American Economic Review: Papers & Proceedings 2016, 106(5): 61–67
http://dx.doi.org/10.1257/aer.p20161071

State Capacity and American Technology: Evidence from the Nineteenth Century†

By Daron Acemoglu, Jacob Moscona, and James A. Robinson*

In *The Rise and Fall of American Growth*, Robert J. Gordon provides a magisterial overview of the economic and human changes that have reshaped US society since the middle of the nineteenth century. In addition to the rich, lively details about the many technologies that have made us so much more prosperous, comfortable, and healthy than our ancestors, Gordon also provides a conceptual framework for thinking about American technology. In this framework, technology progresses at a normal pace during regular times, but then accelerates during distinctive episodes, such as in the aftermath of the second industrial revolution between 1920 and 1970, because of the availability of major "Great Inventions" (Gordon 2016, p. 2), which created opportunities in a wide array of industries and applications. There is no puzzle, according to this framework, as to why productivity growth slowed down after the 1970s: we were simply in a lucky period of accelerated technological opportunities until the 1970s, which then made way for the regular workings of our economy.

In this richly woven narrative, the endogeneity of these technological changes received relatively little attention. Economists have come to think of both the pace and the direction of technological change as endogenous for two related reasons (see Acemoglu 2009, for an overview): (i) within a given institutional environment, technology advances by innovation decisions, affected by such things as taxes, R&D subsidies and tax credits, and the supplies of other factors of production; and (ii) the institutional environment has a key impact on technological progress through the security of property rights, patent laws, educational system, and general competition policy. These factors receive minimal attention in this voluminous tome. Gordon recognizes the importance of certain policies, for instance in his discussion of the New Deal and World War II. He also notes

> *Even in the laissez-faire environment of the nineteenth century, the government intervened in the development of the economy in numerous ways. Government actions included a wide range of legislation, the granting of land to railroads and homesteading settlers, food and drug regulations, the establishment of land grant universities and agricultural research stations, the patent system, deposit insurance, Social Security, and unemployment compensation* (Gordon 2016, p. 289),

and

> *Perhaps the most important government activity to stimulate growth was the patent office and the process of patent approval* (Gordon 2016, p. 312).

But Gordon does not link these important institutional underpinnings to the pace and nature of American innovation or the existence of potential "Great Inventions." Perhaps what enabled the United States to become more innovative than all other countries in the nineteenth century and then come to dominate technology, paving the way to the second industrial revolution, were its policies and the institutional structure lying behind these. If so, it is plausible that the potential for growth of the American economy in the next several decades will also depend not just on exogenous technological constraints,

* Moscona: Harvard University, 468 Cabot Mail Center, Cambridge, MA 02138 (e-mail: jmosconaskolnik@college.harvard.edu); Robinson: Harris School of Public Policy, University of Chicago, 1155 East Sixtieth Street, Chicago, IL 60637 (e-mail: jamesrobinson@uchicago.edu); Acemoglu: Department of Economics, Massachusetts Institute of Technology, E52-380, 50 Memorial Drive, Cambridge, MA 02142 (e-mail: daron@mit.edu).

† Go to http://dx.doi.org/10.1257/aer.p20161071 to visit the article page for additional materials and author disclosure statement(s).

but on these institutions. We do not get answers to these critical questions from Gordon.

Scholars of American innovation, most notably Sokoloff (1988) and Khan (2005), emphasize the critical role of these institutional factors. In this short paper, we go one step further and provide evidence that one particular set of institutions, the presence of the state's infrastructural capacity, played a significant role in nineteenth-century innovations. Though the American state is commonly regarded as being weak in the nineteenth century, a recent historical literature has fundamentally challenged this view (Novak 2008; King and Lieberman 2009). Indeed, the nineteenth-century US state quickly started forming a huge web connecting the country. At the center of this web was the post office, created by the Post Office Act of 1792, which soon became the single most important government employer in the first half of the nineteenth century. In 1816, for example, 69 percent of the federal civilian workforce were postmasters, and by 1841 this number was 79 percent (John 1995). As John (1995, p. 4) puts it, in this period "for the vast majority of Americans the postal system *was* the central government" (italics in original).

The *New York Times* in 1852 described it as the "mighty arm of civil government" (John 1995, p. 10). Its pervasiveness was apparent to de Tocqueville who, during his famous travels in 1831, noted: "There is an astonishing circulation of letters and newspapers among these savage woods … I do not think that in the most enlightened districts of France there is an intellectual movement either so rapid or on such a scale as in this wilderness" (de Tocqueville 1969, p. 283). He also noted how it provides a "great link between minds" and "penetrates" into the "heart of the wilderness" (1969, pp. 384–85, 11).

The presence of the post office was significant for innovation for at least three reasons. First, by facilitating flows of information and knowledge, it helped ideas to spread and facilitated the creation of new ones. Second, for the more prosaic reason that it made patenting and securing intellectual property rights much easier:

> *…rural inventors in the United States could apply for patents without serious obstacles, because applications could be submitted by mail free of postage. The United States Patent and Trademark Office also maintained repositories throughout the country, where inventors could forward their patent models at the expense of the post office. As such, it is not surprising that much of the initial surge in patenting during early American industrialization occurred in rural areas.* Khan (2005, p. 59).

Finally, and no less importantly, the presence of a post office is indicative of a much broader state presence and functionality, for example, via legal services and regulation, access to land, and security of other forms of property rights, which are prerequisites for most innovative activity. Also important was the fact that, already by the 1830s, the post office was a modern bureaucratized institution.

Building on this third observation and in the spirit of the empirical approach in Acemoglu, García-Jimeno, and Robinson (2015), we use the number of post offices in a county as a proxy for the general presence and "infrastructural power" of the state. We argue it was this, as well as the availability of postal services, that made innovation and patenting feasible and desirable. Our main dependent variable is the number of patents granted to residents of a county in a particular period.

In the remainder of this short paper, we first present our post office and patent data, then describe our empirical strategy, and finally present original results for a panel of US counties between 1804 and 1899. Our results show a strong association between the presence and number of post offices in a county and patenting activity, and it appears that it is the opening of postal offices that leads to surges in patenting activity, not the other way around, indicating that it is not a simple reverse causality story underpinning this association. Though our results do not unambiguously establish causality, the interpretation that this statistical relationship mostly reflects the impact of the post offices and the state institutions associated with them on innovative activity receives support from the evidence presented in John (1995) that the expansion of the postal network reflected a range of idiosyncratic factors and motivations.[1]

[1] For example, John (1995, pp. 44–45) notes that pressure from certain segments of society for the state's services

I. Data

Data on the number of post offices in each county were compiled from lists of United States post offices that were assembled for several years during the nineteenth century by the United States Post Office Department (predecessor to the United States Postal Service), headed by the United States postmaster general. Along with the post office name and title of the office's postmaster (one postmaster is assigned to each post office), the lists present the state and county in which the post office was located. The number of post offices is therefore equivalent to the number of postal service employees employed by the federal government. To establish a post office, the federal government appoints a single postmaster to the office. Additional office employees are employed by and under the jurisdiction of each office's postmaster rather than the postmaster general or the federal government. Indeed, in the *Register of Officers and Agents, Civil, Military, and Naval, in the Service of the United States*, a publication established by the US Congress in 1816 to list all US federal employees, the only post office employees listed are the postmasters themselves.

We obtained the number of post offices in each US county for the years 1804, 1811, 1819, 1830, 1837, 1846, 1850, 1855, 1867, 1870, 1879, 1891, and 1899. For the years before 1879, we used United States Post Office Department publications titled *List of the Post Offices in the United States* (in some years, the publication was referred to as *Table of Post Offices in the United States*). In 1874, the federal government began publishing post office information more systematically in a publication titled *The United States Postal Guide*, which is digitized only for some years. This publication is our source for the years 1879, 1891, and 1899.

Patent data for 1836–1900 are from a dataset compiled and shared by Tom Nicholas. The dataset contains patent number, granting year, and county of residence of the first inventor for all patents granted by the United States Patent and Trademark Office between 1836–1900. Inventor location was identified using optimal character recognition techniques (Akcigit, Kerr, and Nicholas 2013). We excluded patents credited to inventors residing outside the United States.

II. Empirical Evidence

Our basic empirical specification is a panel data model with county fixed effects and time effects. We look at two separate samples: a balanced sample of 935 counties over nine time periods (1837, 1846, 1850, 1855, 1867, 1870, 1879, 1891, and 1899) and an unbalanced sample consisting of a maximum of 2,644 counties. We also use information on lags of the post office variable from the years 1804, 1811, 1819, and 1830 so as to maximize the span of our sample. Because there are many county-year pairs with zero patents (4,863 in the balanced panel and 10,188 in the unbalanced panel) and also many counties have no post offices, we start with linear models of the form

$$(1)\quad \log(1 + patents_{ct}) = \sum_{j=1}^{J} \beta_j \log(1 + post\ offices_{ct-j}) + X_{ct}\gamma + \delta_t + \eta_c + \epsilon_{ct}.$$

Here, $patent_{ct}$ is the number of patents issued to residents of county c in time t, $post\ offices_{ct}$ is the number of post offices in county c in time t, X_{ct} is a covariate vector which always includes the log of county population and will later include time interactions with baseline county characteristics. In addition, δ_t denotes a full set of time effects, η_c denotes a full set of county fixed effects, and ϵ_{ct} is an error term. The coefficients of interest are the β_js which measure the potential effect of lagged post offices on patents. Our focus on lagged post offices is motivated by the fact that new post offices cannot plausibly be expected to have an impact on patenting right away. Since patents are a count

"guaranteed that the postal network would expand rapidly into the transappalachian West well in advance of commercial demand."

Similarly Cushing (1893, p. 286) reports "The establishment of post offices in Oklahoma and in other regions recently opened has often been in advance of the actual settlement. Before the Oklahoma counties were named they were called by the Department, A, B, C, D, E, etc. Postmasters were appointed upon recommendations of the delegate from Oklahoma and of Senators Plumb, Paddock, and Manderson. The theory of the Department is that the establishment of an office in a new locality is often the means of educating the people who become its patrons."

Table 1—Baseline Results

	Outcome variable is log(1 + patents)		Outcome variable is patent count			
	OLS (1)	OLS (2)	Poisson (3)	Poisson (4)	Neg. bin. (5)	Neg. bin. (6)
Panel A. Balanced panel of counties existing in 1830						
log (1 + post offices) first lag	−0.004	−0.01	0.055**	0.055**	0.084***	0.075***
	(0.017)	(0.016)	(0.027)	(0.027)	(0.017)	(0.016)
log (1 + post offices) second lag		0.032*		0.043		0.055**
		(0.016)		(0.036)		(0.024)
log (1 + post offices) third lag		0.080***		0.046**		0.048**
		(0.016)		(0.023)		(0.022)
log (1 + post offices) fourth lag		0.126***		0.012		0.034*
		(0.016)		(0.016)		(0.019)
R^2	0.390	0.407				
Observations	8,415	8,415	7,542	7,542	7,542	7,542
Number of counties	935	935	838	838	838	838
Panel B. Unbalanced panel of all counties, 1830–1890						
log (1 + post offices) first lag	−0.001	−0.001	0.074**	0.071**	0.119***	0.103***
	(0.01)	(0.009)	(0.032)	(0.03)	(0.014)	(0.014)
log (1 + post offices) second lag		0.021***		0.067*		0.067***
		(0.008)		(0.035)		(0.017)
log (1 + post offices) third lag		0.061***		0.074***		0.053***
		(0.009)		(0.028)		(0.015)
log (1 + post offices) fourth lag		0.100***		0.016		0.007
		(0.01)		(0.014)		(0.013)
R^2	0.381	0.396				
Observations	16,737	16,737	14,597	14,597	14,597	14,597
Number of counties	2,644	2,644	2,197	2,197	2,197	2,197

Notes: Panel A is a balanced panel of counties established by 1830 and panel B is an unbalanced panel of all counties, 1830–1890, in which counties are added as they are established. Columns 1 and 2 present results from OLS models where log (1 + patent count) is the outcome variable. Columns 3 and 4 present results from Poisson models and columns 5 and 6 present results from negative binomial models. The left-hand side in columns 3–6 is patent count. County and year fixed effects as well as the log of county population are included in all regressions. Standard errors, robust and clustered by county in columns 1–4, are in parentheses.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

variable, we also estimate Poisson and negative binomial models. Because these nonlinear models cannot accommodate the rich set of time interactions we later include to control for other county-level trends, we opt for the OLS models as our baseline. Nevertheless, results from these models show that the exact specification we use does not have a disproportionate impact on the general patterns we document.

Table 1 presents our first set of results. Panel A is for the balanced sample, and panel B for the unbalanced sample. The first two columns are for the OLS estimates of (1) including only log population of the county in that year as a covariate. In column 1, we just have the one period lagged post office variable on the right-hand side. This variable is not significant. This picture changes dramatically when we add four lags of post office in column 2.[2] Now the second lag is positive and

[2] Because the gap between different years in our panel is not even, we also estimated these models normalizing the effects of different lags by the inverse of the distance with the next date. These modified models are not reported

significant at 10 percent. The third and the fourth lags are larger and more precisely estimated, and in consequence, statistically significant at less than 1 percent. This pattern is plausible and suggests that it is mostly the presence of post offices in the previous two decades or so that is most strongly associated with increases in patenting activity. The relationship in column 2 is not only statistically powerful, but economically sizable as well: in the long run, once the full effects of lagged post offices are realized, the opening of a post office in a county that did not previously have a post office or previous patents increases the number of patents, on average, by 0.18.[3] The patterns and quantitative magnitudes are similar in panel B with the unbalanced sample. Columns 3–6 show broadly similar patterns when we estimate the Poisson and negative binomial models, though with some differences. Now the first lag of the post office variable entered by itself is significant, and when all four lags of the post office variable are included together, it is the first and the third lags that are robustly significant. The quantitative magnitudes are a little smaller with the Poisson and slightly larger with the negative binomial model. These differences notwithstanding, the conclusion that there is a strong association between the presence of post offices and subsequent patenting activity in a county remains both with the balanced sample and the unbalanced sample.[4]

An obvious concern is that the results presented in Table 1 reflect not the impact of post offices on patenting, but reverse causality or omitted time-varying factors, such as the expansion of economic activity, which might have simultaneously impacted the location of post offices and patenting. Table 2 investigates these issues, focusing on OLS models. Column 1 replicates column 2 from Table 1 for reference, just for the 930 counties for which we have data for the covariates we will use in columns 2–4. In column 2, we add a whole range of baseline country characteristics, each interacted with a full set of time dummies. These characteristics are: fraction slave population in 1860, fraction of the adult population that are literate in 1850, and the values of farmland and manufacturing output relative to population in 1850 (with both of these variables parameterized in the form of $\log(1 + x)$ since there are some zeros). Counties that differ in terms of these characteristics could have varying potential for industrialization and innovation, and these time interactions flexibly control for any differential trends in patenting related to these differences. Despite the addition of 36 controls, the pattern in column 2 is quite similar to that in column 1, with the only difference being that the second lag is also insignificant now, but the third and fourth flags continue to be strongly significant. The quantitative magnitude is somewhat smaller. The opening of a post office in a county without patents or a post office now increases the number of patents, on average, by 0.05 patents. Column 3 turns to the question of reverse causality. We include the current value and lead of the post office variable in the specification of column 2. Both of these variables are quantitatively very small (for example, less than one-ninth the size of the coefficient of the fourth lags), negative and statistically far from significance. This result suggests that it is unlikely that patenting activity leads to the opening of post offices. In column 4 we add county-level linear trends (thus 930 additional

to conserve space. They yield broadly similar results, and in fact in this case, the equivalent of the specification in column 1 also shows a significant effect comparable in magnitude to the Poisson and negative binomial models.

[3] The opening of a post office in a county with no previous post offices increases the right-hand side variable, $\log(1 + \textit{post offices})$, from $\log 1 = 0$ to $\log 2 = 0.69$. The long-run effect is given by summing the coefficients of the four lags, which gives approximately 0.24, thus increasing the left-hand side variable, $\log(1 + \textit{patent})$, by $0.24 \times 0.69 \simeq 0.166$. Then, the quantitative effect of interest, when evaluated starting from zero patents, is an increase in the number of patents given by $e^{0.166} - 1 \simeq 0.18$. If we were to instead evaluate both the increase in the post office and the impact on patents starting from their mean sample values (respectively, 19.86 and 6.29), then the implied increase of the right-hand side would be approximately 0.047, that is, from $\log(1 + 19.86) \simeq 3.038$ to $\log(1 + 19.86 + 1) \simeq 3.085$. This would raise the left-hand side by $0.24 \times 0.047 \simeq 0.011$, that is, from $\log(1 + 6.292) \simeq 1.987$ to 1.998. This corresponds to an increase of $e^{1.998} - 7.292 \simeq 0.08$, which is about half the effect evaluated at zero post offices and patents.

[4] We also estimated these models using citations of the patents as weights. However, since the citations are available only after 1947, most patents have zero citations, and perhaps because of this reason, the coefficient estimates are less stable, even though the overall pattern of results is quite similar.

Table 2—Robustness and Timing of Effects

Outcome variable: log(1 + patents)	Balanced panel of counties established by 1830				Unbalanced panel of all counties, 1830–1890			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log (1 + post offices) first lead			−0.007				−0.013	
			(0.015)				(0.012)	
log (1 + post offices)			−0.003				−0.004	
			(0.013)				(0.011)	
log (1 + post offices) first lag	−0.028*	−0.01	−0.019	0.021	−0.024**	−0.01	−0.015	0.026**
	(0.015)	(0.014)	(0.016)	(0.013)	(0.012)	(0.011)	(0.012)	(0.01)
log (1 + post offices) second lag	0.013	−0.008	−0.005	−0.006	0.007	−0.015*	−0.011	−0.002
	(0.016)	(0.013)	(0.013)	(0.013)	(0.011)	(0.009)	(0.009)	(0.01)
log (1 + post offices) third lag	0.069***	0.029**	0.022*	0.019	0.058***	0.0235***	0.020**	0.024**
	(0.016)	(0.012)	(0.012)	(0.013)	(0.011)	(0.009)	(0.009)	(0.01)
log (1 + post offices) fourth lag	0.121***	0.057***	0.061***	0.027**	0.116***	0.0611***	0.061***	0.040***
	(0.016)	(0.013)	(0.013)	(0.013)	(0.012)	(0.01)	(0.011)	(0.01)
p-value for slave population share interactions		[0.000]	[0.000]	[0.000]		[0.000]	[0.000]	[0.000]
p-value for literacy interactions		[0.000]	[0.000]	[0.000]		[0.000]	[0.000]	[0.000]
p-value for farm value interactions		[0.000]	[0.000]	[0.000]		[0.000]	[0.000]	[0.000]
p-value for manufactures value interactions		[0.000]	[0.000]	[0.000]		[0.000]	[0.000]	[0.000]
p-value for linear county trends				[0.000]				[0.000]
Observations	8,370	8,370	7,440	8,370	12,842	12,842	11,316	12,842
R^2	0.416	0.583	0.595	0.724	0.559	0.559	0.570	0.713
Number of counties	930	930	930	930	1,526	1,526	1,526	1,526

Notes: The outcome variable is log(1 + patent count). The log of county population and county and year fixed effects are included in all specifications. Columns 2–4 and 5–8 include the characteristics listed at the bottom of the table (slave population share in 1860, adult literacy share in 1850, log farmland value in 1850, and log manufacturing value in 1850) interacted with a full set of time dummies. Columns 4 and 8 include linear county trends. The *p*-values of the corresponding *F*-tests for the controls and county trends are reported in brackets. Robust standard errors clustered by county are in parentheses.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

controls) as a check against differential county-level trends accounting for our results. The results are broadly similar even though now only the fourth lag is statistically significant—while the third lag has a very similar magnitude to before, it is less precise and thus not significant at 5 percent. Finally, columns 5–8 repeat the specifications for the unbalanced sample, with very similar results.

Overall, the results in Tables 1 and 2 show a fairly robust correlation between the presence of the state in the recent past as proxied by the number of post offices in the previous several decades and patenting activity in a county. Moreover, the evidence indicates that it is post offices predicting future patenting, not the other way around. While we cannot claim to have unambiguously estimated the causal effect of post offices—or the presence of the state—on innovative activity, this evidence nonetheless points in the direction of an important role of the infrastructural capacity of the US state in undergirding and supporting innovation and technological change in this critical period of US economic growth.

REFERENCES

Acemoglu, Daron. 2009. *Introduction to Modern Economic Growth.* Princeton, NJ: Princeton University Press.

Acemoglu, Daron, Camilo García-Jimeno, and James A. Robinson. 2015. "State Capacity and Economic Development: A Network Approach." *American Economic Review* 105 (8): 2364–2409.

Akcigit, Ufuk, William R. Kerr, and Tom Nicholas. 2013. "The Mechanics of Endogenous Innovation and Growth: Evidence from Historical U.S. Patents." Unpublished.

Cushing, Marshall H. 1893. *The Story of Our Post Office*. Boston: A.M. Thayer & Co.

Department of Commerce and Labor, Bureau of the Census. 2014. "Register of Officers and Agents, Civil, Military, and Naval, in the Service of the United States." Provo, UT: Ancestry.com Operations, Inc.

Gordon, Robert J. 2016. *The Rise and Fall of American Growth*. Princeton, NJ: Princeton University Press.

John, Richard R. 1995. *Spreading the News: The American Postal System from Franklin to Morse*. Cambridge, MA: Harvard University Press.

Khan, B. Zorina. 2005. *The Democratization of Invention: Patents and Copyrights in American Economic Development, 1790–1920*. New York: Cambridge University Press.

King, Desmond, and Robert C. Lieberman. 2009. "Ironies of State Building: A Comparative Perspective on the American State." *World Politics* 61 (3): 547–88.

Novak, William J. 2008. "The Myth of the 'Weak' American State." *American Historical Review* 113 (3): 752–72.

Sokoloff, Kenneth L. 1988. "Inventive Activity in Early Industrial America: Evidence From Patent Records, 1790–1846." *Journal of Economic History* 48 (4): 813–50.

de Tocqueville, Alexis. (1835) 1969. *Democracy in America*. Edited by J. P. Mayer. Garden City, NY: Doubleday & Co.

US Post Office Department. 1804, 1808, 1828, 1855, 1870. List of Post Offices in the United States. Washington, DC.

US Post Office Department. 1811, 1819, 1822, 1830, 1837, 1850, 1867. Table of Post Offices in the United States. Washington, DC.

US Post Office Department. 1879, 1891, 1899. *United States Official Postal Guide*. Boston: Houghton, Osgood and Co.