Business Dynamism and Innovation: Evidence from Novel 19th-Century Microdata*

Antonin Bergeaud[†]
HEC Paris & CEPR

Ruveyda Nur Gozen[‡]
Cardiff Business School & CEP LSE

Suleyman Faruk Gozen[§]
University of Bristol

Preliminary Draft - October 2025

Abstract

We construct a novel establishment-level dataset combining the nineteenth-century U.S. Census of Manufactures (1850–1880), patent records from CUSP, and population census microdata to trace entrepreneurs, their firms, and inventive activity. This is the first paper to provide systematic micro evidence on the returns to innovation in early Second Industrial Revolution. We examine both the intensive and extensive margin of the return on patenting. We find that patenting establishments were substantially larger by employing more labor, paying higher wages, holding more capital, and producing more output per worker than comparable non-patenters. They had higher survival ability and were less likely to exit their industry. They also operated in more distinct industries, indicating broader scope rather than mere specialization. Estimating historical establishment-level production functions, we find that innovation primarily raised their physical productivity (TFPQ) rather than price-cost markups. The effects are highly heterogeneous: exit reduction is largest for small establishments, TFPQ gains peak among mid-sized establishments, and markups increase only among the very largest establishments. However, women-owned establishments do not experience any observable and comparable returns.

Keywords: Innovation, Entrepreneurship, Establishment Performance, Production Function Estimation, Productivity, Economic Development, Industrial Revolution

JEL Codes: D22, D24, N61, N71, O14, O31, O33, O34, L26, J16

*We thank Richard Hornbeck for sharing the Census of Manufacturers database and Terence Chau for providing their version of early matches. We thank John Van Reenen for his valuable feedback. We thank Jag Grewal for his excellent research assistance. All remaining errors are our own.

[†]Economic Department, HEC Paris, CEP LSE, and CEPR. Email: bergeaud@hec.fr.

[‡]Economics, Cardiff Business School & CEP LSE. Email: gozenr@cardiff.ac.uk.

[§]School of Economics, University of Bristol. Email: suleyman.gozen@bristol.ac.uk.

1 Introduction

How did innovative entrepreneurs perform during the nineteenth century? Were they more likely to survive, and benefit from higher productivity? Did small firms engage in innovation, and if so, how did they benefit from it? Were big firms more inclined to preserve the status quo and slow down their innovative efforts? These questions are central to modern debates in endogenous growth theory, where the gap between leading and lagging firms is believed to shape both the pace and distribution of both incremental and breakthrough innovation.

Yet, despite their contemporary relevance, such questions remain largely unanswered due to data availability for the era that marked the birth of corporate capitalism, leading into the Second Industrial Revolution. This was a period when steam power, mechanization, and managerial innovations creatively destructed production at scale, laying the institutional foundations of modern industry. And yet, we know remarkably little about how innovation translated into firm performance at the micro level during this time. In this paper, we aim to bridge that gap by providing a novel establishment-level[1] data on inventor entrepreneurs and firm dynamics in the early stages of American industrialization.

A rich literature has studied productivity growth, technological change, concentration, and direction of innovation with leader and laggard firms in modern economic contexts (e.g., Akcigit et al. (2021); Akcigit and Kerr (2018); Autor et al. (2020); Philippon (2019); Akcigit and Ates (2023); Graham et al. (2018); Balasubramanian and Sivadasan (2011)), with also more historical focus (e.g., Hornbeck and Rotemberg (2019); Hornbeck and Moretti (2018); Atack (1986); Berger and Ostermeyer (2025)). Yet systematic firm-level evidence on innovation production for the nineteenth century remains scarce. Understanding whether innovation in the 19$^{\text{th}}$ century was driven by large firms or by small and medium-sized businesses, and whether invention translated into higher productivity and longer survival offers fundamental insight into the dynamics of industrial growth and the inclusivity of technological change.

To address these questions, we construct a new and novel dataset that links (i) businesses, (ii) patents, and (iii) demographics. For businesses, we use the newly digitized Census of Manufactures (CMF) data from 1850 to 1880, produced by Richard Hornbeck, his co-authors, and team. The CMF provides the first microeconomic and most comprehensive data on establish-

[1] We use the terms *establishment* and *firm* in our analysis interchangeably.

ments with economic measures of inputs, outputs, capital investments, wages, employment, ownership, and more. We further have the information on gender of the establishment owner in the CMF through Gozen et al. (2024). Since the CMF data has the establishment owner's name and/or company name, we match these establishment owners to the inventor–assignees in the Comprehensive U.S. Patent (CUSP) (Berkes, 2018) database from 1840 to 1900. Furthermore, we link these entrepreneurs to the census of population records from 1840 to 1900, following Andrews et al. (2024). As a result, we provide the first systematic micro-level dataset of entrepreneurs and firms, their patents and economic performance, and demographic information for the period 1850–1880. To our knowledge, this is the first paper to provide such detailed data and analysis on 19$^{\text{th}}$-century businesses with innovations during a critical era of technological advancement and industrialization.

Our analysis yields several core findings. First, innovative firms had larger employment, paid higher wages, produced more output, and exhibited higher labor productivity relative to non-innovative firms[2]. Innovative firms were also substantially more likely to survive than their non-innovative counterparts. We further find that innovation translated strongly along the extensive margin: patenting firms were dramatically less likely to exit their industry, suggesting that intellectual property functioned not only as a productivity enhancer but also as a mechanism for long-run persistence in a period of high entrepreneurial turnover. Moreover, patenting establishments operated across a larger set of distinct industry categories within a given census year, indicating that innovation facilitated scope expansion and diversification, rather than merely deepening existing specialization.

Gains from innovation, however, were heterogeneous and varied systematically with firm size. Although establishments with more than 100 employees represent only 1% of the sample, they consistently generated both incremental and breakthrough innovations at markedly higher rates throughout the period. Contrary to modern findings that dominant firms may reduce innovation effort once they secure market leadership (e.g., Akcigit et al. (2023); Akcigit and Ates (2023)), large 19th-century firms remained actively innovative.

To disentangle these channels, we estimate establishment-level production functions and recover both physical productivity (TFPQ) and historical markups in the spirit of Loecker and

[2]The activities are recorded at the establishment level. Therefore, we refer to these establishments as single-establishment "firms", as in Hornbeck et al. (2024).

Warzynski (2012). We find that, on average, innovation is associated with higher productivity rather than higher price-cost margins: innovative establishments on average were more efficient but did not systematically charge higher markups over marginal cost. However, there is sharp heterogeneity in who benefited from productivity gains and markups. Mid-sized and smaller innovative establishments experienced the strongest productivity gains, suggesting that innovation was a genuine driver of efficiency. By contrast, only the very largest firms were able to translate innovation into higher markups, and even then, these gains were not accompanied by meaningful improvements in TFP. This stands in stark contrast to the modern economy, where innovation by dominant firms is frequently associated with rent preservation and weaker diffusion. In our historical setting, competitive pressure from below remained strong, allowing innovation to operate primarily as a productivity-enhancing force rather than a mere tool of market power consolidation.

And who contributed to the business and innovation dynamism during this period? Using the match between inventor business owners in the Census of Population, we find that both immigrant and domestic entrepreneurs contributed greatly to business and innovation dynamism. More strikingly, immigrant entrepreneurs engaged in innovative activity at an identical rate to domestic business owners even though they were underrepresented in the sample. This finding highlights that immigrants contributed significantly to both business formation and innovation dynamism during the early Second Industrial Revolution. From the gender aspect, we are only able to match a limited number of patents to women-owned establishments, and we do not detect statistically significant performance differences relative to male-owned firms.

Overall, our results suggest that patenting is strongly associated with superior firm performance in the historical context along multiple dimensions: patenting establishments are larger in terms of output, capital, employment, and wages; they are less likely to exit; they operate across a broader set of industries (greater scope); and they exhibit higher physical productivity (TFPQ). However, these average gains mask important heterogeneity in how innovation translates into firm outcomes. Along the extensive margin, patents matter most for small establishments, for whom the risk of exit is highest and intellectual property provides a disproportionately valuable survival advantage. Along the intensive margin, innovation is associated with higher TFPQ primarily among mid-sized firms. By contrast, increases in markups are concentrated only among the very largest establishments. Taken together, the findings reveal that innovation

does not generate uniform returns, but instead interacts systematically with firm size to shape both survival and performance dynamics.

The novelty of our study is fourfold. First, we introduce the most comprehensive micro-level dataset during the early stages of U.S. industrialization to date linking entrepreneurs, establishments, patents, and demographics, creating a resource for future research on innovation and industrialization. By linking innovation to establishment outcomes in the formative decades of the Second Industrial Revolution, our study highlights the micro-foundations of growth and provides a new perspective on how intellectual property interacted with entrepreneurship during one of the most transformative eras of economic development. Second, we provide the first systematic firm-level evidence on the returns to innovation during this episode by documenting evidence along both the intensive and extensive margins of establishment performance. We show that these returns are highly heterogeneous across the firm-size distribution, highlighting the importance of scale and production capacity in realizing the gains from innovation, which also provides crucial insights to the modern-time innovation dynamics. Third, we measure establishment-level markups and physical productivity (TFPQ) using the frontier production function estimation methodology, and to our knowledge, this is the most comprehensive historical study to jointly examine innovation, market power, and productivity at the establishment level. Fourth, we distinguish between the quantity and quality of innovation by separately analyzing breakthrough patents. This allows us to assess not only whether firms invented but whether the type of invention mattered, which is also a relevant discussion for contemporary concerns over innovation quality and technological disruption.

Related Literature Our study connects to several strands of literature in economic history, innovation, and entrepreneurship. First, we build on the long tradition of research on the nineteenth-century U.S. industrialization. Classic work has emphasized the aggregate drivers of productivity growth, technological change, and market integration (e.g., Hornbeck and Rotemberg (2019); Hornbeck and Moretti (2018)). More recent contributions have used linked census data to examine firm-level dynamics, including the diffusion of steam power and the role of railroads in shaping manufacturing productivity (Hornbeck et al. (2023)). By providing systematic micro-level evidence on the returns to patents, we complement this literature by showing how individual establishments and entrepreneurs translated invention into size, wages, and survival

during the critical decades leading to the Second Industrial Revolution.

Second, our paper is also related to the literature on innovation and firm performance. Studies using twentieth- and twenty-first-century data have documented that patents are associated with firm growth, higher wages, and survival advantages (e.g., Balasubramanian and Sivadasan (2011); Graham et al. (2018)). However, much of this evidence comes from modern administrative data, leaving open the question of whether similar relationships held in earlier eras when industrial organization, financing, and the patent system were still evolving. By uncovering parallel patterns in nineteenth-century manufacturing, our results suggest that the links between invention and firm outcomes are both historically deep and institutionally contingent.

Third, we relate our evidence to modern debates on leader versus laggard firms, and the productivity slowdown. Step-by-step growth models emphasize how competitive pressures and the leader–laggard gap shape innovation incentives and diffusion (e.g., Aghion et al. (2005)). A growing empirical literature links widening leader–follower gaps, rising concentration/markups, and weaker diffusion to slower aggregate productivity growth (e.g., Akcigit and Ates (2023); Akcigit et al. (2023); also the concentration/markup debate and its macro implications). Our results provide a historical counterpoint with nuance: large establishments in our period remained actively innovative, but the productivity payoff of innovation was strongest among small–mid-sized establishments, while only the very largest translated innovation into higher markups. This pattern suggests that in a high-dynamism historical setting, where competitive pressure from below remained strong, innovation functioned primarily as a productivity-enhancing force on average rather than a mere vehicle for rent preservation. By contrast, modern concerns center on innovation by dominant firms reinforcing market power and weakening diffusion. We thus contribute a period-specific lens on how the allocation of innovative activity across the size distribution conditions whether innovation scales TFP or instead amplifies price-cost margins.

Finally, we contribute to work on inclusive institutions and the social distribution of innovation. Recent studies examine the roles of women and Black inventors in U.S. innovation history (Andrews (2023); Andrews et al. (2024); Gozen (2024)). We add by examining whether women-owned establishments realized comparable returns to patents as men- or company-owned firms. However, we find muted average returns over innovation, alongside limited sample size.

In a nutshell, our paper is the first to jointly evaluate all of these contemporarily debated mar-

gins in a historical context. Our evidence shows that these mechanisms were indeed present but operated in fundamentally different ways: innovation translated into productivity rather than markups for most firms (with only the largest establishments exhibiting higher markups), exhibited highly size-dependent returns, dramatically improved survival (especially at the bottom of the size distribution), was shaped by demographic inclusivity, and yielded systematically different outcomes when the innovations were breakthrough rather than ordinary patents.

The remainder of the paper is structured as follows. Section 2 introduces the datasets and details the construction of key variables. Section 3 discusses the sample composition and descriptive statistics, while Section 4 presents the core empirical findings. Section 5 presents the robustness checks and outlines the directions for future research. Section 6 concludes by summarizing the main insights and implications of our analysis.

2 Data and Measurement

In this section, we describe the construction of our novel dataset, which links nineteenth-century Census of Manufactures records, historical patent data from CUSP, and individual-level census information. We detail the steps in the matching process and discuss the measurement of key variables, including establishment performance, and patenting activity.

Census of Manufacturers (CMF). The backbone of our dataset is the Census of Manufactures (CMF) panel for 1850, 1860, 1870, and 1880, originally constructed by Richard Hornbeck and co-authors. The CMF records establishment-level information in the manufacturing sector on economic indicators such as output, capital, labor, wages, products, and owner names. This panel has been further developed and used in a series of papers examining productivity, steam power, the role of railroads, and gender aspects (e.g., (Hornbeck and Moretti, 2018; Hornbeck and Rotemberg, 2019; Hornbeck et al., 2023; Gozen et al., 2024)). We use the most recent version of the CMF data (September 2025), which incorporates gender information for each establishment owner from Gozen et al. (2024).

Comprehensive Universe of U.S. Patents (CUSP). We link these establishments from CMF to the universe of U.S. patents from 1840–1900, drawing on the Comprehensive U.S. Patent (CUSP) dataset (Berkes, 2018). CUSP is the current "gold standard" for historical

U.S. patents, providing standardized inventor and assignee names, precise location, technology classifications, and filing details. In the CMF, respondents could report either the name of the company or the name of the establishment owner. For example, an establishment might be listed as "Prattvill Manufacturing Co." or as "William Pickett." In the patent data, two fields can be exploited: the inventor's name and the assignee. Assignees are the companies or individuals who are the owner and controller of the patent and may typically contain company (and/or individual) names. Inventors, on the other hand, include the real inventors and provides person names. Therefore, our matching procedure follows best practices: establishments with company-style names are linked to company assignees using Jaro–Winkler distances with a 0.2 cutoff and subsequent hand verification, while people-named proprietors are linked to inventors by splitting first and last names (using 1880 Census frequencies) and applying a 0.05 cutoff. This procedure yields 11,626 matched establishments (around 3% of CMF) and 27,272 patents (around 5% of all patents).[3] We then track establishments over time using establishment name, county, and industry information (as in (Hornbeck et al., 2023)). If a company appears only in one Census year, we classify it as an exit. If a company retains the same establishment name and county but changes industry classification, we also treat this change as an exit of one firm and the entry of a new one.

Breakthrough Patents. To proxy for the "quality" of a patent, we combine manufacturer-patent data with breakthrough innovation information using the data by Kelly et al. (2021). Breakthrough innovations are calculated based on the backward and forward patent text similarity. If a patent represents a divergence from the prior art and influences future innovations, then it is considered to be an impactful patent. Hence, we analyze the type of firms that had breakthrough patents and their economic performance compared to firms without a breakthrough patent.

Census of Population. Finally, we connect the manufacturer–patent panel to the Census of Population, following the approach of Andrews et al. (2024). This allows us to link patent owners to individual-level census records. Using owner names, county of residence, and demographic identifiers, we recover inventor-entrepreneurs' age, gender, birthplace, literacy, race, marital status, and household characteristics. This step enables us to situate innovation and firm

[3]For the present analysis, we use 1850, 1860, and 1870 data, as the 1880 wave matching is still under construction.

performance within the broader demographic and social context of nineteenth-century America.

The resulting linked dataset of "establishments, patents, and individuals" provides the first systematic micro-level evidence on entrepreneurship, innovation, and performance in the formative decades of U.S. industrialization.

3 Sample Composition and Descriptive Statistics

This section discusses the sample composition of our key variables and related descriptive statistics.

3.1 Sample Composition

Our linked dataset comprises 449,417 manufacturing establishment-year observations from the 1850, 1860, and 1870 Censuses of Manufactures, which is a period of accelerated industrial transition during which steam power, factory organization, and early forms of capital-intensive production diffused unevenly across the United States. As shown in Table B.1, the sample expands meaningfully over time, with 45.3% of establishments recorded in 1870, reflecting both organic industrial growth and improved enumeration capacity.

There is pronounced heterogeneity across industries, as documented in Table B.1. Clothing and Textiles accounts for 17.2% of establishments, followed by Food Products and Lumber (each 12.4%), and Iron and Steel (12.3%), while industries such as Paper and Printing or Other Consumer Goods represent less than 2% of the sample. Overall, the sample displays meaningful variation both over time and across industries, which is critical for the empirical analysis that follows.

In terms of geographical heterogeneity, Table B.2 shows that the sample is highly concentrated in the industrial centres of the Northeast and Midwest: New York and Pennsylvania, accounting for more than one-third of all establishments (17.0% and 16.8%, respectively), followed by Ohio (7.2%) and Massachusetts (5.0%). In contrast, Southern and Western states such as Nevada, Montana, Utah, and Washington each contribute around 0.2% of the sample. This pronounced spatial concentration reflects the historical clustering of nineteenth-century manufacturing ac-

tivity in the Northeastern corridor and Great Lakes region.

Table B.3 reports the summary statistics of our key variables, and it shows that establishments in the sample exhibit substantial variation in size: log revenue, log output, and log capital have means between 5 and 8, but with wide dispersion, especially in output and materials. Employment and wages also display considerable cross-sectional heterogeneity. Patent activity is highly skewed, where the mean number of patents is close to zero, with the median equal to zero and a long right tail reaching up to 271 patents. Breakthrough patents follow an even rarer distribution. The number of observations varies across variables due to availability, with wage data being the most limited. Figure C.1 shows the histogram of our establishment-level economic variables (e.g., revenue, output, capital, employment, and wages), which displays substantial right-skewness, with a long upper tail driven by a relatively small set of very large firms. Figure C.2 confirms that patenting activity is even more extremely skewed, with the vast majority of establishments holding zero patents while a small minority account for very large patent portfolios, a pattern that is even more pronounced for breakthrough inventions.

Table B.4 shows clear and systematic differences in firm characteristics across establishment size groups. As expected, larger establishments (Panels C and D) operate at a substantially higher scale with markedly higher output, capital, employment, and wages relative to very small firms (Panel A). Output per worker and wage levels both rise monotonically with employment size, consistent with increasing productivity and labor quality among larger establishments. Patenting activity displays an even sharper gradient: while firms with fewer than 10 employees hold virtually no patents on average, establishments with more than 100 workers exhibit an order of magnitude higher patenting intensity, including breakthrough patents. Therefore, it shows that establishments with over one hundred employees account for a disproportionate share of both breakthrough and non-breakthrough patents.

Table B.5 shows the raw pairwise correlations between our key variables. Core establishment characteristics (such as revenue, output, employment, capital, and wages) are all positively and strongly correlated with one another, consistent with larger establishments tending to operate at scale across multiple dimensions. Patent activity exhibits only weak correlations with firm size or input measures, but the correlation between total patents and breakthrough patents remains extremely high (0.98), reflecting the fact that breakthrough innovation is heavily concentrated

among establishments with a high volume of patenting activity.

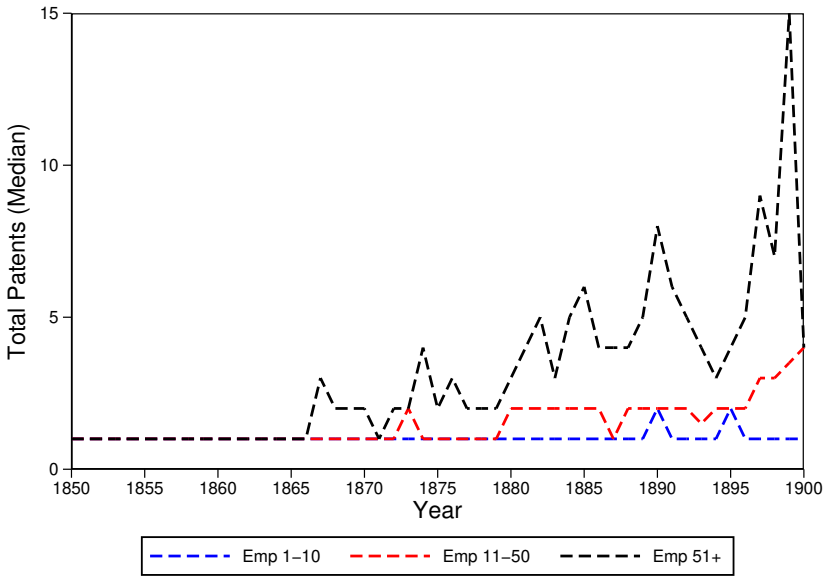
Figures C.3 and C.4 document heterogeneity across industries: average size, capital intensity, and labor scale vary by industry, and patenting activity is higher in a subset of industries such as Other Metal Products, Other Consumer Products, and Iron and Steel Products.

3.2 Descriptive Statistics

In Figure 1, we document how patenting activity evolved over time across the firm-size distribution. We plot the median number of total patents at the establishment level separately for small (1-10 workers), medium (11-50 workers), and large (50+ workers) establishments. We find that patenting rises steadily over time for all size groups, consistent with the broader surge in inventive activity during the Second Industrial Revolution. However, the smallest establishments consistently patent very little throughout the period, with their median patent count remaining close to zero. By contrast, patenting increases much more sharply among medium and especially large establishments, particularly from the 1860s onward.

Overall, the descriptive patterns highlight that larger, and capital-intensive establishments became the driver of the bulk of innovative output despite the numerical dominance of small producers as early as the mid-second half of the 19^{th} century.

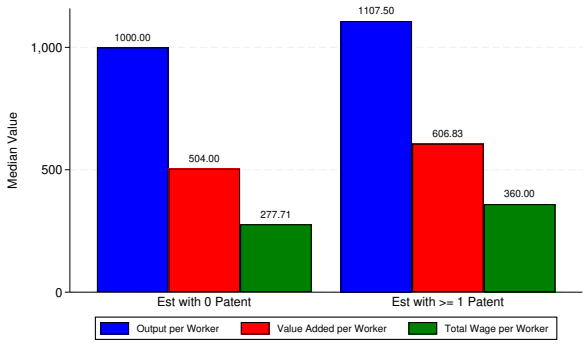
Figure 1: Patenting over Time by Establishment Size



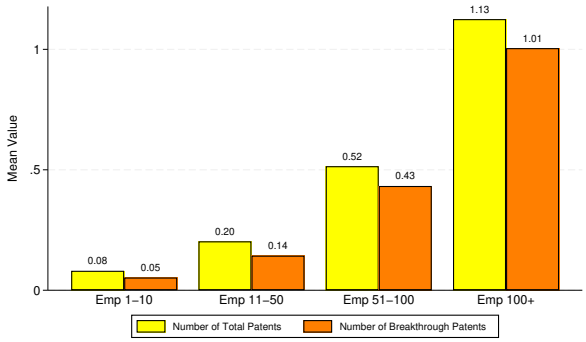
Notes: This figure plots the median number of total patents at the establishment level across establishment size categories. Establishment size categories are defined based on the sample median of establishment-level employment values.

However, the mere emergence of "superstar" establishments does not, by itself, guarantee a productivity-enhancing reallocation. A natural question is therefore whether patenting firms in fact exhibited superior performance, and along which margins such advantages materialized. Figure 2 Panel (a) provides a descriptive fact. Establishments with at least one patent systematically outperform otherwise comparable non-patenting firms across economic outcomes. On median, patenters exhibit higher output per worker, higher value added per worker, and substantially larger total wage bills, which is a pattern consistent with innovation being associated with greater productive capacity and a stronger demand for labor, even within this early industrial setting. Figure 2 Panel (b) further documents the size dependence of inventive activity: both total patents and breakthrough patents rise sharply with establishment employment, with the largest producers (100+ workers) generating the overwhelming share of high-impact inventions. This pattern indicates that inventive activity was far from evenly distributed. Instead, it was already disproportionately concentrated among the upper tail of the firm-size distribution.

Figure 2: Establishment Performance Comparison



(a) Establishment-level Variables by Patent Dummy



(b) Number of Patents by Establishment Size

Notes: Panel (a) reports the sample median establishment-level variables of output per worker, value added per worker, and total wage per worker by patenting status. Panel (b) presents the sample mean of patent variables (the number of total patents, and the number of breakthrough patents) by establishment size categories.

4 Empirical Evidence

In this section, we construct our hypotheses and test them by bringing empirical evidence.

Hypothesis 1. **Innovative establishments exhibit superior economic performance and operate at larger scale than non-innovative establishments.**

To better understand the key differences between innovative and non-innovative firms, we document the differences in terms of size, wages, capital, output, productivity, survival, and exit probability, while also examining heterogeneity across firm sizes.

Table 1 presents baseline regressions comparing the performance of patenting and non-patenting establishments between 1850 and 1870. Across all major economic indicators such as employ-

ment, wages, capital, output, and output per worker, patenting firms are significantly larger and more productive than their non-patenting peers. The coefficient magnitudes imply that patenters employed around 15% more workers, paid slightly higher wages, held 27% more capital, and generated roughly 8% higher labor productivity (measured by output per worker), conditional on year, industry, and county fixed effects.

Table 1: All Businesses with and without Innovation (1850-1870)

	(1)	(2)	(3)	(4)	(5)
	log(employment)	log(avg male wages)	log(capital)	log(output)	log(output per worker)
Has patented (=1)	0.150***	0.125***	0.272***	0.227***	0.0775***
	(0.0412)	(0.0331)	(0.0709)	(0.0588)	(0.0199)
Year FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
R^2	0.157	0.273	0.166	0.135	0.144
N	446906	234638	437705	446906	446906

Notes: This table shows the performance of all establishments with patents compared to the establishments without patents. All columns include a year fixed effects, industry fixed effects, and county fixed effects. Standard errors in parentheses are clustered at the county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The results also speak to mechanisms. Patenting likely enhanced firms' ability to protect rents, secure financing, and expand output without immediately facing competitive erosion. The capital differential, in particular, suggests that innovative establishments were able to mobilize more resources, possibly by leveraging their intellectual property as a signal of quality or by attracting investors. At the same time, the positive association with labor productivity points to efficiency gains, possibly through mechanization, improved production processes, or higher human-capital intensity.

Patents may also differ in terms of impact. While incremental patents can be higher in numbers and still offer an advantage in the market, a single breakthrough invention may yield a disproportionate advantage due to its relatively more innovative and impactful content. To better understand the dynamics between the *quality* of patents and firm performance, we use "breakthrough innovation" data by Kelly et al. (2021). Their approach identifies breakthrough patents based on textual similarity: an innovation is classified as a breakthrough if it diverges significantly from prior art (i.e., exhibits low similarity to past patents) while at the same time

resembles future patents over the subsequent 5- or 10-year window (i.e., it shapes the trajectory of subsequent innovation).

Table 2: All Businesses with and without Breakthrough Innovation (1850-1870)

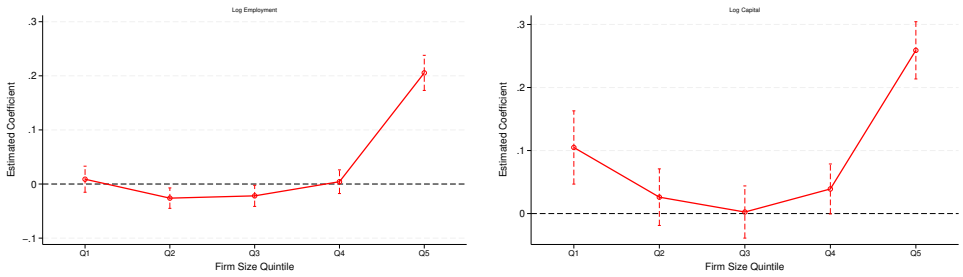
	(1)	(2)	(3)	(4)	(5)
	log(employment)	log(avg male wages)	log(capital)	log(output)	log(output per worker)
Breakthrough patent (=1)	0.217***	0.156***	0.358***	0.318***	0.101***
	(0.0675)	(0.0516)	(0.119)	(0.0986)	(0.0350)
Year FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
R^2	0.157	0.273	0.166	0.135	0.144
N	446906	234638	437705	446906	446906

Notes: This table shows the performance of all establishments with and without breakthrough patents. Breakthrough measure is *p80sim05* (top 20 percentile, 5 year window similarity). All columns include year, industry, and county fixed effects. Standard errors in parentheses are clustered at the county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

In Table 2, we continue testing our first hypothesis by comparing establishments based on whether the establishment has a breakthrough or impactful innovation or not. We find that establishments with breakthrough or more impactful innovations show more superior performance in labor productivity, and producing more output. They are larger in terms of employment, offer higher average monthly wages for workers, and have higher capital investment. Comparing our results with Table 1, we find that establishments with breakthrough innovations yield even more employment, higher wages, capital investment, and labor productivity.

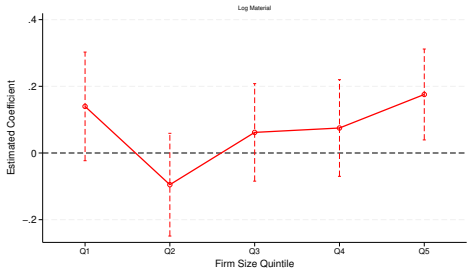
While our baseline results establish an average performance premium for patenting firms, an important question is whether these returns were evenly distributed across the firm size distribution. The beginnings of the Second Industrial Revolution was characterized not only by the spread of new technologies but also by the emergence of larger, more capital-intensive establishments, raising the possibility that economies of scale played a decisive role in capturing the benefits of invention. Examining heterogeneity by firm size therefore allows us to distinguish whether patents provided broad-based advantages to all entrepreneurs or disproportionately favored already larger firms.

Figure 3: Quintile Regressions - Establishment Variables

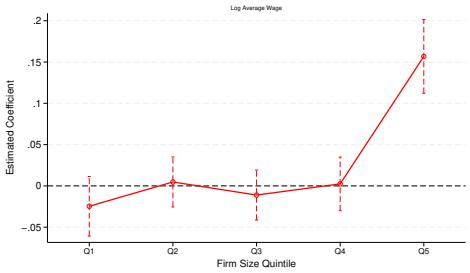


(a) Log Employment

(b) Log Capital



(c) Log Material



(d) Log Average Wage

Notes: This figure plots estimated coefficients from regressions of establishment-level variables on patenting status by firm size (total output) quintile within each census year and industry. Each coefficient measures the relative performance premium of innovative establishments compared to non-patenting peers within the same quintile. All regressions include year, industry and county fixed effects. Standard errors are clustered at the county level. Confidence intervals are shown at the 95% level.

Figure 3 examines whether the returns to innovation differ by firm size quintile. The results reveal strong heterogeneity: innovations are associated with higher gains for larger firms. In particular, the performance premium in employment, capital, and wages rises monotonically across quintiles, with the largest firms (top quintile) experiencing the highest positive association. This pattern suggests that economies of scale were central in translating patents into superior outcomes, amplifying the advantages of larger firms and limiting the returns to sole proprietors and small shops. Figure C.5 replicates the analysis using breakthrough patenting status, and we find robust evidence that larger firms with breakthrough innovations exhibit significantly higher employment, capital, and wages relative to comparable firms without any patents.

This heterogeneity carries important historical and modern implications. In the context of the Industrial Revolution, it suggests that patents may have reinforced the rise of larger enterprises, contributing to the early emergence of "big business" in U.S. manufacturing. In today's economy, the pattern resonates with debates about the scalability of intangible assets and the

dominance of innovation-driven "superstar firms." In both settings, intellectual property appears to magnify the advantages of already larger firms, shaping the distribution of firm growth and survival within industries.

Hypothesis 2. **Innovative establishments were more resilient, their survival probability was higher and exit likelihood was lower.**

In addition to structural differences between innovative and non-innovative establishments in the second half of the 19^{th} century, we also study survival likelihood of an innovative firm as they may have a higher chance to survive in the market. A patent grants monopoly rents over a certain period, which may help firms gain advantage. Therefore, in Table 3, we shift our focus from contemporaneous outcomes to survival across census waves. The survival estimates reveal a strikingly large association: innovative establishments (defined by their patenting status with respect to both total patents and breakthrough patents) are more likely to be observed in multiple census rounds than non-patenting peers, conditional on year, industry, and county. The result is highly statistically significant across specifications. This suggests that innovativeness is not only associated with better contemporaneous performance but also substantial advantages in firm longevity. The survival premium hints towards the long-run payoff to innovation during the formative decades of U.S. industrialization.

Table 3: Survival Probability with and without Innovation (1850–1870)

	(1)	(2)	(3)	(4)
	Survival			
Panel A: Patents				
Has patented (=1)	0.635***	0.642***	0.736***	0.722***
	(0.0934)	(0.0842)	(0.0912)	(0.0901)
Pseudo R^2	0.00186	0.00308	0.0155	0.0327
N	449387	449387	446929	446928
Panel B: Breakthrough Patents				
Has impact (=1)	0.760***	0.768***	0.881***	0.856***
	(0.121)	(0.113)	(0.116)	(0.110)
Pseudo R^2	0.00107	0.00227	0.0145	0.0317
N	449387	449387	446929	446928
Year FE	No	Yes	Yes	Yes
Industry FE	No	No	Yes	Yes
State FE	No	No	No	Yes

Notes: This table reports `logit` estimates of establishment survival on patenting indicators. The dependent variable is the survival of an establishment across census waves within the same county and industry, 1850–1870. Panel A uses an indicator for having any patent; Panel B uses an indicator for having an impact (breakthrough) patent. Standard errors in parentheses are clustered at the county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

We additionally examine the relationship between patenting and establishments' likelihood of exiting their industry. Exit is defined at the broad industry level and captures whether an establishment ceases to appear in its original industry in subsequent census years. For each establishment, we construct indicators of whether it is observed in 1850, 1860, and 1870. An establishment is classified as exiting if it is present in a given census year but does not reappear in the same broad industry in any later census year. For example, an establishment observed in 1850 but absent from its industry in both 1860 and 1870 is coded as an exit, as is an establishment present in 1860 but missing from that industry in 1870. Establishments that continue to appear with the same establishment name, in the same county, and same industry are coded as non-exiting.

Table 4: Exit Likelihood with and without Innovation (1850–1870)

	(1)	(2)	(3)
		Exit	
Panel A: Patents			
Has patented (=1)	-1.369***	-1.297***	-1.155***
	(0.156)	(0.153)	(0.127)
Pseudo R^2	0.0675	0.0731	0.0990
N	240008	240006	224637
Panel B: Breakthrough Patents			
Has impact (=1)	-1.672***	-1.539***	-1.382***
	(0.245)	(0.239)	(0.207)
Pseudo R^2	0.0648	0.0701	0.0967
N	240008	240006	224637
Year FE	Yes	Yes	Yes
Industry FE	No	Yes	Yes
State FE	No	No	Yes

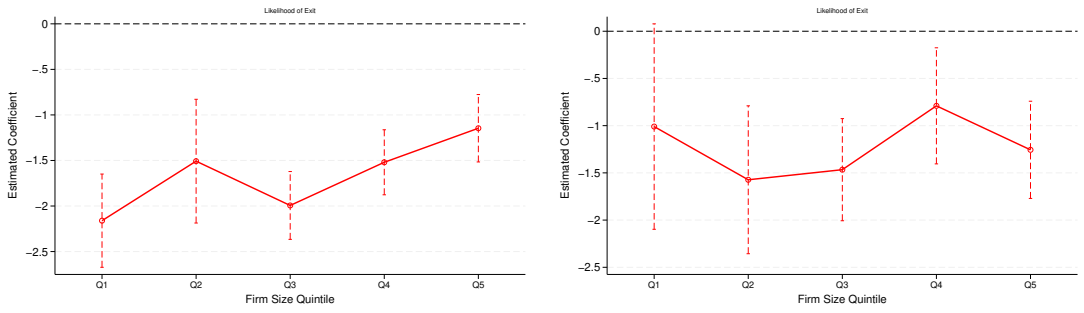
Notes: This table reports logit estimates of establishment exit on patenting indicators. The dependent variable (exit) equals one if an establishment ceases to appear in its original broad industry in subsequent census years. For each establishment, we track its presence in 1850, 1860, and 1870 and classify it as exiting if it is observed in a given census year but does not reappear in the same broad industry thereafter. Panel A uses an indicator for having any patent; Panel B uses an indicator for having an impact (breakthrough) patent. Standard errors in parentheses are clustered at the county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4 shows that innovative establishments (defined by their patenting status with respect to both total patents and breakthrough patents) are substantially less likely to exit the market than comparable non-patenting firms, even after conditioning on year, industry, and state fixed effects. The exit likelihood further declines when an establishment has a breakthrough innovation. Across all specifications, the estimated coefficients are large, negative, and highly statistically significant, implying economically meaningful reductions in exit risk associated with breakthrough-patenting. This result shows that the extensive margin (i.e. the decision of whether a firm continues to exist at all) was a key channel through which innovation translated into business advantage in the nineteenth-century manufacturing economy. Rather than merely expanding output or capital among survivors, innovations (espeically breakthrough innovations)

dramatically increased the likelihood of staying in the market, pointing to intellectual property as a potential and crucial mechanism for resilience during a period of high entrepreneurial turnover.

The exit likelihood may also be heterogeneous across firm size. Figure 4 documents that the protective feature of patenting against exit operates most strongly at the extensive margin precisely where exit risk is highest, i.e., among the smallest firms. The estimated coefficients on exit are negative across all quintiles, but the magnitude of the association is largest for firms in the bottom of the size distribution, and gradually becomes weaker for larger firms. This pattern is economically intuitive: smaller establishments face substantially higher baseline exit risk, and for them, securing a patent provides a disproportionately powerful shield against failure. In contrast, larger firms are already more likely to survive irrespective of patenting, so the marginal survival gain from innovation is smaller. Overall, our finding confirms that the extensive margin is central where patenting serves as a critical survival mechanism for otherwise vulnerable small firms, rather than merely reinforcing persistence among already dominant players.

Figure 4: Quintile Regression - Exit Likelihood



(a) Exit vs. Patents (b) Exit vs. Breakthrough Patents

Notes: This figure plots estimated coefficients of the association between patenting and exit likelihood following the logit estimation, comparing establishments with patents to those without across firm-size distribution. Panel (a) shows the estimates with respect to a patenting status in terms of total patents, and Panel (b) shows the estimates with respect to a patenting status in terms of breakthrough patents. All regressions include industry, year, and state fixed effects. Standard errors are clustered at the county level. Confidence intervals are shown at the 90% level.

Hypothesis 3. **Innovative establishments were more likely to expand their product or market scope.**

In this part, we investigate the association between patenting and establishment scope. We measure scope as the number of broad industry categories in which an establishment is simultaneously active within a given census year. This definition captures the extent to which a firm diversifies across distinct product or market segments, rather than expanding scale within a single line of activity. This measure is well suited for the nineteenth-century setting, where multi-product production and flexible deployment of capital and labor were common, and establishments frequently operated across related industries such as machinery, metals, and transportation goods. Counting the number of industry categories therefore provides a direct and economically meaningful proxy for the breadth of a firm's operations.

Table 5 examines whether patenting is associated with broader product or market scope. Across all specifications, establishments with at least one patent (in terms of both total patent and breakthrough patent counts) operate in significantly more broad industry categories within a census year. The estimated coefficients are remarkably stable and economically meaningful: innovative establishments are likely to have approximately 0.25 additional industry segments on average relative to non-innovative establishments. These results remain robust with the sequential inclusion of industry and state fixed effects, indicating that the relationship is not driven by industrial or regional composition. Overall, the findings suggest that innovation is strongly associated with greater diversification and scope expansion among nineteenth-century establishments.

Table 5: Scope with and without Innovation (1850–1870)

	(1)	(2)	(3)
		Scope	
Panel A: Patents			
Has patented (=1)	0.253***	0.254***	0.254***
	(0.00712)	(0.00720)	(0.00710)
Pseudo R^2	0.00101	0.00101	0.00102
N	435511	435446	435446
Panel B: Breakthrough Patents			
Has impact (=1)	0.245***	0.245***	0.243***
	(0.00886)	(0.00891)	(0.00871)
Pseudo R^2	0.000351	0.000352	0.000359
N	435511	435446	435446
Year FE	Yes	Yes	Yes
Industry FE	No	Yes	Yes
State FE	No	No	Yes

Notes: This table reports logit estimates of establishment scope expansion on patenting indicators. Scope is measured as the number of distinct industry categories reported by an establishment. Panel A uses an indicator for having any patent; Panel B uses an indicator for having an impact (breakthrough) patent. Standard errors in parentheses are clustered at the county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Hypothesis 4. **Innovative establishments had higher total factor productivity (TFP), while the association between innovation and markup is ambiguous and depends on establishment size.**

To further investigate the mechanism through which innovation translates into firm performance, we exploit a unique advantage of the historical *Census of Manufactures* (CMF); unlike most modern datasets, the CMF directly reports both physical quantities produced and output values, allowing us to separately observe quantities and prices at the establishment level. This enables a clean decomposition between physical productivity (TFPQ) and price-cost margins (markups).

We first estimate an establishment-level production function using the Levinsohn and Petrin (2003) method, with materials as the proxy input to address input choice endogeneity, labor as the freely varying input, and capital as the state variable. We assume a Cobb–Douglas

technology and, in our baseline, impose constant returns to scale (CRS), so that the output elasticities of labor, materials, and capital $(\theta^L, \theta^M, \theta^K)$ are directly obtained from the estimated coefficients.

From the estimated production function, we recover input elasticities, which we use to construct Loecker and Warzynski (2012) markups and physical productivity (TFPQ). Establishment-level *material-input* markups are defined as

$$\mu_{it}^{(M)} = \frac{\theta^M}{\alpha_{it}^M}, \qquad \text{where } \alpha_{it}^M \equiv \frac{P_{it}^M M_{it}}{P_{it}^Q Q_{it}} \tag{1}$$

where θ_{it}^M is the estimated output elasticity of materials, and the denominator is the expenditure share of materials in revenue. A higher μ_{it} therefore reflects greater product-market power.

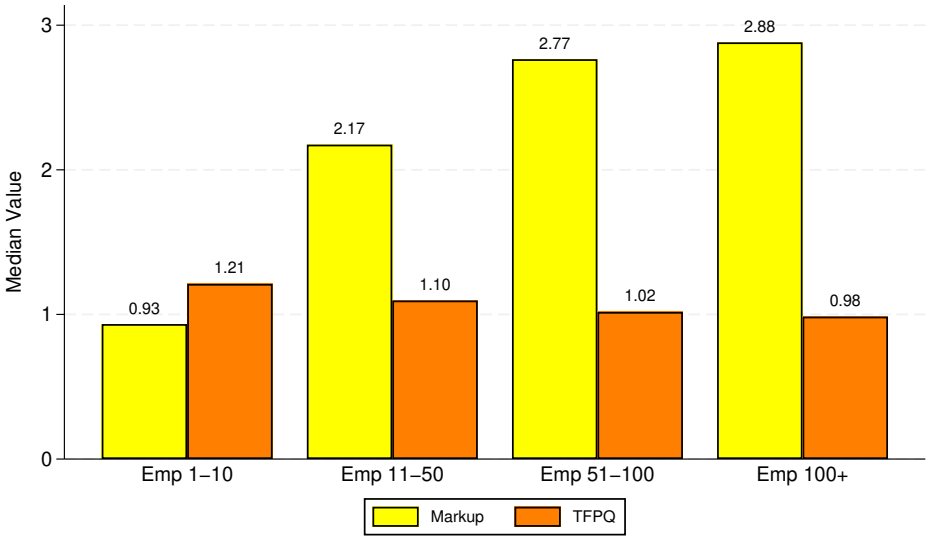
In parallel, we construct TFPQ as the establishment-level residual from the same physical production function, using log physical output (in quantities) net of predicted inputs multiplied by their estimated elasticities:

$$\text{TFPQ}_{it} \equiv \log Q_{it} - \theta_{it}^L \log L_{it} - \theta_{it}^M \log M_{it} - \theta_{it}^K \log K_{it}, \tag{2}$$

where Q_{it} is the physical quantity of output and (L_{it}, M_{it}, K_{it}) denote labor, materials, and capital inputs for each establishment i in Census year t. This ensures that TFPQ captures true technological efficiency, purged of any price or revenue component.

Before examining the role of patenting, we first document how markup and TFPQ behave across the establishment-size distribution. Figure 5 shows a clear divergence between the two margins: markup increases monotonically with establishment size, while TFPQ declines steadily as firms become larger. In other words, larger firms exhibit higher price-cost margins but lower physical productivity relative to smaller establishments. This pronounced heterogeneity across the size distribution suggests that firms may benefit from different performance channels, motivating our next step to investigate whether and how these patterns are systematically related to patenting and innovation activity.

Figure 5: Markup and TFPQ by Establishment Size



Notes: This figure reports the sample median of establishment-level markup and TFPQ by establishment size categories. Markup and TFPQ are constructed following Levinsohn and Petrin (2003) and Loecker and Warzynski (2012).

Table 6 presents the average association between patenting and outcomes along these two margins. We observe a statistically significant increase in TFPQ for innovative (in terms of both patent- and breakthrough patent-holding) establishments, while markups do not increase on average. This asymmetry suggests that innovation during the nineteenth century primarily operated through a productive efficiency channel rather than through immediate price-cost margin expansion in the average firm.

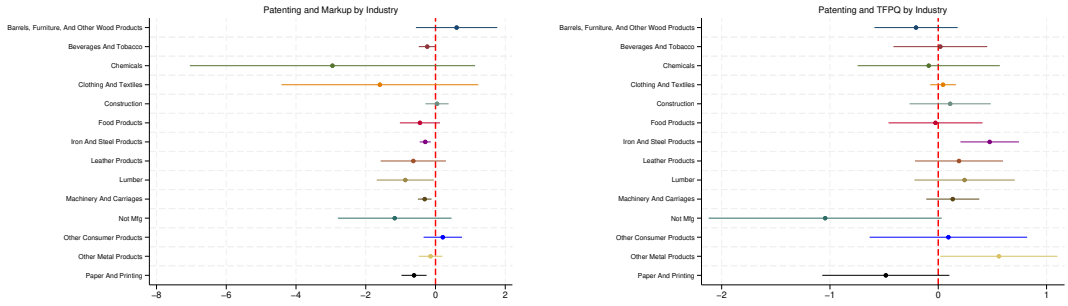
Table 6: Markup and TFPQ with and without Innovation

	(1) Markup	(2) TFPQ
Panel A: Patents		
Has patented (=1)	-0.0138	0.122***
	(0.0243)	(0.0470)
R^2	0.236	0.308
N	160721	159765
Panel B: Breakthrough Patents		
Has impact (=1)	-0.0139	0.158**
	(0.0462)	(0.0790)
R^2	0.236	0.308
N	160721	159765
Year FE	Yes	Yes
Industry FE	Yes	Yes
County FE	Yes	Yes

Notes: This table reports estimates of the association between patenting and establishment-level markup and TFPQ, comparing establishments with patents to those without. Markup and TFPQ are constructed following Levinsohn and Petrin (2003) and Loecker and Warzynski (2012). Standard errors are shown in parentheses are clustered at the county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

We also investigate the industrial heterogeneity for our baseline results. Figure 6 documents substantial heterogeneity in how patenting relates to establishment markup and TFPQ across industries. Panel (a) shows that the estimated association with markups are far from uniform: while industries such as Iron and Steel Products, and Machinery and Carriages exhibit economically meaningful declines in markups associated with patenting, several others display statistically insignificant relationship. Panel (b) similarly reveals heterogeneous associations with TFPQ. In some industries, such as Iron and Steel Products, patenting is linked to positive productivity gains, whereas in industries like Not Manufacturing (No Mfg), the association is negative. Overall, the figure highlights that the implications of patenting are highly industry dependent and differ in both magnitude and sign.

Figure 6: Industry Estimates - Markup and TFPQ



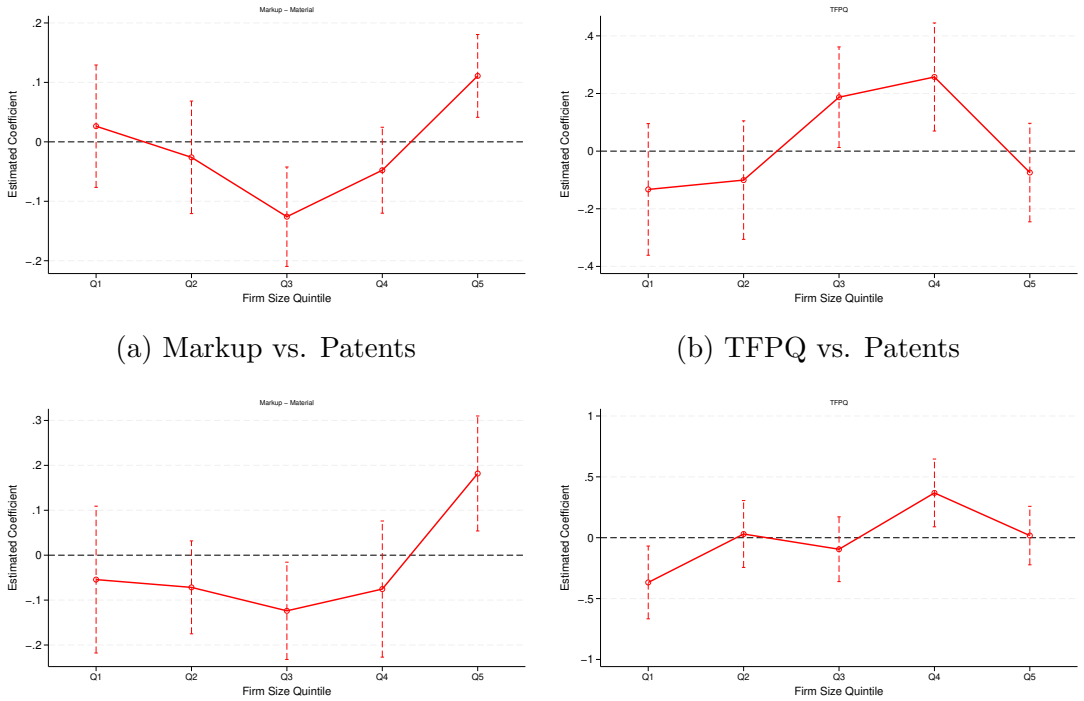
(a) Markup (b) TFPQ

Notes: This figure plots estimated coefficients of the association between patenting and establishment-level markup and TFPQ, comparing establishments with patents to those without across broad industry category. Markup and TFPQ are constructed following Levinsohn and Petrin (2003) and Loecker and Warzynski (2012). All regressions include industry-by-year, and county fixed effects. Standard errors are clustered at the county level. Confidence intervals are shown at the 90% level.

After we document an average estimates, we also examine the heterogeneity of our estimates across firm-size distribution. Figure 7 reveals that this average masks strong and highly nonlinear heterogeneity. Panel (b) shows that TFPQ gains are increasing across the size distribution up to the fourth quintile, while the coefficient in the top quintile is statistically insignificant. Panel (a) shows that the association between patenting (in terms of both total patent and breakthrough patent counts) and markup is close to zero for small and medium firms, but become economically large and have statistical significance in the top quintile. In other words, only the largest establishments that have the capacity to scale convert innovation gains into market power.

Taken together, these results suggest that patenting is positively associated with physical total factor productivity (TFPQ) for medium-scale establishments, while for the very largest establishments it is instead associated primarily with higher markups rather than with additional TFPQ gains. In other words, innovation in nineteenth-century manufacturing appears to raise productivity broadly, but only the largest establishments are able to convert it into product-market power.

Figure 7: Quintile Regressions - Markup and TFPQ



(a) Markup vs. Patents

(b) TFPQ vs. Patents

(c) Markup vs. Breakthrough Patents

(d) TFPQ vs. Breakthrough Patents

Notes: This figure plots estimated coefficients of the association between patenting and establishment-level markups and TFPQ, comparing establishments with patents to those without across firm-size distribution. Markup and TFPQ are constructed following Levinsohn and Petrin (2003) and Loecker and Warzynski (2012). Panel (a)-(b) show the estimates with respect to a patenting status in terms of total patents, and Panel (c)-(d) show the estimates with respect to a patenting status in terms of breakthrough patents. All regressions include year-by-industry, and county fixed effects. Standard errors are clustered at the county level. Confidence intervals are shown at the 90% level.

Hypothesis 5. **The returns to innovation depend on the entrepreneur's demographics: Immigrant entrepreneurs were equally likely as native-born entrepreneurs to operate establishments with breakthrough innovations whereas women entrepreneurs were less likely so.**

We further combine our establishment owner-inventor data with the Census of Population from 1850 to 1900. We use the data from Andrews et al. (2024) where inventors are already matched with the Census of Population. Hence, we provide evidence on only establishment owners with at least one patented innovation with a breakthrough measure included.

Our main comparison includes immigrant versus domestic entrepreneur-inventors. We find that 61% of non-immigrants had at least one breakthrough patent/establishment, whilst 60% of immigrants had at least one breakthrough patent/establishment.

This finding suggests that immigrants and non-immigrants, who were also establishment owners, were equally likely to produce and use breakthrough patents. However, among entrepreneurs with at least one breakthrough innovation, 15.4% were immigrants and 84.6% were non-immigrants. This suggests that although immigrants comprise a much smaller portion, their likelihood of running a business with a breakthrough innovation was equally likely as non-immigrants. In terms of race, white entrepreneurs make up 96.2% in breakthrough patent holding compared to Black entrepreneur-inventors. 54% were married and breakthrough-patent holders had a similar share of marriage rate as well.

Table 7: Women-Owned Businesses with and without Innovation (1850-1870) - Establishment Variables

	(1)	(2)	(3)	(4)	(5)
	log(employment)	log(avg male wages)	log(capital)	log(output)	log(output per worker)
Has patented (=1)	-0.036	-0.064	0.412	0.069	0.106
	(0.107)	(0.210)	(0.308)	(0.190)	(0.172)
Year FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
R^2	0.265	0.411	0.282	0.279	0.279
N	2714	806	2714	2714	2714

Notes: This table shows the performance of women-owned establishments with patents compared to the women-owned establishments without patents. All columns include a year fixed effects, industry fixed effects, and county fixed effects. Standard errors in parentheses are clustered at the county level.

We further extend the analysis to women establishment owners or women entrepreneurs. We find only 17 establishments with a patent out of 2,867 women-owned establishments between 1850 and 1870, which is a very small number compared to men-owned establishments. In Table 7, we restrict the sample to only women-owned establishments with and without patents. Unlike the all-firm results, women-owned businesses with patents do not systematically employ more workers, pay higher wages, or generate higher output per worker compared to their non-patenting women counterparts. The only positive and relatively larger association appears with capital, though the estimate is imprecise. This suggests that the returns to innovation were either muted for women entrepreneurs or that structural constraints limited their ability to convert patents into measurable firm performance advantages during this period.

These muted estimates suggest that the returns to innovation were either more limited for women entrepreneurs or that structural barriers prevented them from fully capitalizing on inventive activity. Legal and institutional barriers, including restricted property rights and limited access to credit, may also have curtailed women's ability to translate intellectual property into firm expansion.

5 Robustness and Future Work

To assess the stability and interpretive credibility of our baseline findings, this section conducts a series of robustness exercises and outlines directions for future research. We begin by examining whether our results are sensitive to alternative regression specifications and fixed-effect structures, ensuring that the innovation premium we document is not an artifact of other potential unobserved heterogeneity. We then incorporate production input controls to address concerns that our estimates may reflect omitted-variable bias rather than capturing the underlying association. Finally, we conclude the section by outlining promising extensions that would further deepen our understanding of the relationship between innovation and establishment dynamics in nineteenth-century industrial America.

5.1 Alternative Specifications

In this section, we provide additional robustness checks for our baseline results by estimating specifications with alternative fixed effect structures through including year-by-industry fixed effects and more granular industry-category fixed effects rather than broader industry fixed effects.

Table 8 strengthens identification by replacing year and industry fixed effects with year-by-industry interactions. The results remain nearly identical to those in Table 1 such that patenting firms are larger, more labor- and capital-intensive, more productive, and pay higher wages. The stability of the estimates suggests that the performance premium of patenters is not driven by time-varying industry shocks or differences in local specialization. Instead, the returns to innovation appear to be robust across industries and time, reinforcing the idea that patenting was systematically linked to superior firm outcomes.

Table 8: All Businesses with and without Innovation (1850-1870) - Year-by-Industry Fixed Effect

	(1)	(2)	(3)	(4)	(5)
	log(employment)	log(avg male wages)	log(capital)	log(output)	log(output per worker)
Has patented (=1)	0.149***	0.125***	0.271***	0.227***	0.0782***
	(0.0411)	(0.0333)	(0.0707)	(0.0585)	(0.0197)
Year-by-Industry FE	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
R^2	0.169	0.274	0.175	0.141	0.146
N	446906	234638	437705	446906	446906

Notes: This table shows the performance of all establishments with patents compared to the establishments without patents. All columns include a year-by-industry fixed effect, and county fixed effect. Standard errors in parentheses are clustered at the county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

This robustness is important for interpretation. It indicates that the premium was not confined to a handful of technologically dynamic industries but reflected a broad-based relationship between invention and performance across nineteenth-century manufacturing.

Table B.6 provides a robustness check with detailed industry instead of broad industry category used in our baseline specification. The results confirm the baseline findings reinforce the robustness of the innovation premium documented in Tables 1 and 8.

Table 9: Women-Owned Businesses with and without Innovation (1850-1870) - Year-by-Industry Fixed Effect

	(1)	(2)	(3)	(4)	(5)
	log(employment)	log(avg male wages)	log(capital)	log(output)	log(output per worker)
Has patented (=1)	-0.037	-0.101	0.386	0.067	0.104
	(0.110)	(0.184)	(0.297)	(0.205)	(0.177)
Year-by-Industry FE	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
R^2	0.287	0.421	0.299	0.295	0.295
N	2714	806	2714	2714	2714

Notes: This table shows the performance of women-owned establishments with patents compared to the women-owned establishments without patents. All columns include a year fixed effects, industry fixed effects, and county fixed effects. Standard errors in parentheses are clustered at the county level.

Table 9 introduces year-by-industry fixed effects to the women-owned subsample. The results again show no consistent performance premium for women innovators, where the coefficients

are small and insignificant for employment, capital, wages, output, and labor productivity.

Table B.7 reports a robustness check for women-owned establishments along with detailed industry category. These results align with the main specifications (Tables 7 and 9), which again reveals muted or absent performance premiums for women innovators.

5.2 Patent Intensity and Establishment Variables

Up to this point, our baseline analysis has focused on whether an establishment holds at least one patent, capturing the presence rather than the scale of innovative activity. However, the intensity of inventive activity may also matter for firm performance. Establishments that generate multiple patents could exhibit stronger productivity and scale advantages than those that merely cross the innovation threshold. To test the robustness of our findings, we therefore replace the patent dummy with the logarithm of total patent counts and breakthrough patent counts, respectively.

Table 10 reports these results. Consistent with our baseline specification, we continue to observe a strong and statistically significant association between patenting and firm performance. A one-percent increase in total patents is associated with approximately 0.14% higher employment, 0.10% higher average male wages, 0.23% higher capital, and 0.20% higher output, with a modest but positive effect on output per worker. The coefficients remain of similar magnitude when focusing on breakthrough patents, suggesting that the productivity and scale advantages of innovation are not limited to incremental inventions.

Overall, this robustness check confirms that the relationship between innovation and firm performance is not driven merely by the binary status of being a patenting firm. Instead, the intensity of inventive activity reinforces the same pattern: establishments that innovate more frequently, and especially those producing breakthrough inventions, display higher output, employment, and capital accumulation, consistent with cumulative innovation effects during the early industrial period.

Table 10: All Businesses with and without Innovation (1850-1870) - Patent Intensity

	(1) log(employment)	(2) log(avg male wages)	(3) log(capital)	(4) log(output)	(5) log(output per worker)
Panel A: Total Patents					
Ln(Total Patents)	0.141***	0.105***	0.232***	0.195***	0.0539***
	(0.0469)	(0.0357)	(0.0777)	(0.0640)	(0.0185)
R^2	0.157	0.273	0.166	0.135	0.144
N	446,906	234,638	437,705	446,906	446,906
Panel B: Breakthrough Patents					
Ln(Br. Patents)	0.147***	0.103***	0.227**	0.195***	0.0486**
	(0.0532)	(0.0400)	(0.0892)	(0.0723)	(0.0208)
R^2	0.157	0.273	0.166	0.135	0.144
N	446,906	234,638	437,705	446,906	446,906
Year FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the association between patenting intensity and establishment-level outcomes. Panel A considers the natural logarithm of one plus total patent counts, while Panel B focuses on the natural logarithm of one plus breakthrough patents. All regressions control for year, industry, and county fixed effects. Standard errors in parentheses are clustered at the county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5.3 Including Control Variables

We exclude control variables from the baseline specifications to avoid conditioning on potential "bad controls" where production inputs such as capital, labor, and materials would be highly correlated. Conditioning on such variables would risk mechanically attenuating the association that we investigate. That said, as a robustness exercise, this section provides an alternative specification that includes the available production input variables to demonstrate that our baseline results are not purely driven by omitted-variable correlations.

Table B.8 and B.9 show the regression estimates, in terms of patenting and breakthrough patenting status. In both tables, we find that our baseline results are still valid such that establishments with patenting and breakthrough patenting are more likely to have higher employment, wages, capital and output compared to their counterparts without patenting conditional on our control variables.

5.4 Future Work and Extensions

Our analysis opens several promising avenues for future research. One natural extension is to exploit the technological classifications (CPC) embedded in patent records to investigate whether the performance premium associated with innovation varies across different technological domains. Such an approach would help reveal whether certain types of inventions were systematically more conducive to entrepreneurial success.

Another important direction involves enriching the analysis with demographic information from the population census (e.g. age, literacy, race, and birthplace) to better understand how individual characteristics shaped both the propensity to innovate and the returns to doing so. This would allow for a more nuanced view of innovation not just as a firm-level outcome, but as a decision embedded within the social fabric of the early U.S. economy.

A third avenue is a more systematic exploration of firm life-cycle dynamics. Examining whether innovation typically emerged at early stages, possibly to enable expansion, or instead followed initial growth would clarify how inventive activity interacted with long-term survival, path dependence, and scaling.

Relatedly, linking family relationships in census records would make it possible to study the role of family businesses and inherited wealth in sustaining inventive activity across generations. This could provide new insights into the intergenerational transmission of entrepreneurial capacity.

Finally, moving beyond descriptive patterns toward designs with stronger causal identification would enable sharper claims about the impact of innovation on firm performance, survival, and broader patterns of industrial development.

Together, these extensions promise to deepen our understanding of how invention and entrepreneurship interacted during the formative decades of U.S. industrialization.

6 Conclusion

This paper has provided the first systematic firm-level evidence on the relationship between invention, entrepreneurship, and performance during the formative decades of U.S. industri-

alization. By linking the Census of Manufactures with patent records and population census microdata, we constructed a novel dataset that allows us to examine not only which firms patented, but also how patents translated into outcomes such as firm survival, diversification, productivity, and market power.

We show that innovative establishments enjoyed a pronounced economic advantage: they were larger in terms of employment, capital, output per worker, and wages, operated across a broader range of industries (greater scope), and were substantially less likely to exit their industry across census waves. Estimating historical production functions, we find that innovation primarily raised physical productivity (TFPQ) rather than markups, indicating that technological efficiency rather than price-setting power was the dominant legacy of invention for most firms. Yet these associations were highly heterogeneous: survival gains were largest for small establishments, TFPQ improvements peaked among mid-sized firms, and only the very largest establishments translated innovation into higher markups, consistent with scale as a prerequisite for market power. By contrast, women-owned firms did not experience comparable returns, highlighting the presence of structural barriers to the diffusion of innovation gains.

These findings highlight the role of intellectual property not just as a productivity-enhancing asset, but as a mechanism shaping which firms survive, expand scope, and ultimately scale. Patents functioned as a form of protection against exit at the bottom of the size distribution and as a platform for productivity-led growth in the middle, while enabling market power only at the top. In doing so, innovation contributed to both opportunity creation and inequality in entrepreneurial outcomes.

The historical perspective carries clear relevance for the present. Many of the patterns we uncover such as the scalability of innovation among large firms, the heightened fragility of small establishments, and the muted returns for under-represented groups mirror current debates on intangible capital, superstar firms, and inclusive growth. By providing a long-run view of how intellectual property shaped the dynamics of entry, survival, and scale, our study contributes not only to economic history, but also to contemporary discussions on the conditions under which innovation fosters both productivity and broad-based economic progress.

References

Aghion, Philippe, Nick Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt, "Competition and innovation: An inverted-U relationship," *The quarterly journal of economics*, 2005, *120* (2), 701–728.

Akcigit, Ufuk and Sina T Ates, "What happened to US business dynamism?," *Journal of Political Economy*, 2023, *131* (8), 2059–2124.

_ **and William R Kerr**, "Growth through heterogeneous innovations," *Journal of Political Economy*, 2018, *126* (4), 1374–1443.

_ **, Douglas Hanley, and Nicolas Serrano-Velarde**, "Back to basics: Basic research spillovers, innovation policy, and growth," *The Review of Economic Studies*, 2021, *88* (1), 1–43.

_ **, Salomé Baslandze, and Francesca Lotti**, "Connecting to power: political connections, innovation, and firm dynamics," *Econometrica*, 2023, *91* (2), 529–564.

Andrews, Michael J., "How do institutions of higher education affect local invention? Evidence from the establishment of U.S. colleges," *American Economic Journal: Economic Policy*, May 2023, *15* (2), 1–41.

_ **, Enrico Berkes, and Ruveyda Gozen**, "Quantifying Patenting by Women in the U.S., 1845-1924," *Working Paper*, 2024.

Atack, Jeremy, "Firm size and industrial structure in the United States during the nineteenth century," *The Journal of Economic History*, 1986, *46* (2), 463–475.

Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen, "The fall of the labor share and the rise of superstar firms," *The Quarterly journal of economics*, 2020, *135* (2), 645–709.

Balasubramanian, Natarajan and Jagadeesh Sivadasan, "What happens when firms patent? New evidence from US economic census data," *The Review of Economics and Statistics*, 2011, *93* (1), 126–146.

Berger, Thor and Vinzent Ostermeyer, "Firm survival and the rise of the factory," *The Economic History Review*, 2025, *78* (1), 62–86.

Berkes, Enrico, "Comprehensive universe of U.S. patents CUSP v2.0," 2018. Data hosted and provided by author.

Gozen, Ruveyda Nur, "Property Rights and Innovation Dynamism: The Role of Women Inventors," 2024. CEP Discussion Papers, London School of Economics.

Gozen, Ruveyda, Richard Hornbeck, Anders Humlum, and Martin Rotemberg, "Historical Differences in Female-Owned Manufacturing Establishments: The United States, 1850-1880," Technical Report, National Bureau of Economic Research 2024.

Graham, Stuart JH, Cheryl Grim, Tariqul Islam, Alan C Marco, and Javier Miranda, "Business dynamics of innovating firms: Linking US patents with administrative data on workers and firms," *Journal of Economics & Management Strategy*, 2018, *27* (3), 372–402.

Hornbeck, Richard and Enrico Moretti, "Who benefits from productivity growth? Direct and indirect effects of local TFP growth on wages, rents, and inequality," Technical Report, National Bureau of Economic Research 2018.

_ **and Martin Rotemberg**, "Railroads, reallocation, and the rise of American manufacturing," Technical Report, National Bureau of Economic Research 2019.

_ **, Shanon Hsuan-Ming Hsu, Anders Humlum, and Martin Rotemberg**, "Technological Stickiness: Switching and Entry in the Long Transition from Water to Steam Power," Technical Report, working paper, University of Chicago 2023.

_ **,** _ **,** _ **, and** _ , "Gaining steam: Incumbent lock-in and entrant leapfrogging," Technical Report, National Bureau of Economic Research 2024.

Kelly, Bryan, Dimitris Papanikolaou, Amit Seru, and Matt Taddy, "Measuring technological innovation over the long run," *American Economic Review: Insights*, September 2021, *3* (3), 303–320.

Levinsohn, James and Amil Petrin, "Estimating production functions using inputs to control for unobservables," *The review of economic studies*, 2003, *70* (2), 317–341.

Loecker, Jan De and Frederic Warzynski, "Markups and firm-level export status," *American economic review*, 2012, *102* (6), 2437–2471.

Philippon, Thomas, *The great reversal: How America gave up on free markets*, Harvard University Press, 2019.

A Patent and CMF Data Matching

- Establishment names from the Census of Manufactures (CMF) were linked to patent assignee names from the Berkes patent database using the Jaro–Winkler string similarity measure.
- Prior to matching, standardised name cleaning was applied to both datasets. This involved converting all names to lowercase and removing punctuation, numbers, and extraneous whitespace to ensure comparability of name strings.
- Common name abbreviations and variants were also standardised. For instance, initials and shortened first names (e.g., 'Wm.' for William or 'Chas.' for Charles) were expanded to their canonical forms, thereby improving the consistency of name matching across both datasets.
- After cleaning, both datasets were partitioned into company-sounding names and people-sounding names. Names containing common firm identifiers (e.g., *company*, *manufacturing*, *brothers*, *limited*) were classified as company-sounding; all remaining names were classified as people-sounding.
- The matching process was implemented in four distinct runs:

1. CMF company-sounding names matched to Berkes company-sounding names;
2. CMF people-sounding names matched to Berkes people-sounding names;
3. CMF company-sounding names matched to Berkes people-sounding names (cross-match);
4. CMF people-sounding names matched to Berkes company-sounding names (cross-match).

- Matching was conducted within geographic blocks, defined primarily by county-level FIPS codes. When county information was unavailable or insufficient, state-level blocks were used.
- Within each geographic block, an iterative matching procedure was applied using progressively lower Jaro–Winkler similarity thresholds. Matches were first accepted at similarity scores greater than or equal to 0.96; remaining unmatched names were subsequently matched at thresholds of 0.94, 0.92, 0.90, 0.88 and 0.86.
- This iterative process ensured that high-confidence matches were prioritised, while allowing lower-similarity matches to be considered only among names that had not previously been matched. These lower-threshold matches were manually checked to reduce false positives in the final dataset for downstream analysis.

B Additional Tables

Table B.1: Sample Composition by Year and Industry

Panel A. Establishments by Year		
Year	Number of Observations	Percentage of Observations
1850	124,873	27.8
1860	121,176	27.0
1870	203,368	45.3
Total	449,417	100.0
Panel B. Establishments by Broad Industry		
Industry	Number of Observations	Percentage of Observations
Barrels, Furniture, and Other Wood Products	30,517	6.8
Beverages and Tobacco	15,655	3.5
Chemicals	13,025	2.9
Clothing and Textiles	76,716	17.2
Construction	39,322	8.8
Food Products	55,230	12.4
Iron and Steel Products	55,153	12.3
Leather Products	28,599	6.4
Lumber	55,492	12.4
Machinery and Carriages	32,629	7.3
Not Mfg	17,554	3.9
Other Consumer Products	7,529	1.7
Other Metal Products	12,605	2.8
Paper and Printing	6,933	1.6
Total	446,959	100.0

Notes: Panel A reports establishment counts by year; Panel B reports counts by broad industry category.

Table B.2: Sample Composition by State

State	Number of Observations	Percentage of Observations
Alabama	4,853	1.1
Arkansas	2,106	0.5
California	9,698	2.2
Colorado	459	0.1
Connecticut	12,702	2.8
District of Columbia	1,870	0.4
Delaware	1,770	0.4
Florida	1,121	0.2
Iowa	9,151	2.0
Idaho	357	0.1
Illinois	13,414	3.0
Indiana	21,902	4.9
Kansas	1,719	0.4
Kentucky	12,585	2.8
Massachusetts	22,431	5.0
Maryland	7,083	1.6
Maine	13,249	2.9
Michigan	10,226	2.3
Minnesota	2,910	0.6
Missouri	17,438	3.9
Mississippi	3,844	0.9
Montana	914	0.2
North Carolina	9,879	2.2
Nebraska	851	0.2
New Hampshire	9,162	2.0
New Jersey	15,505	3.5
Nevada	464	0.1
New York	76,369	17.0
Ohio	32,256	7.2
Oregon	1,516	0.3
Pennsylvania	75,725	16.8
Rhode Island	4,014	0.9
South Carolina	4,792	1.1
Tennessee	5,574	1.2
Texas	3,397	0.8
Utah	680	0.2
Virginia	16,139	3.6
Vermont	7,054	1.6
Washington	361	0.1
Wisconsin	11,375	2.5
West Virginia	2,493	0.6
Total	449,408	100.0

Notes: This table shows the number and percentage of observations for each state included in the Census of Manufacturers 1850, 1860, and 1870.

Table B.3: Summary Statistics

Variable	Mean	STD	P25	Median	P75	Min	Max	Observations
Log Revenue	7.964	1.770	6.909	7.824	8.923	0	18.7267	449,417
Log Output	5.290	4.720	0	5.525	9.190	0	24.41215	449,417
Log Material	5.615	4.134	0	6.535	8.700	0	24.63529	449,417
Log Employment	1.442	0.937	0.693	1.099	1.792	0	9.933532	449,417
Log Capital	7.142	1.670	5.994	6.909	8.102	0	21.89575	439,538
Log Wages	6.992	1.590	5.994	6.909	8.003	0	17.67915	177,999
Number of Patents	0.107	1.812	0	0	0	0	271	449,417
Number of Breakthrough Patents	0.075	1.780	0	0	0	0	271	449,417

Notes: This table reports summary statistics for the key variables used in the analysis.

Table B.4: Summary Statistics by Establishment Size

	Mean	SD	P25	Median	P75	Min	Max	Count
Panel A: Employment 1-10								
Log Output	7.65	1.54	6.91	7.60	8.52	0.00	16.12	394,020
Log Capital	6.81	1.40	5.99	6.91	7.70	0.00	15.38	384,696
Log Employment	1.17	0.53	0.69	1.10	1.39	0.00	2.40	394,020
Log Output per Worker	6.49	1.40	5.91	6.44	7.13	-2.40	14.77	394,020
Log Wages	3.96	0.80	3.37	3.93	4.50	0.00	11.51	203,700
Total Patents	0.08	1.14	0.00	0.00	0.00	0.00	188.00	394,020
Breakthrough Patents	0.05	1.10	0.00	0.00	0.00	0.00	188.00	394,020
Panel B: Employment 11-50								
Log Output	9.86	1.52	9.29	9.90	10.60	0.00	16.93	43,732
Log Capital	9.11	1.36	8.29	9.21	10.04	0.00	21.90	43,311
Log Employment	3.06	0.42	2.71	3.04	3.43	2.48	3.93	43,732
Log Output per Worker	6.80	1.47	6.34	6.89	7.46	-3.93	14.10	43,732
Log Wages	5.83	0.96	5.44	5.98	6.40	0.00	10.40	25,370
Total Patents	0.19	2.35	0.00	0.00	0.00	0.00	236.00	43,732
Breakthrough Patents	0.14	2.31	0.00	0.00	0.00	0.00	236.00	43,732
Panel C: Employment 51-100								
Log Output	11.05	1.52	10.60	11.23	11.74	0.00	16.75	6,772
Log Capital	10.39	1.30	9.62	10.60	11.23	0.00	16.68	6,704
Log Employment	4.28	0.20	4.11	4.26	4.44	3.95	4.62	6,772
Log Output per Worker	6.77	1.51	6.33	6.91	7.45	-4.62	12.59	6,772
Log Wages	6.85	1.15	6.40	7.09	7.60	0.00	10.09	3,858
Total Patents	0.49	6.25	0.00	0.00	0.00	0.00	271.00	6,772
Breakthrough Patents	0.41	6.22	0.00	0.00	0.00	0.00	271.00	6,772
Panel D: Employment 100+								
Log Output	11.79	2.07	11.41	12.06	12.70	0.00	18.73	4,891
Log Capital	11.33	1.59	10.60	11.51	12.21	0.00	18.31	4,825
Log Employment	5.32	0.61	4.88	5.17	5.61	4.62	9.93	4,891
Log Output per Worker	6.47	2.04	6.21	6.82	7.33	-7.35	12.73	4,891
Log Wages	7.59	1.32	7.09	7.82	8.41	2.48	12.46	2,683
Total Patents	1.12	9.62	0.00	0.00	0.00	0.00	236.00	4,891
Breakthrough Patents	1.00	9.59	0.00	0.00	0.00	0.00	236.00	4,891
Panel E: Total Sample								
Log Output	7.96	1.77	6.91	7.82	8.92	0.00	18.73	449,415
Log Capital	7.14	1.67	5.99	6.91	8.10	0.00	21.90	439,536
Log Employment	1.44	0.94	0.69	1.10	1.79	0.00	9.93	449,415
Log Output per Worker	6.52	1.42	5.94	6.50	7.19	-7.35	14.77	449,415
Log Wages	4.25	1.13	3.43	4.01	4.80	0.00	12.46	235,611
Total Patents	0.11	1.81	0.00	0.00	0.00	0.00	271.00	449,415
Breakthrough Patents	0.08	1.78	0.00	0.00	0.00	0.00	271.00	449,415

Notes: This table reports summary statistics for key establishment-level variables by employment size group. For each variable, we present the mean, standard deviation (SD), and selected percentiles (25th, 50th, 75th), along with minimum and maximum values. The final column reports the number of observations per group. Panel E summarizes the full sample across all establishments.

Table B.5: Pairwise Correlation Matrix

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1)	Log Revenue	1.0000							
(2)	Log Output	0.2768*	1.0000						
(3)	Log Material	0.3048*	0.4697*	1.0000					
(4)	Log Employment	0.6037*	0.2293*	0.2265*	1.0000				
(5)	Log Capital	0.6280*	0.3317*	0.3440*	0.6438*	1.0000			
(6)	Log Wages	0.6354*	0.1704*	0.1818*	0.7917*	0.6809*	1.0000		
(7)	Number of Patents	0.0392*	-0.0025*	0.0030*	0.0535*	0.0424*	0.0608*	1.0000	
(8)	Number of Breakthrough Patents	0.0336*	-0.0013	0.0030*	0.0464*	0.0373*	0.0523*	0.9815*	1.0000

Note: This table presents the pairwise correlations among the main variables. * $p < 0.10$.

Table B.6: All Businesses with and without Innovation (1850-1870) - Detailed Industry Fixed Effect

	(1) log(employment)	(2) log(avg male wages)	(3) log(capital)	(4) log(output)	(5) log(output per worker)
Has patented (=1)	0.0988***	0.0831***	0.155***	0.145***	0.0460***
	(0.0307)	(0.0264)	(0.0496)	(0.0435)	(0.0153)
Year FE	Yes	Yes	Yes	Yes	Yes
Detailed Industry FE	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
R^2	0.311	0.371	0.378	0.224	0.181
N	446906	234637	437705	446906	446906

Notes: This table shows the performance of all establishments with patents compared to the establishments without patents with detailed industry fixed effects instead of broad industry categorization. All columns include year, and county fixed effects. Standard errors in parentheses are clustered at the county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.7: Women-Owned Businesses with and without Innovations (1850-1870) - Detailed Industry Fixed Effect

	(1) log(employment)	(2) log(avg male wages)	(3) log(capital)	(4) log(output)	(5) log(output per worker)
Has patented (=1)	0.031	-0.243	0.391	0.113	0.081
	(0.107)	(0.178)	(0.249)	(0.204)	(0.185)
Year FE	Yes	Yes	Yes	Yes	Yes
Detailed Industry FE	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
R^2	0.357	0.534	0.427	0.374	0.374
N	2659	754	2659	2659	2714

Notes: This table shows the performance of women-owned establishments with patents compared to the women-owned establishments without patents with detailed industry fixed effects instead of broad industry categorization. All columns include a year fixed effects, and county fixed effects. Standard errors in parentheses are clustered at the county level.

Table B.8: All Businesses with and without Innovation (1850-1870) - Including Control Variables

	(1) log(employment)	(2) log(avg male wages)	(3) log(capital)	(4) log(output)	(5) log(output per worker)
Has patented (=1)	0.0475***	0.0293***	0.0795***	0.0306***	0.0155
	(0.0158)	(0.00857)	(0.0268)	(0.0116)	(0.0101)
Control Variables	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
R^2	0.576	0.786	0.582	0.532	0.232
N	437705	231350	437705	437705	437705

Notes: This table shows the performance of all establishments with patents compared to the establishments without patents by including control variables in the baseline specification. Control variables for Column (1) are the logarithm of capital and material, for Column (2) are the logarithm of capital, material, and employment, for Column (3) are the logarithm of employment and material, for Column (4) are the logarithm of capital, material, and employment, for Column (5) are the logarithm of capital, and material. All columns include year, and county fixed effects. Standard errors in parentheses are clustered at the county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

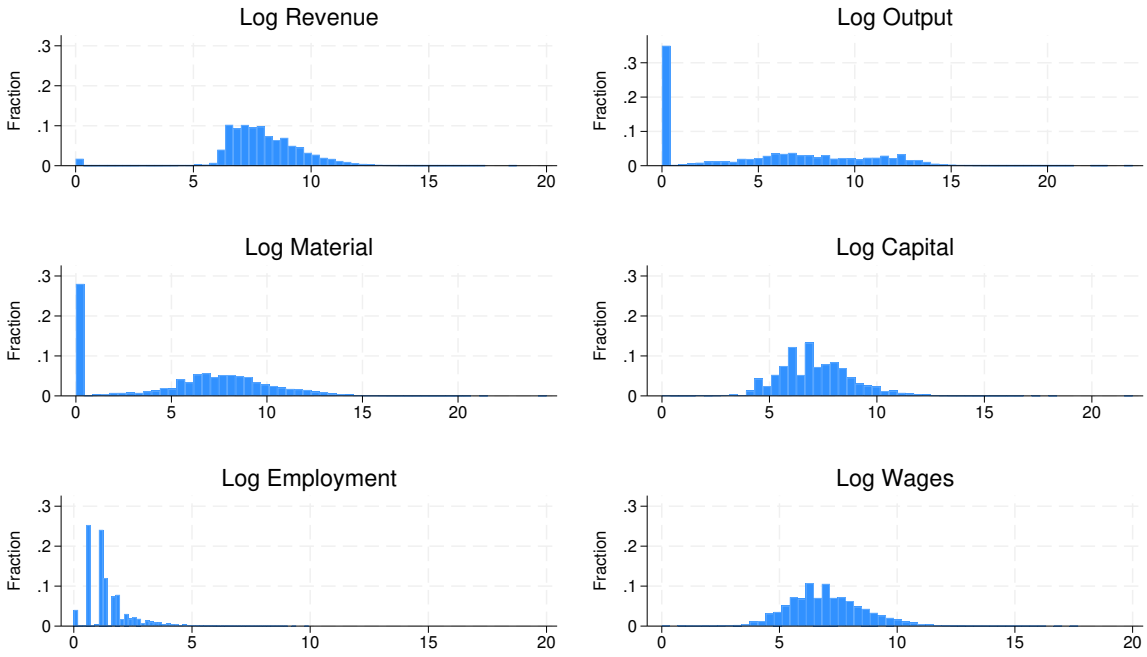
Table B.9: All Businesses with and without Innovation (1850-1870) - Including Control Variables

	(1) log(employment)	(2) log(avg male wages)	(3) log(capital)	(4) log(output)	(5) log(output per worker)
Has impact (=1)	0.0845***	0.0343***	0.0760*	0.0380*	0.0110
	(0.0255)	(0.0117)	(0.0431)	(0.0221)	(0.0194)
Control Variables	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
R^2	0.576	0.786	0.582	0.532	0.232
N	437705	231350	437705	437705	437705

Notes: This table shows the performance of all establishments with breakthrough patents compared to the establishments without breakthrough patents by including control variables in the baseline specification. Control variables for Column (1) are the logarithm of capital and material, for Column (2) are the logarithm of capital, material, and employment, for Column (3) are the logarithm of employment and material, for Column (4) are the logarithm of capital, material, and employment, for Column (5) are the logarithm of capital, and material. All columns include year, and county fixed effects. Standard errors in parentheses are clustered at the county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

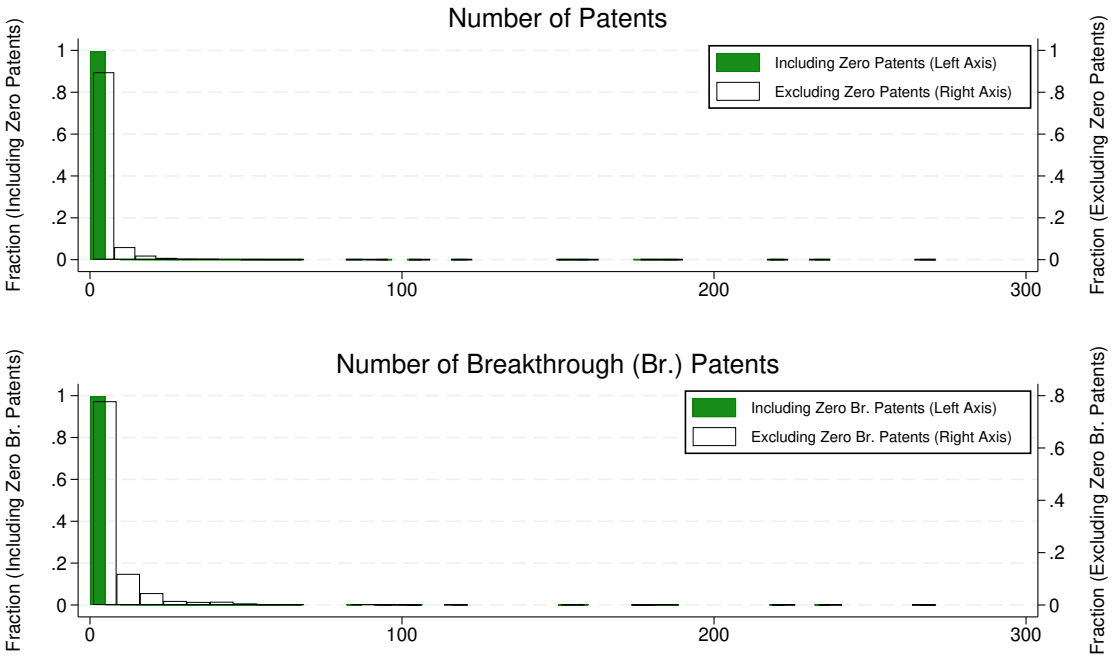
C Additional Figures

Figure C.1: Histogram of Establishment Variables



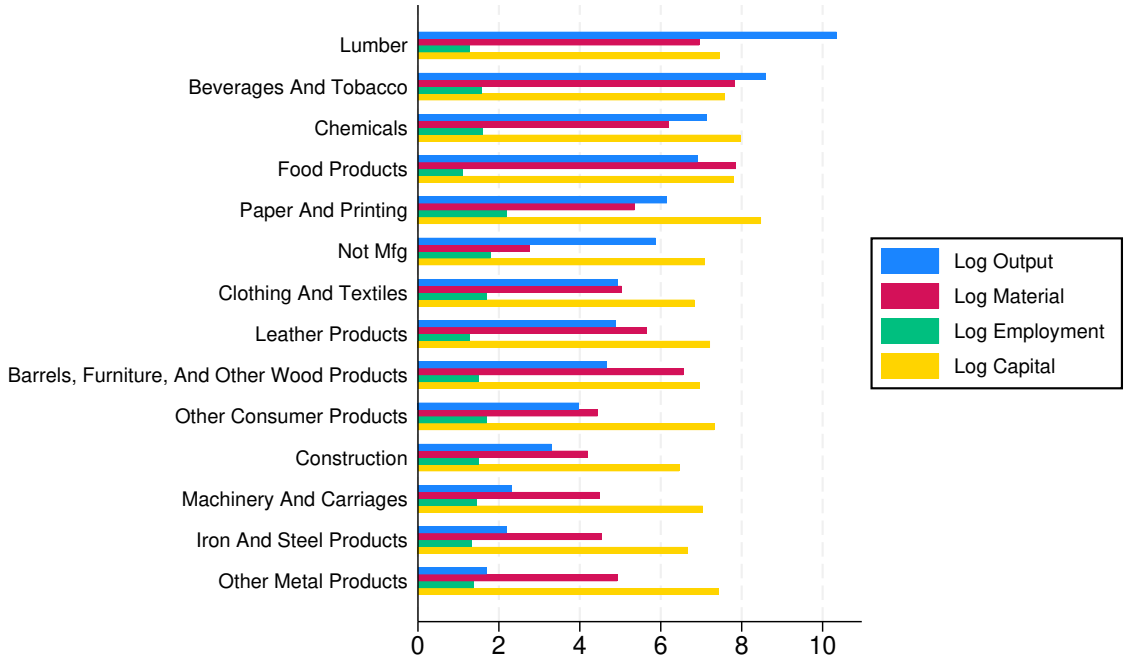
Notes: This figure presents the histogram of the main establishment-level variables.

Figure C.2: Histogram of Patent Variables



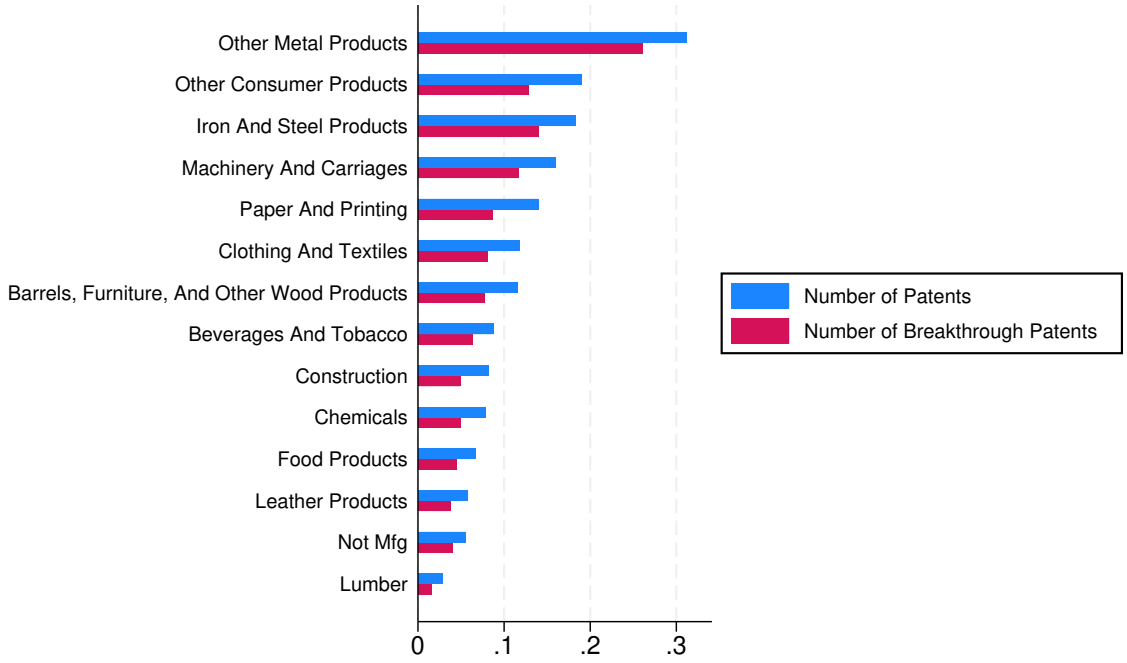
Notes: This figure shows the distribution of firms based on the number of patents and breakthrough patents.

Figure C.3: Average Values of Establishment Variables by Industry



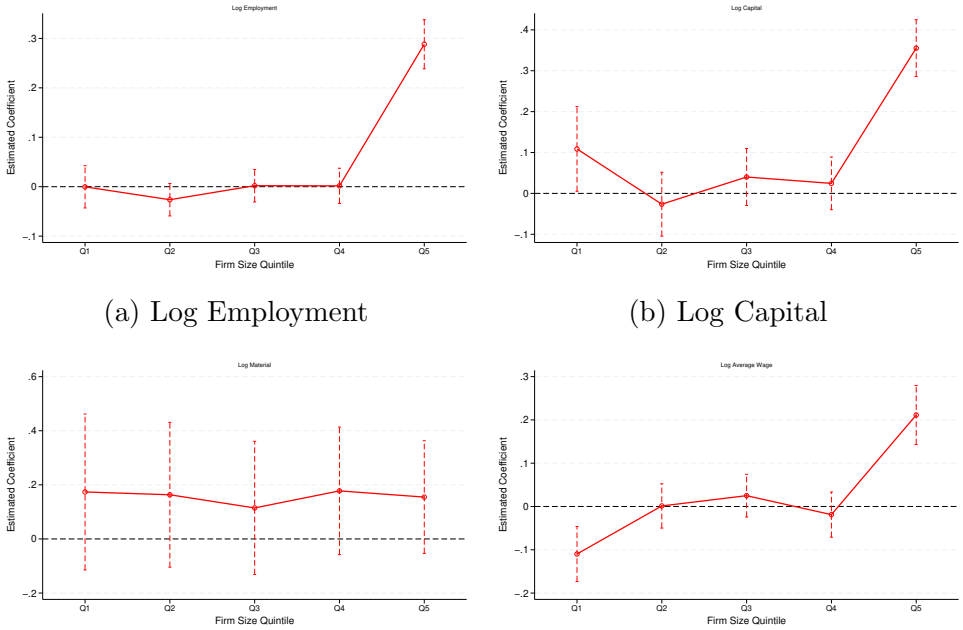
Notes: This figure shows the average values of the main establishment-level variables by broad industry category.

Figure C.4: Average Values of Patent Variables by Industry



Notes: This figure shows the average values of the patent variables by broad industry category.

Figure C.5: Quintile Regressions - Establishment Variables



(a) Log Employment

(b) Log Capital

(c) Log Material

(d) Log Average Wage

Notes: This figure plots estimated coefficients from regressions of establishment-level variables on breakthrough patenting status by firm size (total output) quintile within each census year and industry. Each coefficient measures the relative performance premium of breakthrough patenting establishments compared to non-patenting peers within the same quintile. All regressions include year, industry and county fixed effects. Standard errors are clustered at the county level. Confidence intervals are shown at the 95% level.