Innovation and Industry Development: Lessons from the British Cotton Textile Industry During the U.S. Civil War

William Walker Hanlon

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2012
Abstract

Innovation and Industry Development: Lessons from the British Cotton Textile Industry During the U.S. Civil War

W. Walker Hanlon

This dissertation uses the large shock to the British cotton textile industry in the 19th century, caused by the U.S. Civil War (1861-1865), in order to address three long-running questions about technological progress and industry development. The cotton textile industry was a large and important sector in the British economy during the 19th century. The industry was entirely dependent on imported raw cotton, most of which came from the U.S. South prior to the Civil War. The onset of the war sharply reduced the supply of Southern cotton to the British market, causing a severe downturn in the industry. In response to the shock, cotton textile producers turned to other sources of supply, chiefly India, but also Egypt, Brazil and others, to help meet their raw cotton needs. But cotton from these alternative suppliers, and India in particular, was very different from the U.S. cotton that British producers were used to spinning. As a result, British cotton textile producers were faced with a number of new challenges. The first two chapters of this dissertation describe how the cotton textile industry developed new technology in order to deal with these challenges, and what this response can tell us about the process of innovation. Chapter three then investigates the impact of the recession on other industries in the British economy.
This historical setting has a number of features which makes it a particularly good setting for investigating technological progress. First, the U.S. Civil War caused a shock which was both large and exogenous. The size of the shock ensures that the response will be large enough to clearly observe, while the exogenous nature of the shock means that it can be used as a natural experiment in order to uncover causal relationships. Second, the impact of the U.S. Civil War was largely industry-specific; while the impact of the war on the cotton textile industry was severe, most other sectors of the British economy were not directly impacted. This includes other textile industries based on wool, linen, and silk, which do not show any ill effects during the war. One advantage of this feature is that it allows me to control for other time-varying factors by comparing the cotton textile industry to these other similar industries. I will also be able to uncover evidence of inter-industry connections, since other industries will be affected primarily through their relationship with the cotton textile industry. Another feature of this shock is that, despite the magnitude, there was virtually no government intervention in the affected markets. This unique feature was due to the particularly strong free-market ideology that was dominant in Britain during this period.

The first chapter investigates the theory of directed technical change. The leading theory of directed technical change, developed by Acemoglu (2002), offers two main predictions. First, when inputs are sufficiently substitutable, a change in relative input supplies will generate technical change directed towards the inputs which become more abundant. Second, if technical change is strongly directed towards the more abundant inputs, the relative price of these inputs will increase – the strong induced-bias hypothesis. The chapter provides the first empirical test of these predictions. I extend the theory to a setting in which input quantities are endogenous and affected by international transport cost shocks, such as that caused by the war. Using detailed new patent data, I show that there was a burst of cotton textile innovation in Britain during the war directed towards taking advantage of one input – Indian cotton – which became relatively more abundant. Next, I show that the relative price of Indian cotton first declined and then rebounded, consistent with the strong induced-bias hypothesis. These results provide support for the theory. My extended model also predicts that technical change directed towards the more abundant input will be
magnified by a higher elasticity of input supply. This may explain why inventors chose to focus on innovations for Indian cotton, rather than Brazilian or Egyptian cotton, since I find evidence that the elasticity of supply was higher for Indian cotton.

In the second chapter, I look at whether the stock of available knowledge about a particular type of technology can influence the rate of innovation in that technology. The answer to this question has significant implications for how we think about technological progress and economic growth. This chapter provides a theory which describes how the stock of knowledge can influence the innovation rate, which I call path dependence in innovation. The theory suggests that path dependence in innovation may occur at multiple levels of aggregation, such as specific types of technologies within an industry. This motivates an empirical exercise in which I search for path dependence at multiple technology levels. I introduce an empirical methodology that addresses two potential sources of bias in generating evidence of path dependence in innovation by using a temporary observable shock to innovation rates. My results provide no evidence of path dependence in innovation for cotton textile technologies. However, I do find suggestive evidence of path dependence in innovation for specific subsets of technologies within the cotton textile industry. This illustrates the importance of looking at multiple levels of aggregation when studying path dependence in innovation.

Chapter 3 provides causal evidence that inter-industry connections can influence the geographic location of economic activity. To do so, it compares the impact of the shock caused by the U.S. Civil War on towns in Lancashire County, the center of Britain’s cotton textile industry, to towns in neighboring Yorkshire County, where wool textiles dominated. The results suggest that the shock reduced employment and employment growth in industries related to the cotton textile industry, in towns that were more severely impacted by the shock, relative to less affected towns. The impact still appears over two decades after the end of the U.S. Civil War. This suggests that temporary shocks, acting through inter-industry connections, can have long-term impacts on the distribution of industrial activity across locations.

Each of the three chapters are entirely self contained, so that a reader interested in only one of
these topics need focus on only one chapter. As a result, each chapter contains an overview of the
features of the empirical setting which are relevant for that chapter. There are significant overlaps
between these descriptions. Also, chapters one and two are largely based on the same British
patent data, though each chapter uses some parts of the data which are not used in the other.
# Contents

1  Necessity is the Mother of Invention: Input Supplies and Directed Technical Change  
   1.1 Introduction .................................................. 1  
   1.2 Theory ......................................................... 9  
   1.3 Empirical setting ............................................. 15  
   1.4 Patent data .................................................. 32  
   1.5 Impact on innovation ....................................... 35  
   1.6 Indicators of patent quality ................................ 46  
   1.7 Impact on input prices ..................................... 56  
   1.8 Conclusions .................................................. 65  

2  Is There Path Dependence in Innovation? Theory and Evidence from the 19th Century British Cotton Textile Industry  
   2.1 Introduction .................................................. 67  
   2.2 Theory ......................................................... 71  
   2.3 Empirical setting ............................................. 79  
   2.4 Patent data .................................................. 83  
   2.5 Empirical issues and approach ............................. 87  
   2.6 Analysis ...................................................... 93  
   2.7 Conclusion .................................................. 111  

3  Industry Connections and the Geographic Location of Economic Activity  
   ..................
3.1 Introduction ......................................................... 113
3.2 Empirical setting .................................................. 119
3.3 Data ............................................................... 125
3.4 An Example: Engine & Machine Makers ......................... 130
3.5 Model ............................................................. 133
3.6 Econometric analysis ............................................. 141
3.7 Conclusion ......................................................... 152

References .......................................................... 162

A Appendix to Chapter 1 .............................................. 163
A.1 Definitions of important textile terms .......................... 163
A.2 British cotton and wool imports and re-exports 1861 .......... 164
A.3 Machines for ginning cotton .................................... 165
A.4 Details on the differences between cotton types ................. 166
A.5 Impact of ginning on cotton fiber length ......................... 167
A.6 Background graphs .............................................. 167
A.7 Cotton waste data ................................................ 168
A.8 First-stage regression results for Table 1.2 ...................... 169
A.9 First-stage regression results for Table 1.3 ...................... 170
A.10 Robustness of the estimated impact on textile technologies . 170
A.11 Robustness of the estimated impact on spinning technology subcategories . . . . 172
A.12 Subcategory regressions for other textile inputs ................ 173

B Appendix to Chapter 2 .............................................. 175
B.1 Industry-level Poisson regression results ....................... 175
B.2 Industry-level regression results using only high-quality patents .......................... 177
B.3 Industry-level results without cross-listed patents .............. 178
B.4 Intermediate-level regression results from OLS and IV regressions .............................. 180
B.5 Intermediate-level regression results with Carding machines as a treated technology .................................. 181

C Appendix to Chapter 3 .......................................................... 183
C.1 Empirical setting appendix .................................................. 183
C.2 Data appendix ..................................................................... 186
C.3 Theoretical appendix .......................................................... 190
C.4 Additional statistical results .................................................. 193
List of Figures

1.1 British cotton imports 1815-1910 ......................... 18
1.2 British raw cotton consumption 1815-1910 ................ 18
1.3 British cotton imports by supplier 1850-1880 ............... 19
1.4 Share of imports by supplier 1850-1880 ..................... 20
1.5 Raw cotton prices on the Liverpool market for key varieties 1852-1875 ........ 20
1.6 Share of waste in total raw cotton input 1860-1868 .......... 23
1.7 An Example: Patent No. 2162 from 1862 ..................... 24
1.8 Transportation cost index between New Orleans and Liverpool .......... 29
1.9 Patent counts from BPO technology categories ............... 36
1.10 Count of patents with titles mentioning main textile inputs, 1853-1870 .......... 37
1.11 Monthly patent counts of cotton-related patents ............. 40
1.12 Patent counts in subcategories of the BPO Spinning technology category ........ 43
1.13 Cotton-related and gin/opener/scutcher technology patent renewals ........... 51
1.14 Prices and import quantities for raw cotton by location .............. 58
1.15 Cotton prices relative to the benchmark U.S. variety ............. 60
1.16 Comparing movement of relative cotton prices for different varieties ........ 61
1.17 Estimated impact on the relative price of Indian cotton by year ............. 63
1.18 Comparison of cotton prices on the Bombay and Liverpool markets ........... 65

2.1 British cotton imports 1815-1910 ......................... 80
2.2 Share of imports by supplier 1850-1880 ..................... 80
2.3 British raw cotton consumption 1815-1910 .................................................. 81
2.4 Raw cotton prices on the Liverpool market for key varieties 1852-1875 .......... 82
2.5 British wool and linen textile exports 1815-1910 ........................................... 82
2.6 Relationship between PAT and STOCK with a permanent shock at \( t = 10 \) .... 89
2.7 Patents in each textile industry from 1853-1870 .............................................. 94
2.8 High-quality patents in each textile industry from 1853-1870 ......................... 95
2.9 Inventors with experience in the past 4 years ................................................. 101
2.10 Patents by inventors with experience in the past 4 years ................................ 101
2.11 Gins and Openers/Scutchers patents 1855-1876 ........................................ 103
2.12 Patent counts in subcategories of the BPO Spinning technology category ...... 104
2.13 Gins and Openers/Scutchers patents for which renewal fees were paid .......... 108
2.14 Openers/Scutchers patents for cotton and other textile industries 1855-1876 . . 110
3.1 Maps of the study area ............................................................................... 120
3.2 Population and Occupation Comparisons - 1851-1891 ................................. 121
3.3 British Cotton Imports and Raw Cotton Prices 1815-1910 ............................ 123
3.4 Able-bodied relief seekers as a share of 1861 population ................................ 125
3.5 E&M Industry employment by location ....................................................... 133
A.1 British cotton and wool imports and re-exports in 1861 ................................. 164
A.2 Indian Churka for removing cotton seeds .................................................... 165
A.3 Cottage Saw Gin ............................................................................. 165
A.4 Length of cotton staples for various cotton types .......................................... 166
A.5 A comparison of ginned (left) and hand-cleaned cotton (right) fiber length .... 167
A.6 British imports and exports 1851-1869 ...................................................... 167
A.7 British wool and linen textile exports 1815-1910 ......................................... 168
C.1 Distribution of cotton textile manufacturing stages in Lancashire towns ........ 184
C.2 Total British cotton imports and imports from the U.S. 1815-1910 ............... 184
C.3 British imports and exports 1851-1869 ........................................ 185
C.4 British wool imports and exports 1850-1880 .............................. 185
C.5 Histograms describing the coagglomeration relatedness measure (171 industries) 189
C.6 Histograms describing the input-output relatedness measure (43 sectors) .... 189
C.7 Cotton employment in towns grouped by 1851 cotton employment share .... 194
Acknowledgments

I am grateful to my advisors, Donald Davis and Eric Verhoogen, for their guidance and support. I thank Suresh Naidu, Ronald Findlay, Amit Khandelwal, David Weinstein and Jonathan Vogel for their many thoughtful comments and suggestions. Thanks to Tom Nicholas, Lucie Gadenne, Miguel Morin, Eric Hilt, Jonathan Dingel, Ryan Chahrour, William Kerr, Richard Hornbeck, Henry Overman, Guy Michaels, and the many other seminar participants at Columbia and elsewhere who provided helpful comments. Thanks to Tim Leunig and Humphrey Southall for their help understanding this period in economic history. I thank the staff of the Business and IP Center at the British Library for their help in collecting the data used in Chapters 1 and 2. Matthew Wollard and Ole Wiedenmann at the U.K. Data Archive provided help with the data used in Chapter 3. I acknowledge the support of the National Science Foundation (Grant No. 0962545), the Economic History Association, the Mellon Foundation, and the Kauffman Foundation.
Chapter 1

Necessity is the Mother of Invention: Input Supplies and Directed Technical Change

1.1 Introduction

The idea that a change in the availability or price of inputs to production can play an important role in influencing the rate and direction of technical progress has been used to explain a diverse set of economic phenomena. It has been suggested that the increase in skilled workers in the U.S. in the 1970s caused skill-biased directed technical change, and that this directed technical change allowed the skill premium to increase in spite of the increase in the relative abundance of skilled workers (Acemoglu (1998), Kiley (1999)). Looking at an earlier period, several authors have suggested that a shortage of labor drove the development of labor-saving innovations which played an important role in industrialization in Britain and the U.S. (Habakkuk (1962), Allen (2009)). In the environmental literature, it has been pointed out that the impact of regulations that change the price of inputs, such as a carbon tax, will depend crucially on whether these changes generate directed technical change, and on the direction that this innovation takes (Acemoglu et al. (2012)). Yet, despite the wide application of this idea, several important predictions of leading

\[^1\] Related papers in the environmental literature include Porter (1991), Lanjouw & Mody (1996) and Jaffe & Palmer (1997). The idea of directed technical change has also been applied to consider the impact of high energy prices (Newell et al. (1999), Popp (2002)), the causes of cross-country productivity differences (Acemoglu & Zilibotti (2001), Caselli & Coleman (2006)), and agricultural productivity trends (Hayami & Ruttan (1970), Olmstead &
directed technical change theories remain largely unexplored empirically.

Recently, there have been important advances in the theory of directed technical change by Acemoglu (2002, 2007), building on previous work by Hicks (1932), Kennedy (1964), Samuelson (1965) and Drandakis & Phelps (1966). Acemoglu (2002) presents an endogenous growth model that incorporates multiple inputs and uses it to consider the impact of a change in the supply of inputs, an exogenous parameter, on technological progress. The model provides two new results. First, Acemoglu shows that the direction of technical change depends crucially on the elasticity of substitution between inputs, represented by $\sigma$. When this elasticity is low ($\sigma < 1$), technical change will be directed towards technologies that augment the input which has become relatively scarce. In contrast, when the elasticity of substitution between inputs is high ($\sigma > 1$), technical change will be directed towards technologies that augment the input which has become relatively more abundant. This prediction stands in contrast to previous work, such as Hicks (1932), which hypothesized that technical change would always augment the more scarce input. The second major prediction of Acemoglu (2002) is that, when the elasticity of substitution between inputs is sufficiently high ($\sigma > 2$), technical change will be so strongly directed towards technologies that augment the more abundant input that the relative price of that input can increase. This strong induced-bias hypothesis may explain, for example, how an increase in the supply of skilled workers may increase the skill wage premium.

The aim of this chapter is to test these predictions. To do so, I consider a large exogenous shock to the British cotton textile industry caused by the U.S. Civil War (April 1861 - April 1865). The war, which included a blockade on Southern shipping by the Union navy, sharply increased the cost of transporting U.S. cotton from the South, which provided most of the raw cotton imported into Britain prior to the war (77% in 1860). This forced British producers to turn to raw cotton from alternative suppliers, such as India, Brazil, and Egypt. The cotton available from these alternative suppliers differed from American cotton in important ways. This was particularly true for cotton from India, the second largest supplier, which was a low-quality variety that was more than...
difficult to clean and prepare.

The shock generated a large exogenous shift in the relative cost of providing inputs to production that is used to identify the causal impact on the direction of technical change and input prices. This exogenous change is the first necessary ingredient for testing the theory. I construct an extensive data set covering the price and import quantity of raw cotton supplied by the U.S., India, Brazil, and Egypt before, during, and after the war. Observing data on prices and quantities for multiple inputs is the second crucial ingredient for testing the predictions of the theory, since they allow me to estimate the elasticity of substitution between inputs and track the impact on input prices. In order to track the impact on innovation rates, I gathered previously unexploited data on British patents containing a high level of detail on the types of new technologies being created. Using these patent data it is possible to track patterns of technical change in technologies related to particular inputs, which is a third necessary element for testing the theory. Previous studies have not incorporated all of these necessary elements, and as a result, the main predictions of Acemoglu (2002) have not yet been tested empirically. The novel contribution of this study is to introduce a setting with these features and use it to test the predictions of the theory.

Three other features of the empirical setting are important. First, the impact of the Civil War on the cotton textile industry was large and lasted for several years. There is evidence that output in the industry dropped by as much as 50%. Hundreds of thousands of mill operatives found themselves out of work or working short-time. Thus, this event was large enough to influence innovation rates. Second, I will be able to compare outcomes in the the cotton textile industry to other similar textile industries – based on wool, linen, and silk – which were also important in Britain during this time, but which were not negatively impacted by the Civil War.² This will help me control for other time-varying factors that may be affecting innovation rates. Third, despite the magnitude of the shock, there was virtually no government intervention. This was primarily due to the strong free-market ideology which was dominant in Britain at this time. This reduces the chance that the effects I observe are influenced by government action, which may be a serious

²If anything, these industries benefited somewhat from the reduction in competition from cotton textiles.
I track innovation using British patent data covering 1853-1883. Patent data are one of the most commonly used measures of innovation. Most of the data used in this study were collected from original sources, with over 1500 pages of data digitized, and contain a high level of detail on the types of technologies represented by each patent. While British patent data for this period has been used previously, the novel feature of my data is that each invention has been categorized into one or more of 146 technology categories by the British Patent Office (BPO). These BPO technology categorizations allow me to observe the type of technology represented by each patent. For example, I am able to identify patents for technologies that fit into the two main textile-related BPO categories, spinning and weaving. Using the patent titles, I am also able to identify patents related specifically to the cotton, wool, linen, and silk industries. Even more detailed data were gathered for BPO subcategories within the textile-related categories, which allow me to track patents for specific types of machines, such as cotton gins or spinning mules. Because raw patent counts do not reflect the wide variation in the importance of patented inventions, I gathered three types of additional data in order to adjust for three dimensions of patent quality. Data on whether patent holders paid substantial renewal fees to keep their patents in force are used to indicate the long-term viability of each invention. Data on patent applications in India and the U.S. are used to indicate the wider applicability of the inventions. Listings from a contemporary publication, *Newton’s London Journal*, are used to assess the initial novelty or interest that each patent generated.

The analysis begins with a theory that extends the model of Acemoglu (2002). In his model, final goods are produced using inelastically supplied factors, and results are derived by varying relative factor supplies and observing the impact on technical change. I extend the model by incorporating elastically supplied inputs in place of the fixed factors. Generalizing the model in this way is important for making it empirically testable, since in practice we almost never observe truly fixed factors. In my model, inputs are supplied through trade and each input faces an input-
specific transport cost. This connects the model to the empirical setting, where the U.S. Civil War can be thought of as a large increase in the transport cost of U.S. cotton. The main predictions of the model correspond to those found by Acemoglu (2002) which are described above. It also generates one new prediction: that input supply responses can act to magnify directed technical change and its impact on input prices. So, for example, when the elasticity of substitution between inputs is high, so that we expect technical change to be directed towards the input which becomes more abundant, this effect will be stronger the higher is the elasticity of input supply.

The predictions of the theory depend on the elasticity of substitution between inputs, and the elasticity of input supply for each input, so the first step in the empirical analysis is to estimate these elasticities. To estimate these parameters, I used data on the prices and quantities of the various cotton varieties, most of which I gathered from contemporary periodicals. I estimate the elasticity of input supply for Indian, Brazilian, and Egyptian cotton by comparing the response of import quantities to the change in their prices, where I use the shock as an instrument for the increase in the cotton price. These estimates suggest that Indian cotton was the most responsive to the increase in prices caused by the Civil War, while Brazilian cotton was the least responsive. The elasticity of substitution between inputs is estimated by comparing the response of the price of each variety relative to the U.S. cotton price to the change in relative transport costs caused by the war, while taking into account the price elasticity of each input. The shock provides the necessary instrument for the change in transport costs. I find that the elasticity of substitution between Indian and U.S. cotton was well above 2. The elasticities of substitution for Egyptian and Brazilian cotton, relative to U.S. cotton, were above 1, and were likely above 2 as well. Given these, the predictions of the model are that (1) we should observe technical change that augments these alternative varieties, (2), the long-run price of these inputs should increase, and (3), these changes should be more pronounced for Indian cotton than for the other varieties.

To analyze the direction of technical change, I begin by focusing on patenting patterns in the BPO spinning and weaving technology categories during 1861-1865. These data show that there was an increase in spinning technology patents of 18-21% during this period. In contrast, no
increase is observed in weaving technologies during this period, nor do non-textile technology patents show any significant change. Within the spinning technology category, I can focus specifically on patents related to cotton, wool, linen, and silk. I observe a sharp increase in cotton-related spinning technology patents during the 1861-1865 period, on the order of 67-75%. Spinning patents related to wool, linen, or silk show no similar increase. Data on patents of spinning technology subcategories allow me to look at these changes in even more detail. I find that the increase in cotton textile technology patents was driven by an increase in machines used to prepare the raw cotton for spinning, specifically gins, openers and scutchers – exactly the machines which were needed in order to take advantage of the Indian cotton. All of these results continue to hold when I focus only on high-quality patents using any of my three measures of patent quality. Thus, the patent data provide evidence of technical change which was directed towards those technologies which augment Indian cotton, an input which became relatively more abundant. This is consistent with the predictions of the theory, for Indian cotton, given my elasticity estimates. However, it runs counter to the hypothesis of Hicks (1932) that technical change would be directed towards the input that becomes more scarce.

Next, I consider the impact on input prices. While the price increased substantially for every cotton variety during the Civil War, my focus will be on what happened to the price of Indian, Brazilian, or Egyptian cotton relative to the U.S. cotton price during and after the war. In the absence of directed technical change, the model predicts that the relative price of each of these varieties should have fallen as it became relatively more abundant. On the other hand, the technical change directed towards augmenting Indian cotton may offset this, by increasing the demand for that variety. Graphing the relative price of Indian to U.S. cotton shows a decrease in the first two years of the Civil War, followed by a rapid rebound starting in 1863, around the time when the new technologies were becoming available. In contrast, for Brazilian or Egyptian cotton varieties, where I do not observe evidence of directed technical change, the price relative to U.S. cotton fell at the beginning of the war and remained low throughout the period in which these varieties remained abundant relative to the pre-war period. Econometric results strengthen this finding. I
find that there was a significant decrease in the relative price of Brazilian and Egyptian cotton following the onset of the war. In contrast, the relative price of Indian cotton increased during this period. These results are consistent with the existence of strong induced bias acting on the relative price of Indian cotton, which is the prediction of the model for this variety given my elasticity estimates.

One finding of this study is that there is evidence of technical change directed towards machines which augmented Indian cotton, while no evidence of a similar effect was found for technologies tailored to using Brazilian or Egyptian cotton. A new prediction offered by my theory is that, when inputs are sufficiently substitutable, directed technical change will be more strongly in favor of inputs which are more responsive to price signals. Given that there is some evidence that the elasticity of input supply was higher for Indian cotton than the other varieties, this prediction may help explain why innovation was focused primarily on Indian cotton, though I will discuss other factors which were also likely to have contributed.

This project is closely related to existing work by Newell et al. (1999), Popp (2002), and Aghion et al. (2010). All three of these studies consider the impact of increasing energy prices on energy-efficient innovation using data from the last several decades. All of them find evidence that higher energy prices led to more energy-efficient innovation. Thus, these results provide some evidence that a change in input prices can influence the direction of innovation. Three main features differentiate this project from these existing studies. First, in this study I am able to observe the prices and quantities for multiple input varieties, as well as the technologies related to them. This is important because it allows me to estimate the elasticity of substitution between

---

4Linn (2008) is another recent paper using a similar approach. An alternative approach is taken by Blum (2010) who uses cross-country trade data in an effort to find evidence of directed technical change at a macro level. In particular, he finds that changes in relative factor endowments are negatively correlated with relative factor prices, and that this correlation is larger for factor prices in the long run, which he interprets as evidence of technical change biased toward the factor which became relatively scarce. This approach is potentially complementary to microeconomic studies such as the one presented here. However, standing alone it is difficult to be sure that the changes he observes are truly due to directed technical change rather than other factors, since technology is not observed, and controlling for other potential explanations is difficult in a cross-country context.

5Newell et al. (1999) track energy-efficient innovation by looking at the specifications of air conditioners and water heaters from Sears catalogs, while Popp (2002) and Aghion et al. (2010) track innovation in certain patent categories.
inputs and relate it to technical change and input prices, which is crucial to testing the predictions of the theory. This type of data has not been available in previous studies. Second, this chapter takes advantage of a large exogenous shock in order to estimate the causal impact of a change in the cost of providing inputs on innovation. In contrast, Newell et al. (1999) and Popp (2002) use input prices as their key explanatory variable, something which my results suggest may be endogenous. Aghion et al. (2010) use changes in gas tax rates to predict changes in gas prices. While this may be a cleaner approach than using input price, there is still some concern that changes in gas tax rates may be correlated with other changes that affect innovation in energy-efficient technologies.\footnote{Two plausible ways that this may occur are if changes in gasoline tax rates are generated by changes in government, which are also accompanied by other changes in energy policies or government investment levels, or if energy-related legislation incorporates gas tax changes together with other regulations that affect other energy policies which influence innovation rates in energy-related technologies.} Third, my study considers the impact of directed technical change on input prices, which previous studies have not addressed. It is also helpful that there is virtually no government intervention in my setting, which is a potential concern when looking at energy-related technologies in the past few decades.

While this study is focused on the impact of changes in input supplies on innovation, there are complementary studies that consider the influence of demand factors or competition. On the demand side, Acemoglu & Linn (2004) consider the impact of demand fluctuations on innovation rates in the pharmaceutical industry and find that shifts in demand can be an important driver of new product development. For competition, Bloom et al. (2009) use several measures of technical change, including patents and R&D expenditures, to show that an increase in competition from Chinese producers led European firms to upgrade their technology.

The next section presents the theoretical framework. Section 1.3 details the empirical setting and presents my elasticity estimates. The patent data are described in Section 1.4. I analyze the impact on innovation patterns using raw patent data in Section 1.5. Section 1.6 provides additional evidence using indicators of patent quality. The impact on input prices is analyzed in Section 1.7. Section 1.8 concludes.
1.2 Theory

This section presents an extension of Acemoglu (2002) which introduces inputs that respond to price signals in place of inelastically supplied factors. This extension is important both because it moves the model closer to the empirical setting and because it allows me to derive new theoretical predictions. The model is also modified to include more than two goods and factors, a somewhat trivial extension that allows it to better fit the empirical setting. The exposition of the model is brief; for more details, readers are referred to Acemoglu (2002).

1.2.1 Model setup

The model can be thought of as representing a small sector which is embedded in a larger economy. I.e., it is a partial equilibrium model. It is also dynamic, with continuous time. There are \( i \in 1,...M \) goods in the sector and consumers consume an index of these goods given by:

\[
Y = \left( \sum_{i=1}^{M} Y_{i}^{\epsilon-1} \right)^{\frac{\epsilon}{\epsilon-1}},
\]

where \( \epsilon \in (0, +\infty) \) is the elasticity of substitution between goods. In the empirical setting, we can think of these goods as being various types of cotton textiles which are produced using different varieties of raw cotton. The production function for each good is given by:

\[
Y_{i} = \left( \frac{1}{1-\beta} \right) \left( \int_{0}^{N_{i}} x_{i}(k)^{1-\beta} dk \right) Z_{i}^{\beta},
\]

where \( N_{i} \) is the number of machine types available for producing good \( i \), \( x_{i}(k) \) is the quantity of each machine of type \( k \) specialized for the production of good \( i \), \( Z_{i} \) is an input good which is used to produce good \( i \), and \( \beta \in (0,1) \). In the empirical setting, we can think of these inputs as being varieties of cotton from different locations, such as the U.S., India, or Brazil. Each good is produced using only one input type, so inputs are also indexed by \( i \in 1,...M \).

In previous theories, inputs were exogenous parameters. The main difference in this model
is that inputs are endogenous and will respond to price changes according to the upward-sloping supply function:

\[ Z_i = A_i \tilde{c}_i, \quad (1.2) \]

where \( A_i \) is a fixed factor, \( \tilde{c}_i \) is the price received by producers of input \( i \), and \( \alpha \in [0, +\infty) \) is the price elasticity of input supply. While numerous micro-founded models could underlie this expression, it is sufficient for my purposes that the supply curve is upward sloping. Each input is produced abroad and supplied subject to a transport cost such that for every dollar paid by the input user, only \( 1/\tau_i < 1 \) dollars are received by the input supplier. These transport costs will be the primary focus of my comparative statics. Using these, the relative quantity of any two inputs can be written as:

\[ \frac{Z_i}{Z_j} = \left( \frac{A_i}{A_j} \right) \left( \frac{c_i}{c_j} \right)^{\alpha} \left( \frac{\tau_i}{\tau_j} \right)^{-\alpha}, \quad (1.3) \]

where \( c_i \) is the price paid by the input using firms for input \( i \).

### 1.2.2 Equilibrium

The goods market is competitive, which implies that the ratio of final goods prices for any two goods must satisfy:

\[ \frac{p_i}{p_j} = \left( \frac{Y_i}{Y_j} \right)^{-\frac{1}{\beta}}. \quad (1.4) \]

The price of the index over final goods is chosen as the numeraire. Solving the optimization problem for final goods producing firms delivers the following expressions for machine demands and input prices:

\[ x_i(k) = \left( \frac{p_i}{\chi_i(k)} \right)^{1/\beta} Z_i, \quad (1.5) \]
\[ c_i = \left( \frac{1}{1 - \beta} \right) p_i \left( \int_0^{N_i} x_i(k)^{1 - \beta} dk \right) Z_i^{\beta - 1}, \]  

where \( \chi_i(k) \) is the price for a unit of machines of type \( k \) used for producing good \( i \).

New machines are produced by technology monopolists who face a constant marginal cost \( \psi \). The profit for a monopolist producing a machine type \( k \) used for good \( i \) is \( \pi_i(k) = (\chi_i(k) - \psi)x_i(k) \). Because the demand curve for machines is isoelastic, the optimal price charged by these monopolists is \( \chi_i(k) = \psi/(1 - \beta) \), and to simplify things, I apply the normalization \( \psi = (1 - \beta) \), which implies that equilibrium machine prices are \( \chi_i(k) = 1 \) for all \( i \) and \( k \).\(^7\) It is now possible to use the machine price and machine demand expressions to rewrite production as a function of only the goods price, the level of technology and the input quantity:

\[ Y_i = \left( \frac{1}{1 - \beta} \right) p_i^{1 - \beta} N_i Z_i. \]

This can be substituted into (1.4) in order to rewrite relative prices in terms of relative technologies and relative input quantities:

\[ \frac{p_i}{p_j} = \left( \frac{N_i Z_i}{N_j Z_j} \right)^{-\beta \sigma}. \]  

(1.7)

In this equation, \( \sigma = \varepsilon - (\varepsilon - 1)(1 - \beta) \) is the derived elasticity of substitution between inputs. Using Equation 1.6, relative input prices can be expressed as a function of relative input quantities and relative technology levels:

\[ \frac{c_i}{c_j} = \left( \frac{N_i}{N_j} \right)^{\sigma - 1} \left( \frac{Z_i}{Z_j} \right)^{-\frac{1}{\sigma}}. \]  

(1.8)

This is an important equation for the analysis. It tells us that, holding technologies constant, a change in the relative input quantities will have a negative impact on the relative input prices. However, when the elasticity of substitution between inputs is sufficiently high (\( \sigma > 1 \)), this may

\(^7\)Note that, because machine producers are small, the pricing and production decisions of individual producers will not affect \( Z_i \), so machine producers will not consider the impact of their collective pricing choices on the quantity of input \( i \).
be offset by an increase in the relative technology level. Using (1.3) and (1.8), the relative input quantities can be rewritten as:

\[
\frac{Z_i}{Z_j} = \left( \frac{A_i}{A_j} \right)^{\frac{\sigma}{\sigma+\alpha}} \left( \frac{N_i}{N_j} \right)^{\frac{\alpha(\sigma-1)}{\sigma+\alpha}} \left( \frac{\tau_i}{\tau_j} \right)^{\frac{-\alpha}{\sigma+\alpha}}.
\] (1.9)

Substituting this into 1.8, relative input prices can be expressed as:

\[
\frac{c_i}{c_j} = \left( \frac{A_i}{A_j} \right)^{-\frac{1}{\sigma+\alpha}} \left( \frac{N_i}{N_j} \right)^{-\frac{1}{\sigma+\alpha}} \left( \frac{\tau_i}{\tau_j} \right)^{\frac{\alpha}{\sigma+\alpha}}.
\] (1.10)

### 1.2.3 Incentives for innovation

Given that machine prices equal one, and using the machine demands given by (1.5), instantaneous profits for a technology monopolist firm making machines for industry \(i\) are:

\[
\pi_i = \beta p_i^{1/\beta} Z_i.
\] (1.11)

Machines depreciate fully after use, but machine designs remain available indefinitely. Thus, technology monopolists care about their discounted value of future profits, rather than instantaneous profits, when deciding whether to develop new machines. The net present discounted value can be written using a standard dynamic programming equation as:

\[
rV_i - \dot{V}_i = \pi_i,
\] (1.12)

where \(r\) is the interest rate, \(V\) is the present discounted value of future profits, and \(\pi\) is the flow of profits. Focusing on the steady state, where \(\dot{V} = 0\) and the interest rate is constant, the discounted value of developing a machine of type \(i\) is:

\[
V_i = \frac{\beta p_i^{1/\beta} Z_i}{r}.
\] (1.13)

Taking the ratio of these values for machines used in producing two different final goods, and
substituting for prices, the relative value of machines in each sector can be expressed as a function of relative input quantities:

\[
\frac{V_i}{V_j} = \left(\frac{N_i}{N_j}\right)^{-\frac{\sigma}{\sigma+1}} \left(\frac{Z_i}{Z_j}\right)^{-\frac{\sigma-1}{\sigma+1}}.
\]  

(1.14)

This equation shows that, when the elasticity of substitution between factors is high (\(\sigma > 1\)), an increase in the quantity of input \(i\) will increase the incentive for new inventions in sector \(i\). The opposite will occur when \(\sigma < 1\). Using 1.9 to substitute for the relative input quantities, we have:

\[
\frac{V_i}{V_j} = \left(\frac{A_i}{A_j}\right)^{\frac{\sigma-1}{\sigma+\alpha}} \left(\frac{N_i}{N_j}\right)^{\frac{\alpha(\sigma+1)}{\sigma+\alpha}} \left(\frac{\tau_i}{\tau_j}\right)^{-\frac{\alpha(\sigma-1)}{\sigma+\alpha}}.
\]  

(1.15)

Note that the sign of the exponent on the \(N_i/N_j\) term in this equation may be either positive or negative when \(\alpha > 0\). This matters. If the exponent is positive, then a stable balanced growth path will not exist, in the sense that, starting from an equilibrium in which \(V_i/V_j = 1\), even a slight increase in \(N_i/N_j\) will cause an increase in \(V_i/V_j\). In this case the model will veer into an equilibrium in which one input becomes economically insignificant. The intuition behind this result is that, if goods are highly substitutable and input quantities respond strongly to prices, then it will be optimal to focus all innovation on one good and produce that good very efficiently while neglecting the others. Since I am interested in situations in which both inputs are used in equilibrium, I will rule this possibility out by requiring that \(\alpha \sigma - 2\alpha - 1 < 0\).

1.2.4 Supply of innovation

The production function for new machine varieties is \(\dot{N}_i = \eta R_i\) where \(\eta\) is a parameter which determines the cost of innovation for machines and \(R_i\) represents expenditure on innovation in these machines.\(^8\) On the balanced growth path, \(p_i\) is constant for all \(i\) and \(\dot{N}_i = \dot{N}_j\) for all \(i\) and \(j\). This implies that \(\dot{V}_i = 0\) and \(V_i = V_j\). Using (1.13), this implies that \(\pi_i = \pi_j\). Using (1.7) and

\(^8\)Nothing changes if \(\eta\) is allowed to be different for machines used to produce different goods, as in Acemoglu (2002). However, since I am not interested in running comparative statics on this parameter I have assumed that the cost is the same to simplify things.
I obtain:

\[
\frac{N_i}{N_j} = \left( \frac{Z_i}{Z_j} \right)^{\sigma^{-1}}.
\]  

(1.16)

Substituting for the relative input supply using (1.9) yields:

\[
\frac{N_i}{N_j} = \left( \frac{A_i}{A_j} \right)^{\frac{\sigma-1}{2\omega+1-\omega\alpha}} \left( \frac{\tau_i}{\tau_j} \right)^{-\frac{\omega(\sigma-1)}{2\omega+1-\omega\alpha}}.
\]  

(1.17)

We can now express relative input prices as a function only of the quantity of fixed factors and the trade costs, using 1.10 and 1.17:

\[
\frac{c_i}{c_j} = \left( \frac{A_i}{A_j} \right)^{-\frac{\sigma-2}{2\omega+1-\omega\alpha}} \left( \frac{\tau_i}{\tau_j} \right)^{-\frac{\omega(\sigma-2)}{2\omega+1-\omega\alpha}}.
\]  

(1.18)

Finally, relative input quantities can be expressed as a function only of the quantity of fixed factors and trade costs, using 1.9 and 1.17:

\[
\frac{Z_i}{Z_j} = \left( \frac{A_i}{A_j} \right)^{\frac{1}{2\omega+1-\omega\alpha}} \left( \frac{\tau_i}{\tau_j} \right)^{-\frac{\alpha}{2\omega+1-\omega\alpha}}.
\]  

(1.19)

1.2.5 Predictions

The U.S. Civil War, which included a blockade of Southern ports, can be thought of as a large exogenous increase in the cost of transporting American cotton to England. Thus, I am interested in the predicted impact of a change in trade costs on innovation and input prices. The first prediction to come from the model, which is somewhat trivial, is that a change in transport costs will affect input quantities. This prediction is derived from Equation 1.19.

Impact on input supplies: An increase in \( \tau_i/\tau_j \) will cause a decrease in \( Z_i/Z_j \).

Next, I obtain predictions regarding the impact of a change in relative transport costs on the direction of innovation and on input prices. These are similar to the main results of Acemoglu (2002). Using Equations 1.17 and 1.18, it can be shown that:
**Direction of innovation:** When \( \sigma > 1 \), an increase in \( \tau_i/\tau_j \) will decrease \( N_i/N_j \). I.e., innovation will be directed towards augmenting the input which becomes relatively more abundant.

**Strong induced bias:** When \( \sigma > 2 \), an increase in \( \tau_i/\tau_j \) will decrease \( c_i/c_j \). I.e., directed technical change will lead to an increase in the relative price of the input which becomes relatively more abundant.

The model also allows me to make one additional prediction regarding the role played by the elasticity of input supply, \( \alpha \), in determining the strength of the effects described above. This prediction follows from Equations 1.17, 1.18 and 1.19.

**Strength of directed technical change:** The higher is the elasticity of input supply \( \alpha \), the larger will be the impact of a change in \( \tau_i/\tau_j \) on \( Z_i/Z_j \), the more strongly technical change will be directed towards the more abundant factor when \( \sigma > 1 \), and the stronger will be the positive influence on the relative price of the more abundant input when \( \sigma > 2 \).

The remaining sections will take these predictions to the data. Before doing so, it is worth noting one disconnect between the theory and the empirical analysis. In the model, there is only one elasticity of substitution, and only one elasticity of input supply, and the predictions above are generated by varying these elasticities. In reality, both of these elasticities are likely to vary across inputs, so to more closely match the empirical setting, I would need a model with varying elasticities. However, this significantly increases the complexity of the model beyond what is currently available. In this chapter, I will work with the model as described, and test the prediction by comparing across industries with varying elasticities, while recognizing that these are not the same thing. Specifying a more general model is left to future work.

### 1.3 Empirical setting

During the time period considered in this study, 1855-1883, cotton textiles were Britain’s largest export and raw cotton was Britain’s largest import.\(^9\) In 1860, cotton textile exports were valued at £52 million, dwarfing the next largest export categories (wool textile exports at £15.7

---

\(^9\)Of course, this was not the case during the U.S. Civil War.
million and iron and steel at £13.6 million).\(^{10}\)

### 1.3.1 The Cotton Textile Production Process

It is helpful to have some understanding of the cotton textile production process, and the technologies involved, before proceeding. There are four stages in the cotton textile production process: Preparation, Spinning, Weaving, and Finishing. Preparation involved separating the cotton fibers from the seeds, using gins, opening the cotton fibers using openers, cleaning the cotton by removing leaves, dirt, and other matter using scutchers and carding machines, and preparing it for shipping using balers and packers.\(^{11}\) The earliest step, ginning, generally took place in the cotton producing region, while later steps, such as opening and carding, generally took place in manufacturing centers, such as Britain or the Northern U.S. In the spinning stage, the prepared raw cotton was spun into yarn. This stage took place in manufacturing centers and was generally done using large powered spinning machines. The yarn was then made into fabric, through weaving, the third stage of the production process. Once woven, the fabric could undergo the final stage in the production process, finishing, which included bleaching, dying, or printing.

All of these production stages relied heavily on machinery which was supplied by Britain’s large and innovative textile machinery sector. Table 1.1 shows that the two main textile technology categories, Spinning and Weaving, were among the top ten patent technology categories, out of 146 total categories, based on the number of patents filed from 1855-1883. Spinning and Weaving made up 6% and 5%, respectively, of all British patents during this period, a time at which Britain was a world technology leader.

\(^{10}\)Data from Mitchell & Deane (1962).

\(^{11}\)Definitions of these and other textile-related terms are available in Appendix A.1.
Top ten technology categories, by patent count, out of the 146 total British Patent Office technology categories. “Spinning” includes machinery used in the preparatory and spinning stages of production. “Weaving” includes machinery used in the weaving and finishing stages.

While cotton was the largest textile industry in Britain, textile industries based on wool, linen, and silk were also of significant size. The technology and other inputs used by these industries was generally similar to that used by the cotton textile industry, with the greatest technology differences being in the early stages of production. For example, modified versions of spinning and weaving machinery that were originally developed for the cotton textile industry were also used to spin other fabrics.

### 1.3.2 The impact of the U.S. Civil War

The British cotton textile industry was entirely dependent on imported raw cotton, as growing cotton in Britain was infeasible. At the beginning of the study period, the cotton textile industry was heavily dependent on cotton growers in the U.S. South, as is evident in Figure 1.1. After the beginning of the U.S. Civil War in April of 1861 the North almost immediately declared a naval blockade of Southern ports. While initially ineffective, the blockade became increasingly disruptive to Southern commerce, including the export of raw cotton, as the war continued and the Union Navy expanded. While other suppliers, particularly India, but also Egypt and Brazil, attempted to increase output, they were not able to increase their production rapidly enough to replace the flows from the U.S. Figure 1.2 shows that there was a significant drop in British

---

12This is also illustrated on the map in Appendix A.2.
domestic cotton consumption from 1861-1865, a good indicator of production in the industry.\footnote{It is reasonable to think of the amount of cotton required for a given amount of cotton textiles as being largely fixed, though, of course, small savings could be made. The reduction in production also led to massive unemployment in the cotton textile districts, resulting in the “Lancashire Cotton Famine”. Brady (1963) argues that in fact the drop in production was driven by an oversupply of cotton textile goods on the market in 1860-1861, rather than a drop in the availability of inputs. His argument is based on the fact that the ratio of cotton stocks to imports remained high during the war. However, when one considers the size of the reduction in imports and the drawdown in stocks over the 1861-1865 period, rather than comparing ratios, it is clear that his argument cannot be correct.}

Figure 1.1: British cotton imports 1815-1910

\begin{center}
\includegraphics[width=0.5\textwidth]{cotton_imports.png}
\end{center}

Data from Mitchell & Deane (1962).

Figure 1.2: British raw cotton consumption 1815-1910

\begin{center}
\includegraphics[width=0.5\textwidth]{cotton_consumption.png}
\end{center}

Data from Mitchell & Deane (1962).
Figures 1.3 and 1.4 show, respectively, the impact on the level of imports from each major supplier, and the share of total imports from the U.S., India, and other suppliers.\(^{14}\) It is clear that the shock caused a sharp drop in imports from the U.S. and an increase in imports from other suppliers, particularly India. While imports from the U.S. dropped sharply during the war, significant supplies remained on the market, allowing me to obtain reliable price data for U.S. cotton throughout the shock period.\(^{15}\) Figure 1.5 describes the impact on prices of several cotton varieties on the Liverpool market, which essentially served as the world market for cotton at this time.\(^{16}\) We can see that all prices increased significantly during the 1861-1865 period. The price for Indian cotton was always below the price for U.S. cotton, while the prices for the high-quality Brazilian and Egyptian varieties were generally above the U.S. cotton price. This hints at the differences between these distinct cotton varieties, which are described in the next subsection.

Figure 1.3: British cotton imports by supplier 1850-1880

\(^{14}\)Note that the import data shown in Figures 1.1 and 1.3 come from two different sources.

\(^{15}\)Imports from the U.S. never drop below 70,000 bales per year. For comparison, there were only 100,000 bales of Brazilian cotton imports in 1861.

\(^{16}\)These data, which were collected from original sources for the purposes of this study, are discussed in more detail in Section 1.7.
Figure 1.4: Share of imports by supplier 1850-1880

Data from Ellison (1886).

Figure 1.5: Raw cotton prices on the Liverpool market for key varieties 1852-1875

Quarterly price data from *The Economist*. Upland Middling is the benchmark U.S. cotton variety. Surat is the benchmark Indian cotton variety. Pernambuco is the benchmark Brazilian cotton. Egyptian Fair is the benchmark cotton variety from that location.

Three important points regarding the timing of the shock are visible in these figures. First, the shock was surprising. There was no run-up in price or changes in import levels in anticipation of the outbreak of hostilities. Second, the war caused large changes during the 1861-1865 period.
Third, following the end of the war, conditions began returning to their original equilibrium. The overall level of imports and production rebounded almost immediately, but the re-adjustment of relative input supplies took time. Imports of American cotton remained low through 1870, while imports of Indian, Brazilian, and Egyptian cotton remained high through the mid 1870’s. This is most likely due to delays in rebuilding the transport infrastructure in the Southern U.S.

Another feature of this shock is that it was largely transmitted through the cotton textile industry, rather than being a broad-based economic shock. Once raw cotton imports are removed, total British imports do not appear to be affected during the shock period. Similarly, once textile exports are excluded, British manufacturing exports also fail to show any large effect from the shock. Other main textile industries, based on wool, linen/flax, or silk inputs, showed no negative effects of the shock. If anything, these sectors benefited from the reduced competition from cotton textiles.

### 1.3.3 Differences between U.S. and Indian cotton varieties

This section presents a detailed look at the differences between U.S. and Indian cotton. Understanding these differences is necessary in order to identify technologies which were needed specifically for using low-quality Indian cotton. The raw cotton supplied by the U.S. and India at the time of this study came from biologically distinct varieties. The U.S. varieties were widely considered to be superior and a significant effort had been made in the 1840s and 1850s to introduce them into India, ultimately without widespread success. Thus, the cotton available from India in the 1860s was inferior to U.S. cotton in several important ways, a fact which is reflected in the gap between the price per pound of Indian and U.S. cotton shown in Figure 1.5.

One difference between these varieties was that Indian cotton was more difficult to prepare for

---

17 See Figure A.6 in Appendix A.6.
18 Graphs showing exports in these other sectors are available in Appendix A.6.
19 An article in *The Economist* from February 2, 1861 has this to say about Indian cotton: “Moreover, the native cotton [of India] is the produce of a distinct plant from the American description; and the American seed, though repeatedly introduced and carefully cultivated, has never fairly taken root there, probably from not being so well adapted for the soil and climate as the indigenous article. The Indian cotton, therefore, though it admits of being used largely as a substitute for the American, is decidedly inferior in quality, as well as different.”
spinning. In particular, it was difficult to remove the seeds from the Indian cotton using the cotton gins which were available. This was a result of the unusually small size of the Indian cotton seeds, as well as their strong bond to the cotton plant.\textsuperscript{20} The primary machine used to remove seeds in India was the Churka, a very simple and inexpensive machine. The main alternative, prior to 1860, was the saw gin, which had been developed for processing American cotton. Illustrations of both machines are available in Appendix A.3. However, American saw gins tended to cut up the Indian cotton fibers, reducing their length, and therefore their usefulness. In addition, the saw gins were much more complicated and expensive. For these reasons the saw gin proved ill suited for India. In addition to the difficulty in removing seeds, Indian cotton fibers were also more difficult to open, a process which was done using openers.

The U.S. also had a better developed cotton growing and processing industry than India, which influenced the cleanliness of the cotton. Indian cotton had a difficult journey from the interior to the ports, and passed through the hands of multiple middle-men, who habitually added dirt, salt water, or other substances in order to increase the weight of the cotton.\textsuperscript{21} As a result, the Indian cotton required more cleaning than American cotton, a process that was done using gins, scutchers, and carding machines.

Indian and U.S. cotton also differed in their fiber length. Most of the raw cotton coming from the U.S. was of a medium-length-fiber variety, which was easier to spin than the short-fiber cotton supplied by India.\textsuperscript{22} The fact that Indian cotton was shorter likely compounded the difficulties involved in ginning, since using a gin could significantly shorten the fiber length.\textsuperscript{23}

The difficulty that British producers faced in using Indian cotton is reflected in the share of cotton wasted in the production process, plotted in Figure 1.6. This graph shows that there was

---

\textsuperscript{20}In his \textit{Hand-book the the Cotton Cultivation in the Madras Presidency}, Wheeler (1862) (p. 111) writes, “In fact the seeds of the Indian Cotton are so small, that if the grates of the gin are placed close enough together to prevent the seed from passing through, the saws bring the Cotton so much in contact with the bars, as to cut it to a degree that much injures the staple. Accordingly, Mr. Finnie considered that the American gin was only suited to the American Cotton; that the two must go hand in hand; and where the American cotton failed the gin would prove useless.”

\textsuperscript{21}See, e.g., the description in Wheeler (1862) (p. 125-129) and Mackay (1853).

\textsuperscript{22}Appendix A.4 shows a comparison of fiber lengths from several of the varieties of cotton available to British producers. The Indian varieties, whether from Bengal, Madras, or Surat, are shorter than all other varieties.

\textsuperscript{23}This is illustrated in Appendix A.5 below, which shows the difference between the length of fiber obtained after hand-cleaning and mechanically ginning using a sample of Brazilian cotton.
a sharp increase in cotton waste corresponding to the switch to Indian cotton in 1862. This is particularly surprising given that the high price of raw cotton at this time must have induced producers to take measures to limit such waste. The slow reduction in the waste level after 1862 indicates the impact of improvements in the ability of textile manufacturers to use Indian cotton more efficiently.

Figure 1.6: Share of waste in total raw cotton input 1860-1868

Data from Forwood (1870). These values are calculated by taking the weight of cotton consumed and subtracting the weight of yarn produced, to obtain the weight wasted in the production process.

Another indicator of the differences between U.S. and Indian cotton can be found in the patent descriptions themselves. Though most patents provide only a simple description of the mechanisms involved, a few also mention the motivation behind the new technology. One example is given in Figure 1.7, which describes a patent filed in Britain in 1862 which was specifically designed to open the more tightly-compressed East Indian cotton. This patent was classified in the spinning technology category and the “Openers & Scutchers, etc.” subcategory, and also has “cotton” in the patent title, leading it to be identified as a cotton-related patent.

Qualitative evidence from historians and contemporary observers suggests that the differences between Indian and U.S. cotton was an influential factor during the 1861-1865 period. For example, the historian D.A. Farnie, in his authoritative history of the British cotton textile industry in the 19th century, emphasized the technological changes that using Indian cotton required British producers to undertake. Contemporary observers, such as Ellison (1886) also remarked on the

---

24Farnie (1979) (p. 152-153) writes, “The shortage of American cotton compelled employers to re-equip their
improvements in the quality and usefulness of Indian cotton that took place during this period.\textsuperscript{25}

Figure 1.7: An Example: Patent No. 2162 from 1862

\begin{center}
\includegraphics[width=\textwidth]{fig1_7.png}
\end{center}


\subsection*{1.3.4 Differences between India, Brazil, and Egypt as cotton suppliers}

Part of the analysis will rely on comparisons between cotton supplies from India, Brazil, and Egypt, so it is important to understand the differences between these locations and their potential as sources of cotton supplies. India was Britain’s second largest supplier, making up around 17\% of imports in 1860. Brazil and Egypt, each representing around 3\% of British cotton imports in 1860. Brazil and Egypt, each representing around 3\% of British cotton imports in

\begin{itemize}
\item mills in order to spin Surat [Indian cotton], and especially to improve their preparatory processes. The process of opening the tightly packed raw material become wholly automated through the use of the Crighton Opener, invented in 1861, as was the subsequent process of scutching through the application of the ingenious piano-feed regulator developed in 1862...The reorganization of the preparatory processes entailed such an extensive investment of capital that it amounted almost to the creation of a new industry...Those innovations gave a great stimulus to the textile engineering industry and consolidated the technical supremacy of the Lancashire cotton industry in the world.”
\end{itemize}

\textsuperscript{25}In his book, \textit{The Cotton Trade in Great Britain}, Ellison writes, “The high prices caused by the cotton famine, however, gave an impetus to the culture [of cotton] in India which it would not otherwise have obtained, and thereby secured to Europe a permanent increase in supply. Moreover, the quality of the cotton has been so materially improved by the introduction of better methods of handling the crop, that “Surats” are no longer despised as they were up to within a few years ago.”
1860, were the third and fourth largest suppliers of cotton after the U.S. and India.\textsuperscript{26} There were, however, important differences between these locations and India as cotton suppliers. There is evidence that India had a greater capacity to increase cotton exports than either Brazil or Egypt. The most clear reason was that India was initially a much larger supplier. India had an enormous supply of land suitable for cotton production and abundant labor. There was also a large domestic market for cotton in India, which could be redirected to export as prices increased. In contrast, supplies from Egypt were constrained by land availability, while in Brazil the problem was a shortage of labor.\textsuperscript{27} As a result, many contemporary observers considered that only India, of the three, had the potential to significantly offset a loss of U.S. supplies.\textsuperscript{28}

In addition to this qualitative evidence, I can also compare the ability of each cotton-producing location to respond to a price increase by calculating the elasticity of input supply. The elasticities of input supply for each location can be estimated by taking the logarithm of Equation 1.2 and expressing it as the regression specification, where I allow the elasticity to vary for each cotton supplier:

\[
\log(Z_{it}) = a_i + \alpha_i \log(c_{it}) + \varepsilon_{it}. \tag{1.20}
\]

The elasticity of input supply for each variety can be estimated by comparing imports to Britain of each variety to the Liverpool market price for the variety, using the specification above, where I am assuming that there were no significant changes in the cost of transporting these alternative varieties. These estimates, presented in Table 1.2, are generated by pooling all data and interacting the explanatory variables in each regression with an indicator variable for each cotton variety. Column 1 presents naive regressions of cotton imports on cotton prices. Of course, these results are likely to be biased due to reverse causality, so the remaining columns introduce various in-

\textsuperscript{26}Data from Forwood (1870).

\textsuperscript{27}For Brazil the labor shortage was a result of the end of the Atlantic slave trade in 1850 and competition from the highly profitable sugar industry. See Mann (1860). For Egypt, see Owen (1969).

\textsuperscript{28}In a July 19, 1862 article titled “How is Cotton to be Got?”, The Economist states, “And at the outset, and to clear our ground, we may observe that India and America are practically the only two quarters which need occupy our attention...Nor can the quantity furnished to us regularly from Brazil or Egypt be much increased, either immediately or ultimately, for reasons we have more than once explained.”.
struments for prices. In column 2, U.S. exports to Britain are used as an instrument for the price of alternative cotton varieties. This will be a strong instrument in all cases, and it may satisfy the exclusion restrictions, since contemporary sources indicate that U.S. supplies were not influenced by the availability of other varieties. Columns 3-6 introduce alternative instruments constructed using the Civil War shock which are described in the table notes. In all specifications, the instruments are strong, as shown by the first-stage regression results available in Appendix A.8. As expected, price tends to have a positive relationship to input supply. The estimated elasticity of input supply is higher for Indian cotton in nearly all approaches, and lowest for Brazilian cotton, with Egyptian falling in between. In some case, the difference between the elasticity of supply for Indian and Brazilian cotton is marginally significant.
Table 1.2: Estimates of the elasticities of input supply

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indian price</td>
<td>0.971***</td>
<td>0.649*</td>
<td>0.568</td>
<td>1.067***</td>
<td>0.990***</td>
<td>0.984***</td>
</tr>
<tr>
<td>x India ind.</td>
<td>(0.271)</td>
<td>(0.348)</td>
<td>(0.420)</td>
<td>(0.310)</td>
<td>(0.299)</td>
<td>(0.291)</td>
</tr>
<tr>
<td>Brazilian price</td>
<td>0.314</td>
<td>-0.212</td>
<td>-0.515</td>
<td>0.523</td>
<td>0.370</td>
<td>0.260</td>
</tr>
<tr>
<td>x Brazil ind.</td>
<td>(0.326)</td>
<td>(0.396)</td>
<td>(0.465)</td>
<td>(0.376)</td>
<td>(0.370)</td>
<td>(0.360)</td>
</tr>
<tr>
<td>Egyptian price</td>
<td>0.656**</td>
<td>0.463</td>
<td>0.393</td>
<td>0.969**</td>
<td>1.006***</td>
<td>0.878**</td>
</tr>
<tr>
<td>x Egypt ind.</td>
<td>(0.323)</td>
<td>(0.414)</td>
<td>(0.495)</td>
<td>(0.380)</td>
<td>(0.369)</td>
<td>(0.344)</td>
</tr>
<tr>
<td>Observations</td>
<td>72</td>
<td>72</td>
<td>72</td>
<td>72</td>
<td>71</td>
<td>70</td>
</tr>
</tbody>
</table>

**Differences**

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Indian-Brazilian</td>
<td>0.657</td>
<td>0.861</td>
<td>1.083*</td>
<td>0.544</td>
<td>0.620</td>
<td>0.724</td>
</tr>
<tr>
<td>Chi Sq. test</td>
<td>2.40</td>
<td>2.67</td>
<td>2.98</td>
<td>1.24</td>
<td>1.69</td>
<td>2.58</td>
</tr>
<tr>
<td>Indian-Egyptian</td>
<td>0.315</td>
<td>0.186</td>
<td>0.175</td>
<td>0.098</td>
<td>-0.016</td>
<td>0.106</td>
</tr>
<tr>
<td>Chi Sq. test</td>
<td>0.56</td>
<td>0.12</td>
<td>0.07</td>
<td>0.04</td>
<td>0.00</td>
<td>0.10</td>
</tr>
</tbody>
</table>

**Instruments**

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>U.S. Imports</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shock period</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-onset</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lag shock (t-1)</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lag shock (t-2)</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Estimates are generated by regressing the level of imports on price. Annual data covering 1852-1875. The data are pooled and explanatory variables and instruments are interacted with variety-specific indicator variables. In column 2, U.S. imports are used as an instrument for the cotton price. The shock period instrument used in columns 3-6 is an indicator variable for Q2 1861 - Q1 1865. The post-onset instrument is an indicator variable for the entire period starting in Q2 1861. The lag shock (t-1) instrument is the shock period indicator variable lagged one year. The lag shock (t-2) instrument is the shock period indicator variable lagged two years.

Technological bottlenecks also appear to have been more important for Indian cotton. This was because the high-quality Brazilian and Egyptian cotton varieties were more similar to the American cotton, and so could be handled using the machinery which was already available in 1860. For example, a number of successful ginning mills existed in Egypt prior to 1861 and in Egypt cotton was purchased by Alexandria agents who traveled to the farm districts, which ensured that the cotton arrived in a cleaner condition.  

29See, Owen (1969), who writes, “Then, unlike the situation in India, where cotton changed hands so many times between cultivator and exporter that it was impossible to fix responsibility for dirty or adulterated lint, the Egyptian
ginned using the available saw gins than Indian cotton.\textsuperscript{30}

The importance of these differences is that they rendered Brazilian and Egyptian cotton a less fertile area for technological innovation. This helps explain why, as we will see, innovation during the 1861-1865 period favored Indian cotton rather than these alternative varieties.

\subsection*{1.3.5 Estimating elasticities of substitution}

One of the key parameters in the model is the elasticity of substitution between inputs. This parameter determines the direction of innovation and whether we should expect strong induced bias. In this subsection, I estimate these elasticities. The approach will be to use the estimated impact of the change in transport costs caused by the shock on relative input prices.\textsuperscript{31}

To obtain reliable elasticity estimates, I need a measure of the change in the transport cost for U.S. cotton caused by the Civil War. To measure this, I gathered additional data showing the cotton price in the New Orleans market from a local business periodical, the \textit{New Orleans Price Current}. The data run from the third quarter of 1858 to the second quarter of 1862. New Orleans was the primary cotton exporting city in the southern U.S., and from the initiation of the blockade until the second quarter of 1862, it was within the blockaded zone. It was then captured by the Union. By comparing the price of cotton in New Orleans, prior to its capture, to that in Liverpool, it is possible to obtain a measure of the change in transport costs that took place during the early part of the war.\textsuperscript{32}

Figure 1.8 shows the index of transport costs constructed by comparing the New Orleans and Liverpool cotton prices.\textsuperscript{33} After the onset of the war in the second quarter of 1861, the price of system, whereby the Alexandria merchants sent their agents to buy cotton at the main Delta collection points and made payment according to grade, allowed some control to be exercised over quality.”

\textsuperscript{30}Stein (1957) states that Brazilian producers adopted American-style saw gins to clean their cotton.

\textsuperscript{31}Alternatively, I could look at the impact on relative input quantities, but since I have a relatively small series of transport cost data and the input quantity data is only available annually, this approach would significantly reduce the number of observations available and thus the precision of the estimates.

\textsuperscript{32}I am assuming that there was no significant change in the transport costs from other locations during this period. The price data presented in Figure 1.18 suggest that, at least for India, this was true. There is no qualitative data suggesting otherwise for India or other locations.

\textsuperscript{33}The transport cost index is constructed by assigning the average difference during the years 1858-1859 a value of one.
U.S. cotton in Liverpool increased sharply, while that in New Orleans fell, consistent with the impact we would expect given the increase in the transport costs between these locations due to the blockade.

Figure 1.8: Transportation cost index between New Orleans and Liverpool

The transportation cost index is constructed using the difference in price for Middling Orleans cotton in New Orleans and Liverpool, with the average difference in 1858-1859 set equal to one. The New Orleans price has been adjusted for inflation using the price of Stirling notes in New Orleans in each month.

In order to calculate the elasticity of substitution, I need to estimate the impact of the change in relative transport costs on the relative prices. The estimating equation is:

$$\log\left(\frac{c_{it}}{c_{jt}}\right) = \beta_0 + \beta_1 \log\left(\frac{\tau_{jt}}{\tau_{it}}\right) + \epsilon_{ijt},$$

and regressions are run separately for each alternative cotton variety (Indian, Brazilian, Egyptian) relative to U.S. cotton. Estimation results are presented in Table 1.3 below. Columns 1, 3, and 5 present naive regressions of the price of cotton from each location relative to the price of U.S. cotton on the transport cost index. In columns 2, 4, and 6, I instrument for the transportation cost index using an indicator for the Civil War period and this indicator interacted with a time-trend. The latter variable is included as an instrument to allow for the increasing effectiveness of the Union blockade during the early part of the war. Quarterly indicator variables are included in
all specifications. First-stage results for the IV regressions, available in Appendix A.9, show that these are strong instruments.

Table 1.3: Impact of transport costs on relative cotton prices ($\beta_1$)

<table>
<thead>
<tr>
<th></th>
<th>DV: India/US price</th>
<th>DV: Brazil/US price</th>
<th>DV: Egypt/US price</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS (1)</td>
<td>IV (2)</td>
<td>OLS (3)</td>
</tr>
<tr>
<td>Log tran. costs</td>
<td>-0.0589***</td>
<td>-0.0572***</td>
<td>-0.0834***</td>
</tr>
<tr>
<td>Qtly ind.</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Obs.</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
</tbody>
</table>

Autocorrelation robust standard errors calculated using the Newey-West method with a lag length of 2 shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Quarterly data covering the third quarter of 1858 to the second quarter of 1862. The instruments are an indicator for the Civil War period (Q2 1861-Q2 1862) and that indicator interacted with a time trend that takes the value of 1 in Q2 1861, 2 in Q3 1861, etc.

Because these estimates are from early in the war, it is less likely that a significant amount of biased technical change took place during this period which influenced relative input prices. If I assume that the relative technology level remained fixed during this period, then I can relate the estimation specification to Equation 1.10 from the model. In this case, $N_i/N_j$ is assumed to be constant and therefore included in $\beta_0$ and $\beta_1 = -\alpha/(\sigma + \alpha)$. Combining my preferred estimates of $\beta_1$ obtained in the IV regressions above, and my preferred elasticity of input supply estimates from column 6 of Table 1.2, I can then calculate estimates for the elasticity of substitution using the formula $\sigma = -(\alpha/\beta_1) - \alpha$. The results are shown in the top panel of Table 1.4. For all three cotton varieties, the estimated elasticities of substitution with U.S. cotton are greater than two. The table also includes a bounding exercise conducted using the estimated 95% confidence interval values for $\alpha$ and $\beta_1$. For Indian cotton, even the lower bound estimate is well above 2.

However, we may be concerned that assuming no biased technical change during this period is too strong, and that some directed technical change did take place very rapidly during the 1861-1862 period. If this innovation influenced market prices then it would generate an upward bias in the elasticity estimates in Table 1.4. In order to deal with this concern I estimate lower bound elasticities by assuming that biased technical change took place instantaneously, so that in any
quarter the technologies achieved their steady-state levels. In this case, the estimation specification corresponds to Equation 1.18 and $\beta_1 = \alpha(\sigma - 2)/(2\alpha + 1 - \alpha \sigma)$. Using this approach, I calculate estimates of the elasticity of substitution which are shown in the bottom panel of Table 1.4. Even under this extreme assumption, the estimated elasticity values are well above 1 and, in the case of Indian cotton, very close to 2.

Table 1.4: Estimated elasticity of substitution for each input variety relative to U.S. cotton

<table>
<thead>
<tr>
<th></th>
<th>Assuming no directed technical change</th>
<th>Assuming instantaneous tech. change to SS level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Indian cotton</td>
<td>Brazilian cotton</td>
</tr>
<tr>
<td>Bound exercise</td>
<td>16.29</td>
<td>2.74</td>
</tr>
<tr>
<td>(low, high)</td>
<td>(4.02, 72.4)</td>
<td>(-3.42, 15.6)</td>
</tr>
</tbody>
</table>

Elasticities are calculated using the estimates of $\alpha$ from column 6 of Table 1.2 and the IV estimates of $\beta_1$ from Table 1.3. The estimates assuming no directed technical change use the formula $\sigma = -\alpha/\beta_1 - \alpha$. The bounds exercise presents lower and upper values using the 95% confidence interval estimates of $\alpha$ and $\beta_1$. The estimates assuming instantaneous technical change are calculated using the formula $\beta_1 = \alpha(\sigma - 2)/(2\alpha + 1 - \alpha \sigma)$.

To summarize, the estimates in this section suggest that the elasticity of substitution between Indian and U.S. cotton is above 2. This is a region in which the model predicts that technical change will be directed towards Indian cotton and the strong induced-bias hypothesis will hold. For Brazilian and Egyptian cotton varieties, the estimates are not as high, though they continue to suggest that the elasticities for these varieties are above 1, and I cannot rule out that these elasticities are also above 2. However, the results also suggest that the elasticities of input supply may be lower for these varieties, which would lead us to expect less directed technical change focused on these varieties.
1.4 Patent data

The primary data used to measure innovation in this study come from British patent records. While imperfect, patent data is the best available quantifiable measure of technological advance during this period. Modern patent data has been widely used in recent studies of innovation, building on seminal work by Schmookler (1966), Scherer (1982), Griliches (1984), and Jaffe et al. (1993). Hall et al. (2001) provide a helpful review of the advantages of using patent data, including that (1) patents contain highly detailed information, (2) there are a large number of patents available to study, and (3) patents are provided on a voluntary basis under a clearly defined set of incentives. This study is able to take advantage of thousands of patents and will draw heavily on the detailed information available in the patent descriptions. While British patent laws changed in 1852 and 1883, they were stable during the period of this study.

One disadvantage of using patent data is that it will not capture all types of innovation. Evidence from Moser (2010) shows that a significant fraction of new inventions went unpatented during the period I study. However, her results also suggest that, among all categories, inventions of manufacturing machinery – the primary focus of this study – were the most likely to be patented. The incentive to patent appears to have been particularly strong for textile machinery, which was relatively easy to reverse-engineer. Thus, this concern appears to be less important in the context studied here. A second concern is that patent counts may not reflect the underlying quality of the new inventions, which can vary widely. This concern is addressed in Section 1.6.

Much of the data used in this study was collected for the purpose of this project from around 1,500 pages of printed British patent records. To begin, I constructed a database covering all of the patents granted in Britain between 1855 and 1883, 118,863 in all. These data include both granted patents and those which received provisional protection but where a patent was not ultimately granted.

---

34Moser (2010) writes, “In the 19th century, secrecy appears to have been extremely effective in protecting chemical innovations, but relatively ineffective in manufacturing machinery. Innovations in manufacturing machinery were easy to reverse-engineer. New England mechanics, for example, were able to copy British innovations in textile machinery after they had been able to observe them in Lancashire mills.”

35These data include both granted patents and those which received provisional protection but where a patent was not ultimately granted.
identify each patent, as well as the name of the inventor.

The novel contribution of this data set is that each patent is classified into one or more of 146 technology categories by the British Patent Office (BPO). These classifications allow me to identify the type of technology underlying each patent. The purpose of this categorization was to aid inventors in identifying previously patented technologies in order to determine whether an invention was in fact new. In order to get a sense of these technology categories, Table 1.1 presents the top ten categories by number of patents. My focus will primarily be on the BPO spinning and weaving categories. The spinning category includes technologies related to the preparation of raw cotton, such as cotton gins and carding machines, machines used in the spinning process, such as mules, yarn types, and other related technologies. The weaving category includes technologies such as looms, types of fabrics, and fabric treatments.

These data are supplemented with information from the *A Cradle of Invention* database, which has been used in previous research (e.g., Brunt *et al.* (2008)). This database provides the titles of the patents, which are not available in the patent data I collected. Patent titles can be used to generate more detailed classifications of the technology represented by each patent. Consistent patent titles are available from 1853-1870, after which there was a clear structural change in the naming conventions, with much less detail included in the patent titles available in the data. This database also provides information on the month of the patent application, allowing analysis at the sub-annual level.

Conveniently, the dates given in the data represent the date of the patent application, rather than the date at which the patent was ultimately granted. Thus, the application dates allow me
to identify patents at the earliest stage of the patenting process.

The patent titles allow me to search for keywords in order to identify patents representing particular types of technologies\(^{39}\) In particular, I undertake keyword searches of these titles to identify patents related to the main textile inputs: cotton, wool, linen/flax, and silk. Some summary statistics for these data are provided in Table 1.5. We can see that the majority of those patents listing one of these keywords are also classified into the BPO spinning technology category, while a few are listed in the weaving category, and some others fall into categories other than spinning and weaving. As a quality check, keyword searches were also used to identify those patents with “spinning” or “weaving” in the title. Most patents with spinning in the title are listed in the BPO spinning category, while most of those mentioning weaving are classified in the BPO weaving category. This suggests that the keyword search approach is reliable, though more restrictive, than the BPO categories.

Table 1.5: Summary statistics from patent title keyword searches, 1855-1870

<table>
<thead>
<tr>
<th>Title search term:</th>
<th>Total patents</th>
<th>Number in BPO</th>
<th>Share of BPO Spinning</th>
<th>Share of BPO Weaving</th>
<th>Number in BPO Spinning</th>
<th>Share of BPO Weaving</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cotton</td>
<td>1,230</td>
<td>892</td>
<td>73%</td>
<td>29%</td>
<td>61</td>
<td>5%</td>
</tr>
<tr>
<td>Wool</td>
<td>998</td>
<td>651</td>
<td>65%</td>
<td>21%</td>
<td>57</td>
<td>6%</td>
</tr>
<tr>
<td>Linen</td>
<td>518</td>
<td>397</td>
<td>77%</td>
<td>13%</td>
<td>21</td>
<td>4%</td>
</tr>
<tr>
<td>Silk</td>
<td>392</td>
<td>279</td>
<td>71%</td>
<td>9%</td>
<td>36</td>
<td>9%</td>
</tr>
<tr>
<td>Spinning</td>
<td>976</td>
<td>935</td>
<td>96%</td>
<td>30%</td>
<td>25</td>
<td>3%</td>
</tr>
<tr>
<td>Weaving</td>
<td>1,245</td>
<td>42</td>
<td>3%</td>
<td>1%</td>
<td>1,200</td>
<td>96%</td>
</tr>
</tbody>
</table>

Patents are identified by searching for each title search term, e.g., “cotton”, in the patent titles.

Within each BPO technology category, patents may also be listed in various technology subcategories. For example, within the BPO spinning technology category, it is possible to identify patents falling into subcategories such as “Gins”, “Mules and Twiners”, “Carding Machines”, etc. For this study, I collected data for a number of the more common spinning subcategories. These

---

British patenting system during this period see Van Dulken (1999).

\(^{39}\)This technique has been used for these data previously by Brunt et al. (2008). More extensive patent descriptions are not available in digital form.
data, which are discussed in more detail in Section 1.5.2, allow me to track changes at an even more detailed level.

1.5 Impact on innovation

In this section, I consider whether the direction of innovation in cotton textile technologies changed during the 1861-1865 period in response to the increase in the transport cost of American cotton caused by the U.S. Civil War. To begin, I look at the broad BPO technology categories related to textile production: spinning and weaving. I then consider the impact on technologies related specifically to the primary textile inputs: cotton, wool, linen/flax, and silk. Finally, I focus on specific types of machinery within the BPO spinning technology category using data that classify patents into detailed BPO technology subcategories.

1.5.1 Impact on broad textile technology categories

Our first glimpse of the patent data is presented in Figure 1.9. The top panels graph patent counts for the BPO spinning and weaving technology categories. The left-hand panel contains annual data from 1855-1883 while the right-hand panel shows quarterly data, smoothed using a four-quarter moving average, from 1855-1873. Recall that the month of the patent application is missing for some years after 1873, so consistent quarterly data is available only for 1855-1873. The bottom two panels show similar data for all BPO technology categories except spinning and weaving.

The most striking feature of these graphs is that the onset of the U.S. Civil War corresponded to a sharp increase in patents of spinning-related technologies, and that this unusually high level of spinning patents was sustained throughout most of the 1861-1865 period. This high level of spinning-related technology patents only falls toward the end of the Civil War, which also corresponds with a financial crises that struck England and was particularly severe in the northern cotton manufacturing districts. The drop in innovation is also consistent with decreasing returns
to R&D. No similar increase appears in weaving technologies, nor do other technology categories show similar effects.

Using the results of the keyword searches, it is also possible to focus on those patents mentioning cotton, wool, linen/flax, or silk in the title. Figure 1.10 shows the count of patents mentioning each of the four key textile inputs. Each graph also shows the number of patents mentioning each of these inputs which are also listed in the BPO spinning category. The message from these graphs is clear; there was a sharp and sustained increase in cotton-textile related patents during the 1861-1865 period, while no similar increase appears for the other main textile inputs.

Next, some simple regression procedures are applied. OLS regressions are run based on:
Figure 1.10: Count of patents with titles mentioning main textile inputs, 1853-1870

Quarterly data smoothed using four-quarter moving averages.
\[ P_t = \beta_0 + \beta_1 S_t + \beta_2 T_t + \beta_3 X_t + Q_t \Gamma + \epsilon_t, \]

where \( P_t \) represents the log count of patents in period \( t \), \( S_t \) represents an indicator variable for the shock period (Q2 1861-Q1 1865), \( T_t \) represents a time-trend, \( X_t \) is the log count of non-textile patents in period \( t \), and \( Q_t \) is a set of quarterly indicator variables.\(^{40}\)

Regressions are run separately for each category of patents, e.g., BPO spinning patents. Results are presented in Table 1.6. These results suggest that the shock caused a significant increase in the number of patents listed in the BPO spinning category (column 1), patents mentioning cotton in the title (column 3), as well as BPO spinning patents that also mentioned cotton in the title (column 4). These effects are large; the shock is associated with an increase of spinning technology patents of around 21%, and an increase in cotton-related patents of 53-75%. No similar increase is observed for BPO weaving patents (column 2), or patents mentioning wool, linen/flax, or silk in the title (columns 5-7).\(^{41}\) These results can be checked using a number of alternative procedures and data series, such as using annual data which is available for a longer period. A number of such robustness checks are available in Appendix A.10. The results are robust to all of these alternatives.

One potential worry with these results is that they could be driven, in part, by patenting of technologies which already existed prior the the onset of the war but were not worth patenting until the war altered market conditions. This concern is highlighted in Figure 1.11, which plots the number of cotton-related patents at monthly frequency instead of the quarterly frequency shown in Figure 1.10. While the additional variability observed at the monthly level makes it more difficult

\(^{40}\)Non-textile patents include all patents except those classified in the BPO spinning or BPO weaving categories or those which mention “cotton”, “wool”, “linen”, or “silk” in the patent title. Alternatively, one may be tempted to use the ratio of Indian to U.S. cotton in British cotton imports or the relative price of these two varieties as the primary explanatory variable, which would be closer to the approach used in Newell et al. (1999) and Popp (2002). However, it is reasonable to expect that either of these variables will be influenced by the introduction of new technology, so including them would introduce reverse causality issues into the model and reduce my ability to make causal statements about the influence of market conditions on innovation. Nevertheless, results calculated using this approach are available in Table A.4.

\(^{41}\)The negative coefficient on weaving and the other textile types may suggest substitution of inventive inputs away from these industries and towards cotton textile technologies. However, because much of this drop occurs in 1865, it may also be a symptom of a financial crises that occurred at that time.
Table 1.6: Effect of the shock (1861-1865) on log patent count by patent type

<table>
<thead>
<tr>
<th></th>
<th>BPO Spinning (1)</th>
<th>BPO Weaving (2)</th>
<th>Cotton-related (3)</th>
<th>Cotton-Spinning (4)</th>
<th>Wool-related (5)</th>
<th>Linen-related (6)</th>
<th>Silk-related (7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shock Indicator</td>
<td>0.189***</td>
<td>-0.0208</td>
<td>0.427***</td>
<td>0.560***</td>
<td>-0.0797</td>
<td>-0.112</td>
<td>-0.0614</td>
</tr>
<tr>
<td></td>
<td>(0.0592)</td>
<td>(0.0602)</td>
<td>(0.0448)</td>
<td>(0.0819)</td>
<td>(0.0636)</td>
<td>(0.173)</td>
<td>(0.159)</td>
</tr>
<tr>
<td>Log Total</td>
<td>0.591***</td>
<td>0.258</td>
<td>0.360</td>
<td>0.483</td>
<td>-0.0965</td>
<td>0.785</td>
<td>0.541</td>
</tr>
<tr>
<td>Non-textile Pats.</td>
<td>(0.187)</td>
<td>(0.288)</td>
<td>(0.395)</td>
<td>(0.440)</td>
<td>(0.328)</td>
<td>(0.767)</td>
<td>(0.655)</td>
</tr>
<tr>
<td>Time trend</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Quarter indicators</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Constant</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>72</td>
<td>72</td>
<td>68</td>
<td>60</td>
<td>68</td>
<td>68</td>
<td>68</td>
</tr>
</tbody>
</table>

Newey-West standard errors with a lag length of 4 are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Quarterly data. The data in columns 1-2 cover 1855-1873. The data in columns 3 and 5-7 cover 1853-1870. The data for column 4 covers 1855-1870.

to pick out the underlying trends in patent counts, one feature which is clear is the enormous spike in cotton-related patents in April 1861, the month in which the U.S. Civil War broke out. Given the timing of these patent applications, they could not reflect new innovation undertaken in response to the change in conditions. Rather, it seems clear that this spike was due to an increase in patent applications of already-existing technologies which were previously not worth patenting, but which became worthwhile to patent given the change in conditions caused by the onset of the war.42

Because I am interested in the impact of the cotton shortage on new innovation, it is important to ensure that the results are not being driven by patents of existing ideas. Glancing at Figures 1.9 and 1.10 suggests that it is unlikely that the results are driven by patenting of pre-existing ideas, since there was a sustained high level of new patents throughout most of the 1861-1865 period. One way to address this concern econometrically is to omit data from 1861 from the analysis, which focuses attention on the increase in innovation in the later war years. Results calculated while dropping 1861 are shown in Table 1.7. These results continue to show that BPO spinning category patents increased, by about 18%, in the 1862-1865 period. Cotton-related patents also

---

42 Recall that filing a patent application required only a shorter provisional specification describing the invention, and allowed inventors six months in which to complete the design and file a full patent specification, while still granting them priority based on the application date.
increased, by 52-67%. Further robustness checks are available in Appendix A.10.

Figure 1.11: Monthly patent counts of cotton-related patents

![Figure 1.11: Monthly patent counts of cotton-related patents](image)

Table 1.7: Effect of the shock on log patent count dropping 1861 data

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Shock Indicator</td>
<td>0.165***</td>
<td>0.00289</td>
<td>0.416***</td>
<td>0.512***</td>
<td>-0.0938</td>
<td>-0.151</td>
<td>-0.0752</td>
</tr>
<tr>
<td>Log Total</td>
<td>0.565**</td>
<td>0.300</td>
<td>0.441</td>
<td>0.488</td>
<td>-0.102</td>
<td>0.811</td>
<td>0.522</td>
</tr>
<tr>
<td>Non-textile Pats.</td>
<td>(0.267)</td>
<td>(0.271)</td>
<td>(0.404)</td>
<td>(0.509)</td>
<td>(0.547)</td>
<td>(0.667)</td>
<td>(0.867)</td>
</tr>
<tr>
<td>Time trend</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Quarter indicators</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Constant</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>68</td>
<td>68</td>
<td>64</td>
<td>56</td>
<td>64</td>
<td>64</td>
<td>64</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Quarterly data. The data in columns 1-2 cover 1855-1873. The data in columns 3 and 5-7 cover 1853-1870. The data for column 4 cover 1855-1870. Note that it is not possible to calculate Newey-West standard errors when data from 1861 is dropped.

It may be surprising to think of innovation responding to a change in market conditions within 1-2 years, and it is difficult to estimate the amount of time needed to create a new cotton textile invention. However, some useful information is provided by Lakwete (2003) in her authoritative history, *Inventing the Cotton Gin*. This account details numerous instances in which inven-
tors produced new innovations or patentable improvements on existing inventions within a 1-3 year period. Among these inventors is Eli Whitney, who had invented, patented, and introduced commercially, his famous cotton gin, within two years of first setting foot on a Southern cotton plantation. This information suggests that, at least in the case of gins, it is reasonable to expect innovation to respond rapidly to changing conditions.

The results above suggest that the onset of the U.S. Civil War and the commensurate shortage of raw cotton led to a rapid increase in cotton-related and spinning technology patents, and that this increase was sustained throughout most of the shortage period. While the initial increase was certainly due to the patenting of existing ideas, the sustained high level throughout most of the shock period indicates a burst of innovative activity during these years. Section 1.6 will provide further evidence suggesting that the increase observed in 1861-1865 was not driven by the patenting of pre-existing ideas.

1.5.2 Impact on specific types of spinning technologies

We have already seen that the increase in innovation was concentrated in the BPO spinning technology category, which includes machines used in both the preparatory and spinning stages of production. Now I look in more detail at which spinning technologies were most affected by the shock. In order to do so, I collected additional data which classify patents into BPO technology subcategories within the BPO spinning category. Data were gathered on patents fitting into several of the larger technology subcategories, which are described in in Table 1.8. They can be divided into those related to the preparatory, spinning, or finishing stages of the spinning process. The data are available from 1855-1876. Of the subcategories shown, the most important for adapting to the use of Indian cotton were gins, openers/scutchers, and to a lesser extent, carding machines.

---

43 Two other good examples are McCarthy’s roller gin and Whipple’s cylinder gin, which were both invented in response to the panic of 1837 and patented in the U.S. in 1840.

44 Note that “finishing stages of the spinning process” denotes operations occurring as part of the spinning stage of production, such as bleaching or dyeing yarn, as opposed to the finishing stage of the textile production process as a whole, which involved bleaching, dyeing, etc. of woven fabrics.

45 See Section 1.3.
Table 1.8: Spinning technology subcategories by production stage

<table>
<thead>
<tr>
<th>Preparatory stage</th>
<th>Patents</th>
<th>Spinning stage</th>
<th>Patents</th>
<th>Finishing stage</th>
<th>Patents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gins</td>
<td>122</td>
<td>Mules and twiners</td>
<td>446</td>
<td>Finishing</td>
<td>332</td>
</tr>
<tr>
<td>Openers/scutchers</td>
<td>331</td>
<td>Rollers for spinning</td>
<td>462</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carding</td>
<td>696</td>
<td>Bearings for spinning</td>
<td>242</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combing</td>
<td>354</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Patent counts for BPO spinning technology subcategories, 1855-1876.

I begin the analysis by graphing the count of patents, by year, for each technology subcategory, in Figure 1.12. These graphs show an increase in patents in technology subcategories related to the preparation of raw cotton, particularly gins and openers/scutchers, during the 1861-1865 period. In contrast, technologies related to later stages of the spinning process do not show similar effects. It is particularly interesting that there does not appear to be an increase in combing machine patents. These machines were not used in producing every type of yarn, but when they were used they were the largest source of cotton waste. The next most wasteful stage was carding, which shows only modest evidence of an increase. If innovation had been focused primarily on economizing on waste cotton, I would expect to see an increase in patents of combing and carding technologies. The fact that we do not suggests that innovation was not directed towards economizing on cotton in general.

Next, I analyze these patterns using a regression approach which is similar to those used previously. The main difference is that, because subcategory patents are more sparse, the data include more zero periods. To deal with this I focus on the impact of the shock on the count of patents in each subcategory, rather than the log count. Results are presented in Table 1.9. The first set of results, from OLS regressions, show the change in the count of patents in each subcategory during the shock period. These results may be problematic because of the number of zeros included in the data, so the second set of results are from Negative Binomial regressions.

46See Thornley (1912). Combing machines act somewhat like a standard comb. Their purpose was to remove short fibers and arrange the remaining longer fibers so that they are all pointing the same direction. Combing was generally done when producing higher quality fabrics. While combing machines were used to produce cotton, they were more common in the preparation of wool (worsted) textiles.
Figure 1.12: Patent counts in subcategories of the BPO Spinning technology category

- Gins
- Openers, scutchers, etc.
- Carding machines
- Combing machines
- Mules
- Rollers, etc.
- Bearings, etc.
- Finishing
which are better able to deal with these zero entries.\textsuperscript{47}

These results suggest that the shock period was associated with an increase in innovations in gin, opener/scutcher, and carding technologies. However, there is no strong evidence of an increase in those technologies used in the spinning or finishing stages of yarn production. The effects are quite large in some cases. The shock increased gin patents by over 3 patents per quarter, which is large given that there were on average just under one gin patent per quarter outside of the shock period. For openers/scutchers the results indicate an increase of around 2 patents per quarter compared to an average level of just over 3 patents per quarter outside of the shock period. For carding machines, the shock period saw an increase of just under 3 patents per quarter compared to an average level of just over 7 patents per quarter outside of the shock period.

Table 1.9: Effect on patents in spinning technology subcategories

<table>
<thead>
<tr>
<th>Preparatory</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Gins</td>
<td>Openers/scutchers</td>
<td>Carding machines</td>
<td>Combing machines</td>
<td>Mules/twiners</td>
<td>Rollers, etc.</td>
<td>Bearings</td>
<td>Finishing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. OLS regs. DV: Count of patents</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shock</td>
<td>3.413***</td>
<td>2.073**</td>
<td>2.932*</td>
<td>-0.270</td>
<td>1.528***</td>
<td>-1.009</td>
<td>-0.192</td>
<td>-1.157*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indicator (1.023)</td>
<td>(0.998)</td>
<td>(1.603)</td>
<td>(0.481)</td>
<td>(0.475)</td>
<td>(0.721)</td>
<td>(0.493)</td>
<td>(0.613)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obs. 88</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. OLS regs dropping 1861. DV: Count of patents</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shock</td>
<td>3.829***</td>
<td>2.113***</td>
<td>2.921***</td>
<td>-0.277</td>
<td>1.188</td>
<td>-1.442*</td>
<td>-0.589</td>
<td>-0.793</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indicator (0.478)</td>
<td>(0.671)</td>
<td>(1.042)</td>
<td>(0.616)</td>
<td>(0.730)</td>
<td>(0.861)</td>
<td>(0.466)</td>
<td>(0.599)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obs. 84</td>
<td>84</td>
<td>84</td>
<td>84</td>
<td>84</td>
<td>84</td>
<td>84</td>
<td>84</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. Neg. Binomial regs. DV: Count of patents</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shock</td>
<td>1.736***</td>
<td>0.503***</td>
<td>0.366***</td>
<td>-0.0630</td>
<td>0.298***</td>
<td>-0.207</td>
<td>-0.0447</td>
<td>-0.327**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indicator (0.249)</td>
<td>(0.146)</td>
<td>(0.120)</td>
<td>(0.147)</td>
<td>(0.129)</td>
<td>(0.159)</td>
<td>(0.188)</td>
<td>(0.163)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obs. 88</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D. Neg. Binomial regs dropping 1861. DV: Count of patents</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shock</td>
<td>1.823***</td>
<td>0.510***</td>
<td>0.510***</td>
<td>-0.0659</td>
<td>0.243*</td>
<td>-0.316*</td>
<td>-0.250</td>
<td>-0.209</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indicator (0.267)</td>
<td>(0.163)</td>
<td>(0.163)</td>
<td>(0.159)</td>
<td>(0.141)</td>
<td>(0.181)</td>
<td>(0.227)</td>
<td>(0.169)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obs. 84</td>
<td>84</td>
<td>84</td>
<td>84</td>
<td>84</td>
<td>84</td>
<td>84</td>
<td>84</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Quarterly data for 1855-1876. All regressions include the following controls: count of non-textile patents, a time trend, and a constant. All include a set of quarter indicator variables except that these are dropped from the negative binomial regression on patents of bearings excluding 1861 data in order to obtain convergence. The standard errors shown in panel A have Newey-West standard errors calculated with a lag length of 4. Note that it is not possible to calculate Newey-West standard errors when dropping 1861 from the data.

\textsuperscript{47}Negative Binomial regressions are preferred to Poisson regressions because most of the data series are characterized by overdispersion. Poisson regression results are shown in the robustness table.
It is also possible to look specifically at cotton-related patents in each of these subcategories. These data are limited to the 1855-1870 period by the availability of patent titles. Table 1.10 presents results paralleling those in Table 1.9, but considering only patents which mention “cotton” in the patent title. As before, these results show that the shock was associated with significant increases in gin, opener/scutcher, and carding machine patents, but that spinning and finishing-stage technologies do not show similar increases. Robustness checks calculated using a number of alternative estimation strategies, available in Appendix A.11, show similar results.

Table 1.10: Effect on cotton-related patents in spinning technology subcategories

<table>
<thead>
<tr>
<th></th>
<th>Preparatory</th>
<th></th>
<th>Spinning</th>
<th>Finishing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gins</td>
<td>Openers/ scutchers</td>
<td>Carding</td>
<td>Combing</td>
</tr>
<tr>
<td>A. OLS regs. DV: Count of “cotton” patents</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shock</td>
<td>3.127***</td>
<td>1.611**</td>
<td>2.000*</td>
<td>-0.142</td>
</tr>
<tr>
<td>Indicator</td>
<td>(0.918)</td>
<td>(0.672)</td>
<td>(1.016)</td>
<td>(0.237)</td>
</tr>
<tr>
<td>Obs.</td>
<td>64</td>
<td>64</td>
<td>64</td>
<td>64</td>
</tr>
<tr>
<td>B. OLS regs dropping 1861. DV: Count of “cotton” patents</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shock</td>
<td>3.594***</td>
<td>1.553***</td>
<td>1.762***</td>
<td>-0.312</td>
</tr>
<tr>
<td>Indicator</td>
<td>(0.441)</td>
<td>(0.470)</td>
<td>(0.597)</td>
<td>(0.297)</td>
</tr>
<tr>
<td>Obs.</td>
<td>60</td>
<td>60</td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td>C. Neg. Binomial regs. DV: Count of “cotton” patents</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shock</td>
<td>1.804***</td>
<td>0.676***</td>
<td>0.600***</td>
<td>-0.141</td>
</tr>
<tr>
<td>Indicator</td>
<td>(0.236)</td>
<td>(0.177)</td>
<td>(0.160)</td>
<td>(0.315)</td>
</tr>
<tr>
<td>Obs.</td>
<td>64</td>
<td>64</td>
<td>64</td>
<td>64</td>
</tr>
<tr>
<td>D. Neg. Binomial regs dropping 1861. DV: Count of “cotton” patents</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shock</td>
<td>1.921***</td>
<td>0.657***</td>
<td>0.560***</td>
<td>-0.352</td>
</tr>
<tr>
<td>Indicator</td>
<td>(0.235)</td>
<td>(0.191)</td>
<td>(0.178)</td>
<td>(0.367)</td>
</tr>
<tr>
<td>Obs.</td>
<td>60</td>
<td>60</td>
<td>60</td>
<td>60</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Quarterly data for 1855-1870. All regressions include the following controls: count of non-textile patents, a set of quarter indicator variables, a time trend, and a constant. The standard errors shown in panel A have Newey-West standard errors calculated with a lag length of 4. Note that it is not possible to calculate Newey-West standard errors when dropping 1861 from the data.

If the same procedure is applied to patents related to wool, linen/flax, or silk, no similar patterns emerge. These results are summarized in Table 1.11, which shows the estimated coefficient on the shock indicator variable for patents related to wool, linen/flax, and silk in each technology subcategory. Full regression results are available in Appendix A.12.
Table 1.11: Estimated shock impact on subcategory patents of wool, linen, and silk tech.

<table>
<thead>
<tr>
<th></th>
<th>Preparatory</th>
<th>Spinning</th>
<th>Finishing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gins/Schuters</td>
<td>Carding/Scutchers</td>
<td>Combing/Scutchers</td>
</tr>
<tr>
<td>Wool Patents</td>
<td>0.421</td>
<td>-0.211</td>
<td>-0.230</td>
</tr>
<tr>
<td></td>
<td>(0.940)</td>
<td>(0.340)</td>
<td>(0.213)</td>
</tr>
<tr>
<td>Linen Patents</td>
<td>x</td>
<td>-0.250</td>
<td>-0.0506</td>
</tr>
<tr>
<td></td>
<td>(0.595)</td>
<td>(0.452)</td>
<td>(0.393)</td>
</tr>
<tr>
<td>Silk Patents</td>
<td>0.0115</td>
<td>-0.576</td>
<td>-0.117</td>
</tr>
<tr>
<td></td>
<td>(0.517)</td>
<td>(0.631)</td>
<td>(0.357)</td>
</tr>
<tr>
<td>Obs.</td>
<td>64</td>
<td>64</td>
<td>64</td>
</tr>
</tbody>
</table>

"x" indicates that gins are not used to prepare linen/flax or silk. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Table presents estimated coefficients and standard errors for the shock period indicator variables only. Quarterly data for 1855-1870. All regressions include the following controls: count of non-textile patents, a set of quarter indicator variables, a time trend, and a constant.

The results presented in this section show that there was a sharp increase in cotton-textile-related patents during the 1861-1865 period, and that this increase was concentrated in those early-stage technologies which were most important for taking advantage of Indian cotton. While these results are based on raw patent counts, the next section shows that similar results hold when I focus only on patents of high-quality technologies.

1.6 Indicators of patent quality

This section introduces three measures of patent quality and uses them to evaluate whether the 1861-1865 period was also characterized by an increase in the number of high-quality cotton-textile-related patents.\footnote{While it may also be interesting to consider whether the \textit{share} of high quality patents changed during the 1861-1865 period, this section focuses only on whether the \textit{number} of high quality patents increased, which is ultimately what matters for economic growth.} It is well known that adjusting for quality is important when using patent data because raw patent counts mask the quality of the new technology represented by each individual patent, which may vary widely. Of particular concern is the possibility that a number of patents may represent inventions of limited usefulness. While this is unlikely to be the case given the relatively high patent fees charged in Britain at this time, it is still important to adjust
In this section, I attempt to account for three aspects of patent quality: (1) long-term viability, (2) wider applicability, and (3) initial novelty. By long-term viability, I mean the extent to which the patented invention remains economically important years after its initial introduction. This aspect will be measured using data on the payment of patent renewal fees. Wider applicability means the breadth of different locations and economic environments in which the invention is used. To measure this aspect, I consider patents by British inventors in India and the U.S. The third aspect, initial novelty, is the extent to which the invention was recognized as a significantly new technological contribution. This aspect will be measured by observing whether patents were described in a contemporary periodical focused on new inventions. While it is reasonable to expect these quality measures to be correlated, it is also possible to think of situations in which they may diverge, which is why multiple measures of patent quality are considered.

### 1.6.1 Valuing patents using renewal data

This section uses data on the payment of patent renewal fees in order to assess the long-term viability of patented inventions. Renewal fee data have been used as an indicator of patent quality in previous studies (Schankerman & Pakes (1986), Lanjouw et al. (1998)), including some using historical British patent data (Sullivan (1994), Brunt et al. (2008)). During the period covered by this study, British patents lasted for 14 years, but in order to keep them in force patent holders were required to pay renewal fees of £50 before the end of three years and an additional £100 before the end of seven years. These were substantial sums at the time and the result was that the vast majority of patents were allowed to expire before their full term. My data show that just

---

49 Khan (2005) notes that the barriers for patenting were much higher in Britain than in the United States during this period.

50 For example, an invention that fills a small technological niche may have long-term viability, but may not be broadly applicable and may also fail to arouse the interest of inventors. In contrast, an invention may be widely adopted upon introduction, but may also quickly become obsolete if further technological improvements are relatively straightforward. Finally, a novel but imperfect invention may arouse great interest among contemporary inventors and thereby generate follow-on innovations which soon render the original idea obsolete.

51 For comparison, £100 in 1860 is equivalent to £7,020 2010 pounds using a retail price index deflator, or £65,200 when deflating by average earnings (calculator available through the Measuring Worth project at www.measureingworth.com).
under 18% of patents were renewed at three years, while just over 6% were renewed at seven years. Thus, paying a renewal fee represents a substantial investment which would only have been worth it for a small set of the most successful technologies.

Renewal fee data were gathered from listings in *Mechanics’ Magazine*, a weekly periodical focusing on patents and related topics. The magazine is available from the end of 1858 to the end of 1872, so that data on renewals at year three are available for patents filed from 1856-1869 and data on renewals at year seven are available from 1853-1865. By merging the renewal data with the primary patent data set, it is possible to track renewal patterns for textile-related patents.

Table 1.12 presents regression results describing the number of textile-related patents filed during the 1861-1865 period for which the renewal fee was paid at year three. The results show that there was a significant increase in the number of spinning and cotton-related patents filed during this period which were subsequently renewed. There is no evidence of a similar increase in patents of weaving technologies or those related to wool, linen, or silk.

Table 1.12: Textile patents renewed after three years

<table>
<thead>
<tr>
<th>DV: Count of Patents in Each Technology Sub-category</th>
</tr>
</thead>
<tbody>
<tr>
<td>BPO Spinning</td>
</tr>
<tr>
<td>Shock Indicator</td>
</tr>
<tr>
<td>(0.894)</td>
</tr>
<tr>
<td>Observations</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Quarterly data for 1856-1869. All regressions include the following controls: total count of renewed patents, a time trend, a constant, and a set of quarterly indicator variables.

Table 1.13 present similar results for patents renewed at year seven. Here we do not observe any significant increases. This may suggest that the shock did not influence the number of very high quality inventions, though given the small sample size it is difficult to make any strong conclusions based on these data.
Table 1.13: Textile patents renewed after seven years

| DV: Count of Patents in Each Technology Sub-category |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| BPO Spinning    | BPO Weaving     | Cotton-related Spinning | Cotton-related Spinning | Wool-related Spinning | Linen-related Spinning | Silk-related Spinning |
| Shock Indicator | -0.592          | -0.493          | 0.488           | 0.0438          | -0.493          | 0.525          |
|                 | (0.831)         | (0.690)         | (0.741)         | (0.490)         | (0.690)         | (0.495)         |
| Observations    | 52              | 52              | 52              | 52              | 52              | 52              |

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Quarterly data for 1853-1865. All regressions include the following controls: total count of renewed patents, a time trend, a constant, and a set of quarterly indicator variables.

We can also look at whether the shock affected the number of high-quality patents in the spinning technology subcategories described in Section 1.5.2. Because these data include a number of zero entries, Negative Binomial regressions are used. Table 1.14 presents the results. These indicate that there was a significant increase in the number of gin patents filed during the shock period for which renewal fees were subsequently paid at year three. Table 1.15 presents similar results for patents which were renewed at year seven. Here we see some indications of an increase in gins and openers/scutchers patents during the 1861-1865 period which were subsequently renewed, although these data include relatively few observations, making it difficult to draw strong conclusions.

Table 1.14: Spinning subcategory patents renewed at three years

| DV: Count of patents in each technology sub-category |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| OLS regressions | Gins | Openers/scutchers | Carding machines | Combing machines | Mules/twiners etc. | Bearings | Finishing |
| Shock Indicator | 1.019*** | 0.355 | 0.0190 | 0.392 | 0.546 | -0.0929 | -0.411 | -0.347 |
| Obs.             | 56   | 56   | 56    | 56    | 56    | 56    | 56    | 56    |

Neg. Binomial regressions. DV: Count of patents in each technology sub-category

| Shock Indicator | 2.341*** | 0.369 | 0.00211 | 0.301 | 0.403* | -0.0750 | -0.961* | -0.353 |
| Obs.            | 56     | 56    | 56     | 56    | 56    | 56     | 56     | 56     |

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Quarterly data for 1856-1869. All regressions include the following controls: total count of renewed patents, a time trend, a constant, and a set of quarterly indicator variables.
The impact of the shock on the number of high-quality patents is further explored in Figure 1.13 for cotton-related patents, and patents in the gins and openers/scutchers subcategories. The figure shows the patents, plotted by application date, which were subsequently renewed. The increase in the number of patents filed during the 1861-1865 period which were renewed at year three is visible in both of these graphs. It is also clear that there was an increase in patents of gins and openers/scutchers for which the renewal fee was also paid at year seven.

An important feature of these results is that there was a high level of patents filed in years 1862-1864 which were renewed after three years, and in some cases after seven years. For most of these, the renewal fees would have been paid after the end of the Civil War, during a period in which the markets were returning toward their pre-war equilibrium levels. This suggests that, had these patents been available prior to 1861, they likely would have been worth patenting given that the initial patenting fee was only one-half or one-quarter of the renewal fees. The point is that these technologies were most likely not available prior to 1861, which suggests that there was an increase in new and valuable innovation during the 1861-1865 period.
Figure 1.13: Cotton-related and gin/opener/scutcher technology patent renewals

“At year three” indicates patents for which the renewal fee was paid in to keep the patent in force beyond year three.
“At year seven” indicates that the renewal fee was paid to keep the patent in force beyond year seven.

1.6.2 Valuing patents using foreign patent filings

This section uses patent data from India and the U.S. to assess whether the 1861-1865 period saw an increase in cotton and textile related patents which were widely applicable. This approach has been used previously by Lanjouw et al. (1998). The motivation behind this measure is that observing a British invention which was patented abroad indicates that the invention was viable in a wider range of circumstances. The U.S. and India are used both because data from these locations are available and because they represent significantly different environments in which
the technologies must operate.\textsuperscript{52} India was primarily a producer of low-quality raw cotton at this time. The U.S. was both a major producer of mostly high-quality cotton as well as an important cotton textile manufacturing center. However, patents filed during the Civil War were valid only in the North, which excluded all of the main cotton growing districts, but included most textile manufacturers.\textsuperscript{53}

I begin by analyzing Indian patent data. These data, which I gathered from original printed records, cover 1859-1879. During this period, 1,138 Indian patents were granted, of which 429 went to inventors based in Britain. Each Indian patent was manually reviewed in order to identify textile and cotton related technologies.\textsuperscript{54} Most of these patents are either for cotton gins, or for balers and packers, which were used to prepare the cotton for shipping. Table 1.16 describes how the share of patents made up of all cotton-related technologies, gins, and balers/packers, changed during the 1861-1865 period. The three left columns consider the share of these technologies in all Indian patents, while the right side looks at the share in only Indian patents by inventors based in Britain. There is evidence of a significant increase in the share of patents for gins and cotton-related technologies by British patent holders during the shock period. This is consistent with an increase in inventions in Britain which were also applicable in India.

<table>
<thead>
<tr>
<th>Share of all Indian patents</th>
<th>Share of patents by British inventors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cotton</td>
</tr>
<tr>
<td>Shock Indicator</td>
<td></td>
</tr>
<tr>
<td>(1861-1865)</td>
<td>0.0442***</td>
</tr>
<tr>
<td>(0.0145)</td>
<td>(0.00993)</td>
</tr>
<tr>
<td>Observations</td>
<td>23</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Annual data covering 1859-1879. All regressions include a constant.

\textsuperscript{52}The technologies used by U.S. textile manufacturers tended to differ somewhat from those used by British producers. A classic example is that the British generally used mules for spinning, which could spin finer thread counts and use lower quality cotton, but also required highly skilled operators, while U.S. manufacturers tended to use ring spinning technology that required higher quality cotton but could be operated by less skilled workers.

\textsuperscript{53}There was a separate Confederate Patent Office operating in the South at this time, but given the uncertainty of the war and the difficulty of communication caused by the Union blockade, it was not successful at attracting patent filings by foreigners.

\textsuperscript{54}Patents mentioning “cotton” in the title were coded as cotton patents, patents with “gin” in the title were coded as gins, etc.
The U.S. patent database covers 1857-1873 and includes 94,141 patents, of which 1,160 were held by British inventors.\footnote{These data were generously shared by Tom Nicholas.} Using the inventor name and patent title I attempted to match each of these inventions to a patent filed in Britain, in order to identify a patent family. A total of 974 U.S. patents (84\% of 1,160) were matched to British patents.

My interest is in whether there was an increase in cotton-textile-related patents in the U.S., by British inventors, corresponding to the increase observed in British patents.\footnote{For a comparison of the U.S. and British patent systems, see Khan (2005).} Because there was a reduction in the fees paid by foreign patent holders in the U.S. in 1862, my analysis must focus on the share of textile and cotton-related patents in total U.S. patents by British inventors, rather than the raw number of patents. Table 1.17 presents results for textile-related technologies. These show evidence that there was an increase in the share of cotton-related technologies in U.S. patents by British inventors during the 1861-1865 period.

Table 1.17: Share of textile patents in total U.S. patents by British inventors

<table>
<thead>
<tr>
<th></th>
<th>Spinning</th>
<th>Weaving</th>
<th>Cotton</th>
<th>Wool</th>
<th>Linen</th>
<th>Silk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shock Indicator</td>
<td>0.0317</td>
<td>-0.0313*</td>
<td>0.0338*</td>
<td>0.0191</td>
<td>-0.00484</td>
<td>0.00358</td>
</tr>
<tr>
<td>Indicator</td>
<td>(0.0357)</td>
<td>(0.0174)</td>
<td>(0.0176)</td>
<td>(0.0148)</td>
<td>(0.0191)</td>
<td>(0.00914)</td>
</tr>
<tr>
<td>Obs.</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Annual data covering 1857-1873. All regressions include a constant.

Table 1.18 applies the same exercise to spinning technology subcategories. There appears to have been an increase in the share of patents in the gins subcategory, as well as in bearings. Overall, this provides some evidence of an increase in British cotton-textile-related innovations flowing to the U.S. during the Civil War period.
Table 1.18: Spinning subcategory patents’ share of total U.S. patents by British inventors

<table>
<thead>
<tr>
<th></th>
<th>Gins</th>
<th>Openers/ scutchers</th>
<th>Carding machines</th>
<th>Combing machines</th>
<th>Mules/ twiners</th>
<th>Rollers, etc.</th>
<th>Bearings</th>
<th>Finishing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shock</td>
<td>0.0151***</td>
<td>0.0023</td>
<td>-0.0001</td>
<td>0.0069</td>
<td>-0.0053</td>
<td>-0.0171</td>
<td>0.0227***</td>
<td>-0.0081</td>
</tr>
<tr>
<td>Ind.</td>
<td>(0.0041)</td>
<td>(0.0046)</td>
<td>(0.0098)</td>
<td>(0.0050)</td>
<td>(0.0081)</td>
<td>(0.0169)</td>
<td>(0.0044)</td>
<td>(0.0079)</td>
</tr>
<tr>
<td>Obs.</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1. Annual data covering 1857-1873. All regressions include a constant.

1.6.3 Valuing patents using contemporary publications

A contemporary periodical can be used to highlight the interest or excitement generated by a new patent upon its publication. This approach has previously been used to value historical British patents by Nuvolari & Tartari (2011). This section takes advantage of data that I collected from Newton’s London Journal, a monthly publication devoted to covering new patents and other technology-related topics. This journal was published by William Newton & Sons, one of the preeminent patent agents in London. Among the Journal’s stated goals was making more easily available the information contained in patent filings, and to this end, each issue included abstracts from a selection of recently sealed (i.e., granted) patents, some of which were accompanied by detailed drawings. Though they provide little information about the criteria used to select these patents, presumably they included those patents which were deemed by the editors to be the most important inventions, or those which would be of greatest interest to the readers. Thus, inclusion of a patent abstract in the journal is treated as an indication of the initial novelty of each patent, based on the judgment of a knowledgeable contemporary opinion.

The Journal is available from January 1855 - February 1866, meaning that any patent applied for from 1855-1864 should have been a candidate for inclusion. Matching these patents to the primary patent database allows me to identify patents of textile and cotton related technologies. Because the total number of abstracts may have been limited by space constraints, the analysis

---

57It is worth noting that patent abstracts were only included after the patent had been sealed, so publication was often as long as a year after the initial patent application was filed. This means that the editor would have had some perspective from which to judge the influence of a patent before including it in the journal.
focuses on the share of published abstracts made up of cotton-textile-related technologies. The analysis is based on the date the patent was filed, rather than the publication date, so for example, I look at all patents which were filed in 1861 and then subsequently published, and analyze the share composed of textile-related patents.

Table 1.19 presents results for the main textile technology categories and input types. These results show an increase in the share of abstracts for spinning and cotton-related technologies during the 1861-1865 period, as well as a smaller increase in patents related to wool. Table 1.20 shows similar results for spinning technology subcategories. The only significant result is an increase in the share of patents for gins during the 1861-1865 period. Together these results indicate that the 1861-1865 period was characterized by an increase in the number of cotton-textile-related patents, and particularly patents of cotton gins, which contemporary observers considered to be interesting or novel contributions.

Table 1.19: Share of published abstracts composed of textile-related patents

<table>
<thead>
<tr>
<th></th>
<th>Spinning</th>
<th>Weaving</th>
<th>Cotton</th>
<th>Wool</th>
<th>Linen</th>
<th>Silk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shock Indicator</td>
<td>0.0501***</td>
<td>-0.00765</td>
<td>0.0307***</td>
<td>0.0206***</td>
<td>0.00795</td>
<td>-0.00187</td>
</tr>
<tr>
<td>(0.0112)</td>
<td>(0.0133)</td>
<td>(0.00732)</td>
<td>(0.00337)</td>
<td>(0.00654)</td>
<td>(0.00565)</td>
<td></td>
</tr>
<tr>
<td>Obs.</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 1.20: Share of published abstracts composed of patents in spinning subcategories

<table>
<thead>
<tr>
<th></th>
<th>Gins</th>
<th>Openers/ scutchers</th>
<th>Carding machines</th>
<th>Combing machines</th>
<th>Mules/ twiners</th>
<th>Rollers, etc.</th>
<th>Bearings</th>
<th>Finishing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shock Indicator</td>
<td>0.0087***</td>
<td>0.0034</td>
<td>0.0077</td>
<td>0.0006</td>
<td>0.0033</td>
<td>0.003</td>
<td>0.0055</td>
<td>-0.0019</td>
</tr>
<tr>
<td>(0.0029)</td>
<td>(0.0029)</td>
<td>(0.0059)</td>
<td>(0.0033)</td>
<td>(0.0042)</td>
<td>(0.0022)</td>
<td>(0.0034)</td>
<td>(0.0042)</td>
<td></td>
</tr>
<tr>
<td>Obs.</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

To summarize the results of this section, it appears that there was a significant increase in British patents of high-quality cotton textile technologies, and particularly early stage technolo-
gies such as gins and openers/scutchers, during the U.S. Civil War. This holds whether patent quality means long-term viability, as measured by payment of renewal fees, wider applicability, as measured by patents outside of Britain, or initial novelty, as measured by being mentioned in a contemporary periodical.

1.7 Impact on input prices

This section explores the impact of directed technical change on input prices. In particular, I am interested in whether the technical change directed towards Indian cotton that I have observed is also exerting a positive influence on the relative price of that input, as suggested by the strong induced-bias hypothesis. In order to examine this question, I will compare the patterns observed in the relative price of Indian cotton to those for Brazilian and Egyptian cotton, varieties for which I have found no evidence of directed technical change, at least to the extent that technical change was directed in favor of Indian cotton.

If the strong induced-bias hypothesis is correct for Indian cotton, then I should first observe a drop in the price of this variety relative to U.S. cotton following the onset of the war, as Indian cotton became relatively more abundant. However, as new technologies tailored to the use of Indian cotton became available, there should be a rebound in the relative price of Indian cotton. A similar initial fall should be observed for the relative prices of Brazilian and Egyptian cotton, but without the subsequent rebound.

To evaluate this hypothesis, new price data was gathered from market reports printed in *The Economist* magazine. The data are monthly and cover 1852-1875, but for ease of analysis the following charts show prices averaged by quarter. Price data were available for the following benchmark cotton varieties: Upland Middling from the U.S., Surat Fair from India, Pernambuco Fair from Brazil, and Egyptian Fair. A plot of these data is shown in Figure 1.5.

As a first step in the analysis, Figure 1.14 graphs the benchmark raw cotton price level and import quantity for American, Indian, Brazilian, and Egyptian cotton. For American cotton, it

---

58 Note that Egyptian imports are actually given by the “Mediterranean” category in Ellison (1886), but most of
appears that the price tended to be low when supplies were high and high when supplies were low. In contrast, the price of raw cotton from other locations tended to move along with supplies. This suggests that the market was being driven primarily by the availability of U.S. cotton, a point also made by many contemporary sources. All of these graphs use the same scale, so we can see that the U.S. was generally the largest import supplier, followed by India, while Brazil and Egypt were considerably smaller. It is worth noting that cotton imports from India, Brazil, and Egypt increased during the 1861-1865 period and remained high through the end of 1875. U.S. cotton, in contrast returned to roughly its original level by the end of 1875. Thus, the abundance of cotton from India, Brazil, and Egypt relative to U.S. cotton was higher from 1861 to 1875 than prior to 1861.

These imports came from Egypt.

For example, as one of the four general conclusions of his study of Madras cotton, Wheeler (1862) writes, “The demand for Indian cotton must always depend on the supply of American.” In his analysis of cotton production and trade from 1830-1860, Wright (1971) also comes to the conclusion that the U.S. South would have had substantial monopoly power in the market, if only it had been able to reduce its own supply.
Figure 1.14: Prices and import quantities for raw cotton by location

Price data gathered from *The Economist* magazine. Import data from Ellison (1886) Note that import data for Egypt comes from the “Mediterranean” category in Ellison’s data. Most of these imports came from Egypt. A price was not reported for Egyptian cotton in the first quarter of 1867 and has been interpolated.

Next, Figure 1.15 graphs the price of raw cotton from each location relative to the price of the benchmark U.S. variety. The price of Indian relative to U.S. cotton was unusually low in 1861-1862, the first two years of the war, and a period in which Indian cotton had become relatively more abundant. However, starting in 1863, there was an increase in the relative price of Indian cotton. This upward trend lasted through the early 1870s, despite the fact that the relative quantity of Indian cotton remained higher than prior to 1861. In contrast, the relative price of Brazilian to U.S. cotton fell in 1861-1862 and remained low through 1876, a period during which the relative abundance of Brazilian cotton was high. The relative price of Egyptian to U.S. cotton also fell in 1861, and, though it rallied somewhat in 1866, a year in which supplies fell sharply, it remained
low for most of the post-war period compared to the pre-war period. The patterns observed in Brazilian and Egyptian cotton prices are consistent with what the model would predict in the absence of significant biased technological progress, given the increase in the relative abundance of these varieties after 1861. In contrast, the initial decrease in the relative price of Indian cotton, followed by the increase after 1863, when a significant number of new technologies tailored to the use of Indian cotton were becoming available, is consistent with the strong induced-bias hypothesis.

Figure 1.16 facilitates comparison between movements in the relative price of Indian cotton to those observed for Brazilian and Egyptian cotton. The top panel plots the log prices of Indian and Brazilian cotton relative to U.S. cotton, with the mean value in 1852 set to one. The bottom panel does the same for Indian and Egyptian cotton. In both cases, we can see that the relative price of Indian, Brazilian, and Egyptian cotton varieties move similarly prior to 1861, and that they fall together in 1861, but these relative prices diverge after 1862, with Indian gaining relative to the others. I argue that this divergence is due to the upward pressure on the relative price of Indian cotton exerted by increasing demand caused by the availability of better machines for processing Indian cotton. It is also interesting that this difference fades in the mid-1870’s, which suggests that the influence of the inventions generated during 1861-1865 had faded a decade later.
Figure 1.15: Cotton prices relative to the benchmark U.S. variety

Indian Cotton

Brazilian Cotton

Egyptian Cotton

Price data gathered from *The Economist* magazine.
Comparing the relative prices of Indian/U.S. to Brazilian/U.S. cotton

Comparing the relative prices of Indian/U.S. to Egyptian/U.S. cotton

Price data gathered from The Economist magazine.

Constructing a statistical test of the strong induced-bias hypothesis is more difficult, because of the non-linear nature of the predictions, as well as uncertainty about the time-frame in which new technologies begin influencing the market. One approach is to include all of the period after the start of the war in a single “treatment period”. This biases the results against finding evidence of strong induced bias for Indian cotton, since the early war period, in which the relative price of Indian cotton fell due to the substitution effect, is included together with the later period. The regression specification used to test the impact on relative input prices is obtained by taking the logarithm of Equation 1.18 and converting it into the following regression specification:
\[ \log(c_i/c_j) = \beta_0 + \beta_1 \log(\tau_i/\tau_j) + \varepsilon_{ij}, \]

where an indicator of a ten-year period starting at the onset of the war will be used as a proxy for the change in relative transport costs. I limit the impact to ten years to reflect the fact that innovations which are new in 1861 may not be at the cutting edge of technology a decade later, as reflected by the low rate of patent renewals at year seven. Note that this expression has been derived by substituting out both the relative technology term and the relative input supply term, so that it represents the net effect of a change in transport costs on input prices operating through both of these channels. The top panel of Table 1.21 describes the impact of the shock on relative cotton prices for each variety starting in 1861 and lasting through 1871.\(^6\) These results show that there was a significant decrease in the relative price of Brazilian and Egyptian cotton after the beginning of the war. In contrast, the results indicate that there was a statistically significant increase in the relative price of Indian cotton for the ten years after the onset of the war. This result is particularly striking given that the estimates are likely to be biased downwards due to the inclusion of the first two years of the war, when the relative price of Indian cotton fell, in the impact period.

In order to see these patterns in more detail, Figure 1.17 graphs estimated coefficients generated by regressing the relative price of Indian cotton on individual year indicator variables for each year starting in 1861. This graph shows that the relative price of Indian cotton was significantly lower in 1861-1862. I then observe an increase in the relative price for 8 out of the next 9 years, with statistically significant increases in 5 of these years. The only year between 1863 and 1872 when the estimated coefficient is not positive is 1865, a year in which the end of the war and uncertainty about the stock of cotton remaining in the South led to massive price swings.

\(^6\)If I also include the impact extending through 1875 I continue to obtain positive coefficients for Indian cotton but they are only statistically significant in some specifications. The estimates for Brazilian and Egyptian cotton are always negative and statistically significant. If I consider only the impact starting in 1863, I observe a statistically significant positive impact on Indian cotton and a significant negative impact on Brazilian and Egyptian cotton.
Table 1.21: Impact of the shock on relative cotton prices by variety

<table>
<thead>
<tr>
<th></th>
<th>India/US price</th>
<th>India/US price</th>
<th>Brazil/US price</th>
<th>Brazil/US price</th>
<th>Egypt/US price</th>
<th>Egypt/US price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q2 1861 - Q4 1871</td>
<td>0.0498***</td>
<td>0.0524***</td>
<td>-0.103***</td>
<td>-0.0717***</td>
<td>-0.0851***</td>
<td>-0.0649***</td>
</tr>
<tr>
<td>indicator</td>
<td>(0.0186)</td>
<td>(0.0193)</td>
<td>(0.0199)</td>
<td>(0.0152)</td>
<td>(0.0252)</td>
<td>(0.0245)</td>
</tr>
<tr>
<td>Time trend</td>
<td>-6.03e-05</td>
<td>-0.000740***</td>
<td>-0.000483***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(9.90e-05)</td>
<td>(8.71e-05)</td>
<td>(0.000105)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monthly ind.</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Constant</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>288</td>
<td>288</td>
<td>288</td>
<td>288</td>
<td>288</td>
<td>288</td>
</tr>
</tbody>
</table>

Autocorrelation robust standard errors calculated using the Newey-West method with a lag length of 4 are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All regressions include a constant and a set of monthly indicator variables. Monthly data covering 1852-1875. A very small number of observations are missing in the cotton prices data because they were not reported in The Economist. These values have been interpolated.

Figure 1.17: Estimated impact on the relative price of Indian cotton by year

Estimated coefficients and 95% confidence intervals generated by regressing the relative price of Indian to U.S. cotton on individual year dummies. Regressions use monthly data covering 1852-1875 and include a set of monthly indicator variables. Confidence intervals are generated using autocorrelation robust standard errors calculated using the Newey-West method with a lag length of 4.

The results above provide evidence in favor of the strong induced-bias hypothesis operating for Indian cotton. There is no evidence that a similar effect occurred for Brazilian or Egyptian varieties. This makes sense given that I have observed technical change which was focused primarily on using Indian cotton. It also fits with the elasticity of supply estimates, which suggest that both directed technical change and strong induced-bias will be stronger for Indian cotton, due
to the stronger input supply response expected from this variety.

There is one important caveat to the analysis presented above, which is that the prices used are those on the Liverpool market, rather than farm-gate prices. Thus, they may reflect quality improvements in Indian cotton resulting from the new technologies which took place before the cotton reached the Liverpool market. This is a concern given that the cotton could have benefited from processing by improved machines on their way from India, particularly gins, since most ginning was done in the export country. We may be worried that this increased the wedge between farm-gate and market prices, so that we can’t be sure that the patterns observed in these prices would also be observed in farm-gate prices. This separates the analysis somewhat from the predictions of the theory, which pertain to prices before the input has been acted on by the new technology.

One way to deal with this is to look directly at prices in Bombay, the major Indian export market. While there is not a wealth of price data available, Atkinson (1897) does provide price indexes for three varieties of Indian cotton on the Bombay market. Figure 1.18 graphs these Bombay market prices together with the Liverpool market price, where all prices are presented in logs and normalized so that \(1861=1\).\(^{61}\) We can see that these prices are moving together, which is comforting, since cotton on the Bombay market would not have benefited from improved openers, balers, scutchers, or carding machines. This moves the analysis closer to the farm-gate price, though I cannot rule out that the new technologies generated some quality improvements in Indian cotton before it reached the Bombay market.

\(^{61}\)This is done to eliminate the need to compare in level terms, which is difficult given exchange rate fluctuations.
1.8 Conclusions

This study shows that a large exogenous change in the cost of providing inputs for the 19th century British cotton textile industry led to (1) directed technical change in favor of one input – Indian cotton – which had become more abundant and (2) input price movements consistent with the strong induced-bias hypothesis for this input. Given my elasticity estimates, these results are consistent with the predictions of the directed technical change model of Acemoglu (2002). This provides us with some confidence in the ability of this model to explain the process of technical change and how it is influenced by the cost of providing inputs to production. However, while the predictions of previous theories appear to be correct for Indian cotton, they cannot explain why no similar directed technical change was observed in favor of the Brazilian or Egyptian cotton varieties.

This study adds one new theoretical prediction, which may help explain why technical change was directed primarily towards machines which augmented Indian cotton. Incorporating inputs that respond to price signals into a model of directed technical change, I show that more elastic input supplies will act to magnify technical change that is directed towards inputs which have be-
come relatively more abundant. This prediction may help explain why I observe directed technical change in favor of Indian cotton, but no similar directed technical change in favor of cotton from Brazil and Egypt, where supplies were less responsive to price signals. Thus, this prediction can help us make more precise statements regarding the direction of induced innovation. However, it is likely that other factors also contributed to the inventors’ choice to focus on technologies that augmented Indian cotton. One likely factor was that the market share of Indian cotton was initially larger. Extending existing models to accommodate these alternatives is one direction for future research.

The results of this study lend support to the wide set of theories applying the idea of directed technical change. These include papers using directed technical change and strong induced bias to explain skill-biased technical change and trends in wage inequality (Acemoglu (1998), Kiley (1999)), papers suggesting that directed technical can play an important role in industrialization (Habakkuk (1962), Allen (2009)), and papers which consider how directed technical change may influence the impact of environmental regulations (Acemoglu et al. (2012)). While my results cannot tell us whether directed technical change is operating in any particular setting, they lend plausibility to arguments based on these mechanisms, by providing clearer evidence than was previous available that directed technical change does occur and can meaningfully influence market conditions.
Chapter 2

Is There Path Dependence in Innovation? Theory and Evidence from the 19th Century British Cotton Textile Industry

2.1 Introduction

How does the stock of knowledge in a particular type of technology affect the level of innovation in that technology type? Does current innovation “stand on the shoulders of giants”, benefiting from the lessons learned from previous research? Or have previous innovators already harvested all of the “low hanging fruit”, leading to decreasing returns along particular lines of research? The answer to these questions is central to our understanding of the process of innovation and economic growth, with important policy implications.\(^1\)

In order to address these questions, existing empirical studies, such as Popp (2002) and Aghion \textit{et al.} (2011) construct measures of the rate of innovation and the stock of knowledge for different types of technologies over time using patent data.\(^2\) They find evidence of a positive relationship

\(^1\)Jones (1999) highlights the important role that the relationship between the change in the level of knowledge (technology) and the initial stock of knowledge (represented by \(\phi\) in his framework) plays in growth models. One policy-related paper in which path dependence in innovation plays a central role is Acemoglu \textit{et al.} (2012), which investigates the implications of directed technical change for climate change policy.

\(^2\)Both of these studies are interested in energy-related technologies. Popp (2002) focuses on innovation in several categories of energy-efficient technologies, while Aghion \textit{et al.} (2011) consider “clean” (not carbon emitting) and “dirty” (carbon emitting) automotive technologies.
between the lagged stock of knowledge for a particular technology type and rate of innovation in that technology, which they interpret as evidence in favor of what I will call *path dependence in innovation*.

While these previous studies have highlighted the importance of understanding path dependence in innovation, their empirical approach suffers from several important sources of bias. This study will (1) highlight three important shortcomings of the approach used in previous empirical studies of path dependence in innovation, (2) suggest an approach which addresses these concerns, and (3) implement this approach in one empirical setting in order to generate empirical evidence that is not subject to these sources of bias.

The first concern highlighted by this study has to do with the technology level at which we look for path dependence. For example, suppose that there is an increase in the stock of knowledge for a few technologies used in the cotton textile industry and we are interested in whether this generates path dependence in innovation. We may look for a subsequent increase in innovation in all cotton textile technologies, or an increase in only the types of technologies which experienced the initial increase, or an increase in innovation by only those firms which produced the initial technological improvements. In other words, there are potentially multiple levels at which we could look for path dependence in technology. This study begins by offering a theory showing that path dependence may exist at more than one of these technology levels. If we look for path dependence at only one level, then we may miss it at another level. Moreover, this problem cannot be solved by focusing only on the most disaggregated level (e.g. firms). This motivates an empirical exercise that focuses on multiple technology levels.

The second concern with previous studies is that, in the presence of persistent shocks that increases innovation in a particular firm or technology type, they will estimate a positive relationship between the lagged stock of knowledge and the innovation rate, even when no causal relationship exists. A third concern is that any trends which are specific to the cross-sectional unit of analysis will also generate a positive relationship between the lagged stock of knowledge and the innovation rate when using the empirical approach employed by previous studies.
In order to address these concerns, I suggest an approach that takes advantage of an exogenous temporary shock that affects the stock of knowledge related to a particular set of technologies. If there is path dependence in innovation, then this increase should lead to a higher rate of innovation in these technology types in the period just after the shock. Thus, if the shock is truly temporary, then observing a higher level of innovation in the post-shock period will provide causal evidence of path dependence in innovation. Also, in order to control for trends specific to particular technology types, this approach requires panel data where the time dimension is long enough to allow me to control for time-trends which are specific to the cross-sectional units of the data. Finally, this approach should be implemented at multiple technology levels.

I implement this methodology using a large exogenous shock to the cotton textile industry in 19th century Britain. The shock was caused by the U.S. Civil War, which sharply restricted the supply of U.S. cotton to the British market from 1861-1865. This caused British inventors to turn to other cotton suppliers, particularly India. This led to a surge in innovation during the 1861-1865 period, as British inventors struggled to take advantage of the Indian cotton, which was more available but also more difficult to use. The result was an increase of around 67-75% in cotton textile technology innovation, which is tracked using detailed British patent data. This increase was concentrated in two particular types of cotton textile machinery – gins and openers/scutchers – which were particularly important for dealing with Indian cotton. Thus, the shock caused by the U.S. Civil War was large, exogenous, and significantly affected innovation rates.

One feature of the shock is that it was specific to one industry, cotton textiles, and to a particular subset of technologies within that industry which were related to the use of Indian cotton. While innovation rates in the cotton textile industry jumped during the Civil War, innovation in other similar textile industries based on wool, linen, and silk, do not appear to have been affected. This means that I can control for other time-varying factors by comparing the cotton textile industry to these other similar industries, which were also economically important in Britain during this period. Additionally, because the shock affected only some technology types within the cotton textile industry, I am able to control for time-varying factors at the intermediate level by comparing
technologies categories which were directly impacted by the shock to those that were not.

The most important feature of this shock is that it is largely, though not entirely, temporary. British cotton imports quickly rebounded following the end of the Civil War, as did industry output. However, some changes did persist. Indian cotton remained a much larger fraction of imports in the post-shock period than it was prior to the war. The key point is that I am able to identify the expected direction of the bias that could be generated by these persistent effects, since they are continuations of the changes which drove the increase in innovation during the 1861-1865 period. While this is not a perfect empirical setting in which to implement the methodology, it approximates the ideal setting while allowing me to credibly assess the direction of the remaining bias. This is an improvement over previous studies which are not able to directly address these important sources of bias.

I look for path dependence in innovation at the industry level by comparing patenting patterns in the cotton textile industry to the wool, linen, and silk industries. I find that despite the sharp increase in cotton textile related patents during the 1861-1865 period, there is no evidence of a continuing higher level of innovation after 1865. This result is strengthened because we would expect that any changes caused by the Civil War which persisted after 1865 should cause an increase in innovation in the post-shock period. Thus, this paper provides clear evidence against the existence of path dependence in innovation at the industry level. In contrast, when I use the approach offered in previous studies I find strong evidence in favor of path dependence in innovation.

Next, I look for path dependence at the intermediate level using patents of eight specific types of textile technologies. Two of these eight technology types – gins and openers/scutchers – were specifically related to using Indian cotton inputs. Thus, these two technology types experienced a large increase in patents during the 1861-1865 period. The analysis focuses on whether higher levels were sustained in the years after 1865, while using the remaining six technologies as controls. At this level I find some evidence that there was a sustained higher level of patents in these categories in the three years after 1865. The fact that these results differ from those obtained at
the industry level highlights the importance of considering multiple technology levels.

This study is related to a broader set of literature that is concerned with technology spillovers between firms within and across industries. A recent example is Bloom et al. (2012), which attempts to separate the impact of knowledge spillovers between firms from the impact that a new innovation has through product market rivalry. The main difference between Bloom et al. (2012) and this study is that they focus on spillovers across firms, while this study seeks to capture the net impact of innovation spillovers occurring both within and across firms, including countervailing product market rivalry effects, in order to assess the overall impact of the stock of knowledge on technological progress in particular technology types.

In the next section, I present the theoretical model. Section 2.3 presents the empirical setting that will be used to look for evidence of path dependence in innovation, while section 2.4 describes the patent data. Section 2.5 describes the challenges to identifying path dependence in innovation empirically and introduces an approach that overcomes these concerns. Section 2.6 presents the analysis and section 2.7 concludes.

2.2 Theory

This section presents a theory that motivates the empirical exercise by describing how changes in the existing stock of knowledge can affect the incentives for innovation. The model can be thought of as having three levels. In the “industry” level, final goods are produced by perfectly competitive firms using labor and intermediate inputs. To connect the theory to the empirical setting, we can think of these as corresponding to the cotton, wool, linen, and silk textile industries. In the next level down, which I call the “intermediate level”, intermediate inputs are produced using machines and an input, which I will call raw materials, by perfectly competitive intermediate-goods firms. In the empirical context, we can think of the different intermediate goods as yarn produced using different types of cotton, such as U.S. cotton yarn or Indian cotton yarn. The ma-

---

3For simplicity, a single type of raw materials will be used as an input into all of the different intermediate goods. However, I could also incorporate raw materials which are specific to each intermediate good.
chines used in producing intermediate inputs are built by machine makers. These machine makers are monopolists within a particular technology type. They produce and sell machines using their technology. They also invest in R&D to improve the productivity of the machines they produce.

The model is static, with the initial stock of knowledge (technology level) taken as given for each technology type. My interest will be in understanding how increases in the initial stock of knowledge in a particular technology type will affect further innovation in that technology. The main utility of this theory is that it will help us think about how path dependence might occur and why it might occur for only particular subsets of the available technologies.

I now turn to the specifics of the theory. The economy has many sectors denoted by \( j \in (0, J) \) each of which produces one final good \( Y_j \). Consumption is over an index of these final goods with a constant elasticity of substitution form:

\[
Y = \left( \sum_j Y_j^{\sigma-1} \right)^{\frac{\sigma}{\sigma-1}},
\]

where \( \sigma \in (0, +\infty) \) is the elasticity of substitution between the final goods. The corresponding price index is given by,

\[
P = \left( \sum_j p_j^{1-\sigma} \right)^{\frac{1}{1-\sigma}},
\]

where \( p_j \) is the price of final good \( j \). Using the expressions above, we can write the demand for each final good \( j \) as,

\[
y_j = Y p_j^{\sigma-1} p_j^{-\sigma}. \tag{2.1}
\]

There are infinitely many small identical final goods producers with measure 1 which are indexed by \( f \in (0, 1) \). Producing final goods requires labor and intermediate inputs. Intermediate goods are sector-specific and within each sector there are a continuum of different types of intermediates indexed by \( k \in (0, 1) \). The production function for a final goods producing firm \( f \) in sector \( j \) is,
\begin{align*}
y_{jj} &= \left(\frac{1}{\alpha}\right) L_{jj}^{1-\alpha} \int_0^1 z_{jj}(k)^{\alpha} \, dk, \quad (2.2)
\end{align*}

where \( z_{jj}(k) \) is the quantity of intermediate good \( k \) used, \( L_{jj} \) is the amount of labor employed, and \( \alpha \in (0, 1) \). The first order conditions for a final goods firm’s optimization problem are,

\begin{align*}
q_j(k) &= p_j L_{jj}^{1-\alpha} z_{jj}(k)^{\alpha-1}, \quad (2.3)
\end{align*}

and,

\begin{align*}
w &= \left(\frac{1-\alpha}{\alpha}\right) p_j L_{jj}^{\alpha} \int_0^1 z_{jj}(k)^{\alpha} \, dk, \quad (2.4)
\end{align*}

where \( q_j(k) \) is the intermediate good price and \( w \) is the wage. Using (2.3), I derive the following expressions for the demand for intermediate good \( k \) by firm \( f \), and the total demand for intermediate good \( k \), respectively,

\begin{align*}
z_{jj}(k) &= \left(\frac{p_j}{q_j(k)}\right)^{\frac{1}{1-\alpha}} L_{jj}, \quad (2.5)
\end{align*}

\begin{align*}
z_j(k) &= \left(\frac{p_j}{q_j(k)}\right)^{\frac{1}{1-\alpha}} L_j. \quad (2.6)
\end{align*}

Intermediate goods are produced by many identical perfectly competitive firms indexed by \( i \in (0, 1) \). Producing intermediate goods requires machines and raw materials. The production function for intermediate goods is given by,

\begin{align*}
z_{ji}(k) &= \left(\frac{1}{\beta}\right) K_{ji}(k)^{1-\beta} \int_0^1 A_{ji}(k,l)^{1-\beta} x_{ji}(k,l)^{\beta} \, dl. \quad (2.7)
\end{align*}

In this equation, raw materials are represented by \( K_{ji}(k) \). The set of available machines is indexed by \( l \in (0, 1) \). Each machine type has technology level \( A_{ji}(k,l) \) and the quantity of machines used is given by \( x_{ji}(k,l) \). The parameter \( \beta \in (0, 1) \) determines the substitutability between different types of machines in the production process. The first order conditions for each intermediate good
producer’s optimization problem are, 

\[ \chi_j(k,l) = q_j(k) K_{ji}(k)^{1-\beta} A_j(k,l)^{1-\beta} x_{ji}(k,l)^{\beta - 1}, \] (2.8)

and,

\[ r = \left( \frac{\beta}{1-\beta} \right) q_j(k) K_{ji}(k)^{-\beta} \int_0^1 A_j(k,l)^{1-\beta} x_{ji}(k,l)^{\beta - 1} dl, \] (2.9)

where \( \chi_j(k,l) \) is the price of machine type \( l \) and \( r \) is the price of raw materials. Using (2.8), I derive the following expressions for the demand for machines of each type by each intermediate goods producer, and total demand for each type of machine, respectively:

\[ x_{ji}(k,l) = q_j(k)^{1-\beta} \chi_j(k,l)^{\beta + 1} K_{ji}(k) A_j(k,l), \] (2.10)

and

\[ x_j(k,l) = q_j(k)^{1-\beta} \chi_j(k,l)^{\beta + 1} K_j(k) A_j(k,l). \] (2.11)

Machine producers are monopolists in one particular type of machine in one sector, so they are indexed by their machine type, \( l \). Machines are produced with a constant marginal cost \( \psi \). This means that, holding fixed the machine technology level \( A_j(k,l) \), a machine producer’s profit is given by \( \pi_j(k,l) = x_j(k,l) (\chi_j(k,l) - \psi) \). Solving this, I find that the price charged by a machine producer is a fixed mark-up over marginal cost, \( \chi_j(k,l) = \psi/\beta \). Since I am not interested in conducting comparative statics on the parameters \( \psi \) or \( \beta \), I adopt the normalization that \( \psi = \beta \) so that \( \chi_j(k,l) = 1 \), which will simplify the analysis. I can now rewrite the demand for machines and the profit of machine makers, given technology level \( A_j(k,l) \), as

\[ x_j(k,l) = q_j(k)^{1-\beta} A_j(k,l) K_j(k), \] (2.12)

and,
\[ \pi_j(k) = (1 - \beta) q_j(k)^{\frac{1}{1-\beta}} A_j(k, l) K_j(k). \]

Now consider the machine makers’ decision of how much to invest in R&D in order to improve the quality of their machine. This decision will depend crucially on how we model the costs of innovation. I will use the following general function to represent the relationship between the previous technology level \( A_j^0(k, l) \), R&D expenditure \( R_j(k, l) \), and the level of technology that is produced through these expenditures, \( A_j(k, l) \):

\[ A_j(k, l) = g \left( R_j(k, l), A_j^0(k, l) \right). \] (2.13)

This function is increasing in both \( R_j(k, l) \) and \( A_j^0(k, l) \) and concave in \( R_j(k, l) \). My focus is on path dependence generated through demand channels. To concentrate on this channel, I want to “turn off” the possibility that the level of technology influences innovation through directly affecting the cost of producing new technologies.\(^4\) In order to “turn off” path dependence in innovation which is generated through cost channels, I make the following assumption:

**Assumption 1.**

\[ \frac{\partial^2 g \left( R_j(k, l), A_j^0(k, l) \right)}{\partial R_j(k) \partial A_j^0(k)} = 0. \]

This assumption ensures that a change in the initial technology level will not affect the amount of technological progress which is generated by a given amount of R&D expenditure. The optimal R&D expenditure for a machine making firm is now the solution to,

\[ \max_{R_j(k, l)} (1 - \beta) q_j(k)^{\frac{1}{1-\beta}} K_j(k) g \left( R_j(k, l), A_j^0(k, l) \right) - R_j(k, l). \]

Using the first order condition from this problem, I obtain,

\(^4\)This is not to say that cost channels are not an important. Undoubtedly, cost channels do matter and are worthy of future study. But this study will not allow us to differentiate between these various channels, and so for simplicity I will confine myself to modeling path dependence which is driven through demand channels alone. This follows much of the recent existing work on this topic, such as Acemoglu et al. (2012).
\frac{\partial g}{\partial R_j(k,l)} = \left( \frac{1}{1 - \beta} \right) q_j(k)^{-\frac{1}{\alpha}} K_j(k)^{-1} \tag{2.14}

Note that (2.14) implies that \( R_j(k,l) \) will be increasing in both the price of the intermediate good and the quantity of raw materials allocated to producing the intermediate good. These two influences are akin to the price and market size effects which are familiar from Acemoglu (2002).

Next, we need to solve for \( q_j(k) \) and \( K_j(k) \) and plug these values into (2.14). Using (2.9), (2.12) and \( \chi_j(k,l) = 1 \), I obtain,

\[ q_j(k) = \left( \frac{r}{1 - \beta} \right)^{1-\beta} A_j(k)^{\beta-1}, \tag{2.15} \]

where \( A_j(k) = \int_0^1 A_j(k,l) \, dl \). Using (2.5), (2.7), (2.12) and (2.15), I obtain,

\[ K_j(k) = \left( \frac{r}{1 - \beta} \right)^{\frac{\alpha\beta-1}{1-\alpha}} p_j^{\frac{1}{1-\alpha}} L_j A_j(k)^{\frac{\alpha(1-\beta)}{1-\alpha}}. \tag{2.16} \]

It remains to solve for \( p_j \) and \( L_j \) in terms of technology levels in order to substitute these terms out of (2.16). I solve for \( p_j \) using (2.4), (2.5) and (2.15) to obtain,

\[ p_j = \left( \frac{\alpha}{1-\alpha} \right)^{(1-\alpha)} w^{1-\alpha} \left( \frac{r}{1 - \beta} \right)^{\alpha(1-\beta)} A_j^{-(1-\alpha)}, \]

where \( A_j = \left( \int_0^1 A_j(k,l)^{\frac{\alpha(1-\beta)}{1-\sigma}} \, dk \right) \) is an index over all technologies in sector \( j \). Note that this index is unambiguously increasing in any particular technology \( A_j(k,l) \). To solve for \( L_j \), I use (2.1), (2.2), (2.5) and (2.15) to obtain,

\[ L_j = Y P^{\sigma-1} \left( \frac{r}{1 - \beta} \right)^{\alpha(1-\beta)(1-\sigma)} \left( \frac{1 - \alpha}{\alpha} \right)^{-(1-\alpha)(1-\sigma)} w^{\sigma-\alpha-\sigma} A_j^{-(1-\alpha)(1-\sigma)}. \]

Substituting the four preceding equations into (2.14), I obtain an expression which implicitly defines the machine makers’ optimal R&D expenditure level in terms of only aggregate variables, parameters, and technology levels,
\[
\frac{\partial g \left( R_j(k,l), A_j^0(k,l) \right)}{\partial R_j(k,l)} = \Gamma(r,w,Y,P) A_j^{1+(1-\alpha)(1-\sigma)} A_j^{(\alpha(1-\beta))} ,
\]

where for simplicity I have gathered all of the terms other than the technology terms into the function \( \Gamma(r,w,Y,P) \). Since all of these are aggregate variables, I will think of them as unaffected by the industry-specific changes I will be interested in.

Finally, I want to express the technology indices at the industry and intermediate levels, \( A_j \) and \( A_j(k) \), in terms of initial technology levels and R&D expenditure. Let,

\[
A_j = G_j \left( \{R_j(k,l)\}_{k,l}, \{A_j^0(k,l)\}_{k,l} \right) \quad \text{and} \quad A_j(k) = F_{jk} \left( \{R_j(k,l)\}_{l}, \{A_j^0(k,l)\}_{l} \right).
\]

Note that, given that \( A_j^0(k,l) \) is assumed to be increasing in \( A_j^0(k,l) \), the function \( G_j() \) will be increasing in \( A_j^0(k,l) \) for any technology \((k,l)\) in sector \( j \), and the function \( F_{jk}() \) will be increasing in \( A_j^0(k,l) \) for any technology \( l \) used for intermediate good \( k \) in sector \( j \).

Let \( \phi = 1/(1-\alpha) \) and \( \eta = 1/(1-\beta) \) so that these are, respectively, the elasticity of substitution between intermediate goods and the elasticity of substitution between machines. I can now rewrite (2.17) for technology \( m \) in intermediate good sector \( n \) as,

\[
\frac{\partial g \left( R_j(n,m), A_j^0(n,m) \right)}{\partial R_j(n,m)} = \Gamma(r,w,Y,P) \\
\times \left[ G_j \left( \{R_j(k,l)\}_{k,l}, \{A_j^0(k,l)\}_{k,l} \right) \right]^{1-\sigma \over \phi + 1} \\
\times \left[ F_{jk} \left( \{R_j(k,l)\}_{l}, \{A_j^0(k,l)\}_{l} \right) \right]^{1-\theta \over \eta + 1} .
\]

The main results of the theory, derived from Equation 2.18, are described below.

**Intermediate-level path dependence:** An increase in the initial technology level (stock of knowledge) of some set of technologies related to intermediate good \( k \) will lead to an increase in R&D investments related to all other machines used to produce intermediate good \( k \) when \( \frac{\phi-1}{\eta} > 1 \).
other words, an increase in $A^0_j(k,l)$ for $l \in (a,b)$ and $a \neq b$, will lead to an increase in $R_j(k,m)$ for all $m \in (0,1)$ when $\frac{\phi - 1}{\eta} > 1$.

**Industry-level path dependence:** An increase in the initial technology level (stock of knowledge) of some set of technologies in industry $j$ related to intermediate good $k$ will lead to an increase in R&D investments related to all other machines used in sector $j$ when $\sigma - 1 > 1$. In other words, an increase in $A^0_j(k,l)$ for $l \in (a,b)$ and $a \neq b$ and $k \in (c,d)$, $c > d$ will lead to an increase in $R_j(n,m)$ for all $n$ and $m$ when $\frac{\sigma - 1}{\phi} > 1$.

In both of these results, whether or not we observe path dependence depends on the elasticity of substitution. At the industry level, we should observe path dependence in innovation when the substitutability of final goods is high relative to the elasticity of substitution between intermediate goods. At the intermediate level, we should observe path dependence when the elasticity of substitution between intermediate goods is high relative to the substitutability between particular machine types.

One notable point here is that the model predicts that path dependence may occur solely at the industry level, solely at the intermediate level, at both levels, or at neither level. This implies that we could not rule out the existence of path dependence in technology by looking across only one level of technology, e.g., by comparing across machines related to different parts of the production process within industries. Looking at a less aggregated technology level does not solve the problem; it is possible that path dependence in innovation exists at the industry level, but not at the intermediate level. Moreover, the model indicates that there is a trade-off between path dependence in innovation at different levels. The higher is the elasticity of substitution between intermediate goods $\phi$, the more likely we are to observe path dependence in innovation at the intermediate level, but the less likely we are to observe path dependence in innovation at the industry level. The opposite holds for a low $\phi$. This motivates the multi-level approach to searching for path dependence in innovation which is used in the empirical investigation.
2.3 Empirical setting

The cotton textile industry was one of the largest and most important manufacturing sectors in Britain in the 19th century. In 1860, for example, cotton textiles was Britain’s most valuable export good, and raw cotton was Britain’s largest import.\footnote{Cotton Textile exports were worth £52 million in 1860. The next largest exports were Wool textiles at £15.7 million and Iron and Steel at £13.6 million. The value of raw cotton imports in 1860 was £35.8 million. The next largest import categories were Grain and Flour, at £31.7 million and Oils and Resins at £14.9 million. Data from Mitchell & Deane (1962).} The British cotton textile industry was entirely reliant on foreign suppliers to meet its raw cotton input needs. Prior to the U.S. Civil War (1861-1865), 77% of these inputs were supplied from the Southern U.S. The next largest suppliers were India, at 13% of imports, and Brazil and Egypt, each of which provided around 3% of the market.

This study takes advantage of the large exogenous shock resulting from the U.S. Civil War, which sharply disrupted the supply of raw cotton to Britain from the Southern U.S. during the 1861-1865 period. One indicator of the magnitude of this shortage is shown in Figure 2.1, which plots total British cotton imports, and British cotton imports from the U.S., from 1815-1910. This figure shows that the sharp drop in U.S. cotton imports during the 1861-1865 period led to a drop in overall imports, though a gap also opened up between the lines as other suppliers, chiefly India, increased their exports to Britain.

The sharp change in the share of imports coming from the different suppliers is shown in Figure 2.2. This was an important change because cotton from different locations differed in important ways. Specifically, Indian cotton was more difficult for producers to prepare for the spinning process. The seeds of the Indian cotton were more difficult to remove, a process done using cotton gins. The Indian cotton arrived in a dirtier state and so required more cleaning, a process that was done using openers, scutchers, and carding machines. Also, the Indian cotton was generally of a short-staple variety which was more difficult to spin than the medium or long-staple cotton available from the U.S. and other locations. These are important differences; the increased importance of Indian cotton during the Civil War period led to a burst of innovation.
focused on addressing these challenges.

Figure 2.1: British cotton imports 1815-1910

![Graph showing British cotton imports 1815-1910.](image)

Data from Mitchell & Deane (1962).

Figure 2.2: Share of imports by supplier 1850-1880

![Graph showing share of imports by supplier 1850-1880.](image)

Data from Ellison (1886).

Two other measures of the effects of the cotton shortage are shown in Figures 2.3 and 2.4. Figure 2.3 describes the level of domestic raw cotton consumption in Britain, which is the best available measure of output in the cotton textile industry. This figure shows the severity of the crises that the cotton shortage caused in the industry, with output dropping by as much as a half.
Hundreds of thousands of workers were also left without a job or forced to work short-time. Figure 2.4 shows the impact that the shortage had on cotton prices for the benchmark varieties of American, Indian, Brazilian, and Egyptian cotton. Not surprisingly, the price of all of the major cotton varieties increased sharply during the Civil War period.

Another feature of this shock was that it was largely industry-specific. While the cotton textile industry suffered a severe reduction in output during the Civil War, Britain’s wool, linen, and silk textile industries, which were also important industries in Britain at this time, were largely unaffected. One indicator of this is Figure 2.5, which shows that British wool and linen textile exports did not show any negative effects during the Civil War. If anything, there is some evidence that these textile varieties benefited from the reduction in competition from cotton textiles. While the wool, linen, and silk industries used different input fibers from the cotton textile industry, all of these industries shared many of the same technologies. For example, spinning mules originally designed for the cotton textile industry were also modified for use in the wool, linen, and silk industries.

Figure 2.3: British raw cotton consumption 1815-1910

Data from Mitchell & Deane (1962).
Figure 2.4: Raw cotton prices on the Liverpool market for key varieties 1852-1875

Quarterly price data from *The Economist*. Upland Middling is the benchmark U.S. cotton variety. Surat is the benchmark Indian cotton variety. Pernambuco is the benchmark Brazilian cotton. Egyptian Fair is the benchmark cotton variety from that location.

Figure 2.5: British wool and linen textile exports 1815-1910

Data from Mitchell & Deane (1962)

There are three important points to take away from Figures 2.1-2.5. First, the shock was large. The shortage of raw cotton caused output in Britain’s largest manufacturing industry to drop by as much as half. Second, the shock was, in many but not all respects, temporary. Figure 2.1 shows that the total level of British cotton imports rebounded quickly after 1865, as did the measure
of industry output shown in Figure 2.3. On the other hand, some changes did persist. Figure 2.2 shows that Indian cotton retained a higher share of British imports than during the pre-war period for several years after 1865. Also, Figure 2.4 suggests that cotton price remained relatively high, compared to the pre-war period, for several years after the war. The temporary nature of the shock is an important part of this analysis, and the fact that some of the effects of the shock persist after 1865 will introduce a potential for bias in the empirical results that I will be careful to take into account. The third important feature of the shock is that it was primarily industry and technology-specific. As shown in Figure C.4, the other British textile industries were not negatively impacted by the cotton shortage. This will be important because I will be comparing the innovation patterns observed in the cotton textile industry to innovation patterns in these other, similar, textile industries. This will allow me to control for time-varying factors which may have affected innovation rates in the textile industries.  

2.4 Patent data

In order to track innovation, this study will take advantage of British patent data. Patent data are widely used as a measure of innovation rates, including by the previous studies in this literature. The advantages of using patent data are that there is generally a large number of observations available spanning a wide set of industries and technologies, patent data contain a number of useful pieces of information, and these data are provided voluntarily with clear and consistent incentives.

There are also some potential disadvantages in using patent data. One issue is that patent data will not capture all types of innovation. Using data from the 19th century Worlds Fairs, Moser (2010) shows that a significant fraction of new inventions were not patented. However, her data

---

6For example these industries suffered from periodic financial crisis, including one in 1865-1866, which impacted Britain’s northern industrial districts. Because these shocks are likely to be common to all of the textile industries, comparing across industries allows me to control for such factors.

7Seminal studies using patent data to measure innovation include Schmookler (1966), Scherer (1982), Griliches (1984), and Jaffe et al. (1993).

8See Hall et al. (2001) for a useful review of the advantages and disadvantages of using patent data to track innovation patterns.
also show that, amongst all technology varieties, inventions of manufacturing machinery – the main type of inventions studied here – were the most likely to be patented. A second concern with patent data is that patent counts may not reflect patent quality. I will address this concern by using data on the payment of patent renewal fees, which were expensive fees that British patent holders had to pay in order to keep their patent in force.9

Another concern is that the granting of intellectual property protection on an existing invention may stifle further innovation in a particular technology type. This concern has been highlighted in work by Murray & Stern (2007) and Williams (2010). The focus of this study will be investigating the existence of path dependence in innovation in the context of a system that offers patent protection. This may not reveal whether path dependence in innovation will hold in a context that is free of patent protection. However, because most advanced nations offer patent protection, and this protection is widely utilized, understanding whether path dependence in innovation holds in a context in which patents and other intellectual property protections are available is likely to be the most relevant setting for understanding innovation and economic growth.

The analysis will take advantage of British patent data from two sources. The first are data from the British Patent Office (BPO) which were collected from original printed documents in the British library. This data set covers all patents from 1855-1883, 118,863 in all. In the BPO data, each patent has been classified into one or more of 146 technology categories by the British Patent Office. These categorizations allow me to focus on patents representing particular types of technologies. Most of my analysis will focus on patents in the BPO spinning technology category, which includes textile technologies related to the preparation, cleaning, and spinning of yarn. This was an important area of innovation during this period; BPO spinning patents make up 6% of all British patents during the 1855-1883 period.

The second dataset comes from the A Cradle of Inventions (COI) database, where I use data covering patents from 1853-1880.10 This data set provides two useful pieces of additional infor-

---

9Patent renewal fees have been used as an indicator of patent quality in a number of previous studies, including Schankerman & Pakes (1986), Sullivan (1994), Lanjouw et al. (1998), Brunt et al. (2008).

10I thank Tom Nicholas for suggesting this data source.
First, it includes the patent titles, which were not available in the BPO data. This will allow me to focus on patents which are specifically related to the cotton, wool, linen, and silk industries.\textsuperscript{11} These titles are consistently available for the years 1853-1870. Second, this data set provides more complete inventor names than those available in the BPO spinning data.\textsuperscript{12}

Part of the analysis will take advantage of data on the individual inventors that filed patents. Individual inventors are tracked over time by matching first and last names from the COI database. After matching, I’m left with a total of 3,567 textile inventors who generated 4,164 patents and 6,286 inventory x patent observations from 1853-1870.\textsuperscript{13} One issue in tracking individual inventors is that, for some patents, patent agents are named rather than the actual inventor. This seems to be particularly common for patents filed by foreign inventors. To deal with this issue, I drop all inventors who filed more than 100 patents during the 1853-1870 period from the analysis. Across all spinning patents this eliminates 21 inventors and 441 inventor x patent observations. Most of these individuals are known to have been patent agents.

Within the broad BPO spinning technology category, I collected additional data on eight technology subcategories. These additional data allow me to track innovation patterns in specific types of textile technologies. These eight major spinning technology subcategories include 2,586 patents in the 1855-1876 period out of a total of 5,899 spinning technology patents. Table 2.1 shows the count of patents in each of these subcategories related to the each of the four main textile industries.

\textsuperscript{11}More details on how the patent titles were used to identify patents related to cotton, wool, linen, and silk are available in Chapter 1.

\textsuperscript{12}Specifically, during the 1853-1870 period, the COI database includes the full first and last name of the inventor, while the BPO spinning data includes only the first initial, rather than the first name. Having inventors’ full first name improves my ability to match patents by the same inventor.

\textsuperscript{13}This set includes all inventors with a patent categorized in the BPO spinning technology category or a patent identified as related to cotton, wool, linen, or silk.
Table 2.1: Technology subcategory patents by input type

<table>
<thead>
<tr>
<th>Technology subcategory</th>
<th>Cotton-related</th>
<th>Wool-related</th>
<th>Linen-related</th>
<th>Silk-related</th>
<th>Total patents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gins</td>
<td>109</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>116</td>
</tr>
<tr>
<td>Openers/scutchers</td>
<td>146</td>
<td>57</td>
<td>19</td>
<td>22</td>
<td>244</td>
</tr>
<tr>
<td>Carding machines</td>
<td>196</td>
<td>150</td>
<td>34</td>
<td>21</td>
<td>401</td>
</tr>
<tr>
<td>Combing machines</td>
<td>58</td>
<td>171</td>
<td>43</td>
<td>53</td>
<td>325</td>
</tr>
<tr>
<td>Mules</td>
<td>126</td>
<td>33</td>
<td>9</td>
<td>7</td>
<td>175</td>
</tr>
<tr>
<td>Rollers</td>
<td>117</td>
<td>78</td>
<td>52</td>
<td>24</td>
<td>271</td>
</tr>
<tr>
<td>Bearings</td>
<td>70</td>
<td>28</td>
<td>13</td>
<td>16</td>
<td>127</td>
</tr>
<tr>
<td>Finishing of yarn</td>
<td>23</td>
<td>24</td>
<td>16</td>
<td>40</td>
<td>103</td>
</tr>
</tbody>
</table>

Note that some patents are associated with multiple input types.

All of the analysis in this study is based on the date at which the patent application was originally filed, which represents the earliest date at which the invention entered the patent system. Thus, I will not need to worry about delays in granting of the patent in this analysis. To file an application, inventors had to provide a short preliminary specification, and the filing date became the priority date for the invention. The inventor then had six months in which to file a full specification describing the invention or the patent became void. Because the British patent system did not include a review by a patent examiner at this time, nearly all of these patent applications were granted.

One important issue for this study is the lag between innovative activity and a patent filing. If this lag were, say, two years, then we could not attribute a higher level of innovation during the first two years after 1865 to path dependence in innovation. There are however, several reasons to think that this lag is relatively short – roughly 1-2 years or less – in the present context. One reason is that I am looking at patent applications, which require only a relatively simple preliminary specification and then give the patent filer time to work out the specific details of the invention. Another piece of evidence which pertains directly to cotton gins, an important technology for this study, comes from work by Lakwete (2003). In her book, Inventing the Cotton Gin, she describes numerous instances in which inventions were generated, and patents filed, within a one to two
year period. Thus, I do not expect that the lag between innovation and patent filing to be a major source of bias in this study, though, it may exert some bias in the direction of finding evidence of path dependence in innovation.

There were no major patent law changes in Britain between 1852 and 1883. During this time, a patent application cost £25, which was considered a substantial sum at this time. In addition, patent holders had to pay a fee of £50 to keep their patent in force after three years and an additional fee of £100 to keep the patent in force after 7 years. If both fees were paid, the patent remained in force for 14 years.  

2.5 Empirical issues and approach

This section describes two issues with the empirical strategy used in the leading studies of path dependence in innovation. These issues are likely to generate a spurious positive relationship between the stock of knowledge and the level of new innovation. I then outline an empirical strategy designed to address these concerns.

2.5.1 Issues

I start by considering the empirical specification based on that used in Popp (2002), though the issues I describe are also present in more recent work by Aghion et al. (2011). In Popp’s primary regression specification, reproduced below, $\text{PAT SHARE}_{it}$ represents the ratio of energy-efficient patent type $i$ to total U.S. patents in period $t$, $\text{PRICE}^*_t$ represents the price of energy, which directly

\[ \text{PAT SHARE}_{it} = \frac{\text{energy-efficient patents in period } t}{\text{total U.S. patents in period } t} \]

\[ \text{PRICE}^*_t = \text{price of energy} \]

14For additional information about the British patent system during this period, see Van Dulken (1999) and Khan (2005).

15Other sources of bias may also exist, and some of these may decrease, rather than increase, the chances of observing path dependence in innovation. The two sources of bias I focus on are chosen because existing studies tend to find evidence in favor of path dependence in innovation, and so it seems that sources of bias which make it more likely to observe path dependence in innovation are dominating other types of bias. However, the empirical methodology outlined in this section will be robust to a number of additional sources of bias.

16Aghion et al. (2011) use a similar approach, although it is run on firm-level panel data and they use gas tax rates as an instrument for the price of energy. The use of panel data by Aghion et al. (2011) allows them to control for aggregate time-varying effects, while their use of an instrument for the price of energy reduces reverse causality concerns for this variable. However, unlike this study, they do not include time trends that are specific to the cross-sectional dimension of the panel data, nor do they have an instrument for the stock of knowledge.
affects the incentives for energy-related innovation, $STOCK_{it}$ represents the stock of knowledge of type i, and $X_{it}^*$ is a set of control variables.\(^{17}\)

$$
\log(PAT\ SHARE_{it}) = \alpha + \beta_0 \log(PRICE_{it}^*) + \beta_1 \log(STOCK_{it-1}) + \beta_2 \log(X_{it}^*) + \epsilon_{it}).
$$

Using panel data it is possible to control for aggregate time-varying effects, as well as technology-specific fixed effects, by rewriting this equation as,

$$
\log(PAT_{it}) = \alpha + \beta_0 \log(PRICE_{it}^*) + \beta_1 \log(STOCK_{it-1}) + \beta_2 \log(X_{it}^*) + FE_i + T_t + \epsilon_{it}).
$$

where I now have just the number of patents of type i on the left-hand side, and I have added a set of technology-specific fixed effects $FE_i$ and year indicator variables $T_t$ to the right hand side of the equation.\(^{18}\) This expression is closer to the approach used by Aghion et al. (2011) and allows us to address some potential sources of bias. However, as I will discuss, at least two important sources of bias remain.

One issue with the econometric approach described above is that the presence of persistent shocks can generate a positive correlation between $PAT_{it}$ and $STOCK_{it-1}$, even when no causal relationship between these variables exist. To illustrate this, suppose that we are studying twenty periods and that the data generating process for innovation is $PAT_t = 1 + I[t \geq 10]$ where $I[t \geq 10]$ is an indicator variable that takes a value of one starting in period 10. I.e., there is a persistent shock which increases the level of innovation beginning in period $t = 10$. The left panel of Figure 2.6 plots the relationship between $PAT_t$ and $STOCK_{t-1}$ assuming this simple data generating process. The regression results in the right panel of Figure 2.6 show that, when the permanent shock is unobserved and so not included in the regression, we will estimate a spurious positive relationship

\(^{17}\)The * indicates that the $PRICE$ and control variables were modeled using a distributed-lag approach.

\(^{18}\)Note that $PAT\ SHARE_{it}$ is equal to the ratio of $PAT_t$ to the total number of patents in period t. Taking logs, I can move the total number of patents to the right-hand side of the equation, but because this is an aggregate variable it will be perfectly collinear with $T_t$, so it is dropped from the analysis.
between $PAT_t$ and $STOCK_{t-1}$.

Figure 2.6: Relationship between PAT and STOCK with a permanent shock at $t = 10$

| DV: Count of patents | Lag stock 0.192****  
(0.0344) | Constant -0.0784  
(0.300) | Observations 20 | R-squared 0.634 |

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

When using panel data and including indicator variables for each time period (as in Aghion et al. (2011)), the approach described above will be robust to unobserved persistent shocks which occur at the aggregate level. However, if these shocks occur for particular firms, or particular subsets of technologies, then their results are susceptible to bias from these unobserved persistent shocks. The approach used in Popp (2002) is even more susceptible to this source of bias, since he does not use panel data and so is not able to control for aggregate time-varying factors.

A second issue with the econometric approach described above is that it is susceptible to bias through unit-specific time trends, where unit refers to the cross-sectional units of the panel data. To see why, suppose that the true data generating process is:

$$PAT_{it} = \gamma_i T_t + \phi_i + T_t + \epsilon_{it},$$

where $\gamma_i$ and $\phi_i$ are unit-specific parameters. There are a number of reasons why innovation by a particular firm, industry, or subset of technologies within an industry, may have such a time-trend. For example, this may occur if some firms tend to be expanding and increasing their investments in innovation over the study period while others are shrinking and decreasing their investments. It may also occur if some industries, or some technologies within an industry, are experiencing sustained growth, while others are not. We will see that this will be the case in the empirical
setting I study. Rewriting the expression above, I obtain,

\[ PAT_{it} = \gamma_i T_{t-1} + \phi_i + T_t + \epsilon_{it}, \]

while the stock of knowledge measure is,

\[ STOCK_{it-1} = (1 - \delta)STOCK_{it-2} + \gamma_i T_{t-1}. \]

It is clear from these expressions that there is the potential for bias in a regression of \( PAT_{it} \) on \( STOCK_{it-1} \) when the \( \gamma_i T_{t-1} \) term is omitted from the regression specification.

Both of the source of bias which I have just described are due to conditions which seem likely to occur in reality. An econometric strategy that ignores either of these potential sources of bias is likely to deliver evidence in favor of path dependence in innovation, regardless of whether the stock of existing knowledge has any causal impact on the current level of innovation. In the next section I outline an empirical approach that seeks to deal with these concerns.

2.5.2 Approach

This section introduces an empirical approach that will be used to address these issues. The second issue can be easily addressed by using data with a sufficient time dimension that it is possible to include time-trends specific to each cross-sectional unit in the panel data. The first issue is addressed using an observable temporary shock that affects only a subset of industries or technology types. Consider a temporary shock that increases the incentives for innovation during the shock period. This will also increase the stock of knowledge in the period just after the shock. Then, if we observe a higher level of innovation in the post-shock period, this suggests a causal relationship between the stock of knowledge and the level of innovation. The empirical setting analyzed in this paper is chosen because it approximates these conditions, and, where it departs from the ideal setting it is possible to credibly assess the likely direction of the bias generated by this departure so that it can be taken into account when assessing the empirical results.
I begin the analysis in section 2.6.1 by using panel data in which the cross-sectional dimension spans the four main British textile industries – cotton, wool, linen, and silk textiles. The empirical specification is:

\[
\textit{PAT}_{it} = \beta_0 + \beta_1 \textit{STOCK}_{it} + \beta_2 Y_t + FE_i \Gamma_1 + TT_{it} \Gamma_2 + \textit{SHOCK}_{ct} \Gamma_3 + \epsilon_{it}. \tag{2.19}
\]

In this equation, \textit{PAT}_{it} represents the log count of patents in technology type \(i\) at time \(t\). \textit{STOCK}_{it} represents the log stock of knowledge related to that technology type, which will be measured in several ways, as described below. \(FE_i\) is a set of industry fixed effects, while \(TT_{it}\) is a set of industry-specific time-trends, and \(Y_t\) is a set of year indicator variables. I also include \(\textit{SHOCK}_{ct}\) which is a set of year indicator variables for each year from 1861-1865, interacted with an indicator variable for the cotton industry. The inclusion of this set of variables helps ensure that the results are not driven by the sharp increase in cotton textile patents during the 1861-1865 period.

One issue with running (2.19) as written is that it is likely to be biased by unobserved persistent shocks. To deal with this concern, I use the exogenous shock to the cotton textile industry as an instrument for the stock of knowledge. The first stage regression for this instrumental variables approach is,

\[
\textit{STOCK}_{it} = \alpha_0 + \alpha_1 \textit{POSTSHOCK}_{ct} + \alpha_2 Y_t + FE_i \Lambda_1 + TT_{it} \Lambda_2 + \textit{SHOCK}_{ct} \Lambda_3 + e_{it}, \tag{2.20}
\]

where \(\textit{POSTSHOCK}_{ct}\) is an indicator variable for the four years starting in 1866 interacted with an indicator variable for the cotton textile industry. The identifying assumption for this approach is that the only impact of the Civil War on cotton textile innovation after 1865 occurred through the increase in the stock of experienced inventors. This assumption is likely to be violated, since, as shown in Figure 2.2, the share of Indian cotton in total inputs remained high in the post-shock period. However, given that these persistent effects were largely continuations of those that drove the increase in innovation during the Civil War period, it is reasonable to expect that they will
also exert upward pressure on the level of innovation in the post-war period. Thus, any bias in the IV estimates due to the persistent effects of the war should be in favor of path dependence in innovation.

In section 2.6.2, I look for evidence of path dependence at the intermediate level. To do so, I use data that spans eight different types of textile technologies. The specification for the intermediate-level analysis is:

\[ PAT_{it} = \beta_0 + \beta_1 STOCK_{it} + \beta_2 Y_{it} + FE_i \Gamma_1 + TT_{it} \Gamma_2 + SHOCK_{it} \Gamma_3 + \epsilon_{it}. \]  

(2.21)

This equation differs from the approach used in the industry-level analysis in only minor ways. The cross-sectional dimension of the data spans technology types, rather than industries, with two technology types, gins and openers/scutchers, experiencing the shock directly. The set of \( SHOCK_{it} \) variables will now be the interaction between indicator variables for the years 1861-1865 and indicator variables for the gins and openers/scutchers subcategories. Another difference is that I will use Poisson regressions at the intermediate level because, when cutting the data to such a fine level we are left with fewer observations, so I am more concerned that the discrete nature of the data may bias my results. As before, I will use an IV strategy, where the first stage regression specification is,

\[ STOCK_{it} = \alpha_0 + \alpha_1 POSTSHOCK_{it} + \alpha_2 Y_{it} + FE_i \Lambda_1 + TT_{it} \Lambda_2 + SHOCK_{it} \Lambda_3 + \epsilon_{it}. \]  

(2.22)

Here the instrumental variable is constructed by interacting an indicator variable for the four years after 1865 with an indicator variable for the gins and openers/scutchers subcategories.

In the upcoming analysis I will use four measures of the stock of knowledge. The first two measures are constructed using the identity of the inventors. The first measures, which I will call \( RECENT \), is constructed by taking the stock of inventors who have been active in a particular technology category in the past four years. I will also use the measure, \( EXPER \), which is the
total stock of inventors with experience in a technology type, aggregated starting from the from the first year for which data is available. The third measure, $PSTOCK$ is calculated using the count of patents in the four years prior to $t$ with discount factor $\gamma$, according to: 

$$PSTOCK_{it} = \sum_{\tau=1}^{4} (1 - \gamma)^{\tau-1} PAT_{it-\tau}.$$ 

Following Aghion et al. (2011), I adopt $\gamma = .15$. Finally, I use a measure, $PSUM$, which sums the number of patents of a particular technology type starting from the first year in which data is available. In constructing the $RECENT$ and $PSTOCK$ variables, my use of a lag length of four is chosen to cover several years without sacrificing too much data at the beginning of the time series.

### 2.6 Analysis

The analysis in this section begins by looking for evidence of path dependence at the industry level. This is done by comparing innovation patterns in the cotton textile industry to those in the wool, linen, and silk industries. I then turn to the intermediate level, where I look across a number of specific technology types which are related to particular stages in the production process.

#### 2.6.1 Industry-level analysis

The starting point for the industry level analysis is Figure 2.7, which shows the count of patents, by year, in the cotton, wool, linen, and silk industries. The first point to take from this figure is that there was a sustained high level of patent filing in the cotton textile industry during the 1861-1865 period. This burst of innovation was caused by the shift from U.S. to Indian cotton supplies caused by the Civil War. British innovators responded to this shift by generating new innovations which helped them deal with the challenges involved in using the more difficult to process Indian cotton.\(^9\) Figure 2.7 also shows that the burst of innovation during the 1861-1865 period was confined to the cotton textile industry. None of the other textile industries show any significant increase in innovative output.

\(^{9}\)See Chapter 1 for more details.
Figure 2.7: Patents in each textile industry from 1853-1870

If there is path dependence in innovation at the industry level, how would we expect this to appear in Figure 2.7? Given the large increase in cotton-related innovations during the 1861-1865 period, there should be a significant increase in the stock of knowledge and experienced inventors in the years just after 1865. Thus, we should observe that the increase in cotton-related innovation should have a “broad shoulder”, an elevated level of innovative activity in these technologies in the years after 1865. No such pattern emerges from Figure 2.7, which suggests that path dependence in innovation at the industry level was not an important factor in this empirical setting. This is particularly clear when we compare innovation in the cotton textile industry to that in the wool textile industry. Both of these industries experienced similar levels of patenting both before and after the 1861-1865 period.

It is also worth noting that there is some evidence of industry-specific time trends in Figure 2.7. This is particularly true for patents related to silk textiles, which appear to be declining over the period. In general, the silk textile industry was declining in Britain at this time. Because of this trend, including silk as a control for patenting patterns in the cotton textile industry over this
period without including industry-specific time-trends is likely to lead us to find that there was a relative increase in cotton technology patents in the post-shock period, which would bias our results towards finding evidence of path dependence in innovation. This highlights the importance of accounting for unit-specific time trends in the analysis.

Because we may be concerned about variation in the quality of the technologies represented by these patents, it also worth looking at a similar graph which includes only high quality patents, identified as those for which the patent renewal fee was paid at year three to keep the patent in force. Figure 2.8 presents the count of patents by textile industry focusing only on high-quality patents. It is clear that there was also an increase in the number of high-quality patents filed during the 1861-1865 period, but, as before, these data do not suggest that there is a meaningful level of path dependence in innovation at the industry level.

Figure 2.8: High-quality patents in each textile industry from 1853-1870

Figures 2.7 and 2.8 provide no evidence that the sharp increase in cotton related patents during the Civil War period was followed by a sustained higher level of patents in the years after the war. This pattern is not consistent with the existence of path dependence in innovation at the industry level.
I now analyze these patterns econometrically. To begin, it is interesting to look at the results obtained when I do not account for either of the sources of bias I have described. Table 2.2 presents results obtained using an approach which is similar to that employed by previous studies. In two of the four sets of results I obtain positive and statistically significant coefficients, which suggest path dependence in innovation, including for the PSTOCK measure of knowledge which is the closest to the measure used by previous studies. Based on either of these results, one might have concluded that there was strong evidence in favor of path dependence in innovation. For the other two measures, EXPER and PSUM, I also observe consistent positive coefficient estimates, but these are not statistically significant. A study using one of these measures would likely conclude that there was weak evidence suggesting path dependence in innovation at the industry level.

Table 2.2: Industry-level results obtained when not accounting for sources of bias

<table>
<thead>
<tr>
<th>Knowledge stock measure:</th>
<th>RECENT (1)</th>
<th>EXPER (2)</th>
<th>PSTOCK (3)</th>
<th>PSUM (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log lagged knowledge stock</td>
<td>0.778*** (0.239)</td>
<td>0.638 (0.405)</td>
<td>0.695*** (0.237)</td>
<td>0.626 (0.376)</td>
</tr>
<tr>
<td>Observations</td>
<td>56</td>
<td>68</td>
<td>56</td>
<td>68</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.921</td>
<td>0.907</td>
<td>0.917</td>
<td>0.908</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1. Included controls: input-specific fixed effects, year-specific indicator variables, constant. Included input types: cotton, wool, linen, and silk textiles.

Next, I want to calculate results using the new approach outlined in Section 2.5. Since I will be using an instrumental variables strategy, I begin by estimating first stage regression results based on equation (2.20) for each of my four measures of the stock of knowledge. Table 2.3 presents the results. POSTSHOCK_{ct} is a strong instrument for the stock of recent inventors (RECENT) and the stock of recent patents (PSTOCK). The relationship with the total stock of experienced inventors (EXPER) and the total stock of patents (PSUM) is somewhat weaker, but still positive and, at least for PSUM, close to statistically significant. These results suggest that POSTSHOCK_{ct}\footnote{Recall that POSTSHOCK_{ct} is a variable constructed by interacting an indicator variable for 1866-1869 with an indicator variable for the cotton textile industry.}
will generally be a good instrument for our measures of the stock of knowledge in each technology type.

Table 2.3: First-stage regression results at the industry level

<table>
<thead>
<tr>
<th>DV:</th>
<th>log(RECENT)</th>
<th>log(EXPER)</th>
<th>log(PSTOCK)</th>
<th>log(PSUM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post-shock x cotton indicator</td>
<td>0.253**</td>
<td>0.02997</td>
<td>0.3565***</td>
<td>0.0600*</td>
</tr>
<tr>
<td></td>
<td>(0.0967)</td>
<td>(0.04467)</td>
<td>(0.0905)</td>
<td>(0.03317)</td>
</tr>
<tr>
<td>Observations</td>
<td>56</td>
<td>68</td>
<td>56</td>
<td>68</td>
</tr>
<tr>
<td>F-statistic</td>
<td>61.62</td>
<td>711.72</td>
<td>66.69</td>
<td>1398.50</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Included controls: input-specific fixed effects, input-specific time trends, year-specific indicator variables, constant. Included input types: cotton, wool, linen, and silk textiles.

Table 2.4 presents the main results of the industry-level analysis. Column 1 presents a reduced form regression in which my instrument, an indicator variable for the post-shock period (1866-1869), is the key explanatory variable. Columns 2-5 present instrumental variables regressions in which the POST SHOCK variable is used as an instrument for each of the various measures of the knowledge stock. Since we may worry that these results are influenced by the discrete nature of the data, similar results calculated using Poisson regressions are available in Appendix B.1. Those results are similar to those described here.

In all of these results, I consistently estimate a negative and statistically insignificant coefficient for the relationship between the lagged stock of knowledge and the level of invention. Thus, there is no evidence of path dependence in innovation at the industry level in this context. This result is particularly strong given that we would expect any bias caused by persistent effects of the shock to be in favor of finding path dependence in innovation. These results differ significantly from those obtained without including input-specific time trends and without using the instrumental variables approach, as in Table 2.2. This highlights the importance of accounting for these sources of bias when looking for evidence of path dependence in innovation.
Table 2.4: Main industry-level regression results

<table>
<thead>
<tr>
<th>Knowledge stock measure:</th>
<th>Reduced form</th>
<th>Instrumental variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>DV: Count of patents</td>
<td></td>
<td></td>
</tr>
<tr>
<td>POSTSHOCK</td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Log lagged knowledge stock</td>
<td>-0.245 (0.705)</td>
<td>-1.308 (6.325)</td>
</tr>
<tr>
<td>Post-shock indicator 1866-1869</td>
<td>-0.0392 (0.181)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>68</td>
<td>56</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.935</td>
<td>0.945</td>
</tr>
</tbody>
</table>

|            | (3)          | (4)                    | (5)          |
|--------------------------|--------------|------------------------|
| RECENT                   | 0.174 (0.497) | -0.653 (3.000)         |
| EXPER                    |              |                        |
| PSTOCK                   |              |                        |
| PSUM                     |              |                        |

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Included controls: input-specific fixed effects, input-specific time trends, year-specific indicator variables, constant. Included input types: cotton, wool, linen, and silk textiles. The time period used is 1854-1870 except that, when using the RECENT and PSTOCK variables, the first three years of data must be dropped in order to generate the lagged stock variables.

In order to better understand how the new methodology I have introduced is affecting these results, I calculate additional results where I (1) control for input-specific time trends but do not use the instrumental variables approach and (2) use the instrumental variables approach without including input-specific time trends. The first four columns of Table 2.5 presents results in which I control for input-specific time trends without using the instrumental variables approach. Columns (5)-(8) present the results obtained when I use the instrumental variables approach without controlling for input-specific time trends. These results show that both of the methodological changes I have introduced are having some impact on the estimated coefficients, though they indicate that the inclusion of input-specific time trends is playing a more important role in generating the results observed in Table 2.4 than is the instrumental variables approach.
Table 2.5: Separate impact of methodological innovations on industry-level regression results

<table>
<thead>
<tr>
<th>Knowledge stock measure:</th>
<th>DV: Count of patents</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Input-specific time trends only</td>
</tr>
<tr>
<td></td>
<td>RECENT</td>
</tr>
<tr>
<td>Log lagged knowledge stock</td>
<td>(1)</td>
</tr>
<tr>
<td></td>
<td>-0.0366</td>
</tr>
<tr>
<td></td>
<td>(0.302)</td>
</tr>
<tr>
<td>Obs.</td>
<td>56</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
Included controls: input-specific fixed effects, input-specific time trends, year-specific indicator variables, constant. Included input types: cotton, wool, linen, and silk textiles. The time period used is 1854-1870 except that, when using the RECENT and PSTOCK variables, the first three years of data must be dropped in order to generate the lagged stock variables.

It is also possible to look at similar results generated using only high-quality patents: those for which the patent renewal fee was paid at year three to keep the patent in force. Because the renewal fee was expensive, only about 18% of patents were renewed beyond year three, making this a strong indicator of patent quality. When I calculate results using these renewal fee data, but without dealing with the sources of bias I have described (as in Table 2.2), I find evidence of path dependence in innovation. However, when I do account for these sources of bias I find no strong evidence in favor of path dependence in innovation at the industry level. These results are available in Appendix B.2.

One concern that we may have with the results above is that they include patents which appear in multiple industries because multiple input types are listed in the patent title. A substantial number of patents mentioning cotton, wool, linen, and silk in the title also mention at least one other input type. To deal with this concern, I calculate additional results which drop all patents that mention multiple input types. These results, available in Appendix B.3, are consistent with those described above. When I fail to account for the sources of bias I have identified, I find evidence of path dependence in innovation at the industry level. This evidence disappears as soon as I include input-specific time trends and implement the instrumental variables strategy.
Before leaving the industry-level analysis, it is also interesting to look at whether the increase in the stock of experienced cotton textile inventors in the post-shock period, caused by the shock, led to an increase in patents by experienced inventors. This allows us to consider whether path dependence may occur only for those inventors who gained experience in cotton textile innovation during the shock period. Figure 2.9 shows that the shock led to a sharp increase in the number of inventors who had filed a cotton textile technology patent in the past four years. As a result, the number of inventors with recent experience in this technology type remained high for several years after 1865. Figure 2.10 considers whether this higher level of experienced inventors resulted in additional patents by experienced inventors. Despite the larger number of inventors with recent experience in cotton textile technologies, this figure shows that there was no increase in the number of patents generated by these inventors. So, even among inventors who patented during the 1861-1865 period, there is little evidence that this led to additional patents after 1865. Table 2.6 confirms these observations. A simple regression of the level of innovation on the lagged stock of knowledge (column 1) provides evidence of path dependence in innovation. Once the input-specific time trends are added and the instrumental variables approach is used (columns 2 and 3) there is no longer any strong evidence of path dependence in innovation, even among inventors that gained experience during the shock period.
Figure 2.9: Inventors with experience in the past 4 years

Figure 2.10: Patents by inventors with experience in the past 4 years
Table 2.6: Patents by inventors with recent experience

<table>
<thead>
<tr>
<th>DV: Patents by inventors with recent experience</th>
<th>Baseline</th>
<th>Reduced form</th>
<th>Instrumental variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log stock of inventors with recent experience</td>
<td>1.154**</td>
<td>0.133</td>
<td></td>
</tr>
<tr>
<td>(RECENT)</td>
<td>(0.458)</td>
<td>(1.559)</td>
<td></td>
</tr>
<tr>
<td>Post-shock indicator 1866-1869</td>
<td>0.0344</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.406)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>55</td>
<td>55</td>
<td>55</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.864</td>
<td>0.868</td>
<td>0.870</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Included controls: input-specific fixed effects, year-specific indicator variables, constant. Included input types: cotton, wool, linen, and silk textiles.

2.6.2 Intermediate-level analysis

In this section, I search for evidence of path dependence in innovation at the level of particular types of technologies which are used in specific stages of the textile production process. In terms of the model, we can think of these as sets of technologies which are used to produce different intermediate inputs. This analysis focuses on the eight subcategories of the BPO spinning technology category described in Table 2.1. Of these, two technologies – gins and openers/scutchers – experienced a sharp increase in innovation rates during the 1861-1865 period because of their importance in dealing with the particular challenges of using the lower-quality Indian cotton in production.

My analysis of the intermediate-level innovation patterns begins with Figures 2.12 and 2.11. Figure 2.12 shows the count of patents in all of the eight available technology subcategories from 1855-1876. The most obvious feature of these graphs is that the significant increase in patents of gins and openers/scutchers during the 1861-1865 period is not found in any of the other technology subcategories, with the possible exception of carding machines. Carding machines, which was another technology used in cleaning the Indian cotton, also shows some increase.
detailed look at the level patents of gins and openers/scutchers. In both of these categories, there is some evidence of a continuing higher level of patents in the years just after 1865, a pattern that is consistent with path dependence in innovation at the intermediate level. Specifically, in gins, there is a broad shoulder of patents that extends into 1866 and 1867, while the level of openers/scutchers patents during 1866-1868 is high relative to either the pre-war years or the years after 1868.

Figure 2.11: Gins and Openers/Scutchers patents 1855-1876
Regression results at the intermediate level are generated using Poisson regressions. Poisson regressions are preferred when working with the patent subcategories data because the number of observations in each subcategory is relatively small, which increases the importance of the
discrete nature of the data. Alternative results produced using standard OLS and IV regressions, available in Appendix B.4, are similar to those I will describe below.

As in the previous section, I begin the analysis by calculating results without accounting for the sources of bias I have identified, in the spirit of previous studies. These results are shown in Table 2.7. These results are somewhat mixed, though for three of my measures of the knowledge stock I observe a negative relationship between the stock of knowledge and the level of innovation which is marginally statistically significant. Based on only these results, we would likely conclude that there is no clear evidence of path dependence in innovation at the intermediate good level.

Table 2.7: Intermediate-level results obtained when not accounting for sources of bias

<table>
<thead>
<tr>
<th>Knowledge stock measure:</th>
<th>RECENT</th>
<th>EXPER</th>
<th>PSTOCK</th>
<th>PSUM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lagged knowledge stock</td>
<td>0.00109</td>
<td>-0.000911**</td>
<td>0.00419**</td>
<td>-0.000753*</td>
</tr>
<tr>
<td></td>
<td>(0.00110)</td>
<td>(0.000417)</td>
<td>(0.00186)</td>
<td>(0.000403)</td>
</tr>
<tr>
<td>Observations</td>
<td>144</td>
<td>168</td>
<td>144</td>
<td>168</td>
</tr>
</tbody>
</table>

Poisson regressions
DV: Count of patents
Included controls: subcategory-specific fixed effects, year-specific indicator variables, constant. The time period used is 1856-1876 except that, when using the RECENT and PSTOCK variables, the first three years of data must be dropped in order to generate the lagged stock variables.

To implement the instrumental variables approach at the intermediate level, I focus on gins and openers/scutchers as the treated technology categories. I then construct an instrumental variable which takes the value 1 for the gins and openers/scutchers technologies during the years 1866-1869 and 0 otherwise. Using this variable, I obtain the first-stage regression results in Table 2.8. It is clear that this will be a strong instrument for all four of the knowledge stock measures.
Table 2.8: First-stage regression results at the intermediate level

<table>
<thead>
<tr>
<th>DV:</th>
<th>RECENT</th>
<th>EXPER</th>
<th>PSTOCK</th>
<th>PSUM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post-shock x gins/openers/scutchers indicator</td>
<td>34.6573***</td>
<td>15.9095***</td>
<td>25.6745***</td>
<td>20.1015***</td>
</tr>
<tr>
<td></td>
<td>(7.6632)</td>
<td>(5.4249)</td>
<td>(5.0354)</td>
<td>(4.3655)</td>
</tr>
<tr>
<td>Observations</td>
<td>144</td>
<td>168</td>
<td>144</td>
<td>168</td>
</tr>
<tr>
<td>F-statistic</td>
<td>24.26</td>
<td>648.87</td>
<td>21.23</td>
<td>803.07</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Included controls: subcategory-specific fixed effects, subcategory-specific time trends, year-specific indicator variables, constant. The time period used is 1856-1876 except that, when using the RECENT and PSTOCK variables, the first three years of data must be dropped in order to generate the lagged stock variables.

The primary econometric results at the intermediate level, presented in Table 2.9. My preferred results, generated using the instrumental variable strategy while including technology-specific time trends, are shown in columns 2-5 of Table 2.9, while column 1 presents the corresponding reduced form regression. These results indicate a consistent positive and statistically significant relationship between the stock of knowledge and the level of innovation. Thus, they are suggestive of the existence of path dependence in innovation at the intermediate level. However, these results do not provide conclusive evidence because of the potential for the persistent impact of the shock to positively influence the level of innovation in the gins and openers/scutchers subcategories in the post-shock period, which bias my results towards finding path dependence in innovation.
Table 2.9: Main intermediate-level regression results

<table>
<thead>
<tr>
<th>Knowledge stock measure</th>
<th>DV: Count of patents</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Reduced form</td>
</tr>
<tr>
<td></td>
<td>POSTSHOCK (1)</td>
</tr>
<tr>
<td>Lagged knowledge stock</td>
<td>0.0148**</td>
</tr>
<tr>
<td></td>
<td>(0.00657)</td>
</tr>
<tr>
<td>Post-shock indicator 1866-1869</td>
<td>0.385***</td>
</tr>
<tr>
<td>Observations</td>
<td>168</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
Included controls: subcategory-specific fixed effects, subcategory-specific time trends, year-specific indicator variables, constant. The time period used is 1856-1876 except that, when using the RECENT and PSTOCK variables, the first three years of data must be dropped in order to generate the lagged stock variables.

One feature of these results that is initially puzzling is that the IV regressions tend to suggest a stronger positive relationship between the stock of experienced inventors and the level of innovation than the OLS regression results. One potential reason for this is that the OLS results are also prone to a downward bias due to by repeated unobserved temporary shocks, which, because of the lagged effect of current innovation on the stock of experienced inventors, can generate a negative relationship between these variables. As was true of the positive bias generated by unobserved persistent shocks, the IV results should also be robust to the impact of unobserved temporary shocks as long as they are uncorrelated with the U.S. Civil War.

To separate the importance of the two methodological innovation I have introduced to these results, Table 2.10 presents results when I include only subcategory-specific time trends (columns 1-4) or use only the instrumental variables approach (columns 5-8). These results suggest that both of these innovations are playing a role in generating the consistent positive relationship I have estimated, though it appears that using the instrumental variables approach may be somewhat
more important in this setting than the addition of the subcategory-specific time trends.

As before, we may also worry that raw counts of patents in technology subcategories do not adequately represent the wide variation in the underlying quality of the technologies they represent. One way to address this issue is using data on the payment of patent renewal fees. Unfortunately, this further reduces the number of patent in each category. Still, it is worth looking at these data even if there are too few observations to warrant a full econometric analysis. The left panel of Figure 2.13 shows the count of high quality patents in the gins and openers/scutchers technology subcategories. In both of these categories there is some evidence that more high-quality patents were filed during the years just after 1865. For example, 3 high-quality gin patents were filed in 1866-1867 while only one high-quality gin patent was filed from 1856-1860. On average, five high-quality openers/scutchers patents were filed per year in the four years following the war, while in the four years preceding the war the average was just over two per year. This is more striking when one accounts for the fact that relatively few high quality patent were filed in the other six technology subcategories in the years just after 1865, as shown in the right panel of Figure 2.13.

Figure 2.13: Gins and Openers/Scutchers patents for which renewal fees were paid

It is also possible to look at how these patterns are playing out across the cotton, wool, linen, and silk textile industries. This is impossible to do with gins, which are almost exclusively used
Table 2.10: Separate impact of methodological innovations on intermediate-level regression results

<table>
<thead>
<tr>
<th>Knowledge stock measure:</th>
<th>Subcategory-specific time trends only</th>
<th>Instrumental variables approach only</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RECENT</td>
<td>EXPER</td>
</tr>
<tr>
<td>Lagged knowledge stock</td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td></td>
<td>0.00022</td>
<td>-0.0061***</td>
</tr>
<tr>
<td></td>
<td>(0.0014)</td>
<td>(0.0018)</td>
</tr>
<tr>
<td>Obs.</td>
<td>144</td>
<td>168</td>
</tr>
</tbody>
</table>

Poisson regressions

DV: Count of patents

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Included controls: subcategory-specific fixed effects, year-specific indicator variables, constant. The time period used is 1856-1876 except that, when using the RECENT and PSTOCK variables, the first three years of data must be dropped in order to generate the lagged stock variables.
in preparing cotton, but it is possible for openers and scutchers. Figure 2.14 plots the count of openers/scutchers patents in cotton textile technologies and in technologies related to the other major textile industries. This graph shows that the sharp increase in patents during the 1861-1865 period was exclusive to technologies related to cotton. Moreover, the increased level of openers/scutchers patents after 1865 seems to be driven entirely by a higher level of patents related to cotton. A few of these additional patents were related to other textile inputs as well, but when these cross-listed patents are removed we see no increase in patents that mention one or more of the other textile inputs but not cotton.

Figure 2.14: Openers/Scutchers patents for cotton and other textile industries 1855-1876

The results in this subsection are suggestive of path dependence in innovation which is primarily at the level of individual technologies. These results are subject to two caveats. First, they may biased because of persistent impacts of the Civil War that extend past 1865 and impact the incentives for innovation in particular technology categories. I expect that the direction of this bias is in favor of finding path dependence in innovation. Second, we may be worried that the higher level of patents found in the years just after 1865 are due primarily to innovation that took
place during the Civil War period. Again, this would bias the results in favor of finding evidence of path dependence in innovation. As a result, I interpret these results as providing suggestive but not conclusive evidence of path dependence in innovation.

2.7 Conclusion

This paper highlights three important shortcomings present in previous empirical studies of path dependence in innovation. First, I show theoretically that path dependence in innovation may occur at multiple technology levels. This implies that we should also look for path dependence at multiple levels. Second, I describe how, when using the econometric approach taken in existing studies, persistent shocks that affect the level of innovation will cause us to find evidence of path dependence in innovation when no such effect exists. Third, I show that the empirical approach used in previous studies is also susceptible to bias if there are underlying time-trends in innovation. The use of panel data and fixed effect regression can help control for such time trends at an aggregate level, but this approach is still susceptible to bias if there are trends in the innovation patterns of particular technology types or particular innovative firms, which are likely to be found in practice.

I offer a new approach that addresses these issues by (1) searching for path dependence at multiple technology levels, (2) using a temporary exogenous shock to the level of innovation as an instrument for the stock of knowledge in the post-shock period, and (3) using a panel data regression approach that includes time-trends which are specific to the cross-sectional units of the data. This approach is implemented using the large exogenous shock to the British cotton textile industry caused by the U.S. Civil War. Because this shock was large, exogenous, and largely temporary, it approximates the conditions needed in order to implement my approach, and where it departs from these conditions it is still possible to credibly assess the direction of the remaining bias.

I look for evidence of path dependence in innovation at the industry level by comparing cotton
textile technologies to those related to the wool, linen, and silk textile industries. At this level, I find no evidence of path dependence in innovation; the increase in the stock of knowledge caused by the additional innovation that took place in Britain during the U.S. Civil War does not appear to have generated more innovation during the post-war period. In contrast, I show that, had I used econometric the approach taken in previous studies, I would have found strong evidence of path dependence in innovation.

Within the cotton textile industry, I also look for evidence of path dependence at the intermediate level, by comparing innovation patterns in particular types of textile machinery. At this level, I find some evidence of path dependence in innovation which is consistent across multiple measures of the knowledge stock. Had I used the approach offered in previous studies, I could have found strong evidence either for or against path dependence in innovation, depending on the measure of the knowledge stock I used.

These results highlight the importance of the new econometric approach. The difference between the results I obtain when searching at the industry level and the intermediate level highlights the importance of looking at multiple technology levels. At both level, I find evidence that accounting for the sources of bias I have identified significantly affects the results that are obtained. Finally, it appears that both the inclusion of time-trends specific to the cross-sectional units of the data and the use of the instrumental variables approach are significantly affecting the results.
Chapter 3

Industry Connections and the Geographic Location of Economic Activity

3.1 Introduction

Marshall (1890) suggested that firms in different industries may benefit from locating near to one another, through localized inter-industry spillovers. He identified three channels through which these benefits could flow: input-output linkages, labor market pooling, or technology spillovers. Since then, inter-industry spillovers have been incorporated into theories of industrial linkages and development (Hirschman (1958), Ciccone (2008)), industrial clusters (Porter (1990)), the benefits of urban economies (Jacobs (1969)), the benefits of trade and FDI (Young (1991), Rodriguez-Clare (1996)), and the geography of economic activity (Krugman & Venables (1995)). Policy makers, too, have been influenced by these ideas. The existence of inter-industry spillovers is one of the prime motivations for local industrial policies, such as the tax incentives offered to firms by U.S. municipalities, or the widespread use of special industrial zones in developing countries. The cost of these policies can run into the hundreds of millions of dollars for single U.S. municipalities.¹

Despite this interest, and widespread application, empirical evidence on the role of inter-

¹See Greenstone & Moretti (2004).
industry connections in influencing the geographic location of economic activity is sparse. This is due largely to the difficulty of measuring the patterns of connections between industries, though these measures are improving.\textsuperscript{2} In an important recent study, Greenstone et al. (2010) show evidence of localized productivity spillovers between plants that share similar labor or technology pools.\textsuperscript{3} In a similar vein, Javorcik (2004) finds evidence of spillovers from FDI firms to upstream suppliers.\textsuperscript{4} More related to the current study is Ellison et al. (2010), which uses measures of input-output connections, labor force similarity, and technological spillovers, to provide evidence that the underlying pattern of connections between industries is influencing the geographic location of industries. While these are important contributions, they do not address the most relevant question for policy: can temporary changes in the local availability of inter-industry spillovers, such as those created by economic shocks or policy interventions, have long-term impacts on the geographic location of economic activity?

This study takes the next step, by utilizing a large temporary external shock that altered the spillovers available to certain industries in certain locations, in order to provide causal evidence that changes in the availability of inter-industry spillovers can influence the long-run distribution of economic activity. The shock was caused by the U.S. Civil War and impacted the British economy in the 1860s. During the 19th century, cotton textile production was the largest British manufacturing industry, and cotton textiles were Britain’s largest export. This industry relied entirely on imports of raw cotton, a vital input. Prior to the onset of the U.S. Civil War in 1861, the majority of raw cotton came from the Southern U.S., but the war, and the corresponding Union blockade of Southern ports, sharply reduced raw cotton supplies. The result was a four year depression in the industry, lasting from roughly 1861-1865. My focus will be on how this affected the location of those industries related to (sharing spillovers with) the cotton textile industry.

As with other studies using historical examples (e.g., Donaldson (2010)), the event considered

\textsuperscript{2}There are a number of early studies which looked for the impact of inter-industry connections without accounting for the patterns that these connections take. Important contributions include Glaeser et al. (1992), Kim (1995), Henderson (1997), and Henderson (2003). There is also a much larger literature considering the role of spillovers within industries.

\textsuperscript{3}Another recent contribution is Bloom et al. (2012).

\textsuperscript{4}Similar evidence is also provided by Kugler (2006), Poole (forthcoming), and Balsvik (Forthcoming).
in this study was chosen because it has features that are particularly helpful in identifying the effects of interest, which are unlikely to be found in similar modern events. First, the shock was large, exogenous, and temporary. While the cotton shortage reduced production by about half during the shock period, the cotton textile industry rebounded quickly, attaining its original growth path by roughly 1866-67, only a year or two after the end of the war. Second, the direct effects of the shock were largely industry-specific. British import and export data suggest that, once cotton textiles are removed, the shock had no major effect on other British manufacturing sectors. Thus, changes observed in those industries related to cotton textiles can be attributed to the shock, transmitted through inter-industry connections. Third, despite the large magnitude of the negative effects of the cotton shortage, there was virtually no government intervention. This was largely due to the very strong free-market ideology that was dominant in Britain at the time, particularly in the northern industrial districts that were hardest hit by the recession.

A final important feature is that some locations were severely impacted, while other, economically similar locations, were left nearly untouched. This study compares outcomes in towns from two industrial counties in the north of England, Lancashire and Yorkshire. Lancashire was the heart of Britain’s cotton textile industry at the time. Yorkshire, lying just to the east, was similar to Lancashire in many ways. The key difference between these two counties was that, while towns in Yorkshire also had large textile industries, Yorkshire producers focused primarily on wool-based textiles (woollens & worsted) rather than cotton. Thus, while towns in Lancashire were severely affected by the cotton shortage, towns in Yorkshire were not negatively impacted. Comparing industry growth rates from towns in these two counties thus allows me to better identify the impact of the shock.

The basic hypothesis that I test is that the shock negatively affected employment and employment growth in industries more closely related to the cotton textile industry, in locations more severely impacted by the shock, by reducing the spillover benefits available to these industries. Importantly, I focus on impacts occurring in the years and decades after the Civil War had ended and the cotton textile industry had rebounded. The hypothesis is motivated by a two-country dy-
namic Ricardian trade model building on work by Young (1991) and Matsuyama (1992). In the model, technology growth is driven by localized learning-by-doing spillovers within and between industries. A negative shock to one industry reduces the spillover benefits enjoyed by related industries, in the location in which the affected industry is concentrated. The result is a reduction in technology growth in the related industry, in the more severely affected location relative to the less affected location. Moreover, the model describes channels through which a loss of spillovers in one period can affect employment and employment growth in related industries in future periods. The model is used to derive the empirical specification used in the analysis.

This intuition is perhaps best illustrated using an example from the empirical setting, provided by the Engine & Machinery industry (E&M). This was an important industry in Britain at this time, and one that was connected to the cotton textile industry in Lancashire, as well as the wool textile industry in Yorkshire, through all three of Marshall’s channels. In fact, Marshall himself used the textile and engineering industries to illustrate the possibility of labor market pooling benefits.\(^5\) There is also evidence suggesting that textile machine makers learned from the nearby textile producers that they supplied. Engine and machine makers in Lancashire and Yorkshire competed, both to supply customers in these locations, as well as in the important export market. The data show that the E&M industry had a similar growth path in the two locations prior to the shock, but that E&M firms in Yorkshire towns gained an advantage relative to Lancashire producers in the following decades, allowing them to expand employment more rapidly. This suggests that the recession in the cotton textile industry had persistent impacts on distribution of the E&M industry across locations.

The empirical strategy used to test this hypothesis involves using panel data with two cross-sectional dimension, allowing me to compare impacts across time (pre vs. post shock periods), locations (towns with higher vs. lower shock intensity), and industries (more vs. less related to cotton textiles). The primary data are drawn from the British Census of Production, which were gathered from original sources. These data provide employment disaggregated into 171 industry categories.

\(^5\)Marshall (1920) (p. 226).
groups, spanning nearly the entire private sector economy, available for every ten years from 1851-1891. Thus, I have multiple observations in both the pre- and post-shock period, and are able to observe effects up to 25 years after the end of the recession. These data are available for 11 principal towns, 6 in Lancashire, and 5 in Yorkshire, providing the geographic dimension of the analysis. Additional data from local Poor Law boards, which were the primary source of funds for unemployed workers during this period, allow me to measure the severity of the shock in each town. I find that the share of cotton textiles in a town’s employment in 1851 is a good predictor of the severity of the shock in each location. Thus I am able to strengthen my identification strategy by using each town’s cotton textile employment in 1851 as an instrument for shock intensity.

An important input into the analysis is a measure of the pattern of connections between the cotton textile industry and other industries. Two measures are used. The first is based on the degree to which each industry is geographically coagglomerated with the cotton textile industry, following work by Ellison & Glaeser (1997). Ellison et al. (2010) show that this measure is related to measures of input-output linkages, labor market pooling potential, and technology spillovers. The second measure is an intermediate goods input-output matrix based on Thomas (1987).

The results suggest that industries more closely related to the cotton textile industry, based on the geographic coagglomeration measure, suffered lower employment and employment growth, in more severely affected towns, in the post shock period. The impacts are estimated while controlling for aggregate industry-level and town-level shocks in each year, as well as time-invariant industry-location factors. The impact of the shock on employment and employment growth continues to appear through the 1881-1891 period. The implication is that inter-industry connections can play an important role in affecting the geographic location of economic activity across locations. Moreover, temporary changes in the availability of these connections appear to have the potential to generate long-lasting effects. There appear to be no persistent effects related to the intermediate goods input-output matrix, suggesting that the observed effects are being driven by other types of inter-industry connections.

This project is related to three additional strands of literature. On set of related literature stud-
ies whether temporary shocks have long-term impacts on the geographic location of economic activity, in order to test economic geography models which predict multiple equilibria in the location of economic activity. Most studies focus on the impact of temporary shocks caused by war on city size. Only two studies consider the impact on the location of industries, and these deliver mixed results. This study improves on previous work by (1) considering the role of industry connections and (2) using data that are more detailed and comprehensive. The results of this study are consistent with the existence of multiple equilibria in the geographic location of economic activity, providing support for economic geography models predicting multiple equilibria.

Second, because this project considers the impact of a large temporary trade shock on industry growth over a relatively long time frame, it can help inform our understanding of the relationship between trade and growth. While there is a large empirical literature studying the overall relationship between trade and growth, there are relatively few studies considering the impacts of temporary trade shocks. The contribution of this study is to show that temporary trade shocks can have long-term economic impacts. Moreover, while I have focused on two regions with a country, in order to better identify the effects of interest, my results raise the possibility that trade shocks may also result in reallocations of industries across countries, which could have long-term impacts on country growth rates and the gains from trade. This suggests that a complete understanding of the relationship between trade and growth should take into account the potential effects of increased (or decreased) susceptibility to trade shocks, in addition to the more direct effects of increased overall trade.

Finally, because I study the textile industry in 19th century Britain, there is a natural connec-

---

7Davis & Weinstein (2002), Brakman et al. (2004), Bosker et al. (2007), Bosker et al. (2008b), and Bosker et al. (2008a). In related work, Miguel & Roland (2006) consider the impact of bombing in Vietnam on welfare outcomes.
8Davis & Weinstein (2008) study the effect of the WWII bombing of Japanese cities on the location of eight highly aggregated industrial sectors (e.g., machinery, metals) and find no evidence of persistent effects. In contrast, Redding et al. (forthcoming) study the impact of the division of Germany following WWII on one particular industry, airport hubs, and find evidence of a persistent effect.
9For example, this study considers 171 industry groups, while Davis & Weinstein (2008) consider only eight aggregated industrial sectors and Redding et al. (forthcoming) consider only one, airport hubs.
10See Rodriguez & Rodrik (2000) for a review and critique of recent macroeconomic literature on trade and growth.
11Recent work by di Giovanni & Levchenko (2009) suggests that trade liberalization may increase volatility in an economy.
tion to the debate over the sources of the Industrial Revolution. In particular, there is disagreement over the importance of trade in generating British economic growth.¹² The results of this study support the view, due to Allen (2009) and others, that trade propelled technological progress in the engineering and machinery industries, primary drivers of the industrial revolution, through their connections with the large textile industries.¹³

The next section provides background information on the empirical setting and describes the shock, while Section 3.3 describes the data used. Section 3.4 studies the impact on one industry, Engine & Machine manufacturing. Section 3.5 introduces a model that is used to guide the econometric analysis, which is presented in Section 3.6. Section 3.7 concludes.

### 3.2 Empirical setting

Lancashire County, in the Northwest of England, was the heart of the British cotton textile industry from the end of the 18th century and the cradle of the Industrial Revolution. For the purposes of this study, Cheshire, a smaller cotton textile producing county just south of Lancashire, is treated as part of Lancashire. Lancashire’s main commercial city, Manchester, became synonymous with the cotton textile trade, and the main port, Liverpool, served as the center of the world’s raw cotton market. Yorkshire, the large and historically important county just east of Lancashire had followed Lancashire’s lead in industrialization and was similar to Lancashire in many respects. This study focuses only on the West Riding area of Yorkshire, which is the main industrial area of the county.¹⁴ Figure 3.1 shows the location of these counties in England (left panel) and highlights the principal towns (right panel), over which the analysis will be conducted. Figure 3.2 shows that, while Lancashire was somewhat larger than Yorkshire during the study period, the two

---

¹²Authors such as Deane & Cole (1967) and Rostow (1960) emphasized that trade, by generating the demand which allowed the expansion of the cotton textile industry, was a key “engine of growth”. Mokyr (1985) (p. 22) and McCloskey (1994) (p. 255-258) have disputed this view, arguing that trade, while helpful, did not drive productivity growth, and that gains from trade were relatively small.

¹³Among the most related industries to cotton textiles, based on the geographic coagglomeration measure, are “Engineering and Machinery”, “Millwrights”, and “Iron and Steel”. This argument is made for industrialization more generally by Ciccone (2002).

¹⁴This practice is common for historians studying Yorkshire during this period, because it separates the more modern industrial area of the West Riding from the less developed economies of the North and East Riding.
counties followed similar population growth paths and had relatively similar industrial structures.

However, there was one key difference between Lancashire and Yorkshire. While Lancashire produced cotton textiles, producers in Yorkshire concentrated on wool textile goods (woollen and worsted), industries with a long history in the area. Table 3.1 describes total private employment, and employment in the cotton and woollen/worsted industries in Lancashire and Yorkshire, in 1851, the beginning of the study period. Cotton textile employment made up over 29% of private sector employment in Lancashire, but only 4% in Yorkshire, while woollen/worsted employment made up 1% of private sector employment in Lancashire but over 30% of employment in Yorkshire.\textsuperscript{15}

Figure 3.1: Maps of the study area

\textsuperscript{15}In 1861, just before the onset of the shock, 85% of cotton textile manufacturing workers in England and Wales were located in Lancashire, while 72% of all of the woollen textile manufacturing workers, and 90% of worsted textile workers, were located in Yorkshire.
Despite using different input materials, the cotton and woollen/worsted textile industries were largely similar. For instance, they shared the same three basic production stages. Stage one involved spinning the raw input into yarn. In the second stage, yarn was woven into fabric, while in the final stage the fabric was finished, which often included bleaching, dying, and printing. As a result of this similarity, many technologies developed for one of the industries were also adapted to work in the other. For example, spinning and weaving machines first developed for cotton were modified for use in wool and other textile production. Moreover, as these industries had grown to prominence, first for cotton in Lancashire and later for woollen/worsted in Yorkshire, a large

---

number of subsidiary trades had grown up around them. These related industries included textile machinery producers, chemical manufacturers to produce dyes and bleaches, engineering firms to produce the steam engines that powered the plants, and other textile industries that took advantage of the technological innovations made in the cotton textile industry.\footnote{Farnie (2004) provides a thorough discussion of the ancillary industries. See also Timmins (1998).}

### 3.2.1 Impact of the U.S. Civil War

Prior to 1861, most of the raw cotton used in England’s textile mills was grown in the Southern U.S. The onset of the U.S. Civil War in 1861 created a major disruption of raw cotton supplies.\footnote{A figure showing this drop is available in the Appendix.} While other major suppliers, such as India and Egypt, did increase production, they were unable to adjust rapidly enough to make up for the sharp fall in U.S. exports.\footnote{Furthermore, the cotton produced by these other suppliers, India in particular, was often of a shorter fiber variety that was an imperfect substitute for the high quality long-fiber U.S. cotton. Often, producers were required to mix more expensive American cotton with the Indian cotton in order to make it strong enough to spin. Thus, the fall in imports is likely to understate the impact of the shock on cotton supplies. See Henderson (1969) (pp. 50-51).} Raw cotton prices responded to the tightening of supply by rising. The left-hand panel of Figure 3.3 shows a dramatic spike in cotton prices starting in 1861 and continuing through 1865. Notably, the price increase in 1861 was actually fairly small. J.C. Ollercenshaw (1870, p.112), remarked in his presentation to the Manchester Statistical Society, that, “The American War commenced on April 5th, 1861, but for many months it had little effect on commerce - being generally regarded as merely temporary...” This reflects the commonly held belief that the U.S. Civil War was going to be short and cause relatively little disruption to economic activity. Margins also suffered during the period, with margins on spun cotton yarn becoming negative in 1862 and not recovering pre-shock levels until 1866.\footnote{Data from Forwood (1870). See also, Helm (1869).} In response to the curtailment of supply, rising prices, and falling margins, production in the cotton textile industry fell. One of the best indicators of output in the industry is raw cotton consumption, described in the right-hand panel of Figure 3.3. To summarize, the onset of the U.S. Civil War reduced cotton imports, increased prices, and decreased output. The effects started
in 1861, peaked in 1862-3, and persisted through 1865.  

Figure 3.3 also shows that by the late 1860s, the cotton textile industry had returned to its original growth path. Its expansion then continued with only relatively minor interruptions until WWI. The recovery of production in the cotton textile industry is an important feature of the story, because it suggests that any long-term effect that the shock had on related industries had its roots in the shock period, rather than in persistent changes to the cotton textile industry.

Figure 3.3: British Cotton Imports and Raw Cotton Prices 1815-1910

To argue that the U.S. Civil War created a shock that was primarily industry-specific, I look for evidence of other large direct effects of the war. I would expect any such effects to occur either

---

21 Ollerenshaw (1870) stated that, "...we may consider 1862-5 as the years during which the effect of the American war were really experienced," adding later that, "The year 1866 is to be regarded as an exceptional year [for cotton manufacturers], equally with 1861. The war was over but prices had fallen only 1d to 1.5d per lb."


23 It is worth noting that the impact of the cotton shortage on manufacturers and traders in cotton textiles was more complex than it appears from the foregoing graphs. A number of manufacturers and traders made enormous profits at the beginning of the shock period, by selling stocks of cotton goods or raw cotton at much higher prices. However, later in the shock period manufacturers often found their resources severely strained (See Henderson (1969) (pp. 19-20)). For the rest of England, the early 1860s was generally a period of prosperity, though there were financial crises in 1864 and 1866 as a result of a decline in English gold reserves. This loss of reserves was due in part to a switch from purchasing raw cotton from American in exchange for manufactured goods to purchasing from India, Egypt, or Brazil, where goods were generally purchased with gold or silver. Importantly though, these financial crises affected England as a whole, and should therefore not generated differential impacts between Lancashire and Yorkshire, except to the extent that Lancashire banks and firms were more vulnerable due to the effects of the shock. The late 1860s, on the other hand, was a period of poor economic performance throughout England, with numerous bankruptcies among financial, commercial, railway, and manufacturing firms, a hangover from the earlier financial crises.
through import or export channels. However, British import data from Mitchell (1988) suggest that, once raw material for textiles are removed, British imports show no noticeable effect from the U.S. Civil War. Similarly, once textile exports are excluded, British exports also show no negative effects.24 I may be particularly concerned that the areas I study were directly affected by the U.S. Civil War through armament industries. However, none of the three armaments categories included in the data, “Guns”, “Ordinance”, and “Ships”, make up more than 0.2% of employment in any of towns included in this study, during the study period, with the exception Liverpool, which is not included in the analysis for this reason.25

Unlike cotton, the woollen and worsted industries show little effect from the shock.26 It is not surprising that imports of raw wool are unaffected, since most of these imports come from Spain, Australia, South Africa, or South America. While there was some effect on wool textile exports to the U.S., this was made up for by an increase in exports to European markets during the period, particularly to France following a new trade agreement in 1860, and through increased demand for military woolens.27

Given that I observe a strong negative shock to the cotton textile industry but little direct effect on the woollen and worsted industries, I expect the direct effects to be much stronger in Lancashire than in Yorkshire. This is confirmed by Figure 3.4, which shows the number of able-bodied workers seeking relief from local Poor Law Boards as a fraction of the total 1861 population, in Lancashire and Yorkshire, over the shock period.28 These Poor Law Boards were the main apparatus through which relief was provided to paupers and the unemployed in England during this period.29

---

24 Graphs of British import and export data are available in the Appendix.
25 Shipbuilding made up roughly 2-4% of employment in Liverpool during the study period, and Liverpool was active in providing ships used in the Civil War.
26 Data from Mitchell & Deane (1962). Graphs are available in the Appendix.
27 Helm (1869) states, “in some of our foreign markets, linen and woolen goods have, as at home, taken the place of cotton.” See also Jenkins & Ponting (1987) (p. 157-163) for a broad discussion of the effect of the U.S. Civil War on wool and worsted textiles. While the U.S. imposed an additional tariff on wool textile imports under The Morrill Tariff Act of 1861, wool textile exports to the United States rose steadily during the 1860’s (Jenkins & Ponting (1987) p. 157-163) due to weak competition from domestic suppliers and the need for heavy woolen goods for military uniforms and supplies.
28 Data are drawn from several of the major districts within each county, but do not cover the entire county.
29 Southall (1991) shows that Lancashire had the highest rate of pauperage of any English county in 1863, with
To summarize, while the U.S. Civil War generated an enormous negative shock to cotton textiles, a shock that fell heavily on Lancashire, it appears that other direct effects on the British economy were of limited size. In contrast to the experience of Lancashire, Yorkshire’s mainstay wool textile industry shows little effect from the shock, and may have actually benefited somewhat from substitution away from higher priced cotton textiles. It is this large differential impact that is exploited in this study in order to pinpoint the effects of the shock on those industries that were related to cotton textiles.

Figure 3.4: Able-bodied relief seekers as a share of 1861 population

Data from Southall et al. (1998).

### 3.3 Data

The primary database used in this study is drawn from the the British Census of Population. These decennial Censuses are the best available source of information on the structure of the British economy over the period of this study, 1851-1891. For each census, occupation data was collected from respondents by trained registrars, and each census report presents summaries of 10.3 percent of the population receiving Poor Law Relief, while the West Riding was among the lowest. A similar pattern holds when Southall looks only at able-bodied relief seekers. In contrast, in other cyclical downturns occurring during the study period, such as those 1868 and 1879, Southall finds that relief rates in Lancashire and Yorkshire were similar.

See Lee (1979) p. 3.
employment by occupational category.\textsuperscript{31} The number of occupational categories in these reports varies somewhat over the period studied, with a high of 478 categories in 1861 and a low of 348 in 1891.

One feature of these occupational categories is that they generally closely correspond to industries. For example, there are occupational categories such as “Cotton textile manufacturer”, “Chair maker”, and “Nail manufacture”. This feature allows me to treat occupational categories as industries. Over the four decades covered by this study, some adjustments were made to the reported occupational categories. Linking these categories over time was a time-consuming task that eventually yielded occupations gathered into 234 occupational groups (hereafter, just “groups” or “industries”) of which 204 groups are available for all years.\textsuperscript{32} Of these, 171 private sector occupational categories are used in the analysis.\textsuperscript{33}

Data are available at three levels of geographic specificity. The most specific is the town level, for towns with populations over 50,000. My main analysis takes advantage of 11 towns for which consistent data can be obtained for 1851-1891.\textsuperscript{34} Six of these towns are located in Lancashire and five are in Yorkshire. Data are also available at the district level. These districts are larger than towns and data is available for the entire area of the two counties, but only for 1851-1861. Data are also available at the county level.

The towns used in the analysis are listed in Table 3.2, together with town population, a measure of the intensity of the shock in each location based on the increase in relief seekers as a share of the 1861 population, and the share of cotton textile employment in each town in 1851.\textsuperscript{35} The primary measure of the severity of the shock in each town will be the percentage point increase

\textsuperscript{31}Woolard (1999) compares these summary tables to the original enumerators books from the Isle of Man in 1881 and finds that “the original process of classification was remarkably accurate in light of the rules applied” (p. 29), particularly in the manufacturing and industrial categories.

\textsuperscript{32}Scholars familiar with these data have noted that it is nearly impossible to perfectly match occupational categories over time due to changes in occupational classifications (see, e.g., Lee (1979)). Even within the aggregated industry groups that were constructed for this project, it is sometimes not the case that employment values are comparable over time. However, the differences-in-differences approach used in this study relies on comparisons across locations within years, so the results should be robust to most changes in occupational classifications.

\textsuperscript{33}These 171 industries exclude 6 cotton based industries as well as government occupations and occupations that do not correspond to industries, such as “Labourer”.

\textsuperscript{34}Details related to the construction of the database are available in the Appendix.

\textsuperscript{35}A map showing the location of these towns is available in Appendix.
in able-bodied workers seeking relief over the shock period, based on data from local Poor Law Boards, shown in the fourth column of Table 3.2. I will also want to apply an instrument for the shock severity measure, for reasons discussed later. The instrument will be the share of cotton textile employment in total employment in each town in 1851, which are shown in the last column of Table 3.2.

Table 3.2 reveals significant variation in cotton employment and the intensity of the shock across towns. Even within Lancashire, there is variation in the impact of the shock, despite the fact that all towns were major cotton textile centers. Contemporary reports suggest that some of this variation was due to local specialization in different product categories (e.g., heavy vs. fine cotton fabrics) or stages of the production process (e.g., spinning vs. weaving), which were impacted differently by the shortage of raw cotton. Still, a casual glance at the last two columns Table 3.2 suggests that those towns in which cotton textiles made up a significant share of production in 1851 tend to have experienced a more intense shock effect. This relationship can also be tested more formally; regression results are presented in Table 3.3. These results indicate that 1851 cotton textile employment share will be a strong instrument for the severity of the shock in each town.

Table 3.2: Towns used in the analysis

<table>
<thead>
<tr>
<th>County</th>
<th>Town</th>
<th>1861 Pop.</th>
<th>Increase in relief seekers/1861 pop.</th>
<th>Cotton emp. /1851 pop.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lancashire</td>
<td>Blackburn</td>
<td>63,126</td>
<td>7.75%</td>
<td>34.00%</td>
</tr>
<tr>
<td></td>
<td>Bolton</td>
<td>70,395</td>
<td>1.55%</td>
<td>22.59%</td>
</tr>
<tr>
<td></td>
<td>Manchester</td>
<td>460,428</td>
<td>8.96%</td>
<td>11.63%</td>
</tr>
<tr>
<td></td>
<td>Oldham</td>
<td>94,344</td>
<td>3.59%</td>
<td>5.01%</td>
</tr>
<tr>
<td></td>
<td>Preston</td>
<td>82,985</td>
<td>9.39%</td>
<td>25.12%</td>
</tr>
<tr>
<td></td>
<td>Stockport*</td>
<td>54,681</td>
<td>3.70%</td>
<td>12.63%</td>
</tr>
<tr>
<td>Yorkshire</td>
<td>Bradford</td>
<td>106,218</td>
<td>-0.12%</td>
<td>0.58%</td>
</tr>
<tr>
<td></td>
<td>Halifax</td>
<td>37,014</td>
<td>0.29%</td>
<td>0.71%</td>
</tr>
<tr>
<td></td>
<td>Huddersfield</td>
<td>34,877</td>
<td>0.32%</td>
<td>1.28%</td>
</tr>
<tr>
<td></td>
<td>Leeds</td>
<td>207,165</td>
<td>0.34%</td>
<td>0.06%</td>
</tr>
<tr>
<td></td>
<td>Sheffield</td>
<td>185,172</td>
<td>0.85%</td>
<td>0.04%</td>
</tr>
</tbody>
</table>

* Stockport, Cheshire is treated as part of Lancashire in this study
Poor Law relief data from Southall et al. (1998).

---

36 See Arnold (1864) (p. 102) and Henderson (1969) (p. 2). A map showing the distribution of cotton textile manufacturing activities across Lancashire towns is available in the Appendix.
Table 3.3: Relationship between cotton textile employment and shock intensity

<table>
<thead>
<tr>
<th></th>
<th>Shock intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(increase in relief seekers/ 1861 pop.)</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>Cotton textile</td>
<td>0.728**</td>
</tr>
<tr>
<td>employment share 1861</td>
<td>(0.228)</td>
</tr>
<tr>
<td>Constant</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>11</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.530</td>
</tr>
<tr>
<td>Standard errors in parentheses</td>
<td>*** $p&lt;0.01$, ** $p&lt;0.05$, * $p&lt;0.1$</td>
</tr>
</tbody>
</table>

While the data set spans 171 industries and 11 towns, not all industries have positive employment in all towns in all years. Because I work with log employment, location-industries with no employment will result in missing entries. For the analysis, I will include only location-industries that have positive employment levels for all five years. The result will be an unbalanced panel with 1,543 complete location-industries. In essence, this analysis will be on the intensive margin of employment, i.e., changes in employment levels in industries present in a location, while ignoring the extensive margin, i.e., industries emerging or disappearing in a location. This makes sense given that my focus is on the role of inter-industry spillovers, which presumably exist primarily when both industries are present in a location.

### 3.3.1 Industry relatedness

One important input into the analysis is a measure of relatedness between the cotton textile industry and other industries in the economy. Two approaches are taken to measuring the pattern of relatedness. The first uses the British Census data to construct a relatedness measure based on the geographic coagglomeration of industries. This measure is inspired by Ellison & Glaeser (1997), as well as, Ellison et al. (2010), who find that geographic coagglomeration patterns are correlated with measures of technological spillovers, occupational similarity, and input-output flows.\(^{37}\)

---

\(^{37}\)One critique of this approach is Helsley & Strange (2010), who show that coagglomeration patterns across locations may be inefficient. This should, if anything, cause Ellison et al. (2010) to understate the strength of the
Thus, the advantage of this measure is that it may reflect many forms of relatedness, though this comes at the cost of not being able to identify the particular types of industry connections that matter the most. Geographic coagglomeration is measured for each pair of industries and reflects the propensity of the two industries to concentrate production in the same location, where concentration implies that the size of the industry is in excess of the size that would be predicted given the location’s overall size (population). Specifics of the calculation of the coagglomeration measure are available in the Appendix.

It is important to note that the geographic coagglomeration measure is calculated using the district level census data, which are different from the town-level data used as the primary outcome variable. The district-level data are significantly more geographically comprehensive than the town-level data, giving me more regions to work with (71 districts vs. 11 towns). Furthermore, I can calculated an alternative coagglomeration dropping all of the districts containing towns available in the town-level data and show that this alternative coagglomeration measure delivers similar results.

One way to check the reasonableness of the geographic coagglomeration relatedness measure is to consider the least and most related industries. At the top of the related list is cotton textile production, followed by cotton textile finishing. Other textile industries, such as woollen, worsted, thread, and miscellaneous weaving, are also among the top 15 most related. This may reflect technological spillovers, labor market pooling, or other forms of inter-industry connections. The third most related industry is paper manufacturing, which may seem odd at first, but an important input for this industry at the time was waste cotton. The sixth most related industry, coal mining, produced the second most important material input for the cotton textile industry. Engine and machine makers, an industry discussed in more detail in Section 3.4, ranks thirteenth. Toll collector and road construction also rank among the top 15, reflecting the importance of road transport in the cotton districts. Among the fifteen least coagglomerated industries, there are none that one would naturally expect to be related to cotton textiles. Examples include ship transport, sugar relationship between their measures of inter-industry connections and coagglomeration patterns.

38A table showing the most and least coagglomerated industries is available in the Appendix.
refining, cooper, lodging-house keeper, and tobacconist. Overall, it appears that coagglomeration is providing me with a reasonable measure of relatedness between industries.

A coagglomeration measure of relatedness to the wool textile industry, calculated using data from Yorkshire, gives relatedness patterns that are fairly similar to that observed for cotton in Lancashire. The correlation between these two coagglomeration measures is 0.3.\(^{39}\)

The second measure of relatedness is based on an input-output matrix for intermediate goods. This matrix is based on one constructed by Thomas (1987), and divides the economy into 42 industries.\(^{40}\) The primary source used by Thomas to construct his input-output matrix was the 1907 Census of Production, Britain’s first industrial census, though he also drew on a wealth of supplementary information. Because this input-output matrix was constructed using data from after the study period, there is some worry that input-output relationships may have changed between the beginning of the study period and the 1907 Census of Production. One way to test for such changes is to compare this input-output matrix to a less detailed matrix constructed by Horrell et al. (1994) for 1841, which divides the economy into 17 categories. Once the categories in these matrices are matched, I find that the correlation between their entries is high (>|.96), giving me confidence that the input-output matrix is relevant for the period covered by this study.

### 3.4 An Example: Engine & Machine Makers

This section explores the impact of the shock on one related industry, “Engine & Machine Makers” (E&M).\(^{41}\) There are two reasons to give the E&M industry special attention. First, a number of scholars have argued that this industry played a central role in the industrial revolution, and in Britain’s economic success throughout the 19th century.\(^{42}\) Second, this industry appears to be closely related to the cotton textile industry in Lancashire, and also to the woollen and worsted

---

\(^{39}\)This figure exclude cotton, woollen, and worsted textile manufacturing.

\(^{40}\)Details of the adjustments made to the original Thomas (1987) input-output matrix are available in the Appendix.

\(^{41}\)Other highly related industries display similar patterns to those described for the E&M industry in this section, including “Worsted Manufacturing”, “Stone Workers”, “Brick Making”, “Dying and Calendering”, “Millwrights”, “Iron Manufacture”, “Hoisery”, “Silk”, and “Flax and Linen”.

\(^{42}\)For example, Allen (2009) argues that “the great achievement of the British Industrial Revolution was, in fact, the creation of the first large engineering industry that could mass-produce productivity-raising machinery.”
industries in Yorkshire. Farnie (2004) writes that “Textile engineering became the most impor-
tant of all the ancillary trades [to cotton textiles]. Its light engineering section supplied spinning
machines and looms and a whole succession of related equipment, while its heavy engineering in-
dustry supplied steam engines, boilers, and mechanical stokers.” The coagglomeration relatedness
measure indicates a high level of relatedness between textile manufacturing and E&M. According
to this measure, E&M is the 9th most related industry to cotton textiles in Lancashire (out of the
171 included in the analysis dataset), and the 26th most related to wool textiles in Yorkshire. The
intermediate goods input-output matrix does not show significant flows between these industries,
since the E&M industry primarily produced capital goods.

Connections between the E&M and cotton textile producers likely took multiple forms. The
most obvious connection was direct backward demand linkages from textile producers to the firms
supplying their machinery, at least for a subset of E&M firms. Two-way knowledge and technol-
ogy flows between these industries may have also been important, though evidence of such flows
is necessarily sparse. One indication of their importance is the fact that machinery firms often
specialized in machinery catered to the needs of producers in their local area. For example,
Bolton machine makers dominated the market for machinery to spin fine thread counts, which was
mainly produced in the Bolton area, while Oldham producers were dominant in the machinery
for spinning of heavier thread counts, where Oldham producers dominated production. Given the
relatively close proximity of these locations, there seem to be few reasons, other than knowledge
flows between textile and E&M producers, that explain this specialization pattern. Furthermore,
it is well documented that operators in the textile mills were active in “tweaking” their machines,
to the point where, for example, “no two pairs of [spinning] mules worked in precisely the same
way”. It seems reasonable to expect that productivity advances made through such tweaking
would eventually be incorporated by local machine producers. The textile and engineering in-
dustries may have also been linked through labor market connections. Marshall (1920) (p. 226)
suggests that textile firms located near E&M firms (or vice-versa) because these industries used

---

44Lazonick (1990) (p. 96).
complementary sets of labor.\textsuperscript{45} Given these connections, I may expect the shock to reduce the growth rate of the E&M industry in locations that were more severely impacted by the cotton shortage.

The pattern of growth of the E&M industry across towns over the study period is explored in Figure 3.5. The left-hand panel of this figure presents the sum of log employment for towns in Lancashire (high shock intensity) compared to the sum over towns in Yorkshire (low shock intensity).\textsuperscript{46} The right-hand panel presents the sum of log employment in towns, where the towns have been grouped based on the share of their employment in cotton textiles in 1851, the instrument for the intensity of the shock.\textsuperscript{47} These figures suggest that the shock led to a slowdown in the growth of the E&M industries in locations in which cotton textiles were initially more important, which tended to be the locations that were more severely impacted by the shock. While the increase in log employment was similar in towns with high and low cotton textile employment shares prior to 1861, there was divergence in 1871-1891, with the towns in which cotton textiles were less important gaining a relative advantage in the E&M industry.

Given the available data, it is not possible to pinpoint which types of inter-industry connections are driving the observed effects. However, given that I observe effects which persist several decades after the end of the cotton shortage and the recovery of the cotton textile industry, it is clear that these effects cannot be driven by contemporaneous demand linkages alone. Rather, the appearance of the persistent divergence following the shock suggests that the temporary recession in cotton textile production was transmitted through inter-industry connections and generated long-

\textsuperscript{45}Textile production employed many women and children, while E&M firms employed almost exclusively adult males. \textit{Marshall (1920)} argued that by co-locating, these firms could pay lower wages and still achieve similar total household incomes levels. Note that this is the opposite of how most economists think of the labor market pooling effect today, which involves benefits to firms employing similar labor forces through co-locating. Marshall acknowledged both of these potential benefits, but less attention has been paid to the benefits of co-location for industries employing different labor forces.

\textsuperscript{46}Graphs showing the sum of log employment are shown rather than those showing the log of the sum of employment across towns because the former reduces the influence of outlier towns in the graph and also because this corresponds more directly to the econometric methodology. The main outlier is Oldham which was the home of the most dominant engineering firm over the period, Platt Bros. of Oldham, which does not appear to have suffered as much from the shock, perhaps due to the benefits of economies of scale or market power.

\textsuperscript{47}The groups are: high – Blackburn, Bolton, and Preston; medium – Manchester, Oldham, and Stockport; low – Bradford, Halifax, Huddersfield, Leeds, Sheffield.
lasting impacts on the relative productivity of E&M firms in more severely affected locations. In the next section, I present a model describing how such long-term effects can be generated through one type of inter-industry connections: learning-by-doing spillovers.

Figure 3.5: E&M Industry employment by location

3.5 Model

This section presents stylized dynamic Ricardian trade model that describes how a shock can be transmitted across related industries and affect outcomes during and after the shock period. Industries are linked through learning-by-doing spillovers and is closely related to work by Young (1991). There is also a larger related set of literature, including work by Grossman & Helpman (1990), Rivera-Batiz & Romer (1991), Matsuyama (1992), and Feenstra (1996), that considers spillovers using a more limited set of industries and spillover patterns. While learning-by-doing technology spillovers are only one of the set of potential types of inter-industry connections, I have chosen to focus on them because this approach allows a simple representation of inter-industry connections. This is, perhaps, why previous studies in this literature have chosen this approach as well.

The model economy is composed of two locations, called home (H) and foreign (F), and indexed by \( l \in \{ H, F \} \). To relate the model to the empirical setting, I can think of home as representing a town in Lancashire, while foreign corresponds to a town in Yorkshire. These locations
produce output, trade, and consume. The labor force in each location, $L_l$, is equal to one. Trade is costless, implying that prices will equalize across locations.

The model is dynamic, but it can be solved iteratively as a series of separate one-period steps, with technology taken as given in any period. The periods are then linked through technological spillovers, which, being external to firms, do not affect the equilibrium in a particular period. Therefore, I solve a one-period version of the model first and then describe how outcomes from one period determine technology levels in the next. Time subscripts are suppressed until the dynamic elements of the model are introduced.

There are two sectors in the economy, a differentiated good sector, and a homogeneous good sector. The homogeneous good is numeraire, so $p_G = 1$. Individuals’ utility over these sectors is given below, where $D$ is an index of differentiated good consumption, $G$ is consumption of the homogeneous good, and $\mu \in (0, 1)$.

$$U = D^\mu G^{1-\mu}$$

Within the differentiated goods sector, there are a fixed number $N$ goods, indexed by $i \in \{1, \ldots, N\}$, which also correspond to industries. Individuals’ utility is such that the index of differentiated goods consumption takes a Cobb-Douglas form, where $d_i$ is the quality of consumption of good $i$, $\gamma_i \in (0, 1)$, and $\sum_{i=1}^{N} \gamma_i = 1$.

$$D = \prod_{i=1}^{N} d_i^{\gamma_i}$$

Within each industry, home and foreign each produce one variety, so there are $2N$ total varieties of differentiated goods. Preferences over these varieties take a CES form, where $x_{il}$ is consumption of the variety of good $i$ produced in location $l$, and $\rho \in (0, 1)$. The corresponding price index for each differentiated good sector is given by $p_i$, where $q_{il}$ denotes the price of the variety of good $i$ produced in location $l$, and $\sigma = 1/(1-\rho) \in (1, +\infty)$. 
Thus, the economy has three levels of product specificity: the sector level, the goods level, and the variety level. This utility setup may seem over-complicated at first glance, yet it provides some advantages that will help simplify the analysis. First, this framework will allow firms in both locations to produce positive quantities and compete in each industry (though with different varieties), even with perfect competition between firms. Second, the Cobb-Douglas formulation means that shocks to one industry will not impact other industries through general price index effects. This is an important simplification that will allow me to focus only on inter-industry impacts through technology spillovers. Letting $E$ represent total expenditures in the economy, it can be shown that expenditures in each differentiated goods industries ($E_i$), and for the homogeneous good ($E_G$), respectively, are $E_i = \mu r_i E$ and $E_G = (1 - \mu)E$.

Next, consider the production side of the economy. The homogeneous good sector is composed of many perfectly competitive firms that produce output using labor only, with the production function $G = L_G$. If a location produces the homogeneous good, then the wage in that location is $w_l = p_G = 1$. Production in a differentiated goods industry $i$ in location $l$ depends on technology $A_{il}$ and labor $L_{il}$. Within each differentiated good industry and location there are many perfectly competitive firms, denoted with the subscript $f$. Thus, the production function is $x_{ilf} = A_{il} L_{ilf}$. Firms can freely copy new technologies from one another, so all firms in the same industry and location share the same technology level. However, this information does not flow across locations, so technology in home may differ from that in foreign in the same industry.\footnote{This is consistent with results indicating that spillovers are sharply attenuated with distance. See, e.g., Rosenthal & Strange (2001), Bottazzi & Peri (2003), and Arzaghi & Henderson (2008). Feenstra (1996) shows the importance of this assumption and provides a review of some related empirical literature.}

Given these production functions, perfect competition implies the following price levels for each differentiated good variety.

\[
q_{il} = \frac{w_l}{A_{il}} \quad (3.1)
\]
The differentiated goods production function, price, and demand equations, can be used to derive output and employment for each industry and variety in the differentiated goods sector, even though output for any particular firm in an industry is indeterminate.

\[ x_{il} = E_i p_i^{\sigma-1} q_{il}^{-\sigma} \]  

(3.2)

\[ L_{il} = E_i p_i^{\sigma-1} q_{il}^{-\sigma} A_{il}^{-1} \]  

(3.3)

Using the price levels for a variety from Equation 3.1, it is possible to express the price index for each good, and for all goods, in terms of wages, technology, and exogenous parameters only. This in turn allows me to express output and employment in terms of only wages, technology, and exogenous parameters, using Equations 3.2 and 3.3.

3.5.1 Solving within a period

Solving for outcomes within a particular period, taking technology as given, involves assuming perfect competition and labor market clearing. The assumption of perfect competition has already been used to derive the price equations above. Labor market clearing requires that the sum of all of the individual labor demands for a location equal the total labor force in that location, i.e., such that \( L = L_{Gl} + \sum_{i=1}^{N} L_{il} \). Given Equations 3.1 and 3.3, I can write employment in the differentiated goods industry as a function of only wages and expenditures. Expenditures can also be written as a function of wages since, with all income being spent in any given period, it must be the case that \( E = L(w_H + w_F) \). Thus, finding the equilibrium in a particular period amounts to finding the wages that clear the labor market.

In order to simplify the problem, I make the common assumption that the homogeneous good sector is sufficiently large that each location produces at least some homogeneous goods, i.e., let \( 1 - \mu > 1/2 \). Under these circumstances \( w_H = w_F = 1 \). This is an important simplification for

\footnote{This approach has been used in existing papers including Krugman & Venables (1995) and Fujita et al. (1999) (ch. 14).}
the analysis, since it will rule out shocks to one industry affecting other industries through wages. While this will aid the theoretical analysis, I will need to be careful to control for potential impacts through wages in the empirical exercise.

Plugging these wages into Equation 3.1 I obtain prices for each variety and good. I can then calculate employment and output in each differentiated goods industry. Once employment in all differentiated goods industries is calculated, the remainder must be employed in producing the homogeneous good. Before moving on, it is useful to establish one additional fact, which is derived by dividing Equation 3.3 for home by the same expression for foreign, to obtain Equation 3.4. This expression shows that, in a particular industry, the location with the better relative technology will have higher employment.

\[
\frac{L_{iH}}{L_{iF}} = \left( \frac{A_{iH}}{A_{iF}} \right)^{\sigma^{-1}}
\]

(3.4)

### 3.5.2 Linking multiple periods

The model becomes dynamic when outcomes in one period are linked to technology levels in the next. Technological improvements occur as a result of learning-by-doing, as in Arrow (1962) and Lucas (1988). Technological advances occur only in the differentiated good sector; the homogeneous good sector produces no spillovers. The amount of technological advance in an industry depends on the amount of learning generated in the previous period that the industry benefits from. An industry can benefit from both learning generated within the industry (within-industry spillovers) as well as from learning spillovers from related industries (inter-industry spillovers). The amount of learning generated in industry j that industry i benefits from depends on the parameter \( \tau^{ij} \), as shown in Equation 3.5. Thus, there is an \( n \times n \) matrix of \( \tau^{ij} \) parameters that represent the extent to which learning generated by employment in industry j benefits industry i. The only restriction on these parameters is that \( \tau^{ij} \geq 0 \) for all i and j.\(^{50}\) Note that in this expression spillovers

\(^{50}\)It may be surprising at first that I do not assume that within-industry spillovers are larger than cross-industry spillovers, but there are reasons to think that cross-industry spillovers may be larger. For example, firms may have more incentive to hide knowledge from competitors in their industry but to collaborate with firms in other industries.
depend on $L_{jlt} + 1$ rather than simply $L_{jlt}$. This ensures that spillovers from industry $j$ will be zero when $L_{jlt} = 0$ and positive whenever $L_{jlt} > 0$ and $\tau^{ij} > 0$.

$$\ln(A_{ilt+1}) - \ln(A_{ilt}) = S_{ilt} = \sum_{j=1}^{N} \tau^{ij} \ln(L_{jlt} + 1)$$ (3.5)

This expression gives the technology level in any period, given outcomes in the previous period. Thus, given a set of initial technology levels, the model can now be solved for all future periods.

### 3.5.3 A cost shock

The economic shock is introduced into the model as an additional cost $\phi > 0$ that must be paid for production in one differentiated good industry, denoted $i = C$ (for Cotton) for one period, hereafter labeled period $s$.\(^{51}\) The price for the variety of good $C$ from location $l$ in period $s$ is as follows.

$$q_{Cls} = \frac{w_{ls}}{A_{Cls}} + \phi$$ (3.6)

Observation 1 describes the effects of the shock in the shock period, where I presume that industry $C$ has better technology in home than in foreign at the beginning of the shock period.

**Observation 1.** A cost shock affecting industry $C$ in period $s$, with $A_{CHs} > A_{CFs}$, will have the following effects.

1. Employment in industry $C$, in the shock period, will fall in the location with better technology in industry $C$ (home). The log difference $\ln(L_{CHs}) - \ln(L_{CFs})$ will also fall.

$$\frac{dL_{CHs}}{d\phi} < 0 \quad \text{and} \quad \frac{d(\ln(L_{CHs}) - \ln(L_{CFs}))}{d\phi} < 0 \quad \text{when} \quad A_{CH} > A_{CF}$$

See Kugler (2006) for more on this topic.

\(^{51}\)This reduced-form approach can be motivated by a model that explicitly incorporates intermediate goods in production, as shown in the Appendix.
2. Employment in all other industries will be unaffected by the shock in period $s$,

$$\frac{dL_{il{s}}}{d\phi} = 0$$

The intuition behind part (1) is that, because home is initially more productive in industry C, $\phi$ will be a larger share of the production cost in home than in foreign. In part (2), other industries are not impacted because I have ruled out wage and price index effects between industries, as can be seen in Equation 3.3, in order to focus on effects occurring through non-pecuniary channels. A formal proof is available in the Appendix.

### 3.5.4 Impact of the shock on related industries

Using Equations 3.3 and the expression for $S_{il{s}}$ given in Equation 3.5, I can make some observations on the effect of the shock on related industries in future periods. First, consider the impact of a reduction in employment in the industry receiving the negative cost shock on related industries. The result, stated below, follow directly from Equations 3.3 and 3.5.

**Observation 2.** If industry $i$ benefits from spillovers from industry C ($\tau_{iC} > 0$), then a decrease in $L_{Cl{s}}$, will cause a decrease in $S_{il{s}}$, $A_{iil{s}+1}$, and $L_{iil{s}+1}$. The greater is $\tau_{iC}$, the larger will be the effect.

Given that employment in industry C falls, at least in H, where that industry is most productive, I know that related industries will suffer a loss of spillovers in H. Next, I am interested in how the magnitude of these effects in H compare with those in F, where the impact of the shock, in terms of the total employment lost in industry C, is less severe.
Observation 3. The loss of spillovers in industries related to industry C in the shock period s will be larger in the location in which industry C is initially more productive. I.e., for some industry i with $\tau^C > 0$,

$$\frac{dS_{ihs}}{d\phi} > \frac{dS_{ifs}}{d\phi} \quad \text{when} \quad A_{CHs} > A_{CFs}$$

The intuition here is that the loss of spillovers in related industries will be larger in the location in which the loss of employment in industry C is greater. A formal proof is available in the Appendix. Naturally, the differential impact on the level of spillovers in period s will lead to differential impacts on the level of technology and employment in period s+1. The results above describe the impact of the shock on technology and employment in related industries in period s+1. I can also speculate about the impacts in periods beyond s+1, though here the effects will become more complicated.

Observation 4. If industry i benefits from spillovers from industry C, then a shock to industry C can affect employment in industry i in future periods through three channels: (1) through reducing $L_{ilt}$, which, in the presence of within-industry spillovers, reduces the spillovers available to industry i and therefore reduces $A_{ilt+1}$ and $L_{ilt+1}$; (2) through long-term effects on employment in industry C which affect the spillovers from industry C to industry i in future periods; (3) through affecting employment in other industries, which, if these industries are also related to i, affects the spillovers available to industry i in the future.

These three channels can be seen more clearly by decomposing the spillovers term. In the equation below, the first term on the right-hand side represents channel (1), the second term represents channel (2), and the third term represents channel (3).

$$S_{ilt} = \tau^iln(L_{ilt} + 1) + \tau^Cln(L_{Clt} + 1) + \sum_{j \neq i,C}^{N} \tau^jln(L_{jlt} + 1)$$

The nature of these long-run impacts on related industry i are not clear, due to the complex effects that the shock can have on employment in other related industries, which can impact future...
spillover levels in unexpected ways. However, it is clear that, in the presence of within-industry spillovers, the impact of the shock through channels (1) and (2) will be negative for more related industries in more severely impacted locations. The impact through channel (3) is indeterminate, but if industries related to industry $i$ are also related to industry $C$, as I would expect if there are clusters of related industries, then the impact through channel (3) is also likely to be negative for more related industries in more severely impacted locations.

Of these three channels, channel (3) is the most elusive, but it may well be the most important. Previous studies looking at the impacts of spillovers through input-output linkages due to FDI, such as Aitken & Harrison (1999), have found little evidence of spillovers between firms within the same industry. On the other hand, studies such as Javorcik (2004), Kugler (2006), and Amiti & Cameron (2007) have found evidence of spillovers between firms in different industries. Kugler (2006) argues that this makes sense, because firms are likely to hide information from their competitors while cooperating with firms that they do not compete with, such as their suppliers and customers.

### 3.6 Econometric analysis

The model suggests that the shock should have a negative impact on those industries related to the cotton textile industry and that this impact should be larger in those locations which were more severely impacted by the shock, which also happen to be the locations which were initially more productive in the cotton textile industry. This section tests this prediction. The empirical strategy involves a fixed effect regression approach where the panel has two cross sectional dimensions (towns and industries). Thus, I am comparing across time (pre- vs. post-shock), industries (more vs. less related to cotton textiles), and locations (more vs. less severely impacted by the shock).

The first step is to derive a usable empirical specification from the model. The basis for my empirical specification is Equation 3.3, which describes employment in industry $i$ and location $l$ as a function of expenditures on goods in industry $i$, the price index for goods in industry $i$, and the
productivity of location l in producing its variety of good i. Taking logs and substituting out $q_{ilt}$ using Equation 3.1 and $A_{ilt}$ using the technology growth expression from Equation 3.5, I obtain the following.

$$ln(L_{ilt}) = ln(\mu \gamma_i E) + (\sigma - 1)ln(p_{it}) + (\sigma - 1)ln(A_{ilt-1}) + (\sigma - 1)S_{ilt-1}$$

This expression will motivate the empirical exercises. The key to the empirical strategy will be using the shock, interacted with industry relatedness, as a proxy for the spillovers term, while controlling for factors that vary at the industry-period, location-period, and industry-location level.

Our first test of the model is based on the specification given in Equation 3.7 below, where $R_i$ represents industry i’s relatedness to the cotton textile industry, $V_l$ represents the intensity of the shock in town l, Post indicates the post-shock period. I control for factors that vary at the industry-year level, such as $p_{it}$ by including industry-year dummy variables, $\theta_{it}$. Factors varying at the location-period level, though not present in the stylized model above, must also be considered. These are controlled for by including location-year dummy variables, $\psi_{lt}$. Finally, I allow for variation at the industry-location level by including industry-location fixed effects, $\xi_{il}$. Note that because I am estimating a fully saturated fixed-effect model, lower-level interaction terms are not included, since they would be perfectly correlated with the fixed effects.

$$ln(L_{ilt}) = \beta_0 + \beta_1 (R_i V_l Post_t) + \theta_{it} + \psi_{lt} + \xi_{il} + \epsilon_{ilt}$$ (3.7)

While Equation 3.7 will be useful for studying whether the shock had an impact on the level of employment in related industries, I also want to consider the possibility that the growth rate of related industries may have also been impacted. A simple way to look for this is using Equation 3.8 below, where $T_t$ is a time-trend variable for the post-shock period ($T_{1851} = 0, T_{1861} = 0, T_{1871} = 1, T_{1881} = 2, T_{1891} = 3$).

$$ln(L_{ilt}) = \beta_0 + \beta_1 (R_i V_l T_t) + \theta_{it} + \psi_{lt} + \xi_{il} + \epsilon_{ilt}$$ (3.8)
The main coefficient of interest will be $\beta_1$. A negative $\beta_1$ coefficient estimate would suggest that the shock had a negative impact on more related industries in more severely impacted towns, supporting the predictions of the model. Identification relies on the assumption that there are no omitted variables that change between the pre and post-shock time periods, affect related industries more than unrelated industries, and are stronger in those towns more severely hit by the shock. One concern that I may have with these identification assumptions is that the severity of the shock, as measured by the number of able-bodied workers seeking relief from Poor Law boards in each town, may reflect changes in the economic conditions in these towns occurring at the same time as the shock which are not driven by the shock. To strengthen the identification strategy against this concern, for each specification I will calculate results in which the employment share of cotton textile production in 1851 in each location is used as an IV for the severity of the shock (interacted with the other terms). I would not expect that having a larger share of cotton textile employment in 1851 would cause a change in the performance of related industries after 1861, but not before, other than through the effects of the shock. The results in Table 3.3 suggest that this will be a very strong instrument.

Because I am interested in spillovers across industries, I exclude cotton textile production and other cotton-based industries from the analysis. These industries differ fundamentally from all others in that they were directly impacted by the shock regardless of their location. In order to better handle serial correlation issues, standard errors are clustered at the industry-location level. This allows for standard errors to be correlated over time within an industry-location, in order to deal with potential serial correlation issues. In order to make the results easier to interpret, the continuous variables, the shock intensity measure and relatedness measures, are standardized to

---


53 Alternatively, I may have chosen to cluster at the industry or the town level. Results, available in the Appendix, suggest that my results are robust to these alternatives.

54 Bertrand et al. (2004) show that serial correlation issues can be important in differences-in-differences estimation, though the structure of the data used in this study, with its large cross section and relatively short time dimension, makes serial correlation less of a worry. See also Angrist & Pischke (2009). Another potential issue, pointed out by Donald & Lang (2007), is that t-statistics may suffer when the number of clustered groups is small. However, because this study uses data from a relatively large number of clustered groups (171), this should not be a major concern.
have a mean of zero and standard deviation of one. Summary statistics for the data used in the analysis are available in the Appendix.

Baseline regression results for Equations 3.7 are given in Table 3.4. These results are calculated using a standard fixed-effects approach at the industry-location level, for the 1,543 unique industry-locations in the data, while including industry-year and location-year dummy variables. Columns (1) and (2) calculate results using the coagglomeration and input-output relatedness measures, respectively, while both measures are included in column (3). Column (4) presents “reduced form” results, where the share of cotton textile employment to total employment in each location is used in place of the shock severity measure based on number of relief seekers. Columns (5)-(7) present IV results in which the share of cotton textile employment to total employment in each location is used as an IV for the severity of the shock in each location in a Two-Stage Least Squares regression.

The results in Table 3.4 suggest that log employment was negatively impacted by the shock in more related industries, based on the coagglomeration measure, in more severely affected towns. No similar negative impacts appear for those industries linked to cotton textiles through input-output connections, and the coagglomeration measure results continue to be significant when the input-output measure is included. This suggests that the negative employment impacts are driven by some channel other than direct input output linkages. In terms of magnitude, these results suggest that in a location that is one standard deviation above the mean in terms of shock severity, an industry that is one standard deviation above the mean in relatedness would suffer a relative reduction in employment of 4.5-7%. Put another way, the predicted impact on an industry with a relatedness measure equal to that of the Engine & Machine Makers (1.27 s.d.), in the most severely impacted town Preston (1.77 s.d.) is a relative employment reduction of 10-16% in the post-shock period.

Table 3.5, which has the same format as Table 3.7, shows results corresponding to Equation 3.8. This allows me to study whether I still observe the same results when I allow for simple linear growth trend in the impact of the shock on related industries in more severely impacted
Table 3.4: Fixed effect regression results for Eq. 3.7

<table>
<thead>
<tr>
<th></th>
<th>Coag. only</th>
<th>IO only</th>
<th>Both</th>
<th>Reduced form</th>
<th>IV Coag</th>
<th>IV IO</th>
<th>IV Both</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coag. x Shk x Post</td>
<td>-0.0472***</td>
<td>-0.0486***</td>
<td>-0.0600***</td>
<td>-0.0797***</td>
<td>-0.0806***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IO x Shk x Post</td>
<td>0.0480</td>
<td>0.0536*</td>
<td>0.0238</td>
<td>0.0217</td>
<td>0.0320</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 3.5: Fixed effect regression results for Eq. 3.8

<table>
<thead>
<tr>
<th></th>
<th>Coag. only</th>
<th>IO only</th>
<th>Both</th>
<th>Reduced form</th>
<th>IV Coag</th>
<th>IV IO</th>
<th>IV Both</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coag. x Shk x TT</td>
<td>-0.0231***</td>
<td>-0.0235***</td>
<td>-0.0287***</td>
<td>-0.0385***</td>
<td>-0.0386***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IO x Shk x TT</td>
<td>0.0156</td>
<td>0.0183</td>
<td>0.00189</td>
<td>-0.00252</td>
<td>0.00244</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

locations. The main variable of interest is then Coag. x ShockInt. x TT. The coefficients on this term continue to be negative and strongly statistically significant, providing some evidence that the shock may have impacted the growth rate in log employment, in addition to the level.

An alternative approach to the more standard fixed effects is to use first-differencing. First differencing will be efficient when error terms follow a random walk, which seems likely in my setting. First difference results are calculated using the following two specifications. In Equation 3.9, T_{1871} is an indicator variable for the year 1871, the first post shock year. Thus, this expression allows me to test for impact on the level of employment in related industries. In Equation 3.10, Post_{1} is an indicator for the post shock period. This expression allows me to test for persistent
Table 3.6: First-difference regression results for Eq. 3.9

<table>
<thead>
<tr>
<th></th>
<th>Coag. only</th>
<th>IO only</th>
<th>Both</th>
<th>Reduced form</th>
<th>IV Coag</th>
<th>IV IO</th>
<th>IV Both</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td>(7)</td>
<td></td>
</tr>
<tr>
<td>Coag. × Shk</td>
<td>-0.0250</td>
<td>-0.0266</td>
<td>-0.0238</td>
<td>-0.0294</td>
<td>-0.0320</td>
<td></td>
<td></td>
</tr>
<tr>
<td>× Post</td>
<td>(0.0174)</td>
<td>(0.0174)</td>
<td>(0.0187)</td>
<td>(0.0238)</td>
<td>(0.0237)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IO × Shk</td>
<td>0.0573*</td>
<td>0.0603**</td>
<td>0.0644*</td>
<td>0.0830*</td>
<td>0.0871**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>× Post</td>
<td>(0.0305)</td>
<td>(0.0304)</td>
<td>(0.0348)</td>
<td>(0.0439)</td>
<td>(0.0440)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ind-year Ind.</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Loc-year Ind.</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>6,172</td>
<td>6,172</td>
<td>6,172</td>
<td>6,172</td>
<td>6,172</td>
<td>6,172</td>
<td></td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses *** p < 0.01, ** p < 0.05, * p < 0.1

impacts on the change in log employment in related industries.

\[
\Delta \ln(L_{it}) = \beta_0 + \beta_1 (RV_{i1871}) + \theta_{it} + \psi_{it} + \Delta \epsilon_{ilt}
\]  \hspace{1cm} (3.9)

\[
\Delta \ln(L_{it}) = \beta_0 + \beta_1 (RV_{it \text{Post}_t}) + \theta_{it} + \psi_{it} + \Delta \epsilon_{ilt}
\]  \hspace{1cm} (3.10)

Results for OLS regressions based on Equation 3.9 are presented in Table 3.6. These results are qualitatively similar to those obtained using the fixed effects, but the magnitudes are somewhat smaller and the coefficients are not statistically significant. Results for OLS regressions based on Equation 3.10 are presented in Table 3.7. Here, I observe fairly strong evidence that the shock had a negative impact on the growth in log employment in more related industries and more severely impacted locations in the post-shock period. It appears that both the fixed effects and first differences approaches indicate that the shock had persistent impacts.

It is also interesting to consider the impact of the shock year-by-year. This can be done by allowing by estimating separate coefficients on the interaction term for each post-shock year, rather than using the post-shock flag, which requires that the coefficient be the same for all years. Table 3.8 presents year-by-year results. These results indicate that the shock had a negative impact in all post-shock periods on related industries, based on the coagglomeration measure, and that the impact was growing over time. As before, no consistent impacts are observed in those industries.
related through input-output linkages, though it does appear that input-output related industries may have experienced some positive impacts in 1871.

Table 3.8: Year-by-year regression results

<table>
<thead>
<tr>
<th>Coag.</th>
<th>IO</th>
<th>Both</th>
<th>Reduced form</th>
<th>IV Coag</th>
<th>IV IO</th>
<th>IV Both</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td>(7)</td>
</tr>
<tr>
<td>Coag. × Shk</td>
<td>-0.0346**</td>
<td>-0.0361**</td>
<td>-0.0426***</td>
<td>-0.0551***</td>
<td>-0.0572***</td>
<td></td>
</tr>
<tr>
<td>× 1871</td>
<td>(0.0151)</td>
<td>(0.0151)</td>
<td>(0.0162)</td>
<td>(0.0210)</td>
<td>(0.0209)</td>
<td></td>
</tr>
<tr>
<td>Coag. × Shk</td>
<td>-0.0288</td>
<td>-0.0300</td>
<td>-0.0435**</td>
<td>-0.0580**</td>
<td>-0.0585**</td>
<td></td>
</tr>
<tr>
<td>× 1881</td>
<td>(0.0207)</td>
<td>(0.0206)</td>
<td>(0.0214)</td>
<td>(0.0279)</td>
<td>(0.0276)</td>
<td></td>
</tr>
<tr>
<td>Coag. × Shk</td>
<td>-0.0782***</td>
<td>-0.0796***</td>
<td>-0.0939***</td>
<td>-0.126***</td>
<td>-0.126***</td>
<td></td>
</tr>
<tr>
<td>× 1891</td>
<td>(0.0269)</td>
<td>(0.0269)</td>
<td>(0.0262)</td>
<td>(0.0340)</td>
<td>(0.0337)</td>
<td></td>
</tr>
<tr>
<td>IO × Shk</td>
<td>0.0549**</td>
<td>0.0590**</td>
<td>0.0524*</td>
<td>0.0634</td>
<td>0.0707*</td>
<td></td>
</tr>
<tr>
<td>× 1871</td>
<td>(0.0262)</td>
<td>(0.0262)</td>
<td>(0.0303)</td>
<td>(0.0388)</td>
<td>(0.0390)</td>
<td></td>
</tr>
<tr>
<td>IO × Shk</td>
<td>0.0434</td>
<td>0.0469</td>
<td>0.0112</td>
<td>0.00744</td>
<td>0.0104</td>
<td></td>
</tr>
<tr>
<td>× 1881</td>
<td>(0.0427)</td>
<td>(0.0429)</td>
<td>(0.0455)</td>
<td>(0.0587)</td>
<td>(0.0587)</td>
<td></td>
</tr>
<tr>
<td>IO × Shk</td>
<td>0.0457</td>
<td>0.0549</td>
<td>0.00798</td>
<td>-0.00580</td>
<td>0.0104</td>
<td></td>
</tr>
<tr>
<td>× 1891</td>
<td>(0.0401)</td>
<td>(0.0394)</td>
<td>(0.0461)</td>
<td>(0.0599)</td>
<td>(0.0586)</td>
<td></td>
</tr>
</tbody>
</table>

Ind-loc FE: Yes Yes Yes Yes Yes Yes Yes
Ind-year Ind. Yes Yes Yes Yes Yes Yes Yes
Loc-year Ind. Yes Yes Yes Yes Yes Yes Yes
Obs. 6,172 6,172 6,172 6,172 6,172 6,172 6,172

Robust standard errors in parentheses  *** p<0.01, ** p<0.05, * p<0.1

We have shown that the results tend to be robust to changes in the estimation specification. I can also check their robustness to changes to the data used to estimate the results. Table 3.9 presents results that assess the robustness of the findings to several changes in the data used in the estimation. Columns (1) and (2) present results calculated when data from Liverpool is included.
Column (1) presents fixed effects results including both relatedness measures, and column (2) presents IV results. Recall that Liverpool was excluded because industries other than cotton textiles were directly impacted by the civil war there. Including Liverpool does not appear to alter the results.

Another potential issue arises from the fact that the set of towns used to obtain the results may also influence the coagglomeration measure. To see why this might cause an issue, suppose that the coagglomeration measure simply reflects a random chance that an industry ends up co-agglomerated with cotton textiles, rather than underlying connections between these industries. Furthermore, suppose that when cotton textiles is shocked in a location all industries in that location suffer, with the industries more agglomerated in the location suffering more. This scenario may cause errant results if the coagglomeration measure is driven by the towns used in the analysis. This is unlikely because I have calculated coagglomeration using district-level, rather than town-level data, which covers many areas other than the towns used in the analysis. However, if one is still worried that the towns used are driving the coagglomeration measure, then results can be recalculated using a coagglomeration measure calculated using only districts that do not include towns used in the subsequent analysis. Results using this alternative coagglomeration measure are presented columns (3) and (4) of Table 3.9.

Columns (5) and (6) of Table 3.9 presents results when all wool-related industries have been dropped from the analysis. While these are an important class of industries which are closely related to the cotton textile industry, I may be worried that the fact that they are initially much larger in Yorkshire towns could lead them to grow faster there, in the presence of economies of scale or within-industry spillovers. While dropping these industries does reduce the size of the estimated impact, it does not alter the basic message. Finally, columns (7) and (8) include results in which all industries sharing intermediate good input-output connections with the cotton textile industry have been excluded from the analysis. This appears to strengthen the results, providing further evidence that the observed impacts were not driven by intermediate goods input-output connections.
Table 3.9: Some robustness checks

<table>
<thead>
<tr>
<th></th>
<th>With Liverpool</th>
<th>Ait. Coag. Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Reg. IV</td>
<td>Reg. IV</td>
</tr>
<tr>
<td>Coag × Shk</td>
<td>-0.0516***</td>
<td>-0.0464**</td>
</tr>
<tr>
<td>× Post</td>
<td>(0.0173)</td>
<td>(0.0194)</td>
</tr>
<tr>
<td>IO × Shk</td>
<td>0.0401</td>
<td>0.0501*</td>
</tr>
<tr>
<td>× Post</td>
<td>(0.0284)</td>
<td>(0.0304)</td>
</tr>
<tr>
<td>Ind-loc FEs</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Ind-year Ind.</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Loc-year Ind.</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Obs</td>
<td>8,535</td>
<td>7,715</td>
</tr>
</tbody>
</table>

|                      | With Wool      | No IO Inds.       |
|                      | Reg. IV        | Reg. IV           |
| Coag × Shk           | -0.0283*       | -0.0785***        |
| × Post               | (0.0160)       | (0.0244)          |
| IO × Shk             | 0.0431         | 0.0208            |
| × Post               | (0.0300)       | (0.0445)          |
| Ind-loc FEs          | Yes            | Yes               |
| Ind-year Ind.        | Yes            | Yes               |
| Loc-year Ind.        | Yes            | Yes               |
| Obs.                 | 7,435          | 4,320             |

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1
Three innovations offered by this study, vis a vis existing studies, is the use of geographically disaggregated data, the use of more detailed industry-level data, and accounting for the pattern of inter-industry connections. Given that my results differ from previous studies which find that temporary shocks have few long-term impacts, such as Davis & Weinstein (2008), it is interesting to look at which of these innovations appears to be driving these results. To analyze the impact of using geographically disaggregated data, I can apply my estimation strategy at a less disaggregated level, by comparing outcomes in the county of Lancashire as a whole to outcomes in Yorkshire. This analysis follows that shown in Equation 3.7 except that a flag for Lancashire is used in place of the shock severity measure. Results calculated using these county-level data are presented in Table 3.10, columns (1)-(2). While these results have the same sign as those found with the full data set, they are not statistically significant. Thus, had I used less geographically disaggregated data I would not have found strong evidence that the shock impacted related industries.

In order to assess the importance of using data that are more disaggregated at the industry level, I can repeat the analysis across more aggregated industry sectors. For comparability, I use the eight broad industrial sectors used in Davis & Weinstein (2008). These sectors are Machinery, Metals, Chemicals, Textiles and Apparel, Processed Food, Printing and Publishing, Lumber and Wood, and Ceramics. Industries not fitting into one of these categories are dropped. The relatedness between the remaining categories and the Textile and Apparel category are then calculated using the geographic coagglomeration measure. Then the Textiles and Apparel sector is excluded and results are calculated using the approach in Equation 3.7. Results, shown in Table 3.10, columns (3)-(6), suggest that had I used data aggregated to these eight industrial sectors I would have found only weak evidence that the shock affected the level of employment in related industries.

55 As before, Lancashire includes neighboring Cheshire county while only the West Riding is included in Yorkshire.
Finally, I can calculate results while ignoring industry connections. This is done using the specifications in Equations 3.11 and 3.12, which estimate the impact of the shock based only on the severity in each location. For these specification, it makes more sense to cluster errors at the location level, since the shock to each location is assumed to impact all of the industries in that location.

\[
\ln(L_{ilt}) = \beta_0 + \beta_1(V_lPost_t) + \theta_{ilt} + \xi_{ilt} + \epsilon_{ilt}
\]  
\[\text{(3.11)}\]

\[
\ln(L_{ilt}) = \beta_0 + \beta_1(V_lT_t) + \theta_{ilt} + \xi_{ilt} + \epsilon_{ilt}
\]  
\[\text{(3.12)}\]

Regression results are shown in Table 3.11, where the top set of results correspond to Equation 3.11 and the bottom shows results from separate regressions using Equation 3.12. Columns (1)-(2) present baseline and IV results, respectively, ignoring inter-industry connections. As before, I observe a negative effect on industries in more severely impacted towns, but these results are not statistically significant. Thus, had I ignored the pattern of inter-industry connections, I would not have found strong evidence that the shock had a long-term impact. To learn more, I run separate regression for only those industries with positive coagglomeration relatedness mea-

Table 3.10: Regression results using less disaggregated data

<table>
<thead>
<tr>
<th>Coag. × Lanc</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coag. × Lanc</td>
<td>-0.0203</td>
<td>-0.0569</td>
<td>-0.0959</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>× Post</td>
<td>(0.0376)</td>
<td>(0.0400)</td>
<td>(0.0640)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IO × Lanc</td>
<td>0.0127</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>× Post</td>
<td>(0.0759)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Industry-location FEs | Yes | Yes | Yes | Yes | Yes | Yes |
| Industry-year Ind.    | Yes | Yes | Yes | Yes | Yes | Yes |
| Location-year Ind.    | Yes | Yes | Yes | Yes | Yes | Yes |

Observations 1,710 1,710 440 440 440 440

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1
Table 3.11: Results calculated ignoring inter-industry connections

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Shock Int. × Post</td>
<td>-0.101* (0.0515)</td>
<td>-0.105 (0.0655)</td>
<td>-0.162 (0.0924)</td>
<td>-0.211 (0.140)</td>
<td>-0.0879* (0.0438)</td>
<td>-0.0819 (0.0558)</td>
</tr>
<tr>
<td>Shock Int. × TT</td>
<td>-0.0334 (0.0251)</td>
<td>-0.0310 (0.0303)</td>
<td>-0.0592 (0.0447)</td>
<td>-0.0776 (0.0671)</td>
<td>-0.0279 (0.0214)</td>
<td>-0.0210 (0.0247)</td>
</tr>
<tr>
<td>Industry-location FEs</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Industry-year Ind.</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>7,715</td>
<td>7,715</td>
<td>1,385</td>
<td>1,385</td>
<td>6,330</td>
<td>6,330</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

sures, in columns (3)-(4) and only those with negative coagglomeration relatedness measures, in columns (5)-(6). I observe that the coefficients measuring the impact of the shock are much larger (more negative) for the more related industries than for the less related industries, though the results for the more related industries are still not statistically significant, likely due to the much reduced sample size. In sum, these results indicate the importance of accounting for the pattern of inter-industry connections in obtaining my results.

Together, the results shown in Tables 3.10 and 3.11 indicate that obtaining my results depended both on the use of detailed data, both in terms of industries and geographic locations, as well as paying attention to the patterns of connections between industries. The combination of these elements may help explain why my results differ from those found in previous studies. I cannot, however, rule out other factors, such as the fact that I consider a shock driven by economic forces while existing studies consider shocks generated by war.

3.7 Conclusion

This project describes a large, exogenous, industry-specific shock to the 19th century British economy and shows that it affected the distribution of industries across locations up to 25 years after the end of the shock. In particular, it shows that those industries more closely related to the
industry which was directly affected, cotton textiles, suffered long-term reductions in employment and employment growth, in those towns which were more severely affected by the shock. This provides causal evidence that inter-industry spillovers can transmit negative shocks with long-term effects. These effects are of significant magnitude and longevity given the transient nature of the shock considered.

These results have implications for two types of policies. First, while this study focuses on a large negative shock caused by exogenous factors, it seems possible that localized industrial policy interventions may be able to generate similar effects. This may provide some justification for the widespread use of these policies. However, these results also suggest that the effectiveness of such policy interventions depends crucially on the pattern of connections between industries, which are currently not well understood.

These results also have implications for policies that may influence the vulnerability of an economy to economic shocks, because they suggest that even temporary shocks, if large enough, can have long-lasting effects. While this study shows how a temporary shock can reallocate industries within a country, the same forces are also likely to be at work across countries. Thus, policy makers may want to consider the potential effects of volatility when considering policies that increase and economy’s vulnerability to temporary shocks.

While this study tells me that inter-industry connections can transmit economic forces between industries, with long-term implications, I am unable to identify the particular types of connections that are most important, though it does not appear that intermediate good input-output linkages are driving the observed results. This is an important direction for future research.


Aghion, Philippe, Dechezlepretre, Antoine, Hemous, David, Martin, Ralph, & Van Reenen, John. 2011 (July). *Carbon Taxes, Path Dependency and Directed Technical Change : Evidence from the Auto Industry*. Working paper. 2.1, 2, 2.5.1, 16, 2.5.1, 2.5.2


Atkinson, Fred J. 1897. Silver Prices in India. *Journal of the Royal Statistical Society*, 60(1), 84–147. 1.7, 1.18


Bloom, Nicholas, Schankerman, Mark, & Van Reenen, John. 2012 (March). *Identifying technology spillovers and product market rivalry*. Working paper. 2.1, 3


Habakkuk, HJ. 1962. *American and British Technology in the Nineteenth Century: Search for Labor Saving Inventions*. Cambridge University Press. 1.1, 1.8


Helm, Elijah. 1869. The Cotton Trade of the United Kingdom, During the Seven Years, 1862-68, as Compared with the Seven Years, 1855-61; With Remarks on the Return of Factories Existing in 1868. *Journal of the Statistical Society of London*, 32(4), pp. 428–437. 20, 27


Horrell, Sara, Humphries, Jane, & Weale, Martin. 1994. An Input-Output Table for 1841. *Economic History Review*, 47(3), 545–566. 3.3.1


Kugler, Maurice. 2006. Spillovers from foreign direct investment: Within or between industries? *Journal of Development Economics*, 80(2), 444–477. 4, 50, 3.5.4


Mackay, Alexander. 1853. *Western India*. London: Nathaniel Cooke. 21


Ollersnshaw, J.C. 1870. Our Export Trade in Cotton Goods to India. *Transactions of the Manchester Statistical Society, 1869-1870*, p. 112. 21


Popp, David. 2002. Induced Innovation and Energy Prices. *American Economic Review*, **92**(1), 160–180. 1, 1.1, 5, 40, 2.1, 2, 2.5.1, 2.5.1, A.10


Thomas, Mark. 1987. *An Input-Output Approach to the British Economy, 1890-1914*. Ph.D. thesis, Oxford University. 3.1, 3.3.1, 40, C.2.2


Williams, Heidi. 2010 (August). *Intellectual property rights and innovation: Evidence from the human genome*. Working paper. 2.4


Appendix A

Appendix to Chapter 1

A.1 Definitions of important textile terms

The following definitions were constructed with the aid of The “Mercury” Dictionary of Textile Terms. 1950. Textile Mercury Limited: Manchester, England.

**Bale**- Raw cotton or wool are packed in varying weights in bale form for shipping.

**Baler or Packer**- Machines used to compress raw cotton into bales.

**Bolls**- The seed pod of cotton and has from three to five cells, each of which contains from six to twelve seeds, the seeds being covered with cotton fibers.

**Carding**- A very thorough opening-out and separating of the fibers of cotton, together with an effective cleaning. This machine is the last where cleaning the cotton takes place (unless the cotton has to be combed).

**Combing**- This term is used literally and denotes the combing of fibrous materials in sliver form by mechanically actuated combs or by hand-operated combs. In general, the objects in combing are two, namely (1) to obtain the maximum parallelization of the fibers and (2) to remove impurities and undesired short fibers.

**Gin**- A cotton cleaning machine with the primary purpose of separating the cotton seeds from the cotton fibers.

**Opening cotton**- This is done on machines (openers) which beat the cotton into a more fleecy
condition and also remove a good proportion of the dirt and heavier impurities.

**Scutching**- An operation in preparing cotton for spinning that has three objects, to reduce the cotton to a loose open condition by beating it, removal of impurities remaining in the cotton after opening, and the formation of a continuous lap or web of cotton wound on to a rod—which laps go forward to the carding engine.

### A.2 British cotton and wool imports and re-exports 1861

Figure A.1: British cotton and wool imports and re-exports in 1861

Map drawn by Charles Joseph Minard from the collections of the Library of Congress access through cartographia.wordpress.com. Blue represents cotton and wool from the United States, the orange from British territories in South Asia, and brown from the Levant (the East Mediterranean). Pink represents cotton and wool imported to Britain that was subsequently re-exported to Europe. There is also a small sliver of imports from Brazil, also in a light blue. One millimeter represents 5,000 tons of cotton or wool. According to Mitchell & Deane (1962), cotton imports in 1861 totaled 1257 million lbs. while wool imports totaled only 147.2 million lbs., so most of the imports shown on this map will be of cotton.
A.3 Machines for ginning cotton

Figure A.2: Indian Churka for removing cotton seeds

Reproduced from Wheeler (1862).

Figure A.3: Cottage Saw Gin

Reproduced from Wheeler (1862).
A.4 Details on the differences between cotton types

Figure A.4: Length of cotton staples for various cotton types

Reproduced from Wheeler (1862).
A.5  Impact of ginning on cotton fiber length

Figure A.5: A comparison of ginned (left) and hand-cleaned cotton (right) fiber length

Reproduced from Pearse (1921).

A.6  Background graphs

Figure A.6: British imports and exports 1851-1869

British imports

British exports of finished manufactures

Data from Mitchell & Deane (1962).
Figure A.7: British wool and linen textile exports 1815-1910

Wool textile exports

Linen textile exports

Data from Mitchell & Deane (1962)

A.7 Cotton waste data

Table A.1: Cotton consumption and waste data 1860-1868

<table>
<thead>
<tr>
<th></th>
<th>1860</th>
<th>1861</th>
<th>1862</th>
<th>1863</th>
<th>1864</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cotton consumed</td>
<td>1,079,321</td>
<td>1,005,477</td>
<td>449,821</td>
<td>476,445</td>
<td>561,196</td>
</tr>
<tr>
<td>Waste in spinning</td>
<td>113,328</td>
<td>105,575</td>
<td>76,469</td>
<td>71,466</td>
<td>78,567</td>
</tr>
<tr>
<td>Waste share</td>
<td>10.50%</td>
<td>10.50%</td>
<td>17.00%</td>
<td>15.00%</td>
<td>14.00%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1865</th>
<th>1866</th>
<th>1867</th>
<th>1868</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cotton consumed</td>
<td>718,651</td>
<td>890,721</td>
<td>954,517</td>
<td>996,197</td>
</tr>
<tr>
<td>Waste in spinning</td>
<td>100,611</td>
<td>115,793</td>
<td>114,533</td>
<td>119,544</td>
</tr>
<tr>
<td>Waste share</td>
<td>14.00%</td>
<td>13.00%</td>
<td>12.00%</td>
<td>12.00%</td>
</tr>
</tbody>
</table>

Data from Forwood (1870). Cotton consumption and waste are in thousands of lbs.
## A.8 First-stage regression results for Table 1.2

Table A.2: First-stage results for Table 1.2

<table>
<thead>
<tr>
<th></th>
<th>Indian price</th>
<th>Brazilian price</th>
<th>Egyptian price</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>For column (2)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U.S. Imports</td>
<td>-0.416***</td>
<td>-0.366***</td>
<td>-0.350***</td>
</tr>
<tr>
<td></td>
<td>(0.0487)</td>
<td>(0.0353)</td>
<td>(0.0399)</td>
</tr>
<tr>
<td><strong>For column (3)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shock period</td>
<td>0.993***</td>
<td>0.896***</td>
<td>0.843***</td>
</tr>
<tr>
<td></td>
<td>(0.161)</td>
<td>(0.121)</td>
<td>(0.133)</td>
</tr>
<tr>
<td><strong>For column (4)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shock period</td>
<td>0.732***</td>
<td>0.726***</td>
<td>0.647***</td>
</tr>
<tr>
<td></td>
<td>(0.115)</td>
<td>(0.100)</td>
<td>(0.106)</td>
</tr>
<tr>
<td>Post-onset</td>
<td>0.547***</td>
<td>0.356***</td>
<td>0.409***</td>
</tr>
<tr>
<td></td>
<td>(0.081)</td>
<td>(0.071)</td>
<td>(0.075)</td>
</tr>
<tr>
<td><strong>For column (5)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shock period</td>
<td>0.400***</td>
<td>0.431***</td>
<td>0.388***</td>
</tr>
<tr>
<td></td>
<td>(0.107)</td>
<td>(0.089)</td>
<td>(0.102)</td>
</tr>
<tr>
<td>Post-onset</td>
<td>0.494***</td>
<td>0.295***</td>
<td>0.367***</td>
</tr>
<tr>
<td></td>
<td>(0.068)</td>
<td>(0.057)</td>
<td>(0.064)</td>
</tr>
<tr>
<td>Lag shock</td>
<td>0.578***</td>
<td>0.512***</td>
<td>0.452***</td>
</tr>
<tr>
<td></td>
<td>(0.105)</td>
<td>(0.087)</td>
<td>(0.100)</td>
</tr>
<tr>
<td><strong>For column (6)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shock period</td>
<td>0.498***</td>
<td>0.547***</td>
<td>0.518***</td>
</tr>
<tr>
<td></td>
<td>(0.097)</td>
<td>(0.070)</td>
<td>(0.082)</td>
</tr>
<tr>
<td>Post-onset</td>
<td>0.452***</td>
<td>0.231***</td>
<td>0.312***</td>
</tr>
<tr>
<td></td>
<td>(0.064)</td>
<td>(0.046)</td>
<td>(0.052)</td>
</tr>
<tr>
<td>Lag shock</td>
<td>0.310***</td>
<td>0.198**</td>
<td>0.099</td>
</tr>
<tr>
<td></td>
<td>(0.103)</td>
<td>(0.077)</td>
<td>(0.089)</td>
</tr>
<tr>
<td>Lag shock t-2</td>
<td>0.366***</td>
<td>0.429***</td>
<td>0.483***</td>
</tr>
<tr>
<td></td>
<td>(0.096)</td>
<td>(0.069)</td>
<td>(0.081)</td>
</tr>
</tbody>
</table>

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1
A.9 First-stage regression results for Table 1.3

Table A.3: First-stage regression results for Table 1.3

<table>
<thead>
<tr>
<th></th>
<th>DV: Transport cost index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shock period ind.</td>
<td>-0.1738</td>
</tr>
<tr>
<td></td>
<td>(0.3251)</td>
</tr>
<tr>
<td>Shock x Time-trend</td>
<td>0.6298***</td>
</tr>
<tr>
<td></td>
<td>(0.0928)</td>
</tr>
<tr>
<td>Obs.</td>
<td>16</td>
</tr>
<tr>
<td>F-stat</td>
<td>29.56</td>
</tr>
</tbody>
</table>

Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1.

A.10 Robustness of the estimated impact on textile technologies

While the results contained in Section 1.5 are calculated using quarterly data and a fairly simple estimation strategy, a number of alternatives data sets and approaches are available. Coefficient estimates from some of these alternatives are presented in Table A.4. Rows 1-2 show results calculated using annual data which allow me to extend the time period used in the BPO spinning and BPO weaving category regressions to 1855-1883. Regressions done using monthly data are shown in rows 3-4. Rows 5-6 present Poisson regression results, an approach which is commonly used to analyze patent data. In these rows, the dependent variable is the count of patents of each type. Rows 7-8 present results in which the dependent variable is the share of patents of each type in the total number of patents, which is similar to the approach used in Popp (2002). In row 9, the ratio of India to U.S. raw cotton imports to Britain is used as an independent variable in place of the shock period indicator. All of these results suggest that the shock period was characterized by a significant increase in the number of spinning technology patents, particularly those which were cotton-related. Finally, in row 10, the price of U.S. cotton is used as the main independent
variable, an approach which is similar to that taken in Newell et al. (1999) and Popp (2002). Using price as an explanatory variable is likely to introduce bias due to reverse causality or omitted variables into the regressions, since input prices will be influenced by the availability of new technologies.\(^1\) One indicator of the importance of this bias is that the results shown in row 12 are weaker than all of the other results, though they are still qualitatively similar.

### Table A.4: Regressions using alternative approaches

<table>
<thead>
<tr>
<th>BPO</th>
<th>BPO Spinning</th>
<th>BPO Weaving</th>
<th>Cotton-related</th>
<th>Cotton-Spinning</th>
<th>Wool-related</th>
<th>Linen-related</th>
<th>Silk-related</th>
</tr>
</thead>
<tbody>
<tr>
<td>DV: Log of patent count</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1) Annual with 1861</td>
<td>0.189*** (0.0498)</td>
<td>0.0230 (0.0649)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) Annual no 1861</td>
<td>0.178*** (0.0566)</td>
<td>0.0449 (0.0735)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) Monthly with 1861</td>
<td>0.173*** (0.0486)</td>
<td>-0.0137 (0.0538)</td>
<td>0.432*** (0.0772)</td>
<td>0.516*** (0.0863)</td>
<td>-0.123 (0.0891)</td>
<td>-0.0789 (0.105)</td>
<td>0.0169 (0.102)</td>
</tr>
<tr>
<td>(4) Monthly no 1861</td>
<td>0.144*** (0.0536)</td>
<td>0.00276 (0.0597)</td>
<td>0.427*** (0.0844)</td>
<td>0.470*** (0.0938)</td>
<td>-0.139 (0.0993)</td>
<td>-0.0621 (0.116)</td>
<td>0.00522 (0.113)</td>
</tr>
<tr>
<td>DV: Number of patents</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5) Poisson with 1861</td>
<td>0.178*** (0.0388)</td>
<td>-0.00805 (0.0448)</td>
<td>0.409*** (0.0592)</td>
<td>0.536*** (0.0698)</td>
<td>-0.119 (0.0736)</td>
<td>-0.137 (0.103)</td>
<td>0.0139 (0.117)</td>
</tr>
<tr>
<td>(6) Poisson no 1861</td>
<td>0.178*** (0.0388)</td>
<td>-0.00805 (0.0448)</td>
<td>0.401*** (0.0651)</td>
<td>0.478*** (0.0773)</td>
<td>-0.118 (0.0806)</td>
<td>-0.154 (0.114)</td>
<td>0.0449 (0.128)</td>
</tr>
<tr>
<td>DV: Log share of total patents</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(7) Share DV with 1861</td>
<td>0.165*** (0.0474)</td>
<td>-0.0313 (0.0497)</td>
<td>0.424*** (0.0687)</td>
<td>0.533*** (0.0855)</td>
<td>-0.0802 (0.0905)</td>
<td>-0.109 (0.114)</td>
<td>-0.0679 (0.145)</td>
</tr>
<tr>
<td>(8) Share DV no 1861</td>
<td>0.143*** (0.0522)</td>
<td>-0.00135 (0.0542)</td>
<td>0.417*** (0.0723)</td>
<td>0.485*** (0.0917)</td>
<td>-0.0865 (0.102)</td>
<td>-0.148 (0.125)</td>
<td>-0.0813 (0.163)</td>
</tr>
<tr>
<td>DV: Log of patent count</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(9) India/US imports</td>
<td>0.0051*** (0.00542)</td>
<td>0.00346 (0.0690)</td>
<td>0.0335** (0.0116)</td>
<td>0.0401** (0.0156)</td>
<td>-0.00709 (0.00822)</td>
<td>-0.0110 (0.0134)</td>
<td>0.0103 (0.0126)</td>
</tr>
<tr>
<td>(10) Log price US cotton</td>
<td>0.0913 (0.0575)</td>
<td>-0.0481 (0.0621)</td>
<td>0.344** (0.128)</td>
<td>0.350* (0.186)</td>
<td>-0.0864 (0.0877)</td>
<td>-0.239* (0.133)</td>
<td>0.0340 (0.139)</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. \(* p<0.01, \** p<0.05, \* p<0.1\). All regressions include include the log of total non-textile patents and a time trend. Annual regressions use data from 1855-1883. The shock period is 1861-1864 for row 1. In row 2, 1861 is dropped from the data and the shock period is 1862-1864. Because the war lasted for only a small part of 1865 it has not been included in the shock period when using annual data. Regressions using monthly data, shown in rows 3-4, include a full set of monthly indicator variables. In row 3 the shock period is April 1861-March 1865. In row 4 1861 is dropped from the data and the shock period is January 1862-March 1865. The regressions in rows 5-6 have Newey-West standard errors with a lag length of 4. In rows 5 and 7 the shock period is Q2 1861-Q1 1865. In rows 6 and 8, 1861 has been dropped from the data. The import data used in the regressions in row 9 are from Ellison (1886). The price data used in the regressions in row 10 are from Mitchell & Deane (1962).

\(^1\)See analysis in Section 1.7.
A.11 Robustness of the estimated impact on spinning technology subcategories

Table A.5 presents regression results paralleling those shown in Table 1.10 but generated using several alternative estimation approaches. The results presented in rows 1-2 are generated using Poisson regressions with and without dropping data for 1861. Row 3 includes results where the share of British imports from India to imports from the U.S. is used as the key explanatory variable in place of the shock period indicator. In row 4, the key explanatory variable is the price of U.S. cotton. The last two rows present regressions where the dependent variable is the share of patents in each subcategory to total patents. All of these specifications provide evidence that the shock period was characterized by an increase in patents of gins and openers/scutchers, and potentially also an increase in carding machines. On the other hand, there is no strong evidence of a similar impact on later-stage cotton textile technologies.

Table A.5: Cotton textile subcategory impacts using alternative specifications

<table>
<thead>
<tr>
<th>Preparatory</th>
<th>Spinning</th>
<th>Finishing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Girs</td>
<td>Openers/scutchers</td>
<td>Carding machines</td>
</tr>
<tr>
<td>Preparatory data. DV: Count of patents in each technology sub-category</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1) Poisson with 1861</td>
<td>1.815***</td>
<td>0.676***</td>
</tr>
<tr>
<td>(2) Poisson drop 1861</td>
<td>1.923***</td>
<td>0.657***</td>
</tr>
<tr>
<td>Annual data. DV: Count of patents in each technology sub-category</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) India/US Imports</td>
<td>1.005***</td>
<td>0.671***</td>
</tr>
<tr>
<td>(4) Log Price US Cotton</td>
<td>13.37***</td>
<td>5.479*</td>
</tr>
<tr>
<td>Quarterly data. DV: Share of total patents x 100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5) Share reg. with 1861</td>
<td>0.374***</td>
<td>0.193***</td>
</tr>
<tr>
<td>(6) Share reg. drop 1861</td>
<td>0.431***</td>
<td>0.183***</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1. Quarterly data for 1855-1870. All regressions include the following controls: count of non-textile patents, a time trend, and a constant. Regressions using quarterly data include a set of quarter indicator variables. The import ratio and price regressions use annual data because these explanatory variables are only available at annual frequency.
A.12 Subcategory regressions for other textile inputs

This appendix presents regression results paralleling those given in Table 1.10, but calculated for wool, linen/flax, or silk-related patents, rather than cotton. Some technologies, such as gins, were not used for certain types of textile inputs.

Table A.6: Effect on wool-related patents in spinning technology subcategories

<table>
<thead>
<tr>
<th>Preparatory</th>
<th>Spinning,</th>
<th>Finishing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS regs. DV: Count of “wool” patents in each subcat.</td>
<td></td>
</tr>
<tr>
<td>Shocks</td>
<td>Obs.</td>
<td></td>
</tr>
<tr>
<td>A. OLS regs.</td>
<td>Shock</td>
<td>Indicator</td>
</tr>
<tr>
<td></td>
<td>0.0267</td>
<td>(0.107)</td>
</tr>
<tr>
<td></td>
<td>-0.166</td>
<td>(0.263)</td>
</tr>
<tr>
<td></td>
<td>-0.551</td>
<td>(0.410)</td>
</tr>
<tr>
<td></td>
<td>-0.549</td>
<td>(0.423)</td>
</tr>
<tr>
<td></td>
<td>0.0727</td>
<td>(0.253)</td>
</tr>
<tr>
<td></td>
<td>-0.302</td>
<td>(0.307)</td>
</tr>
<tr>
<td></td>
<td>0.149</td>
<td>(0.164)</td>
</tr>
<tr>
<td></td>
<td>-0.0759</td>
<td>(0.152)</td>
</tr>
<tr>
<td>B. OLS regs dropping 1861. DV: Count of “wool” patents in each subcat.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shocks</td>
<td>Obs.</td>
<td></td>
</tr>
<tr>
<td>B. OLS regs.</td>
<td>Shock</td>
<td>Indicator</td>
</tr>
<tr>
<td></td>
<td>0.0628</td>
<td>(0.120)</td>
</tr>
<tr>
<td></td>
<td>-0.191</td>
<td>(0.292)</td>
</tr>
<tr>
<td></td>
<td>-0.719</td>
<td>(0.455)</td>
</tr>
<tr>
<td></td>
<td>-0.348</td>
<td>(0.470)</td>
</tr>
<tr>
<td></td>
<td>0.133</td>
<td>(0.283)</td>
</tr>
<tr>
<td></td>
<td>-0.285</td>
<td>(0.340)</td>
</tr>
<tr>
<td></td>
<td>0.0415</td>
<td>(0.175)</td>
</tr>
<tr>
<td></td>
<td>-0.0515</td>
<td>(0.169)</td>
</tr>
<tr>
<td>C. Neg. Binomial regs. DV: Count of “wool” patents in each subcat.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shocks</td>
<td>Obs.</td>
<td></td>
</tr>
<tr>
<td>C. Neg. Binomial regs. DV: Count of “wool” patents in each subcat.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shocks</td>
<td>Obs.</td>
<td></td>
</tr>
<tr>
<td>D. Neg. Binomial regs dropping 1861. DV: Count of “wool” patents in each subcat.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shocks</td>
<td>Obs.</td>
<td></td>
</tr>
</tbody>
</table>

Standard errors in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1. Quarterly data for 1855-1870. All regressions include the following controls: count of non-textile patents, a set of quarter indicator variables, a time trend, and a constant.
Table A.7: Effect on linen-related patents in spinning technology subcategories

<table>
<thead>
<tr>
<th>Preparatory</th>
<th>Spinning</th>
<th>Finishing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gins</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Openers/</td>
<td>Carding</td>
<td>Combing</td>
</tr>
<tr>
<td>scutchers</td>
<td>machines</td>
<td>machines</td>
</tr>
<tr>
<td>Shock</td>
<td>-0.0588</td>
<td>-0.0190</td>
</tr>
<tr>
<td>Indicator</td>
<td>(0.172)</td>
<td>(0.242)</td>
</tr>
<tr>
<td>Obs.</td>
<td>64</td>
<td>64</td>
</tr>
</tbody>
</table>

A. OLS regs. DV: Count of “linen” patents in each subcat.

B. OLS regs dropping 1861. DV: Count of “linen” patents in each subcat.

C. Neg. Binomial regs. DV: Count of “linen” patents in each subcat.

D. Neg. Binomial regs dropping 1861. DV: Count of “linen” patents in each subcat.

x- Gins are not used for preparing linen. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Quarterly data for 1855-1870. All regressions include the following controls: count of non-textile patents, a set of quarter indicator variables, a time trend, and a constant.

Table A.8: Effect on silk-related patents in spinning technology subcategories

<table>
<thead>
<tr>
<th>Preparatory</th>
<th>Spinning</th>
<th>Finishing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gins</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Openers/</td>
<td>Carding</td>
<td>Combing</td>
</tr>
<tr>
<td>scutchers</td>
<td>machines</td>
<td>machines</td>
</tr>
<tr>
<td>Shock</td>
<td>0.00352</td>
<td>-0.157</td>
</tr>
<tr>
<td>Indicator</td>
<td>(0.149)</td>
<td>(0.156)</td>
</tr>
<tr>
<td>Obs.</td>
<td>64</td>
<td>64</td>
</tr>
</tbody>
</table>

A. OLS regs. DV: Count of “silk” patents in each subcat.

B. OLS regs dropping 1861. DV: Count of “silk” patents in each subcat.

C. Neg. Binomial regs. DV: Count of “silk” patents in each subcat.

D. Neg. Binomial regs dropping 1861. DV: Count of “silk” patents in each subcat.

x- Gins are not used for preparing silk. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Quarterly data for 1855-1870. All regressions include the following controls: count of non-textile patents, a set of quarter indicator variables, a time trend, and a constant. Negative Binomial regression results for rollers dropping 1861 are omitted because they fail to converge.
Appendix B

Appendix to Chapter 2

B.1 Industry-level Poisson regression results

This section presents industry-level regression results mirroring those described in the main text except that they are calculated using Poisson regressions rather than standard OLS or IV regressions. Table B.1 presents results obtained when I do not account for the sources of bias I have identified. Regardless of which knowledge stock measure I use, these results suggest a positive and statistically significant relationship between the lagged stock of knowledge and the level of innovation. Table B.2 presents the results obtained when I account for the sources of bias I have identified. Once I control for these sources of bias, these Poisson regression results provide no evidence of path dependence in innovation at the industry level.
Table B.1: Industry-level Poisson regression results obtained when not accounting for sources of bias

<table>
<thead>
<tr>
<th>Knowledge stock measure:</th>
<th>DV: Count of patents</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RECENT (1)</td>
</tr>
<tr>
<td>Lagged knowledge stock</td>
<td>0.00195***</td>
</tr>
<tr>
<td></td>
<td>(0.000715)</td>
</tr>
<tr>
<td>Observations</td>
<td>56</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Included controls: input-specific fixed effects, year-specific indicator variables, constant. Included input types: cotton, wool, linen, and silk textiles.

Table B.2: Industry-level Poisson regression results

<table>
<thead>
<tr>
<th>Knowledge stock measure:</th>
<th>DV: Count of patents</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Reduced form</td>
</tr>
<tr>
<td></td>
<td>POSTSHOCK (1)</td>
</tr>
<tr>
<td>Lagged knowledge stock</td>
<td>-0.000164</td>
</tr>
<tr>
<td></td>
<td>(0.000346)</td>
</tr>
<tr>
<td>Post-shock indicator</td>
<td>-0.0740</td>
</tr>
<tr>
<td>1866-1869</td>
<td>(0.135)</td>
</tr>
<tr>
<td>Observations</td>
<td>68</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Included controls: input-specific fixed effects, input-specific time trends, year-specific indicator variables, constant. Included input types: cotton, wool, linen, and silk textiles. The time period used is 1857-1870 except that, when using the RECENT and PSTOCK variables, the first three years of data must be dropped in order to generate the lagged stock variables.
B.2 Industry-level regression results using only high-quality patents

This appendix presents results which are similar to those presented in Section 2.6.1 except that instead of using patent counts to measure the level of innovation, I use the count of high-quality patents for which the patent renewal fee was paid at year three. The renewal fee data is available from 1856-1869 which limits somewhat the set of data I can use, particularly for the PSTOCK measure of the knowledge stock, which requires that I drop the first four years of the data. Also because only a small number of patents were renewed, I am concerned about the discrete nature of the data influencing my results, so Poisson regression are used. Table B.3 presents results in which I do not account for the sources of bias I have identified. Here I find a positive relationship between the lagged knowledge stock and the level of innovation which is statistically significant in most specifications.

Table B.3: Results obtained when not accounting for sources of bias

<table>
<thead>
<tr>
<th>Knowledge stock measure:</th>
<th>RECENT (1)</th>
<th>EXPER (2)</th>
<th>PSTOCK (3)</th>
<th>PSUM (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lagged knowledge stock</td>
<td>0.00341*</td>
<td>0.00137**</td>
<td>0.0120</td>
<td>0.00512**</td>
</tr>
<tr>
<td></td>
<td>(0.00179)</td>
<td>(0.000556)</td>
<td>(0.0115)</td>
<td>(0.00247)</td>
</tr>
<tr>
<td>Observations</td>
<td>52</td>
<td>52</td>
<td>40</td>
<td>52</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Included controls: input-specific fixed effects, year-specific indicator variables, constant. Included input types: cotton, wool, linen, and silk textiles.

Next, I present industry-level Poisson regression results in which I account for the sources of bias I have identified by including input-specific time trends and using an instrumental variables strategy. Table B.4 describes these results. These results provide no clear evidence in favor of path dependence in innovation, in line with the findings described in Section 2.6.1.
Table B.4: Main industry-level regression results with renewal fee data

<table>
<thead>
<tr>
<th>Knowledge stock measure</th>
<th>Reduced form</th>
<th>Instrumental variables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>POSTSHOCK (1)</td>
<td>RECENT (2)</td>
</tr>
<tr>
<td>Lagged knowledge stock</td>
<td>0.00103</td>
<td>0.00171</td>
</tr>
<tr>
<td></td>
<td>(0.00181)</td>
<td>(0.00309)</td>
</tr>
<tr>
<td>Post-shock indicator</td>
<td>0.174</td>
<td></td>
</tr>
<tr>
<td>1866-1869</td>
<td>(0.991)</td>
<td></td>
</tr>
</tbody>
</table>

Observations 52 52 52 40 52

Standard errors in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1

Included controls: input-specific fixed effects, input-specific time trends, year-specific indicator variables, constant. Included input types: cotton, wool, linen, and silk textiles. The time period used is 1857-1870 except that, when using the RECENT and PSTOCK variables, the first three years of data must be dropped in order to generate the lagged stock variables.

B.3 Industry-level results without cross-listed patents

This appendix presents regression results calculated while dropping all patents which mention multiple input types. These patents make up a substantial fraction of the total number of textile related patents. For example, out of the 1423 patents which mention cotton in the title, 546 also mention either wool, linen, or silk. In the analysis shown in Section 2.6.1, these patents are counted towards both the number of cotton patents and the number of patents for the other textile category. We may be worried that including these cross-listed patents is biasing our results if, for example, a large number of cotton patents also mention other input types, so that improvements in cotton textile technology spill over into the other input categories. Dropping these cross-listed patents should reduce this concern.

Table B.5 presents results in which I do not account for the sources of bias I have identified, as in Table 2.2 in the main text. As before, these results suggest that there is path dependence in innovation, though this finding is not significant for all of the different measures of the knowledge
stock. Table B.6 presents results in which I account for these sources of bias while dropping all cross-listed patents. Here I find no strong evidence of path dependence in innovation, which supports the results presented in the main text.

Table B.5: Results obtained without accounting for sources of bias – without cross-listed patents

<table>
<thead>
<tr>
<th>Knowledge stock measure:</th>
<th>DV: Count of patents</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Log lagged knowledge stock</td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td></td>
<td>0.603**</td>
<td>0.304</td>
<td>0.639**</td>
<td>0.621</td>
</tr>
<tr>
<td></td>
<td>(0.247)</td>
<td>(0.449)</td>
<td>(0.279)</td>
<td>(0.427)</td>
</tr>
<tr>
<td>Observations</td>
<td>56</td>
<td>68</td>
<td>56</td>
<td>68</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.895</td>
<td>0.871</td>
<td>0.893</td>
<td>0.875</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Included controls: input-specific fixed effects, year-specific indicator variables, constant. Included input types: cotton, wool, linen, and silk textiles.

Table B.6: Main industry-level regression results without cross-listed patents

<table>
<thead>
<tr>
<th>Knowledge stock measure:</th>
<th>DV: Count of patents</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Log lagged knowledge stock</td>
<td>POSTSHOCK</td>
<td>RECENT</td>
<td>EXPER</td>
<td>PSTOCK</td>
<td>PSUM</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
</tr>
<tr>
<td></td>
<td>-0.857</td>
<td>-3.791</td>
<td>-0.681</td>
<td>-2.927</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.898)</td>
<td>(5.304)</td>
<td>(0.630)</td>
<td>(2.877)</td>
<td></td>
</tr>
<tr>
<td>Post-shock indicator 1866-1869</td>
<td>-0.308</td>
<td>(0.285)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>68</td>
<td>56</td>
<td>68</td>
<td>56</td>
<td>68</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.922</td>
<td>0.907</td>
<td>0.822</td>
<td>0.927</td>
<td>0.912</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Included controls: input-specific fixed effects, input-specific time trends, year-specific indicator variables, constant. Included input types: cotton, wool, linen, and silk textiles. The time period used is 1854-1870 except that, when using the RECENT and PSTOCK variables, the first three years of data must be dropped in order to generate the lagged stock variables.
B.4 Intermediate-level regression results from OLS and IV regressions

This section provides OLS and IV regression results which mirror the Poisson results presented in the main text analysis at the intermediate level. Table B.7 presents results calculated without accounting for potential sources of bias. Table B.8 presents results calculated when I account for these sources of bias. In both cases, the results are similar to those described in the main text except that my preferred regression results in Table B.8 columns 5-9 are not statistically significant.

Table B.7: Intermediate-level results obtained when not accounting for sources of bias in levels (OLS and IV)

<table>
<thead>
<tr>
<th>Knowledge stock measure:</th>
<th>DV: Count of patents</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RECENT (1)</td>
<td>EXPER (2)</td>
<td>PSTOCK (3)</td>
<td>PSUM (4)</td>
</tr>
<tr>
<td>Lagged knowledge stock</td>
<td>0.0111</td>
<td>-0.00840</td>
<td>0.0586</td>
<td>-0.00727</td>
</tr>
<tr>
<td></td>
<td>(0.0264)</td>
<td>(0.00933)</td>
<td>(0.0425)</td>
<td>(0.00962)</td>
</tr>
<tr>
<td>Observations</td>
<td>144</td>
<td>168</td>
<td>144</td>
<td>168</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.751</td>
<td>0.724</td>
<td>0.755</td>
<td>0.724</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Included controls: subcategory-specific fixed effects, year-specific indicator variables, constant. The time period used is 1856-1876 except that, when using the RECENT and PSTOCK variables, the first three years of data must be dropped in order to generate the lagged stock variables.
Table B.8: Main intermediate-level regression results in levels (OLS and IV)

<table>
<thead>
<tr>
<th>Knowledge stock measure:</th>
<th>Reduced form</th>
<th>Instrumental variables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>POSTSHOCK</td>
<td>RECENT</td>
</tr>
<tr>
<td>Lagged knowledge stock</td>
<td>0.102</td>
<td>0.191</td>
</tr>
<tr>
<td></td>
<td>(0.0823)</td>
<td>(0.193)</td>
</tr>
<tr>
<td>Post-shock indicator</td>
<td>3.032</td>
<td></td>
</tr>
<tr>
<td>1866-1869</td>
<td>(2.611)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>168</td>
<td>144</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.749</td>
<td>0.754</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
Included controls: input-specific fixed effects, input-specific time trends, year-specific indicator variables, constant. The time period used is 1856-1876 except that, when using the RECENT and PSTOCK variables, the first three years of data must be dropped in order to generate the lagged stock variables.

B.5 Intermediate-level regression results with Carding machines as a treated technology

In the main text I have treated carding machines as a non-treated technology. However, Figure 2.12 suggests that the shock may have led to some increase in innovation in carding machine technologies. In this section, I show that my intermediate-level regression results are essentially unchanged if I classify carding as a treated technology category, along with gins and openers/scutchers. Table B.9 shows that when I fail to account for the sources of bias I do not find any clear relationship between the stock of knowledge and the level of innovation at the intermediate level.
Table B.9: Intermediate-level results with Carding as a treated technology category obtained when not accounting for sources of bias (Poisson regressions)

<table>
<thead>
<tr>
<th>Knowledge stock measure:</th>
<th>DV: Count of patents</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RECENT (1)</td>
</tr>
<tr>
<td>Lagged knowledge stock</td>
<td>0.00194</td>
</tr>
<tr>
<td></td>
<td>(0.00120)</td>
</tr>
<tr>
<td>Observations</td>
<td>144</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1.
Included controls: subcategory-specific fixed effects, year-specific indicator variables, constant. The time period used is 1856-1876 except that, when using the RECENT and PSTOCK variables, the first three years of data must be dropped in order to generate the lagged stock variables.

Table B.10: Intermediate-level regression with Carding as a treated technology category (Poisson regressions)

<table>
<thead>
<tr>
<th>Knowledge stock measure:</th>
<th>DV: Count of patents</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Reduced form</td>
</tr>
<tr>
<td></td>
<td>POSTSHOCK (1)</td>
</tr>
<tr>
<td>Lagged knowledge stock</td>
<td>0.0185**</td>
</tr>
<tr>
<td></td>
<td>(0.00880)</td>
</tr>
<tr>
<td>Post-shock indicator 1866-1869</td>
<td>0.338***</td>
</tr>
<tr>
<td></td>
<td>(0.102)</td>
</tr>
<tr>
<td>Observations</td>
<td>168</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1
Included controls: subcategory-specific fixed effects, subcategory-specific time trends, year-specific indicator variables, constant. The time period used is 1856-1876 except that, when using the RECENT and PSTOCK variables, the first three years of data must be dropped in order to generate the lagged stock variables.
Appendix C

Appendix to Chapter 3

This appendix provides additional information that may be helpful in understanding and replicating the results presented in the paper. It is divided into four sections. The first section provides some additional background information on the empirical setting. The second section provides additional information on the procedures used to construct the data. The third section presents additional results and proofs related to the theory. The fourth section provides supplementary statistical results.

C.1 Empirical setting appendix

Figure C.1 was adapted from Catling (1986) and shows the variation in different production activities across major cotton-textile producing towns considered in this study. This map shows that weaving was concentrated in the northern towns, Preston and Blackburn, spinning was concentrated in the southern towns of Bolton, Oldham, and Stockport, and most finishing was done in Manchester.
Figure C.1: Distribution of cotton textile manufacturing stages in Lancashire towns

Figure C.2 presents data describing total British cotton imports, and British cotton imports from the U.S., from 1815-1910. This chart shows that, just prior to the start of the Civil War, the majority of Britain’s cotton supplies were coming from the U.S., but that these imports dropped nearly to zero during most of the 1861-1865 period. While imports from other suppliers increased, they were not able to make up for the drop in U.S. cotton, leading to a sharp drop in overall cotton imports during the Civil War. However, imports quickly rebounded following the end of hostilities.

Figure C.2: Total British cotton imports and imports from the U.S. 1815-1910

This paper argues that the shock was primarily industry-specific. Figure C.3 provides data
supporting this argument. The left-hand panel shows that there was no visible effect on total British imports, or British raw material imports, once raw materials for textiles are excluded. The right-hand panel shows that, once textile exports are excluded, the shock does not appear to have effected British exports of manufactured goods.

Figure C.3: British imports and exports 1851-1869

![British Imports vs British Exports](image1)

Data from Mitchell (1988).

Figure C.4: British wool imports and exports 1850-1880

![British Wool Exports vs Imports](image2)

Data from Mitchell & Deane (1962).
C.2 Data appendix

C.2.1 Census data

The census data were taken from the original Census Enumerator Reports for 1851-1891, which are available at the British Library.

Town-level data

Town level data for these years and counties was available only for the towns used in the analysis, plus Liverpool. Salford is treated as part of Manchester, consistent with their very close proximity. Data for males and females were combined in all analysis. The data for 1851-1861 are available divided into workers over 19 and workers under 20, by occupation. In 1871, data by occupation are available only for workers over 19, while in 1881-1891 data are only available for all workers. It is, therefore, necessary to estimate values for 1871 employees under 20, as this was an important fraction of the labor force at this time. This is done by calculating the average ratio of all employees, to employees over 20, in each industry and location, in 1851 and 1861, and then multiplying this value by the number of employees in each industry and location in 1871, to obtain 1871 values that are consistent with the other years.

Matching occupations/industries over time

Industry categories changed over time, so it was necessary to combine multiple industries in order to construct more consistent industry groupings over the study period. Individual categories in the years were combined into industry groups based on (1) the census’ occupation classes and (2) the name of the occupation. For some occupations, it was not possible to form consistent groups over all years, and these occupations are omitted from the analysis. It is still the case that some occupation groups may not be perfectly consistent over time, however, because my identification strategy controls for aggregate industry-year effects, it should be less affected by this source of measurement error. The occupational categories prior to 1851 and after 1891 change significantly, relative to the period that I study, which motivates my choice of study period.
Calculating geographic coagglomeration

The measure of geographic coagglomeration is based on Ellison et al. (2010) and is calculated using data from 71 district within Lancashire. The use of data from only Lancashire here is important. There are two good reasons for taking this approach. First, it is the pattern of relatedness in Lancashire that determines which industries are affected by the shock. Thus far, little is known about the extent to which deep patterns of industry relatedness vary across locations. Using only one location minimizes this concern. Second, there is a danger in calculating coagglomeration measures over too broad an area. To see why, consider some industry x which is related to both the cotton and wool textile industries and is coagglomerated with cotton textiles in Lancashire and wool textiles in Yorkshire. If the coagglomeration measures is calculated using only Lancashire county, I find that industry x is coagglomerated with cotton textiles. If data from Yorkshire are added, there will be a sharp reduction in the level of coagglomeration of industry x with cotton textiles, because there are now concentrations of industry x in Yorkshire districts with very little cotton textile production (but a lot of wool textile production). To calculate the measure, I suppose that there are M districts, and let $s_{mi}$ be the share of industry i’s employment in some district m. Let $x_m$ be district m’s share of workers. Geographic coagglomeration between industries i and j is given by $R_{ij}$, where,

$$R_{ij} = \frac{\sum_{m=1}^{M} (s_{mi} - x_m)(s_{mj} - x_m)}{1 - \sum_{m=1}^{M} x_m^2}$$

Table C.1 gives the fifteen most and least related industries to the cotton textile industry, according to the coagglomeration measure. Note that, for illustration, this table includes cotton-based industries which are later excluded from the analysis.
Table C.1: Most and least coagglomerated industries with cotton textiles

<table>
<thead>
<tr>
<th>Rank</th>
<th>Most coagglomerated</th>
<th>Least coagglomerated</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Cotton textile manufacturing</td>
<td>Ship transport</td>
</tr>
<tr>
<td>2</td>
<td>Cotton textile finishing</td>
<td>Sugar refining</td>
</tr>
<tr>
<td>3</td>
<td>Paper manufacturing</td>
<td>Shipbuilding</td>
</tr>
<tr>
<td>4</td>
<td>Hat making</td>
<td>Hemp, sacking, and sailcloth</td>
</tr>
<tr>
<td>5</td>
<td>Worsted and stuff manufacturing</td>
<td>Cooper</td>
</tr>
<tr>
<td>6</td>
<td>Coal mining</td>
<td>Silk mercer</td>
</tr>
<tr>
<td>7</td>
<td>Fuller (wool textile finishing)</td>
<td>Soap maker</td>
</tr>
<tr>
<td>8</td>
<td>Woollen textile manufacturing</td>
<td>Hemp, jute manufacture</td>
</tr>
<tr>
<td>9</td>
<td>Stone worker</td>
<td>Musical instrument maker</td>
</tr>
<tr>
<td>10</td>
<td>Toll collector</td>
<td>Lodging house keeper</td>
</tr>
<tr>
<td>11</td>
<td>Weaver, misc.</td>
<td>Tobacconist</td>
</tr>
<tr>
<td>12</td>
<td>Thread manufacturing</td>
<td>Messenger, porter</td>
</tr>
<tr>
<td>13</td>
<td>Engine and machine maker</td>
<td>Artificial flower maker</td>
</tr>
<tr>
<td>14</td>
<td>Road construction</td>
<td>Timber cutting</td>
</tr>
<tr>
<td>15</td>
<td>Brick making</td>
<td>Watch making</td>
</tr>
</tbody>
</table>

C.2.2 Input-output matrix

Relatedness between industries based on the input-output matrix can be measured using either upstream, downstream, or both types of connections. Downstream connections from cotton textiles to another industry are measured by the share of that industry’s inputs composed of cotton textile outputs. Similarly, upstream connections between cotton textiles and its supplier industries are measured as the share of the supplier industry’s output sold to the cotton textile industry. These two measures are combined to obtain the input-output connections measure used in the analysis. This matches the approach used in Ellison et al. (2010).

The input-output matrix used in this project is derived from that produced by Thomas (1987), which includes 41 categories. While this matrix is generally suitable for the project at hand, one revision is necessary. Thomas’ matrix combined the cotton and silk textile industries. If I treat silk as if it were the same as cotton, I end up with an unrealistically high measure of the input-output linkages between these two industries. Therefore, it was necessary to separate the silk and cotton industries, creating one additional entry in the input-output matrix. Thomas’ sources and methodology were followed as closely as possible in separating these two industries. The primary
information used to create the separate silk entry came from the same source used by Thomas in the original matrix, the 1907 Census of Production.

**Relatedness measure distributions**

The following four figures describe the distribution of industry relatedness arising from the coagglomeration and input-output relatedness measures. The left-hand panels provide the raw distribution, while the right-hand panels present the distribution where each industry has been weighted by its employment in 1851.

Figure C.5: Histograms describing the coagglomeration relatedness measure (171 industries)

![Histogram of Industries](image1)

![Histogram Weighted by 1851 Employment](image2)

Figure C.6: Histograms describing the input-output relatedness measure (43 sectors)

![Histogram of Industries](image3)

![Histogram Weighted by 1851 Employment](image4)

**Summary statistics for analysis data**

The table below presents simple summary statistics for the data used in the primary analysis. Note that the continuous variables have been standardized to have a mean of zero and standard deviation of one.
Table C.2: Summary statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ln(Emp.)- Town level</td>
<td>6907</td>
<td>3.88</td>
<td>2.03</td>
<td>0</td>
<td>10.48</td>
</tr>
<tr>
<td>Ln(Emp.)- County level</td>
<td>1710</td>
<td>7.08</td>
<td>1.80</td>
<td>0</td>
<td>12.15</td>
</tr>
<tr>
<td>Coag Input-Output</td>
<td>42</td>
<td>0</td>
<td>1.00</td>
<td>-0.35</td>
<td>4.37</td>
</tr>
<tr>
<td>Shock severity</td>
<td>11</td>
<td>0</td>
<td>1.00</td>
<td>-0.86</td>
<td>1.77</td>
</tr>
<tr>
<td>1851 cotton emp. share</td>
<td>11</td>
<td>0</td>
<td>1.01</td>
<td>-0.88</td>
<td>1.29</td>
</tr>
</tbody>
</table>

C.3 Theoretical appendix

This section presents some additional work related to the theory. The first subsection provides a more rigorous proof of equilibrium existence for the theoretical model. Next, I show that the cost shock formulation used in the model can be derived from a micro-founded model. Finally, I present more rigorous proofs supporting the assertions made in Obs. 1.

C.3.1 Derivation of the shock period cotton textile price equation

This subsection describes how the reduced form approach used to introduce the cost shock into the model can be motivated by a model that explicitly incorporates intermediate inputs. The formulation used to introduce the cost shock into the model is given in Equation 3.6 and reproduced below (omitting time and location subscripts).

\[ q_C = \frac{w}{\bar{A}_C} + \phi \]

This formulation can be derived from a model in which production requires both labor, L, and an intermediate input good, I, such as raw cotton. In particular, consider a model in which these inputs are both necessary and must be used in fixed proportions, adjusted for labor productivity \( A_C \). The new production function is \( X_C = \min(\bar{A}_C L_C, \beta I_C) \) where \( \beta \) determines the ratio of intermediate input to labor. Let \( \psi_C \) be the exogenous price of the intermediate input and let \( \tilde{q}_C \) be the sales price of the finished good. Producer’s f’s optimization problem is then,
\[
\max_{L_{cf}, L_{cf}} \tilde{q}_C X_C - wL_C - \psi_C I_c
\]

Solving this maximization problem, I obtain the following price charged by perfectly competitive producers in industry C.

\[
\tilde{q}_C = \frac{w}{A_C} + \frac{\psi_C}{\beta}
\]

Decomposing the intermediate input price into some baseline price \(\psi_{C0}\) and some price increase due to the shock \(\Delta \psi_C\), I have the following.

\[
\tilde{q}_C = \frac{w}{A_C} + \frac{\psi_{C0}}{\beta} + \frac{\Delta \psi_C}{\beta}
\]

Now let \(\phi = \Delta \psi_C / \beta\) and define \(q_e = \tilde{q}_e - \frac{\psi_{C0}}{\beta}\) to obtain the formulation used in Equation 3.6.

Using a fixed-input proportions production function here may seem somewhat restrictive, but it is a reasonably good approximation of the empirical setting. Given the nature of cotton textile production, it was hard for producers to significantly reduce the share of raw cotton inputs in production, though efforts were made to reduce waste cotton in the production process during the period in which raw cotton prices were high. Note that in this formulation, technological progress acts to reduce the ratio of labor to raw cotton inputs required in production, which also fits the empirical setting well, and increases the flexibility of the production function.

C.3.2 Proof of Obs. 1

Part 1 – Effect of the shock on industry C in the shock period

Begin by noting that \(dq_{Cl}/d\phi = 1\). Taking this derivative and reorganizing I obtain the following, where the inequality follows from the fact that \(A_{CH} > A_{CF}\) implies \(q_{CH} < q_{CF}\).
\[
\frac{dL_{CH}}{d\phi} = \frac{E_i q_i \alpha A_{CH}^{-1}}{(q_{CH}^{-\sigma} + q_{CF}^{-\sigma})^2} \left[ -q_{CH}^{-\sigma} \frac{\sigma - 1}{q_{CF}^{-\sigma}} - \frac{\sigma}{q_{CF}^{-\sigma} q_{CH}^{-\sigma}} \right] < 0
\]

Second, I want to show that \( \frac{d(\ln(L_{CH}) - \ln(L_{CF}))}{d\phi} < 0 \). I start with the difference in log employment between the two locations.

\[
\ln(L_{CH}) - \ln(L_{CF}) = -\sigma (\ln(q_{CH}) - \ln(q_{CF})) - (A_{CH} - A_{CF})
\]

Taking the derivative gives me the following.

\[
\frac{d(\ln(L_{CH}) - \ln(L_{CF}))}{d\phi} = -\sigma \left[ \frac{1}{q_{CH}} - \frac{1}{q_{CF}} \right] < 0
\]

### C.3.3 Proof of Obs. 3

We must show that when \( \tau^{iC} > 0 \) and \( A_{CH} > A_{CF} \), then \( \frac{dS_{iH}}{d\phi} > \frac{dS_{iF}}{d\phi} \), where I ignore subscripts since all variables are for period s. Taking the derivatives, this will be true when,

\[
\frac{\tau^{iC} L_{CH}}{L_{CH} + 1} \frac{dL_{CH}}{d\phi} < \frac{\tau^{iC} L_{CF}}{L_{CF} + 1} \frac{dL_{CF}}{d\phi}
\]

Given that \( L_{CH} > L_{CF} \), which follows from \( A_{CH} > A_{CF} \) and Equation 3.3, it will be sufficient to show the following.

\[
\frac{dL_{CH}/d\phi}{L_{CH}} < \frac{dL_{CF}/d\phi}{L_{CF}}
\]

Using Equations 3.3 and taking derivatives, this is equivalent to showing,

\[
\frac{1}{q_{CH}^{-\sigma}} + \frac{1}{q_{CF}^{-\sigma} q_{CH}^{-\sigma}} > \frac{1}{q_{CF}^{-\sigma}} + \frac{1}{q_{CH}^{-\sigma} q_{CF}^{-\sigma}}
\]

Multiplying through by \( q_{CF}^{\sigma} \) I obtain the following, where the inequality must hold given that \( A_{CH} > A_{CF} \) implies \( q_{CH} < q_{CF} \).
\[
\left( \frac{q_{CF}}{q_{CH}} \right)^{\sigma} + \left( \frac{q_{CF}}{q_{CH}} \right) > 1 + \left( \frac{q_{CF}}{q_{CH}} \right)^{\sigma-1}
\]

C.4 Additional statistical results

C.4.1 Impact on the cotton textile industry

We have seen that overall output in the cotton textile industry quickly rebounded following the end of the Civil War. Here, I explore whether the distribution of this industry across geographic locations experienced long-term changes as a result of the cotton shortage. I begin with Figure C.7, which describes cotton textile employment in towns, where towns have been grouped into low, medium, and high groups, based on each towns 1851 cotton textile employment share (our instrument for the shock intensity). I see that those towns with a large cotton textile production share in 1851 (Blackburn, Bolton, and Preston), which were also those experiencing some of the most severe shock effects during 1861-1865, saw a continued rapid expansion of cotton textile production, though the graph also suggests that the pace of expansion may have slowed somewhat in these towns after 1861. Towns with a moderate share of initial cotton textile production in employment in 1851 (Manchester, Oldham, and Stockport) show a decline, driven largely by manufacturing being pushed out of Manchester by other commercial activities. Those with a low level of initial cotton textile employment (Yorkshire towns) experience continued growth, albeit from a much lower level. Thus, no clear pattern of geographic redistribution in this industry appears.
Next, I look for long-term effects statistically using an approach similar to that applied to the related industries. In particular, I run a regression where the dependent variable is log employment in each location. The key independent variable is the intensity of the shock in each location. Results for these regressions are presented in Table C.3 below. They provide weak (not statistically significant) evidence that growth in the cotton textile industry may have also been slower in more severely impacted towns in the post shock period. This may indicate that the shock had a long-term impact on the cotton textile industry in these locations. However, there are a couple of facts that caution against taking these results too seriously. First, the 1860-1861 period was one of the most successful ever for this industry, which may inflate the 1851-1861 growth estimates while reducing the 1861-1871 estimates. Second, these results may simply indicate that because of its high concentration in high-impact industries, these results may simply indicate that the industry had reached the point of diminishing returns to concentration in these locations. For example, large mills were unlikely to open in the city of Manchester because of the difficulty and expense of finding the amount of space needed for the large modern mills at this time. Instead, mills often preferred to open in smaller surrounding towns where they could still benefit from their proximity to markets and supporting industries but at lower cost. Third, it may be that the more severely impacted towns experienced broad negative impacts from the shock. Because these regressions
involve only one industry, I am unable to control for aggregate town-level changes that may have been caused by the shock.

Table C.3: Effects of the shock on the cotton textile industry in each location

<table>
<thead>
<tr>
<th>DV: Industry Log Emp.</th>
<th>DV: Chg. in Industry Log Emp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline (1)</td>
<td>Reduced Form (2)</td>
</tr>
<tr>
<td>Shock Int. × Post</td>
<td>-0.261</td>
</tr>
<tr>
<td>(0.172)</td>
<td>(0.187)</td>
</tr>
<tr>
<td>Shock Int. × TT</td>
<td>-0.0904</td>
</tr>
<tr>
<td>(0.0746)</td>
<td>(0.0849)</td>
</tr>
<tr>
<td>Year dummies</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>55</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses  *** p<0.01, ** p<0.05, * p<0.1

C.4.2 First-stage regression results

This section presents first-stage regression results for the main fixed effect and first-differences regressions presented in the main body, as well as the main robustness regressions. As all of these results show, using each location’s 1851 cotton textile employment share provides a strong instrument for the shock intensity in each location.

Table C.4: First stage regression results for IV’s in Table 3.4

<table>
<thead>
<tr>
<th>Coag Only Reg.</th>
<th>IO Only Reg.</th>
<th>Both Reg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coag × Shk × Post</td>
<td>IO × Shk × Post</td>
<td>Coag × Shk × Post</td>
</tr>
<tr>
<td>Coag Only Reg.</td>
<td>.7440591***</td>
<td>.7441289***</td>
</tr>
<tr>
<td>(0.0340811)</td>
<td>(.0340234)</td>
<td>(.0513354)</td>
</tr>
<tr>
<td>IO Only Reg.</td>
<td>.7386922***</td>
<td>-.0022852</td>
</tr>
<tr>
<td>(.0513013)</td>
<td>(.0322577)</td>
<td>(.0513354)</td>
</tr>
<tr>
<td>Both Reg.</td>
<td>7,715</td>
<td>7,715</td>
</tr>
<tr>
<td>476.64</td>
<td>207.33</td>
<td>239.77</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses  *** p<0.01, ** p<0.05, * p<0.1
Table C.5: First stage regression results for IV’s in Table 3.6

<table>
<thead>
<tr>
<th></th>
<th>Coag Only Reg.</th>
<th>IO Only Reg.</th>
<th>Both Reg.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coag<em>Shk</em>Post</td>
<td>IO<em>Shk</em>Post</td>
<td>Coag<em>Shk</em>Post</td>
</tr>
<tr>
<td>Coag × Shk × Post</td>
<td>.7440527*** (.0345327)</td>
<td>.7441208*** (.0344749)</td>
<td>-.0003971 (.0076608)</td>
</tr>
<tr>
<td>IO × Shk × Post</td>
<td>.7385605*** (.0520102)</td>
<td>-.002227 (.0326903)</td>
<td>.7386125*** (.0520462)</td>
</tr>
<tr>
<td>Obs</td>
<td>6,172</td>
<td>6,172</td>
<td>6,172</td>
</tr>
<tr>
<td>F-stat</td>
<td>464.25</td>
<td>201.65</td>
<td>233.52</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses  *** p<0.01, ** p<0.05, * p<0.1

Table C.6: First stage regression results for IV’s in Table 3.9 column 2

<table>
<thead>
<tr>
<th></th>
<th>Coag * Shock Int. * Post</th>
<th>IO * Shock Int. * Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coag × Shock Int. × Post</td>
<td>.7624628*** (.0341529)</td>
<td>-.0003798 (.0075552)</td>
</tr>
<tr>
<td>IO × Shock Int. × Post</td>
<td>-.002289 (.0322721)</td>
<td>.756377*** (.0514424)</td>
</tr>
<tr>
<td>Observations</td>
<td>8,535</td>
<td>8,535</td>
</tr>
<tr>
<td>F-stat</td>
<td>249.84</td>
<td>108.27</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses  *** p<0.01, ** p<0.05, * p<0.1

Table C.7: First stage regression results for IV’s in Table 3.9 column 4

<table>
<thead>
<tr>
<th></th>
<th>Coag * Shock Int. * Post</th>
<th>IO * Shock Int. * Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coag × Shock Int. × Post</td>
<td>.7319098*** (.0283006)</td>
<td>-.0009566 (.008672)</td>
</tr>
<tr>
<td>IO × Shock Int. × Post</td>
<td>-.0037146 (.0389476)</td>
<td>.7387571*** (.0513059)</td>
</tr>
<tr>
<td>Observations</td>
<td>7,715</td>
<td>7,715</td>
</tr>
<tr>
<td>F-stat</td>
<td>342.30</td>
<td>103.67</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses  *** p<0.01, ** p<0.05, * p<0.1
Table C.8: First stage regression results for IV’s in Table 3.9 column 6

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Coag × Shock Int. × Post</td>
<td>.7445251***</td>
<td>-.0004399</td>
</tr>
<tr>
<td></td>
<td>(.0349962)</td>
<td>(.0080869)</td>
</tr>
<tr>
<td>IO * Shock Int. * Post</td>
<td>-.0025286</td>
<td>.7384876***</td>
</tr>
<tr>
<td></td>
<td>(.0323037)</td>
<td>(.0512362)</td>
</tr>
</tbody>
</table>

Observations 7,435 7,435
F-stat 226.72 104.14

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table C.9: First stage regression results for IV’s in Table 3.9 column 8

<table>
<thead>
<tr>
<th>Without IO Inds. (col 8)</th>
<th>Coag * Shock Int. * Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coag × Shock Int. × Post</td>
<td>.761904***</td>
</tr>
<tr>
<td></td>
<td>(.0421449)</td>
</tr>
</tbody>
</table>

Observations 4,320
F-stat 326.82

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

C.4.3 Alternative clustering of standard errors

The results presented in the main text were calculated while clustering standard errors at the industry-location level. This allowed for correlation between the standard errors within each industry-location over time. However, it does not allow for correlation in the standard errors of the same industry across locations, or for different industries in the same location. The motivation for this approach was to cluster at the industry-location level to deal with serial correlation issues while hoping that the inclusion of industry-period and location-period dummy variables reduced correlation among error terms within industries or locations. I can check the robustness of this approach by calculating additional results while clustering at either the industry level or the location level. These approaches will be more restrictive, since they will reduce the number of clusters to 171 industries, or 11 towns, respectively. Table C.10 presents these robustness checks. Columns (1)-(4) show results with standard errors clustered at the industry level. Baseline (columns (1) and
(3)) and reduced form results (columns (2) and (4)) are included, but IV results are not calculated because of the technical difficulties involved in calculating fixed effect IV results using clustered standard errors that include multiple units of the panel variable. These columns show that clustering at the industry level does not meaningfully change the results. Columns (5)-(8) present results when standard errors are clustered at the town level for the 11 towns. In this much more restrictive specification the estimated coefficients are now only marginally statistically significant, though the estimated coefficients continue to show a negative impact that is statistically significant at the 90% confidence level.

Table C.10: Checking robustness to alternative clustering of standard errors

<table>
<thead>
<tr>
<th></th>
<th>Clustered by industry</th>
<th></th>
<th>Clustered by location</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline</td>
<td>Reduced form</td>
<td>Baseline</td>
<td>Reduced form</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Coag × Shk × Post</td>
<td>-0.0486**</td>
<td>-0.0600***</td>
<td>-0.0486**</td>
<td>-0.0600*</td>
</tr>
<tr>
<td></td>
<td>(0.0223)</td>
<td>(0.0213)</td>
<td>(0.0217)</td>
<td>(0.0273)</td>
</tr>
<tr>
<td>IO × Shk × Post</td>
<td>0.0536*</td>
<td>0.0238</td>
<td>0.0536</td>
<td>0.0238</td>
</tr>
<tr>
<td></td>
<td>(0.0290)</td>
<td>(0.0354)</td>
<td>(0.0328)</td>
<td>(0.0394)</td>
</tr>
<tr>
<td>Coag × Shk × TT</td>
<td>-0.0235**</td>
<td>-0.0287***</td>
<td>-0.0235*</td>
<td>-0.0287*</td>
</tr>
<tr>
<td></td>
<td>(0.0118)</td>
<td>(0.0109)</td>
<td>(0.0110)</td>
<td>(0.0136)</td>
</tr>
<tr>
<td>IO × Shk × TT</td>
<td>0.0183</td>
<td>0.00189</td>
<td>0.0183</td>
<td>0.00189</td>
</tr>
<tr>
<td></td>
<td>(0.0132)</td>
<td>(0.0165)</td>
<td>(0.0140)</td>
<td>(0.0160)</td>
</tr>
<tr>
<td>Obs</td>
<td>7,715</td>
<td>7,715</td>
<td>7,715</td>
<td>7,715</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1