

NBER WORKING PAPER SERIES

TECHNOLOGY ADOPTION AND PRODUCTIVITY GROWTH:  
EVIDENCE FROM INDUSTRIALIZATION IN FRANCE

Réka Juhász  
Mara P. Squicciarini  
Nico Voigtländer

Working Paper 27503  
<http://www.nber.org/papers/w27503>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
July 2020

We thank Sandy Black, Bruno Caprettini, Diego Comin, Tarek Hassan, Michael Huberman, Felix Iglhaut, Markus Lampe, Nils Lehr, Bob Margo, Joel Mokyr, Tomasso Porzio, Diego Puga, Eric Verhoogen and numerous seminar audiences at Bocconi, Boston University, Cologne, Columbia, Leuven, Lund, Montreal University, Namur University, Penn State, Yale, UCL, UCLA, and participants at the 2020 NBER Summer Institute Meetings, the 2019 SED meetings, the 1st Workshop of Economic History in Bologna, “The Transnational History of French Industrialisation Before 1914” conference, and the IGC-Stanford conference on Firms, Development and Trade 2019 for helpful comments and suggestions. We thank Gaia Dossi, Johanna Gautier, Angelo Krüger, Lorine Labrue, Simon Margolin, Marie Michaille, Alberto Mola, and Sebastian Ottinger for excellent assistance throughout the construction of the dataset, Guillaume Daudin for kindly sharing data, and Eugene White for sharing his expertise of hyperinflation during the French Revolution. Financial support from the National Science Foundation (Grant # 1628866) is gratefully acknowledged. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2020 by Réka Juhász, Mara P. Squicciarini, and Nico Voigtländer. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Technology Adoption and Productivity Growth: Evidence from Industrialization in France  
Réka Juhász, Mara P. Squicciarini, and Nico Voigtländer  
NBER Working Paper No. 27503  
July 2020, Revised in October 2020  
JEL No. F63,N23,O14

### **ABSTRACT**

We construct a novel plant-level dataset to examine the process of technology adoption during a period of rapid technological change: The diffusion of mechanized cotton spinning during the Industrial Revolution in France. We document new stylized facts that can help explain why major technological breakthroughs tend to be adopted slowly and – even after being adopted – take time to be reflected in aggregate productivity statistics. Before mechanization, cotton spinning was performed in households, while production in plants only emerged with the new technology around 1800. This allows us to isolate the plant productivity distribution of new technology adopters in mechanized cotton spinning. We find that this distribution was initially highly dispersed. Over the subsequent decades, mechanized spinning experienced dramatic productivity growth that was almost entirely driven by a disappearance of plants in the lower tail. In contrast, innovations in other sectors (with gradual technological progress) shifted the whole productivity distribution. We document rich historical and empirical evidence suggesting that the pattern in cotton spinning was driven by the need to re-organize production under the new technology. This process of ‘trial and error’ led to widely dispersed initial productivity draws, low initial average productivity, and – in the subsequent decades – to high productivity growth as new entrants adopted improved methods of operating the mechanized technology.

Réka Juhász  
Department of Economics  
Columbia University  
420 West 118th Street  
New York, NY 10027  
and NBER  
rj2446@columbia.edu

Nico Voigtländer  
UCLA Anderson School of Management  
110 Westwood Plaza  
C513 Entrepreneurs Hall  
Los Angeles, CA 90095  
and NBER  
nico.v@anderson.ucla.edu

Mara P. Squicciarini  
Bocconi University  
Department of Economics  
Via Roentgen, 1  
Milan 20136  
Italy  
mara.squicciarini@unibocconi.it

A data appendix is available at <http://www.nber.org/data-appendix/w27503>

[T]here were strong pairwise complementary relations between factory organization and machinery [...] employers needed to simultaneously determine the choice of technique, the level of worker effort, and the way incentives were set up and communications and decisions flowed through the firm hierarchy. [...] Factories were the repositories of useful knowledge ... but they were also the places in which experimentation took place. – Mokyr (2010, pp. 345-46)

## 1 Introduction

The diffusion of innovation is at the core of aggregate productivity growth in the long run. Despite its importance for economic development, understanding the determinants and effects of technology adoption has proven difficult. As a consequence, the literature faces a number of open questions. For example, many technologies that ended up being widely adopted – such as hybrid corn seed in agriculture (Griliches, 1957) or by-product coke ovens in the iron and steel industry (Mansfield, 1961) – were slow to diffuse across firms.<sup>1</sup> This slow adoption is particularly puzzling given that new technology can provide a substantial boost to firm productivity (Syverson, 2011; Bloom, Eifert, Mahajan, McKenzie, and Roberts, 2013; Giorcelli, 2019). There is also a second, well-documented puzzle: When major innovations such as information technology (IT) or electricity spread across firms, the widely expected boost in aggregate productivity proved hard to document in the data. This prompted Robert Solow to remark in 1987 that “[...] what everyone feels to have been a technological revolution, a drastic change in our productive lives, has been accompanied everywhere, including Japan, by a slowing-down of productivity growth, not by a step up. You can see the computer age everywhere but in the productivity statistics.”<sup>2</sup>

A natural lens to study aggregate effects of technology adoption is the firm (or plant) productivity distribution – an approach that has gained prominence over the last two decades (c.f. Melitz, 2003; Hsieh and Klenow, 2009; Combes, Duranton, Gobillon, Puga, and Roux, 2012; Perla and Tonetti, 2014). However, applying this framework to technology adoption is difficult for numerous reasons: The specific technology used by firms is rarely directly observed, and even if it is known, new and old technologies typically coexist within narrowly defined sectors, or even within plants. In addition, the productivity distributions under the old and new technologies are not independent – a plant’s productivity with the old technology can affect its propensity to adopt innovations. These factors render it difficult to isolate aggregate productivity growth among adopters of the new technology. One approach to tackle these challenges is the use of randomized control trials (RCTs), which provide clean identification of the effects of technology at the firm level (Bloom et al., 2013; Atkin, Chaudhry, Chaudry, Khandelwal, and Verhoogen, 2017; Hardy and McCasland,

---

<sup>1</sup>The more general observation that technology is often slow to diffuse is attributed to Rosenberg (1976). See Hall and Khan (2003) and Hall (2004) for surveys of the literature on technology diffusion. Comin and Hobijn (2010) document substantial lags in the adoption of new technologies, estimating that the variation in adoption lags across countries can account for at least one-quarter of per capita income disparities.

<sup>2</sup>New York Times, July 12, 1987, p. 36. Such productivity puzzles are not restricted to the introduction of computers: David (1990) documents similar trends following the diffusion of electricity earlier in the 20th century.

2016). However, the relatively small sample size and short time horizon do not allow for a systematic analysis of the firm productivity distribution.

This paper bypasses the typical limitations by exploiting a unique historical setting – the adoption of mechanized cotton spinning in France during the First Industrial Revolution. Mechanized cotton spinning had been invented in Britain and – if operated efficiently – promised huge productivity improvements. We hand-collect novel plant-level data from historical surveys covering mechanized cotton spinning and two comparison sectors – metallurgy and paper milling – at two points in time, around 1800 and in 1840. Importantly, in 1800, mechanized cotton spinning had only recently been adopted in France. Four decades later, the technology had reached maturity (Pollard, 1965). Thus, the time period that we study encompasses both the initial phase of adoption and the period when the new technology had widely spread.

Our empirical strategy relies on a number of features of the historical setting. First, before the emergence of the new technology, cotton spinning was performed in home production with hand-operated spinning wheels (see the illustration in Appendix Figure A.1). Plants – or let alone factories – did not exist in cotton spinning prior to its mechanization. The new technology – the famous spinning jenny and the subsequent spinning mule – required the organization of production in cotton mills (see Appendix Figure A.2). Thus, all cotton spinning facilities in our dataset operated mechanized spinning in a central location, that is, in a *plant*.<sup>3</sup> This allows us to isolate the productivity distribution specific to the new technology and to examine its evolution over time. Second, the mechanized spinning technology changed relatively little between 1800 and 1840. We can thus narrow down potential drivers of productivity differences across plants and over time: With the underlying technology being very similar, productivity differences likely resulted from the more or less efficient *operation* of the new technology.<sup>4</sup> This goes beyond existing work, where productivity changes due to technology adoption reflect both the productivity differential of the new technology itself and the efficiency with which this technology is operated. Third, we document that owners who set up mechanized cotton spinning plants typically had a background in banking and finance as opposed to handspinning, suggesting that productivity under the new and old technologies were not systematically related.

---

<sup>3</sup>Throughout the paper, we refer to mechanized cotton spinning production organized in a central location as a *plant*. This is in order to delineate the organization of mechanized production from earlier organizational forms. In particular, under the pre-industrial putting-out system across much of France, local merchant-manufacturers organized home production across a multitude of households. They typically owned the capital equipment (the spinning wheels), which they gave to spinners together with the raw material, then collecting the spun yarn and marketing it (Huberman, 1996). This system was akin to ‘firms’ employing many *external* workers who performed production in scattered locations, typically for a piece-rate as opposed to a wage. Our survey of cotton spinning in 1806 allows us to eliminate the few cases of home production that were recorded by the enumerators. We thus focus on producers that organized *mechanized* production in plants.

<sup>4</sup>We do not claim that there was no technological progress in mechanized cotton spinning whatsoever. The technology did experience incremental changes. We show, using a number of robustness checks across plants and over time, that these do not explain our results.

We document three main findings for mechanized cotton spinning plants: 1) we observe a highly dispersed productivity distribution in the initial period (1806) relative to 1840; 2) we estimate that the mechanized cotton spinning industry underwent a substantial (82%) increase in plant productivity between 1806 and 1840 *after* mechanization had already been adopted; 3) this aggregate productivity growth in mechanized cotton spinning had a strong lower-tail bias: It was largely driven by the disappearance of lower-tail plants.

We compare the evolution of the plant productivity distribution in cotton spinning with two comparison sectors – metallurgy and paper milling. Importantly, in these sectors, production was already organized in plants before the Industrial Revolution because of their reliance on water power and high-fixed-cost machinery (see the illustrations in Figures A.4 – A.7 in the appendix). Thus, operational and organizational knowledge had already diffused. Technological progress in these sectors was more gradual and took the form of integrating new capital vintages into existing production setups – as opposed to the radical shift from home-based to factory-based production in cotton spinning. We find that the whole productivity distribution shifted to the right in the comparison sectors.

To rationalize these findings, we build on a mechanism that has been emphasized by historians and back it with new empirical evidence: Adopting mechanized cotton spinning required plants to learn best-practice methods along multiple dimensions. This is consistent with a setting in which plants learn about the optimal use of multiple inputs or tasks that in turn exhibit complementarities in the production function. We show that these features initially (when plants have little experience in the optimal performance of tasks) lead to a fat lower tail in the plant productivity distribution. Over time, as plants learn about the efficient use of multiple inputs (tasks), the lower tail disappears. We provide rich historical evidence consistent with this mechanism. Mechanization required the move to factory-based production, which represented an important organizational challenge: How should the layout of the mill be designed? How should machines be powered and how should the work flow be organized? How should workers – not used to the hierarchy and discipline of factory production – be recruited and managed? These organizational innovations proceeded via a process of trial and error, and the diffusion of organizational knowledge took time. As Allen (2009, p.184) writes: “The cotton mill, in other words, had to be invented as well as the spinning machinery *per se*.”

We provide several pieces of evidence in line with the proposed mechanism. First, we show that the exit rate of plants in mechanized cotton spinning was substantially higher than in other sectors between 1800 and 1840. This is consistent with the idea that early adopters faced considerable uncertainty about best-practice methods in production. Coherently, we show that exiting plants were substantially less productive than those that survived. Second, using plant age, we show that cotton spinning plants that entered the market later had higher productivity. This finding

supports our argument that knowledge about the optimal organization of mechanized cotton spinning diffused slowly over time. Moreover, the productivity advantage of younger plants is only a robust feature in cotton spinning around 1800. Younger spinning plants were *not* more productive in 1840; and in metallurgy, young plants were as productive as older ones in both periods.<sup>5</sup> This is consistent with the idea that later entrants in mechanized cotton spinning could draw from a better pool of organizational knowledge, which evolved over time and reached maturity around 1840. This process was muted in metallurgy, where plant-based production methods had been established much earlier. Finally, we examine our data for patterns compatible with learning effects. We find evidence supporting the spatial diffusion of knowledge: Plants located closer to other high-productivity plants were themselves more productive, and this relationship is strong only for cotton spinning plants and only during the initial period of technology adoption. We show, using a rich set of controls and placebo tests, that these results are unlikely to be driven either by selection into productive locations or by omitted variables.

Overall, our findings allow us to shed light on the two puzzles of slow adoption of major innovations and slow aggregate productivity growth. In combination with rich historical evidence, the highly dispersed initial productivity distribution points to information disparities across cotton plants: The organization of factory production evolved via a process of trial and error such that only the most productive adopters, those that ‘stumbled’ onto the efficient organization of factory production, survived. By 1840, information about the efficient organization of cotton mills had dispersed, leading to a much narrower distribution of plant productivity. Uncertainty about the optimal organization of production with a new technology can thus explain why *aggregate* productivity gains for technologies are initially small. Early adopters need to experiment with the efficient organization of production, and most of them will inadvertently operate the new technology inefficiently. Consequently, the potential benefits of the new technology materialize relatively slowly for the average plant. Our findings thus lend support to the view in the literature that the effects of major new technologies, such as electricity and IT, appear slowly in industry productivity “due to the need to re-organize the operation of the entire manufacturing facility to make effective use of this innovation” (Hall and Khan, 2003, p.8).

At the same time, potential adopters of a new technology expect that the knowledge about its optimal operation will improve over time. This creates incentives to delay entry. Indeed, we observe that later entrants in cotton spinning had significantly higher productivity than early adopters. These findings can help to explain the slow diffusion of major innovations, especially if they require a re-organization of production.

Our paper is closely related to a large literature on technology adoption. The most developed strand of this literature explains the low rates of technology adoption in agriculture in developing

---

<sup>5</sup>Due to data limitations, we cannot conduct this exercise for paper milling.

countries.<sup>6</sup> Given its importance in driving long-run sustained growth, understanding technology adoption in manufacturing is arguably equally important. However, this strand of the literature is much smaller, perhaps due to the difficulty of observing and categorizing the use of diverse technologies in manufacturing. The majority of recent papers circumvent this issue by using RCTs to understand the determinants of technology adoption (c.f. Bloom et al., 2013; Hardy and McCasland, 2016).<sup>7</sup> In line with our results, some of the evidence from experimental settings points to the importance of organizational barriers (Atkin et al., 2017). We contribute to this literature by studying the full plant productivity distribution for one of the most famous innovations in history, both at the initial phase of technology adoption and once the industry had reached maturity. Interestingly, our results suggest that the complex nature of production processes in manufacturing relative to agriculture may be important. In cases where technology adoption requires a substantial re-organization of production, adoption may be slow and productivity effects may take time to materialize.

Second, theoretical and empirical research has pointed to the importance of firm dynamics for understanding aggregate productivity (Jovanovic, 1982; Melitz, 2003; Syverson, 2011). A key insight of this literature is that – as there are large differences in productivity even within narrowly defined sectors – examining the entire distribution is critical for understanding aggregate productivity (Hsieh and Klenow, 2009). We apply this insight to the arguably most important structural break in economic history: The First Industrial Revolution, which saw unprecedented growth in manufacturing productivity (Crafts, 1985; Crafts and Harley, 1992; Galor, 2011). So far, productivity growth during this period has been studied mostly at the country level, or – in some cases – at the aggregate sectoral level. To the best of our knowledge, our paper is the first to study the contribution of plant dynamics to manufacturing productivity improvements during the Industrial Revolution.<sup>8</sup> In addition, we examine the puzzle of the slow diffusion of major technological innovations through the lens of the firm-productivity literature: Our findings suggest that focusing on the entire distribution of plant productivity can help deepen our understanding of

---

<sup>6</sup>See for example Foster and Rosenzweig (1995), Munshi (2004), Bandiera and Rasul (2006), Conley and Udry (2010), Duflo, Kremer, and Robinson (2011), Suri (2011), Hanna, Sendhil, and Schwartzstein (2014), Ben Yishay and Mobarak (2018), Beaman, Magruder, and Robinson (2014), and Emerick, de Janvry, Sadoulet, and Dar. (2016).

<sup>7</sup>One notable exception is Giorcelli (2019), who exploits a historical natural experiment to identify the productivity effects of adopting modern management practices and modern machinery by Italian firms after WWII.

<sup>8</sup>Similar to our setup, but in the modern context, Collard-Wexler and De Loecker (2015) examine how a specific sector was affected by a major innovation: the introduction of the minimill in the steel industry starting in the 1960s. This led to large aggregate productivity increases via both technology upgrading of surviving producers and exit of old incumbents. In contrast, our setting isolates productivity gains *within* the new technology, by separating it from pre-industrial home spinning. In other words, upgrading of existing ‘old’ producers is not present in our data, so that exit, entry, and improvements all occur among operators of the new, mechanized cotton spinning technology. A related paper in the historical context is Braguinsky, Ohyama, Okazaki, and Syverson (2015), who study the Japanese cotton spinning industry in the late 19th century and early 20th century. Rather than technology adoption, this paper focuses on the effects of acquisitions on acquired plants. These see a rise in productivity due to lower inventories and higher capacity utilization – which is likely a result of improved demand management.



technology *adoption*. This is in line with Foster, Grim, Haltiwanger, and Wolf (2018), who provide empirical support for the argument by Gort and Klepper (1982) that periods of rapid innovation are associated with a surge in firm entry, followed by a period where experience with the new technology is accumulated, eventually leading to a shakeout where unsuccessful firms (or plants) exit.<sup>9</sup>

Third, our paper is related to a strand of the literature that studies the role of organizational changes during the Industrial Revolution. In a set of papers Sokoloff (1984, 1986) shows that in the first half of the 19th century in the United States, both mechanizing and non-mechanizing industries displayed i) marked increases in firm scale as measured by the number of employees, and ii) substantial increases in productivity. Sokoloff speculates that these results are consistent with two important sources of productivity growth during the early stages of industrialization: division of labor and modifications in routines. However, a crucial distinction between the US and French context is that the former did not have an important cotton handspinning sector prior to mechanization. Thus, mechanized cotton spinning firms not only had to introduce the new technology to the US, but they also had to create the necessary infrastructure to supply raw materials and to market their final products (Ware, 1931).<sup>10</sup> In the European context, the seminal work by Pollard (1965) documents important organizational innovations (in what we would call ‘management’ today) that took place in the first half of the 19th century in Britain.<sup>11</sup> We contribute to this literature in the following ways. First, we show that, similar to the US, plant size and industry productivity were growing in France across a variety of sectors. This suggests that these patterns are a general feature of the early process of industrialization. Second, our main interest lies in understanding productivity growth during the process of technology adoption. While our results are consistent with Sokoloff’s conjecture that changes in firm organization contributed to rapid productivity growth during this period, our findings highlight that these productivity gains took time to materialize. Third, our unique setting allows us to isolate the effect of technology adoption from other factors such as the development of input or output markets. Finally, we analyze

---

<sup>9</sup>The literature has typically measured innovation via firm-level R&D expenditures or patents. Since these may miss important dimensions of firm activity (such as learning), Foster et al. (2018) rely on broader indicators for the presence of innovation: firm entry and productivity growth. Due to the indirect nature of these indicators, Foster et al. (2018) are reluctant to conclusively tie their results to technological innovation. Our unique setting has the advantage that we can isolate all plants in the cotton industry that operated a major new technology and track their productivity distribution over time.

<sup>10</sup>This type of confounder is not present in the French setting, which had a long history of handspinning before the Industrial Revolution. Raw cotton and cotton yarn markets were well-developed across France by 1800. Thus, one cannot replicate our results in the US context, despite the availability of US firm level data across a broader set of industries. In addition, the first US manufacturing census with wide coverage is from 1820. At this point, the US mechanized sector was arguably more mature than the French one from 1806. In other words, the early first cotton industry survey in France allows a unique look at its infancy period.

<sup>11</sup>Subsequent research in this area is discussed in Mokyr (2010). Squicciarini and Voigtländer (2015) show that upper-tail human capital played an important role in the industrialization of France.



the overall plant productivity distribution and are thus able to shed new light on how productivity growth evolved.

The paper is structured as follows. The next section discusses the historical context. Section 3 describes the construction of our novel dataset from archival sources, while Section 4 presents and discusses our empirical results. Section 5 draws conclusions from our results for technology adoption during periods of major innovation.

## 2 Historical Background

Early nineteenth century France presents an ideal setting for our study of technology adoption. While England was the first country to industrialize, France was a close follower and structurally similar to England (Crafts, 1977; Voigtländer and Voth, 2006). The flagship inventions of the Industrial Revolution – most notably the spinning jenny – were developed in Britain. France adopted these widely during the first half of the 19th century and witnessed a similar acceleration in industrial output as Britain (Rostow, 1975).<sup>12</sup> By focusing on France, we thus study the effects of technology adoption in the context of an industrializing economy that was (at least initially) mostly adopting technology developed elsewhere.

Our empirical analysis builds on three pillars that emerge from the historical context. First, the new cotton spinning technology entailed a move from home-based to factory-based production. This allows us to isolate the productivity distribution of adopters of the new technology. Second, we compare the productivity distribution in cotton spinning to two other sectors: metallurgy and paper milling. Production in these sectors was already organized in plants prior to the Industrial Revolution and experienced mostly incremental technological change that did not entail a drastic reorganization of the production process. Using these comparison sectors helps us to distinguish the effect of adopting a major new technology in cotton spinning from other trends that may have affected productivity distributions more broadly. Third, we discuss the historical evidence for the mechanism that arguably explains our empirical findings: The move to large-scale factory-based production in cotton spinning required the development of best-practice methods for organizing production through a process of trial and error.

### 2.1 Mechanized Cotton Spinning

*Development of Mechanized Spinning in Britain.* Cotton textiles was the flagship industry of the First Industrial Revolution, contributing one-quarter of TFP growth in Britain during the period 1780-1860 (Crafts, 1985). Cotton spinning is the process by which raw cotton fiber is twisted into yarn. Traditionally, this task was performed mostly by women in their homes, using a simple spinning wheel (see Figure A.1 in the appendix). With this old technology, each spinner was able to spin only one thread of yarn. The industry was rurally organized and generally centered around a

---

<sup>12</sup>Appendix A.1 provides a more detailed discussion of the Industrial Revolution in France.

local merchant-manufacturer who would supply spinners with the raw cotton, collect their output, take care of the marketing, and often also owned the spinning wheels (Huberman, 1996).

The breakthrough “macro-invention” in spinning was forged in Britain in 1765, when James Hargreaves invented the spinning jenny. This made it possible to spin multiple threads simultaneously, as twist was imparted to the fibre not by using the worker’s hands, but rather by using spindles. In 1779, Samuel Crompton further improved the design, inventing the water-powered spinning mule (see the left panel in Figure A.2). The productivity effect of these innovations was enormous. Allen (2009) estimates that the first vintage of the spinning jenny alone led to a three-fold improvement in labor productivity. Correspondingly, the price of yarn declined rapidly in the late 18th century (Appendix Figure A.3), especially for the highest-quality yarn, where prices declined from 1,091 pence per pound to 76 pence per pound in real terms between 1785 and 1800 (Harley, 1998).<sup>13</sup>

The mechanization of cotton spinning required production to be organized within factories for two reasons. First, the dependence on high fixed-cost inanimate power sources (such as water and steam) led to the concentration of production in one location.<sup>14</sup> Second, mechanized production increased the need for monitoring workers, who were now paid factory wages rather than for their home-produced output (Williamson, 1980; Szostak, 1989).<sup>15</sup> As production moved to factories (see the illustration in the right panel of Figure A.2), cotton spinners faced a set of new challenges in organizing production effectively. We return to this issue in the next section.

*Adoption of Mechanized Spinning in France.* Mechanized spinning was adopted with some lag in France. Efforts to adopt the technology had begun with state support during the *Ancien Régime*. By the beginning of our sample period in 1806, the large-scale expansion of the industry documented in Juhász (2018) had just begun: The technology was known throughout the country (Horn, 2006), and a number of domestic spinning machine makers had been established (Chassagne, 1991). The spinning machinery itself was produced mainly in France (using British blueprints) because of a

---

<sup>13</sup>Harley (1998) collected price data for three different qualities of yarn from British sources: 18, 40 and 100 count yarn (the count is an industry-wide standard that refers to the length per unit mass, implying that higher counts are finer). While all counts saw striking price declines (see Figure A.3), this trend was most pronounced for the finest, highest-quality varieties. Machine spinning had the largest impact on the fine high-quality yarn, which British hand-spinners had not been able to effectively produce and to which the mule-jenny (a subsequent vintage of the machine introduced in the late 18th century) was well-suited (Riello, 2013). Our data include information on the type of yarn produced, allowing us to account for quality differences across plants.

<sup>14</sup>The initial spinning jenny was hand-powered, but newer vintages of machinery that used inanimate sources of power (the throstle and the mule) were rapidly developed.

<sup>15</sup>Huberman (1996, p. 11) describes the need for monitoring in mechanized cotton spinning: “If there were multiple breakages of yarn on the larger machines, the mule had to come to a complete stop to piece the broken threads. There was also doffing, when the reels were full of spun cotton, the mule had to be stopped and the reels removed. Finally, there was cleaning. At all times, the spinners could expend effort as they were motivated to, and without proper supervision or incentives they could disguise how hard they could in fact work.” This created a strong need for monitoring, so that even early hand-powered machines (a particular vintage of spinning machinery) were housed in the “garrets of cottages and later in sheds” (Huberman, 1996, p. 11).

ban on exporting machinery from Britain until 1843 (Saxonhouse and Wright, 2004).

A number of features in the historical setting allow us to study how technology adoption affected productivity growth. First, mechanized spinning technology operated in centralized locations (plants), while the old technology relied on home production. This allows us to identify all producers that used the new technology, i.e., all *plants* observed in our data.<sup>16</sup> Consequently, we are able to isolate the productivity distribution for plants that used the new, mechanized technology.

Second, the sharp break in organizational requirements under the new technology rendered experience with the old technology effectively useless: The type of skills necessary for successfully running mechanized cotton spinning factories were very different from those required under the old cottage-industry technology. Correspondingly, historical evidence suggests that most owners who operated the new technology in France around 1800 did not have a background in the old technology. Table 1 presents information on the socio-economic background of the owners of mechanized cotton spinning establishments for the early period of technology adoption (1785-1815). Based on these figures, the vast majority of owners were “traders, bankers and commercial employees” (62.5%), while only a small fraction (10.2%) came from the production side, i.e. “workers and mechanics,” and three-quarters of this latter group were in fact highly skilled mechanics Chassagne (1991, p. 274).<sup>17</sup>

Third, during our sample period, the vintages of capital in cotton spinning and the technology remained largely unchanged.<sup>18</sup> This renders it unlikely that our analysis is confounded by the introduction of novel technology *after* the onset of mechanized spinning.<sup>19</sup> We can confirm this in our data, because the plant survey in 1806 asked which vintage of capital plants used. We find that both the throstle and the mule-jenny were widely used by existing cotton mills in France. The next vintage to appear (the self-acting mule – a fully automated machine) did not spread until the 1840s, i.e., after our sample period ended (Huberman, 1996). Moreover, in contrast to Britain, mechanized cotton spinners in France did not switch from water to steam power to a large extent,

---

<sup>16</sup>We follow Mokyr (2010, p. 339) in defining factory-based production as “the precise circumscription of work in time and space, and its physical separation from homes.” This definition is solely based on the *organization* of production; it does not rely on the use of inanimate sources of power. This is important because it allows us to refer to ‘factory’ or ‘plant’ production even when our data do not include specific information on power sources. Section 3.1 describes how we clean the data of a small number of observations that mix features of home and factory production.

<sup>17</sup>It is possible that some merchant-manufacturers previously involved with handspinning are reported under the category “traders, bankers and commercial employees.” However, Chassagne (1991, p. 274) clarifies that “Two-thirds of entrepreneurs came from a general trade background, predominantly in fabric and cloth, which proves the predominance of commercial factors in the launching and success of industrial enterprises.” This highlights the importance of marketing skills (as opposed to previous experience with handspinning) in setting up cotton spinning factories.

<sup>18</sup>There were three main types of machinery in use: the spinning jenny, the throstle and the mule-jenny. These were not good substitutes; in particular, the mule-jenny was well-suited to spinning higher-count, finer yarns (Riello, 2013).

<sup>19</sup>This is not to say that there were no improvements of the existing technology, or that plants did not increase their capital intensity. In robustness checks, we show that our main results are unlikely to be driven by these types of incremental improvements to the technology.

owing to the fact that France was not particularly well-endowed with coal (Cameron, 1985).<sup>20</sup>

Fourth, the historical literature summarized in Juhász (2018) underlines that the French needed to figure out many aspects of operating the technology efficiently themselves. That was partly because of the ban on machine exports and emigration of engineers and skilled workers from Britain throughout our sample period (Saxonhouse and Wright, 2004), and partly because a lot of the learning that needed to be done was tacit (Mokyr, 2001, 2010).<sup>21</sup>

## 2.2 Comparison Sectors: Metallurgy and Paper Milling

We compare the evolution of the plant productivity distribution in mechanized cotton spinning to two other sectors: metallurgy and paper milling. The key feature of both comparison sectors is that they had already organized production in plants since well before the Industrial Revolution. In metallurgy, plant production was mostly due to the reliance on high fixed-cost machinery such as the furnaces used both in smelting and refining. In paper milling, production was organized in plants because of the beating engine's reliance on water-power.<sup>22</sup>

Technological change in the two comparison sectors was introduced gradually and within the existing organization of production. The switch from charcoal to coal in metallurgy could be introduced by modifying a plant's existing machines and ovens (see the illustration in Figures A.4 and A.5 in the appendix). In paper milling, the mechanization of forming paper (one step in the production process) did not substantially alter the layout of the factory or other parts of the production process (see Figures A.6 and A.7). Moreover, the adoption of new technologies in both sectors was slow in France. Adoption began in the 1820s, and it was fairly modest until 1840. In metallurgy, the switch to coal was delayed because – in contrast to Britain – the relatively low price of charcoal kept the old (charcoal-based) technique profitable until the new (coal-based) technology's efficiency had risen (Allen, 2009). The new production processes using coal – coke smelting and the puddling process in refining – were both introduced around the 1820s in France and gradually adopted thereafter. By 1827, there were only four French departments where iron was smelted with coke. The adoption of the puddling process was more widespread, with 149 puddling furnaces in use across 15 out of 86 departments (Pounds and Parker, 1957).<sup>23</sup>

---

<sup>20</sup>This is confirmed in our data for 1840, showing that the majority of cotton spinning plants were still using water power (see Table D.3 in the appendix).

<sup>21</sup>Some British involvement in technology transfer took place *despite* the legal bans and the wars that occurred during this time period (Horn, 2006; Chassagne, 1991). However, while foreign technology transfer likely played *some* role in French learning, machines had to be produced domestically, workers were local, and some fundamental factors were different (such as the relative scarcity of coal).

<sup>22</sup>The beating engine breaks down the raw input vegetable matter into cellulose fiber. The production process is described in more detail in Appendix A.3.

<sup>23</sup>We can also verify indirectly that the new coal-based technology was introduced slowly in metallurgy: Adopting coal required a switch from water to steam as power source (Pounds and Parker, 1957), and the latter is reported by plants in 1840. Only 16% of plants in metallurgy report using steam-power, suggesting that the vast majority of them still used the older charcoal-based technology (see Table D.3 in the appendix).

In paper milling, the adoption of mechanized paper formation was similarly slow. While evidence on the adoption of mechanized paper-making is relatively scarce, our plant-level data from 1840 suggest that the technology had not been widely adopted. Only 10 (out of 348) plants explicitly mention having mechanized their production process, and only 42 plants report using steam power, which was necessary for the mechanized machine.

Given these features, the plant productivity distributions in metallurgy and paper milling reflect the same mix of old and new technology vintages that is typically observed in standard data sources. Observing two industries that were already organized in plants prior to the Industrial Revolution allows us to disentangle the productivity dynamics that were specific to mechanized cotton spinning from broader trends that affected the productivity distributions of all industries simultaneously.

### **2.3 The Challenging Transition to Factory-Based Production in Cotton Spinning**

The transition to large-scale factory-based production has been characterized as “one of the most dramatic sea changes in economic history” (Mokyr, 2010, p. 339). While cotton spinning was not the first sector to organize production in factories, the organizational changes during its industrialization went far beyond the experience made by other sectors. By the time cotton spinning mills reached maturity around 1830 in Britain (Pollard, 1965), they were larger, with a finer division of labor and a greater concentration of capital than production units in other sectors that had been organized as plants earlier (Chapman, 1974). However, the biggest change was the development of *flow production* – the production of standardized goods in huge quantities at low unit costs by “arranging machines and equipment in line sequence to process goods continuously through a sequence of specialized operations” (Chapman, 1974, p. 470). It is this synchronisation of highly specialized machines that distinguishes the cotton spinning factory as a “fully-evolved factory” from earlier developments (Chapman, 1974, p. 471).

Developing efficient cotton spinning mills meant solving organizational challenges along multiple dimensions. As Allen (2009, p.184) writes:

“The spinning jenny, water frame and mule were key inventions in the mechanization of cotton spinning, but they were only part of the story. [...] the machines had to be spatially organized, the flows of materials coordinated, and the generation and distribution of power sorted out. A corresponding division of labour was needed. The cotton mill, in other words, had to be invented as well as the spinning machinery *per se*.”

Successful mill designs were observed and copied. Chapman (1970, p. 239) shows that early mills in England had a remarkably similar structure because plants quite literally copied the original design of the Arkwright mills over and over again. Moreover, since few millwrights were qualified to build the power units of their mills, these were typically built by the same handful of engineers, qualified from experience (Chapman, 1970, pp. 239-40). It took time for design defects to be

improved; for example, contemporaries were aware of ventilation problems in the Arkwright-style mills, but continued to use the same layout regardless (Fitton and Wadsworth, 1958, p. 98).

Beyond organizing the production process efficiently, the factory setting and the use of inanimate power itself produced a host of new, unanticipated challenges. For example, fire hazards were a particularly pressing issue in the case of cotton spinning because of the highly flammable cotton dust (Langenbach, 2013). A process of trial and error eventually led to best-practice mill design that reduced fire hazards: Cotton textile mills introduced the so-called “fire-proof building” in Britain in the late 18th century, which entailed leaving no timber surfaces exposed by using cast-iron columns instead of wood (Johnson and Skempton, 1955). However, it quickly became apparent that fireproof mills were not indeed fireproof, because “steel or wrought iron, when heated, will fail by buckling or bending very much sooner than the equivalent beam of post or wood” (Boston Mutual Fire, 1908, p. 3). US textile mills developed what became known as “slow-burning mills” in the 1820s, recognizing that fires could not be prevented but their effects could be minimized by better mill design. Partly, this entailed moving back to using wood: “Timber posts offer more resistance to fire than either wrought-iron, steel, or cast iron pillars, and in mill construction are preferable in many respects (Boston Mutual Fire, 1908, p. 3). Chassagne (1991, p. 340) posits that early 19th century French mills consisted of multiple buildings and covered vast spaces (as opposed to building vertically) partly to minimize the fire hazard. Similarly, building structures needed to withstand the stress they faced from the vibrations of machines (Chassagne, 1991, p. 435). Iron rods with plates held beams to the masonry walls to prevent the vibrations of machines from shaking the walls apart (Langenbach, 2013).

Besides the need to develop the optimal mill layout and building structure, plants faced a host of other challenges that stemmed from concentrating workers under one roof and implementing a division of labor. On the one hand, workers had to adapt to the discipline and economic hierarchy of factory work. Following instructions, showing up to work on time, or getting along with other employees was new to workers who largely had experience with a domestic system (Pollard, 1963). Huberman (1996) describes that monitoring worker effort was a huge problem in cotton spinning mills.<sup>24</sup> He estimates that it took two generations for efficient labor management practices to be developed. Once again, progress was made via trial and error. Firms in Britain initially experimented with dismissals, which led to disastrously high turnover rates, and later with replacing male with female spinners in the hope that the latter could be more easily disciplined. Finally, around 1830, the industry settled on efficiency wages to motivate unobservable effort.

---

<sup>24</sup>As machines were not yet standardized, managers (overseers) lacked the technical information necessary to monitor effort. “[T]he operative spinner was firmly of the opinion that no two mules could ever be made alike. As a consequence, he proceeded to tune and adjust each of his own particular pair of mules with little respect for the intentions of the maker or the principles of engineering. Before very long, no two mules ever were alike” (Catling, 1970, quoted in Huberman, 1996, p. 59).



Managing teamwork efficiently became an issue when continuous flow processes were developed, which happened first in cotton spinning (Mokyr, 2001, 2010). Manufacturers did not have the experience, training, or access to knowledge to effectively manage labor or the mills in general. As Mokyr (2010, p. 350) emphasizes:

“[...] ‘management’ was not a concept that was known or understood before the Industrial Revolution. Military and maritime organization, the royal court, and a few unusual set-ups aside, the need for organizations in charge of controlling and coordinating large numbers of workers and expensive equipment was rare anywhere before 1750. British managers fumbled and stumbled into solutions, some of which worked and some did not.”

In fact, Pollard (1965) argues that the lack of modern management techniques limited the size of firms. In his seminal book on the topic, he shows that large firm size (above 100-200 workers) was seen as undesirable:

“up to the end of the eighteenth century at least...management was a function of direct involvement by ownership, and if it had to be delegated..., the business was courting trouble. This was a powerful argument against the enlargement of firms beyond the point at which an intermediate stratum of managers became necessary. [...] In the centuries preceding the Industrial Revolution, firms engaged in production were unable to cope with size, essentially because they could not cope with the problems of management which it involves.” (Pollard, 1965, p. 23)

The fact that owners were directly involved in management in the late 18th century is also confirmed in our data. In the paper milling survey of 1794, we have information on the name of the owner and of the manager for 174 plants in the sample. In 135 of these plants (78%), the owner and the manager were the same person.

In summary, the first generation of mechanized cotton spinners faced many organizational challenges all at once, along multiple dimensions. Developing efficient factory-based production proceeded via a process of trial and error, and it took decades for best-practice methods to emerge. According to Pollard (1965), the process was more or less complete by about 1830 in Britain: “a cotton mill was so closely circumscribed by its standard machinery, and there was so much less scope for individual design, skill or new solutions to new problems, by 1830, at least, ... that little originality in internal layout was required from any but a handful of leaders” (Pollard, 1965, p. 90). The initial lack of knowledge about best practice along multiple dimensions is a crucial element in the discussion of our empirical findings.

### **3 Data Construction**

Our analysis is based on a novel plant-level dataset constructed from handwritten historical industrial surveys. The data have a panel-like structure covering three industries: mechanized cotton spinning, metallurgy, and paper milling. We observe plants in these sectors at two points in time: around 1800 and again in 1840. Below, we discuss the main features of the data and the vari-



ables used in our analysis. Appendix B provides a more detailed description of the data, including sources.

### 3.1 Industrial Surveys Around 1800

Our data from the turn of the 19th century are based on three industry-specific surveys that were conducted by the French government. The survey for paper milling was implemented in 1794 during the French Revolution; it contains data on 593 plants. The most important survey for our analysis – cotton textiles – was conducted by the Napoleonic regime in 1806, covering 389 plants.<sup>25</sup> Finally, the survey in metallurgy in 1811 covers 477 plants.<sup>26</sup> Each of the three surveys provides hundreds of pages of handwritten returns that are available in the French National Archives in Pierrefitte-sur-Seine. Figures B.1, B.2, and B.3 in the appendix show sample pages from the three surveys. Although these data have not been systematically used for quantitative analyses,<sup>27</sup> the quality of French record-keeping in this period is well-known. The period is referred to as the “Golden Age of French Regional Statistics” (Perrot, 1975). Grantham (1997, p. 356) observes that “the quality equals that of any estimate of economic activity for a century to come.” Though the surveys were conducted at different points in time, we refer to their date henceforth as 1800.

*Distinguishing Mechanized Cotton Spinning Taking Place in Plants.* The rich data collected in the cotton spinning survey of 1806 allow us to identify production units that used the new, mechanized technology, organized in a central location, i.e., in a *plant* – as distinct from producers using handspinning technology. This distinction can be made with relatively high confidence because the 1806 survey specifically deals with *mechanized* cotton spinners. Thus, with the exception of a handful of cases, handspinning was not typically enumerated. We can identify these cases because establishments were asked about the vintage of mechanized capital that they used. We drop observations where handspinning wheels were reported. In addition, the 1806 survey also asked for the location of the plant. This helps us to filter out a few merchant-manufacturers who

---

<sup>25</sup>In France, cotton spinning and weaving were generally not vertically integrated during this time period. Weaving, particularly in the early 19th century, was rurally organized. This implies less of an incentive to locate the workers in a common location, i.e. in a plant. Nevertheless, our dataset contains a few examples of vertically integrated spinning and weaving plants. We deal with these integrated plants in the following way. In the 1806 survey, enumerators were instructed to separately collect data for spinning and weaving activities (which is indicative of the lack of integration across these sectors in general). In the few cases where both took place under the same roof, we observe labor and output reported by activity and can thus estimate productivity separately for the spinning activities. In the 1840 survey (for which we only observe total labor and revenues), only 7% of plants that spun cotton yarn reported activities in both spinning and weaving. We follow the classification in Chanut, Heffer, Mairesse, and Postel-Vinay (2000) and use only plants that reported *exclusively* cotton spinning.

<sup>26</sup>Bougin and Bourgin (1920) compiled an enormously rich overview of the metallurgy sector in 1788 using data from a wide variety of archival sources, including some recall data that was asked of plants in the 1811 survey. Unfortunately, since about 80% of plants do not report employment in Bougin and Bourgin (1920) for 1788, we cannot use these data in our baseline analysis. However, we do use the data as a validation check on plant survival.

<sup>27</sup>The only exception that we are aware of is Juhász (2018), who uses the data from 1806 on the mechanized cotton spinning industry.

placed early vintages of the spinning jenny (i.e., a mechanized vintage) in workers' homes: In these cases, the survey does not report one location for production, but many.<sup>28</sup>

### 3.2 Industrial Census around 1840

The second period in our study is based on data from the first industrial census in France, conducted in 1839-47 and digitized by Chanut et al. (2000). For simplicity, we refer to these data as the '1840 census.' While this census covers all manufacturing establishments, we only use data for cotton spinning (528 plants), metallurgy (839 plants), and paper milling (348 plants).<sup>29</sup> Figure 1 shows the spatial distribution of plants around 1800 and in 1840 for the three industries.

The different surveys contain remarkably rich information, although the exact set of variables varies from survey to survey. Here, we discuss only the variables used across all sectors, as well as additional, sector-specific variables when they are used in the empirical analysis. Our main variable of interest is labor productivity measured at the plant level and defined as the log of revenues per worker. We use this measure because it can be constructed for all sectors and in both time periods.<sup>30</sup> Additionally, we are interested in estimating plant survival across the different industries over time. Given the importance of these variables for our analysis, we discuss the key assumptions made in constructing them for each sector.

### 3.3 Constructing Plant Productivity

We face two challenges in constructing consistent productivity measures across plants and time. First, while the surveys for the three sectors around 1800 report output quantity (and some information on product-specific prices and quality), the census in 1840 directly reports plant-specific revenues (but not output quantities). To render productivity measures comparable over time, we have to construct revenues for 1800. Second, worker categories are not consistently reported across all plants in 1800 in metallurgy and paper milling. We discuss how we deal with each of these issues below.

---

<sup>28</sup>Juhász (2018, Appendix pp. 21-22) contains a detailed description of the cleaning process for these data. Of the 626 entries, only 43 plants were dropped for these reasons. This is not because these types of production units were so insignificant, but rather because the survey was not designed to capture them, so they were typically *not* enumerated. Note that the number of observations in Juhász (2018) differs substantially from that used in this paper, because the former contains all plants active in the French Empire.

<sup>29</sup>One potential concern with the 1840 census is that plants with less than 10 workers may have been systematically under-reported (Chanut et al., 2000). This is mostly relevant for paper milling, where plant size is the lowest. In robustness checks, we show that our baseline results hold even when only using plants with at least 10 workers in both periods.

<sup>30</sup>Our revenue-based productivity measure reflects both product prices and quantities. It is thus potentially affected by changes in markups (Garcia-Marin and Voigtländer, 2019). However, this is unlikely to be quantitatively important because all three sectors in our analysis produced standardized, often intermediate products.

### ***Estimating Plant Revenues in 1800***

*In cotton spinning*, the 1806 survey reports the quantity of yarn spun (in kilograms) as well as the minimum and maximum count of yarn spun, where the count of yarn is the standardized measure of quality in the sector.<sup>31</sup> We construct plant-level revenue by multiplying the quantity of plant-level output by the price of the average quality of yarn produced by the plant. We use a schedule of prices for different counts of yarn reported by the French government.<sup>32</sup>

*In metallurgy*, the 1811 survey asked for the quantity of output produced (by product) as well as the price charged by the plant, by product type.<sup>33</sup> While product-specific output quantity is reported by all plants, the product-specific price is only reported by a subset of plants. We compute the average price for each product using the subset of plants where this information is available. We obtain plant revenues by multiplying product-specific plant output by the average price for each product and summing across products.

*In paper milling*, the 1794 survey reports the total quantity of paper products produced (in metric quintals), but it does not provide plant-specific output prices. To construct revenues, we multiply plants' output quantity with the average price of paper products produced in the corresponding department, as reported in the *Tableaux du Maximum* – an extraordinary data source compiled in 1794 during the French Revolution that provides detailed data on goods prices and trade links across French regions. We use the department-specific price in order to accurately capture the product mix produced by plants in this area (see Appendix B.4 for detail). In robustness checks, we use the country-wide sectoral price.

*Price deflators* Finally, to compare revenues in the earlier periods and in 1840, for all three sectors, we deflate revenue data using the producer price index (PPI) for the respective survey years from Mitchell (2003). Appendix B.5 provides detail on the underlying assumptions and approximations in constructing these deflators. We note in passing that potential errors in the deflators would affect our estimates for *average* growth rates in the three sectors between 1800 and 1840, but they would not change the growth pattern across the plant distribution (e.g., the lower-tail bias in cotton spinning).

### ***Constructing Consistent Labor Variables***

*In cotton spinning*, the data provide consistent information on the number of workers employed by the plant.

---

<sup>31</sup>We use the (unweighted) average of the minimum and maximum count of the yarn produced by the plant as a proxy for its average output quality. The maximum and minimum count is the only information that plants provided on the quality of yarn that they produced.

<sup>32</sup>Source: Document number AN F12/533 from the French National Archives. In robustness checks, we use a single sector-level price, which we define based on the average quality of yarn produced across all cotton spinning plants.

<sup>33</sup>The survey includes the following products: iron of first quality, iron of second quality, iron of third quality, steel using the cementation process, natural steel, and pig iron.

*In metallurgy*, about 40% of the plants reported both ‘internal’ and ‘external’ labor in the 1811 survey, while the remainder of plants reported only total labor. Woronoff (1984, p. 138) describes external labor as only having very loose ties to the plant. These workers did not typically work at the location of the plant, their work was not supervised by the manager, and their identity was often not even formally known to the manager. They performed tasks such as driving, collecting charcoal for the plant, or performing other jobs without belonging to the hierarchy or reporting to superiors in the chain of command. These types of workers were highly unlikely to be considered formal salaried employees of the plant in the 1840 census. The challenge is thus to construct a consistent measure of labor in 1811, given that approximately 60% of the observations report only total labor, with no indication of whether this includes external labor. For these plants, we need an estimate of the size of their internal labor force. We use a nearest neighbor matching algorithm to determine whether plants that only report total labor are more likely to be reporting internal labor only or the sum of internal and external labor.<sup>34</sup> When our algorithm suggests that the plant is reporting internal and external labor together, we estimate the number of internal workers by using the mean proportion of internal labor from all plants that report both types (the internal labor share is 20%).

*In paper milling*, the vast majority of plants only reported male labor in 1794. We impute the total number of employees in each plant by scaling male labor (reported by each plant) in 1794 by the average proportion of total employees to male employees in 1840 (where we observe both). The validity of this method hinges on the assumption that the proportion of male employees remained constant over time. We are able to check this using the subset of plants that report all types of workers in 1794. We find that the proportions are remarkably consistent.<sup>35</sup> Moreover, we show that our results are robust to using only male employees in both periods.

### 3.4 Linking Plants over Time

It is possible to link plants over time given that all surveys report the name of the owner and the location up to the commune, which is the lowest administrative unit in France.<sup>36</sup> We use two pieces of information to link plants over time: First, we match plants by their owner names in a given location in the respective industry.<sup>37</sup> Since the name of the owner may change even if the physical structure of the plant is the same, we also match by location in a second step: We match locations

---

<sup>34</sup>We match each plant that reports only total labor to its nearest neighbor that reports internal and external labor, where matching is based on capital, output, and the stage of production. We then classify a plant as “reporting only internal labor” (“reporting total labor”) if its reported total labor is closer to the matched plant’s internal labor force (internal plus external labor force).

<sup>35</sup>The proportion of total employees to male employees is 2.26 in 1794 for the subset of 20 plants that report all types of labor, while in 1840 it is 2.28 (among all plants).

<sup>36</sup>In bigger cities such as Paris, the *arrondissement* is also reported.

<sup>37</sup>We use a fuzzy string match to allow for differences in spelling as well as for different first names of owners, in cases where the plant was passed on within a family. Ambiguous matches were verified by hand.

where there is *only* one plant in the respective sector in 1800 and where there is at least one plant active in the same sector in 1840. This turns out to be fairly common in the data. An obvious concern is whether this ‘local matching’ indeed identifies the same plant. This is likely, given a fortuitous feature across all three of our industries: their reliance on water power. Only a small number of locations in a particular commune will be suitable for setting up a water-powered mill, as rapid stream flow is needed to yield sufficient power. Moreover, the backwater created by one mill means that another mill cannot be located in close proximity. Consistent with this, Crafts and Wolf (2014) argue that agglomeration in the cotton textile industry was not observed until steam became the common source of power in Britain. Consequently, our ‘local matching’ arguably identifies plants that have the same location within communes. Whether these were owned by the same entrepreneur (or their descendants), or whether they had passed on to a different owner is not crucial for our analysis.

One way to validate the assumptions underlying our ‘local matching’ is to examine how frequently communes with a single plant active in the sector in 1800 show up in 1840 with multiple plants active in the same sector. If this occurs frequently in the data, it would suggest that in fact there are multiple suitable locations for production in that sector for a particular commune. This is not the case in our data: For the vast majority of single-plant communes that we identify in the initial period, there either continues to be one plant or no plants in the subsequent survey. It is exceedingly rare (6% of cases in both paper milling and cotton spinning, and 8% in metallurgy) across all three surveys for single plant communes to ‘add’ additional plants (despite the large increase in the overall number of plants in metallurgy and cotton spinning). As an additional check on our methodology, it is also possible to compare plant survival in metallurgy to that reported in Woronoff (1984) for this sector over the period 1788-1811. If our strategy of ‘local matching’ led to too many plant matches over time, we would expect an exaggerated survival rate. The contrary is true: Our estimates of plant survival rates for the period 1811-1840 are well below (one half or less) those that we calculate for the period 1788-1811. This suggests that it is unlikely that we systematically overestimate plant survival.

### **3.5 Plant Survival Rates**

Our main measure of plant survival is based on the combination of matching by owner name and ‘local matching’ that we described above. We define the survival rate as the percentage of plants from the initial period that survive into the later period. The numerator counts all plants that fulfill at least one of the following two conditions: i) the plant has the same owner in both periods; ii) there is only one plant in the respective sector in the location in the initial period and *at least* one plant in the same sector in 1840. The denominator is the sum of *all* plants in the given sector in the initial period. Note that this rate does not adjust for the fact that the number of plants located

in communes that have only one plant varies across the three sectors in our sample.<sup>38</sup> Thus, we may mechanically find higher survival rates in a sector where single-plant communes are relatively more frequent. To address this issue, we also construct the ‘restricted sample’ survival rate as a robustness check. This measure is based solely on single-plant locations. The numerator counts the number of communes that had only one plant in the respective sector, in both the initial period and in 1840 (indicating plant survival). The denominator counts the number of communes that had a single plant in the respective sector in the initial period and either one or no plant in 1840 (indicating plant survival and plant exit, respectively).<sup>39</sup>

## 4 Empirical Analysis

In this section, we use our data to study the evolution of the plant size and productivity distribution in mechanized cotton spinning after the new technology had been adopted. We contrast the patterns in cotton spinning to those observed in two comparison sectors – metallurgy and paper milling. After documenting our main result, we propose and investigate a mechanism that can rationalize the observed patterns. Finally, we consider a set of alternative mechanisms that could account for the results and test them empirically.

### 4.1 The Evolution of Plant Scale

What did plants look like in 1800 across the three sectors, and to what extent did they undergo change during our study period? Table 2 provides an overview, reporting the evolution of plant size, measured by the number of workers. A few points stand out. First, as early as 1806, cotton spinning plants were strikingly large. The average spinning plant in this period had 63 employees (the median was 30).<sup>40</sup> Despite the recent introduction of mechanized cotton spinning in France, plants were already much larger than in the two comparison sectors, which had a much longer tradition of factory-based production: Plants in metallurgy (reported in 1811) had on average 20 workers; paper milling plants had on average 13 employees.<sup>41</sup> In sum, despite the late start of

---

<sup>38</sup>Among the 593 plants in paper milling in 1794, 218 (36.8%) were the only plants active in their commune in this sector. For cotton spinning in 1806, the proportion is 25.4% (99 out of the 389 plants), and in metallurgy in 1811, 69% (329 out of 477 plants).

<sup>39</sup>Based on this sample definition, we exclude plants that were the only ones in their commune in 1800, and where there was more than one plant in 1840. As discussed above, the number of these “uncertain” observations is very small across all sectors, which we consider a validation of our methodology.

<sup>40</sup>Recall that mechanization, which triggered the move to factory-based production, was invented only in the late 18th century. Moreover, the machines were only adopted sporadically across France prior to the 1800s. Consistent with these facts, the median plant in cotton spinning was three years old in 1806.

<sup>41</sup>One caveat with making this comparison is that the paper milling survey dates from 1794. Thus, plant size may have grown by 1806 – the year of the cotton spinning survey. In addition, we had to extrapolate the overall number of workers in paper milling in 1794 (including women and children – see Section 3.3). However, it is unlikely that true plant scale would have been very different in 1806. This is because even in 1840, the average plant size in paper milling was only 43 (including women and children, which are reported in this year). We can thus be confident that paper milling plants in 1806 were substantially smaller than cotton plants. Finally, there is a concern that the 1840 census did not enumerate all plants with less than 10 employees. Table D.13 shows that plant scale increased across



factory-based production in cotton spinning, the average plant was large compared to other sectors already at the turn of the 19th century.<sup>42</sup>

Table 2 also shows that all sectors underwent significant growth in plant size over the period 1800-1840. Average plant size doubled in cotton spinning and grew threefold in the other two sectors. In 1840, the average cotton spinning plant had 112 employees (median 72), while metallurgy and paper milling had 57 and 43 workers on average, respectively. The level and increase in plant scale is consistent with Sokoloff's (1984, Table 1, p. 5) findings for the United States using establishment level data from the 1820 and 1850 manufacturing censuses. The data for the US cover 10 industries. Similar to our data for France, the US data include both mechanizing sectors (such as cotton textiles) and non-mechanizing sectors such as tanning.<sup>43</sup> While their coverage is richer, the US data are less suitable for isolating productivity changes in mechanized cotton spinning over the early phase of industrialization. The reasons are twofold: First, there was no handspinning tradition before mechanization in the US, so that local input and output markets had to be established at the same time as plants (Ware, 1931). This is prone to lead to differences in local production costs and markups, thus affecting (revenue-based) productivity estimates (Garcia-Marin and Voigtländer, 2019). Second, the US coverage begins at a later point in time, in 1820, when mechanization was already well underway (Ware, 1931).

## 4.2 The Pattern of Productivity Growth

Next, we examine the evolution of plant productivity in France during the first decades of the 19th century. We begin by examining average annual labor productivity growth. Column 1 in Table 3 shows that all three sectors experienced a significant increase in labor productivity. The largest productivity gains were achieved in cotton spinning (2.4% per year), followed by metallurgy

---

all three sectors even when we limit the sample in both years to only include plants with more than 10 employees.

<sup>42</sup>Granted, our data only cover two comparison sectors. However, these two sectors had a long tradition of factory-based production and relied on high fixed-cost capital as well as inanimate power sources. This makes it likely that metallurgy and paper milling plants had relatively large plant size, as compared to plants in other sectors around 1800 for which we do not have data. We can also use the 1840 Census, where we know the size of plants in *all* other sectors to gauge support for this assertion: The average plant size in cotton spinning in 1840 is in the 85th percentile of all plants, while both metallurgy and paper plants are in the upper tercile of the plant size distribution across all sectors. In addition, all three sectors belonged to the top 90th percentile in terms of the share of plants using "any power," where the 81 sectors in the census are ranked by the share of plants using inanimate power sources.

<sup>43</sup>For comparison, the average size of cotton textile establishments was 34.6 in the US in 1820 and 97.5 in 1850. Similar to France, but based on a larger set of comparison sectors, cotton textiles was the biggest in terms of plant scale already in 1820. The one exception to this is glass-making, but Sokoloff (1984) urges those numbers to be treated with caution as they are based on a sample of only 8 establishments. "Iron and iron products" in the US had 19.5 and 24.2 employees in the two periods, respectively, while paper milling had an average of 14.3 and later, 22.4 employees. Since employment and productivity grew in both mechanizing and non-mechanizing sectors, Sokoloff (1986) concluded that mechanization itself cannot be the only driving force behind these trends. Our findings for the French sectors support this conclusion and add the additional finding that even *after* the initial adoption of mechanized spinning, productivity and plant size grew substantially.



(1.9%) and paper milling (0.7%).<sup>44</sup> Remarkably, the estimated productivity increase is largest in spinning, despite the fact that all plants in this sector already used the new technology in 1800. In other words, because we only compare plants that used mechanized cotton spinning, the observed productivity increase must be due to efficiency gains *within* the new technology. This is in contrast to the two comparison sectors, where innovations replaced older technology vintages in existing plants – most prominently, coal as an energy source in metallurgy, and the Foudrinier machine in paper milling (which mechanized the formation of paper). Thus, labor productivity growth in those sectors reflects not only improvements in operating existing vintages, but also gains from the adoption of new technology vintages.

In which part of the productivity distribution were these gains concentrated? Figure 2 plots the distribution of labor productivity in the three sectors at the beginning and at the end of our sample period, illustrating our main results. In cotton spinning, two features stand out. First, the initial dispersion in labor productivity was large in 1800 relative to that in 1840. Second, the productivity gains are almost exclusively concentrated in the lower tail – the lower tail disappeared over our sample time period, while increases in productivity at the upper tail were modest. The contrast between cotton spinning and our two comparison sectors is striking. In metallurgy and paper milling, the entire productivity distribution shifted to the right between 1800 and 1840. Quantile regressions confirm this pattern. Columns 2-6 in Table 3 report these results for the three sectors, estimating regressions for productivity growth at different quantiles of the productivity distribution. Figure 3 displays the corresponding coefficients. In cotton spinning, the bias towards productivity growth in the lower tail is striking. Productivity growth at the 25th percentile was twice as large as that at the 75th percentile (3.3% per year relative to 1.65%), and the difference is more than fourfold between the 10th and the 90th percentile (3.9% and 1.0%, respectively). In the comparison sectors, the differences are more modest across the distribution; if anything, growth was concentrated in the upper tail: In both metallurgy and paper milling, the productivity growth at the 25th percentile was marginally *lower* than at the 75th percentile.

Could the different pattern in cotton spinning be driven by output quality? Recall that our data enable us to use quality-adjusted prices in 1800 for cotton spinning, while quality adjustments are not possible in the two comparison sectors. Panel B in Table D.1 in the appendix presents quantile regressions without quality-adjustments, i.e., using the same sector-level price across all plants in cotton spinning. The magnitude of the lower-tail bias is slightly smaller, but it remains striking.

---

<sup>44</sup>Given that we discount revenues using price indices, all our productivity calculations reflect price-adjusted revenue-based productivity. To obtain the average annual growth rates between the two time periods (around 1800 and 1840), we first regress log output per worker  $\ln(Y/L)$  on a dummy for 1840 in each sector. This coefficient measures the percentage growth in output per worker over the entire time period between the respective survey years. We then annualize these values (and corresponding standard errors) by dividing by the number of years between the surveys in each sector. Note that this method delivers average annual growth figures, not accounting for compound growth. In cotton spinning, the overall growth over the period 1806-40 amounts to 82% (2.42% per year x 34 years).

The difference in productivity growth between the 10th and the 90th percentile is 3.4% relative to 1.6%. Thus, differential trends in output quality do not drive the observed differences in the productivity distributions across the three sectors.

Our baseline productivity measure is log output per worker. For cotton spinning, we can also compute TFP, using detailed data on physical machinery (number of spindles) – see Appendix D for detail. Panel C in Table D.1 confirms the lower-tail bias of productivity growth in cotton spinning using TFP. Table D.2 presents additional robustness checks on data choices we have made in the paper milling industry. We show that our quantile regression results in paper milling are robust to i) using only male labor and ii) to pricing output in 1794 using the country-wide average price as opposed to the departmental price used in the baseline analysis.

Summing up, the results in this section show that *after* the adoption of the new technology in mechanized cotton spinning, the industry witnessed major increases in productivity that were driven by a disappearance of the lower tail of the productivity distribution. Over time, the plant productivity distribution became less dispersed. This is in contrast to the patterns observed in the comparison sectors, where mean productivity growth was more modest and occurred relatively evenly across the productivity distribution.

### 4.3 Proposed Mechanism: Learning About Best Practice in Factory-Based Production

What explains the lower-tail bias of productivity growth in mechanized cotton spinning? The historical narrative points to an important role of learning about the efficient organization of factory-based production and adequate handling of new machinery. As documented in Section 2, early adopters in cotton spinning needed to engage in trial and error along multiple dimensions. This involved the development of best-practice methods for operating the new technology efficiently and integrating it into other, similarly new aspects of factory-based production.

*A Stylized Framework.* We formalize this process in a simple stylized framework that gives rise to lower-tail biased productivity growth. We summarize the key features here; Appendix C presents the model and shows simulation results. The essential ingredients are: i) a production function that involves multiple *complementary* inputs (tasks); ii) independent, random draws in the efficiency associated with each input; and iii) the lower bound of each input’s efficiency draw increases over time (i.e., very bad draws disappear).

In cotton spinning, initially, very low input efficiency draws are possible. The complementarity across inputs implies that a low efficiency draw for only one input diminishes output substantially.<sup>45</sup> This gives rise to a fat lower tail of the productivity distribution. Over time, we let the lower bound for the input efficiency draws in cotton spinning increase. This reflects improved knowledge about organizing and handling each input – either due to learning within plants or due

---

<sup>45</sup>In the extreme complementarity case – a Leontief production function – output drops in proportion to the minimum efficiency draw, no matter how large the draws for the other inputs are.

to knowledge diffusion across plants.<sup>46</sup> At the same time, we assume that the upper bound for input efficiency draws does not change, i.e., the technological frontier in mechanized cotton spinning remains the same. Intuitively, learning (within plants or from other producers) eliminates the worst mistakes in organizing production; but the best possible draws remain unchanged because the underlying technology does not change. This setup leads to productivity growth concentrated in the lower tail. Figure C.1 in the appendix illustrates the simulated productivity distributions.<sup>47</sup> The left panel of Figure C.1 shows the resulting evolution of the plant productivity distribution over time: The initially fat lower tail disappears in the second period – intuitively, very bad draws that pull down output towards zero (even if all other draws are high) have been eliminated. This productivity pattern mirrors the one for cotton spinning in the data (see Figure 2).

We also adapt our simulation to the comparison sectors. Since plant-based production in these sectors was already well-established around 1800, we choose a higher lower bound to reflect that learning had already occurred. To represent the incremental technological progress that improved these technologies over time, we shift both the upper and lower bound outward in the second period. The corresponding simulation results are shown in the right panel of Figure C.1. They resemble the pattern observed for our control sectors, metallurgy and paper milling, where the whole productivity distribution shifted to the right.

While this stylized theoretical framework is not the only one that gives rise to lower-tail bias in productivity growth, it is a simple setup that represents the key features of the historical evidence. In what follows, we use this stylized framework to guide our discussion of possible mechanisms.

*Plant Survival Across the Sectors.* The historical evidence discussed in Section 2 suggests that learning best-practice methods was an important dimension during the shift to factory-based production in cotton spinning. We now provide evidence that this mechanism can explain the observed plant productivity patterns. First, we examine plant survival across sectors. If early adopters of mechanized cotton spinning technology had to experiment with best-practice methods, we would expect initially low plant survival rates relative to the other two sectors. Experimentation was arguably costly. For example, the layout of the factory was to a large extent a sunk investment. Changes and extensions could of course be made, but at substantial cost. Thus, initial design flaws were hard to correct and likely led to exit. Indeed, we find evidence of substantially larger exit rates in cotton spinning relative to the other two sectors. Table 4 reports plant survival rates over our sample period, using the two measures defined in Section 3.5 in each of the three sectors. Based on our baseline measure, survival rates in spinning (7%) were slightly lower than in paper

---

<sup>46</sup>While our stylized theory is agnostic as to which of these mechanisms dominated, the empirical results below suggest that the diffusion of organizational practices *across* plants was an important dimension.

<sup>47</sup>We use a CES production function with three inputs and choose the elasticity of substitution across inputs to be 0.5, indicating a strong degree of complementarity. Efficiency in each of the inputs is drawn from a uniform distribution with support  $[0, 1]$ . Over time, the lower bound for each distribution increases to 0.1.

milling (9%) and much lower than in metallurgy (34%). Note that the paper milling survey was conducted in 1794, more than 10 years earlier than the cotton spinning survey (1806). If we adjusted for the longer horizon in paper milling, the implied survival rates for 1806-1840 would be significantly higher. This implies that the differences in survival rates between cotton spinning and paper milling are probably higher than reflected in Table 4.

Turning to the ‘restricted sample’ survival rates, the differences across the three sectors are even starker. By this measure, the survival rate in spinning is still about 7%, but it is much higher in the comparison sectors: 20% in paper milling and 49% in metallurgy. Recall that this second survival rate is based on single-plant locations. Thus, the low survival rate observed in cotton spinning means that many locations lost their (only) cotton mill. This is consistent with a mechanism in which owners that invested in a cotton spinning mill with poor layout had to exit the market, and the structure of the mill was not subsequently used by other firms in cotton spinning.<sup>48</sup>

*Plant Exit and Productivity.* In Table 5 we examine the extent to which plants that eventually exited the market by 1840 had lower initial productivity around 1800, as compared to surviving plants. This pattern is strong in cotton spinning, where churn was the highest: Exiting plants were 53% less productive than survivors, and this difference is highly statistically significant. Exiting plants were also less productive in the comparison sectors, but there, the pattern is less pronounced: exiting plants were about 15% less productive in both sectors, and this number is not statistically different from zero. Columns 2 and 3 in Table 5 also show that exiting plants were significantly smaller, both in terms of employees and output, and that this pattern is particularly pronounced in cotton spinning.

Overall, we find that exiting plants in cotton spinning were particularly unproductive compared to survivors, and – likewise – the survival rate was significantly lower than in other sectors. These findings – in combination with the widely dispersed productivity distribution in cotton spinning – are consistent with large organizational challenges and low initial guidance in switching to factory-based production in cotton spinning.

*Age Profile of Plant Productivity.* Next, we examine whether the age profile of plant productivity is consistent with a mechanism of learning best-practice methods. We expect younger plants to have higher productivity, because they had a larger set of previously established plants from whom they could learn. We exploit the richness of our data to test this in both 1800, when best practice mill design was still evolving, and in 1840, when according to Pollard (1965), the industry had reached maturity – at least in Britain. The 1806 survey in cotton spinning contains the year of

---

<sup>48</sup>Lower survival rates in spinning could also be consistent with the fact that the industry moved towards steam power to a larger extent than other sectors and hence moved away from water power. However, note in Table D.3 that even in spinning, steam power seemed to be used *in addition to* water power: 66% of cotton spinning plants still used water power as a source in 1840. Moreover, cotton mills using steam power were somewhat less productive (see Table D.4), suggesting that in France plants did not face a strong profit incentive to move away from water power.

foundation of plants. This allows us to compute a dummy for ‘young’ plants, defined as below-median age (with the median age in 1806 being three years). We first examine whether plant age is systematically correlated with productivity in cotton spinning. Column 1 in Table 6 shows that the unconditional association is strongly positive: ‘Young’ plants were 58% more productive in 1806. This could be driven by mechanisms other than the one we examine. For example, new entrants may have used the most recent vintage of capital, leading to higher physical productivity (Foster, Haltiwanger, and Syverson, 2008). To address this issue, we control for several important plant characteristics in columns 2-7 of Table 6. These include the average quality (count) of the yarn spun, the capital intensity of the plant (measured as log spindles per worker), the number of workers in the plant, and the vintage of machinery (a binary variable for using different vintages of machinery – these are not mutually exclusive categories)<sup>49</sup> The coefficient of interest remains large and highly significant when we add these controls one-by-one.<sup>50</sup> The only control that notably changes the magnitude of the coefficient of interest is the quality of yarn (column 2). However, the quality of the yarn that a plant could spin partly reflected its level of learning, and is thus arguably also capturing our mechanism. This is because the quality of yarn that could be spun by a plant depended on the quality of the machine and the precision of its operation (Huberman, 1996). This, in turn, may have reflected improvements of the machinery achieved by tinkering, and the acquired skill of the spinners, both of which are a central part of our learning mechanism.

Next, we turn to the productivity-age pattern in 1840. While the data for this second period are generally more comprehensive, we do not observe plant age. However, we can perform a similar – albeit weaker – test based on the comparison of surviving and entrant plants: In Table 7 we regress log output per worker on an indicator for whether the plant was an ‘entrant’ in 1840 (as opposed to a surviving plant by our definition from Section 3.5). The coefficient on the ‘entrant’ dummy thus reflects the average productivity differential for plants that entered between the initial survey year (1806) and 1840. Best-practice mill design evolved over this period, and it had largely converged by 1840 (Pollard, 1965). Correspondingly, we find that ‘young’ plants were not more productive; the coefficient is in fact negative, but not statistically different from zero. This holds also when we control for the use of water power, steam power, any other power source (wind or animal power used by a small subset of plants), and for the number of workers. We perform an additional exercise in the appendix, using data on metallurgy plants in both periods (Tables D.5 and D.6). For this sector, in both periods, the best measure of plant age that we observe is a binary indicator of plant survival from 1788 to 1811, and for survival from 1811 to 1840 (the latter being the same procedure as for cotton spinning in Table 7). Overall, the results do not point to younger plants in metallurgy having a strong productivity advantage: In 1811, ‘young’ metallurgy plants

---

<sup>49</sup>The three different vintages of machinery are the spinning jenny, the throstle, and the mule-jenny.

<sup>50</sup>We do not include all controls together because of multicollinearity concerns.

were only marginally more productive, and in 1840 ‘young’ metallurgy plants were somewhat *less* productive.<sup>51</sup>

*Spatial Diffusion of Knowledge.* In what follows we shed light on *how* the learning process took place. The historical background discussed in Section 2 suggests that plants copied successful designs and setups of the production process from each other. To examine this channel, we estimate whether a plant’s own productivity was higher in the proximity of other high-productivity plants. We use the following specification:

$$\ln(Y/L)_{ij} = \beta_0 + \beta_1 Dist_{ij}^{p90} + FE_j + \epsilon_{ij} ,$$

where  $\ln(Y/L)_{ij}$  is labor productivity (log output per worker) for plant  $i$  located in department  $j$ ;  $Dist_{ij}^{p90}$  is log distance to the nearest plant (in the same sector) with productivity in the 90th percentile (in the distribution of *all* plants in the sector across France). Plants that are themselves in the top productivity decile are excluded from the sample to avoid introducing a mechanical relationship. Our preferred specification includes department fixed effects ( $FE_j$ ) to absorb unobserved location characteristics that may make all plants in a given region more productive, irrespective of local spillovers. Thus, the coefficient of interest  $\beta_1$  reflects the extent to which plants in the same department benefit from being located closer to a high-productivity plant (which may be located in the same or in another department). We do not interpret these correlations as causal effects, but as evidence that is compatible with spatial spillovers of production knowledge. We estimate the specification separately for the three sectors, and in both time periods. Standard errors are clustered at the department level to account for spatial correlation.

Before presenting the results, we first examine the spatial distribution of high-productivity plants across our sectors and time periods. Figure D.1 in the appendix plots the spatial distribution of cotton spinning, metallurgy, and paper milling plants, distinguishing those in the 90th percentile of the productivity distribution. Unsurprisingly, some regions have a larger concentration of high-productivity plants than others, but there are no marked differences in this pattern across the three sectors or over time. Since we exploit within-department variation in distance to high-productivity plants, the cross-regional differences in the concentration of high-productivity plants do not affect our results.

Figure 4 visualizes our baseline results on spatial diffusion, and Table D.7 in the appendix reports the corresponding regressions. To allow for direct comparability, we report the standardized beta coefficients of  $Dist^{p90}$  for all three sectors in the two periods. The estimated coefficient for

---

<sup>51</sup>In metallurgy in 1811, ‘young’ plants were 22% more productive when no controls are added (as compared to almost 60% for the same specification in cotton spinning). Once we control for plant size (log workers) in column 2 in Table D.5, the productivity advantage disappears. Note that the metallurgy survey has sparser information on this dimension; plant size is the only control that can be added in 1811.



cotton spinning in 1806 is negative, statistically significant and large in magnitude: A one standard deviation increase in distance to a high productivity plant in the sector is associated with a 0.81 standard deviation decline in labor productivity. The pattern is much weaker in the two control sectors, metallurgy in 1811 and paper milling in 1794 – the coefficients are less than one-third in magnitude as compared to cotton spinning. In addition, in 1840, there is essentially no relationship between labor productivity and distance to top-plants in any of the three sectors. Thus, proximity to high-productivity plants mattered the most in cotton spinning in 1806, i.e., in the sector that saw the most dramatic change during industrialization, and in the period before knowledge about the optimal organization of production had spread widely. While this pattern is consistent with a learning mechanism, it could also be driven by a number of other forces. We turn to examining possible alternative explanations below.

While department fixed effects capture unobserved differences that vary at the department level, they cannot capture unobserved differences at a finer spatial level. To test for these, we use four approaches. First, we control directly for some prominent location fundamentals at the commune level such as the availability of fast-flowing streams (as a source of water power), proximity to coal (which matters for steam power), and the share of forest cover (which matters for access to charcoal – a major input in metallurgy).<sup>52</sup> Table D.8 in the appendix shows the results for specifications that control for all of these location fundamentals. The pattern of the coefficients of interest does not change, and the estimated magnitudes remain very similar. Moreover, the location fundamentals themselves are mostly small and statistically insignificant. This is probably driven by the fact that the department fixed effects already account for the most important spatial differences in location fundamentals. Second, our results could be affected by more general agglomeration externalities, as opposed to learning. In particular, our findings may be driven by high-productivity plants emerging (within departments) where the density of production is large due to agglomeration forces. We investigate this explanation directly by adding a control for the density of production at the commune level (measured as the log of total output in the sector, excluding a plant's own output). Table D.9 in the appendix shows that controlling for the local density of production barely affects our results: The estimated coefficient on distance to high-productivity plants in cotton spinning in 1806 decreases by about 10% to -0.74 (se 0.19), while the distance coefficients in the other sectors and in 1840 remain relatively small. The coefficient on local production density itself is generally small, positive, and never statistically different from zero.

Appendix Table D.10 presents our third approach to probe for unobserved location fundamen-

---

<sup>52</sup>The sources for these data are as follows. Data on the stream flow of rivers are from EURO-FRIEND (<http://ne-friend.bafg.de/servlet/is/7397/>). These data report water flow rates from thousands of geocoded collection points across France. Data on charcoal sources are from the highly detailed 'Cassini maps' produced in the late 18th century that contain information on forest cover. These maps were geo-referenced by Vallauri, Grel, Granier, and Dupouey (2012). Finally, data on the location of coal deposits – both in France and near its borders – were geo-referenced from maps in Tarr and McMurry (1993) and Guiollard (1993).



tals (within departments). We conduct a placebo exercise that examines whether plant productivity in 1800 was also related to the distance to high-productivity plants that only emerged later, i.e., plants in the top-90th percentile of productivity in 1840. Reassuringly, the estimated coefficient in cotton spinning is close to zero and statistically insignificant. In other words, productivity in cotton spinning in 1806 was not related to high-productivity locations three decades later. This suggests that the large estimated coefficient in our baseline specification is not driven by persistent location fundamentals within departments.

Finally, we examine the extent to which the estimated distance coefficient in cotton spinning in 1806 may be driven by plant selection. It is possible that we estimate a large (negative) coefficient in cotton spinning in 1806 not because plants were learning from their high-productivity neighbors but rather because *ex-ante* high-productivity plants selected into ‘productive locations’ (i.e., chose to locate near existing high-productivity plants). Given that we observe plant age in cotton spinning in 1806, we can examine selection patterns, building on our findings from Table 6 that ‘young’ plants in cotton spinning were particularly productive. In Table 8, we show that our coefficient of interest is robust to controlling for log plant age (column 1) and to adding the interaction of plant age and distance to high productivity plants (column 2). This suggests that (conditional on department fixed effects) plant entry did not vary systematically with the location of high-productivity plants. A different, arguably more conservative approach is to estimate the coefficient of interest using only the subsample of plants that entered *before* the nearest high-productivity plant. The timing of entry of these plants rules out systematic selection. In column 3 of Table 8 we show that the coefficient on distance to high-productivity plants remains statistically highly significant, although it is somewhat smaller than in the baseline sample (-0.44, se 0.153). Columns 4 and 5 show that the coefficient of interest remains stable and statistically significant as controls for plant age and the interaction between plant age and distance to high-productivity plants are added in this subsample.

In summary, the consistently larger distance coefficient estimated in cotton spinning in 1806, in combination with a series of robustness checks, points to the spatial diffusion of knowledge as one mechanism through which learning across plants took place.

#### **4.4 Robustness to Alternative Explanations**

In the final part of this section we consider potential alternative mechanisms that could explain our results. Recall that we observe the lower-tail bias in productivity growth only in cotton spinning, and not in the other two sectors. This suggests that the mechanisms driving the effect are either specific to cotton spinning or they affect this sector differentially. Below, we consider some competing mechanisms that would fit this pattern.

*Market Integration.* Could increased market integration explain our results in cotton spinning? As the French economy became more integrated over time, it is possible that lower-productivity plants

faced tougher competition and had to exit the market.<sup>53</sup> However, we do not observe a lower-tail bias in productivity growth in the two comparison sectors. Thus, for market integration to explain our results, it must have affected the cotton sector differently.

We first note that the exit pattern in Table 5 speaks against such a mechanism: If market integration increased particularly strongly in cotton spinning, we would expect even relatively productive plants to exit the market, because of an increased productivity threshold that allows for profitable production (Melitz, 2003). However, the contrary is true: Exiting plants in cotton spinning had a particularly *low* productivity, as compared to the other two sectors.

Second, data on market integration also speak against a confounding role of this channel. Cotton yarn (and textiles more generally) are high value-to-weight products, which makes them more easily tradable over long distances than iron or paper. This suggests that cotton spinners may already have faced tougher competition through more integrated markets in the early 1800s. Thus, we would not necessarily expect cotton spinning to be the most affected sector by increased market integration after 1800. Consistent with this reasoning, we present evidence for relatively high market integration in cotton yarn in the late 18th century. We use data in 1794 from Daudin (2010) on the number of districts across France that reported consuming cotton textiles, iron, or paper products from any district in a given department.<sup>54</sup> Intuitively, higher market integration means lower price differentials across departments, which in turn implies that highly productive areas could dominate the market throughout France. Consequently, we can infer high market integration from the data if we observe that a few (presumably highly productive) departments sold to many other departments, while the majority of departments produced no output, or did so only for local consumption. Figure D.2 in the appendix shows that this pattern is particularly strong in cotton textiles. Many departments produced mostly for themselves if at all (these are the zeros and ones), while a few departments supplied cotton to a large number of districts. The top tercile of departments exported cotton textiles to 30 or more districts. In the two comparison sectors, there is less specialization and less evidence for market integration: Fewer departments report not supplying to anyone (particularly in paper), and the top decile of departments supplied only to 6 (paper) and 7 (iron) districts in total. This suggests that cotton textile plants were already competing in a bigger market than the comparison sectors around 1800.

The appendix presents two further robustness checks that probe the extent to which market integration may explain our results. We introduce two controls in our quantile productivity regressions (see the baseline results in Table 3). Table D.11 controls for market potential, and Table

---

<sup>53</sup>Market integration arguably increased during our sample period both for policy reasons such as the abolition of internal barriers to trade during the French Revolution (Daudin, 2010), and because of infrastructure improvements that reduced transport costs such as the introduction of railways in the late 1820s.

<sup>54</sup>Districts are administrative units that stayed in place for approximately five years, from 1790 to 1795 – with each department including from a minimum of 3 to a maximum of 10 districts.

D.12 introduces region fixed effects.<sup>55</sup> Controlling for market potential changes the coefficients of interest only marginally, while adding region fixed effects dampens the lower-tail bias in cotton spinning: Productivity growth at the 10th percentile is now only twice as large as in the 90th percentile (as compared to a factor of almost 4 in the baseline regressions in Table 3). This suggests that a part of the pattern is driven by reallocation across regions, but not in a way systematically related to market potential.

*Plant Size and Early Spinning Workshops.* Another potential concern is that our results may be driven by the disappearance of small cotton spinning plants (or ‘workshops’). While our baseline sample filters out small establishments that used *handspinning* wheels (see Section 3.1), it may include relatively small plants that operated early vintages of mechanized spinning jennies that did not necessarily need inanimate sources of power. This small-scale setup may have been inherently different from the larger-scale factories powered by inanimate power sources. While the move to factory-based production was swift, systematic differences of smaller mechanized cotton workshops could account for the lower-tail bias of productivity growth in this sector. Our data does not differentiate between these two types of plants (we observe the capital vintage, but not the power source in the 1800 data). However, we can examine the extent to which our results could be driven by these forces in a number of ways.

One way to examine the influence of smaller plants on our results is to adopt a stricter definition of ‘factory-production’ and use only those plants that have more than 10 employees. This should exclude the majority of the smaller workshops that may have been organized as factory-based production along some but not all dimensions. Figure D.3 and Table D.14 in the appendix show that the lower-tail bias in productivity growth is robust to using only larger plants. The magnitudes remain similar at most quantiles, while the lower-tail bias of productivity growth remains unique to mechanized cotton spinning. These findings also address the concern that smaller plants may have been under-sampled in 1840.

Next, we implement an even more conservative definition of factory-production. We drop any plant from the 1806 survey that reports using the earliest vintage of machinery – the spinning jenny. These were the types of machines that could, in principle, have been operated also in small workshops without inanimate power sources. While the introduction of spinning jennies – even in workshops – would have faced challenges in re-organizing production, dropping them from the survey allows us to assess the extent to which these may have driven the lower-tail bias of productivity growth. Table D.15 contains the quantile regression results. The lower-tail bias of productivity growth remains striking. Thus, it is unlikely that early spinning workshops that share

---

<sup>55</sup>Market potential is computed as the sum of inverse distance-weighted city population in 1800. Regions are larger than departments: There are 22 regions in France and 86 departments. Our data in the three sectors do not have enough observations to include *department* fixed effects, i.e., to estimate meaningful productivity distributions within departments.

some, but not all features of factory-based production drive the lower-tail bias of productivity growth.

Finally, our results could also be driven by increasing plant scale. Recall, however, that this is unlikely, as all sectors witnessed an increase in plant scale. Indeed, as the results in Table D.16 show, controlling for the number of workers in each plant does not alter our findings.

*Effects of the Napoleonic Blockade on the Cotton Spinning Sector.* Juhász (2018) shows that temporarily higher trade protection from British competition shifted the location of the mechanized cotton spinning industry within France. To what extent does varying trade protection explain the lower-tail bias of productivity growth in cotton spinning? This is unlikely, given that our results hold within regions, where the pattern of protection was very similar (see Table D.12). Figure D.4 presents further evidence that varying trade protection does not drive our results, by splitting the sample into plants in northern and southern regions in France (corresponding to the main dimension along which protection varied). The productivity distributions in the north and south are remarkably similar, and in both regions, productivity growth until 1840 was due to a disappearing lower tail.

Could the blockade explain the lower plant survival rates observed in cotton spinning relative to the other sectors? In Table D.17 we split mechanized cotton spinning plants into the same northern and southern regions and report survival rates separately. Indeed, consistent with Juhász (2018), survival rates are lower in the south than in the north (which experienced a relative increase in trade protection during the blockade). However, survival rates remain low relative to the other two sectors even when we narrow our sample to only northern plants.<sup>56</sup> This suggests that an important part of low survival rates in mechanized cotton spinning are not driven by the uneven effects of the Napoleonic blockade across France.

*Improvements in the Technology.* Another possible alternative mechanism is that productivity growth in cotton spinning was driven by technological improvements of the mechanized machinery after 1800. This is unlikely for two reasons. First, as we discussed in Section 2.1, the vintages of capital in cotton spinning remained largely unchanged during our sample period. Second, it is unclear why the successive adoption of better capital vintages would be concentrated in the lower tail of the plant-productivity distribution. In fact, in our two comparison sectors – where new vintages became available – the whole distribution shifted to the right.

In Table D.18 we examine whether capital deepening can account for our results by controlling directly for the capital-labor ratio of the plant (measured as the number of spindles per employee).

---

<sup>56</sup>Note that the raw baseline survival rate for northern plants is now slightly higher than in paper milling (11% and 9%, respectively). Recall, however, that the cotton spinning survey was conducted much later than paper milling (1806 as opposed to 1794). Moreover, the restricted sample survival rate in the north is half that measured in paper milling (10% and 20%, respectively). Taken together, this suggests that even comparing only cotton spinning plants in the north of France to the comparison sectors (calculated for all of France), survival rates for northern cotton spinning plants were low.

The lower-tail bias for productivity growth remains robust and similar in magnitude, suggesting that neither differential capital deepening nor changes in the vintages of capital drive our results.<sup>57</sup> In sum, our results are unlikely to be driven exclusively by technological improvements over time that were accessible to all plants.

*Quality Upgrading.* We assess the extent to which the lower-tail bias in mechanized cotton spinning could be driven by quality upgrading across plants over time. The concern is that if plants shifted their product mix to higher-quality products over time (or more precisely, plants producing higher-quality yarn entered the market), this could explain our result. We have no way of directly testing for this because we do not observe any information about the quality of output produced in the 1840 census. However, we can use our data in 1806 to examine the extent to which plants producing low quality account for the fat lower tail in that period. Unsurprisingly, quality is positively correlated with our measure of labor productivity (see column 2 in Table 6). However, quality does not account for a significant part of the lower-tail bias in productivity growth (see Panel B of Table D.1): The lower-tail bias of productivity growth in cotton spinning continues to hold even when using prices in 1806 that do not account for quality differences across plants. The negative gradient is not as steep when we do not adjust for quality, but productivity growth at the 10th percentile remains twice that of productivity growth at the 90th.

Finally, quality could still drive our results indirectly if it led to higher sales and thus larger plant size. To examine this possibility, we estimate the quantile regressions *not* adjusting for quality differences in prices across plants in 1806, and controlling for plant size by the number of employees (Table D.19). The lower-tail bias of productivity growth continues to hold. Taking all pieces of evidence together, it seems unlikely that quality upgrading alone could explain our results.

## 5 Conclusion

The unique setting examined in this paper allows us to shed light on some of the open questions in the technology adoption literature. First, our findings speak directly to why the *aggregate* productivity effect of major technological breakthroughs such as IT and electricity may be hard to pin down in the data. Based on our results, the full effects of a new technology may take significant time to materialize, as firms need to learn how the new technology can be operated efficiently once the technology has been widely adopted. In our context, adopting mechanized cotton spinning required producers to reorganize production from households to factories. In other settings, the specific challenges may be different, but they are plausibly subject to similar mechanisms. For mechanized cotton spinning, we estimate that productivity for the average plant using essentially

---

<sup>57</sup>Newer versions of the different machines on the market had more spindles. We use the number of spindles per worker as our measure of the capital-labor ratio of the plant (as opposed to the number of machines per worker). Thus, the results in Table D.18 capture the effect of newer machines having more spindles.

identical new production methods increased by about 82% until 1840, relative to when the technology was in its infancy in 1806. Put differently, observers estimating the productivity effect of switching from handspinning to mechanized spinning would significantly underestimate the long-run aggregate productivity gain if they only looked at the initial data around 1800.

Second, our results also shed light on the slow adoption of major new technologies. When there is uncertainty about how to operate a new technology efficiently, and the knowledge – once acquired – is observable to competitors, firms face a strategic incentive to delay adoption. The high exit rates observed in cotton spinning relative to other sectors, alongside the higher productivity observed for younger plants in 1806, suggest that plants that entered later were at an advantage. If firms understand the significant uncertainty they face when setting up a spinning mill at early stages of adoption, they have an incentive to delay the switch to the new technology in order to take advantage of the learning externalities generated by other early adopters.

Finally, our unique setting allows us to isolate a dimension of productivity growth that is usually hidden: Productivity differences across firms (or plants) reflect both the underlying technology and the efficiency with which the respective technology is used. Both features play important roles in firms' decisions to adopt new technologies: What are the potential productivity gains of a new technology (i.e., its frontier), and is the operational knowledge needed to achieve these gains (i.e., operate at the frontier) readily available? Separating these features empirically is difficult because of data limitations. In our context, the frontier in the mechanized cotton spinning technology changed relatively little between 1800 and 1840. At the same time, all plants in mechanized cotton spinning operated the new technology using identical or very similar vintages. This allows us to observe productivity differences that are driven by the (more or less efficient) *operation* of the new technology. Our results suggest that this dimension is important: Approaching the frontier of a given new technology via operational improvements can take a long time, and it can explain some of the salient features in the adoption of major innovations.

## References

- Allen, R. C. (2009). *The British Industrial Revolution in Global Perspective*. Cambridge University Press.
- Atkin, D., A. Chaudhry, S. Chaudry, A. K. Khandelwal, and E. Verhoogen (2017). Organizational Barriers to Technology Adoption: Evidence from Soccer-Ball Producers in Pakistan. *Quarterly Journal of Economics* 132(3), 1101–1164.
- Bandiera, O. and I. Rasul (2006). Social Networks and Technology Adoption in Northern Mozambique. *Economic Journal* 116, 869–902.
- Beaman, L., J. Magruder, and J. Robinson (2014). Minding Small Change Among Small Firms in Kenya. *Journal of Development Economics* 108, 69–86.

- BenYishay, A. and A. Mobarak (2018). Social Learning and Incentives for Experimentation and Communication. *Review of Economic Studies* 86, 976–1009.
- Bloom, N., B. Eifert, A. Mahajan, D. McKenzie, and J. Roberts (2013). Does Management Matter? Evidence from India. *Quarterly Journal of Economics* 128(1), 1–51.
- Boston Mutual Fire (1908). *Report No. 5: Slow-Burning or Mill Construction*. Boston Manufacturers Mutual Fire Insurance Co. 3rd ed., rev. and enl. Boston: The Company. .
- Bougin, H. and G. Bourgin (1920). *L'Industrie Sid'erurgique en France au D'ebut de la R'evolution*. Paris: Imprimerie Nationale.
- Braguinsky, S., A. Ohyama, T. Okazaki, and C. Syverson (2015). Acquisitions, Productivity, and Profitability: Evidence from the Japanese Cotton Spinning Industry. *American Economic Review* 105(7), 2086–2119.
- Cameron, R. (1985). A New View of European Industrialization. *Economic History Review* 38(1), 1–23.
- Catling, H. (1970). *The Spinning Mule*. David & Charles, Newton Abbott.
- Chanut, J., J. Heffer, J. Mairesse, and G. Postel-Vinay (2000). Editions de l'école des hautes études en sciences sociales, Paris.
- Chapman, S. (1970). Fixed Capital Formation in the British Cotton Industry, 1770-1815. *Economic History Review* 23(2), 235–253.
- Chapman, S. (1974). The Textile Factory Before Arkwright: A Typology of Factory Development. *Business History Review* 48(4), 451–478.
- Chassagne, S. (1991). *Le coton et ses Patrons: France 1760-1840*. Editions de l'Ecole des Hautes Etudes en Sciences Sociales.
- Collard-Wexler, A. and J. De Loecker (2015). Reallocation and Technology: Evidence from the US Steel Industry. *American Economic Review* 105(1), 131–171.
- Combes, P., G. Duranton, L. Gobillon, D. Puga, and S. Roux (2012). The Productivity Advantages of Large Cities: Distinguishing Agglomeration from Firm Selection. *Econometrica* 80(6), 2543–94.
- Comin, D. and B. Hobijn (2010). An Exploration of Technology Diffusion. *American Economic Review* 100(5), 2031–59.
- Conley, T. G. and C. R. Udry (2010). Learning About a new Technology: Pineapple in Ghana. *American Economic Review* 100, 35–69.
- Crafts, N. (1977). Industrial Revolution in England and France: Some Thoughts on the Question, 'Why Was England First'? *Economic History Review* 30(3), 429–441.
- Crafts, N. (1985). *British Economic Growth During the Industrial Revolution*. Oxford: Oxford University Press.
- Crafts, N. and K. Harley (1992). Output Growth and the British Industrial Revolution: A Restatement of the Crafts-Harley View. *Economic History Review* 45(4), 703–730.
- Crafts, N. and N. Wolf (2014). The Location of the UK Cotton Textiles Industry in 1838: A Quantitative Analysis. *Journal of Economic History* 74(4), 1103–1139.
- Daudin, G. (2010). Domestic Trade and Market Size in Late-Eighteenth-Century France. *Journal of Economic History* 70(3), 716–743.
- David, P. (1990). The Dynamo and the Computer: An Historical Perspective on the Modern Productivity Paradox. *American Economic Review* 80(2), 355–361.



- Duflo, E., M. Kremer, and J. Robinson (2011). Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya. *American Economic Review* 101(6), 2350–90.
- Emerick, K., A. de Janvry, E. Sadoulet, and M. Dar. (2016). Technological Innovations, Downside Risk, and the Modernization of Agriculture. *American Economic Review* 106(6), 1537–61.
- Fitton, R. and A. Wadsworth (1958). *The Strutts and the Arkwrights, 1758-1830: A Study of the Early Factory System*. Manchester University Press.
- Foster, A. and M. Rosenzweig (1995). Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture. *Journal of Political Economy* 103(6), 1176–1209.
- Foster, L., C. Grim, J. Haltiwanger, and Z. Wolf (2018). Innovation, Productivity Dispersion, and Productivity Growth. NBER Working Paper 24420.
- Foster, L., J. Haltiwanger, and C. Syverson (2008). Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability? *American Economic Review* 98(1), 394–425.
- Galor, O. (2011). *Unified Growth Theory*. Princeton University Press.
- Garcia-Marin, A. and N. Voigtländer (2019). Exporting and Plant-Level Efficiency Gains: It's in the Measure. *Journal of Political Economy* 127(4), 1777–1825.
- Giorcelli, M. (2019). The Long-Term Effects of Management and Technology Transfers. *American Economic Review* 109(1), 1–33.
- Gort, M. and S. Klepper (1982). Time Paths in the Diffusion of Product Innovations. *Economic Journal* (367), 630–653.
- Grantham, G. (1997). The French Cliometric Revolution: A Survey of Cliometric Contributions to French Economic History. *European Review of Economic History* 1(3), 353–405.
- Griliches, Z. (1957). Hybrid Corn: An Exploration in the Economics of Technological Change. *Econometrica* 25(4), pp. 501–522.
- Guiollard, P. (1993). *Les chevalements des houillères françaises*. Pierre-Christian Guiollard.
- Hall, B. (2004). Innovation and Diffusion. NBER Working Paper 10212.
- Hall, B. and B. Khan (2003). Adoption of New Technology. NBER Working Paper 9730.
- Hanna, R., M. Sendhil, and J. Schwartzstein (2014). Learning Through Noticing: Theory and Experimental Evidence in Farming. *Quarterly Journal of Economics* 129(3), 1311–1353.
- Hardy, M. and J. McCasland (2016). It Takes Two: Experimental Evidence on the Determinants of Technology Diffusion. Working Paper.
- Harley, C. (1998). Cotton Textile Prices and the Industrial Revolution. *Economic History Review* 51(1), 49–83.
- Horn, J. (2006). *The Path Not Taken: French Industrialization in the Age of Revolution, 1750-1830*. Cambridge MA: MIT Press.
- Hsieh, C. and P. Klenow (2009). Misallocation and Manufacturing TFP in China and India. *Quarterly Journal of Economics* 124(4), 1403–1448.
- Huberman, M. (1996). *Escape from the Market*. Cambridge University Press.
- Johnson, H. and A. Skempton (1955). William Strutt's Cotton Mills, 1793-1812. *Transactions of the Newcomen Society* 30(1).
- Jovanovic, B. (1982). Selection and the Evolution of Industry. *Econometrica* 50(3), 649–670.

- Juhász, R. (2018). Temporary Protection and Technology Adoption: Evidence from the Napoleonic Blockade. *American Economic Review* 108(11), 3339–76.
- Langenbach, R. (2013). Better than Steel? Part 2: Tall Wooden Factories and the Invention of “Slow-burning” Heavy Timber Construction. Article URL accessed June 20, 2020.
- Mansfield, E. (1961). Technical Change and the Rate of Imitation. *Econometrica* 29(4), 741–766.
- Melitz, M. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica* 71(6), 1695–1725.
- Mitchell, B. (2003). *International Historical Statistics. Europe, 1750-2000*. Palgrave Macmillan.
- Mokyr, J. (2001). The Rise and Fall of the Factory System: Technology, Firms, and Households since the Industrial Revolution. *Carnegie-Rochester Conference Series on Public Policy* 55, 1–45.
- Mokyr, J. (2010). *The Enlightened Economy: An Economic History of Britain 1700-1850*. Yale University Press.
- Munshi, K. (2004). Social Learning in a Heterogeneous Population: Technology Diffusion in the Indian Green Revolution. *Journal of development Economics* 73, 185–213.
- Perla, J. and C. Tonetti (2014). Equilibrium Imitation and Growth. *Journal of Political Economy* 122(1), 52–76.
- Perrot, J. (1975). *L'âge d'Or de la Statistique Régionale Française (An IV-1804)*. Paris: Société des Etudes Robespierriéristes.
- Pollard, S. (1963). Factory Discipline in the Industrial Revolution. *Economic History Review* 16(2), 254–271.
- Pollard, S. (1965). *The Genesis of Modern Management: A Study of the Industrial Revolution in Great Britain*. Harvard University Press.
- Pounds, N. and W. Parker (1957). *Coal and Steel in Western Europe: The Influence of Resources and Techniques on Production*. Bloomington: Indiana University Press.
- Riello, G. (2013). *Cotton: The Fabric that Made the Modern World*. Cambridge University Press.
- Rosenberg, N. (1976). Factors Affecting the Diffusion of Technology. *Explorations in Economic History* 10(1), 3–33.
- Rostow, W. (1975). *How it All Began: Origins of the Modern Economy*. Methuen and Co., London.
- Saxonhouse, G. and G. Wright (2004). Technological Evolution in Cotton Spinning. In D. Farnie and D. Jeremy (Eds.), *The Fibre that Changed the World: The Cotton Industry in International Perspective, 1600-1990s*, pp. 129–152. Oxford University Press.
- Sokoloff, K. (1984). Was the Transition from the Artisanal Shop to the Nonmechanized Factory Associated with Gains in Efficiency? Evidence from the US Manufacturing Censuses of 1820 and 1850. *Explorations in Economic History* 21(4), 351–382.
- Sokoloff, K. (1986). Productivity Growth in Manufacturing During Early Industrialization: Evidence from the American Northeast, 1820-1860. In *Long-Term Factors in American Economic Growth*, pp. 679–736. University of Chicago Press.
- Squicciarini, M. and N. Voigtländer (2015). Human Capital and Industrialization: Evidence from the Age of Enlightenment. *Quarterly Journal of Economics* 30(4), 1825–1883.
- Suri, T. (2011). Selection and Comparative Advantage in Technology Adoption. *Econometrica* 79, 159–209.
- Syverson, C. (2011). What Determines Productivity? *Journal of Economic Literature* 49(2), 326–365.

- Szostak, R. (1989). The Organization of Work: The Emergence of the Factory Revisited. *Journal of Economic Behavior and Organization* 11(3), 343–358.
- Tarr, R. and F. McMurry (1993). *New Geographies: Second Book*. Macmillan Company.
- Vallauri, D., A. Grel, E. Granier, and J. Dupouey (2012). *Les Forêts de Cassini. Analyse Quantitative et Comparaison Avec les Forêts Actuelles*. Marseille: WWF/INRA.
- Voigtländer, N. and H. Voth (2006). Why England? Demographic Factors, Structural Change and Physical Capital Accumulation During the Industrial Revolution. *Journal of Economic Growth* 11(4), 319–361.
- Ware, C. (1931). *Early New England Cotton Manufacture; A Study in Industrial Beginnings*. Houghton Mifflin Company.
- Williamson, O. (1980). The Organization of Work: A Comparative Institutional Assessment. *Journal of Economic Behavior and Organization* 1(1), 5–38.
- Woronoff, D. (1984). *L'industrie sidérurgique en France pendant la Révolution et l'Empire*. Paris: École des hautes études en sciences sociales.

# FIGURES

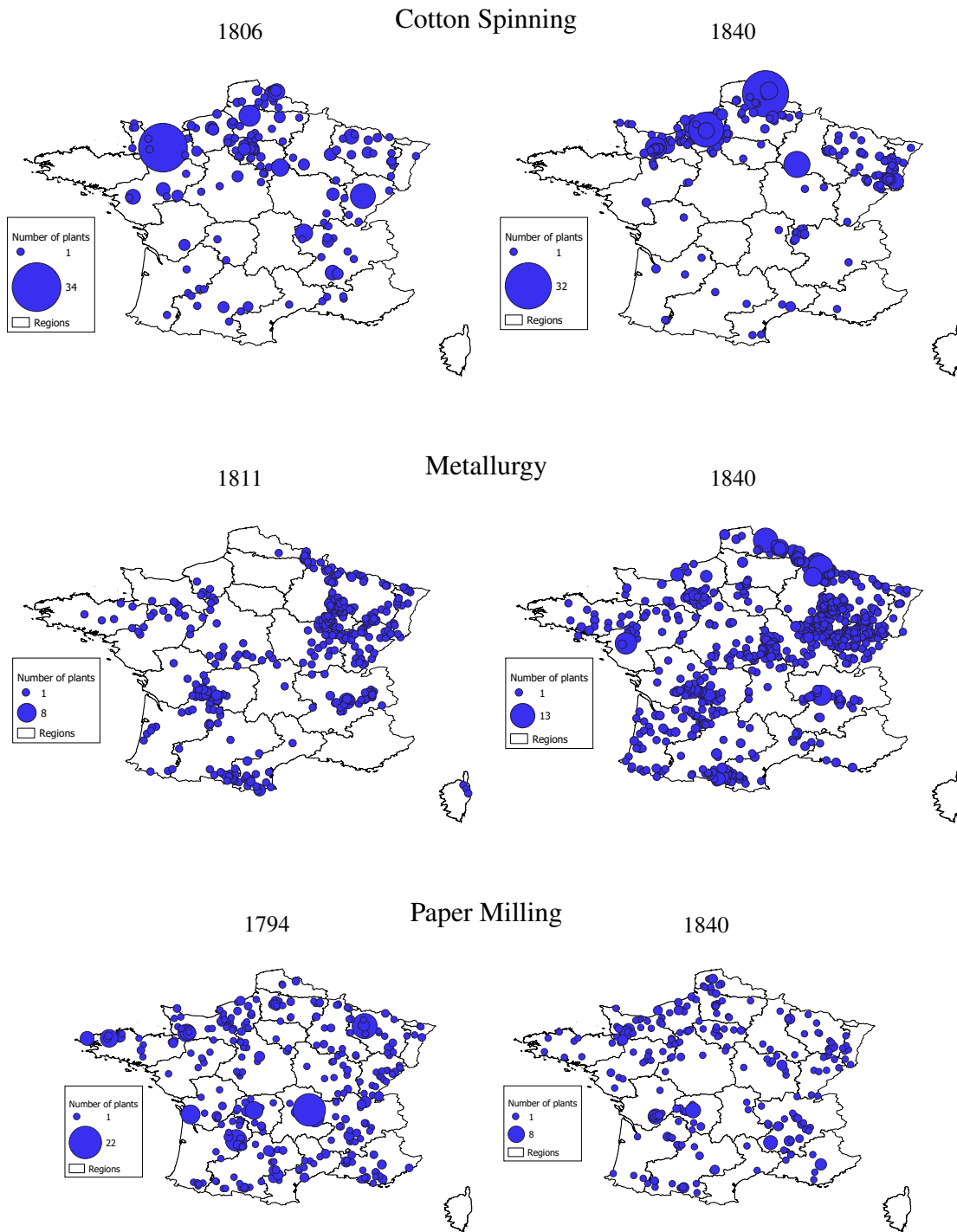


Figure 1: Spatial Distribution of Plants Across France in the Three Sectors

Note: The figure shows the spatial distribution of plants in cotton spinning (top), metallurgy (middle), and paper milling (bottom). Dot sizes reflect the number of plants per commune.

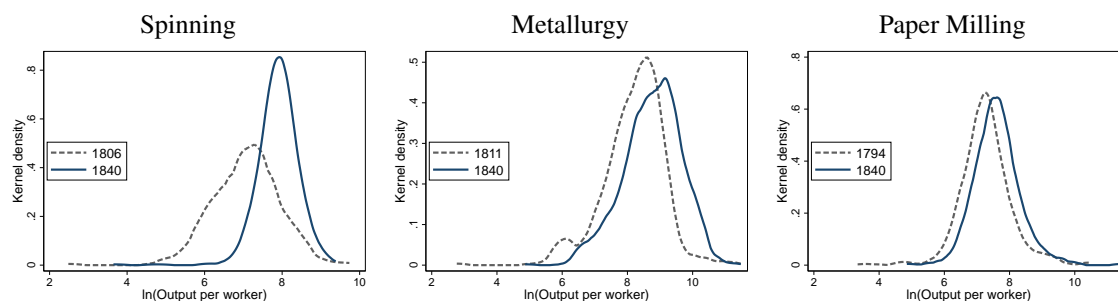


Figure 2: Changes in the Productivity Distributions in the Three Sectors

Notes: The figure shows the distribution of  $\log(\text{output per worker})$  for the three sectors at the beginning of our sample period (around 1800) and in 1840. Productivity growth in spinning was mainly due to a disappearing lower tail. In contrast, in metallurgy and paper milling, the whole distribution shifted to the right.

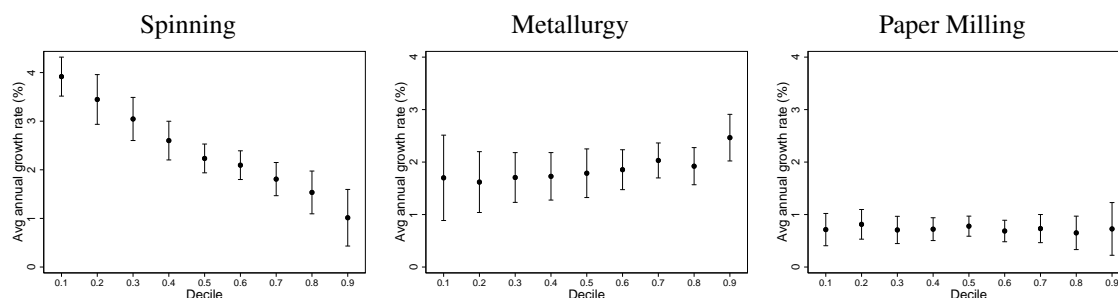


Figure 3: Productivity Growth at Different Quantiles of the Distribution

Notes: The figure visualizes the results of quantile regressions for growth in  $\ln(\text{output per worker})$  for the three sectors, estimated at each decile. Productivity growth in spinning was concentrated in the lower tail of plant productivity. In contrast, in metallurgy and paper milling, productivity growth occurred relatively evenly across the distribution.

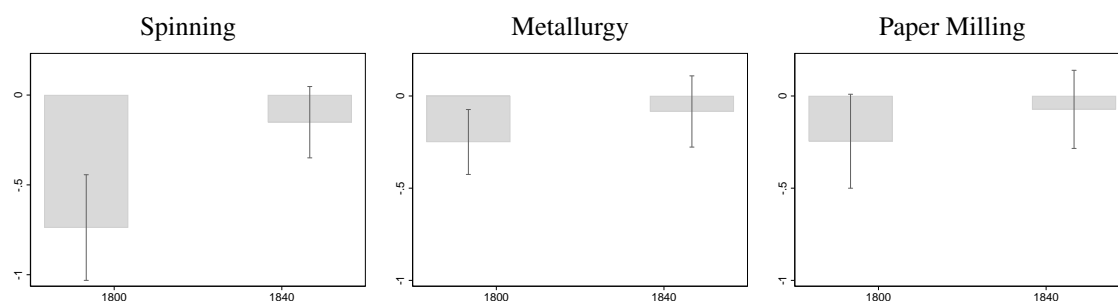


Figure 4: Proximity to High-Productivity Plants

Notes: The figure shows that proximity to high-productivity plants mattered the most in mechanized cotton spinning at the beginning of our sample period (around 1800), when the technology had just been introduced in France. The figures plots the standardized beta coefficients of  $Dist^{p90}$ , which measures the log distance to the closest plant (in the same sector) with productivity in the 90th percentile. The dependent variable is  $\log(\text{output per worker})$ . All regressions include department fixed effects (see Table D.7). Whiskers indicate 90% confidence intervals.

## TABLES

Table 1: Background of Owners of Mechanized Cotton Spinning Establishments

<i>Owners active between 1785-1815</i>	
Traders, bankers, and commercial employees	62.5%
Nobility or administrator pre-1789	10.2%
Workers and mechanics	10.2%
Industrialists	9.5%
Liberal profession	6%
Other	3.5%

*Notes:* Data are from Chassagne (1991, p. 274). The author collected data on the owners from a variety of archival source. The sample covers 185 establishments in mechanized cotton spinning.

Table 2: Summary Statistics: Plant Scale in the Three Sectors

Sector	year	(1)	(2)	(3)	(4)	(5)	(6)
		mean	sd	median	10perc	90perc	N
Cotton spinning	1806	63	(101)	30	4	150	372
	1840	112	(148)	72	28	210	528
Metallurgy	1811	20	(23)	11	4	46	457
	1840	57	(114)	22	7	135	839
Paper milling	1794	13	(19)	11	5	23	550
	1840	43	(58)	19	5	112	348

*Notes:* The table reports summary statistics on the number of workers per plant in the three sectors covered by our analysis. The year of the first survey varies across the sectors, while information in 1840 is available for all sectors. Data sources: See Section 3.



Table 3: Annual Productivity Growth (in %) at Different Quantiles of the Distribution

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Average	At the following quantiles:					N
		0.1	0.25	0.5	0.75	0.9	
Spinning (1806-1840)	2.420*** (0.154)	3.917*** (0.204)	3.293*** (0.229)	2.234*** (0.151)	1.651*** (0.167)	1.014*** (0.297)	868
Metallurgy (1811-1840)	1.949*** (0.185)	1.700*** (0.415)	1.776*** (0.271)	1.787*** (0.236)	2.025*** (0.187)	2.465*** (0.226)	1296
Paper milling (1794-1840)	0.734*** (0.111)	0.713*** (0.157)	0.681*** (0.137)	0.779*** (0.098)	0.759*** (0.137)	0.726*** (0.256)	868

*Notes:* The table reports the average annual productivity growth (in %) between the initial sample period (around 1800) and 1840 for the three sectors (column 1), and annual productivity growth estimated at different quantiles (columns 2-6). Column 7 reports the number of observations. Robust standard errors in parentheses. Notation for statistical significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 4: Survival Rates Across Sectors

	Spinning	Metallurgy	Paper milling
Period	1806-1840	1811-1840	1794-1840
Survival rate	7.5%	34%	9%
Number of plants	389	477	593
Restricted sample survival rate	6.5%	49%	20%
Number of plants	93	303	218

*Notes:* The “survival rate” is defined as the percentage of plants from the initial period that survived to the later period based on matching either by name or location (see Section 3.5 for details). The “restricted sample survival rate” adjusts for the fact that different sectors have single-plant communes to a varying degree. It is based on the subset of plants located in communes that had only one plant in the initial period and that had either exactly one plant in 1840 (‘survival’) or no plant at all in the 1840 data (‘exit’).

Table 5: Productivity of Exiting Relative to Surviving Plants

	(1)	(2)	(3)
Dependent variable	log(Y/L)	log(L)	log(Y)
Spinning (exit = 1)	-0.533*** (0.165)	-0.869*** (0.218)	-1.402*** (0.251)
Metallurgy (exit = 1)	-0.139 (0.087)	-0.439*** (0.089)	-0.578*** (0.097)
Paper milling (exit = 1)	-0.179 (0.150)	-0.151 (0.131)	-0.331* (0.172)

*Notes:* The dependent variables are log(output per worker) in column 1, log(total labor) in column 2, and log(output) in column 3. Exit is a dummy variable equal to one for plants that existed in the initial period and that had exited the market by 1840 (based on the baseline survival rate – see Section 3.5). In cotton spinning, there were 340 plants in 1806 with information on output and labor, and 314 of these had exited by 1840. In metallurgy, there were 457 plants with data to compute productivity in 1811, and 292 had exited by 1840. In paper milling, there were 520 plants with information on output and labor in 1794, 473 of which had exited by 1840. Robust standard errors in parentheses. Notation for statistical significance: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

Table 6: Cotton Spinning in 1806: Productivity and Plants' Age Profile

Dependent variable: log(Output per worker)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Young plant	0.575*** (0.088)	0.374*** (0.079)	0.543*** (0.083)	0.575*** (0.089)	0.493*** (0.085)	0.608*** (0.086)	0.534*** (0.085)
log(Yarn quality)		0.673*** (0.074)					
Spinning jenny			-0.626*** (0.087)				
Throstle				-0.003 (0.092)			
Mule jenny					0.481*** (0.086)		
log(Workers)						0.107*** (0.025)	
log(Spindles per worker)							0.336*** (0.070)
R <sup>2</sup>	0.11	0.32	0.20	0.11	0.18	0.14	0.17
N	340	323	340	340	340	340	340

*Notes:* The table shows that mechanized cotton spinning plants that had just entered the market by 1806 had a significantly higher productivity. 'Young' is a dummy variable equal to one for cotton spinning plants with below-median age (with the median age in 1806 being three years). log(Yarn quality) is the log (unweighted) average of the minimum and maximum count of the yarn produced by the plant. Spinning-jenny, throstle and mule-jenny are binary indicators equal to one for plants using the earliest (spinning-jenny), intermediate (throstle), and latest (mule-jenny) vintage of spinning machinery, respectively. The number of spindles is a standard measure of a spinning machine's production capacity, irrespective of vintage. Robust standard errors in parentheses. Notation for statistical significance: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table 7: Cotton Spinning in 1840: Productivity and Plants' Age Profile

Dependent variable: log(Output per worker)					
	(1)	(2)	(3)	(4)	(5)
Entrant 1840	-0.084 (0.077)	-0.029 (0.077)	-0.080 (0.078)	-0.078 (0.077)	-0.144** (0.065)
Water power		0.327*** (0.062)			
Steam power			-0.045 (0.076)		
Other power				-0.193** (0.090)	
log(Workers)					-0.373*** (0.027)
R <sup>2</sup>	0.00	0.03	0.00	0.01	0.24
N	839	839	839	839	839

Notes: The table shows that in 1840, when the mechanized cotton spinning technology had reached maturity, new entrant plants did not have a productivity advantage over existing plants anymore. Entrant 1840 is a binary indicator equal to one for plants that entered the market after 1806. Water power, steam power, and other (wind or animal) power are binary indicators equal to one for plants using the respective source of power. Robust standard errors in parentheses. Notation for statistical significance: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.

Table 8: Testing for Spatial Selection of New Plants in Cotton Spinning in 1806

Dependent variable: log(Output per worker)					
	(1)	(2)	(3)	(4)	(5)
			Only plants entering before high-productivity plants		
$Dist^{p90}$ (1806)	-0.791*** (0.136)	-0.845*** (0.129)	-0.439*** (0.153)	-0.393** (0.153)	-0.481** (0.196)
Plant age (in 1806)	-0.046 (0.085)	-0.203 (0.135)		-0.153 (0.133)	-0.388* (0.205)
Plant age $\times$ $Dist^{p90}$ (1806)		0.237 (0.203)			0.365 (0.258)
Department FE	✓	✓	✓	✓	✓
R <sup>2</sup>	0.56	0.57	0.66	0.66	0.67
N	284	284	176	176	176

Notes: The table shows that our results for proximity to high-productivity plants (see Figure 4) are not confounded by new, more productive plants entering near existing high-productivity plants.  $Dist^{p90}$  (1806) is the log distance to the nearest plant in cotton spinning with productivity in the 90th percentile in 1806. Plant age is the log number of years since the plant had been established, reported in 1806. Standard errors (clustered at the departmental level) in parentheses. Notation for statistical significance: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1.