Vertical and Horizontal Integration in Imperial Russian Cotton Textiles, 1894-1900

Amanda G. Gregg*,†

Draft: March 25, 2016

Preliminary and incomplete. Please do not cite or circulate.

Abstract

When do firms produce their own inputs instead of purchasing them on the market? In one explanation firms engage in vertical integration to save the cost of transacting on the market, especially when markets are thinner and therefore price risk is greater (Coase 1937). On the other hand, firms that wish to vertically or horizontally integrate may be unable to do if they face financial constraints, because integration requires additional capital. This paper finds evidence for a thin markets explanation of integration within the Russian cotton textile industry in 1894 and 1900, though capital-intensive industries like spinning required financial resources. The 1894 data describe firms' horizontal and vertical integration in especially rich detail. Vertically and horizontally integrated factories were larger in terms of number of workers and tended to be located outside of European Russia in Siberia or the Far East, where markets were thinner. Vertically integrated firms were older, had more workers and machine power, and produced more revenue per worker given the same machine power.

*Middlebury College Department of Economics. Warner Hall Room 305F, 303 College Street, Middlebury, VT 05753. agregg@middlebury.edu.

†This research was supported by grants from the Economic History Association, the National Science Foundation, the Yale MacMillan Center, and the Yale Program in Economic History. Thanks to Timothy Guinnane, Naomi Lamoreaux, and participants of the Yale Economic History Lunch and the 2015 World Economic History Congress for helpful comments. The staff of the National Library of Russia and the Russian State Historical Archive also provided kind and patient assistance.

1 Introduction

Vertical integration can reduce transaction costs and alleviate uncertainty, but its effects are not always positive: integration requires additional coordination within the firm, which may be costly, and it may reduce competition. In the Russian Empire, large vertically integrated factories and firms coexisted with many highly specialized, atomistic factories. This paper seeks to explain patterns of factory organization in the Russian Empire by focusing on cotton textiles, a capital-intensive industry that possessed remarkable variation in factory organization.

I document the characteristics of horizontally and vertically integrated cotton textile factories and firms using a newly collected database of manufacturing establishments. I find that vertically integrated factories and firms were older, had more workers and more machine power, and tended to be located outside of European Russia, i.e. in Siberia or the Russian Far East, regions far removed from denser markets. Vertically integrated firms and factories produced more revenue given the same workers and machine power, and the variance of the distribution of productivity for vertically integrated factories and firms was smaller than for non-integrated factories or firms. The paper also links factories to the firms that owned them; I find that factories owned by multi-factory firms had more workers and were more likely to be located outside of European Russia.

Theories explaining vertical integration emphasize transaction costs, for example contracting costs, and financial market development. According to Coase (1937), firms engage in vertical integration to save the cost of transacting on the market, and they expand until transactions cost savings equals the cost of managing a larger firm. Williamson's (1985)

explanation is similar: greater transactions costs, or less perfect markets, encourage vertical integration. Brown's (1992) study of German cotton textile firms largely confirms the Coase hypothesis. Brown argues that German firms before World War I tended to be highly vertically integrated since a protective tariff wall kept German markets thin, exposing firms to price risk for inputs and outputs. The Russian Empire also enforced high protective tariffs and included some very distant, less-integrated markets: indeed, I find that factories and firms located on the Empire's periphery, where markets were thinner, tended to be more vertically and horizontally integrated.

Much recent literature argues that vertical integration relates to capital markets imperfections, but the effect of better capital markets could be positive or negative.[1] Flawed capital markets may encourage integration if one result of little capital is fewer firms and hence thinner markets. On the other hand, if firms are unable to obtain capital, firms cannot acquire down-stream or up-stream processes, so there may be less integration. This paper's results demonstrate that that factories and firms outside of European Russia were more horizontally and vertically integrated: assuming capital markets were less developed outside of European Russian, those firms that managed to integrate outside of European Russia do not appear to have faced insurmountable capital shortages. Furthermore, I find little evidence that corporations were more vertically or horizontally integrated in the cotton industry overall, despite their access to additional capital markets, which further emphasizes the im-

[1] Acemoglu et al (2009) argue in a cross-country regression framework that countries with both greater contracting costs and great financial development have more vertical integration. Macchiavello (2012) takes into account that the size distribution of firms varies by industry. In this theory, entry leads to more competition, which reduces vertical integration in the largest firms but also forces smaller, dis-integrated firms to exit. The author predicts that "higher financial development reduces vertical integration in industries where a high share of output is produced by small firms" (Page 1).

portance of thin markets.[2] However, corporations dominated particularly capital-intensive branches of the cotton industry like spinning. Shortages of long-term capital in Russia may have limited vertical integration into certain stages of cotton production.

This paper proceeds as follows: Section 2 provides background on the Russian industrial sector and the cotton textile industry in particular. Section 3 outlines predictions about the shape of vertical and horizontal integration in the Russian textile industry based on theories of integration. The next section, Section 4, describes the Imperial Russian factory database. Sections 5, 6, and 7 present results and additional estimates. Because I observe these factories and firms at two cross sections, I cannot argue that certain features of the Russian economy or certain factory characteristics caused vertical or horizontal integration, nor can I argue that integration produced certain characteristics. The goal of this paper, rather, is to document the characteristics of the largest and most integrated factories and firms. The two striking facts that emerge from this description, that factories on the Russian periphery tended to be integrated and that incorporation drove integration into spinning, motivate several lines of future research, which I outline in Section 8 .

2 The Russian Cotton Textile Industry: History and Patterns of Organization

Cotton textile production was among the most productive, valuable, and technologically advanced industries in the Russian Empire. The industry emerged in the eighteenth century

[2]Detailed information on the development of Russian credit and capital markets could shed light on the debate.

and grew rapidly throughout the nineteenth century, and unlike many Russian industries, the Imperial government interfered with the cotton industry relatively little.[3] Table 1 compares the cotton industry to other industries in the Russian Empire based on the population of factories from the 1894 and 1900 manufacturing censuses. Although the cotton industry represented only 4.21 percent of factories in the data, these factories accounted for more more twenty percent of the total revenue. The industry's factories were some of the most productive as measured by revenue per worker, and only the paper, foods, and wood industries had more machine power per worker.

This study focuses on Russian cotton factories, because the industry was more vertically and horizontally integrated on average and displayed a great deal of variation in integration as well. Table 1 shows the cotton industry's remarkable level of vertical integration. The cotton industry possessed the highest average number of different productive activities per factory in the 1894 census.[4] Cotton factories also displayed great variation in integration: the cotton industry had the highest standard deviation in the number of activities per factory.

The major branches of Russia's cotton textile industry included weaving, spinning, and chintz printing. Russia's cotton textile industry developed according to a similar pattern seen in many contexts: the final stage of cloth production, fabric printing, developed first, and the earlier stages such as the production of yarn (spinning) appeared only much later.[5] The village of Ivanovo in Vladimir province, part of the Central Industrial region, represented the

[3]Tugan-Baranovsky, "Russian Factory," 49-50. Tugan-Baranovsky also documents how, even in the eighteenth century, the cotton industry used relatively little serf labor and argues that use of free labor contributed to its high level of development (64).

[4]In a Tobit regression left-censored at zero of log number of activities on log number of workers and a dummy variable for the cotton industry, the cotton coefficient is large and statistically significant (.20 with a .05 standard error.)

[5]Tugan-Baranovsky, "Russian Factory," 48.

Empire's center of cotton weaving. Cloth printing and other branches of the cotton industry also tended to be located in the Central Industrial Region.[6] Finally, the development of the Russian railroad network in the second half of the nineteenth century enabled the movement of raw cotton from Turkestan to the central provinces.[7] The production of raw cotton was located primarily outside of European Russia.

Certain branches of cotton textile production required large, long-term capital investments. Imperial Russian firms faced significant barriers to acquiring long-term capital.[8] Even Gerschenkron (1962) argued that the Russian industrial sector had been so held back by weak capital markets and by serfdom that the state had to substitute for private capital to stimulate industrial growth.[9] For many firms, incorporation provided access to important additional sources of long term capital. Corporations could sell shares on stock markets, and incorporation guaranteed limited liability to investors. From the eighteenth century to the first World War, however, Russian firms only acquired corporate charters with special permission granted by the Ministry of Finance and, in the final step of the process, by the Tsar himself. Firms that incorporated chose to bear the costs of costly incorporation in return for access to scarce long-term capital.

The Russian cotton industry had the highest proportion of corporations of any industrial branch, and corporations displayed many differences when compared to non-incorporated

[6]Ibid., 173.

[7]Ibid., 292.

[8]Russia's developing banking sector, for example, focused on short-term and medium-term credit, not long-term investments (Gregg 2015, Crisp 1976).

[9]Gerschenkron, p. 20: "The scarcity of capital in Russia was such that no banking system could conceivably succeed in attracting sufficient funds to finance a large-scale industrialization; the standards of honesty in business were so disastrously low, the general distrust of the public so great, that no bank could have hoped to attract even such small capital funds as were available." Kahan (1989) assessed the state's role in less optimistic terms: some policies like industrial excise taxes contradicted the state's industrialization efforts (see Chapter 2: Government Policies and the Industrialization of Russia).

firms such as partnerships and single proprietorships. On average, corporations produced greater output per worker with more capital-intensive technologies (Gregg 2015a). Incorporation, therefore, may have been critical to firms wishing to integrate to include certain branches of cotton textile production, for example mechanized spinning, since incorporation was costly, disentangling selection into incorporation from the effects of incorporation itself is difficult.[10]

This paper will not assess the Russian cotton textiles industry's competitiveness. Cotton firms seemed to have formed cartels rather than combinations to restrain trade. Owen (1991) and Bovykin (1984) describe cartels and monopolies in the cotton textile industry and document in particular the price-fixing arrangements among the large cotton weavers in Ivanovo. Unfortunately I cannot observe cartels in the data (described below in Section 4).

3 Model and Predictions

This section outlines predictions concerning the relationship between factory or firm characteristics and vertical or horizontal integration. I define a vertically integrated factory or firm as one that possessed more than one specialization, or activity. For example, a cotton factory that had weaving as well as spinning operations was vertically integrated with two activities. In the case of firms, the definition is more broad: if a firm owned two factories, one of which spun yarn and the other of which wove cloth, the firm would be considered vertically integrated with a total number of activities of two. In the first part of the paper, I consider vertical integration within factory establishments alone. Horizontal integration

[10]Disentangling selection from the effects of incorporation is the primary concern of Gregg (2015a).

is the grouping together of factories into firms; the measure of horizontal integration is the number of factories per firm. I study horizontal integration in the second half of the paper, beginning in Section 8.

We first consider which factories and firms were more likely to be vertically integrated. Important dimensions include age, scale, and enterprise form. Older firms should have been more vertically integrated, since older firms enjoyed more opportunities to pursue projects. Second, corporations should have been more vertically integrated, since corporations had access to extra sources of capital, though corporations' advantages should be most apparent in especially capital-intensive activities. Finally, since vertical integration required larger factories, there should be a positive relationship between vertical integration and a measure of scale such as number of workers or total revenue. Similar dimensions should have been relevant for firms that combined plants, i.e. horizontally integrated. Horizontally integrated firms should have been older and more likely to be incorporated.

Furthermore, since market density varied by geography across the Empire, we should observe different patterns of vertical and horizontal integration in the different regions of the Empire. I divide the Russian Empire into two groups of regions. European Russia included most regions of the Empire, including Congress Poland and most of the Caucasus; Non-European Russia denotes the Eastern Siberia, Western Siberia, and Turkestan regions.[11] Regions outside of European Russia likely had thinner markets and weaker capital markets. Factories located in Siberia and the Far East had few alternatives should their input markets have been compromised. Such firms likely chose to produce their own inputs. If factories or

[11]See Gregg (2015b) for a more precise definition of European Russia as defined by the statistical volumes. Provinces with number codes greater than 68 are considered members of Non-European Russia.

firms were more vertically integrated outside of European Russia, I would find support for Coase's hypothesis.

I test the capital markets' importance by examining corporations, since corporations had additional access to long-term capital. If corporations were more vertically or horizontally integrated, capital market development may have been a crucial determinant of integration. I will also examine a particularly capital-intensive industry, spinning, to assess whether incorporation may have been critical for certain branches of cotton production in particular. The state of Russian capital and credit markets, then, may have been important determinants of vertical and horizontal integration.

I estimate the following regression equations to study the relationship between vertical integration and factory characteristics:

$$\begin{aligned} log(na_{ij}) = \quad & \alpha + \beta_l log(Workers_{ij}) + \beta_a log(Age_{ij}) + \beta_{ER} 1[Prov = ER]_{ij} \\ & + \beta_C 1[E.Form = Corp]_{ij} + \beta_k log(Power_{ij}) + \epsilon_{ijt} \end{aligned} \quad [1]$$

$$\begin{aligned} log(na_{ijt}) = \quad & \alpha + \gamma_l log(Workers_{ijt}) + \gamma_a log(Age_{ijt}) \\ & + \gamma_C 1[E.Form = Corp]_{ij} + \eta_t + \mu_{ijt} \end{aligned} \quad [2]$$

Equation 1 is estimated using 1894 data alone, and Equation 2 is estimated with both years of data. Here, $log(na_{ijt})$ is the natural log of the number of activities in factory i in province j in year t. The term $log(Workers_{ijt})$ is the natural log of the factory's number of workers, $1[Prov = ER]$ is a dummy variable that equal one of the factory was located in European Russia, $1[E.Form = Corp]$ is a dummy variable that equals one if the factory was owned by a corporation, and $log(Power)_{ij}$ is the factory's log total machine power

(measured in horsepower), a measure of physical capital. The term η_t controls for the year of the observation. Notice that the European Russia dummy only appears in Equation 1, because only the 1894 volume includes factories outside of European Russia, and that the year control only appears in Equation 2, since this equation includes more than one year.

I predict that $\beta_l > 0$ and $\gamma_l > 0$, since larger factories had more workers; $\beta_a > 0$ and $\gamma_a > 0$, since older factories had more opportunities to take on additional operations; and $\beta_C > 0$ and $\gamma_C > 0$, since corporations had additional capital to purchase additional operations. The sign of β_{ER} more difficult to predict: if capital markets inside European Russia were better integrated, factories in European Russia could have been more vertically integrated since developed capital markets would have allowed them to purchase additional functions, or, it could mean that factories were less vertically integrated, since the market was thicker and price risk lower (following Brown 1992).

Vertically integrated factories of firms should enjoy performance advantages if that integration allowed them to save on transaction costs. Factories that had more activities should be more productive. I measure this in two ways. First, I measure productivity as the residual of a log Cobb-Douglas production function. I also directly estimate the determinants of labor productivity (log revenue per worker):

$$
\begin{aligned}
log(Revenue/Worker_{ij}) = \quad & \alpha + \delta_l log(Workers_{ij}) + \delta_k log(Power_{ij}) \\
& + \delta_C 1[E.Form = Corp]_{ij} + \delta_{ER} 1[Prov = ER] \\
& + \delta_{na} log(na_{ij}) + e_{ij}
\end{aligned}
\qquad [3]
$$

$$
\begin{aligned}
log(Revenue/Worker_{ijt}) = \quad & \alpha + \phi_l log(Workers_{ijt}) + \phi_C 1[E.Form = Corp]_{ijt} \\
& + \phi_{ER} 1[Prov = ER]_{ij} \upsilon_{ijt} + \phi_{na} log(na_{ijt}) + \eta_t + \upsilon_{ijt}
\end{aligned}
\qquad [4]
$$

Similarly to Equations 1 and 2, Equation 3 is estimated using only 1894 data, and Equation 4 is estimated using both years of data. In these equations, $log(Revenue/Worker_{ij})$ or $log(Revenue/Worker_{ijt})$ is the log revenue per worker of factory i in province j in year t as measured in rubles; $log(Workers_{ijt})$ is a factory's log number of workers; $log(Power_{ij})$; $1[E.Form = Corp]$ is a dummy variable that equals one if the factory was owned by a corporation; $1[Prov = ER]_{ij}$is a dummy variable that equals one if the factory was located in European Russia; η_t controls for the year of the observation in Equation 4; na_{ijt} is the log number of activities that take place within the factory.

Factories with more machine power should be more productive in terms of revenue per worker, so we should find $\delta_k > 0$. If factories have significant economies of scale, we should also find $\delta_l > 0$ and $\phi_l > 0$. Gregg (2015a) finds that corporation-owned factories in the Russian Empire were more productive on average, so I should find that $\delta_C > 0$ and $\phi_C > 0$. If vertical integration increases productivity, we should find that $\delta_{na} > 0$ and $\phi_{na} > 0$.

4 Data

Data for this paper comes from the cotton industry chapters from surveys of manufacturing establishments conducted by the Russian Empire in 1894 and 1900. The 1894 factory-level volume lists for every factory in the Russian Empire: name; street address; number, type, and power of all machines; number of workers by age and gender; ruble value of mineral and plant-based fuel sources; and total value of output in rubles. Also, most significantly for this paper, the volume lists every production activity that took place within the factory.

Figure 1 shows an example from the 1894 factory list for the Nikolskaia Manufaktura Partnership of Savva Morozov Son and Company, a particularly complex factory. Savva Morozov was a serf who, because of his success in silk textiles, purchased his family's freedom from the Sheremetev estate in 1820 (Tugan-Baranovsky 1970, p. 77). This example shows the impressive amount of detail the 1894 volume gives for each individual factory. The Morozov factory produced in almost every branch of the cotton industry. The factory produced cotton wool, yarn, woven cloth, looms and shafts, and chemicals; they dyed and bleached their own cloth; and they even baked their own bread, butchered their own meat, and repaired their own tools.

Similarly, the Empire published a factory-level volume for the 1900 factory survey, though this volume lists much less detail than that for 1894. The volume lists each factory's name and street address, number of workers, total value of output, and a description of the factory's products. For this census, the authors devised a standard classification system for the subindustries within the cotton industry (See Table 4). From these, we can break down which factories had which activities to approximate the information given in the 1894 factory-level

volume. For example, if a factory is classified as belonging to the Cotton Spinning and Weaving subindustry, the factory would be listed as having the activities "spinning" and "weaving." Figure 2 shows two entries from the 1900 factory list as an example. The two entries are for two factories owned by Aleksei Vasilievich Smirnov. Unlike the entries in the 1894 factory list, the entries in 1900 are short paragraphs that describe each factory, and much less detail is provided about the factory's activities.

Furthermore, for both years, I am able to match factories to the firms that own them. The RUSCORP Database (Owen 1992), a list of all corporations founded in the Russian Empire, enables matching corporation-owned factories to the corporations that owned them using the factory's name, location, and industry. When factories were not owned by corporations, they were listed under the personal names of the partners or single proprietors that own them. In this case, I match factories to firms if the first, middle, and last names are the same for the single owner or for all partners and if the factories were located in the same province.

The dataset for this paper consists of every cotton factory from the 1894 and 1900 factory lists.[12] Table 2 shows the breakdown of the number of factories and firms in the dataset by region for both years. There are 666 factories in the 1894 cotton data and 731 factories in the 1900 cotton data.

[12]The Database of Imperial Russian Manufacturing Establishments (Gregg 2013) contains only a sample of cotton factories from 1894 and 1900, while this paper uses data from the population of cotton factories from these years.

5 Overview of Vertical and Horizontal Integration in the Cotton Industry

This section presents counts of factories and firms and descriptive statistics to show how much vertical and horizontal integration there was in the Russian cotton industry in 1894 and 1900 and how integrated factories and firms differed from non-integrated factories and firms. Integrated factories were not evenly distributed about the Empire, and integrated factories and firms had higher revenue, more workers, were older, had more machine power, and had higher revenue per worker.

Table 2 breaks down factories and firms in 1894 and 1900 by the regions of the Russian Empire. The region Turkestan does not appear in the 1900 data, because Turkestan was not part of European Russia. The regions which contained the most cotton firms and factories are the Central Industrial, Previslitskii (Poland), Turkestan, and Central Blacksoil regions. Many cotton factories also located in the Prebaltic regions, which contained St. Petersburg.

Tables 3 and 4 show the number of factories in 1894 and 1900 by activity and by the subindustry classification system used in the 1900 factory-level volume. The categories in Table 4 are broken down into their components to form the column titled "Number of Factories (1900)" shown in Table 3. As shown in both tables, the 1894 data provide much more detail on factories' activities. Still, both years of data describe some of the major categories of cotton activities: cotton wool production, thread-making, spinning, weaving, dyeing, and finishing. The data for 1894, however, include not only additional activities in the cotton industry but also activities outside of the cotton industry. 145 factories, for example, included a repair shop: this would normally be classified in the metals and machines

industry. Many factories also produced their own gas, peat, or electric fuel, and 9 factories made bricks.

Different activities were concentrated in different regions of the Russian Empire. Table 5 shows how the activities weaving, spinning, and cotton cleaning were distributed about the Empire in 1894 and 1900. Weaving and Spinning were concentrated in the industrial regions of the Empire: Central Industrial, Previslitskii, Central Blacksoil, and Prebaltic. Cotton cleaning, an early stage in raw cotton processing, represented a prominent industry in Turkestan, a region that largely overlaps with modern-day Uzbekistan.[13]

Table 6 shows how the characteristics of factories varied by the number of activities contained within the factory. I have grouped factories into three categories: factories with one activity (non-integrated factories), factories with between 2 and 4 activities (moderately integrated factories), and factories with 5 or greater activities (highly integrated factories). As the number of activities in factories increased, revenue, number of workers, machine power, age, and revenue per worker increased (though revenue per worker decreased slightly in 1894 between moderately integrated and very integrated factories). These results are consistent with predictions: integrated factories were larger, older, and more productive.

The numbers of revenue, number of workers, machine power, age, and revenue per worker are all larger in 1900 in Table 6 for several reasons. First, factories grew over time, and many of the factories survived more than one period. Second, and probably more significantly, the 1900 volume reports much less detail about the activities taking place without factories than the 1894 volume. Thus, a factory with 5 activities as listed in 1900 was probably a larger factory than a factory with 5 activities in 1894.

[13]Tugan-Baranovsky, "Russian Factory," 292.

6 Determinants of Vertical Integration, Horizontal Integration, and Revenue per Worker

Taking the number of activities in a factory or firm as a measure of vertical integration, how does vertical integration vary with factory or firm size, geography, enterprise form, and factory age? The regressions shown in Table 7 address the major dimensions determining vertical integration. Table 9 shows that more vertically integrated firms were more productive, and Figure 3 shows that the distributions of total factor productivity become tighter about the mean as the number of activities within the factory increases.

6.1 Which factories were vertically integrated?

Factories were less integrated within European Russia, where capital markets were likely more developed and markets thicker. These results suggest that thin markets encouraged vertical integration. The first regression results are presented in Table 7. Columns 1 through 5 of Table 7 show how the log number of activities in cotton factories varied with log number of workers, log factory age, whether the factory was located in European Russia, whether the factory was owned by a corporation, and the machine power in the factory. In Column 1, the coefficient on log number of workers is positive and statistically significant at the .001 level: factories with larger labor forces included more activities within the factory. The coefficient on log factory age is small relative to its standard error in Column 2, but it becomes much larger once I control for whether the factory was located in European Russia. In Column 3, the coefficient on log factory age is large and positive, and the coefficient on the European Russia dummy is large and negative.

Column 4 adds a dummy variable for whether the factory was owned by a corporation, and the coefficient is positive but small relative to its standard error. When I add a control for the machine power in the factory in Column 6, the coefficient on the corporation dummy becomes even smaller; Gregg (2015a) argues that incorporation allows factories to purchase large machines, controlling for machine power absorbs much of the difference between corporations and non-corporations.

The results change slightly for non-cotton activities, however. As with the log total number of activities, the number of log non-cotton activities increases with number of workers and is smaller if a factory was located in European Russia. The coefficient on log factory age, however, is now negative and small. Factories with more machine power had more non-cotton activities (Column 6). Also, the coefficient on the corporation-owned dummy variable is positive, a bit larger, but still statistically insignificant in Column 7. Factories that had more powerful machines were more likely to participate in industries outside of cotton.

Table 7 Columns 8 and 9 show similar regressions including factories from both the 1894 and 1900 volumes. The relationships shown in the previous regressions are similar here: older factories with more workers performed more functions. In this table, I introduce a control for the year of observation. The coefficient for a 1900 factory is negative and large, as could be predicted: the 1900 volume provides a much less detailed description of factory activities.

6.2 A Close Examination of Capital-Intensive Production: Spinning

Mechanized spinning required large machine purchases, which only certain Russian firms may have been able to finance. Very few Russian cotton firms produced yarn as a final product. In this section, therefore, I examine those Russian firms that vertically integrated to produce their own yarn. I show that, indeed, spinning was capital intensive and that corporations dominated spinning. For certain industries, access to capital represented a first-order consideration for factories deciding whether to vertically integrate.

Table 8 outlines results that compare factories that include spinning with those that do not. A number of striking differences emerge. In 1894, 92 factories out of 666 cotton factories (13.81 percent) spun yarn.[14] Spinning factories were on average much larger than factories that did not spin yarn: spinning factories on average had over 1,800 workers (Panel A). The 99th percentile for number of workers for factories across all industries in the Russian Empire in 1894 was 808; in the cotton industry, the 90th percentile for number of workers in 1894 was 1,342. Thus, cotton spinning factories possessed some of the largest workforces of any kind of factory in the Russian Empire in 1894. These cotton spinning factories also tended to be more vertically integrated: they had 3.26 activities on average compared to 1.78 for factories that did not spin yarn.

Most crucially, Panel A demonstrates that cotton spinning was remarkably capital intensive. The average total machine horsepower of cotton spinning factories was 1508.44, many times greater than the average for factories that did not spin yarn. Perhaps unsurprisingly,

[14]Imperial Russian Manufacturing Database (2015)

factories owned by corporations dominated cotton spinning: over sixty percent of cotton spinning factories compared to twelve percent of factories without spinning were owned by incorporated firms.

Panel B documents the characteristics of cotton spinning factories, fixing size, enterprise form, and location by estimating probit regressions. Even controlling for size (log number of workers), spinning factories were more likely to be owned by corporations. In the case of spinning, location did not strongly determine integration: factories located outside of European Russia were not more likely to include spinning. Once a factory's total machine power is included in the probit regression, the coefficient on enterprise form is no longer significant (Column 5). Incorporation and capital are strongly correlated, and the capital itself, not the enterprise form per se, determined whether a factory could spin yarn. These results suggest that in this case of capital-intensive industries, financial market obstacles were more important than input price variation.

6.3 Matching Factories to Firms and Determinants of Firm Size

I first match factories in the 1894 and 1900 data to the firms that owned them. If a factory was owned by a partnership or single proprietorship, I match factories whose owners had the same first, middle (patronymic), and last names. If a factory was owned by a corporation, I match all factories that had the same corporation name. Table 10 Panel A shows how many factories can be matched to firms. Here, "firm size" means the number of factories that belonged to a firm. Thus, in 1894 there were 580 firms that owned one factory. Also in 1894, there were 64 factories that belonged to two-factory firms, and hence there were

32 two-factory firms, and so on. In both 1894 and 1900, there was only one firm that had seven members, the Karl Sheibler Corporation in Petrokovskaia Province (part of modern-day Poland, named after Piotrków). There were slightly more many-member firms in 1900 than there were in 1894.

Table 10 Panel B shows the number of corporations and non-corporations by firm size. The largest firms tended to be corporations, especially in 1900, though there were not very many large firms in either year.

What were the characteristics of factories that belonged to large, horizontally integrated firms? Table 11 presents Tobit regressions in which the dependent variable is the size of the firm. This tobit is left-censored at zero, because each firm must own at least one factory (and the log of one is zero). The unit of analysis in this regression is the factory: thus, the dependent variable is the size of the firm to which each factory belongs. The most important determinants of whether a factory belonged to a large firm are the number of workers and whether the factory was located in European Russia. Factories that had more workers and that were located outside of European Russia were more likely to belong to multi-member firms.

6.4 Were vertically integrated factories more productive?

This section documents productivity differences between more and less vertically integrated factories. Vertically integrated factories could appear to be more productive either because vertical integration improved these factories' efficiency or because more productive (and, perhaps, better-managed) factories tended to vertically integrate. In this paper, I will not

be able to evaluate the direction of causation. However, finding that vertically integrated factories were in fact significantly less productive would raise doubts about how well Imperial Russian markets functioned. I find that vertically integrated factories were indeed more productive, even controlling for factory size and enterprise form.

Table 9 presents estimates of production functions for the 1894 data separately and the 1894 and 1900 combined. The dependent variable in these regressions is the log of the total value of output in rubles divided by the factory's number of workers. Column 1 shows a correlation between a factory's productivity as measured as revenue per worker and the factory's level of vertical integration (number of activities). Column 2 controls for factory's inputs: even controlling for a factory's size in terms of number of workers and machine power, factories with more activities produced more revenue per worker. Columns 3 and 4 introduce controls for the factory's age and whether the factory was owned by a corporation. Factories owned corporations produced more revenue per worker given the same number of workers, amount of machine power, and number of activities. This is consistent with my previous work: Gregg (2015a) showed that corporations were more productive.

Columns 5 through 8 present similar regressions over both years of data, though a control is introduced for the year. Factories in 1900 had substantially less revenue per worker than factories in the 1894 data: perhaps the 1900 factory-level volume captured more small factories than did the 1894 volume. Also, the coefficient on log factory age is now much smaller but remains statistically insignificant. The coefficients for log number of activities and whether the factory is owned by a corporation are both positive and statistically significant at the .001 level.

Productivity can also be measured as the residual of a log Cobb-Douglas production function. Kernel density estimates of residuals by number of activities are presented in Figure 3. The first figure shows a density plot of residuals from a regression of log Revenue on log Workers using data from both 1894 and 1900. The second figure plots the density of residuals from a regression of log Revenue on log Workers using data from 1894 only (since machine power is only measured in 1894). As the number of activities increases, the density of residuals becomes tighter about the median. More specifically, while there were highly unproductive factories with fewer activities, factories that had more activities do not appear on the lower tail of the distribution. However, the difference is smaller when I control for machine power: much of the productivity difference between vertically integrated and non-integrated factories can be explained by differences in physical capital.

7 Additional Functional Forms and Alternative Definitions of Integration

Table 12 presents several additional functional forms and specifications. I present results that use OLS instead of Tobit, robust or clustered standard errors, and nonlinear terms. Regressions using OLS decrease the size of the estimates, all standard error forms give similar results, and regressions using nonlinear terms perform somewhat worse than regressions without nonlinear terms. I also present results that use alternative definitions of integration.

Table 12 shows results using 1894 data alone and 1894 and 1900 combined, estimating the main regression equations using Tobit, OLS, and robust and clustered standard errors. In

general, the results are the same as observed in Table 7: older factories with more workers and more machine power outside of European Russia had more activities. Regressions estimated with OLS have coefficients that are smaller in absolute value, and clustering standard errors reduces statistical significance somewhat. Overall, however, the results are unchanged.

So far in the paper, I have only examined regressions using linear terms. Table 12Columns 2 and 6 displays the results of regressions that include some squared terms. Column 2 introduces log workers squared and log factory age squared: the coefficient on log workers squared is now statistically significant, though smaller than the log workers coefficient in Column 1, and log workers has lost statistical significance. The regression in Column 6 is similarly difficult to interpret: now the only statistically significant coefficient is the year dummy.

Finally, I examine an alternative definition of integration. In the paper so far, I have used the number of activities within the factory or firm as the dependent variable in a Tobit regression. There are at least two other ways to define vertical integration using the variables available in the 1894 and 1900 volumes. The first way is to define a vertically integrated factory or firm as one that had more than one activity. Table 10 Panel C shows a breakdown of factories and firms in both years using these two definitions of integration. About half of factories and firms in 1894 and a slightly smaller proportion of firms and factories in 1900 had more than one activity.

In Table 13, I present estimates from a probit regression in which the dependent variable is whether the factory or firm had more than one activity. The probit regressions presented in 13 show similar patterns as observed previously in the paper: factories or firms that were older, had more workers, and were located outside of European Russia were more likely to be

integrated. In these regressions, corporate ownership does not strongly predict integration, nor does having more machine power. Additional capital, then, may be helpful for acquiring more activities overall but not for becoming integrated in the first place.

8 Conclusion and the Path Forward

This paper has documented the characteristics of horizontally and vertically integrated factories and firms in the Imperial Russia cotton industry, one of the Empire's most technologically advanced and productive industries. I find that older factories with more workers located that were outside of European Russia tended to be the most vertically integrated. Factories that had more workers and that were located outside of European Russia were more likely to belong to horizontally integrated firms. And there is evidence that more vertically integrated factories and firms were also more productive.

The results of this paper emphasize some of the main characteristics of the Russian industrial sector in general. The various branches of the Russian cotton industry were distributed about the geography of the Empire, where market thickness and access to credit and capital varied greatly. There were some very large, highly vertically integrated firms and many single-factory, single-proprietor-owned firms as well. Further research is needed to establish how prices varied throughout the Russian Empire, to pinpoint the sources of firm and factory growth, to explore how the Russian governments' disparate industrial policies contributed to the size distribution of firms in the industrial sector, and to evaluate the competitiveness of Russian industry at the turn of the twentieth century.

References

[1] Acemoglu, Daron, Simon Johnson, and Todd Mitton. "Determinants of Vertical Integration: Financial Development and Contracting Costs." *The Journal of Finance 64.3 (2009)*: 1251-1290.

[2] Bovykin, V. I. *Formirovanie finansovovo kapitala v Rossii: Konets XIX v. - 1908 g.* Moskva: Nauka, 1894.

[3] Brown, John C. "Market Organization, Protection, and Vertical Integration: German Cotton Textiles before 1914." *The Journal of Economic History* 52.2 (1992): 339-351.

[4] Coase, Ronald H. "The Nature of the Firm." *Economica.* 4.16 (1937): 386-405.

[5] Crisp, Olga. *Studies in the Russian Economy before 1914.* London: Macmillan, 1976.

[6] Dennison, Tracy. *The Institutional Framework of Russian Serfdom.* Cambridge: Cambridge University Press, 2011.

[7] Gerschenkron, Alexander. *Economic Backwardness in Historical Perspective: A Book of Essays.* Cambridge: Belknap Press, 1962.

[8] Goldsmith, Raymond W. "The Economic Growth of Tsarist Russia 1860-1913. *Economic Growth and Social Change* 9 No. 3 (April 1961): 441-475.

[9] Gregg, Amanda. "Factory Productivity and the Concession System of Incorporation in Late Imperial Russia." Unpublished Manuscript (2015a).

[10] Gregg, Amanda. "Code Book and Data Appendix: Imperial Russian Manufacturing Establishments Database: 1894, 1900, and 1908". Unpublished Manuscript (2015b).

[11] Gregory, Paul R.. "Some Empirical Comments on the Theory of Relative Backwardness: the Russian Case." *Economic Development and Cultural Change* 22 No. 4, 1974: 654-665

[12] Timothy Guinnane, Ron Harris, Naomi R. Lamoreaux, and Jean-Laurent Rosenthal. "Putting the Corporation in its Place." *Enterprise and Society* 8 No. 3 (2007): 687-729.

[13] Kahan, Arcadius. *Russian Economic History: The Nineteenth Century.* Chicago: University of Chicago Press, 1989.

[14] Macchiavello, Rocco. "Financial Development and Vertical Integration: Theory and Evidence." Journal of the European Economic Association 10.2 (2012): 255-289.

[15] Nafziger, Steven. "Peasant communes and factor markets in late nineteenth-century Russia." *Explorations in Economic History* 47.4 (2010), 381-402.

[16] Owen, Thomas C. *The Corporation under Russian Law.* Cambridge: Cambridge University Press, 1991.

[17] Owen, Thomas C. Codebook for RUSCORP: A Database of Corporations in the Russian Empire, 1700-1914. Third Release. Baton Rouge, LA, 1992 [Producer]. Ann Arbor: Inter-university Consortium for Political and Social Research, 1992 [Distributor].

[18] Owen, Thomas C. RUSCORP: A Database of Corporations in the Russian Empire, 1700-1914. Third Release. Baton Rouge, LA, 1992 [Producer]. Ann Arbor: Inter-university Consortium for Political and Social Research, 1992 [Distributor].

[19] Polnoe sobranie zakonov rossiskoi imperii (aka PSZ, Volumes 1 and 2). [Complete collection of laws of the Russian Empire]. Available: http://www.nlr.ru/e-res/law_r/search.php

[20] Shepelev, Leonid V. *Aktsionernye kompanii v Rossii.* [Corporations in Russia] Leningrad: Nauka, 1973

[21] Svod zakonov rossiiskoi imperii. [Code of Laws of the Russian Empire]. St. Petersburg: Official Publication of the Russian Empire, 1914.

[22] Tugan-Baranovsky, Mikhail Ivanovich. *The Russian Factory in the Nineteenth Century.* Translated by Arthur Levin and Claora S. Levin under the supervision of Gregory Grossman. (Homewood, Illinois: Richard D. Irwin, Inc. for the American Economic Association, 1970).

[23] Williamson, Oliver. The Economic Institutions of Capitalism. New York: Free Press, 1985.

Original Data Sources

[1] Ministry of Finance Department of Trade and Manufacturing of the Russian Empire. *Russian Factory Production: List of Factories and Plants* [Фабрично-заводская промышленность России: Перечень фабрик и заводов]. St. Petersburg: Tipografia E.A Efrona, 1897.

[2] Ministry of Finance of the Russian Empire. *List of Factories and Plants of European Russia* [Список фабрик и заводов Европейской России]. St. Petersburg: Tipografia Kirshbaum, 1903.

[3] Ministry of Trade and Industry of the Russian Empire, Industrial Division. *List of Factories and Plants of the Russian Empire.* [Список фабрик и заводов Российской Империи]. V.E Varzar, ed. St. Petersburg: Tipografia Kirshbaum, 1912.

[4] Ministry of Finance of the Russian Empire. Statistical Results on Factories and Plants by Industries Not Subject to the Excise Tax for 1900. [Статистические сведения о фабриках и заводах по производствам необложенным акцизом за 1900 год.] V.E. Varzar, ed. St. Petersburg: Tipografia Kirshbaum, 1903.

[5] Ministry of Trade and Industry of the Russian Empire, Industrial Division. Statistical Results on Factory Production in the Russian Empire for 1908. [Статистические сведения по обрабатывающей фабрично-заводской промышленности Российской Империи за 1908 год]. V.E. Varzar, ed. St. Petersburg: Tipografia Kirshbaum, 1912.

Archive

Russian State Historical Archive [Российский государственный исторический архив] (Abbreviated RGIA). St Petersburg, Russia.

Tables

Table 1: Revenue, Labor, and Degree of Vertical Integration by Industry (1894 and 1900)

	Total Factories	%	Total Revenue	%	Total Rev./ Workers	Total Power/ Workers (1894)	Avg. Num. of Activities (1894)
Cotton	**1,397**	**4.21**	**931,654,477**	**21.85**	**1,377.70**	**.27**	**2.31 (Std. Dev. 2.05)**
Animal Products	3,042	9.17	203,952,736	4.78	2,233.09	.11	1.48 (0.81)
Chemicals	1,263	3.81	253,318,493	5.94	2,729.14	.22	1.46 (0.92)
Flax, Hemp, and Jute	780	2.35	115,024,057	2.70	896.57	.19	1.90 (1.36)
Foods	12,853	38.74	1,404,825,064	32.94	2,124.98	.29	1.50 (0.88)
Metals and Machines	3,318	10.00	542,674,191	12.73	1,477.45	.10	1.36 (0.67)
Mineral Products	2,864	8.63	120,441,589	2.82	580.27	.07	1.33 (0.73)
Mixed Materials	762	2.30	54,442,893	1.28	1,064.81	.06	1.44 (0.94)
Paper	1,923	5.80	130,893,215	3.07	1,104.40	.31	1.90 (1.08)
Silk	577	1.74	45,336,931	1.06	935.19	.05	1.35 (0.76)
Wood	2,434	7.34	145,715,856	3.42	1,301.03	.30	1.39 (0.77)
Wool	1,962	5.91	315,875,978	7.41	1,269.69	.22	1.93 (1.59)
Total	33,175	100	4,264,155,462	100			

Source: Imperial Russian Manufacturing Database (2015). The "Foods" category includes both factories subject to the excise tax and factories not subject to the excise tax. Percentages are in the Revenue category are rounded and may not add up to 100. The final column reports standard deviations in parentheses.

Table 2: Number of 1894 Cotton Factories by Region

	1894		1900	
	Number of Factories	Number of Firms	Number of Factories	Number of Firms
Caucasus	1	1	2	2
Central Blacksoil	40	39	89	88
Central Industrial	427	404	431	401
Eastern	5	5	4	4
Northwestern	2	2	2	2
Prebaltic	39	35	37	31
Previslitskii	75	68	154	143
Southern	17	17	12	12
Turkestan	60	58	(N/A)	(N/A)
Total	666	618	731	683

Source: Imperial Russian Manufacturing Database (2015). Factories are matched to firms by first and last names and by province, unless the factory is owned by a corporation, in which case they are matched by the corporation name.

Table 3: Number of Cotton Factories by Activity (1894 and 1900)

Activity (Cotton)	Number of Factories (1894)	Number of Factories (1900)	Activity (Other Ind.)	Number of Factories (1894)
Weaving	366	426	Repair Shop	**145**
Dyeing	219	120	Gas	**70**
Finishing	145	282	Bricks	**9**
Spinning	92	117	Peat	**6**
Printed Fabric	87		Looms and Shafts	**5**
Cotton Wool	78	49	Bread Baking	4
Bleaching	77		Chemicals	4
Cotton Cleaning	59		Wood	**4**
Cotton Baling	40		Electricity	3
Engraving	28	2	Butchery	3
Thread	13	21	Looms	3
Mech. Spinning	9		Flour	3
Spinning (Fallen)	8	21	Butter	3
Wicks	4		Alizarine	2
Hydrocyanic Dyeing	4		Book Binding	1
Velvet	2	6	Mechanical	1
Twine	2		Soap	1
Bands	2		Artificial Wool	1
Printed Skirts	1		Boxes	1
Machine Belts	1		Brushes	1
Heald	1		Foundry	1
Cordage	1		Rubber Weaving	1
Belts	1		Machine Building	1
Rugs	1		Lime	1
Knitted Fabric	1		Flax Spinning	1
Waste Yarn	1		Hemp Scutching	1
Lacework	1		Wool Cleaning	1
Spun Yarn	1		Worsted	1
Brocade	1		Fire Hoses	1
Talc Padding	1		Mineral Paints	1
Chintz	1		Carpentry	1
Cordage (alt)	1			
Worsted Vicuna	1			

Source: Imperial Russian Manufacturing Database (2015). Activities are indicated in the entry for each factory in 1894 (See Figure 1). For 1900, the categories are compiled by decomposing the categories listed in Table 4.

Table 4: Number of Factories in Each 1900 Cotton Subindustry

Subindustry Classification	Number of Factories
Cotton Weaving	152
Cotton Dyeing and Finishing	124
Cotton Weaving Distribution Offices	116
Cotton Weaving with Dyeing and Finishing	67
Cotton with Weaving, Dyeing, and Distribution	53
Cotton Quilting	49
Cotton Spinning	42
Cotton Weaving, Dyeing, and Finishing	38
Cotton Spinning and Weaving	37
Vicuna	24
Cotton Thread	21
Velvet	6
Cotton Engraving	2
Total	731

Source: Imperial Russian Manufacturing Database (2015)

Table 5: Number of Cotton Factories by Activities and Region

	Weaving		Spinning		Cleaning	
	1894	1900	1894	1900	1894	1900
Caucasus	1	0	1	1	0	n/a
Central Blacksoil	17	62	1	2	0	n/a
Central Industrial	298	285	42	58	0	n/a
Eastern	0	0	1	0	0	n/a
Northwestern	1	0	0	0	0	n/a
Prebaltic	15	8	21	21	0	n/a
Previslitskii	34	71	24	33	0	n/a
Southern	0	12	1	2	0	n/a
Turkestan	1	n/a	1	n/a	59	n/a
Total	367	426	92	117	59	

Source: Imperial Russian Manufacturing Database (2015).
The categories for 1900 come from decomposing the categories listed in the 1900 volume (See Table 4). The last column says "n/a" because the 1900 volume gives no information on cotton cleaning.

Table 6: Descriptive Statistics for Cotton Factories by Number of Activities, 1894 and 1900

Number of Activities	**Revenue** 1894	1900	**Workers** 1894	1900	**Machine Power** 1894	1900
1	89,370.84	328,866.3	70.28	283.74	41.31	n/a
	(297,779.8)	(936,864.4)	(180.54)	(653.40)	(197.92)	
$1 < x < 5$	561,634.8	1,228,719	393.08	882.25	254.95	n/a
	(1,154,009)	(2,362,472)	(716.69)	(1678.57)	(676.78)	
$x > 4$	2,970,454		1,981.90		1372.82	n/a
	(2,896,532)		(2120.30)		(1670.88)	

Number of Activities	**Age** 1894	1900	**Revenue per Worker** 1894	1900
1	19.33	18.99	1277.22	876.31
	(18.43)	(18.56)	(1926.88)	(1138.05)
$1 < x < 5$	26.70	29.95	2157.94	1231.96
	(23.08)	(24.83)	(3238.71)	(1873.44)
$x > 4$	42.55		2055.27	
	(28.73)		(1753.23)	

Source: Imperial Russian Manufacturing Database (2015)

Table 7: Determinants of Number of Cotton Factory Activities (1894 and 1900)

Years:	1894 Only							1894 and 1900	
Tobit	Dep. Var: Log Number of Activities					Log Number of Non-Cotton Activities		Log Number of Activities	
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
Log Num. of Workers	.35***	.34***	.36***	.42***	.41***	.48***	.22***	.26***	.26***
	(.020)	(.022)	(.022)	(.051)	(.051)	(.036)	(.059)	(.018)	(.023)
Log Factory Age		.064	.16***	.098*	.099*	-.014	-.0064	.18***	.18***
		(.036)	(.037)	(.039)	(.039)	(.046)	(.043)	(.030)	(.030)
European Russia			-.97***	-.74***	-.72***	-.69***	-.63**		
			(.14)	(.16)	(.17)	(.19)	(.19)		
Log Machine Power				-.045	-.050		.18***		
				(.16)	(.043)		(.051)		
Owned by Corporation					.082		.14		.054
					(.11)		(.11)		(.090)
Year = 1900								-.36***	-.36***
								(.062)	(.062)
Intercept	-1.25***	-1.36***	-.87***	-.93***	-.91***	-2.06***	-1.45***	-1.48***	-1.46***
	(.10)	(.13)	(.13)	(.16)	(.16)	(.21)	(.20)	(.12)	(.13)
Chi2	288.37	259.16	312.20	223.02	223.76	292.99	220.53	335.47	335.26
R2	.1942	.1909	.2300	.2323	.2335	.3229	.3008	.1274	.1275
N	626	571	571	380	379	571	379	1201	1200

*** $p < .001$, ** $p < .01$, * $p < .05$. Source: Imperial Russian Manufacturing Database (2015). Tobit is left-censored at zero, since a firm has one or more activities. Standard errors in parentheses. Regressions in columns 8 and 9 include only European Russia.

Table 8: Capital-Intensive Production: Spinning

Panel A: Factories that Spun Yarn Were Larger, Had More Powerful Machines, and Tended to be Corporations

	Factories That Spin Yarn	Factories That Do Not Spin	|t-stat|
Total Machine Power	1508.44	81.63	19.07
	(175.69)	(10.52)	
Total Workers	1821.82	192.16	16.60
	(211.89)	(20.02)	
Proportion Incorporated	.64	.12	19.35
	(.033)	(.0096)	
Total Num. of Activities	3.26	1.78	13.02
	(.18)	(.035)	

Panel B: Probit Regressions for Factories that Included Spinning

Probit	Dependent Variable: Factory Spun Yarn				
	[1]	[2]	[3]	[4]	[5]
Owned by	1.45***	.50**	.43**	.44	.098
Corporation	(.15)	(.18)	(.18)	(.24)	(.28)
	[.25]	[.074]	[.072]	[.023]	[.010]
Log (Number		.46***	.47***	.77***	-.21
of Workers)		(.054)	(.055)	(.099)	(.17)
		[.054]	[.054]	[.030]	[-.021]
Factory was Located			-.064	-1.15	
in European Russia			(.45)	(.81)	
			[-.0077]	[-.14]	
Log (Total					1.12***
Machine Power)					(.20)
					[.11]
Intercept	-1.57***	-3.66***	-3.62***	-3.90***	-5.21***
	(.087)	(.30)	(.47)	(.66)	(.81)
Chi2	99.93	135.66	143.38	123.82	81.93
R2	.2170	.3749	.3750	.5393	.5795
N	665	625	625	544	341
Province Controls	No	No	No	Yes	Yes

*** $p < .001$, ** $p < .01$, * $p < .05$. Source: Imperial Russian Manufacturing Database (2015). Robust standard errors of the coefficients in parentheses. Marginal effects (dF/dx) in brackets.

Table 9: Determinants of Cotton Factory Productivity (Revenue per Worker), 1894 and 1900

OLS	Dependent Variable: Log Revenue per Worker							
	1894 Only				1894 and 1900			
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Log (Number of Activities)	.30*** (.063)	.22** (.076)	.18* (.075)	.16* (.079)	.33*** (.051)	.36*** (.062)	.33*** (.059)	.23*** (.066)
Log (Machine Power)		.20*** (.044)	.16*** (.046)	.16** (.048)				
Log (Number of Workers)		-.31*** (.054)	-.33*** (.054)	-.34*** (.055)		-.026 (.022)	-.13*** (.026)	-.087** (.028)
Form = Corporation			.42*** (.11)	.41** (.12)			.75*** (.090)	.68*** (.094)
Log (Factory Age)				.078 (.43)				-.0028 (.032)
YEAR = 1900								-.34*** (.078)
Intercept	6.94*** (.076)	7.85*** (.15)	7.98*** (.16)	7.85*** (.18)	6.67*** (.049)	6.78*** (.11)	7.10*** (.12)	7.16*** (.13)
F-Stat	22.67	11.26	11.30	9.75	40.67	20.51	34.04	23.94
R2	.0347	.1098	.1448	.1507	.0327	.0342	.0856	.0972
N	504	344	343	320	1012	1012	1011	939

*** $p < .001$, ** $p < .01$, * $p < .05$. Source: Imperial Russian Manufacturing Database (2015).
Regressions with 1894 and 1900 data include European Russia only. Robust standard errors in parentheses.

Table 10: Number of Cotton Factories, Firms, and Corporations

Panel A: Number of Factories and Firms by Firm Size

	1894		1900	
Firm Size	Number of Factories	Number of Firms	Number of Factories	Number of Firms
1	580	580	653	580
2	64	32	42	21
3	15	5	15	5
4	0	0	8	2
6	0	0	6	1
7	7	1	7	1

Panel B: Number of Firms by Enterprise Form (Corporations vs. Non-Corporations)

	1894		1900	
Firm Size	Number of Non-Corporations	Number of Corporations	Number of Non-Corporations	Number of Corporations
1	498	82	554	99
2	18	13	11	10
3	0	5	2	3
4	0	0	0	2
6	0	0	0	1
7	0	1	0	1

Panel C: Number of Integrated and Non-Integrated Cotton Factories and Firms

	1894		1900	
	Factories	Firms	Factories	Firms
Number of Activities = 1	333	315	412	524
Number of Activities > 1	333	304	319	159
Total	666	619	731	683

Source: Imperial Russian Manufacturing Database (2015). The author matched factories to firms, as described in the text: factories that do belong to corporations belong to the same firm if the first, middle, and last names match and if they are located in the same province in the same industry. Factories that belong to corporations are matched by corporation name. Factories are matched to corporations using the RUSCORP Database (Owen 1992).

Table 11: Factory Characteristics Determining Cotton Firm Size, 1894 and 1900

Tobit	Dependent Variable: Log Firm Size						
	1894 Only				1894 and 1900		
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Log (Total Workers)	.44*** (.084)	.42*** (.085)	.45*** (.089)	.32* (.16)	.62*** (.086)	.61*** (.089)	.68*** (.11)
Log (Factory Age)		.090 (.12)	.20 (.13)	.20 (.14)		.030 (.10)	.00071 (.10)
European Russia			-1.46** (.43)	-1.59*** (.45)			
YEAR = 1900							-.37 (.19)
Form = Corporation				-.11 (.28)			-.25 (.24)
Log (Machine Power)				.12 (.14)			
Intercept	-4.25*** (.72)	-4.31*** (.77)	-3.46*** (.66)	-3.01*** (.65)	-5.77*** (.74)	-5.74*** (.76)	-5.73*** (.79)
F-Stat	61.71	59.28	72.91	52.25	145.32	144.90	149.78
R2	.1900	.1863	.2291	.1872	.2526	.2559	.2646
N	579	532	532	348	1208	1125	1124

*** $p < .001$, ** $p < .01$, * $p < .05$. Source: Imperial Russian Manufacturing Database (2015). The unit of observation is the factory. Thus the dependent variable is the size of the firm to which the factory belongs. Tobit is left-censored at zero, since a firm has one or more factories. Standard errors in parentheses. Regressions that include 1900 exclude factories outside European Russia.

Table 12: Determinants of Number of Cotton Factory Activities (1894 and 1900): Alternative Specifications

	Dep. Var: Log Number of Activities							
	1894				1894 and 1900			
	Tobit [1]	Tobit [2]	OLS [3]	OLS [4]	Tobit [5]	Tobit [6]	OLS [7]	OLS [8]
Log Num. of Workers	.41***	.091	.23***	.23***	.24***	.24**	.14***	.14***
	(.085)	(.095)	(.030)	(.057)	(.031)	(.076)	(.011)	(.017)
Log Num. of Workers^2		.028**				.00094		
		(.0091)				(.0073)		
Log Factory Age	.099*	-.0080	.068*	.068*	.14**	.058	.062***	.062**
	(.043)	(.13)	(.027)	(.028)	(.048)	(.10)	(.013)	(.017)
Log Factory Age^2		.030				.016		
		(.025)				(.019)		
Year = 1900					-.44**	-.44***	-.27***	-.27***
					(.14)	(.058)	(.030)	(.063)
European Russia	-.72*	-.92***	-.33***	-.33*				
	(.30)	(.11)	(.083)	(.16)				
Corporation	.082		.11	.11	.079		.072	.072
	(.11)		(.084)	(.10)	(.15)		(.047)	(.093)
Log (Machine Power)	-.050		.0053	.0053				
	(.073)		(.029)	(.047)				
Intercept	-.91**	-.16	-.30***	-.30	-1.15***	-1.08***	-.16***	-.16*
	(.30)	(.26)	(.084)	(.16)	(.024)	(.21)	(.044)	(.072)
Std. Errors	C	R	R	C	C	R	R	C
Chi2 / F	26.27	104.53	78.91	43.12	26.70	82.98	106.08	50.88
R2	.2335	.2379	.4600	.4600	.1157	.1157	.2714	.2714
N	379	571	379	379	1256	1257	1256	1256

*** $p < .001$, ** $p < .01$, * $p < .05$. Source: Imperial Russian Manufacturing Database (2015). Tobit is left-censored at zero, since a firm has one or more activities. Standard errors in parentheses: "C" means clustered by Province, "R" means robust. Regressions for that include both 1894 and 1900 use observations only from European Russia. Chi2 statistics are reported for Tobit regressions; F-statistics are reported for OLS.

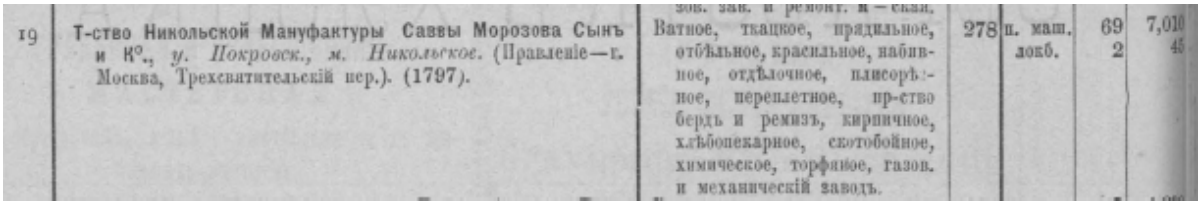
Table 13: Determinants of Cotton Factory Integration: 1894 and 1900

Probit	Definition of Integration: $na_{ijt} > 1$			
	Dep. Variable: Factory is Integrated		Dep. Variable: Firm is Integrated	
	1894	1894 and 1900	1894	1894 and 1900
	[1]	[2]	[3]	[4]
Log Num. of Workers	.66***	.25***	.66***	.34***
	(.11)	(.028)	(.12)	(.032)
Log Age	.12	.15***	.15*	.14**
	(.069)	(.034)	(.074)	(.040)
Year = 1900		-.43***		-1.14***
		(.076)		(.087)
European Russia	-1.16***		-1.26***	
	(.28)		(.30)	
Corporation	.12	.13	.16	-.067
	(.23)	(.12)	(.28)	(.13)
Log (Machine Power)	-.20*		-.15	
	(.089)		(.091)	
Intercept	-1.30***	-1.38***	-1.42***	-1.68***
	(.29)	(.13)	(.32)	(.15)
Wald Chi2	96.42	216.93	99.67	325.60
Pseudo R2	.2799	.1235	.3329	.2334
N	379	1256	348	1172

*** $p < .001$, ** $p < .01$, * $p < .05$. Source: Imperial Russian Manufacturing Database (2015). Tobit is left-censored at zero, since a firm has one or more activities. Robust standard errors in parentheses. Regressions for that include both 1894 and 1900 use observations only from European Russia.

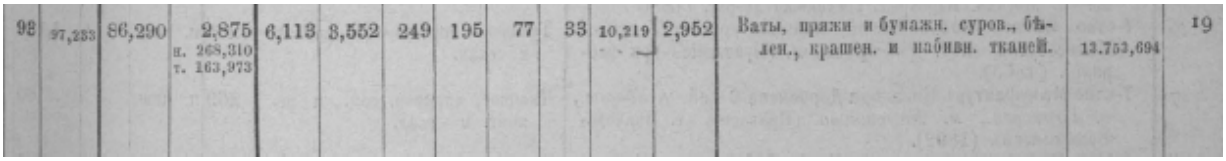
Figures

Figure 1: 1894 Factory List Example



19	Т-ство Никольской Мануфактуры Саввы Морозова Сынъ и К°., *у. Покровск., м. Никольское.* (Правленіе—г. Москва, Трехсвятительскій пер.). (1797).										Ватное, ткацкое, прядильное, отбѣльное, красильное, набивное, отдѣлочное, плисорѣзное, переплетное, пр-ство бердъ и ремизъ, кирпичное, хлѣбопекарное, скотобойное, химическое, торфяное, газов. и механическій заводъ.	278	п. маш. локб.	69 2	7,010 45

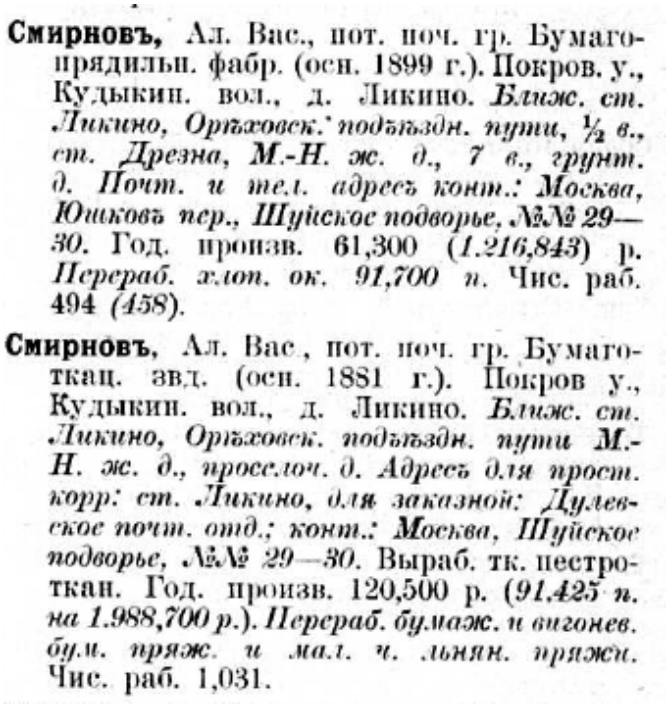
Source: Perechen' fabrik i zavodov (1896), Page 2. Translation: Nikolskaia Manufaktura Partnership of Savva Morozov Son and Co. Located in Pokrovsk. uezd, (m) Nikol'skoe (Board of Directors: Moscow, Trekhsiatitel'skii Lane. Founded: 1797. Activities: cotton wool, weaving, spinning, bleaching, dyeing, printed fabric, finishing, velvet, book binding, looms and shafts, bricks, bread baking, butchery, chemicals, peat, gas, mechanical. The factory operates 278 days per year. It has 69 steam engines with a total power of 7,010 horsepower and 2 locomotives with a total horsepower of 45.



93	97,233	86,290	2,875 н. 268,310 т. 163,973	6,113	3,552	249	195	77	33	10,219	2,952	Ваты, пряжи и бумажн. суров., бѣлен., крашен. и набивн. тканей.	13.753,694	19

Source: Perechen' fabrik i zavodov (1896), Page 3. The factory has 93 steam boilers with a total square footage of 97,233. It used 86,290 rubles in plant-based fuels, 2,875 rubles in general mineral fuels, 268,310 rubles in oil, 163,973 rubles in peat. They employ 6,113 adult men, 3,552 adult women, 249 adolescent men, 195 adolescent women, 77 boys, and 33 girls, for a total number of workers of 10,219. They also employ 2,952 workers outside of the factory. They produce cotton wool, yarn, and woven and dyed fabrics. Their total annual output is 13,753,694 rubles.

Figure 2: 1900 Factory List Example



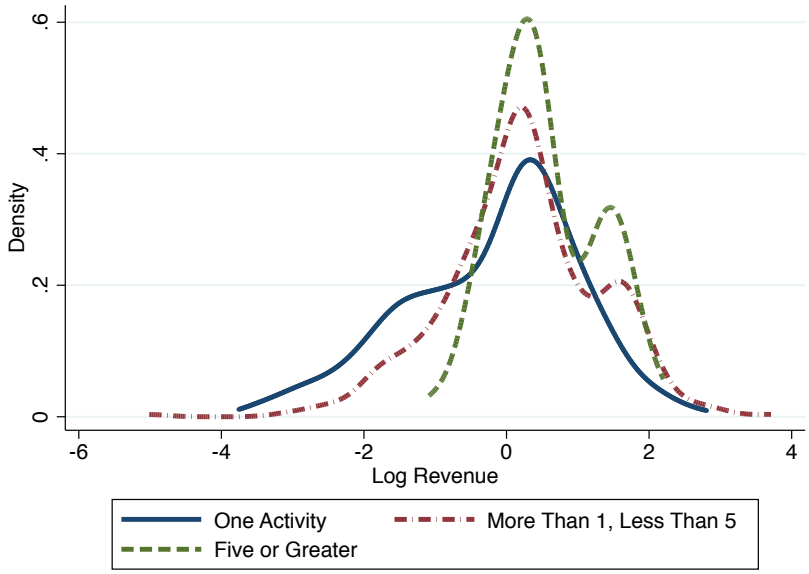
Смирновъ, Ал. Вас., пот. поч. гр. Бумагопрядильн. фабр. (осн. 1899 г.). Покров. у., Кудыкин. вол., д. Ликино. *Ближ. ст. Ликино, Орѣховск. подъѣздн. пути, ½ в., ст. Дрезна, М.-Н. ж. д., 7 в., грунт. д. Почт. и тел. адресъ конт.: Москва, Юшковъ пер., Шуйское подворье, №№ 29—30.* Год. произв. 61,300 (*1.216,843*) р. *Перераб. хлоп. ок. 91,700 п.* Чис. раб. 494 (*458*).

Смирновъ, Ал. Вас., пот. поч. гр. Бумаготкац. звд. (осн. 1881 г.). Покров у., Кудыкин. вол., д. Ликино. *Ближ. ст. Ликино, Орѣховск. подъѣздн. пути М.-Н. ж. д., проселоч. д. Адресъ для прост. корр: ст. Ликино, для заказной: Дулевское почт. отд.; конт.: Москва, Шуйское подворье, №№ 29—30.* Выраб. тк. шестротк. Год. произв. 120,500 р. (*91,425 п. на 1.988,700 р.*). *Перераб. бумаж. и вигонев. бум. пряж. и мал. ч. льнян. пряжи.* Чис. раб. 1,031.

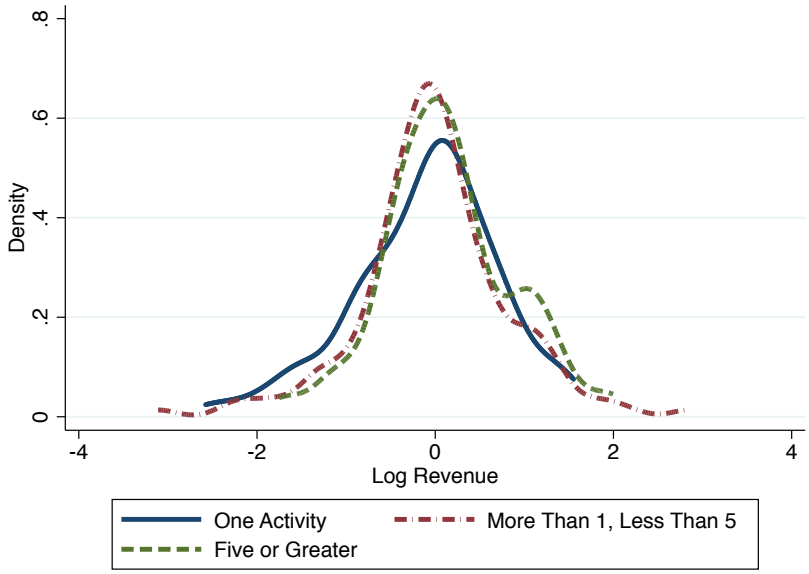
Source: 1900 Factory List, Page 7 (Vladimirskaia Guberniia). Translation: First entry: "Smirnov, Al. Vac. Hereditary citizen. Cotton spinning factory. (Founded 1899). Pokrov. uezd, Kudykin volost, Likino road. Close to Likino station, Orelkhovsk. route, 1/2 verst, st. Dresna, M-N Railroad, 7 versts. Mail and telephone address: Moscow, Iushkov Lane, Shiuskoe Compound, Numbers 29-30. Yearly output 61,300 (1,216,843) rubles. Processing of about 91,700 puds of cotton. Number of workers: 494 (458)." Second entry: "Smirnov, Al. Vac. Hereditary citizen. Cotton spinning factory. (Founded 1881). Pokrov. uezd, Kudykin volost, Likino road. Close to Likino station, Orelkhovsk. route, M-N Railroad. Address for correspondence: Likino Station, for orders: Dulevskoe mail department. Distribution: Moscow, Shuiskoe Compound, Numbers 29-30. Yearly output: 120,500 rubles (91,425 pounds and 1,988,700 rubles). Processing of cotton and vicuna yarn and (small gauge) linen yarn. Number of workers: 1,031."

Figure 3: Production Function Residuals

Residuals from Regressing Revenue on Labor (1894 and 1900)



Residuals from Regressing Revenue on Labor and Capital (1894)



Source: Imperial Russian Manufacturing Database (2015).

Appendix

Determinants of Firm Activities and Firm-Level Production Functions

This section repeats much of the analysis presented above at the level of the firm instead of the factory. A more common vertical integration model considers division of activities across establishments owned by the same firm. For example, a vertically integrated firm can own three factories, each of which perform a distinct stage in their production process. In the case of Russian textiles, many processes occurred in single factory buildings. Firm-level analysis will confirm that integration patterns in Russia are similar when the unit of analysis is the factory or firm

Firm-Level Vertical Integration and Production Functions

In this section I aggregate factories into the firms that owned them to perform analysis at the firm level rather than the factory level. The value of revenue, number of workers, and machine power for the firm are taken to be the sum of those for each factory. The age of the firm is the age of the oldest factory that belongs to the firm. In all regressions, I control for the number of factories that belong to the firm (firm size).

Results for a tobit regression showing the determinants of number of activities for firms is shown in Table A1. The results are similar to those shown in factory-level analysis. Column 1 controls for the number of establishments (factories) within the firm (the firm size), the firm's number of workers, and the firm's age. As expected, the number of activities in the

firm increased in the number of establishments owned by the firm. The number of activities increased in the log number of workers and firm's age, though the coefficient on the log of firm age is small and not statistically significant. Column 2 shows that firms located in European Russia tended to have fewer activities. Though corporate firms had more activities, the difference is not statistically significant.

Columns 2 and 3 present the same regressions presented in Columns 1 and 2, though now the dependent variable is the number of non-cotton activities taking place within the firm. Neither the number of establishments within the firm nor the firms's age are significant determinants of the number of non-cotton activities within the firm. Log firm size has a similarly-sized coefficient as it did for the regressions using overall number of activities. Firms outside European Russia tended to have more non-cotton activities. Also, corporations tended to have more non-cotton activities within firms than non-corporations, though the difference is not statistically significant.

Table A1 Columns 5 through 7 shows similar tobit regressions for both years combined. As in Section 6.1, firms outside European Russia are excluded from these regressions, since the 1900 volume lists only factories in European Russia. Columns 5 and 6 show regressions for all activities that can take place within the firm, while Column 7 shows only cotton activities, which places both 1894 and 1900 on the same basis since the 1900 volume only includes cotton activities. The number of total activities or cotton activities increases in firm size, log number of workers, and factory age; and the number of activities or cotton activities is smaller for firms in 1900.

The previous section presented results that suggested that vertically integrated Russian factories produced more revenue per worker. Table A2 shows similar patterns at the firm

level. Column 1 through 4, which use only the 1894 data, show that firms with more overall activities, more machine power, and fewer workers tended to produce more revenue per worker. Corporations were also more productive. Columns 5 through 8 use both the 1894 and 1900 data: in these regressions, firms with more factories produce more revenue per worker, though the result is fragile: when I control for whether the firm is a corporation, the coefficient on firm size loses statistical significance. Vertical integration and corporate ownership are strongly related to firms' revenue per worker.

Determinants of Number of Products

The 1894 factory-level volume provides information not only on the activities taking place within the factory but also the factory's final products. Table A3 shows similar regressions to those presented in Table 7. The number of products, whether defined as total number of products, cotton products only, or cotton with bricks and chemicals, increases with the number of workers and decreases if the factory was located in European Russia. The coefficient on log factory age is small and changes sign when I control for whether a factory was located in European Russia.

Table A1: Determinants of Cotton Firm Activities, 1894 and 1900

Years:	1894				1894 and 1900		
Tobit	Dep. Var: Log Number of Activities		Dep. Var: Log Num. of Non-Cotton Activities		Dep. Var: Log Number of Activities		Dep. Var: Log Num. of Cotton Activities
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
No. Firm Establishments	.24**	.13	.015	-.019	.20*	.20*	.28**
	(.094)	(.092)	(.095)	(.092)	(.081)	(.082)	(.084)
Log Num. of Workers	.34***	.34***	.36***	.39***	.33***	.35***	.30***
	(.023)	(.027)	(.062)	(.070)	(.024)	(.030)	(.031)
Log Factory Age	.067	.15***	.052	.083	.19***	.19***	.20***
	(.037)	(.039)	(.076)	(.076)	(.039)	(.039)	(.041)
European Russia		-.86***		-.93*			
		(.14)		(.36)			
Corporation		.12		.077		-.082	-.13
		(.12)		(.15)		(.12)	(.12)
Year = 1900					-.98***	-.99***	-.77***
					(.081)	(.082)	(.085)
Intercept	-1.62***	-1.00***	-2.43***	-1.85***	-2.09***	-2.14***	-2.26***
	(.15)	(.17)	(.42)	(.41)	(.17)	(.18)	(.19)
Chi2	287.78	330.09	53.95	60.31	449.62	449.07	341.59
R2	.2227	.2559	.1910	.2135	.1998	.1998	.1613
N	532	531	141	141	1123	1122	1125

*** $p < .001$, ** $p < .01$, * $p < .05$. Source: Imperial Russian Manufacturing Database (2015). Tobit is left-censored at zero, since a firm has one or more activities. Standard errors in parentheses.

Table A2: Cotton Industry Firm-Level Production Functions (1894 and 1900)

OLS	Dependent Variable: Log Revenue per Worker							
	1894 Only				1894 and 1900			
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
No. Factories in the Firm	-.059	.028	.019	.028	.082**	.094**	.058	.067
	(.076)	(.064)	(.063)	(.062)	(.031)	(.033)	(.035)	(.037)
Log (Number of Activities)	.30***	.20*	.18*	.17	.27***	.29***	.26***	.15*
	(.067)	(.082)	(.082)	(.085)	(.051)	(.058)	(.056)	(.066)
Log (Machine Power)		.19***	.16**	.17**				
		(.046)	(.047)	(.049)				
Log (Number of Workers)		-.31***	-.32***	-.34***		-.014	-.11***	-.069*
		(.057)	(.057)	(.058)		(.022)	(.026)	(.029)
Form = Corporation			.35**	.33**			.81***	.74***
			(.12)	(.13)			(.088)	(.092)
Log (Firm Age)				.072				-.019
				(.045)				(.034)
YEAR = 1900								-.34***
								(.082)
Intercept	6.97***	7.82***	7.93***	7.82***	6.61***	6.66***	7.02***	7.09***
	(.10)	(.15)	(.16)	(.19)	(.059)	(.11)	(.12)	(.13)
F-Stat	10.32	7.17	7.35	6.87	29.36	19.93	34.66	23.44
R2	.0354	.1112	.1344	.1425	.0281	.0286	.0848	.0986
N	463	312	311	295	936	936	935	873

*** $p < .001$, ** $p < .01$, * $p < .05$. Source: Imperial Russian Manufacturing Database (2015).
Regressions using data from 1894 and 1900 (Column 5 through 8) include European Russia only.

Table A3: Determinants of Number of Cotton Factory Products (1894)

Tobit	Dep. Var: Log Number of Products			Log Number of Products, Cotton Only			Log Number of Cotton Products with Bricks and Chemicals		
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
Log Num. of Workers	.30*** (.036)	.30*** (.041)	.36*** (.045)	.30*** (.036)	.31*** (.041)	.38*** (.047)	.29*** (.035)	.30*** (.040)	.36*** (.044)
Log Factory Age		-.047 (.059)	.069 (.062)		-.070 (.058)	.045 (.060)		-.057 (.057)	.054 (.060)
European Russia			-1.36*** (.23)			-1.40*** (.226)			-1.32*** (.220)
Intercept	-2.26*** (.24)	-2.17*** (.27)	-1.57*** (.24)	-2.28 (.24)	-2.15 (.27)	-1.57 (.24)	-2.21 (.24)	-2.10 (.26)	-1.52 (.24)
Chi2	94.70	81.64	123.44	97.74	85.49	131.43	95.20	82.48	124.37
R2	.1226	.1156	.1748	.1304	.1251	.1922	.1249	.1185	.1787
N	625	570	570	625	570	570	625	570	570

Source: Imperial Russian Manufacturing Database (2015). Tobit is left-censored at zero, since a firm has one or more activities. Standard errors in parentheses. Chemicals includes Alizarine.