**Online Appendix**

Maps of Soil Fertility and Working Days by Province

Figures A1 and A2 provide province-level maps that help visualize the geography of Russian agriculture and factory working days. In particular, Figure A1 provides a map by province of mean factory working days, and Figure A2 shows a map by province of mean soil fertility variable provided by Dower et. al (2018). We can observe that Figure A2, which plots the average proportion of fertile soil by Province, is almost inverse of Figure A1, which plots the average number of working days by province. Provinces with fertile soil tended to have fewer working days. Note also that, in Figure A2, colder Northern provinces have more working days: factories did not shut down more in cold places due to climate constraints.

Additional Results by Industrial Category

The timing of a factory’s working year can vary greatly due to the agricultural growing season and also due the particular productive processes of each industry. The lumber industry, for example, is more seasonal than other industries both due to the industry’s tendency to locate factories outside cities and because logging is easier in the winter, when trees are free of leaves and when sledges can be employed to transport timber. Results in this appendix thus consider some of the key regression results in the paper by industrial split samples to evaluate whether the relationships between working days, productivity, and other characteristics vary by industrial category.

Table A1 documents the relationship between productivity as measured by revenue per worker and a factory’s total working days, where each column in the table only includes factories in a particular industrial category. These regressions show that, consistently across each industry, working days are positively correlated with revenue per worker. Moreover, urban location and age are also positively correlated with revenue per worker, though in the paper industry, urban factories are less productive.

In Table A2, we examine whether the relationship between a factory’s total working days and other factory characteristics such as urban location, number of nearby factories, age, and total capital differ by industry. Like Table A1, for the most part, the relationships we observed in the main text are stable across industrial categories, though there are some interesting differences in the metals and wood industries. In particular, in metals and wood, the relationship between a factory’s total machine power and its total working days is negative. We speculate that, because most metals factories operated full-year anyway, total machinery may not have been an important source of variation. For the wood industry, it may be the case that the relationship between working days and mechanization is fundamentally different. In the logging industry, more machine power may have allowed a lumber mill to complete a finite amount of work in a season in a shorter amount of time. Again, however, here we can only speculate without more detailed information about the exact timing of a factory’s working year.

Results Including Seasonal Wage Variation

 Table A3 repeats some of the analysis presented in Table 7 of the main text, with the addition of district-level agricultural wage data collected from a source published in Tsentralnii Statisticheskii Komitet (1913), which describes the year 1910. The source provides agricultural wage rates for men and women in the winter and summer, allowing us to calculate, for example, the percentage difference in male agricultural wages for summer vs. winter by district. Unfortunately, and especially because these data describe such a later period, the regressions presented in Table A3 suffer from a major source of endogeneity: likely, industrial activity could have relocated in response to extreme wage variation. Thus, we see here a positive and statistically insignificant relationship between wage variation and working days, where we would have expected a negative and significant relationship if factories shut down in response to high summer wages. These results may mean that by our period, wages in agriculture and industry may have aligned, as suggested by Mironov (2010).

Additional Results for Correlates of Working Days: More Refined Industrial Categories and Survival

In the main text we use a relatively broad categorization of the industries provided by the chapter headings in the original source. However, the volume provides additional highly detailed descriptions of all activities taking place within each factory, from which it is possible to construct finer industrial categories, though with some error, given the inevitability of overlapping categories from factories that perform many disparate functions. Therefore, we regard the enumerators’ classification as our main reference point. At the same time, we recognize that twelve industries could be too few and aggregate factories that are not fully comparable. Therefore, we create a new classification of industries, mainly based on the industry and the first listed activity of a factory. We construct twenty-one industries, all of which have >250 observations. The largest industry, Flour, has 2,844 observations, but sixteen of the industries have less than one thousand observations.

Table A4 replicates Table 3, Column 2 with different specifications. Column 1 copies the original column, Column 2 introduces the new industry controls, and Column 3 does the same excluding the foods industry, which is more dependent on agricultural output than others. We see that all rows other than log machine power are basically identical. In Column 2, we see that the sign on log machine power has reversed. We find that this is largely caused by new foods subgroups, specifically flour and drinks. These industries are themselves seasonal, and any relationship we see with this should be taken with caution. When we exclude all foods in Column 3, we find an almost exact 0, which makes us more confident that we are dealing with measurement error associated with smaller groups and error associated with using listed activities instead of assigned industries. In Column 4, we include an additional control for whether firms exit before 1900. In this specification, we exclude factories with fewer than fifteen workers to mirror our baseline survival results. We see very similar results to the original regression. In Column 5, we do not make these exclusions for demonstration purposes, and find a 0 in the relationship between machine power and working days. Given that in this specification the definition of survival is biased to deem all small firms to have exited, when they in fact may have not been enumerated in 1900 by construction, the relationship between working days and machine power is thus obscured by the incorrectly measured survival dummy.

Additional Survival Estimates: Naïve Matching, More Refined Industrial Categories, and an Alternative Definition of Full-Time Work

Table A5 presents additional robustness checks for the survival regressions presented in the main text in Table 5 Panel B. Table A5 Column 1 replicates Table 5 Panel B Column 4, a central survival result, for reference. In Column 2, we control for the more refined industrial categories, as explained above. The results are similar, though the estimate for the coefficient on log Working Days has become much noisier. Excluding categories within the foods industry, the coefficient more than triples, again confirming that the foods industry was particularly volatile.

Columns 4 through 7 of Table A5 implement a measure of survival in which all industrial categories and factories of all sizes are included, unlike in the main text, where we make a greater effort to create parity between the 1894 and 1900 factory lists. Compared to the original estimates presented in the main text in Table 5 Panel B, the coefficient estimates are similar in magnitude and direction, if just slightly larger. Thus the decision to exclude small factories does not drive the difference in survival between factories that operate more vs. fewer working days.

 In the main text, we remain agnostic about the number of days that represent full-time factory operation, and we perform analysis that split the data by the median number of working days in each Industry and Region pairing (Table 2). Here, we attempt to construct a measure of full-time equivalent months similar to that used in Atack, Bateman and Margo (2002). Although we are unable to reconstruct this exact variable using our data, we can use an alternative definition of full-time provided by contemporary literature. Dement’ev (1897) found an average number of working days in mechanized factories to have been 276 days. We recreate Table 2 and Table 5 Panel A in Tables A6 and A7 respectively defining full-year operation as 276+ working days. The differences we find between factories working greater than vs. fewer than 276 days are similar but smaller compared to what we find splitting the data by the median Industry-Region number of working days.

Additional Description of the 1900 Factory Census

Table A8 provides descriptive statistics for the 1900 factory census, to which the 1894 factories were matched in order to assess their probability of survival. Although this paper has not used the 1900 data except as an indication of survival, we provide summary statistics on factories in the 1900 database to provide a sense of how the factories described by each source differ.

As explained in the main text, there are important ways the sample frames differ between the 1894 and 1900 factory censuses. First, while the 1894 data covers the entire Russian Empire, the 1900 census only covers European Russia. Second, in practice, the 1894 data seem to include more small factories. For these reasons, the number of observations covered by the 1900 database is smaller than that included in 1894. Furthermore, the 1900 database includes far fewer variables for each factory, which is why so few variables are included in Table A8 compared to Table 1. Most importantly, the 1900 database does not include a measure of factory working time.

Compared to factories in the 1894 database, factories in 1900 tended to be larger on average in terms of number of workers and were more likely to be located in urban areas. The distribution of factories across industries and regions, however, is roughly similar between the two years. Regressions and other results that match 1894 factories to 1900 include adjustments to account for differences in industrial composition and factory size.

Additional Regressions with 1900 Wage Aggregates

In the main text, Table 3 includes regressions with measures of wages on the right-hand side to demonstrate that, while factories with more working days paid more wages overall, factories with fewer working days tended to pay more per day. These exercises partly replicate similar results in Atack et. al (2002) Table 3, with a few caveats: first, wages are only measured at the province-industry level, and second, the model employed by Atack et. al (2002) places the wage measures on the left-hand side.

Table A9 Panels A and B present OLS regressions with measures of factory wages on the left-hand side and with a similar set of controls to those used in Atack et. al (2002) Table 3, though still at the factory level. We report these regressions in the appendix rather than the main text, because with this functional form, the left-hand side variable does not vary much across factories. Panel A considers a continuous measure of working days as the main variable of interest on the right-hand side. We see that, in all columns, the number of working days is positively correlated with log wages and negatively correlated with log wages per working day. Note that Row 1 of Columns 4-6 is equal to Columns 1-3 minus one, respectively (We present these columns to provide a convenient comparison to the results presented in Atack et al. (2002)). Panel B more closely follows the regressions presented in Atack et. al (2002) Table 3 by using a binary indicator of longer working time, whether a factory’s working days exceeds its region-industry median. Once again, results are similar: factories that work more than the median working days in an industry-region cell have higher wages and lower wages per day in the province and industry in which they are located.

Descriptive Statistics for Young and Exiting Factories

 A complete picture of the long-run trajectory of part-year operation in the Russian Empire requires an understanding of the relationship between factory entry and exit and working days. In the main text, we present evidence that factories that tended to work fewer working days per year were more likely to exit. Tables A10 and A11 enhance this picture by providing descriptive statistics of very young and exiting factories. To keep these results comparable and consistent to our survival estimates, we include only factories in European Russia and with 15 or more employees.

 Table A10 includes four panels describing young factories, those with an age of less than three years. Panels A through C provide general description, while Panel D compares younger and older factories. Panel D highlights the many differences between younger and older factories: younger factories work far fewer days per year, are much smaller in terms of number of workers or total machine power, are more capital intensive in terms of power per worker, and are roughly similar in terms of workforce composition, urban vs. rural location, and number of factories in the same district and industry. Many of these differences are similar as those between surviving and exiting factories.

 Comparisons between exiting and surviving factories are highlighted in Table A11. Factories that did not survive to 1900 worked fewer days per year, had much less machine power (and were less capital intensive), were much smaller in terms of number of workers, were younger, were more likely to employ children, employed slightly more women, were located near a larger number of other factories in the same industry, and were more likely to be located in urban locations.

 Placing the descriptive statistics on entering and exiting factories in parallel presents a more complex picture of the evolving Russian industrial sector’s relationship with working time. Factories with more machine power had more working days and were more likely to survive. Entering factories were more capital intensive but smaller overall. It seems likely that entering as a full-year factory required large capital expenditure but factories that were able to make this expenditure were more likely to survive. Gregg (2020) highlights how limited access to long-term finance created a variety of distortions in the Imperial Russian industrial sector. Furthermore, factories located near other factories, which perhaps enjoyed more robust and less varying labor markets, were more likely to survive.

Given how different European and non-European Russia were on many relevant characteristics, we replicate some of our main results split on this variable in Table A12. In the first two Columns, we observe the same positive relationship between machine power and working days that we see for the combined sample. Note that the non-European Russian sample and thus power we have in these columns are considerably smaller. In Columns 3 and 4 we replicate Table 3, Column 2 and find largely similar results. Most notably, in European Russia, the relationship between machine power and working days appears not to be significant with the inclusion of controls. This could be due to less variation in machine power, as well as correlation between controls causing sensitivity of our results. In Columns 5, we re-do this analysis excluding firms with fewer than 15 workers and operating outside European Russia, as we do when matching the 1894 census to the 1900 census. In this specification, all coefficients in European Russia are significant and go in the same direction as the results in Table 3. Lastly, we repeat the productivity analysis by replicating Table 4, Column 1 focusing on the European Russian sample in Columns 6. The results are again very similar to the original table in direction and significance of coefficients.

Table A1. Revenue per Worker and Working Days, Split Samples by Industry

|  |  |
| --- | --- |
|   | Dependent Variable: Revenue per Worker |
| Industry Included: | Animal | Chemical | Cotton | Flax | Foods | Metals and Machines | Mineral |
|  | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Working days | 0.0028\*\*\* | 0.0066\*\*\* | 0.0081\*\*\* | 0.0014 | 0.0058\*\*\* | 0.0036\*\*\* | 0.0046\*\*\* |
|  | (0.0005) | (0.0020) | (0.0016) | (0.0017) | (0.0005) | (0.0005) | (0.0003) |
| City | 0.3347\*\* | 0.5571\*\* | 0.0928\* | 0.2845\*\*\* | 0.1564 | 0.0453 | 0.2385\*\* |
|  | (0.1367) | (0.2154) | (0.0424) | (0.0512) | (0.0920) | (0.0912) | (0.1065) |
| Age / 100 | 0.4179 | -0.0628 | 0.8247\*\* | 0.8141\*\* | 0.1328 | 0.1054 | 0.0900 |
|  | (0.3595) | (0.7833) | (0.3160) | (0.2533) | (0.1541) | (0.3581) | (0.1701) |
| Age squared /1000 | -0.0153 | -0.0035 | -0.0713\*\* | -0.0180 | -0.0163 | 0.0253 | -0.0293 |
|  | (0.0203) | (0.0916) | (0.0246) | (0.0189) | (0.0121) | (0.0423) | (0.0183) |
| Constant | 6.5034\*\*\* | 6.0310\*\*\* | 4.6959\*\*\* | 5.9629\*\*\* | 5.9922\*\*\* | 5.6150\*\*\* | 5.0061\*\*\* |
|  | (0.2030) | (0.4463) | (0.4998) | (0.3196) | (0.1486) | (0.1801) | (0.1218) |
|  |  |  |  |  |  |  |  |
| Observations | 1,428 | 616 | 464 | 276 | 4,704 | 1,320 | 1,011 |
| *R*-squared | 0.172 | 0.440 | 0.171 | 0.148 | 0.286 | 0.136 | 0.333 |
| Regional controls | YES | YES | YES | YES | YES | YES | YES |

*Notes*: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. The outcome variable is the revenue per worker. This table mirrors the results in Table 4, Column 1, by industry. Robust standard errors in parentheses. *Source*: Ministry of Finance, *List of Factories and Plants* (1897).

Table A1 (cont…)

|  |  |
| --- | --- |
|  | Dependent Variable: Revenue per Worker |
| Industry Included: | Mixed Materials | Paper | Silk | Wood | Wool |
|  | (8) | (9) | (10) | (11) | (12) |
|  |   |   |   |   |   |
| Working days | 0.0078\*\* | 0.0010 | 0.0083\*\*\* | 0.0026\*\*\* | 0.0033\*\*\* |
|  | (0.0025) | (0.0010) | (0.0010) | (0.0006) | (0.0008) |
| City | 0.5377\*\*\* | -0.2939\* | 0.2434 | -0.0392 | 0.3046\*\* |
|  | (0.1441) | (0.1494) | (0.2219) | (0.1709) | (0.1287) |
| Age / 100 | 0.5937 | 0.6216\*\* | 1.6789\*\*\* | -0.3919 | 0.1538 |
|  | (0.8053) | (0.2097) | (0.2497) | (0.3722) | (0.4838) |
| Age squared /1000 | -0.0644 | -0.0275\*\* | -0.0827\*\*\* | 0.0249\* | -0.0122 |
|  | (0.1396) | (0.0094) | (0.0115) | (0.0125) | (0.0248) |
| Constant | 3.3709\*\*\* | 5.9476\*\*\* | 4.2123\*\*\* | 6.9235\*\*\* | 6.1793\*\*\* |
|  | (0.4352) | (0.2844) | (0.3075) | (0.2265) | (0.1762) |
|  |  |  |  |  |  |
| Observations | 315 | 657 | 211 | 803 | 686 |
| *R*-squared | 0.505 | 0.137 | 0.391 | 0.119 | 0.323 |
| Regional controls | YES | YES | YES | YES | YES |

*Notes*: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. The outcome variable is the revenue per worker. This table mirrors the results in Table 4, Column 1, by industry. Robust standard errors in parentheses. *Source*: Ministry of Finance, *List of Factories and Plants* (1897).

Table A2. Working Days and Factory Characteristics, Split Samples by Industry

|  |  |
| --- | --- |
|   | Dependent Variable: Number of Working Days |
| Industry Included: | Animal | Chemical | Cotton | Flax | Foods | Metals and Machines | Mineral |
|  | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| *Log*  | 0.078\*\*\* | -0.012 | 0.005 | 0.063\*\*\* | 0.023\*\*\* | -0.021\*\*\* | -0.015 |
| (machine power +1) | (0.018) | (0.014) | (0.010) | (0.018) | (0.006) | (0.006) | (0.012) |
| *Log* workers | 0.016 | 0.090\*\*\* | 0.060\*\*\* | 0.027 | 0.021\*\*\* | 0.069\*\*\* | 0.201\*\*\* |
|  | (0.015) | (0.017) | (0.014) | (0.023) | (0.007) | (0.008) | (0.012) |
| City | 0.126\*\*\* | 0.130\*\*\* | -0.007 | 0.124\*\*\* | 0.266\*\*\* | 0.087\*\*\* | 0.144\*\*\* |
|  | (0.036) | (0.041) | (0.033) | (0.047) | (0.019) | (0.019) | (0.030) |
| Age /100 | 0.458\*\*\* | 0.760\*\*\* | 0.092 | -0.165 | 0.258\*\*\* | 0.190\*\*\* | 0.535\*\*\* |
|  | (0.161) | (0.260) | (0.148) | (0.185) | (0.084) | (0.070) | (0.133) |
| Age squared /1000 | -0.026\* | -0.068\* | -0.009 | 0.006 | -0.005 | -0.016\*\*\* | -0.026\*\* |
|  | (0.015) | (0.036) | (0.016) | (0.014) | (0.008) | (0.006) | (0.012) |
| *N* of factories in  | 0.206\*\*\* | 0.268\*\* | 0.253\*\*\* | -0.560\*\*\* | 0.110\*\*\* | 0.010 | -0.012 |
| district-ind./100 | (0.063) | (0.136) | (0.064) | (0.207) | (0.034) | (0.013) | (0.121) |
| Constant | 5.264\*\*\* | 4.896\*\*\* | 5.139\*\*\* | 5.015\*\*\* | 4.997\*\*\* | 5.217\*\*\* | 4.627\*\*\* |
|  | (0.185) | (0.104) | (0.309) | (0.081) | (0.068) | (0.058) | (0.096) |
|  |  |  |  |  |  |  |  |
| Observations | 1,479 | 628 | 574 | 303 | 5,697 | 1,380 | 1,021 |
| Regional controls | YES | YES | YES | YES | YES | YES | YES |
| Pseudo *R2* | 0.0792 | 0.144 | 0.599 | 0.345 | 0.0393 | 0.712 | 0.285 |

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1 The outcome variable is the number of working days of a firm. This table mirrors the results in Table 3, Column 2, by industry. Standard errors in parentheses. *Source*: Ministry of Finance, *List of Factories and Plants* (1897).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Table A2 (cont…) |  |  |  |  |
|  | Dependent Variable: Total Working Days |
| Industry Included: | Mixed Materials | Paper | Silk | Wood | Wool |
|  | (8) | (9) | (10) | (11) | (12) |
| *Log* machine power +1 | 0.003 | 0.004 | -0.005 | -0.045\*\*\* | -0.006 |
|  | (0.010) | (0.006) | (0.012) | (0.012) | (0.008) |
| *Log* workers | 0.010 | 0.034\*\*\* | 0.055\*\*\* | 0.180\*\*\* | 0.112\*\*\* |
|  | (0.012) | (0.008) | (0.017) | (0.018) | (0.012) |
| City | 0.085\*\* | 0.055\*\* | 0.024 | 0.209\*\*\* | -0.028 |
|  | (0.035) | (0.023) | (0.033) | (0.042) | (0.028) |
| Age / 100 | 0.377\* | 0.071 | 0.180 | 0.061 | 0.175 |
|  | (0.216) | (0.053) | (0.210) | (0.158) | (0.124) |
| Age squared / 1000 | -0.046 | -0.003 | -0.018 | 0.001 | -0.006 |
|  | (0.036) | (0.003) | (0.028) | (0.008) | (0.013) |
| Num. of factories in | 0.058 | -0.008 | -0.063 | 0.335\*\*\* | 0.046\*\* |
| district-ind. / 100 | (0.038) | (0.020) | (0.061) | (0.122) | (0.020) |
| Constant | 5.281\*\*\* | 5.525\*\*\* | 4.899\*\*\* | 4.599\*\*\* | 4.896\*\*\* |
|  | (0.068) | (0.042) | (0.063) | (0.109) | (0.237) |
|  |  |  |  |  |  |
| Observations | 366 | 714 | 229 | 851 | 967 |
| Region controls | YES | YES | YES | YES | YES |
| Pseudo *R2* | 1.570 | -0.148 | 8.494 | 0.195 | 0.423 |

Notes: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1 The outcome variable is the number of working days of a firm. This table mirrors the results in Table 3, Column 2, by industry. Standard errors in parentheses. *Source*: Ministry of Finance, *List of Factories and Plants* (1897).

Table A3: Working Days, District Characteristics, and Seasonal Wage Variation

|  |  |
| --- | --- |
|  OLS | *Log* Working Days |
| Dependent Variable | (1) | (2) | (3) | (4) |
|   |   |   |   |   |
| Fertile soil |  |  | -0.0320 | -0.0615 |
|  |  |  | (0.0451) | (0.0577) |
| Urban/population |  |  | 0.593\*\*\* | 0.647\*\*\* |
|  |  |  | (0.0929) | (0.0774) |
| Population |  |  | 0.0109 | 0.00811 |
|  |  |  | (0.0152) | (0.0168) |
| Number of strikes 1890-1894 |  |  | -0.108\* | -0.309 |
|  |  |  | (0.0605) | (0.235) |
| Latitude |  |  | 0.00818 | 0.00920 |
|  |  |  | (0.00494) | (0.00593) |
| Male summer wage, living independently | 0.00630\*\*\* |  | 0.00144 |  |
|  | (0.00236) |  | (0.00228) |  |
| Percentage difference in summer vs. winter wages for males living independently |  | 0.0296 |  | 0.0163 |
|  |  | (0.0254) |  | (0.0307) |
| Constant | 5.152\*\*\* | 5.236\*\*\* | 4.728\*\*\* | 4.698\*\*\* |
|  | (0.0402) | (0.0224) | (0.278) | (0.336) |
|  |  |  |  |  |
| Observations | 536 | 392 | 422 | 305 |
| *R*-squared | 0.009 | 0.002 | 0.079 | 0.116 |
| *Notes*: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Standard errors clustered by province in parentheses. *Sources*: Ministry of Finance *List of Factories and Plants* (1897), Dower et al. (2018), Central Statistical Committee (1910).  |
|  |  |  |  |  |
|  |  |  |  |

|  |
| --- |
| Table A4. Firm Characteristics and Operating Time |
|  | Dependent Variable: *Log* (Working Days), Tobit |
|  | (1) | (2) | (3) | (4) | (5) |
|  |  |  |  |  |  |
| *Log* (total machine  | 0.008\*\* | -0.017\*\*\* | -0.004 | 0.011\*\*\* | -0.000 |
| power + 1) | (0.003) | (0.003) | (0.004) | (0.003) | (0.003) |
| *Log* workers | 0.059\*\*\* | 0.092\*\*\* | 0.082\*\*\* | -0.008 | 0.054\*\*\* |
|  | (0.004) | (0.004) | (0.004) | (0.006) | (0.004) |
| City  | 0.164\*\*\* | 0.189\*\*\* | 0.112\*\*\* | 0.157\*\*\* | 0.160\*\*\* |
|  | (0.010) | (0.010) | (0.010) | (0.011) | (0.010) |
| Age | 0.249\*\*\* | 0.220\*\*\* | 0.189\*\*\* | 0.232\*\*\* | 0.212\*\*\* |
|  | (0.040) | (0.039) | (0.040) | (0.041) | (0.039) |
| Age squared | -0.009\*\*\* | -0.012\*\*\* | -0.011\*\*\* | -0.008\*\* | -0.007\*\*\* |
|  | (0.004) | (0.003) | (0.003) | (0.003) | (0.003) |
| Number of factories | 0.058\*\*\* | 0.036\*\*\* | 0.036\*\*\* | 0.028\*\* | 0.068\*\*\* |
| in district/ind. / 100 | (0.011) | (0.011) | (0.011) | (0.011) | (0.011) |
| Survived to 1900 |  |  |  | 0.057\*\*\* | 0.074\*\*\* |
|  |  |  |  | (0.010) | (0.009) |
| Constant | 5.016\*\*\* | 4.672\*\*\* | 4.726\*\*\* | 5.217\*\*\* | 5.016\*\*\* |
|  | (0.035) | (0.038) | (0.036) | (0.040) | (0.034) |
|  |  |  |  |  |  |
| Observations | 14,209 | 14,209 | 8,512 | 7,497 | 13,786 |
| Industry controls | YES(old) | YES | YES | YES | YES |
| Regional controls | YES | YES | YES | YES | YES |
| *R2* or pseudo R2 | 0.109 | 0.156 | 0.250 | 0.173 | 0.118 |

*Notes*: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Column 1 copies Column 2 of Table 3 for comparison; Column 2 includes our new definition of industries; Column 3 does the same and excludes the Foods industry; Column 4 controls for firm survival/exit, based on whether firms survived to 1900 (1894 firms with fewer than 15 employees dropped) Column 5 does the same as Column 4, but does not drop 1894 firms described in Column. The outcome variable is the log of the factory’s number of working days. The Tobit model in all regressions is right censored at 6, because log(365) is approximately equal to 6. Standard errors in parentheses. *Source*: Ministry of Finance, *List of Factories and Plants* (1897).

Table A5. Probit Regressions: Probability of 1894 Factory to Appear in 1900 (European Russia Only)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|   | (1) | (2)  | (3) | (4) | (5) | (6) | (7) |
|  | *Probit* (Table 5, Column 4) | *Probit* (New Industries) | *Probit* (New Industries, no Foods) | *Probit* (Crude comparison) | *Probit* (Crude comparison) | *Probit* (Crude comparison) | *Probit* (Crude comparison) |
|   |   |  |  |  |  |  |  |
| *Log* (working  | 0.081\*\* | 0.049 | 0.188\*\*\* | 0.366\*\*\* | 0.346\*\*\* | 0.208\*\*\* | 0.252\*\*\* |
| days) | (0.041) | (0.043) | (0.057) | (0.019) | (0.020) | (0.025) | (0.026) |
|  | [0.031]\*\*\* | [0.031] | [0.073]\*\*\* | [0.145]\*\*\* | [0.137]\*\*\* | [0.083]\*\*\* | [0.101]\*\*\* |
| *Log* (revenue per  | 0.143\*\*\* | 0.179\*\*\* | 0.170\*\*\* |  |  | 0.139\*\*\* |  |
| worker) | (0.018) | (0.020) | (0.022) |  |  | (0.013) |  |
|  | [0.055]\*\*\* | [0.055]\*\*\* | [0.066]\*\*\* |  |  | [0.056]\*\*\* |  |
| TFP |  |  |  |  |  |  | 0.078\*\*\* |
|  |  |  |  |  |  |  | (0.013) |
|  |  |  |  |  |  |  | [0.031]\*\*\* |
| Constant | -1.856\*\*\* | -2.221\*\*\* | -2.819\*\*\* | -2.075\*\*\* | -2.255\*\*\* | -2.687\*\*\* | -1.943\*\*\* |
|  | (0.243) | (0.258) | (0.329) | (0.101) | (0.136) | (0.162) | (0.160) |
|  |  |  |  |  |  |  |  |
| Observations | 7,048 | 7,048 | 5,172 | 16,487 | 16,487 | 12,111 | 12,111 |
| Industry controls | YES | YES(new) | YES(new) | NO | YES | YES | YES |
| Regional controls | YES | YES | YES | NO | YES | YES | YES |
| Age, Age2 controls | YES | YES | YES | NO | NO | YES | YES |
| Pseudo *R2* | 0.0493 | 0.0619 | 0.0663 | 0.0165 | 0.0380 | 0.0496 | 0.0446 |

*Notes*: \*\*\* p<0.01, \*\* p<0.05, \*p<0.10. Robust standard errors in parenthesis. Marginal effects (dprobit) in brackets. Regressions estimate the probability that an 1894 factory appears in the 1900 database. TFP in Column 7 is the residual of a log Cobb Douglas production function that includes total workers and total machine power as inputs. The regressions in Columns 4 through 7 include all industrial categories and factories of all sizes and thus represents a crude comparison between sources. *Source*:Ministry of Finance, *List of Factories and Plants* (1897)and Ministry of Finance, *List of Factories and Plants* (1903).

Table A6: Operation Duration and Firm Characteristics, using Dement’ev’s definition of Full-Year

|  |  |  |  |
| --- | --- | --- | --- |
| Variable | Operating More Than 276 Days | OperatingFewer Than 276 days | Test Statistic |
|  |  |  | Two-Sample t-test, |t| |
| Total machine power | 51.75 | 27.93 | 7.4610 |
|  | (3.62) | (1.36) |  |
|  |  |  |  |
| Total workers | 86.99 | 52.01 | 8.0425 |
|  | (4.61) | (2.04) |  |
|  |  |  |  |
| Total machine power  | 0.77 | 1.19 | 9.2580 |
| per worker | (0.030) | (0.028) |  |
|  |  |  |  |
| Age | 21.62 | 19.32 | 6.3314 |
|  | (0.31) | (0.21) |  |
|  |  |  |  |
| Number of factories | 46.38 | 29.15 | 23.6387 |
| in district-industry | (0.75) | (0.35) |  |
|  |  |  |  |
| Number of Women  | 21.56 | 10.61 | 6.3691 |
| employed | (1.94) | (0.74) |  |
|  |  |  |  |
| Number of children | 1.98 | 0.79 | 8.0144 |
| employed | (0.172) | (0.061) |  |
|  |  |  |  |
| Women employed /  | 0.10 | 0.069 | 10.8706 |
| total workers | (0.002) | (.002) |  |
|  |  |  |  |
| Children employed /  | 0.0221 | 0.0127 | 8.3698 |
| total workers | (0.0011) | (0.00059) |  |
|  |  |  | Two-Proportion z-test, |z| |
| Urban | 0.65 | 0.33 | 39.2010 |
|  | (0.0064) | (0.0044) |  |

*Notes*: 276 days represents the number of days Dement’ev counted as average for full-year machine-powered factories in 1897. Total machine power measures the total amount of horsepower in a firm. Standard errors in parentheses. Number of factories in district-industry is the number of factories in the same district and industry. *Source*: Ministry of Finance, *List of Factories and Plants* (1897).

Table A7: Factory Working Days and Survival to 1900 using Dement’ev’s definition of Full-Year

Panel A: Factory Survival and Part-Year Operation (European Russia and Factories with Greater than 15 Workers Only)

|  |  |  |  |
| --- | --- | --- | --- |
|  | All Factories in European Russia |  <276 days | >=276 days |
|  |  |  |  |
| Number of factories in 1894 | 8,448 | 4,814 | 3,563 |
|  |  |  |  |
| Number of factories that survive to 1900 | 5,000 | 2,832 | 2,132 |
|  |  |  |  |
| Survivors to 1900 / factories in 1894 | 59.19 | 0.5883 | 0.5984 |

*Sources*:Ministry of Finance, *List of Factories and Plants* (1897)and Ministry of Finance, *List of Factories and Plants* (1903)

Table A8: Descriptive Statistics from the 1900 Factory Census

Panel A: Descriptive Statistics

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | *N* | *Mean* | *Med.* | *Standard Dev.* | *Min.* | *Max.* |
| Number of Workers | 15,649 | 100.79 | 25.00 | 384.58 | 0.0 | 12,346 |
| Urban | 15,626 | 0.47 | 0.00 | 0.50 | 0.0 | 1 |

Panel B: Frequency, Location, Mean Number of Workers by Industry

|  |  |  |  |
| --- | --- | --- | --- |
| Industry | Observations | Percentage Urban | Mean Number of Workers |
| Animal | 1,254 | 0.568 | 41.410 |
| Chemical | 530 | 0.543 | 103.023 |
| Cotton | 731 | 0.392 | 544.919 |
| Flax | 414 | 0.604 | 189.440 |
| Foods (Untaxed) | 2,501 | 0.446 | 28.673 |
| Foods (Taxed) | 2,797 | 0.212 | 75.853 |
| Metals and machines | 1,806 | 0.736 | 136.380 |
| Mineral products | 1,590 | 0.375 | 84.028 |
| Mixed materials | 341 | 0.821 | 80.070 |
| Paper | 1,072 | 0.816 | 68.064 |
| Silk | 308 | 0.383 | 103.591 |
| Wood | 1,426 | 0.413 | 51.741 |
| Wool | 879 | 0.437 | 141.270 |
| Total | 15,649 | 0.474 | 100.764 |
|  |  |  |  |

Panel C: Frequency, Location, Mean Number of Workers by Region

|  |  |  |  |
| --- | --- | --- | --- |
| Region | Observations | Percentage Urban | Mean Number of Workers |
| Caucasus | 388 | 0.767 | 58.601 |
| Central Black Soil  | 2,662 | 0.429 | 78.300 |
| Central Industrial | 3,373 | 0.406 | 171.484 |
| Eastern | 935 | 0.513 | 62.809 |
| Northern | 641 | 0.463 | 104.740 |
| Northwestern | 1,548 | 0.400 | 30.866 |
| Prebaltic | 1,370 | 0.667 | 161.685 |
| Previslitskii | 2,188 | 0.466 | 82.774 |
| Southern | 1,574 | 0.646 | 56.942 |
| Southwestern | 970 | 0.270 | 104.877 |
| Total | 15,649 | 0.474 | 100.786 |

*Notes*: *Source*: Ministry of Finance, *List of Factories and Plants* (1903).

Table A9: Additional Regressions with 1900 Wage Aggregates

Panel A: Correlates of Province-Industry Log Wages and Log Wages per Working Day

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |
| Dep. Variable: | *Log* Wage | *Log* Wage / Working Days |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
|   |   |   |   |   |   |   |
| *Log* of working days | 0.198\*\*\* | 0.164\*\*\* | 0.089\*\*\* | -0.802\*\*\* | -0.836\*\*\* | -0.911\*\*\* |
|  | (0.007) | (0.008) | (0.006) | (0.007) | (0.008) | (0.006) |
| *Log* machine power +1 |  | 0.028\*\*\* | 0.012\*\*\* |  | 0.028\*\*\* | 0.012\*\*\* |
|  |  | (0.002) | (0.002) |  | (0.002) | (0.002) |
| *Log* workers |  | -0.007\*\* | -0.002 |  | -0.007\*\* | -0.002 |
|  |  | (0.003) | (0.003) |  | (0.003) | (0.003) |
| City |  | 0.246\*\*\* | 0.115\*\*\* |  | 0.246\*\*\* | 0.115\*\*\* |
|  |  | (0.008) | (0.006) |  | (0.008) | (0.006) |
| Age / 100 |  | -0.087\*\*\* | 0.024 |  | -0.087\*\*\* | 0.024 |
|  |  | (0.033) | (0.025) |  | (0.033) | (0.025) |
| Age squared / 1000 |  | 0.003 | -0.004\* |  | 0.003 | -0.004\* |
|  |  | (0.003) | (0.002) |  | (0.003) | (0.002) |
| Constant | -2.865\*\*\* | -2.811\*\*\* | -2.180\*\*\* | -2.865\*\*\* | -2.811\*\*\* | -2.180\*\*\* |
|  | (0.037) | (0.041) | (0.043) | (0.037) | (0.041) | (0.043) |
|  |  |  |  |  |  |  |
| Observations | 16,388 | 13,700 | 13,700 | 16,388 | 13,700 | 13,700 |
| *R*-squared | 0.055 | 0.133 | 0.503 | 0.490 | 0.513 | 0.721 |
| Industry controls | NO | NO | YES | NO | NO | YES |
| Regional controls | NO | NO | YES | NO | NO | YES |

*Notes*: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Robust standard errors in parentheses. Log wage is the log of total wages divided by total enterprises in a province-industry cell, divided by the average number of workers in that province-industry cell. *Sources*: Ministry of Finance, *List of Factories and Plants* (1897), and Ministry of Finance, *Statistical Results on Factories and Plants by Industries Not Subject to the Excise Tax for 1900* (1903).

Panel B: Correlates of Province-Industry Log Wages per Working Day: Alternative Specification with Binary Definition of Working Time

|  |  |  |
| --- | --- | --- |
|  Dep. Variable | *Log* Wage | *Log* Wage / Working Days |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
|   |   |   |   |   |   |   |
| Factory is full year  | 0.076\*\*\* | 0.031\*\*\* | 0.054\*\*\* | -0.540\*\*\* | -0.524\*\*\* | -0.537\*\*\* |
| (reg. med. def.) | (0.007) | (0.008) | (0.006) | (0.009) | (0.009) | (0.008) |
| *Log* machine  |  | 0.027\*\*\* | 0.011\*\*\* |  | 0.044\*\*\* | 0.017\*\*\* |
| power +1 |  | (0.002) | (0.002) |  | (0.003) | (0.003) |
| *Log* workers |  | 0.005\* | -0.000 |  | -0.056\*\*\* | -0.028\*\*\* |
|  |  | (0.003) | (0.003) |  | (0.004) | (0.004) |
| City |  | 0.278\*\*\* | 0.121\*\*\* |  | 0.133\*\*\* | 0.051\*\*\* |
|  |  | (0.008) | (0.006) |  | (0.010) | (0.010) |
| Age /100 |  | -0.059\* | 0.033 |  | -0.143\*\*\* | -0.072\*\* |
|  |  | (0.034) | (0.025) |  | (0.042) | (0.036) |
| Age squared /1000 |  | 0.002 | -0.004\*\* |  | 0.004 | -0.000 |
|  |  | (0.003) | (0.002) |  | (0.003) | (0.003) |
| Constant | -1.859\*\*\* | -2.010\*\*\* | -1.739\*\*\* | -6.834\*\*\* | -6.797\*\*\* | -6.696\*\*\* |
|  | (0.005) | (0.010) | (0.027) | (0.007) | (0.014) | (0.033) |
|  |  |  |  |  |  |  |
| Observations | 16,388 | 13,700 | 13,700 | 16,388 | 13,700 | 13,700 |
| *R*-squared | 0.007 | 0.101 | 0.497 | 0.184 | 0.201 | 0.358 |
| Industry controls | NO | NO | YES | NO | NO | YES |
| Regional controls | NO | NO | YES | NO | NO | YES |

*Notes:* \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Robust standard errors in parentheses. Log wage is the log of total wages divided by total enterprises in a province-industry cell, divided by the average number of workers in that province-industry cell. Sources: Ministry of Finance, *List of Factories and Plants* (1897), and Ministry of Finance, *Statistical Results on Factories and Plants by Industries Not Subject to the Excise Tax for* (1903)*.*

Table A10: Additional Descriptive Statistics about Young Factories (Age < 3) (Excluding Non-European Russia and Factories with Fewer than 15 Employees)

Panel A: Descriptive Statistics

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | *N* | *Mean* | *Med.* | *Standard Dev.* | *Min.* | *Max.* |
| Working days | 726 | 209.56 | 232.00 | 79.19 | 8.0 | 363 |
| Total machine power | 742 | 30.91 | 6.00 | 84.68 | 0.0 | 1,103 |
| Number of workers | 742 | 50.09 | 29.00 | 73.45 | 15.0 | 973 |
| Age | 742 | 1.22 | 1.00 | 0.71 | 0.0 | 2 |
| Urban | 742 | 0.54 | 1.00 | 0.50 | 0.0 | 1 |
| Factories in dist-Ind | 742 | 42.16 | 17.00 | 59.36 | 1.0 | 202 |
| Employed women | 742 | 0.37 | 0.00 | 0.48 | 0.0 | 1 |
| Employed children | 742 | 0.13 | 0.00 | 0.34 | 0.0 | 1 |
| Number of women | 742 | 7.49 | 0.00 | 27.12 | 0.0 | 517 |
| Number of children | 742 | 1.25 | 0.00 | 4.92 | 0.0 | 85 |

Panel B: Frequency, Age, Location, and Operation Duration of Young Factories by Industry

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Industry | Observations | Average Firm Age (Years) | Percentage Urban | MedianWorkingDays | Mean Working Days |
| Animal | 38 | 1.50 | 0.68 | 265.00 | 220.49 |
| Chemical | 31 | 1.23 | 0.48 | 230.00 | 210.16 |
| Cotton | 35 | 1.37 | 0.60 | 250.00 | 228.49 |
| Flax | 17 | 1.18 | 0.53 | 230.00 | 211.00 |
| Foods | 175 | 1.19 | 0.46 | 198.50 | 187.37 |
| Metals and machines | 96 | 1.20 | 0.78 | 265.00 | 237.60 |
| Mineral products | 79 | 1.05 | 0.27 | 140.00 | 156.75 |
| Mixed materials | 31 | 1.23 | 0.55 | 240.00 | 197.83 |
| Paper | 46 | 1.39 | 0.83 | 280.00 | 262.51 |
| Silk | 14 | 1.57 | 0.64 | 247.50 | 227.07 |
| Wood | 95 | 1.17 | 0.35 | 200.00 | 198.54 |
| Wool | 85 | 1.19 | 0.67 | 280.00 | 240.99 |
| Total | 742 | 1.22 | 0.54 | 232.00 | 209.56 |
|  |  |  |  |  |  |

Panel C: Frequency, Age, Location, and Operation Duration of Young Factories by Region

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Region | Observations | Average Firm Age (Years) | Percentage Urban | MedianWorkingDays | Mean Working Days |
| Caucasus | 17 | 0.88 | 0.76 | 200.00 | 209.24 |
| Central Black Soil  | 137 | 1.21 | 0.33 | 175.00 | 178.05 |
| Central Industrial | 134 | 1.29 | 0.46 | 240.00 | 214.02 |
| Eastern | 57 | 1.40 | 0.35 | 177.00 | 177.86 |
| Northern | 9 | 1.33 | 0.33 | 233.00 | 207.78 |
| Northwestern | 64 | 1.13 | 0.63 | 260.00 | 228.31 |
| Prebaltic | 63 | 1.21 | 0.86 | 280.00 | 246.25 |
| Previslitskii | 138 | 1.30 | 0.66 | 275.00 | 239.12 |
| Southern | 94 | 1.06 | 0.70 | 200.00 | 198.18 |
| Southwestern | 29 | 1.10 | 0.31 | 195.00 | 181.66 |
| Total | 742 | 1.22 | 0.54 | 232.00 | 209.56 |

*Notes*: *Total Machine Power* presents the total amount of horsepower in a firm. *Factories in Dist-Ind* is the number of factories located in the same district and industry. *Source*: Ministry of Finance, *List of Factories and Plants* (1897).

Panel D: T-Test Comparisons of Young (Age < 3) and Older (Age $\geq $ 3) Factories

|  |  |  |  |
| --- | --- | --- | --- |
| Variable | Older Factories (Age $\geq $ 3) | Younger Factories (Age < 3) | Test Statistic |
|  |  |  | Two-Sample t-test, |t| |
| Total working days | 243.33 | 209.56 | 12.4904 |
|  | (0.83) | (2.94) |  |
|  |  |  |  |
| Total machine power | 69.79 | 30.91 | 3.6363 |
|  | (3.51) | (3.11) |  |
|  |  |  |  |
| Total workers | 132.48 | 50.09 | 5.7657 |
|  | (4.71) | (2.70) |  |
|  |  |  |  |
| Total machine power  | 0.52 | 0.67 | 3.6109 |
| per worker | (0.013) | (0.046) |  |
|  |  |  |  |
| Age | 24.32 | 1.22 | 29.2363 |
|  | (0.26) | (0.26) |  |
|  |  |  |  |
| Number of factories | 41.27 | 42.16 | 0.4308 |
| in district-industry | (0.64) | (2.18) |  |
|  |  |  |  |
| Number of women  | 31.63 | 7.49 | 4.2159 |
| employed | (1.89) | (1.00) |  |
|  |  |  |  |
| Number of children | 2.47 | 1.25 | 2.4135 |
| employed | (0.17) | (0.18) |  |
|  |  |  |  |
| Women employed /  | 0.13 | 0.11 | 2.1195 |
| total workers | (0.0026) | (0.0069) |  |
|  |  |  |  |
| Children employed /  | 0.023 | 0.032 | 2.6319 |
| total workers | (0.00097) | (0.0039) |  |
|  |  |  | Two-Proportion z-test, |z| |
| Urban | 0.46 | 0.43 | 2.2652 |
|  | (0.0044) | (0.012) |  |

*Notes*: Total machine power measures the total amount of horsepower in a firm. Standard errors in parentheses. *Source*: Ministry of Finance, *List of Factories and Plants* (1897).

Table A11: Additional Descriptive Statistics about Exiting Factories (Excluding Non-European Russia and Factories with Fewer than 15 Employees)

Panel A: Descriptive Statistics

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | *N* | *Mean* | *Med.* | *Standard Dev.* | *Min.* | *Max.* |
| Working days | 3,413 | 235.96 | 265.00 | 72.93 | 6.0 | 365 |
| Total machine power | 3,448 | 30.49 | 2.00 | 101.30 | 0.0 | 2,316 |
| Number of workers | 3,448 | 71.78 | 31.00 | 176.03 | 15.0 | 5,512 |
| Age | 3,023 | 19.34 | 14.00 | 20.44 | 0.0 | 262 |
| Urban | 3,445 | 0.59 | 1.00 | 0.49 | 0.0 | 1 |
| Factories in dist-ind | 3,444 | 44.93 | 20.00 | 56.04 | 1.0 | 202 |
| Employed women | 3,448 | 0.42 | 0.00 | 0.49 | 0.0 | 1 |
| Employed children | 3,448 | 0.16 | 0.00 | 0.37 | 0.0 | 1 |
| Number of women | 3,448 | 13.34 | 0.00 | 52.55 | 0.0 | 1230 |
| Number of children | 3,448 | 1.57 | 0.00 | 6.60 | 0.0 | 194 |

Panel B: Frequency, Age, Location, and Operation Duration of Exiting Factories by Industry

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Industry | Observations | Average Firm Age (Years) | Percentage Urban | Median Working Days | MeanWorkingDays |
| Animal | 302 | 24.42 | 0.69 | 273.00 | 231.74 |
| Chemical | 150 | 13.53 | 0.48 | 256.00 | 226.62 |
| Cotton | 130 | 20.21 | 0.46 | 250.00 | 237.90 |
| Flax | 67 | 17.28 | 0.49 | 180.00 | 179.21 |
| Foods | 737 | 20.98 | 0.54 | 250.00 | 224.84 |
| Metals and machines | 429 | 21.27 | 0.75 | 280.00 | 263.64 |
| Mineral products | 411 | 19.73 | 0.32 | 200.00 | 206.25 |
| Mixed materials | 251 | 16.75 | 0.72 | 265.00 | 239.94 |
| Paper | 251 | 22.96 | 0.84 | 280.00 | 276.83 |
| Silk | 61 | 14.80 | 0.49 | 240.00 | 228.38 |
| Wood | 290 | 12.60 | 0.51 | 250.00 | 224.00 |
| Wool | 369 | 17.09 | 0.66 | 279.00 | 255.86 |
| Total | 3,448 | 19.34 | 0.59 | 265.00 | 235.96 |
|  |  |  |  |  |  |

Panel C: Frequency, Age, Location, and Operation Duration of Exiting Factories by Region

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Region | Observations | Average Firm Age (Years) | Percentage Urban | MedianWorkingDays | Mean Working Days |
| Caucasus | 91 | 13.19 | 0.66 | 260.00 | 234.46 |
| Central Black Soil  | 560 | 20.91 | 0.45 | 218.00 | 204.08 |
| Central Industrial | 893 | 21.17 | 0.53 | 260.00 | 240.49 |
| Eastern | 213 | 21.28 | 0.48 | 200.00 | 186.46 |
| Northern | 66 | 30.62 | 0.34 | 252.00 | 225.27 |
| Northwestern | 182 | 14.09 | 0.56 | 275.00 | 245.45 |
| Prebaltic | 434 | 22.99 | 0.84 | 280.00 | 265.22 |
| Previslitskii | 528 | 14.83 | 0.69 | 287.00 | 268.89 |
| Southern | 305 | 12.72 | 0.75 | 259.50 | 225.63 |
| Southwestern | 176 | 24.50 | 0.34 | 240.00 | 217.38 |
| Total | 3,448 | 19.34 | 0.59 | 265.00 | 235.96 |

*Notes*: *Total Machine Power* presents the total amount of horsepower in a firm. *Source*: Ministry of Finance, *List of Factories and Plants* (1897).

Panel D: T-Test Comparisons of Exiting and Surviving Factories

|  |  |  |  |
| --- | --- | --- | --- |
| Variable | Exiting Factories | Surviving Factories  | Test Statistic |
|  |  |  | Two-Sample t-test, |t| |
| Total working days | 235.96 | 241.40 | 3.5009 |
|  | (1.25) | (0.96) |  |
|  |  |  |  |
| Total machine power | 30.49 | 87.85 | 9.4405 |
|  | (1.73) | (4.90) |  |
|  |  |  |  |
| Total workers | 71.78 | 154.70 | 10.2134 |
|  | (3.00) | (6.42) |  |
|  |  |  |  |
| Total machine power  | 0.45 | 0.60 | 5.8431 |
| per worker | (0.018) | (0.016) |  |
|  |  |  |  |
| Age | 19.34 | 23.86 | 8.9761 |
|  | (0.37) | (0.33) |  |
|  |  |  |  |
| Number of factories | 44.93 | 37.58 | 6.3199 |
| in district-industry | (0.95) | (0.71) |  |
|  |  |  |  |
| Number of women  | 13.34 | 38.53 | 7.7128 |
| employed | (0.89) | (2.64) |  |
|  |  |  |  |
| Number of children | 1.57 | 2.77 | 4.2774 |
| employed | (0.11) | (0.22) |  |
|  |  |  |  |
| Women employed /  | 0.13 | 0.12 | 1.2674 |
| total workers | (0.0034) | (0.0031) |  |
|  |  |  |  |
| Children employed /  | 0.033 | 0.018 | 8.2865 |
| total workers | (0.0017) | (0.00093) |  |
|  |  |  | Two-Proportion z-test, |z| |
| Urban | 0.59 | 0.53 | 5.4078 |
|  | (0.0084) | (0.0071) |  |

*Notes*: Total Machine Power measures the total amount of horsepower in a firm. Standard errors in parentheses. *Source*: Ministry of Finance, *List of Factories and Plants* (1897).

|  |  |
| --- | --- |
| Table A12. Robustness check with sample restrictions. |  |
|  |  |
| Dependent Variable | *Log* (Working Days) | *Log*(Rev/Worker) |
| Model | *Tobit* | *Tobit* | *Tobit* | *Tobit* | *Tobit* | *OLS* |
| Sample Restrictions: | European Russia | Non-European Russia | European Russia | Non-European Russia | European Russia, >15 workers | European Russia |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
|  |  |  |  |  |  |  |
| *Log* (total machine  | 0.021\*\*\* | 0.038 | 0.001 | 0.081\*\*\* | 0.012\*\*\* |  |
| power + 1) | (0.002) | (0.025) | (0.003) | (0.032) | (0.003) |  |
| *Log* workers |  |  | 0.062\*\*\* | -0.093\*\* | -0.006\*\* |  |
|  |  |  | (0.004) | (0.041) | (0.006) |  |
| City  |  |  | 0.159\*\*\* | 0.038 | 0.157\*\*\* | 0.236\*\*\* |
|  |  |  | (0.010) | (0.081) | (0.011) | (0.064) |
| Age /100 |  |  | 0.227\*\*\* | 1.475\*\*\* | 0.248\*\*\* | 0.088 |
|  |  |  | (0.039) | (0.504) | (0.041) | (0.106) |
| Age squared /1000 |  |  | -0.008\*\* | -0.057 | -0.008\*\*\* | 0.002 |
|  |  |  | (0.003) | (0.050) | (0.003) | (0.007) |
| *N.* of factories |  |  | 0.065\*\*\* | -2.180\*\*\* | 0.025\*\* |  |
| in dist-ind /100 |  |  | (0.011) | (0.305) | (0.011) |  |
| *Log* working days |  |  |  |  |  | 0.698\*\*\* |
|  |  |  |  |  |  | (0.065) |
| Constant | 5.235\*\*\* | 4.957\*\*\* | 5.018\*\*\* | 5.097\*\*\* | 5.229\*\*\* | 3.490\*\*\* |
|  | (0.006) | (0.039) | (0.035) | (0.139) | (0.040) | (0.426) |
| Observations | 16,487 | 635 | 13,786 | 423 | 7,497 | 12,098 |
| Industry controls | NO | NO | YES | YES | YES | YES |
| Regional controls | NO | NO | YES | YES | YES | YES |
| R2 or pseudo R2 | 0.003 | 0.001 | 0.115 | 0.137 | 0.169 | 0.319 |

*Notes*: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Columns 1 and 2 replicate Table 3 , Column 1, but with factories located in only European or non-European Russia; Column 3 and 4 do the same for Table 3, Column 2. Columns 5 and 6 restrict the sample to factories with 15 or greater employees. Columns 7 and 8 replicate Table 4, Column 1. In all Tobit models, the regressions are right censored at 6, because log(365) is approximately equal to 6. Standard errors in parentheses. *Source*: Ministry of Finance, *List of Factories and Plants* (1897).

Figure A1: Map of Average Working Days by Province


*Sources*: Ministry of Finance, *List of Factories and Plants* (1897). Map of Imperial Russian provinces provided by Andre Zerger.

Figure A2: Map of average soil fertility by Province.



*Sources*: Dower et al. (2018). Map of Imperial Russian provinces provided by Andre Zerger.

Notes: Original soil fertility data is on Uezd level. We weight soil fertility by Uezd area (1905 data) to calculate average Province soil fertility.