Management, Productivity, and Technology Choices: Evidence from U.S. Mining Schools

Michael Rubens*

February 10, 2022

Abstract

A key difference between managers and other production inputs is that managers choose the other inputs. Modeling management as a Hicks-neutral productivity shifter, which is a common practice, omits the productivity returns from these input decisions. I illustrate this through a historical episode in which technology choices were important and managers plausibly influenced those choices. I study the entry of the first mining college graduates into coal mine management positions in Pennsylvania. Whereas the Hicks-neutral productivity effect of these managers was negative and not significantly different from zero, their indirect productivity effect through electrical locomotive adoption was 3% on average.

Keywords: Management, Productivity, Technology Adoption, Higher Education

JEL Codes: L11, J24, N52

---

*KU Leuven, michael.rubens@outlook.com.

A previous version of this article was titled 'Managerial knowledge and technology choices: Evidence from U.S. mining schools'. I thank Jo Van Biesebroeck, Jan De Loecker, Frank Verboven, John Asker, Otto Toivanen, Chad Syverson, Penny Goldberg, Naomi Lamoreaux, Mike Andrews, and participants at various workshops and conferences. I thank the editor, Allan Collard-Wexler, and three anonymous referees for their helpful comments and suggestions. Financial support from the Research Foundation Flanders (grant 11A4618N), the ERC Consolidator grant 816638 MPOWER, the Belgian American Educational Foundation and the Fulbright Commission is gratefully acknowledged. Any errors are my own.

1 Introduction

Managers are crucial drivers of firm performance. The standard practice in the literature on management and productivity is to consider management as a Hicks-neutral productivity shifter (Bloom and Van Reenen, 2007; Bloom et al., 2017). In this article, I argue that if we do standard productivity estimation with management as a Hicks-neutral productivity shifter, we miss any effect of managers on productivity via input choices. This matters because managerial duties often include the selection of input bundles, for instance through technology choices.

I illustrate this through a historical episode in which technology choices were important and managers plausibly influenced those choices: the entry of the first mining college graduates into coal mine management positions in Pennsylvania between 1900 and 1914. This case study combines three interesting features. First, the arrival of electricity during this time period fundamentally changed mining technology. Second, the introduction and spread of technical higher education programs in the U.S.A. provides an important shock to managerial characteristics. Anecdotal evidence suggests that mining college-educated managers played an important role in the selection of new electrical technologies. Third, anthracite coal is a nearly homogeneous product, and the Pennsylvanian anthracite mining industry was unconcentrated at the time. This offers a stylized setting in which firm performance is determined by physical productivity, which simplifies the empirical exercise. The broader point made is, however, relevant beyond this case study. In any setting where managers select input bundles, it is crucial to take these decisions into account when quantifying the effects of managers on firm performance and productivity.

I assemble a novel mine-level data set that tracks output and input quantities, the educational background of managers, and detailed technology choices of all anthracite mines in Pennsylvania. I use this data set to examine how managers with mining college engineering degrees differ from other managers both in terms of their technology choices and in the total factor productivity (TFP) of their mines. I start with the technology choices: I focus on the decision of how to haul coal to the surface, which is a crucial determinant of mine productivity. There are four alternative technologies: (i) mules, the traditional technique, (ii) steam locomotives, (iii) electrical locomotives, and (iv) compressed air locomotives. I find that mining college graduates are 50% more likely to use electrical locomotives compared to other managers, but not more likely to use other new or established locomotive types. Using an event study methodology that exploits the panel structure of the data, I find that the arrival of mining college graduates is followed by increased electrical locomotive adoption, rather than the other way around. This is also consistent with historical anecdotal evidence of mining engineering graduates being instrumental in choosing coal haulage technologies.

Second, I use a production and cost model to estimate the Hicks-neutral effects of mining college graduates on total factor productivity, taking into account the endogeneity of input choices. Mines that hire mining college graduates are likely to be different in unobserved ways. I find that mining college graduates are not significantly different from other managers in terms of their Hicks-neutral productivity effects. However, the production function estimates also reveal that electrical locomotives have a higher marginal product compared to the other new haulage technology, compressed air locomotives, and these are being held fixed when calculating the Hicks-neutral productivity effects of mining college graduates. I find that by selecting better locomotives, mining college graduates increased productivity by 3.0%. This comes in addition to any Hicks-neutral productivity effects these managers might have. Based on the data, I cannot tell with sufficient certainty whether mining college graduates increased mining productivity or not, but can tell that the Hicks-neutral productivity estimates are an underestimate of their total effect on productivity. This holds more in general: as soon as choice variables of managers, such as technology choices, enter the production function as inputs, the Hicks-neutral productivity effects of managers do not capture the productivity returns of managers that come from better input decisions.

Finally, I explore the underlying reasons of why mining engineering graduates differ from other managers in terms of their electrical locomotive choices. I distinguish three mechanisms. First, the productivity-increasing effects of these locomotives could be higher when being operated by mining college graduates. Second, it could be that mining college graduates operate or acquire electrical locomotives at a lower cost. Third, mining college graduates may have superior information about the benefits and costs of electrical locomotives prior to adoption. I find suggestive evidence against the informational differences channel. I propose a test to distinguish between the differential benefits and costs channels, but limited statistical power prevents me from being conclusive on this front.

This article contributes to three sets of literature. First, it builds on the management and productivity literature, which usually models the effects of managers on firm performance through their effects on Hicks-neutral productivity (Bloom and Van Reenen, 2007; Bloom et al., 2013, 2017; Adhvaryu et al., 2199). I contribute to this literature by showing that better input decisions by managers are an additional, and potentially large, channel through which they affect productivity, in addition to their Hicks-neutral productivity effects.

Second, I contribute to the literature on higher education and productivity. Most of the empirical work in this literature estimates the effects of the educational background of managers on productivity, profitability, or management practices (Bertrand and Schoar, 2003; Moretti, 2004; Braguinsky et al., 2015; Braguinsky and Hounshell, 2016). I contribute by showing that technically educated managers make better technology adoption decisions compared to other managers. In contrast to related research by Toivanen and Väänänen (2016), Andrews (2021), and Bianchi and Giorcelli (2020), I focus on the adoption of new technologies, rather on their invention. Although there is a literature that studies the relationship between higher education and technology adoption (Wozniak, 1984, 1987; Lleras-Muney and Lichtenberg, 2005), these articles do not focus on managers.

Third, I contribute to a literature that studies the historical role of engineers in fueling technological change and economic growth. Maloney and Valencia Caicedo (Forthcoming) find that the historical presence of engineers in the U.S.A. is associated with higher patenting activity and income today. Hanlon (2020) finds that engineers were central actors in invention during the British industrial revolution.[1] An important difference with these articles is that I simultaneously observe plant-level production data, the presence of engineers in management positions, and technology choices, which allows investigating the short-term effects of engineer-managers. The resulting evidence paints a nuanced view of the role played by engineers. On the one hand, engineers' short-term productivity effects seem rather limited, and they may even have been worse managers than others based on their Hicks-neutral productivity records. On the other hand, engineers were instrumental in selecting better technologies, which on their turn fueled productivity growth, which is consistent with historical accounts of the entry of engineers in extractive industries (Spence, 1970; Hovis and Mouat, 1996).

The remainder of this article is structured in three parts. In section 2, I discuss the historical background on management and technological change in the early 20th century Pennsylvania coal mining industry. In section 3, I examine how managers with mining engineering degrees differ from other managers in terms of their technology choices and in terms of the total factor productivity of the mines under their supervision. Finally, in section 4, I search for the underlying mechanism of why college-educated engineers differ from other managers in terms of their technology choices.

2 Industry background

The case study in this article is about management and productivity in the Pennsylvanian coal mining industry between 1900 and 1914. I examine how a new type of manager, the college-educated mining engineer, affected firm productivity. The purpose of this section is to provide the necessary background to support the empirical model in the next section. I start by discussing the business environment of coal firms in order to conceptualize firm performance in this industry, and technological change in coal mining. Second, I discuss management in this industry, and the arrival of mining college graduates.

[1] In contrast to this article, he does not define engineers based on educational background but based on how inventors refer to themselves.

Business environment

Demand

All mines in the data set produced anthracite coal, which is a nearly homogeneous product. It was mainly used for residential heating purposes, and did not require any form of processing before being used, unlike most other coal types. Coal markets were integrated: 98% of the mines in Pennsylvania were connected to the railroad network, and on average 84% of mine output was transported by train to the final markets. The remaining output was partly sold locally at the mine gate, or re-used as an energy input. Given that there were 322 unique anthracite firms in the sample, and that coal markets were integrated due to railroads, coal markets were unconcentrated. Assuming a state-level coal market, the average firm had a market share of barely 0.9%, and the median firm a share of 0.2%. 95% of firms had a market share below 5%. The largest firm, the Philadelphia & Reading Coal & Iron Co., had a market share of 20% in 1900, which diminished to 15% by the end of the sample. Combined with the absence of product differentiation, this implies that coal markets were very competitive. Hence, firm performance was mainly driven by cost-side efficiency in this industry, rather by market power.

Input markets

Even if coal markets were perfectly competitive, firms could have had monopsony power on labor markets, especially in geographically isolated locations. Boal (1995) found no evidence for labor monopsony power in West Virginian coal mines, but Rubens (2022) does for coal miners in Illinois between 1884 and 1902. Given that mine-level wages and coal prices are not observed in the context of this article, I will assume that both these prices were exogenous from the perspective of individual mines in the baseline model. Especially for managers, the core study object of the article, exogenous wages seem reasonable given the high labor mobility observed in the managerial background data, and the many outside options available to these workers.

Costs

The extraction process of an underground coal mine consisted of three main steps. First, a shaft or tunnel had to be dug to reach the coal seam. Next, coal had to be excavated using either picks and black powder or using cutting machines. Finally, it had to be hauled back to the surface, using mules or underground mining locomotives. Both extraction and haulage were gradually mechanized due to the invention of cutting machines and underground mining locomotives, as will be explained more in detail below. Coal firms' costs differed both because of the cutting and haulage technology used, and because of differences in total factor productivity.

Technological change

Transporting coal to the surface, 'coal haulage', was a crucial part of the coal extraction process. In a mining journal article, Hodges (1905) asserts that *"the problem of getting coal from the working face to the surface in the most economical way is one of the most serious which the mine manager has to solve."* Mules were traditionally used for haulage, but were gradually replaced by underground mining locomotives from the 1880s onward. Three main locomotive types existed: steam-powered, electrical and compressed air locomotives.[2]

Steam locomotives were invented first, and their usage was already widespread by the start of the panel: in 1900, 70% of the mines in the data set operated at least one steam locomotive. They were more efficient compared to mules, but had two main disadvantages. First, using steam locomotives below the surface implied that their exhaust had to be evacuated from the mine, which was costly (Schlesinger, 1890). Second, steam locomotives could cause explosions when mine gas was present (Randolph, 1905). These concerns led to the development of two new locomotive types for underground usage during the 1890s: electrical and compressed-air locomotives. Electrical locomotives were a very successful technology, hauling loads at more than twice the speed than other locomotive types on steep grades, such as inside coal mines, and four times the speed on

[2]Images of these locomotive types are in the Online Appendix.

horizontal tracks (Brewer, 1915). A downside was the risk of electrocution when mines got flooded (Gairns, 1904). Compressed air locomotives were a safe technology, but had to be refilled frequently, resulting in a limited range. For this reason, they were found to be less economical compared to electrical locomotives (Schlesinger, 1890).

Electricity was generated by dynamos, which were usually coal-powered. These dynamos are observed in the data as well, and are almost entirely collinear with the presence of electrical locomotives. In other words, in order to operate electrical locomotives, dynamos had to be installed. The same held for compressed-air engines: air compressors which used steam power had to be installed first. (U.S. Census Bureau, 1895).

Figure 2 depicts the total number of locomotives used of each type in Pennsylvanian anthracite mines. Mines could use several locomotive types concurrently, and 40% did. In 1900, around 2000 steam locomotives were already being used, whereas barely any of the other two types were in use. Up to 1904, the number of electrical and air locomotives grew at similar rates, after which electricity took over to become the standard technology. The share of electrified mines increased from 10% to 60% between 1900 and 1908. Compressed air engines were never used in more than 40% of the mines.

[Figure 2 here]

Besides the mechanization of the coal hauling process, the coal cutting process was mechanized as well, using mechanical cutting machines. Cutting machines were introduced in the U.S.A. in 1882 and spread throughout the last two decades of the 19th century. In contrast, the adoption of electrical underground locomotives mainly started after 1900, so variation in their usage rates is likely to play a more important role driving productivity growth during the time period considered in this article.

Management

I will examine how changes in management affected coal mining productivity, both through technology choices and through Hicks-neutral productivity. To do so, it is important to first define management and discuss the tasks and responsabilities of coal mine managers. Second, I discuss the relevant changes in managerial characteristics observed throughout the sample period.

Tasks and responsabilities

There were three layers of management in coal firms. At the top, there was a 'general manager', often also the firm owner, who was usually based in a city and delegated daily mine management to 'mine superintendents'. In a fourth of all mines, mainly the smaller ones, the general manager would assume the role of superintendent himself. Mine superintendents are the main object of interest in this article. They had a wide range of responsibilities, including technical procurement, human resources management, production line design, financial analysis and cost reporting (Ochs, 1992). Investment decisions were made by firm owners, or boards of directors, but were informed by superintendents. A diary of a superintendent in a Pennsylvania anthracite mine who discusses infrastructure work to deepen the mine shaft mentions:

> *"I made up a set of careful estimates for work to be done in the mines during the winter, [...] , a new barn and alterations at slope. They were approved by the directors, and I was ordered to proceed with the work."*[3]

Both the technical analysis and financial calculations that underlied technology investment decisions were made by the superintendents (Ochs, 1992). As none of the general managers had an engineering background, it is logical that they relied on the technically schooled superintendents when making their technology choices. I present and discuss anecdotal evidence for this delegation

[3]From https://wynninghistory.com/2017/01/19/life-of-a-coal-mine-superintendent1/

of technology choices in section 3. Finally, the lowest level of operational management was carried out by 'foremen', two thirds of whom worked below the surface.

Managerial shock: mining college graduates

I rely on the introduction of college-educated mining engineers into coal management positions as an observable managerial shock. Whereas different continental European nations already had specialized engineering colleges from the early 18th century onward, American universities such as MIT and Columbia only started offering engineering degrees during the 1860s (Lundgreen, 1990). Rapid technological change during the second industrial revolution increased the demand for educated engineers. Lehigh University, which was home to an important mining college in Pennsylvania, phrased its 1872 mission statement as follows:

> *"To introduce branches which have been heretofore more or less neglected in what purports to be a liberal education [...] especially those industrial pursuits which tend to develop the resources of the country."*

Mining engineering was such a 'neglected' branch. The first mining college graduates only started entering managerial positions in the coal mining industry during the early 1900s. The solid line in figure 1 shows that the share of mines with a college-level mining engineer as superintendent increased from none in 1898 to 6% in 1914. All these mining engineers graduated from two local mining colleges: Lafayette College and Lehigh University. The fraction of mines managed by a graduate with other college degrees, the dashed line, grew from 2 to 6% as well. They slowly replaced an older generation of non-educated superintendents who had usually entered the mines around the age of twelve, and worked their way up through the ranks.

The total number of college-level managers was still low. Among the 432 mine superintendents in the data, only 17 obtained at least a college-level undergraduate degree, of which 7 were degrees in mining engineering. None of the general managers or foremen in the data set were educated at mining colleges. Superintendents managed, however, multiple mines simultaneously, and the data set comes at the mine level. Of the 604 mines in the data set, 61 were managed by a mining college graduate at some point, and 17 by a graduate from another college-level degree.

[Figure 1 here]

Why would mining college graduates be different?

Why would college-educated mining engineers have been different in terms of their technology choices or productivity? First, their abilities and knowledge about mining technology was likely different due to their education. The rise of electricity was anticipated by mining colleges: by the year 1900, most mining engineering programs had compulsory courses in electrical engineering and applied electricity in their junior or senior years. The mining colleges attended by the managers in the data set offered courses on coal haulage technologies and electrical engines, which could plausibly have affected their haulage technology choices. More detailed information on the curricula of the mining colleges present in the dataset is in Online Appendix O.3. Secondly, more technically able individuals could have self-selected into mining engineering programs. When discussing differences between mining college graduates and other managers in this article, I will not be able to distinguish this selection effect due to pre-college differences in managers from the treatment effect of mining colleges.

Data

Output and inputs

Mine output and input data are obtained from the Annual *Report of the Bureau of Mines* of the Department of Internal Affairs of Pennsylvania. It includes 604 Pennsylvanian coal mines between 1898-1914. I observe annual coal extraction in tons, and its breakdown into output that is transported over the railroad network, sold locally, and re-used as input. The average number of workers employed is observed, as are intermediate inputs, such as powder kegs and coal. All these variables are measured in quantities. I also observe the town in which the mine was located. A map with mine locations is in the Online Appendix.

Managerial education

Information on the educational background of managers is obtained by matching the manager names in the data set with mining college alumni registries. I define mining engineers based on whether they obtained a mining engineering degree at a college, and will henceforth use the terms 'mining engineer' and 'mining college graduates' interchangeably. I rely on alumni yearbooks and college catalogs which cover approximately 90% of mining engineering graduates in the U.S.A., 99% of mining engineering graduates in non-Western U.S. states, and 100% of mining engineering graduates in Pennsylvania. More information on this matching procedure and data coverage is in Appendix Appendix A.

Technology

I observe the usage of mules and each of the three mining locomotive types from 1900 onward. Restricting the dataset to the period 1900-1914 reduces it to 572 mines and 4,469 observations. A complication is that the number of locomotives of each type is observed at the county-firm-year level, whereas all other variables are observed at the mine-year level. The average county in the data set contained 28 mines, the median county just four. Both attributing locomotives to the mine-level and mine managers to the firm-level requires ad-hoc weights. I choose to bring the entire data set to the mine-year level and assign locomotive usage evenly to all mines in a given county-firm-year pair. As such, it is assumed that upon adopting a locomotive, firms install them in all mines in a given county. Omitting the observations for which output or labor is either zero or unobserved reduces the dataset further to 4,079 observations, which is the sample on which the model is estimated.

Prices and wages

I obtain daily wages and coal prices from the *Annual Report of the Secretary of Internal Affairs of the Commonwealth of Pennsylvania*, for which I observe average miner wages and anthracite prices between 1902 and 1913. I impute wages and prices for the other years by using the historical price deflator from the U.S. Department of Labor.[4] Further details concerning the data sources and data cleaning are in Appendix Appendix A.

3 Empirical analysis

The key point of this article is to distinguish the *direct* productivity effects of managers through Hicks-neutral productivity from their *indirect* effects through the selection of inputs. As soon as choice variables of managers enter the production function as inputs, conditioning on these inputs when estimating the production function fails to capture managers' indirect productivity effects from better input choices.[5] In this section, I illustrate this argument for mining college-educated managers by carrying out an empirical analysis in two steps. I start by examining whether mining college graduates differed from other managers in their locomotive technology choices, without distinguishing the reasons *why* this would be the case; these underlying mechanisms are discussed in Section 4. Second, I examine the Hicks-neutral effects of managers on total factor productivity, keeping haulage technologies and all other input choices fixed.

[4]Accessed through https://www.minneapolisfed.org/about-us/monetary-policy/inflation-calculator/consumer-price-index-1800-

[5]In Online Appendix O.1, I present a more general model of production and management to illustrate this point.

Mining engineers and technology adoption

Environment and notation

I study the haulage technology choices of a manager of a coal mine i in a year t. I assume that managers make decisions on inputs, rather than mine owners, in line with the anecdotal evidence presented earlier. Mining requires some fixed capital inputs $\mathbf{K}_{it} = (K_{it}^{st}, K_{it}^{el}, K_{it}^{air})$, which is a vector composed of dummies indicating the usage of at least one steam, electrical or compressed-air locomotive, denoted $\tau \in \{st, el, ca\}$. Besides hauling coal to the surface, some capital was required to reach the coal seam. These costs are considered as sunk, and hence do not enter the production function. A third of the mines operated without any locomotive, using mules instead, meaning that $\mathbf{K}_{it} = (0, 0, 0)$.

Coal mines produce Q_{it} tons of coal per year t, and require labor inputs L_{it}, and intermediate inputs M_{it}, besides the capital technologies introduced above. Managers can be alumni from a mining college, which is indicated by the dummy ${X^{mc}}_{it} \in \{0, 1\}$, or another type of college, indicated by ${X^{oc}}_{it} \in \{0, 1\}$. Together, these two dummies compose the vector $\mathbf{X}_{it}$. The mine production function is given by $Q_{it} = F(\mathbf{K}_{it}, \mathbf{X}_{it}, L_{it}, M_{it}, \omega_{it}; \boldsymbol{\beta})$. Total factor productivity is denoted ω_{it}, and the production function is parametrized by $\boldsymbol{\beta}$.

Coal is sold at a price P_t, labor and intermediate inputs are purchased at prices W_t^L, W_t^M, and managers at wages $\mathbf{W}_t^{\mathbf{X}}$. In the baseline model, I assume that there is a single market for these inputs in Pennsylvania, and that wages are the same for all mines in a given year.[6] The three locomotive types have mine-level prices $\mathbf{W}_{it}^{\mathbf{K}}$, which are all latent but are allowed to differ across mines and time. All these prices are assumed to be exogenous from each mine's point of view.

[6] In Online Appendix O.2, I relax this assumption.

Choosing locomotives

The key technology choice faced by managers was whether to replace underground steam locomotives by either electrical or compressed air locomotives. As was discussed earlier, the engineering literature points to the superiority of electrical engines over compressed-air engines in terms of efficiency, which I also verify empirically in the next section. In order to use compressed-air or electrical locomotives, an air compressor or electricity generator was required, and both these power sources and the locomotives had to be purchased, transported, and installed. Hence, the choice of the locomotive types to be used in year t is assumed to be made at time $t-1$. Given that coal markets are assumed to be perfectly competitive, I assume that managers make their input decisions on a mine-by-mine basis. The relevant production unit for which all decisions are modeled is hence the mine, rather than the firm. I assume that managers choose the combination of locomotive technologies that minimize expected costs at the mine, subject to the constraint that a certain coal output Q_{it} has to be delivered in every time period. All input prices and output elasticities are assumed to be observed by the superintendents. As electrical and compressed-air locomotives were a new technology, their returns may have been private information during at least a part of the panel. I test for informational differences between managers in section 4.

I assume that intermediate inputs and non-managerial labor can be adjusted flexibly every year. Labor markets were very flexible and unregulated at the time (Naidu and Yuchtman, 2018). Data from Illinois on coal labor contracts from the same time period shows that most miners were offered weekly, semi-monthly or monthly contracts, which is in line with the flexible labor assumption (Rubens, 2022) . The main intermediate input was coal itself, which was re-used from the firm's own output, and could therefore be flexibly adjusted. Every year, superintendents choose the level of variable inputs that minimize their current costs, taking the presence of locomotives as given. They choose the combination of locomotives for the next year that minimizes their expected costs, conditional on the desired output level being produced in that year. Denoting marginal costs as λ_{it}, this implies that the manager faces the following choice problem for locomotives:

$$\mathbf{K}_{it+1} = \arg\min \mathbb{E}_t \Big[W^M_{t+1} M_{it+1} - W^L_{t+1} L_{it+1} \\ - \mathbf{W}^{\mathbf{K}}_{it+1} \mathbf{K}_{it+1} - \lambda_{it}((Q_{it+1}) - F(M_{it+1}, L_{it+1}, \mathbf{K}_{it+1}, \mathbf{X}_{it+1}, \omega_{it+1})) \Big]$$

I assume that productivity follows a first-order Markov process, Equation (1), with a productivity shock υ_{it} that is i.i.d. across mines and over time.

$$\omega_{it} = g(\omega_{it-1}) + \upsilon_{it} \tag{1}$$

This implies that managers can forecast inputs and productivity in the next period up to the shock υ_{it}. Hence, the investment function for all three locomotive types is given by equation (2). Usage of each locomotive type in year t depends on the variable production inputs, productivity and input prices in year t, and on the productivity shock ν_{it}. The manager's educational background $\mathbf{X}_{it}$ enters the locomotive choice model because I allow managers to make different locomotive choices depending on their education.

$$K^{\tau}_{it} = a^{\tau}(L_{it}, M_{it}, \mathbf{X}_{it}, \mathbf{W}_{it}, \omega_{it}, \nu_{it}) \qquad \text{for } \tau \in \{st, el, air\} \tag{2}$$

Comparing superintendents' capital choices

I implement the locomotive demand function (2) by regressing the usage of each capital type on managerial education, log productivity, log intermediate inputs, and log labor, and include both mine and year fixed effects. Total factor productivity is latent, but is estimated in the next section. The residual u^{τ}_{it} contains all unobserved variation in type-τ locomotive costs and returns. I include mine and year fixed effects because a part of the costs and returns to each locomotive type may be due to locational differences between mines, and may also change over time as mining technology progressed. In this baseline specification, I rule out that locomotive fixed costs $\mathbf{W}^{\mathbf{K}}_{ft}$ are shared across different types, or that their output elasticities depend on each other, which I relax in Appendix Appendix B. The prices of labor, coal, and intermediate inputs are subsumed into the year fixed effects as they are assumed not to vary across mines.

$$K^{\tau}_{it} = \alpha^{\tau}_1 \omega_{it} + \boldsymbol{\alpha}^{\tau}_2 \mathbf{X}_{it} + \alpha^{\tau}_3 l_{it} + \alpha^{\tau}_4 m_{it} + \delta^{\tau}_i + \delta^{\tau}_t + u^{\tau}_{it} \tag{3}$$

The coefficients of interest, $\boldsymbol{\alpha}^{\tau}_2$, measure how the usage of type-τ locomotives differ between mining college graduates, other college graduates, and other managers. In order to interpret these coefficients as causal effects of managers' educational backgrounds on locomotive choices, this educational background needs to be orthogonal to the locomotive return/cost shocks u^{τ}_{it}, which includes the unexpected productivity shock ν_{it}. The opposite causality could hold if mine owners anticipate acquiring a new locomotive, and hire a mining or other college graduate to carry this out. Secondly, it could be that the latent locomotive return/cost shocks affect the utility of a manager to work at mine i. For instance, if mine i experiences a flood, this both increases the cost of installing an electrical locomotive and may decrease the attractiveness of that mine to a manager. It has to be noted, however, that many of these shocks will already feature in the mine productivity residual ω_{it}, which is controlled for. In the next two subsections, I will investigate these potential endogeneity issues. For now, I assume that u^{τ}_{it} is conditionally independent of the presence of college graduates $\mathbf{X}_{it}$.

I estimate equation (3) for each locomotive type using a linear probability model, in which standard errors are clustered at the superintendent level, as managerial characteristics vary at this level. I use a discrete choice model as a comparison in Appendix Appendix B. For the productivity estimates, I rely on the Cobb-Douglas specification of the production model from the next subsection. The estimates are in panel (a) of table 1. Superintendents with mining degrees are 19.9 p.p. more likely to use electrical locomotives than non-educated managers, which is a relative increase of 47.8% compared to the average usage rate of electrical locomotives. The increase in electrical locomotive usage is statistically significant with a 95% confidence interval between 8.5 and 31.4 p.p.

In contrast, mining college graduates are 14.7 p.p. *less* likely to use compressed air locomotives, although this difference is not statistically significant: the air locomotive 95% confidence interval is [-33.7 p.p.; 4.3 p.p.]. Hence, the difference in electrical locomotive usage between mining college graduates and non-educated managers is significantly larger than the difference in compressed air locomotive usage. Mining college graduates are also 4.8 p.p. more likely to use steam locomotives than non-educated managers, with a 95% confidence interval of [-1.8 p.p.; 11.4 p.p.]. The difference in steam usage between managers is not significantly above zero, but is significantly smaller than the difference in electrical engine usage at the 10% significance level.

[Table 1 here]

Event study

Was the arrival of mining college graduates followed by the adoption of electrical locomotives, or was it the other way around? In order to shed more light on this ordering, I conduct an event study that makes use of the panel structure of the data. I estimate how changes in superintendent characteristics in the past and future correlate with changes in the usage of each locomotive type. Denoting the change in the usage of locomotive type τ as $A_{it}^{\tau} \equiv K_{it}^{\tau} - K_{it-1}^{\tau}$ and the change in whether the superintendent has a mining college degree as $A_{it}^{mc} \equiv X_{it}^{mc} - X_{it-1}^{mc}$, I estimate equation (4) using up to three years before and after the arrival of a mining college graduate. I still include year dummies, but not mine dummies as these are differenced out.

$$A_{it}^{\tau} = \sum_{p=-3}^{3} \gamma_p^{\tau} A_{it+p}^{mc} + \tilde{\delta}_t^{\tau} + \tilde{u}_{it}^{\tau} \tag{4}$$

I again use a linear probability model with standard errors clustered at the superintendent level. The estimated coefficients γ_p^{τ} and their 95% confidence intervals are plotted in figure 3. Panels 3(a)

- 3(c) plot the estimates for electrical, steam and compressed-air engines, which are the technology choices of interest. Panel 3(d) does the same, but uses the difference in log labor as the dependent variable of equation (4), by means of a comparison. Panel 3(a) shows that prior to the arrival of a mining college graduate, electrical locomotive adoption was not significantly different compared to mines where no mining college graduate would be hired, and the estimate of this difference is around 5p.p. The arrival of a mining college graduate was, however, associated with an increase in electrical locomotive adoption of 40 p.p. one year after the arrival of the mining college graduate, and the 90% confidence interval lies between 5 and 65 p.p. The absence of a pre-trend in electricity adoption prior to hiring the mining school superintendents shows that the arrival of mining college graduates was followed by a change in locomotive usage, rather than being preceded by it.

Panel 3(b) shows the same estimates for compressed-air locomotives, the other new technology: adoption was similar three years before a mining college graduate arrived, but lower two and one year before the mining college graduate arrived. Mines with lower air adoption were hence less likely to hire a mining college graduate. The adoption of compressed air engines increased after mining college graduates arrived, just as for electrical engines. Steam locomotive adoption, in panel (c), was not significantly different between mines that hired and did not hire mining college graduates before the arrival of these graduates. There seems to be an increase in steam adoption after their arrival, but this is not significantly different from zero in every year, except for the third year, in which it is borderline significantly positive. Standard errors around this event study are wide because there is little adoption of steam engines. Labor usage, in panel and 3(d), did not change significantly after mining college graduates were hired.

[Figure 3 here]

Anecdotal evidence supporting the technology choice mechanism

The finding that the entry of mining college graduates was followed by increased uptake of electrical locomotives could mean that mining college graduates decided to reorganize the production process and choose better haulage technologies after they arrived, but it could equally be the case that mine owners anticipated the adoption of electrical haulage technology, and hired mining college graduates for this purpose. It is hard to distinguish these two mechanisms solely using the available data, so I use anecdotal evidence from the arrival of mining engineers in mine management positions to shed more light on this issue. In his discussion of electrical mining locomotive adoption in Illinois, mining engineer Leonard V. Newton mentions in 1913 that “the engineering staff of the Madison Coal Corporation is planning on the installation of tandem locomotives [the parallel usage of two locomotive types, in this case steam and electrical locomotives], their claim being that this will avoid enlarging their track gauge" (Newton, 1913). This suggests that the initiative of locomotive adoption lied with the mining engineers. In the context of a different mining industry, copper mining, Spence (1970) argues the following about mining engineers, based on historical anecdotal evidence:

> “The next step was to convince the mine promoters and miners alike that scientific methods were so superior to empirical development that engineers should supervise every step from claim surveying to smelting.” (Spence, 1970)

This again suggests that the initiative for technical change lied with the engineers, not the mine owners. Along the same lines, in his business history of the South Canyon coal mine in the early 1900s, which is based on archival records, Twitty (2017) mentions that “The directors were not experienced in mine development, and so hired experts who were” and that “the manager faced an uphill battle of educating the officers and board, so they had context to understand his regular reports to them”. About the design of the haulage infrastructure in the coal mine, he notes the following:

> “The engineers used 60-pound rails (weight per yard of rail) spaced 3’ apart, and charted a very gentle 3.5 percent climb. [...] they could have run regular narrow-gauge trains on the line, but opted for an electric locomotive [...]"

All of this evidence is in line with managers/mining engineers taking the initiative to adopt new technologies, such as electrical locomotives, rather than the mine owners. It is unsurprising that the largest differences between managers applied to electrical locomotives: electricity was a new technology, and electricity experts were rare. In the *Transactions of the Institute of Mining Engineers* from 1905, it is noted that:

> *"The machinery used with compressed air so closely resembles that used with steam, that mechanics familiar with the one have little to learn in managing the other. [...] men competent to manage pneumatic plants are easily obtained, while experts in electricity are scarce."* (Randolph, 1905)

Mining engineers and Hicks-neutral productivity

The previous section established that mining engineers were more likely to use and adopt electrical locomotives compared to other managers. In this section, I examine whether mining engineers also increased total factor productivity, keeping technology choices and all other input quantities fixed. This is the Hicks-neutral productivity effect of management, which does not take into account different technology choices by managers.

Production model

In order to study the productivity effects of managers, a functional form for the production function $Q_{it} = F(\mathbf{K}_{it}, \mathbf{X}_{it}, L_{it}, M_{it}, \omega_{it}; \boldsymbol{\beta})$ needs to be specified. I start by studying the Hicks-neutral productivity effects of mining college graduates by using a Cobb-Douglas production function, which in logarithms is given by Equation (5). The output elasticities of all inputs $\boldsymbol{\beta}$ are assumed to be constant across firms and over time. The coefficient $\boldsymbol{\beta}^X = (\beta^{mc}, \beta^{oc})$ contains the output elasticities of mining college and other college graduates. The productivity residual is assumed to be a scalar and is denoted ω_{it}. Given that the production function has the Cobb-Douglas form, the

Hicks-neutral productivity effect of any input is given by the corresponding output elasticity. For instance, the output elasticity of mining college graduates, β^{mc}, measures the effect of having a mining college graduate manager on output, keeping all other inputs fixed.

$$q_{it} = \beta^l l_{it} + \beta^m m_{it} + \boldsymbol{\beta}^k \mathbf{K}_{it} + \boldsymbol{\beta}^x \mathbf{X}_{it} + \omega_{it} + \beta^t t + \varepsilon_{it} \tag{5}$$

Besides having a Hicks-neutral effect on productivity, mining college graduates could also affect productivity by changing the output elasticities of the other inputs, which is ruled out using the Cobb-Douglas specification. I will allow for such directed effects of management in the mechanisms discussion in Section 4. I also rule out that locomotives have directed effects, which I relax in Online Appendix O.2. I allow for measurement error in output, which is denoted ε_{it}. I add a linear time trend, captured by the coefficient β^t, in order to allow for changes in productivity over time due to reasons outside of the model. Labor and intermediate inputs are continuous variables and measured in quantities: labor is measured by the average number of workers during the year, and intermediate inputs by the tons of coal used. Black powder was also used as an intermediate input, but it was purchased and brought by the miners, rather than by the firms, and is therefore assumed to be a perfect complement to workers, and excluded from the firms' cost function.

Input choices: timing assumptions

In order to estimate the Hicks-neutral productivity effects of managers, the production function needs to be identified. I follow the literature on production function estimation by imposing timing assumptions on the input choices (Olley and Pakes, 1996; Ackerberg et al., 2015). It was already assumed in section 3 that labor and intermediate inputs are variable and static inputs, meaning that they can be flexibly adjusted in a every time period and only affect current profits. Locomotives are, in contrast, assumed to be fixed inputs, which have to be chosen one year ahead. Variable input and technology choices are made by mine managers. The objective functions of the managers and mine owners are hence assumed to be aligned, which is in line with historical evidence (Hovis and

Mouat, 1996).

Managers are, in contrast, selected by the mine owners. I allow for adjustment costs in managerial selection: given that mine superintendents, and especially mining college graduates, were scarce, there was probably some search friction on the market for managers. Managers earn a wage $\mathbf{W}_t^{\mathbf{X}}$ that depends on their educational background. Each year, the owner of mine i chooses a mine superintendent with educational background $\mathbf{X}_{it}$ for the next year, by minimizing the expected cost of each mine. Online Appendix O.1 contains a more detailed model of how managers choose which mine to join.

Output elasticities of the variable inputs

I start with the identification of the output elasticities of the variable inputs. Given that both input and goods prices are assumed to be exogenous to firms, and using cost minimization, it can be shown that the output elasticities of labor and intermediate inputs are equal to their revenue shares (Foster et al., 2008). In order to implement such an approach, sales and input expenditure needs to be observed, but I only observe output and input *quantities*. However, under the assumption that coal prices and input prices are homogeneous across firms in each given year t, the revenue shares can be calculated using annual average wages and coal prices in the Pennsylvanian anthracite industry.[7] The revenue shares of labor and materials are given by $\frac{W_t^L L_{it}}{P_t Q_{it}}$ and $\frac{W_t^M M_{it}}{P_t Q_{it}}$. I multiply the daily wages with the average number of days worked in the mines to get an annual wage bill, and multiply the coal price with the amount of coal consumed to get intermediate input costs. I estimate the output elasticity of labor and materials as the median revenue share of each input, as in Collard-Wexler and De Loecker (2016).

$$\begin{cases} \hat{\beta}^L & = \text{med}[\frac{W_t^L L_{it}}{P_t Q_{it}}] \\ \hat{\beta}^M & = \text{med}[\frac{W_t^M M_{it}}{P_t Q_{it}}] \end{cases}$$

[7] In contrast, in Online Appendix O.2, I use an approach that does not impose homogeneous input and output prices.

I block-bootstrap the standard errors within mine blocks with 250 draws to get standard errors around these median revenue shares. The results are in Table 2(a): the output elasticities of labor and materials are 0.677 and 0.090, respectively.

[Table 2 here]

Output elasticities of the fixed inputs

Second, I discuss the identification of the fixed inputs' output elasticities. This is crucial, because the main objective of this section is to estimate the Hicks-neutral productivity effects of mining college graduates. In contrast to the variable inputs, the revenue share approach cannot be used to recover the output elasticities of managers and locomotives. Simply regressing output on mining college graduates and locomotives would require the assumption that locomotives and managers are randomly assigned across mines independently of productivity levels, which is a strong assumption. In order to identify the output elasticities of the fixed inputs, I therefore rely on the timing assumptions made above. I impose a linear specification for the productivity transition from Equation (1), which results in Equation (6), with serial correlation ρ and unexpected random productivity shocks υ_{it}. This is in the spirit of Blundell and Bond (2000) and Shenoy (2021). In Online Appendix O.2, I allow for a more flexible first-order Markov process.

$$\omega_{it} = \rho\omega_{it-1} + \upsilon_{it} \tag{6}$$

Taking ρ differences of the production function results in the moment conditions in (7), which are used to estimate the coefficient vector $(\boldsymbol{\beta}^k, \boldsymbol{\beta}^x, \beta^t, \rho)$. In line with the earlier made timing assumptions, the key identifying assumption is that mining engineers and locomotives are chosen prior to the arrival of the transient productivity shock υ_{it} and the change in measurement error ε_{it}, whereas labor and intermediate inputs are chosen afterwards. I use the lagged instruments up to only one time lag.

$$\mathbb{E}\Big[\upsilon_{it} + \varepsilon_{it} - \rho\varepsilon_{it-1} | \left\{\begin{matrix} \mathbf{K}_{it} \\ \mathbf{X}_{it} \\ l_{it-1} \\ m_{it-1} \\ t \end{matrix}\right\}\Big] = 0 \tag{7}$$

The main advantage of this production function identification strategy over 'control function' approaches that rely on inverting input demand, as in Ackerberg et al. (2015), is that the latter either requires that locomotive and manager prices are observed, which they are not, or that they are serially uncorrelated, which is unlikely. The auto-regressive productivity model does not require these assumptions. A drawback of Equation (6) is that it does not allow for productivity to depend on cumulative past output, which would be the case if learning by doing or increasing marginal costs with mine depth are important. I discuss these types of productivity dynamics in Appendix Appendix B.

Results

The estimated output elasticities of the fixed inputs using the auto-regressive estimator are in column (III) of Table 2(b). The coefficient on mining college graduates is -0.046, which corresponds to an output elasticity of -4.5%, but is not significantly different from zero. Although mining college graduates are estimated to perform slightly worse than others in terms of Hicks-neutral productivity, we cannot rule out that they had a small positive Hicks-neutral productivity effect. The coefficient on managers who graduated from college with a different degree than mining engineering is -0.015 with a confidence interval [-0.244;0.214]. Hence, the data do not allow to say much about the difference in productivity between these types of managers and non-educated managers.

The estimated locomotive coefficients imply that electrical locomotives increase productivity by 17.5%, compared to 21.3% for steam locomotives and 8.1% for air locomotives. The output elasticities of electrical and steam locomotives were not significantly larger than the output elasticity of compressed air locomotives. As was discussed before, the engineering literature found that electrical engines were more efficient compared to compressed air engines, which is consistent with the magnitudes of the production coefficients, even if a lack of statistical power prevents these to be distinguished from each other at a high probability. Also, if compressed air locomotives were cheaper than electrical engines, they could have been the more profitable technology even if having a smaller effect on productivity. However, this was not the case: Gairns (1904) presents cost estimates for installing a compressed air locomotive and electrical locomotive in the same mine, and reports a cost of $7,062 and $6,687, respectively.

[Table 2 here]

For comparison reasons only, I also included the estimates when simply regressing the productivity residual $\hat{\omega}_{it} \equiv q_{it} - \hat{\beta}^L l_{it} - \hat{\beta}^M m_{it}$ on locomotive and manager dummies and a linear time trend using OLS, in column (I), and when also including mine fixed effects, in column (II). These estimates point to small output elasticities of mining college graduates: in the OLS specification, the mining college coefficient is estimated to be -0.041, which corresponds to an output elasticity of -4.0%, and it lies significantly below zero. In the model with mine fixed effects, the mining college coefficient is -0.116 and again significantly negative, which implies that mining college graduates would *decrease* Hicks-neutral productivity, at least when not taking into account their different technology choices.

Implications for the returns to mining college graduates

Using the estimates from the previous two sections, I now compare mining college graduates' direct, Hicks-neutral, productivity effects with their indirect productivity effects through locomotive choices. The total productivity effect of mining college graduates is in equation (8). The Hicks-neutral effect of mining college graduates on productivity is equal to $\exp(\beta^{mc})$, but this keeps all other inputs fixed, and hence does not take into account different locomotive choices by mining college graduates. In order to estimate the total productivity effect of mining college graduates, the output elasticity of electrical locomotives $\exp(\beta^{el})$ hence needs to be added to the Hicks-neutral productivity effect, weighted by the difference in electrical locomotive choice probabilities between mining college graduates and other managers:

$$
\begin{aligned}
\frac{\mathbb{E}(Q_{it}|X_{it}^{mc}=1)}{\mathbb{E}(Q_{it}|X_{it}^{mc}=0)} = 1 + \frac{\partial Q_{it}}{\partial X_{it}^{mc}}\frac{X_{it}^{mc}}{Q_{it}} \\
+ \frac{\mathbb{E}(Q_{it}|K_{it}^{el}=1)\mathbb{E}(K_{it}^{el}=1|X_{it}^{mc}=1)+\mathbb{E}(Q_{it}|K_{it}^{el}=0)\mathbb{E}(K_{it}^{el}=0|X_{it}^{mc}=1)}{\mathbb{E}(Q_{it}|K_{it}^{el}=1)\mathbb{E}(K_{it}^{el}=1|X_{it}^{mc}=0)+\mathbb{E}(Q_{it}|K_{it}^{el}=0)\mathbb{E}(K_{it}^{el}=0|X_{it}^{mc}=0)} \\
\approx \underbrace{\exp(\beta^{X})}_{\text{Direct effect}} + \underbrace{\frac{\exp(\beta^{el})\mathbb{E}(K_{it}^{el}=1|X_{it}^{mc}=1)+\mathbb{E}(K_{it}^{el}=0|X_{it}^{mc}=1)}{\exp(\beta^{el})\mathbb{E}(K_{it}^{el}=1|X_{it}^{mc}=0)+\mathbb{E}(K_{it}^{el}=0|X_{it}^{mc}=0)}}_{\text{Indirect effect}}
\end{aligned}
\tag{8}
$$

Using the imputed choice probabilities for electrical locomotives from the estimated equation (3), the 'indirect' productivity effect from employing a mining college graduate is now 3.0%, which comes in addition to any Hicks-neutral productivity effect mining college graduates might have. In comparison, mine output grew by 2.0% per year on average, and output per worker by 1.8% per year. If the point estimate of the output elasticity of mining college graduates in column (III) of Table 2 is correct, this would imply that the productivity effect of mining college graduates was still negative, at -1.5%, but three times less negative than the -4.5% when only considering Hicks-neutral productivity effects. In the scenario that corresponds to the upper bound of the 95% confidence interval on the mining college graduate output elasticity, these managers have a Hicks-neutral productivity return of 3.9%, which is 43% lower than their total productivity return of 6.9%. Due to the limited variation in mining college graduates in the data, there is insufficient statistical power to be conclusive on whether they increased or decreased productivity. The main point is that their returns would be underestimated when merely looking at their Hicks-neutral productivity effects without considering their different technology choices.

Settings with unobserved technology

The distinction between Hicks-neutral productivity effects and input choices depends on the level of detail in which input choices are observed in the data. In the context of this article, specific locomotive technology types (K^{el}, K^{st}, K^{ca}) are observed, and included in the production function. I now discuss how less detailed technology data would affect the analysis. A first possibility is that detailed technology types, in this case locomotive types, are be unobserved, but that their value would be included in a scalar capital stock $K \in \mathbb{R}^{+}$. Assume that mining college graduates are more likely to use electrical locomotives, which have a higher output elasticity than other locomotive types. Hiring a mining college graduate would then change the production function from $Q = K^{\beta^K} L^{\beta^L} M^{\beta^M} X^{\beta^X} \Omega$ to $Q = K^{\theta^K} L^{\beta^L} M^{\beta^M} X^{\beta^X} \Omega$, with $\beta^K < \theta^K$, because electrical locomotives have a higher output elasticity. In such a setting, it would seem as if mining college graduates have a capital-augmenting effect, although they simply change the unobserved mix of technology types that compose the capital stock. This type of mechanism features, for instance, in Van Biesebroeck (2003), which studies technological change in the automotive industry.

Secondly, consider a setting in which locomotives would not be included in the capital stock K at all, with a production function $Q = L^{\beta^L} M^{\beta^M} K^{\beta^K} X^{\beta^X} \Omega$. Different locomotive choices by mining college graduates would then be measured as a Hicks-neutral productivity effect of mining college graduates through the output elasticity β^X, because mining locomotives would be part of the Hicks-neutral productivity residual Ω. Hence, the distinction between Hicks-neutral productivity effects of managers and their productivity returns through input choices only exists when these different input choices enter the production function in some observable way. This can be either as separate inputs, as is the case in this article, or as a component of the total capital stock, as discussed in the previous paragraph.

4 Why did mining engineers select different technologies?

Managers with mining engineering degrees may have made different locomotive choices compared to other managers for three reasons. They may have had better information about the returns and costs of each locomotive type before purchasing them, if this information was not common knowledge, they may have been able to obtain higher returns from electrical locomotives, or they may have operated electrical locomotives at a lower cost. I now discuss each of these mechanisms more in detail.

Information

There is a large literature that highlights the importance of imperfect information for technology adoption. A frequent hypothesis is that education plays a role in shaping managers' prior beliefs about technology returns or costs (Rosenzweig, 1995). I test for such informational differences by comparing the mining college graduate coefficient in equation (3) between two sets of mines that are likely to have different information about the locomotives. In a first set of mines, locomotives of a given type have not already been adopted in other mines of the same firm that are located in different counties. In a second set of mines, they have already been adopted. Assuming that the information about locomotive returns and costs are shared within a firm, the informational benefit of a mining school graduate would disappear as soon as the same locomotive type has already been used within the firm. The results in tables 3(b)-(c) show that this is not the case. Mining college graduates are 0.083 p.p. more likely to use electrical locomotives for the first time in the firm, but 28 p.p. more likely to use them if already present in the firm. The difference between both coefficients is statistically significant. This is the opposite result as what one would expect under the informational differences mechanism. Moreover, technical studies that demonstrated the benefits of electrical locomotives were widely published and disseminated by the start of the panel in 1900, for instance in Schlesinger (1890), which goes against the private information hypothesis.

[Table 3 here.]

Locomotive returns

Secondly, it is possible that mining engineers were able to increase the returns from electrical mining locomotives, as measured by the output elasticity of this locomotive type. This hypothesis can be tested by estimating an interaction term between the mining college dummy and the electrical locomotive dummy, β^{xke}. This interaction effect quantifies how the output elasticity of electrical locomotives differs between mines managed by mining college graduates and those managed by other managers.

$$q_{it} = \beta^l l_{it} + \beta^m m_{it} + \boldsymbol{\beta}^k \mathbf{K}_{it} + \boldsymbol{\beta}^x \mathbf{X}_{it} + \beta^{xke} K_{it}^{el} X_{it}^{mc} + \omega_{it} + \beta^t t + \varepsilon_{it} \tag{9}$$

I use the same identification approach as above, but now add the product of the electrical locomotive and mining college dummy, and its lagged value, as instruments in Equation (7). Panel 3(a) reports that the estimate of β^{xke} is negative, which suggests that mining college graduates did not alter the returns from electrical locomotives. However, this interaction term is very imprecisely estimated. There is hence a lack of power to conclude with certainty that different locomotive returns across managers was not the main driver of the differences in locomotive usage.

Locomotive costs

Finally, it could be that the costs to procure and/or operate electrical locomotives was lower for mining college graduates, which enters through the term $\mathbf{W}_{it}^{\mathbf{K}}$. The maintenance costs of electrical engines could, for instance, be lower in the presence of a mining engineer. Locomotive costs can also be non-monetary, such as search costs. It is hard to test for differences in these fixed costs between different managers, but if locomotive returns and information would not differ between managers, the difference in choices should come from the cost side. For variable locomotive costs, there is a direct test: if mining college graduates would mainly change the variable costs of locomotives, we would expect mining college graduates to use more electrical locomotives. I examine this intensive margin by re-estimating equation (3) with the log number of locomotives as the left-hand side variable, which restricts the sample to the mines in which at least one locomotive of type τ was used. The estimates are in panel 1(b). When using at least one electrical locomotive, mining engineers do not use significantly more electrical locomotives than the other superintendents, the coefficient is -0.076. However, this coefficient is very imprecisely estimated, and ranges between -0.38 and 0.23. Insufficient statistical power again prevents to conclude that mining college graduates did not use more electrical locomotives.

5 Conclusion

A key difference between managers and other production inputs is that they make active decisions regarding the bundle of inputs used by firms. In this article, I distinguish the 'direct' effects of managers on total factor productivity through Hicks-neutral productivity from their 'indirect' effects through input choices. I illustrate this distinction using a case study of the introduction of mining college graduates in coal mine management positions in Pennsylvania during the early 20th century. I find that mining college graduates selected better coal haulage technologies than other managers. Solely estimating the Hicks-neutral productivity effects of mining college graduates would lead to missing these indirect productivity gains through better technology choices.

References

Ackerberg, D., Caves, K., & Frazer, G. "Identification properties of recent production function estimators" *Econometrica*, Vol. 83 (2015), pp. 2411–2451.

Adhvaryu, A., Nyshadham, A., & Tamayo, J. A. “Managerial Quality and Productivity Dynamics” *Review of Economic Studies*, (forthcoming).

Aguirregabiria, V. & Luengo, A. “A microeconometric dynamic structural model of copper mining decisions” 2015 , Mimeo.

Andrews, M. “The role of universities in local invention: Evidence from the establishment of US colleges” 2021 , Mimeo.

Asker, J., Collard-Wexler, A., & De Loecker, J. “(Mis) Allocation, Market Power, and Global Oil Extraction” *American Economic Review*, Vol. 109 (2019), pp. 1568–1615.

Benkard, C. L. “Learning and forgetting: the dynamics of aircraft production” *American Economic Review*, Vol. 90 (2000), pp. 1034–1054.

Bertrand, M. & Schoar, A. “Managing with Style: The Effect of Managers on Firm Policies” *Quarterly Journal of Economics*, Vol. 118 (2003), pp. 1169–1208.

Bianchi, N. & Giorcelli, M. “Scientific education and innovation: from technical diplomas to university STEM degrees” *Journal of the European Economic Association*, Vol. 18 (2020), pp. 2608–2648.

Bloom, N., Sadun, R., & Van Reenen, J. “Management as a Technology?” Working Paper 22327, National Bureau of Economic Research, 2017 .

Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., & Roberts, J. “Does management matter? Evidence from India” *Quarterly Journal of Economics*, Vol. 128 (2013), pp. 1–51.

Bloom, N. & Van Reenen, J. “Measuring and Explaining Management Practices Across Firms and Countries” *Quarterly Journal of Economics*, Vol. 122 (2007), pp. 1351–1408.

Blundell, R. & Bond, S. “GMM estimation with persistent panel data: an application to production functions” *Econometric reviews*, Vol. 19 (2000), pp. 321–340.

Boal, W. M. “Testing for employer monopsony in turn-of-the-century coal mining” *RAND Journal of Economics*, Vol. 26 (1995), pp. 519–536.

Braguinsky, S. & Hounshell, D. A. “History and Nanoeconomics in Strategy and Industry Evolution Research: Lessons from the Meji-Era Japanese Cotton Spinning Industry” *Strategic Management Journal*, Vol. 37 (2016), pp. 45–65.

Braguinsky, S., Ohyama, A., Okazaki, T., & Syverson, C. “Acquisitions, productivity and profitability: evidence from the Japanese cotton spinning industry” *American Economic Review*, Vol. 105 (2015), pp. 2086–2119.

Brewer, C. B. “The Passing of the Steam Locomotive” *Scientific American*, Vol. 113 (1915), pp. 488–489.

Collard-Wexler, A. & De Loecker, J. “Production Function Estimation and Capital Measurement Error” Working Paper 22437, National Bureau of Economic Research, 2016 .

Foster, L., Haltiwanger, J., & Syverson, C. “Reallocation, firm turnover, and efficiency: selection on productivity or profitability?” *American Economic Review*, Vol. 98 (2008), pp. 394–425.

Gairns, J. *Industrial Locomotives for Mining, Factory and Allied Uses*, Vol. 26, pp. 291–309: Cassier’s Magazine, 1904 .

Hanlon, W. W. “The Rise of the Engineer: Inventing the Professional Inventor During the Industrial Revolution” 2020 , Mimeo.

Hodges, C. “Compressed air locomotives for mine haulage” *Mining Magazine*, Vol. 11 (1905), pp. 305–310.

Hovis, L. & Mouat, J. “Miners, engineers, and the transformation of work in the western mining industry, 1880-1930” *Technology and Culture*, Vol. 37 (1996), pp. 429–456.

Lleras-Muney, A. & Lichtenberg, F. R. “Are the more educated more likely to use new drugs?” *Annales d’Économie et de Statistique*, pp. 671–696.

Lundgreen, P. "Engineering Education in Europe and the U.S.A., 1750-1930: The Rise to Dominance of school culture and the engineering professions" *Annals of Science*, Vol. 47 (1990), pp. 33–75.

Maloney, W. F. & Valencia Caicedo, F. "Engineering growth: innovative capacity and development in the Americas" *Journal of the European Economic Association*, (forthcoming).

Moretti, E. "Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data" *Journal of Econometrics*, Vol. 121 (2004), pp. 175–212.

Naidu, S. & Yuchtman, N. "Labor Market Institutions in the Gilded Age of American History" *The Oxford Handbook of American Economic History*, Vol. 1 (2018).

Newton, L. V. "Electric Locomotive Haulage in and About Mines" Thesis for degree of Bachelor of Science in Mining Engineering, University of Illinois, 1913 .

Ochs, K. H. "The Rise of American Mining Engineers: A Case Study of the Colorado School of Mines" *Technology and Culture*, Vol. 33 (1992), pp. 278–301.

Olley, S. & Pakes, A. "The dynamics of productivity in the telecommunications equipment industry" *Econometrica*, Vol. 64 (1996), pp. 1263–1297.

Randolph, B. S. "Comparison of Electric and Compressed-Air Locomotives in American Mines" in Brown, M. W. ed. *Transactions of the Institution of Mining Engineers*, Vol. 27: Institution of Mining Engineers, 1905 .

Rosenzweig, M. R. "Why are there returns to schooling?" *American Economic Review*, Vol. 85 (1995), pp. 153–158.

Rubens, M. "Oligopsony Power and Factor-Biased Technology Adoption" 2022 , Mimeo.

Schlesinger, W. M. "The Advantages of Electricity for Mining Purposes" *Ohio Mining Journal*, Vol. 19 (1890), pp. 16–23.

Shenoy, A. "Estimating the Production Function under Input Market Frictions" *The Review of Economics and Statistics*, Vol. 104 (2021), pp. 666–679.

Spence, C. C. *Mining engineers & the American west: the lace-boot brigade, 1849-1933*: New Haven: Yale University Press, 1970 .

Toivanen, O. & Väänänen, L. "Education and invention" *Review of Economics and Statistics*, Vol. 98 (2016), pp. 382–396.

Twitty, E. "History and Summary of South Canyon Coal Mine Historic Site, Garfield County, CO" Report for the City of Glenwood Springs, CO, December, 2017 .

U.S. Census Bureau *Reports of the Inspectors of Coal Mines of Pennsylvania*, Vol. 1895: Clarence M. Busch, 1895 .

Van Biesebroeck, J. "Productivity dynamics with technology choice: An application to automobile assembly" *Review of Economic Studies*, Vol. 70 (2003), pp. 167–198.

Wozniak, G. D. "The adoption of interrelated innovations: A human capital approach" *Review of Economics and Statistics*, Vol. 66 (1984), pp. 70–79.

——— "Human capital, information, and the early adoption of new technology" *Journal of Human Resources*, Vol. 22 (1987), pp. 101–112.

Table 1: Manager education and technology usage

	(I) 1(Elec. loc.)		(II) 1(Air. loc.)		(III) 1(Steam loc.)	
(a) Extensive margin	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
1(Mining col. grad.)	0.199	0.058	-0.147	0.097	0.048	0.034
1(Other grad.)	0.057	0.093	-0.005	0.022	0.068	0.151
Average usage	.468		.303		.754	
Observations	4079		4079		4079	
Within R-squared	.303		.093		.072	

	log(Elec. loc.)		log(Air. loc.)		log(Steam loc.)	
(b) Intensive margin	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
1(Mining col. grad.)	-0.076	0.155	0.098	0.105	0.284	0.129
1(Other grad.)	-0.050	0.143	<0.001	<0.001	0.099	0.062
Observations	1889		1244		3065	
Within R-squared	.664		.66		.18	

Notes: The dependent variables in panel (a) are dummies indicating the usage of each locomotive type in year t, and the logarithm of the number of locomotives of each type in panel (b). By taking logs, panel (b) omits mines which do not use any locomotives of a certain type, which is why the number of observations varies by type. Regressors include a dummy indicating whether the superintendent attended a mining or other college, and the logs of the number of employees, intermediate inputs, and total factor productivity, all measured in year t. Mine and year fixed effects are controlled for as well. Standard errors are clustered at the superintendent level.

Table 2: Production function

	(I)		(II)		(III)	
(a) Variable inputs	Estimate	S.E.				
Labor	0.677	0.011				
Materials	0.090	0.003				
			log(TFP)			
(b) Fixed inputs	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
1(Mining col. grad.)	-0.041	0.039	-0.116	0.041	-0.046	0.043
1(Other grad.)	-0.274	0.102	0.135	0.135	-0.015	0.117
1(Elec. loc.)	0.214	0.023	0.111	0.036	0.161	0.027
1(Air loc.)	0.090	0.023	-0.069	0.042	0.078	0.028
1(Steam loc.)	0.271	0.027	0.133	0.057	0.193	0.034
Model	OLS		FE		AR(1)	
Observations	4079		4079		3429	
R-squared	.344		.453		.332	

Notes: Panel (a) reports the output elasticities of labor and intermediate inputs using the median factor share approach, with standard errors being block-bootstrapped with 250 iterations. Panel (b) reports the output elasticities of managers and locomotives using (I) OLS, (II) mine fixed effects, and (III) the autoregressive productivity GMM estimator. Column (III) has less observations as lagged values of all inputs and output are needed to estimate this specification.

Table 3: Mechanisms

	(I)		(II)		(III)	
(a) Different returns	Estimate	S.E.				
1(M.C. grad)*1(Elec. loc.)	-0.017	0.094				
Observations	3429					

	1(Elec. loc.)		1(Air. loc.)		1(Steam loc.)	
(b) Loc. not yet used	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
1(Mining col. grad.)	0.083	0.057	-0.060	0.024	0.050	0.062
Observations	2019		2760		1848	
Within R-squared	.156		.06		.101	

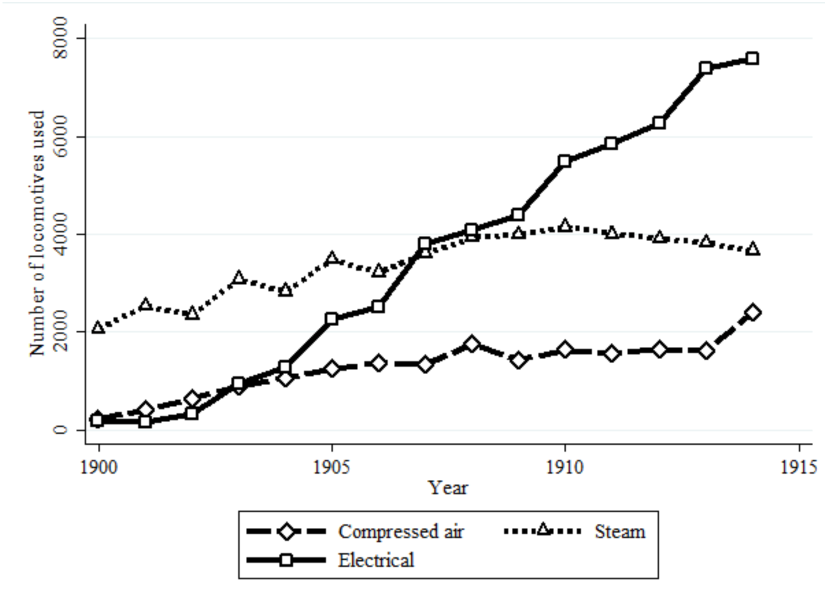
	1(Elec. loc.)		1(Air. loc.)		1(Steam loc.)	
(c) Loc. already used	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
1(Mining col. grad.)	0.280	0.037	-0.232	0.034	0.014	0.025
Observations	2060		1319		2231	
Within R-squared	.423		.183		.069	

Notes: Panel (a) reports the interaction term coefficient between the mining college graduate dummy and electrical locomotive dummy in the production function, using the auto-regressive productivity GMM estimator described in the main text. Panels (b) and (c) report the mining college graduate coefficient if the firm does not already use the same locomotive type in other counties, and if it does. The same controls as in the prior locomotive usage regressions are used. A linear probability model is used with standard errors clustered at superintendent level.

Figure 1: Educational background of mine superintendents

Notes: The solid line plots the share of Pennsylvania anthracite mines with a manager with a college-level mining engineering (E.M.) degree, the dashed line does the same for other college degrees.

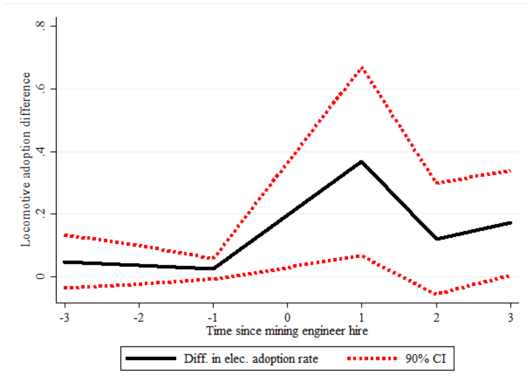
Figure 2: Usage of mining locomotives



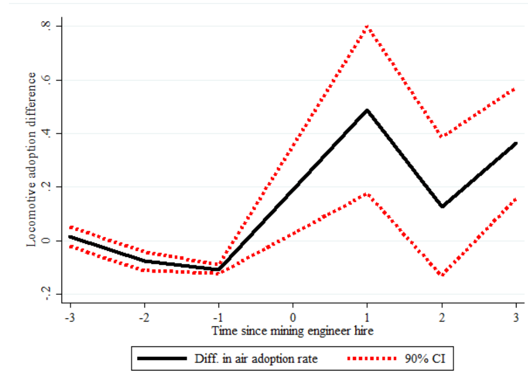
Notes: This graph shows the total number of locomotives of each type used in all anthracite mines of Pennsylvania.

Figure 3: Mining engineer arrival: event study

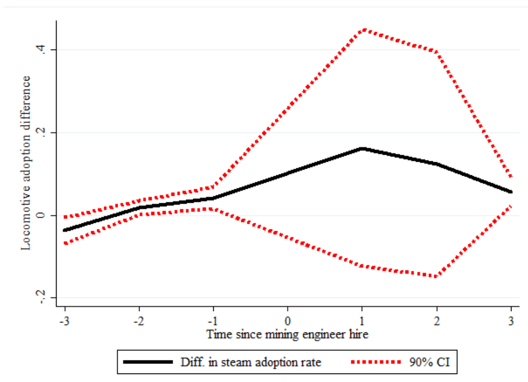
(a) Electrical locomotives



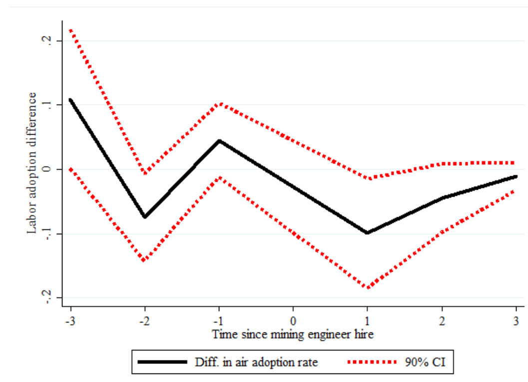
(b) Compressed air locomotives



(c) Steam locomotives



(d) Labor



Notes: The estimated effect of mining college graduate hires at times $t - p$ on the probability of locomotive adoption at time t is plotted for each locomotive type. Confidence intervals at the 95% level are indicated by the dotted red lines.

Appendix

Appendix A Data sources and cleaning

Production and cost data

Data on output, inputs, managers, technical characteristics, ownership and locations of mines were obtained from the *Report of the Bureau of Mines* by the Department of Internal Affairs of Pennsylvania. This data was collected by government-appointed mine inspectors. The data structure is unchanged between 1900 and 1914 and is composed of four tables per county. A first table lists all mines, their owners, the managers, a post office location and the railroad to which it is connected. A second table provides output and input data at the mine level. Technology choices are reported in a third table, at the firm-county-year level. Fourthly, the occupational breakdown of labor is given, again at the firm-county-year level.

I construct unique mine identifiers by tracking mine name changes over time. It happens that mines have multiple sub-units which are reported separately in some years, but not in others. I aggregate these sub-units to the mine-year level in order to have an observation unit that remains stable over time. Locomotive counts of all three types are observed at the county-firm-year level, rather than the mine level. I assign locomotives evenly to all mines belonging to the same firm-county-year pairs in the baseline specification. 4 observations have no county name listed, which I drop because they cannot be linked to the machine data. I also drop 318 observations for which no mine superintendent is listed (no information on superintendents can be linked) and 298 observations for which the same mine-year combinations entered twice in the report tables. I do not observe the mine locations, but observe the town in which the mine superintendent was located. I assume that this is also the location of the mine. I use google maps to link the town names to geographical coordinates. Panel (a) of table A1 shows some summary statistics on the mines.

Annual extraction was on average 0.21 Mton. The average mine had 509 employees.

[Table A1 here]

Management data

Mining engineering graduates

I obtain a list of all mining engineering programs offered in the U.S.A. in 1912 from the *Report of the Commissioner of Education for the Year 1912*, volume 2. This results in a list of 40 mining engineering programs, which are listed in the Online Appendix. Four mining engineering programs were offered in Pennyslvania, at Lehigh University, Lafayette College, Pennsylvania State College, and the University of Pittsburgh. Information on the graduates from these programs can be found from two sources: college alumni registries which include all alumni up to a certain date, and college yearbooks, which include all students in a certain year. I access these alumni registries and catalogs through two sources: *Ancestry.com* and self-collected alumni registries and catalogs. This way, I collect a database with mining college graduates containing their last, middle, and given name. I match the coal mine superintendents in the mining data set with the names in the alumni registry, and correct for false positives by comparing the graduation years, manager birth years, and addresses in the various datasets. This matching procedure results in 17 college graduates in the coal mining data set, of which 7 graduated with a degree in mining engineering, from *Lehigh University* and *Lafayette College*, both located in Pennsylvania. More details on this matching procedure and on the mining college alumni database can be found in Online Appendix O.3.

Summary statistics on the managers

Summary statistics on the managers are in panel (b) of table A1. Superintendents with mining college degrees were on average 31, and those with other college degrees 33, whereas the average non-educated superintendent was 49. Mining college graduates were put in charge of more mines:

10 on average, compared to only 3 for non-educated managers, and 1 for other college graduates. Their mines were, however, slightly smaller in size than those managed by non-educated superintendents. The number of superintendents per firm ranged between 1 and 4, and was on average 1.04. Multi-mine firms employed on average 1.17 superintendents.

Appendix B Robustness checks

Common fixed cost component across locomotive types

It could be that a part of the fixed cost to use a certain locomotive type is also useful for other locomotive types. If this is the case, then the presence of other locomotive types at the mine should decrease the cost of using a given locomotive type. In table A2(a), I add the usage of the other two locomotive types as a regressor to the choice function from equation (3). The results are very similar compared to the specification in the main text: mining college graduates are associated with a significant and large increase in electrical locomotive usage of 19.7 percentage points and an increase in steam locomotive usage by 5.9 percentage points. In contrast, air locomotive usage was 16.1 percentage points lower for mining college graduates. The difference in electricity usage between mining college graduates and others is still hence still significantly above zero and significantly larger than both the compressed air and steam usage differences, at the 5% and 10% confidence level respectively.

Discrete-choice model

When estimating the technology choice model (3), a linear probability model was used. In panel (b) of table A2, I estimate a logit model instead. I report the marginal effects at the mean. The estimates are somewhat higher compared to the baseline model, but the conclusions are similar: mining college graduates are 26.5 p.p. more likely to use electrical locomotives, 13.5 p.p. more likely to use steam locomotives, which are both significantly positive effects. The compressed air coefficient is now positive too, at 6.7 p.p., but not significantly larger than zero. An important difference compared to the baseline model is that mine fixed effects are no longer included in this specification, which could explain the slight difference in estimates. It is likely that time-invariant mine characteristics, such as the geographical location of the mine, are correlated with both locomotive adoption and managerial hires.

Cost dynamics

If cost dynamics are important, which is often the case in extractive industries, then productivity should depend on the total amount of coal extracted in the previous periods, which may also have affected the need for mining locomotives.[8] In order to examine whether cost dynamics affect the estimated effects of mining college graduates on locomotive usage, I re-estimate equation (3) with log cumulative output as an additional regressor. The results are in table A2(c). The mining college coefficients are very similar to those from the baseline model in table 1(a), and the R-squared barely changes.[9] Cumulative output itself has barely any effect on the usage probability of any locomotive type.

[8]This could have been the case for coal mines. Coal that can be reached at the lowest cost is usually mined first. Marginal costs are hence likely to increase as more coal is extracted (Aguirregabiria and Luengo, 2015; Asker et al., 2019). On the other hand, there could have been some 'learning by doing', as in Benkard (2000).

[9]The R-squared is slightly *lower* compared to the model without cumulative output, but the sample size differs as at least one lagged period of output needs to be observed to calculate cumulative past output.

Table A1: Summary statistics

(a) Mines	Mean	Std. Dev.	Min.	Max.	Observations
Coal extracted, Mtons	0.21	0.19	0	3.52	5,029
Output shipped, share	0.84	0.18	0	1	4,868
Employees	509.19	427.13	0	6595	5,029
Powder, 1000 kegs	51.96	137.06	0	1748.43	5,029
Coal inputs, Ktons	20.56	21.56	0	494.48	5,029
Mining locomotives	25.35	35.50	0	172	5,029
Man. has M.C. degree	0.06	0.24	0	1	5,029
Man. has other col. degree	0.01	0.1	0	1	4,393

(b) Managers	Mean	Std. Dev.	Min.	Max.	Observations
Age					
if mining degree	30.77	6.94	21	45	30
if other degree	32.53	5.02	20	43	38
if no degree	48.88	10.96	19	80	1,179
# Mines managed					
if mining degree	9.83	11.77	1	30	30
if other degree	1.21	0.53	1	3	38
if no degree	3.32	6.10	1	41	1,179
Output, Mtons					
if mining degree	0.22	0.16	0.025	0.67	30
if other degree	0.17	0.26	0.04	0.72	38
if no degree	0.14	0.13	0	0.81	1,179

Table A2: Technology choice: robustness

(a) Common fixed costs	1(Elec. loc.)		1(Air. loc.)		1(Steam loc.)	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
1(Mining col. grad.)	0.197	0.060	-0.161	0.092	0.059	0.041
1(Other grad.)	0.042	0.081	-0.019	0.036	0.062	0.145
Observations	4079		4079		4079	
Within R-squared	.329		.143		.149	

(b) Logit model	1(Elec. loc.)		1(Air. loc.)		1(Steam loc.)	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
1(Mining col. grad.)	0.265	0.051	0.067	0.073	0.135	0.049
1(Other grad.)	-0.117	0.167	-0.434	0.109	0.064	0.118
Observations	4079		4079		4079	

Notes: Compared to table 1, panel (a) adds the usage of both other locomotive types as additional control variables. Panel (b) uses a logit model, and reports the marginal effects. Mine fixed effects are excluded from the logit model. Panel (c) adds log cumulative past output as a regressor, which reduces the number of observations as only observations with lagged values are considered.