0066)
$r_{ij}\gamma_j\times$ Covariate	-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)
$r_{ij}(1-\gamma_j)\times$ Covariate	-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)
Network Interaction	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)
Network Interaction $\times$ Covariate	-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	29,592	79,133	29,592	79,134	29,590	79,129	29,592	79,133

Table A.17: Testing for Heterogeneity in the Rail-Telegraph Interaction Effects. The table presents estimation results of the cluster event-study approach based on equation (2), 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 sex ratio. 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 network 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: 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)
Rail and Telegraph	0.358	-0.021	0.022	-0.064	0.196	-0.018	-0.236	0.322	-0.214	-0.014
$r_{ij}\gamma_j$	(0.139)	(0.019)	(0.133)	(0.054)	(0.220)	(0.010)	(3.760)	(0.160)	(0.098)	(0.071)
Rail and No Telegraph	0.0013	-0.0044	0.0078	-0.0163	0.0706	-0.0206	-0.0170	0.0131	-0.0080	0.0291
$r_{ij}(1-\gamma_j)$	(0.046)	(0.005)	(0.041)	(0.009)	(0.039)	(0.008)	(0.0188)	(0.0046)	(0.0068)	(0.0190)
No Rail and Telegraph	0.154	0.0191								
$(1-r_{ij})\gamma_j$	(0.022)	(0.007)								
Network Interaction	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	20,399	20,399	20,399	20,399	52,399	52,399	123,140	123,140	238,097	238,097

Table A.18: Rail-Telegraph Interaction Effects: Placebo Event Studies using Close Match Signal-Generating Towns. The table presents estimation results of the cluster event-study approach based on equation (2), 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, sex ratio, alcohol vendors per capita, religious ascriptions Herfindahl index, Presbyterian sittings per capita, and log population. The dependent variable is a dummy 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 network 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 network 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-Telegraph Complementarities: Transitions to a First Crusade Event

	2 weeks			3 weeks			4 weeks		
	Meeting	Petition	March	Meeting	Petition	March	Meeting	Petition	March
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Railroad and No Telegraph	0.15	-0.10	-0.36	0.23	0.41	-0.35	0.25	0.64	-0.47
$r_{ij}(1-\gamma_j)$	(0.16)	(0.42)	(0.25)	(0.14)	(0.35)	(0.20)	(0.13)	(0.26)	(0.19)
No Railroad and Telegraph	0.90	1.07	1.74	0.88	1.35	1.83	0.87	1.61	1.80
(1- r_{ij})γ_j	(0.18)	(0.45)	(0.25)	(0.17)	(0.37)	(0.21)	(0.16)	(0.31)	(0.20)
Railroad and Telegraph	1.63	0.16	2.33	1.78	0.79	2.41	1.99	0.23	2.54
$r_{ij}\gamma_j$	(0.31)	(0.94)	(0.51)	(0.29)	(0.77)	(0.45)	(0.29)	(0.74)	(0.41)
Signal-recipient distance	0.03	0.34	-0.08	0.11	0.72	-0.14	0.14	0.39	-0.16
	(0.09)	(0.29)	(0.16)	(0.08)	(0.28)	(0.13)	(0.08)	(0.23)	(0.12)
If the signal is meeting	-0.22	0.03	-0.23	-0.17	0.17	-0.17	-0.14	0.05	-0.16
	(0.12)	(0.29)	(0.17)	(0.10)	(0.24)	(0.14)	(0.10)	(0.10)	(0.13)
If the signal is petition	-0.16	0.180	0.001	-0.15	0.08	0.001	-0.17	-0.12	0.03
	(0.15)	(0.31)	(0.20)	(0.14)	(0.28)	(0.17)	(0.14)	(0.25)	(0.15)
State FE		Yes			Yes			Yes	
Observations		28,168			28,168			28,168	

Table A.19: Rail-Telegraph Interaction Effects: Transitions to a First Crusade Event The table presents estimation results of the cluster event-study approach based on equation (2). The dependent variable classifies four un-ordered binary responses: i) whether a town within the cluster radius experienced a no event within the time window, ii) whether a town within the cluster radius experienced a meeting within the time window, iii) whether a town within the cluster radius experienced a petition within the time window, and iv) whether a town within the cluster radius experienced a march within the time window. Effects are reported relative to the baseline no event category. Columns (1)-(3) report coefficient estimates in the logistic probabilities for the models using a 2-week window event study. Columns (4)-(6) report coefficient estimates in the logistic probabilities for the models using a 3-week window event study. Columns (7)-(9) report coefficient estimates in the logistic probabilities for the models using a 4-week window event study. All models include state fixed effects, signal in original fixed effects, signal date fixed effects, and the distance between generating and recipient towns. Standard errors are robust and clustered at the event-cluster town levels.

B Online Appendix: Selection and Mis-classification

Here we discuss formally the possibility (and the econometric implications) of measurement error in our data on Temperance Crusade events. We then present some empirical evidence to assess the quantitative effect measurement error may have on our main estimates. Two main issues are a possibility in our setting: i) Selection into our data set, which, moreover, may be correlated with the network structure. For example, some protests may have happened but were never recorded in the sources that historians (and we) used to build our data set. If this is the case, we might expect its likelihood to depend on railroad and telegraph access. As we will illustrate here, this is a form of measurement error "from the right" (it may bias our IV estimates by generating mis-classification in the explanatory variables).[1] ii) Accidents and other exogenous disruptions in the railroad network may be correlated with measurement error in our Crusade event data. For example, the newspaper reporting of train accidents may have crowded out their reporting on simultaneous Crusade events, subsequently reducing the likelihood that these events appear in our data set. As we will illustrate here, this is a form of measurement error "from the left" (it may bias our IV estimates by generating mis-classification in the outcome variable, i.e., it is a violation of the instrument exclusion restriction).

To illustrate these econometric issues, we introduce some notation. Define $a^*_{i,t}$ to be a dummy variable taking the value of 1 if a crusade event took place in town i at time t. Define $a_{i,t}$ to be a dummy variable taking the value of 1 if a crusade event is recorded in our data set for town i at time t. In our setting, it is safe to assume that there is no "upwards mis-classification": $\mathbb{P}(a_{i,t} = 1|a^*_{i,t} = 0) = 0$. This is, if a Crusade event did not happen, our data set will never record a Crusade event as having happened. On the other hand, there may be "downward mis-classification": $\mathbb{P}(a_{i,t} = 0|a^*_{i,t} = 1) \equiv \alpha_i \geq 0$. This is, Crusade events that did happen may appear in our data set as not having happened. We index the probability of downward mis-classification by i to emphasize that towns with different characteristics (i.e., network access) may have different mis-classification probabilities. We present our discussion (whose conclusions all generalize) in a simplified version of equation (1), where only a railroad network is in place, town i has just two neighbors ($R_i = \{j, k\}$), and railroad-mediated information matters only at lag 1. Because we estimate town-fixed effects models that effectively average over the time series-variation town by town, consider the time-series model for town i (because our panel is almost perfectly balanced, the coefficient estimates from the panel regression effectively weight each town's "own regression" coefficient estimate almost uniformly):

$$a^*_{i,t} = \beta_0 + \beta_r(r_{ij,t}a^*_{j,t-1} + r_{ik,t}a^*_{k,t-1}) + \epsilon_{i,t} \tag{B.1}$$

[1]Notice that this is not an attrition problem, as our data set is a panel of all existing towns in 1870.

The feasible regression, however, uses $a_{h,t}$ instead of $a^*_{h,t}$ for $h = i, j, k$. Consider an IV estimator of this model using $(r_{j\ell,t-1} + r_{km,t-1})$ as an instrument (suppose for simplicity that j has just one other neighbor, ℓ, other than i, and k has just one other neighbor, m, other than i). The probability limit of the IV estimator is

$$\beta_r^{IV} = \frac{\mathrm{Cov}(a_{i,t}, r_{j\ell,t-1} + r_{km,t-1})}{\mathrm{Cov}(r_{ij,t}a_{j,t-1} + r_{ik,t}a_{k,t-1}, r_{j\ell,t-1} + r_{km,t-1})} \tag{B.2}$$

Denote as w_{it} the random variable define by $a_{i,t} = a^*_{i,t} + w_{i,t}$. The conditional distribution of w_{it} is thus

$$\mathbb{P}(w_{i,t} = 0 | a^*_{i,t} = 0) = 1 \tag{B.3}$$

$$\mathbb{P}(w_{i,t} = 0 | a^*_{i,t} = 1) = 1 - \alpha_i \tag{B.4}$$

$$\mathbb{P}(w_{i,t} = -1 | a^*_{i,t} = 1) = \alpha_i \tag{B.5}$$

Consider first the numerator in equation (B.2). It is equal to

$$N_i \mathrm{Cov}(a^*_{i,t}, r_{j\ell,t-1}) + N_i \mathrm{Cov}(w_{i,t}, r_{j\ell,t-1}) \tag{B.6}$$

where N_i is the number of rail neighbors of town i (in our example, $N_i = 2$), and j represents any rail neighbor of town i. Notice that the first term in (B.6) is the covariance of i's Crusade activity with the rail link variation of i's neighbors with their neighbors. The second term is the covariance of i's measurement error with rail link activity of i's neighbors with their neighbors. Consider now the first covariance in this expression, using equation (B.1):

$$\mathrm{Cov}(a^*_{i,t}, r_{j\ell,t-1}) = \beta_r \mathrm{Cov}(r_{ij,t}a^*_{j,t-1}, r_{j\ell,t-1}) + \beta_r \mathrm{Cov}(r_{ik,t}a^*_{k,t-1}, r_{j\ell,t-1}) + \mathrm{Cov}(\epsilon_{i,t}, r_{j\ell,t-1}) \tag{B.7}$$

Our instrument exclusion restriction implies that the last covariance of this expression is zero. For the first covariance of (B.7), using iterated expectations,

$$\mathrm{Cov}(r_{ij,t}a^*_{j,t-1}, r_{j\ell,t-1}) = \mathbb{E}\left[\mathbb{E}\left[r_{ij,t}r_{j\ell,t-1}a^*_{j,t-1}|a^*_{j,t-1}\right]\right] - \mathbb{E}\left[\mathbb{E}\left[r_{ij,t}a^*_{j,t-1}|a^*_{j,t-1}\right]\right]\mathbb{E}\left[r_{j\ell,t-1}\right]$$

Notice that for the first inner expectation, the covariation in rail link activity between contiguous links does not depend on previous Crusade activity. For the second inner expectation, similarly, rail link activity does not depend on previous Crusade activity of neighbors. Thus, we have

$$\begin{aligned}\mathrm{Cov}(r_{ij,t}a^*_{j,t-1}, r_{j\ell,t-1}) &= \mathbb{E}\left[r_{ij,t}r_{j\ell,t-1}\right]\mathbb{E}\left[a^*_{j,t-1}\right] - \mathbb{E}\left[r_{ij,t}\right]\mathbb{E}\left[a^*_{j,t-1}\right]\mathbb{E}\left[r_{j\ell,t-1}\right] \\ &= \mathrm{Cov}(r_{ij,t}, r_{j\ell,t-1})\mathbb{P}(a^*_{j,t-1} = 1) \\ &= \rho_1\sigma_r\mathbb{P}(a^*_{j,t-1} = 1)\end{aligned}$$

where ρ_1 is defined as the correlation in rail link activity between pairs of contiguous links, and σ_r^2 is the variance of rail link activity. For the second covariance in (B.7), an analogous argument shows that

$$\mathrm{Cov}(r_{ik,t}a^*_{k,t-1}, r_{j\ell,t-1}) = \rho_2\sigma_r\mathbb{P}(a^*_{k,t-1} = 1)$$

where ρ_2 is defined as the correlation in rail link activity between pairs of links one link apart (finally notice that j and k here stand for any two neighbors of i). Combining these results, we have that

$$\mathrm{Cov}(a^*_{i,t}, r_{j\ell,t-1}) = \beta_r\mathbb{P}(a^*_{j,t-1} = 1)\sigma_r(\rho_1 + \rho_2)$$

We can now look at the second term in (B.6):

$$\mathrm{Cov}(w_{i,t}, r_{j\ell,t-1}) = \mathbb{E}\left[w_{i,t}r_{j\ell,t-1}\right] - \mathbb{E}\left[w_{i,t}\right]\mathbb{E}\left[r_{j\ell,t-1}\right]$$

Using iterated expectations,

$$\begin{aligned}\mathbb{E}\left[w_{i,t}r_{j\ell,t-1}\right] &= \mathbb{E}\left[\mathbb{E}\left[w_{i,t}r_{j\ell,t-1}|a^*_{i,t}\right]\right] \\ &= \mathbb{E}\left[w_{i,t}r_{j\ell,t-1}|a^*_{i,t} = 0\right]\mathbb{P}(a^*_{i,t} = 0) + \mathbb{E}\left[w_{i,t}r_{j\ell,t-1}|a^*_{i,t} = 1\right]\mathbb{P}(a^*_{i,t} = 1) \\ &= \mathbb{E}\left[w_{i,t}r_{j\ell,t-1}|a^*_{i,t} = 1\right]\mathbb{P}(a^*_{i,t} = 1)\end{aligned}$$

where the last line follows by noticing that $w_{i,t} = 0$ whenever $a^*_{i,t} = 0$. The same argument also implies that

$$\mathbb{E}\left[w_{i,t}\right] = \mathbb{P}(a^*_{i,t} = 1)\mathbb{E}\left[w_{i,t}|a^*_{i,t} = 1\right]$$

Putting these together,

$$\mathrm{Cov}(w_{i,t}, r_{j\ell,t-1}) = \mathbb{P}(a^*_{i,t} = 1)\mathrm{Cov}(w_{i,t}, r_{j\ell,t-1}|a^*_{i,t} = 1)$$

Collecting all these results, we conclude that the numerator of the IV estimator is

$$N_i\beta_r\mathbb{P}(a^*_{j,t-1} = 1)\sigma_r(\rho_1 + \rho_2) + N_i\mathbb{P}(a^*_{i,t} = 1)\mathrm{Cov}(w_{i,t}, r_{j,t-1}|a^*_{i,t} = 1). \tag{B.8}$$

Consider now the denominator of (B.2). It is equal to

$$\mathrm{Cov}(r_{ij,t}a_{j,t-1}, r_{j\ell,t-1}) + \mathrm{Cov}(r_{ij,t}a_{j,t-1}, r_{km,t-1}) + \mathrm{Cov}(r_{ik,t}a_{k,t-1}, r_{j\ell,t-1}) + \mathrm{Cov}(r_{ik,t}a_{k,t-1}, r_{km,t-1})$$

Notice that the first and fourth terms represent the same covariance (between rail-mediated information about i's neighbors' Crusade activity and those neighbor's rail link activity with their own neighbors), and that the second and third terms represent the same covariance (between rail-mediated information about i's neighbors' Crusade activity and i's other neighbors' rail link activity with their own neighbors). Thus, assuming for simplicity (we will relax this assumption below) that i's neighbors mis-classification rates are the same ($\alpha_j = \alpha_k$), the denominator takes the form

$$N_i\mathrm{Cov}(r_{ij,t}a_{j,t-1}, r_{j\ell,t-1}) + N_i\mathrm{Cov}(r_{ij,t}a_{j,t-1}, r_{km,t-1}) \tag{B.9}$$

Using iterated expectations, the first covariance in this expression can be written as

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{E}\left[r_{ij,t}a_{j,t-1}r_{j\ell,t-1}|a^*_{j,t-1}\right]\right] - \mathbb{E}\left[\mathbb{E}\left[r_{ij,t}a_{j,t-1}|a^*_{j,t-1}\right]\right]\mathbb{E}\left[r_{j\ell,t-1}\right] \\
=&\mathbb{P}(a^*_{j,t-1}=1)\mathbb{E}\left[r_{ij,t}r_{j\ell,t-1}a_{j,t-1}|a^*_{j,t-1}=1\right] - \mathbb{P}(a^*_{j,t-1}=1)\mathbb{E}\left[r_{ij,t}a_{j,t-1}|a^*_{j,t-1}=1\right]\mathbb{E}\left[r_{j\ell,t-1}\right] \\
=&\mathbb{P}(a^*_{j,t-1}=1)\left(\mathbb{E}\left[r_{ij,t}r_{j\ell,t-1}a^*_{j,t-1}|a^*_{j,t-1}=1\right] + \mathbb{E}\left[r_{ij,t}r_{j\ell,t-1}w_{j,t-1}|a^*_{j,t-1}=1\right]\right) \\
&+\mathbb{P}(a^*_{j,t-1}=1)\left(\mathbb{E}\left[r_{ij,t}a^*_{j,t-1}|a^*_{j,t-1}=1\right] + \mathbb{E}\left[r_{ij,t}w_{j,t-1}|a^*_{j,t-1}=1\right]\right)\mathbb{E}\left[r_{j\ell,t-1}\right]
\end{aligned}$$

where the second line follows because $a_{j,t-1} = 0$ whenever $a^*_{j,t-1} = 0$, and the third line follows from the definition of $w_{j,t}$. Notice now that whether a Crusade event was mis-classified in period $t-1$ cannot depend on rail link activity of contemporary or subsequent periods (it can only depend on rail link disruptions in previous periods). This implies that we can break the conditional expectations involving $w_{j,t-1}$ from the expression above to obtain

$$\begin{aligned}
&\mathbb{P}(a^*_{j,t-1}=1)\left(\mathbb{E}\left[r_{ij,t}r_{j\ell,t-1}|a^*_{j,t-1}=1\right] - \alpha_j\mathbb{E}\left[r_{ij,t}r_{j\ell,t-1}|a^*_{j,t-1}=1\right]\right) \\
&+\mathbb{P}(a^*_{j,t-1}=1)\left(\mathbb{E}\left[r_{ij,t}|a^*_{j,t-1}=1\right] - \alpha_j\mathbb{E}\left[r_{ij,t}|a^*_{j,t-1}=1\right]\right)\mathbb{E}\left[r_{j\ell,t-1}\right] \\
&=\mathbb{P}(a^*_{j,t-1}=1)(1-\alpha_j)\mathrm{Cov}(r_{ij,t}, r_{j\ell,t-1}) \\
&=\mathbb{P}(a^*_{j,t-1}=1)(1-\alpha_j)\rho_1\sigma_r
\end{aligned}$$

which follows from $\mathbb{E}\left[w_{j,t-1}|a^*_{j,t-1}=1\right] = -\alpha_j$.

An analogous argument implies that the second covariance in (B.9) is

$$\mathbb{P}(a^*_{j,t-1}=1)(1-\alpha_j)\rho_2\sigma_r.$$

Putting these together, the denominator of (B.2) is

$$N_i \mathbb{P}(a^*_{j,t-1} = 1)(1-\alpha_j)\sigma_r(\rho_1 + \rho_2). \tag{B.10}$$

Replacing (B.8) and (B.10) in (B.2), the IV estimator is

$$\beta_r^{IV} = \frac{1}{1-\alpha_j}\beta_r + \frac{\mathrm{Cov}(w_{i,t}, r_{j\ell,t-1} | a^*_{i,t} = 1)}{(1-\alpha_j)\sigma_r(\rho_1 + \rho_2)}$$

Notice also that

$$\mathrm{Cov}(w_{i,t}, r_{j\ell,t-1} | a^*_{i,t} = 1) = \rho_{wr}\sigma_r\sqrt{\mathrm{Var}(w_{i,t}|a^*_{i,t} = 1)},$$

where ρ_{wr} is the correlation between mis-classification of Crusade events in our data and railroad accidents. Moreover, $\mathrm{Var}(w_{i,t}|a^*_{i,t} = 1) = \alpha_i(1-\alpha_i)$. Recall that to simplify our derivation above we assumed all neighbors of i had the same mis-classification rate α_j. More generally, these may vary with town characteristics such as network access. Allowing for α to vary across i's neighbors, the expression for the probability limit of β_r^{IV} depends instead on the average of these mis-classification rates ($\overline{\alpha}_i \equiv (1/N_i)\sum_{j\in R_i}\alpha_j$). Thus, we obtain

$$\beta_r^{IV} = \pi_i\beta_r + \frac{\rho_{wr}}{\rho_1 + \rho_2}\pi_i\sqrt{\alpha_i(1-\alpha_i)} \tag{B.11}$$

where $\pi_i \equiv 1/(1-\overline{\alpha}_i)$. Finally, averaging across i, the fixed effects IV estimator is

$$\beta_r^{IVFE} = \overline{\pi}\beta_r + \frac{\rho_{wr}}{\rho_1 + \rho_2}\overline{\pi\sqrt{\alpha_i(1-\alpha_i)}} \tag{B.12}$$

Several points are worth discussing about equation (B.11), which clearly illustrate the potential sources of bias in this setting: i) Notice that $\pi_i \geq 1$ is an inflation factor, and cannot switch the sign of the IV estimator relative to the true coefficient. This is the standard bias from mis-classification of a binary regressor (see DiTraglia and García-Jimeno (2019)), and constitutes the source of bias that arises if some of the Crusade activity failed to be recorded in our data set. Variation in mis-classification rates across towns matters only insomuch as this changes the average mis-classification rate across neighboring towns. Notice, however, that only the average of these inflation factors across towns matters for the fixed effects IV estimator. Moreover, this inflation factor will be bounded above by the largest mis-classification rate across all of i's neighbors. ii) The second term in (B.12) is the source of bias that arises from a particular form of violation of our instrument exclusion restriction: when railroad accidents affect the likelihood that, for example, the media reports on Crusade events, and this leads to those events not being recorded in our data set, $\rho_{wr} \neq 0$. Notice, moreover, that we expect $\rho_{wr} > 0$: an active rail link ($r_{ij,t-1} = 1$) makes it more likely that a given Crusade event appears in our data set ($w_{i,t} = 0$), while a disrupted rail link ($r_{ij,t-1} = 0$) makes it more likely that a given Crusade event is not recorded ($w_{i,t} = -1$). Below we will present some indirect evidence suggesting that $\rho_{wr} \approx 0$. iii) We can compute estimates of ρ_1 and ρ_2 directly from our data: the correlation between rail activity of pairs of adjacent rail links is $\rho_1 = 0.67$, and the correlation between rail activity of pairs of rail links one link apart is $\rho_2 = 0.52$. iv) Notice also that $\Lambda(\pi) \equiv \sqrt{\alpha_i(1-\alpha_i)}$ is also bounded above.

Thus, a *lower* bound for β_r must be

$$\underline{\beta}_r = \frac{1}{\overline{\pi}}\beta_r^{IVFE} - \Lambda(\pi)\frac{\rho_{wr}}{0.67 + 0.52} \tag{B.13}$$

By pinning down the inflation factor $\overline{\pi}$ and the correlation between railroad access and mis-classification ρ_{wr}, we can then bound the bias of our estimator.

B.1 Backing out mis-classification using newspaper reporting

We first discuss the possibility that Blocker (1985), our main source for Crusade activity information, may have missed Crusade events based on his newspaper and archival research. We show that under two mild assumptions we will make explicit below, we can pin down the average rate of mis-classification of towns in our data (i.e., the fraction of towns that experienced Crusade activity but were not classified by Blocker (1985) as Crusading towns) using information from newspaper reports before and during the Temperance Crusade. The idea is as follows: we know with certainty that the set of towns identified by Blocker (1985) as having experienced a Crusade are in fact all Crusade towns. In what follows we refer to these as Blocker towns. In contrast, if mis-classification is present, the remaining set of towns includes both towns that were truly not affected by Crusade activity, and towns where Crusade events did take place. We refer to these as non-Blocker towns. Prior to the beginning of the Crusade, truly Crusading towns (c), and truly non-crusading towns (n) may have been reported on newspapers at different baseline rates. With the onset of the Crusade, newspaper coverage of Crusading towns should increase differentially more for Crusading towns. Thus, comparing the overall rate of newspaper article mentions of Blocker towns gives us a measure of the increase in the reporting rate. Part of this differential may entail a change in overall newspaper behavior across all towns, and part may be in response to the Crusading activity. Consider now the set of towns that did not experience Crusade activity. The differential rate at which they may have been reported during the Crusade period should reflect the overall changes in newspaper behavior during that period, but not the changes directly related to the Crusade. If the set of non-Blocker towns contains a fraction of truly Crusading towns, then the increase in newspaper reporting for this group of towns should partly reflect the increased coverage of the mis-classified towns. Put another way, in the absence of mis-classification we should not expect to see a change in reporting

Period	Type of town	Crusade	Non-Crusade	Blocker Crusade	Blocker non-Crusade
		True unobserved average news reporting rates		Observed average news reporting rates under mis-classification	
Jan, 1893- Dec, 1874	Non-unique names	r_c	r_n	(A) $r_c = 0.021$	(B) $\frac{1}{1+\alpha} r_n + \frac{\alpha}{1+\alpha} r_c = 0.021$
	Unique names	$\delta_{uc} r_c$	$\delta_{un} r_n$	(C) $\delta_{uc} r_c = 0.004$	(D) $\frac{1}{1+\alpha}\delta_{un} r_n + \frac{\alpha}{1+\alpha}\delta_{uc} r_c = 0.002$
Jan, 1874- Jul, 1874	Non-unique names	$\beta_c \gamma r_c$	γr_n	(E) $\beta\gamma r_c = 0.219$	(F) $\frac{1}{1+\alpha}\gamma r_c + \frac{\alpha}{1+\alpha}\beta_c \gamma r_c = 0.169$
	Unique names	$\beta_u \gamma \delta_{uc} r_c$	$\gamma \delta_{un} r_n$	(G) $\beta_u \gamma \delta_{uc} r_c = 0.048$	(H) $\frac{1}{1+\alpha}\gamma\delta_{un} r_n + \frac{\alpha}{1+\alpha}\beta_u \gamma \delta_{uc} r_c = 0.017$

Inferring Mis-Classification based on Rates of Newspaper Reporting

Table B.1: Inferring Mis-classification: Average Newspaper Reporting Rates by Type of Town. The left panel of the table presents average rates of news reporting for four groups of towns, before and during the Temperance Crusade: Truly crusading towns with non-unique and unique names, and truly non-crusading towns with non-unique and unique names. These rates are unobserved in the presence of mis-classification. r_c is the baseline rate for truly crusading towns with non-unique names, r_n is the baseline rate for truly non-crusading towns with non-unique names, δ_{uc} is the differential rate of reporting for truly crusading towns with unique names, δ_{cn} is the differential rate of reporting for truly non-crusading towns with unique names, γ is the differential rate of reporting common to all towns during the Crusade period, β_c is the differential rate of reporting during the Crusade period for truly crusading towns with non-unique names, and β_u is the differential rate of reporting during the Crusade period for truly crusading towns with unique names. α is the average mis-classification rate (the probability that a truly crusading town was mis-classified by Blocker (1985) as non-crusading). The right panel of the table presents the corresponding observed average rates of news reporting implied by mis-classification at rate α. Each cell presents the corresponding newspaper reporting rate estimated from our text analysis exercise. Cells are labeled A through H for ease of reference. Reporting rates are computed as the fraction of newspaper articles referring to Temperance-related topics mentioning a town in the corresponding group relative to all articles referring to Temperance-related topics.

on this set of towns beyond any overall differences specific to the Crusade period; the larger the extent of mis-classification, the larger the signal we should detect in that group.

Comparing the newspaper reporting rates of Blocker and non-Blocker towns, however, faces a difficulty: a significant fraction of 1870 towns in the US share a name. In fact, from the 15971 towns in the 1870 census, there are only 2386 unique town names. Despite our best efforts, our text-scraping code is likely to make errors when distinguishing between news articles reporting on towns with the same name, either as a result of errors in the scraping itself, or because the article text does not mention the state or county of the corresponding town. We leverage this difficulty, however, by comparing news-reporting rates along a second dimension, by dividing towns into those with a unique name and those with a non-unique name. This comparison is useful because the set of towns with unique names should not face changes in its news reporting rate caused by the increased reporting of homonym towns (as they have no homonym). This allows us to implement a "triple-differences" comparison (pre-Crusade vs. during the Crusade, Blocker vs. non-Blocker, and unique name vs. non-unique name). Under the assumptions that mis-classified towns faced similar news coverage rates as correctly classified towns of similar characteristics (unique or non-unique name) before the Crusade began, and that the mis-classification probability was similar for towns with unique and non-unique names, we show that the mis-classification rate α is identified.

Table B.1 illustrates our comparison groups and the underlying newspaper reporting rates. The panel on the left presents the (unobserved) correctly classified groups, while the panel to the right presents the resulting observed mixture rates from Blocker (1985)'s classification. For

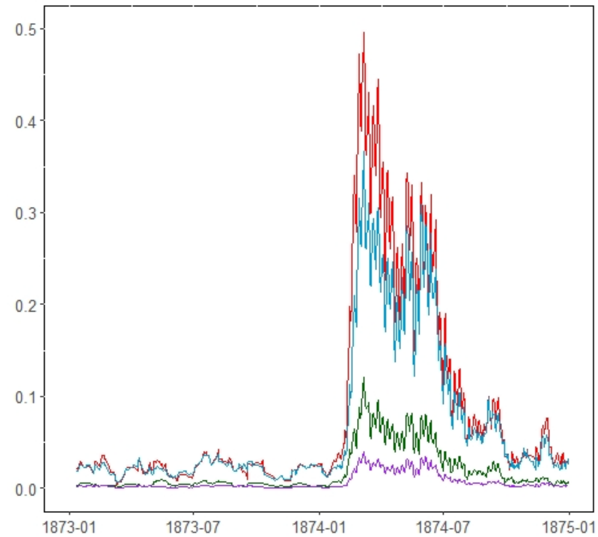


Figure B.1: Inferring Mis-classification: Rates of Newspaper Reporting over Time. The figure plots average rates of newspaper reporting of towns in articles related to temperance, between January 1873 and December 1874, based on our text-scraping of the Library of Congress' online newspaper repository. The curves are 10-day moving averages. The red curve represents towns with non-unique names classified in our data as having experienced Crusade activity. The blue curve represents towns with non-unique names classified in our data as not having experienced Crusade activity. The green curve represents towns with unique names classified in our data as having experienced Crusade activity. The purple curve represents towns with unique names classified in our data as not having experienced Crusade activity.

convenience, we have labeled the cells on the right with the letters A through H. The top panel corresponds to the pre-Crusade period (we scraped the newspaper archive for all of 1873), while the bottom panel corresponds to the Crusade period (covering the first semester of 1874). r_c denotes the baseline rate of news reporting of Crusading towns with non-unique names, and r_n the baseline rate of news reporting of non-Crusading towns with non-unique names. These are allowed to be different, as these two sets of towns likely differed along many dimensions. In turn, δ_{uc} denotes the differential rate of baseline news reporting for Crusading towns with unique names, and δ_{un} the differential rate of baseline news reporting of non-Crusading towns with unique names. γ is the overall differential rate of reporting during the Crusade period. β_c is the differential rate of reporting during the Crusade period for Crusading towns with non-unique names, while β_u is the differential rate of reporting during the Crusade period for Crusading towns with unique names. Under mis-classification at rate α, the observed news reporting rates for the set of non-Blocker towns are mixtures of the corresponding rates for truly Crusading and non-Crusading towns. The right panel of Table B.1 also reports the corresponding rates we computed based on our text scraping of the Library of Congress's online newspaper archive averaging over the relevant time period, while Figure B.1 presents the time series of these rates for each of the four groups of towns. The news reporting process appears stationary before the Crusade began. Reassuringly, average news reporting rates in this period are very similar for both types of towns with non-unique names, suggesting the plausibility of the assumptions we pointed out above. Naturally, reporting of alcohol/temperance-related articles explodes with the onset of Crusade activity, which we can see clearly in the figure. Although the increase in reporting of towns with non-unique names is much larger than for towns with unique names, the proportionate increase is larger for Blocker towns (green relative to red) compared to non-Blocker towns (purple relative to blue), suggesting a much higher signal-to-noise ratio among the towns classified as Crusading by Blocker (1985). To assess the extent of mis-classification consistent with this difference relative to the pre-Crusade period, we turn to Table B.1, where we average each curve over the pre-Crusade and Crusade periods.

Our first observation is that $A = r_c = r_n = B = 0.021$, since any $\alpha > 0$ would otherwise require cells A and B in Table B.1 to take different values. Indeed, Figure B.1 illustrates that in the pre-Crusade period, news reporting rates of Blocker and non-Blocker towns with non-unique names (red and blue curves) are pretty much identical. Not surprisingly, the average news reporting rates of towns with unique names are considerably lower. These do differ between Blocker and non-Blocker towns (although harder to see in the figure, the baseline rate for Blocker towns in green is around twice as large as the baseline rate for non-Blocker towns in purple). Notice that the ratio of A to C identifies δ_{uc}. Thus, from cell D we have

$$\frac{D}{A} + \frac{(D - C)}{A}\alpha = \delta_{un} \tag{B.14}$$

Replacing for $\beta_c \gamma r_c = E$ in the expression in cell F,

$$F = \frac{1}{1+\alpha}\gamma A + \frac{\alpha}{1+\alpha}E$$

Solving for γ,

$$\gamma = \frac{F}{A} + \frac{F - E}{A}\alpha \tag{B.15}$$

Finally, replacing for $\beta_u \gamma \delta_{uc} r_c = G$ in the expression for H,

$$H = \frac{1}{1+\alpha}\gamma\delta_{un} A + \frac{\alpha}{1+\alpha}G$$

Replacing for δ_{un} from (B.14) and solving for γ,

$$\gamma = \frac{H + (H - G)\alpha}{D + (D - C)\alpha} \tag{B.16}$$

Equating (B.15) and (B.16), we obtain a quadratic equation in α that depends only on the data moments:

$$(F - E)(D - C)\alpha^2 + [F(D - C) + D(F - E) + A(G - H)]\alpha + (FD - AH) = 0$$

The positive root is the relevant solution. Using the moments from Table B.1, we find

$$\alpha = 0.05$$

We obtained this estimate under the assumption that mis-classified and correctly classified towns faced similar news reporting rates. One may conjecture, however, that mis-classified towns may have been so precisely because they were less prominent in the news. Notice from the expression in cell B on Table B.1, that the smaller the rate multiplying $\alpha/(1 + \alpha)$, the smaller α must be to rationalize A = B. Thus, our estimate of $\alpha = 0.05$ is an *upper bound* for mis-classification, implying an upper bound for the inflation factor $\bar{\pi}$ of

$$\bar{\pi} < \frac{1}{1 - 0.05} = 1.05$$

Below we will consider the implications of this upper bound on the extent of mis-classification, and further consider an extreme scenario where mis-classification is twice as large ($\alpha = 0.1$).

B.2 Assessing the magnitude of the correlation between mis-classification and railroad accidents

Our identification strategy relies on the existence of a (negative) relationship between railroad accidents and observed Crusade activity, arising because accidents in the railroad lead to communication disruptions reducing information flows fueling protest diffusion. In the presence of mis-classification in our data set related to newspaper coverage of Crusade events, an alternative channel leading to the same negative relationship between railroad accidents and observed Crusade activity may arise, constituting a violation of the exclusion restriction: if a crowding-out effect is present such that newspaper reports of railroad accidents lead to less newspaper reporting of Crusade activity, the occurrence of railroad accidents will be negatively related to the recording of Crusade events in our data set. Furthermore, it will imply a positive ρ_{wr}.

Because the existence of such a channel requires a crowding-out mechanism to be in place, we begin this subsection implementing an indirect test of crowding out effects by looking at the relationship between train accident reporting in newspapers and differential reporting of a battery of different topics commonly covered by newspapers. Our presumption is that if railroad accidents induce crowding out effects, there is little reason to expect them to show up over some topics but

News Topics Crowding-out Effects from Railroad Accident News

Dependent Variable: Word Count of			
Politics	-0.036	Disasters	-0.014
	(0.034)		(0.019)
Economics	-0.051	Religion/Family	0.041
	(0.032)		(0.034)
Business	-0.047	Health/Education	0.010
	(0.032)		(0.049)
Sports	0.013	World	0.015
	(0.013)		(0.029)
Farming	0.053	Entertainment	-0.034
	(0.034)		(0.028)
Weather	0.050		
	(0.031)		

Table B.2: Railroad Accident News Crowding-Out Effects. The table presents estimates of equation (B.17) for each of a batter of alternative news topics. Each coefficient and associated standard error correspond to a different regression. The main explanatory variable in all cases is the count of railroad accident news reports. All models include newspaper and period fixed effects, the log word count of the newspaper, the page count of the newspaper, a quadratic polynomial in the total count of keyword matches across all topics, and a constant. A period is defined as a five-day interval, and the panel covers the period Jan. 1872 - Dec. 1874. The keywords included in each topic are described in Table D.1. All regressions include 30,141 observations and cover 282 newspapers.

not over others. We scraped the content of all articles in our data set of newspapers covering the period Jan. 1872 - Dec. 1874, based on a list of keywords which we use as signals of coverage on a host of different topics. Table D.1 describes the set of keywords we used for each topic. Based on this search, we created a count of mentions of each group of words in a given newspaper-period. In parallel, we scraped the same set of newspaper articles based on a list of keywords signaling the reporting of railroad-related accidents, and computed a count of the number of railroad accident mentions in a given newspaper-period.[2] For each topic of interest j we regress the count of keyword matches y^j_{nt} on the count of railroad accident news r_{nt}, newspaper and period fixed effects, and a vector of control regressors that includes a fourth-order polynomial on the total number of keyword matches across all topics ($Y_{nt} = \sum_j y^j_{nt}$), the log word count of the newspaper, the number of pages of the newspaper, and a constant:

$$y^j_{nt} = \alpha_n + \tilde{\alpha}_t + \beta r_{nt} + \gamma \text{Log word count}_{nt} + \delta \text{Page count}_{nt} + \mathcal{P}(Y_{nt}) + \epsilon_{nt} \tag{B.17}$$

We report the results of this exercise in Table B.2. Across all topics, we find no statistically

[2]A positive hit in our railroad accident search corresponds to the finding of an article where any word in the set {rail, train, passenger car, engine car, locomotive, railroad, railway, wagon} appears simultaneously with any word in the set {accident, break, broke, turned over, explosion, exploded, explode, derail, derailed, derailment, ran into, collision, collided, obstruct, cattle on track, misplaced switch, defective wheel, defective frog, land slide, falling break-beam, overloading car, burned, ran off track, wreck, washed out, sink, demolish}.

significant relationship between periods with higher than average railroad accident reports and differential reporting of any of the news topics we considered. The sample size in these exercises is large, so we do not believe this to be an under-powered test. Moreover, these results are robust to alternative classifications of keywords into topics, and to the omission of alternative sets of keywords from different topics. Overall, we find no evidence of crowding-out effects from railroad accident news reports. If anything, out of the 11 topics we considered, only the model for weather-related topics shows a marginally significant coefficient (at the 10 percent level). The effect is, however, of positive sign.

We undertake a second empirical exercise to assess whether we can detect any change in the underlying relationship between newspaper coverage of railroad accidents and the actual occurrence of railroad accidents during the Temperance Crusade. If newspaper reporting behavior about rail accidents is different during the months of the Crusade (for example through a form of crowding out), this could induce a correlation between the reporting of Crusade activity and railroad accidents. Thus, we aggregate counts y_{st} of newspaper mentions of rail accidents either at the state-month level or at the state-half month level (matching newspaper locations to their corresponding states), and counts x_{st} of rail accident occurrences from the *Railroad Gazettes* for the period Jan.-1872-June 1874. We estimate models of the form

$$y_{st} = \alpha_s + \tilde{\alpha}_t + \beta_t x_{st} + \gamma x_{st} + \epsilon_{st} \tag{B.18}$$

While γ measures the average (across states and time) rate at which news article mentions are generated per railroad accident happening, the β_t captures any period-specific difference in this rate. We can then compare the period-specific slope differences β_t during the Crusade (Dec. 1873 to July 1874) to the period prior to its onset (Jan. 1872 to Nov. 1973). Each railroad accident is reported an average of 1.5 times across all newspapers from the corresponding state during the month of the accident (0.86 times within the two-week windows). Figure B.2 plots the β_t coefficient estimates from (B.18) over time, aggregating the data either at the month level (left-hand side) or at the half-month level (right-hand side), showing that the average rate at which newspapers reported on railroad accidents was no different before the Crusade began (white) or during the Crusade period (pink), irrespective of the time period definition. Neither of the exercises we presented here suggest a correlation between mis-classification of Crusade events and railroad accidents operating through newspaper reporting.

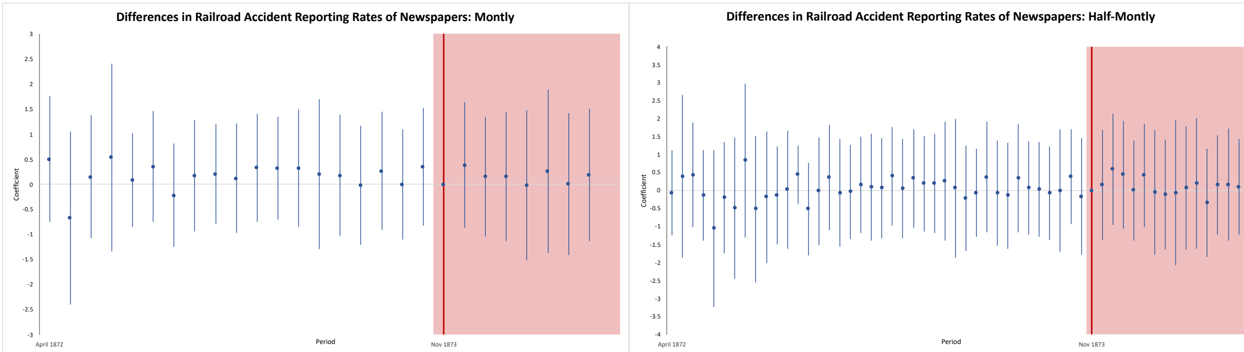


Figure B.2: Differences over Time in Newspaper Reporting of Railroad Accidents. β_t coefficients from equation (B.18). The figures plot the coefficients β_t from (B.18) (and associated 95 percent confidence intervals), measuring the difference over time (relative to November 1873) in the reporting rate of railroad accidents by newspapers of the state where the accident took place. The left-hand side figure reports the results from a regression where accidents and news reports are aggregated at the monthly level. The right-hand side figure reports the results from a regression where accidents and news reports are aggregated at the half-month level. The month-level regression is based on 840 state-month observations. The half-month regression is based on 1680 state-half month observations. The pink shade denotes the period of active Temperance Crusade activity.

B.3 Sensitivity Exercise: Lower Bounds on IV Estimates under alternative Measurement Error Scenarios

The evidence from subsection B.1 suggests mis-classification, if present, is likely to be small. Under plausible assumptions, it also suggests a 5% upper bound on the mis-classification rate. The evidence from subsection B.2 similarly suggests that any correlation between measurement errors and railroad accidents, if present, is likely to be small. Table B.3 reports lower bounds from equation (B.13) for our causal effects of interest. These are based on the IV coefficients from our benchmark specification (column 2 from Table 2), in a sensitivity analysis where we allow the key parameters governing the IV bias, α and ρ_{wr}, to take values in the sets $\{0.05, 0.1\}$ and $\{0.01, 0.05\}$. Notice that under this sensitivity analysis we are allowing for values larger than the upper bounds we estimated above.

The first panel of Table B.3 reports lower bounds for the effect of the first lag of rail-mediated signals. The first row considers an extremely large mis-classification rate of 10 percent –twice as large as the upper bound we estimated–. Compared to our point estimate of 0.037, we find lower bounds of 0.012 or 0.029 depending on whether ρ_{wr} is 0.05 or 0.01. When we instead consider a mis-classification rate of 5 percent –the upper bound we estimated for mis-classification–, we find lower bounds for the causal effect of 0.014 and 0.031 depending on whether ρ_{wr} is 0.05 or 0.01. Across all of these scenarios, the lower bounds we obtain are positive. The second panel then reports lower bounds for the effect of the first lag of telegraph-mediated effect signals, across the same range of scenarios. We find lower bounds between 0.13 and 0.16, all positive and close to out point estimate of 0.172. Finally, the third panel reports lower bounds for the effect of the

Sensitivity of IV Estimates to Mis-classification of Crusade Events

Extent of mis-classification	Implied Inflation Factor $\overline{\pi}$	Lower Bounds on Casual Effects						
			β_r^0		β_{γ}^0		β_d^2	
		ρ_{wr}	0.05	0.01	0.05	0.01	0.05	0.01
Extreme: 10% un-reported	1.11		0.012	0.029	0.134	0.151	-0.016	0.001
Upper bound: 5% un-reported	1.05		0.014	0.031	0.143	0.160	-0.015	0.002
Point estimates			0.037		0.172		0.006	

Table B.3: Sensitivity of IV Estimates to Mis-classification of Crusade Events. The table reports lower bounds for the benchmark social interaction effects using equation (B.13), for alternative values of the mis-classification rate (α) and the correlation between mis-classification and rail link variation (ρ_{wr}). The columns under panel β_r^0 report lower bounds for the first lag of rail-mediated signals. The columns under panel β_{γ}^0 report lower bounds for the first lag of telegraph-mediated signals. The columns under panel β_d^2 report lower bounds for the third lag of distance-mediated signals.

third lag of distance-mediated signals, across the same range of scenarios. In this case, the small magnitude of our point estimate (0.006) leads to negative lower bounds when ρ_{wr} is very large. For the scenarios where $\rho_{wr} = 0.01$, however, the lower bounds are positive even under the extreme case where $\alpha = 0.1$. All together, this table considers very conservative scenarios, strongly suggesting that the causal effects of rail and telegraph-mediated information flows are positive, and that bias caused by the two measurement-error channels considered in this appendix is small.

C Online Appendix: Aggregate Dynamics: Testing Models of Social Interactions

In this appendix, we 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 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 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 must be characterized by a *concave* adoption curve.[3] The top left panel in Figure C.1 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.[4] 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 C.1 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

[3] 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.

[4] 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 share of non-adopters adopting at a given instant has both an inertial component and a component proportional to the share who have already adopted.

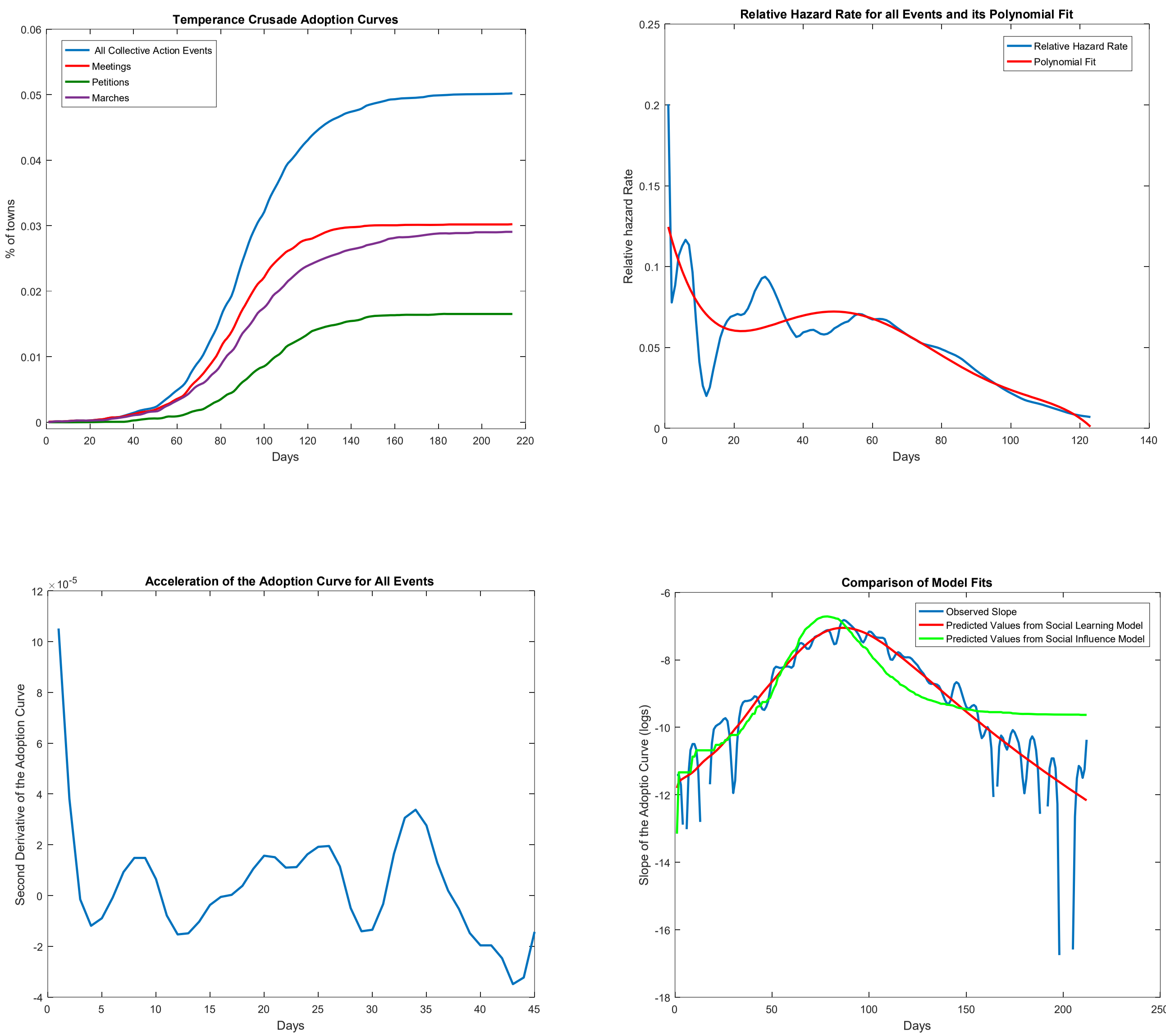


Figure C.1: Tests of Alternative Protest Diffusion Signatures, based on Young (2009). The figure in the top left presents the adoption curves $p(t)$ of the Crusade. The blue line includes all types of Temperance Crusade events –meetings, petitions, and marches–. The red line includes only meetings, the 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 C.1 (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 C.1 (in green) and on the coefficients in column (3) of Table C.1 (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 (2017) 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) > (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 C.1: 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 (2017), comparing the models from columns (2) and (3).

polynomial in time.[5] 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 C.1 depicts both the relative hazard rate (in blue), and the fitted values based on the estimates from the model in column (1) of Table C.1. 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.[6]

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

[5] 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.

[6] 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.

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 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.[7] In Figure 1 we already illustrated the slow and bumpy start of the Crusade. In the bottom left panel of Figure C.1 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 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 C.1. 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 (2017). 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.[8] The much better fit of the model in column (3) of the table can also be seen graphically. In the bottom

[7] 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)

[8] The Schennach and Wilhelm (2017) test requires providing a tuning parameter ε_n. We follow their advice and compute ε_n based on their suggested optimal choice.

right panel of Figure C.1 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 C.1. 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.[9]

[9]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.

D Online Appendix: Supplementary Data Description

D.1 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 historical newspapers from 1690 to present. The archive's interface allows a researcher to carry out keyword searches.

We searched for the following keywords (or combination of keywords, when a keyword is likely to generate numerous 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. The output of any keyword search is an image of a newspaper page containing one or more of these keywords. The text generated by processing this image can be downloaded. We downloaded any text that contained any of these keywords in its body. We also downloaded meta-information about the newspaper publishing the text - such as its name and location. This step resulted in several thousand articles which contained at least one of these keywords, some of which may be duplicates.

To reduce problems due to image-to-text processing, we implemented the following steps:

1. We removed punctuation and signs that were likely to be included in the output due to imperfect image processing, such as \ or |.

2. We searched for words which may have been unintentionally separated, creating two consecutive unintelligible words. For example, if word "development" was separated into two consecutive words like "deve" and "lopment," we tried to combine them 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 heuristic.

To reduce the number of false positive mentions of town names, we tried to identify the text of any article mentioning a Crusade event in a newspaper and then carried out the search of census town names only in this text. Unfortunately, in addition to the problems generated by imagine-to-text processing, working with historical newspapers is challenging because they rarely have indicators for where an article begins and where it ends. To cover the approximate body of an article, we used the following steps. If several consecutive pages turned up in the keyword search from the same date and newspaper, we assumed they were coming from the same article spread across multiple pages. We combined such texts on consecutive pages to form an article. If there are multiple hits from the same newspaper on the same day, but they on nonconsecutive pages, we assumed they belong to multiple articles. In this combined body, we then treated the locations of the very first and the very last keyword hits as indicators of where an article may lie. It is unlikely that these words coincide with the exact first and last words of any article, so we supplemented the text between these two keywords with an additional text of 100 words before and after, to increase the chance of covering the full article. We carried out a town name search in this large combined text by looking for a match to any of the nearly 15,000 towns in the U.S., as we detail next.

Searching for the Names of Towns

We searched for mentions of towns within the above-mentioned combined text that is likely to capture an article body. We matched the list of recovered town names from all articles to our list of Crusading towns.

This procedure can generate false positives for two reasons. First, some town names can correspond to words with alternative meanings. For example, “Union” is a town in NY, as well as an English word. To reduce such false positives, we checked if the word indicating the town name started with a capital letter. Second, there may be multiple towns in the US with the identical name in different counties or states. To deal with such cases, we checked if any state names are mentioned in the article text. If only one state was mentioned, we marked the town in this state as mentioned in our dataset. If there were multiple states with a possible match, we assigned 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 in the US was coded as unmentioned (0), town + state name mentioned or unique town (1), or as partial information or multiple town/town+state names mentioned as a possible match (a number between 0 and 1).

Because newspapers were smaller in page sizes and had fewer pages for the duration of interest, newspapers’ likelihood of publishing multiple articles on the same topic on any given day is small. However, our approach may still generate town names that are not relevant to the women’s protests. Unfortunately, after these steps, there were still some town names with false positives. To reduce the likelihood of misclassification, we checked if articles which cover stories about the Crusade mention the U.S. Census towns which were not mentioned by Blocker (1985). For the articles that did not contain a town reported by Blocker to have a Crusade event, but included a subset of the keywords listed related to the Crusade, we carried out a manual text search in the retained text. Four research assistants manually checked the text to ensure that the articles mentioned of the Crusade events and retained those relevant to the Crusade.

Using this output we 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.

D.2 Newspaper Search for Crowding Out of Topics

To identify the news articles on different topics, we first generated a list of topics which were commonly reported in the newspapers of the Crusade period: politics, economics, business, sports, farming, weather, disasters, religion and family, health and education issues, world news, and finally news on entertainment. We created an extensive list of keywords, provided in Table D.1, to detect the mention of each topic. We searched for these keywords in the "Chronicling of America" Newspaper database of the Library of Congress, for each day, going back to January 1st of 1872 to June 30th of 1874. We downloaded all articles related to the keywords, along with the number of times each keyword is mentioned, and retained the metadata about the newspaper - such as its name and location.

Keywords by Topic

Topic	Keywords
Politics	election, party, Republican, Democrat, general, captain, president, politics, legislative, council, elector, congress, mayor, municipality, governor
Economics	price/prices, economy, economics, industrial, industry, gold, silver, coal
Business	restaurant, store, shop, butcher, office, business, goods, sale, hotel, insurance, assurance, shoes, boots
Sports	sport/sports, sportive, player, cricket
Farm	livestock, crop/crops, wheat, corn, drought, farm/farming, animal/animals, veterinary, cattle, cow/cows, chicken/chickens, horse/horses, mill, furnace
Weather	storm, rain, rainy, thunder, wind, snow, snowy, freezing, heat, ice, icy
Disasters	earthquake, tornado, fire, burned down, collapse, mudslide, people, gathering, meeting
Religion/Family	church, reverend, synagogue, temple, pastor, bible, religion, father, children, child, baby, mother, family, marriage, wedding, divorce, parent
Health/Education	medical, doctor, disease, contagion, sickness, illness, health, nurse, school, university, college, teacher, student, educate, education, learning
World	England, France/French, Germany, Japan, Europe, Asia, Britain/British, war
Entertainment	theater, festival, tournament, game, saloon

Table D.1: List of Keywords Used to Construct News Topics for the Crowding Out Exercise. The table lists the keywords used to classify newspaper articles into topics, for the railroad accident crowding out exercise reported in table Table B.2.

D.3 Search of the Railroad Accident Coverage in Newspapers

We search for railroad accident mentions in the "Chronicling of America" newspaper database of the Library of Congress, for each day, between the dates of Jan 1st, 1872 to June 30th of 1874.

We searched for two groups of keywords – first group indicating an event around railroads and the second indicating forms of accidents – and retained the article if a newspaper article contains at least one keyword from each group:

Group 1: "rail", "train", "passenger car", "engine car", "locomotive", "railroad", "railway", "wagon".

Group 2: "accident", "break", "broke", "turned over", "explosion", "explode", "derail", "derailed", "derailment", "ran into", "collision", "collided", " cattle on track", "misplaced switch", "defective", "wheel", "defective frog", "land slide", "falling brake-beam", "overloading car ", "burned", "ran off track", "wreck", "washed out", "sink", "demolish".

We downloaded all of the articles that were identified to mention a railroad accident, along with a dummy indicating which keywords are mentioned, and retained the metadata about the newspapers, including their name and locations.

D.4 Search of the Railroad Accident Records

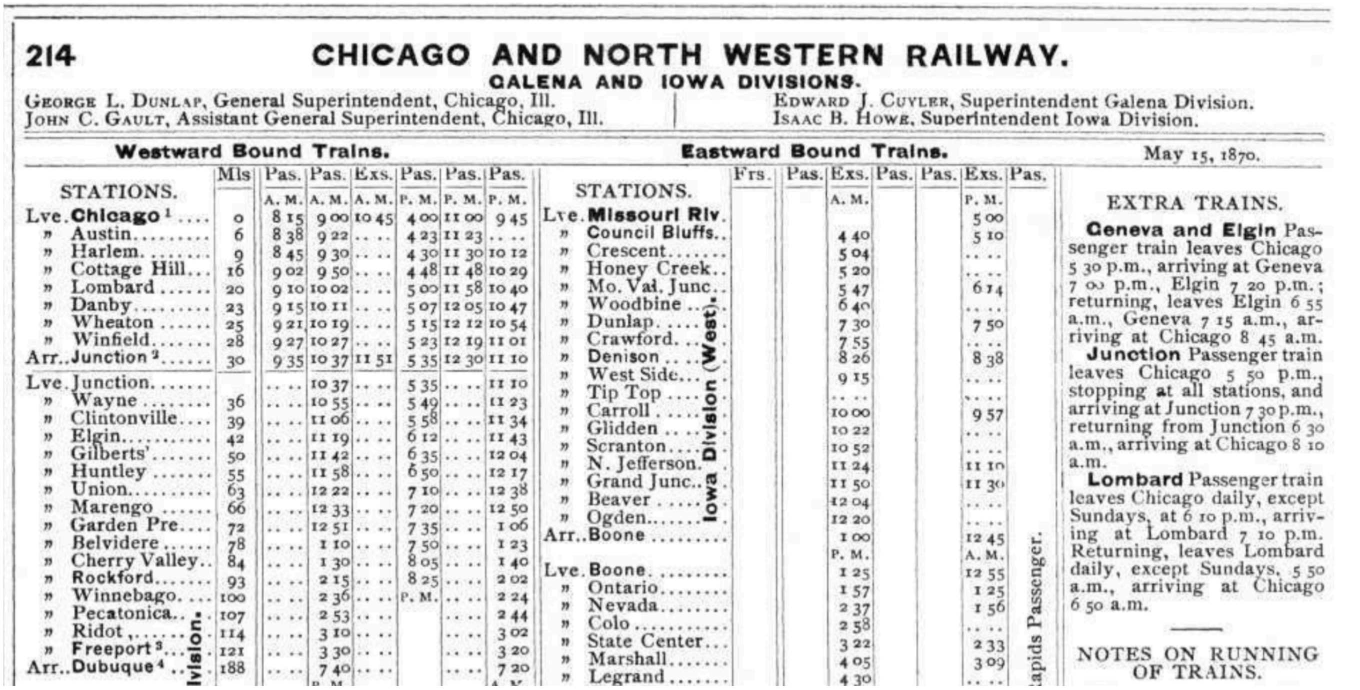
We search for railroad accidents reported in the monthly Railroad Gazette between Jan 1st, 1872 and June 30th of 1874. Railroad Gazette volumes were accessible via Hathitrust.org and Googlebooks.com. In the monthly gazette volumes, there is a reporting of accidents that took place within the month. We recorded all accidents mentioned in the gazettes, along with date, location, and railroad company information if they were reported. This search resulted in 2,186 accident data points.

There were no accidents reported for two months in the gazettes: January and April 1872. To address the possibility that this was a reporting oversight, for these months, we carried out a search in newspaper archives. For 12 accidents, gazettes only provided the year of the accident, but no other date. We removed these accidents from our analyses.

For other accidents reported in the gazette volumes, the location information was missing, or it was provided, but it did not correspond to any of the town names from the 1870 Census, or it corresponded to multiple town names. To identify the location of each accident using the 1870 railroad map, we use the following procedure on the Railroad Gazette accident data and Jeremy Atack's railroad archive[10]:

1. Search the town and state name in Atack's 1870 railroad map.

[10] We use Atack's 1870 ArcGIS shape files from the Vanderbilt University which cover all rail lines in the continental U.S. as of 1870. The collection can be accessed at https://my.vanderbilt.edu/jeremyatack/data-downloads/



214 **CHICAGO AND NORTH WESTERN RAILWAY.**

GALENA AND IOWA DIVISIONS.

George L. Dunlap, General Superintendent, Chicago, Ill.
John C. Gault, Assistant General Superintendent, Chicago, Ill.

Edward J. Cuyler, Superintendent Galena Division.
Isaac B. Howe, Superintendent Iowa Division.

Westward Bound Trains.

STATIONS.	Mls	Pas. A. M.	Pas. A. M.	Exs. A. M.	Pas. P. M.	Pas. P. M.	Pas. P. M.
Lve. **Chicago** [1]	0	8 15	9 00	10 45	4 00	11 00	9 45
» Austin	6	8 38	9 22	....	4 23	11 23	....
» Harlem	9	8 45	9 30	....	4 30	11 30	10 12
» Cottage Hill	16	9 02	9 50	....	4 48	11 48	10 29
» Lombard	20	9 10	10 02	....	5 00	11 58	10 40
» Danby	23	9 15	10 11	....	5 07	12 05	10 47
» Wheaton	25	9 21	10 19	....	5 15	12 12	10 54
» Winfield	28	9 27	10 27	....	5 23	12 19	11 01
Arr. **Junction** [2]	30	9 35	10 37	11 51	5 35	12 30	11 10
Lve. Junction		....	10 37	....	5 35	....	11 10
» Wayne	36	....	10 55	....	5 49	....	11 23
» Clintonville	39	....	11 06	....	5 58	....	11 34
» Elgin	42	....	11 19	....	6 12	....	11 43
» Gilberts'	50	....	11 42	....	6 35	....	12 04
» Huntley	55	....	11 58	....	6 50	....	12 17
» Union	63	....	12 22	....	7 10	....	12 38
» Marengo	66	....	12 33	....	7 20	....	12 50
» Garden Pre	72	....	12 51	....	7 35	....	1 06
» Belvidere	78	....	1 10	....	7 50	....	1 23
» Cherry Valley	84	....	1 30	....	8 05	....	1 40
» **Rockford**	93	....	2 15	....	8 25	....	2 02
» Winnebago	100	....	2 36	....	P. M.	....	2 24
» Pecatonica	107	....	2 53	....		....	2 44
» Ridot	114	....	3 10	....		....	3 02
» **Freeport** [3]	121	....	3 30	....		....	3 20
Arr. **Dubuque** [4]	188	....	7 40	....		....	7 20

ivision.

Eastward Bound Trains.

STATIONS.	Frs.	Pas.	Exs. A. M.	Pas.	Pas.	Exs. P. M.	Pas.
Lve. **Missouri Riv.**						5 00	
» **Council Bluffs**			4 40			5 10	
» Crescent			5 04			....	
» Honey Creek			5 20			....	
» Mo. Val. Junc.			5 47			6 14	
» Woodbine			6 40			....	
» Dunlap			7 30			7 50	
» Crawford			7 55			....	
» Denison			8 26			8 38	
» West Side			9 15			....	
» Tip Top			....			....	
» Carroll			10 00			9 57	
» Glidden			10 22			....	
» Scranton			10 52			....	
» N. Jefferson			11 24			11 10	
» Grand Junc.			11 50			11 30	
» Beaver			12 04			....	
» Ogden			12 20			....	
Arr. Boone			1 00			12 45	
			P. M.			A. M.	
Lve. Boone			1 25			12 55	
» Ontario			1 57			1 25	
» Nevada			2 37			1 56	
» Colo			2 58			....	
» State Center			3 22			2 33	
» Marshall			4 05			3 09	
» Legrand			4 30			....	

Iowa Division (West).

apids Passenger.

May 15, 1870.

EXTRA TRAINS.

Geneva and Elgin Passenger train leaves Chicago 5 30 p.m., arriving at Geneva 7 00 p.m., Elgin 7 20 p.m.; returning, leaves Elgin 6 55 a.m., Geneva 7 15 a.m., arriving at Chicago 8 45 a.m.

Junction Passenger train leaves Chicago 5 50 p.m., stopping at all stations, and arriving at Junction 7 30 p.m., returning from Junction 6 30 a.m., arriving at Chicago 8 10 a.m.

Lombard Passenger train leaves Chicago daily, except Sundays, at 6 10 p.m., arriving at Lombard 7 10 p.m. Returning, leaves Lombard daily, except Sundays, 5 50 a.m., arriving at Chicago 6 50 a.m.

NOTES ON RUNNING OF TRAINS.

Figure D.1: Travelers' Guide Example

- If there is a unique town identified in a state, we search for the name of the railroad company in the Travelers' Official Railway Guide for the United States and Canada from 1870 (Vernon, 1870). A page from the guide is provided as an example in Figure D.1. We find the starting station and the end station of the railroad path on which the town is located to confirm the town location and state.

  For example, if a Railroad Gazette record states that "An accident took place on the Chicago & Northwestern Railway in Chicago, IL on 1873-01-24", we search in the Travelers' Guide information about the railroad company "Chicago & Northwestern Railway". Figure D.1 shows detailed information about "Chicago & Northwestern Railway." If we find Chicago as a station listed on this railroad's path, we confirm "Chicago" "IL" as the town and state where the accident took place.

- If the town location and railroad company do not match, we undertake a second search in historical newspaper archives, detailed in Step 4.

2. If the accident town is not in Atack's 1870 railroad map, but there is a station with the identical name in the Travelers' Guide, then we try to identify the railroad path that this town is on. To do this, using the Travelers' Guide, we first locate the nearest towns before and

after this town, go back to the railroad map, and check if there are other towns between these towns.

- If there is one town between the nearest towns, we set accident location to this town.
- If there are multiple towns between the nearest towns, we choose the town with the largest population among them and set it as the accident location.
- If there are no towns, we find the town that is the closest in distance to this segment of railroad path and mark it as the town of the accident location.

3. If an accident town is found in neither the 1870 map nor the Travelers' Guide:

   - if accident location refers to a river, lake, mountain, mill, valley or another landform, then we search in the railroad map the location of the mentioned landform and search nearby towns. If there is a unique town near the landform (e.g., a town nearby a mountain), we set this town as the accident town. If there are multiple towns nearby the landform (e.g., a river may go through multiple towns), we set the nearby town with the highest population as the accident location.
   - If the accident location refers to a village, city, neighborhood, area rather than a town, we find the nearest town in our database to this location and set it to the accident location.

4. Additional Search on Newspaper Archives:

   Even after Steps 1-3, the town and state of some accidents could not be identified. To find the location of these remaining accidents, we searched the Newspaper Archive database accessible at https://access.newspaperarchive.com/.

   Specifically, if the accident location does not match any of the towns through which the railroad company of the accident is going through; or if the date and the railroad company of the accident are reported in the Railroad Gazette without an accident location:

   - We search for mentions of railroad accidents using the phrases "railroad accident," "railway accident," and "train accident" in the Newspaper Archive. We check if there is a mention of an accident matching the name of the accident railroad company or the state where the accident took place published within three days of the accident date reported in the Gazette. If we find a newspaper article matching the accident description, we record the mentioned town as the town of accident. Otherwise, we record "NOT FOUND."

Note that no accidents were reported in the Railroad Gazette for January and April 1872. Since this may be a reporting error, we searched for railroad accident records for these two months in Newspaper Archive and recorded all accidents found as well.

Illustration of a Telegraph Map from WesternUnion (1874)

Figure D.2 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.

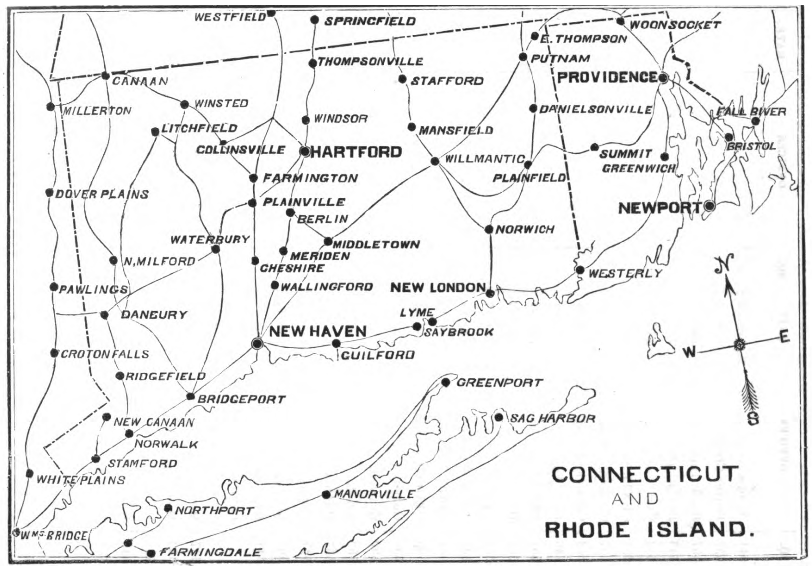


Figure D.2: The Telegraph Network in Connecticut and Rhode Island, 1874. The figure reproduces the Western Union telegraph lines map in WesternUnion (1874) for the states of Connecticut and Rhode Island.