Managerial knowledge and technology choice: evidence from US mining schools

Michael Rubens*

November 22, 2019

Latest version here

Abstract

How do managers affect firm performance? I propose a model in which managers choose the production function, and compare it to the predominant view of managers as production inputs. I apply this model to understand how the introduction of mining engineering degrees in the US changed management and productivity in the coal mining industry. I find that conditionally on all inputs and technology choices, mines managed by managers with mining degrees were not more productive than other mines. Mining college graduates did, however, tend to select better technologies, which led output to increase by 5.5% more per year compared to mines managed by other managers. The distinctive feature of mining college graduates was that they had ex-ante knowledge about the returns to various new technologies, while other managers had to acquire this information through trial-and-error.

Keywords: Management, Productivity, Technology Adoption, Higher Education
JEL Codes: L25, O14, J24, I23, M51

*KU Leuven and Research Foundation Flanders, michael.rubens@kuleuven.be
I thank Jo Van Biesebroeck, Jan De Loecker, Frank Verboven, Otto Toivanen, Chad Syverson, Penny Goldberg, Naomi Lamoreaux and participants at the KU Leuven IO lunch, Yale IO prospectus workshop, Business History Conference and RES Conference. This is the third chapter of my Ph.D. dissertation at KU Leuven. Financial support from the Research Foundation Flanders, Belgian American Educational Foundation and Fulbright Commission is gratefully acknowledged. Any errors are my own.

1 Introduction

"With no means of educating miners to their work, the conduct of mines in this country is a lamentable story of mismanagement, energy wrongly directed, and consequent great losses." John A. Church, mine superintendent, 1871

Recent evidence points to the importance of managers as drivers of firm performance. 'Good management' is usually viewed either as a well-defined set of best practices (Bloom & Van Reenen, 2007; Bloom, Sadun, & Van Reenen, 2017), or as a time-invariant individual trait (Hoffman & Tadelis, 2018; Bertrand & Schoar, 2003; Lazear, Shaw, & Stanton, 2015).[1] The exact mechanism through which managers affect firm performance is, however, less clear. Prior research tended to view managers as inputs in the production function or as shifters of the production isoquant.[2]

Managers could, however, also choose the production function itself, by adopting different technologies. In this paper, I examine how the human capital level of managers affects both their technology choices and their output elasticities. This distinction is crucial for the empirical evaluation of the productivity effects of managers. Suppose managers with high human capital levels systematically choose production functions with a higher production possibilities frontier, but are not inputs to production themselves. Conditional on all input choices, the presence of such managers will not correlate with total factor productivity (TFP), even if they do enhance output by choosing better technologies.

I illustrate this theoretical distinction by examining how the introduction of college-level mining engineering degrees in the U.S. affected productivity and technology usage in the coal mining industry. I track the first mining engineering graduates as they entered managerial positions in Pennsylvania coal mines between 1900 and 1914. This offers a stylized setting to think about management and technology choice which is unique for three reasons. First, the spread and rise of mining engineering colleges throughout the United States during this period provides a large shock to managerial human capital. Second, the first decade of the 20th century saw rapid technological change, as electrification fundamentally changed many industries, including coal mining. I observe detailed technology choices down to the mine level, together with standard production and cost variables. Finally, physical productivity was the main driver of firm performance in this industry, as coal is a homogeneous product that was sold on competitive markets. This offers a stylized setting to define and measure

[1] An example of such traits is the possession of good interpersonal skills. These two views of management have very different implications for how to 'manage managers'. In the first case, selection and allocation of managers is crucial (Benson, Li, & Shue, 2018; Terviö, 2009), and their education and training merely serves as a signaling mechanism. In the second case, managers be can actively improved, for instance by attending a business school or by hiring external consultants (Bloom, Eifert, Mahajan, McKenzie, & Roberts, 2013).

[2] With a Cobb-Douglas production function, both amount to the same. Managers could, however, also have factor-augmenting effects (Caroli & Van Reenen, 2001; Van Biesebroeck, 2003; Bender, Bloom, Card, Van Reenen, & Wolter, 2018).

firm performance.

The underlying question is relevant, however, beyond this historical setting. Technological innovations, such as artificial intelligence, and educational innovations, such as 'Tech MBAs',[3] usually happen simultaneously. When evaluating the effects of the educational innovation on firm performance, viewing managers only as production inputs will lead to different conclusions compared to a model where these managers also choose production technologies. As the returns from new technologies are often not immediately known, managers need to decide on technologies under uncertainty, and this is exactly where new educational programs could make a difference.

I specify and estimate an empirical model of coal extraction in which mine managers either hold a mining engineering degree, another college degree, or no degree at all. The educational background of the manager can both enter the production function as an input, shifting the isoquant, or influence production technology choices. I focus on how to haul coal to the surface, which was a crucial determinant of mine productivity. There were four alternative technologies (i) mules, the traditional technique, (ii) steam locomotives, (iii) electrical locomotives, and (iv) compressed-air locomotives. I find that mining school graduates do not shift the production frontier significantly conditional on all other inputs. In other words, they were not an input in the production function themselves. They were, however, more than twice as likely to choose electrical locomotives compared to other managers. These types of engines had a higher marginal product compared to other engines, while their costs were similar.

Finally, I examine the underlying mechanism why mining engineering graduates differed from other managers in terms of their technology choices. I find that the distinctive feature of mining college graduates lied in how they acquired information about new technologies. They knew on beforehand that electrical engines were superior to alternative innovative technologies. Ordinary managers, in contrast, had to learn about the returns to every new technology through trial-and-error and by observing adoption at other mines. The informational advantage of mining engineers was hence temporary: as time progressed, locomotive returns became common knowledge.

This paper makes three main contributions to the literature. First, I contribute to the literature on management and productivity, such as (Bloom & Van Reenen, 2007; Bloom et al., 2013, 2017). This literature tends to view managers as inputs in the production function.[4] I allow, in contrast, the production function itself to be endogenous. The implication from this is that the returns from managers is not limited to their 'direct' effects on total factor productivity. When controlling for all inputs and technology choices, and when addressing endogenous managerial hiring, mines managed

[3] https://www.economist.com/whichmba/tech-mbas-catching-up

[4] An alternative approach is to deduct the gains from good managers from variation in managerial compensation, using a revealed-preferences argument (Terviö, 2009). I do not use such an approach as I do not observe data on managerial compensation.

by mining college graduates were not more productive than other mines. By selecting the optimal production function through technology choices, these graduates did, however, increase output by 5.5% per year.

Secondly, I contribute to the literature on the productivity effects of higher education. The usual approach in this literature is to correlate TFP measures with managerial education (Bertrand & Schoar, 2003; Braguinsky, Ohyama, Okazaki, & Syverson, 2015). This again views managers as production inputs. I find, in contrast, that the main effect of mining engineering degrees was that it changed technology choices.[5] Interestingly, these effects held uniquely for mining college degrees. Managers with other, mostly liberal arts, college degrees were not different from uneducated managers in terms of their technology choices, even if they had attended elite institutions. My findings hence imply that the diversification of the US higher education landscape during the late 19th century was instrumental in driving innovation.

Finally, I contribute to the literature on technology adoption and diffusion. The role of information spillovers in technology adoption was studied in a many papers, including Levin et al. (1987); Besley and Case (1993); Jovanovic and Nyarko (1996); Foster and Rosenzweig (1995); Munshi (2004); Conley and Udry (2010). I contribute to the understanding of information spillovers and technology diffusion in two ways. First, I show that the importance of information spillovers decreased as managers became technically more educated. Informational spillovers did not affect technology choices of mining college graduates, but did for other managers. The hypothesis that education and information spillovers could be substitutable was already hypothesized by Rosenzweig (1995), but not found empirically. The reason for this is probably that they described settings in which agents enjoyed at most primary schooling, which is unlikely to provide concrete knowledge of new technologies.[6] Second, the benefits from technically educated managers were transitory and conditional on technological change: as information spread, all managers ultimately found out which technology to use.

The remainder of this paper is structured as follows. Section 2 discusses the industry background and data sources. A model of managerial education, learning and technology choice is constructed in section 3. Section 4 estimates this model and discusses the results, and is followed by conclusions.

[5]Other papers have found higher education increases technology adoption (Wozniak, 1984, 1987; Skinner & Staiger, 2005; Lleras-Muney & Lichtenberg, 2005), but did not examine whether the type of education matters, and did not contrast this view to the 'manager as input' channel. Toivanen and Väänänen (2016); Bianchi and Giorcelli (2017) finds that technical higher education led to more technology invention, but does not examine the effects on technology adoption.

[6]The educational system was also stationary in these papers, in contrast to the important educational shock which I exploit in this paper, and which allows to understand the interaction of educational change and technology diffusion better.

2 Industry background

2.1 Pennsylvania anthracite mining, 1900-1914

This paper studies anthracite extraction in Pennsylvania between 1900 and 1914. Anthracite is the coal type with the highest energy content, and the main coal product in Pennsylvania at the time. The extraction process consisted of three main steps. First, a shaft or tunnel had to be dug to reach the underground coal seam. Next, coal was excavated either using picks and black powder or using mechanized cutting machines. Finally, it had to be hauled back to the surface, either using mules or underground mining locomotives. Anthracite is a nearly homogeneous product, with limited differentiation in caloric content. It was sold as a fuel after extraction without any further processing, unlike other coal types.[7] Pennsylvania mines transported on average 85% of their output over the railroad network towards urban markets, with the remainder being either sold locally or re-used as an input. Product markets were hence likely to be in perfect competition. There are 287 unique anthracite firms and 615 unique mines in the dataset.

Coal mine management

Coal firms were headed by a general manager, often also the owner, who was based in a nearby city. Daily management at the mine level was delegated to 'superintendents' (henceforth 'managers'), who are the main object of interest in this paper. The lowest level of operational management was carried out by 'foremen', of which two thirds worked below the surface. Managers had a wide range of responsibilities, including technical procurement, human resources management, production line design, financial analysis and cost reporting (Ochs, 1992). As coal markets were competitive, low marginal costs were the main driver of firm performance.

Technological change: underground mine locomotives

Transporting coal to the surface 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 hauling, but were gradually replaced by underground mining locomotives from the 1880s onwards. Three main locomotive types existed: steam-powered, electrical and compressed air locomotives. Images of all three types are shown in figure 1.

[Figure 1 here]

[7] Bituminous coal is often degassed into 'cokes' before usage in order to improve its properties.

Steam locomotives were invented first, and were already relatively common by 1900. They were more efficient compared to mules, but also more dangerous: underground air quality deteriorated and there was a risk of explosions due to mine gas (Randolph, 1905). These concerns led to the development of electrical and compressed air locomotives during the 1890s. Compressed air locomotives were a relatively safe technology, but had the disadvantage of having to be refilled frequently, resulting in a limited range. Electrical locomotives, finally, required the installation of overhead lines, and could lead to electrocution in humid or flooded mines (Gairns, 1904). Their massive adoption indicates their usefulness. Electrification was a new technology, and electricity experts were rare, as the *Transactions of the Institute of Mining Engineers* mentioned:

> *"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)

Figure 2 depicts the total number of locomotives used of each type, and the share of mines using them. As mines could use several locomotive types concurrently, the shares add up to more than one. In 1900, around 2000 steam locomotives were already being used, while 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 as 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. Weighting these shares by mine output delivers somewhat higher shares as mines using locomotives were larger, but the evolution per type is very similar to the the unweighted version in figure 2.

[Figure 2 here]

Educational change: mining engineering programs

Up to the 1880s, the USA lagged far behind Europe in terms of technical higher education. While different continental European nations already had specialized engineering colleges from the early 18th century onwards, 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 demand for engineers. Lehigh University was home to an important mining college in Pennsylvania and 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 annual number of U.S. graduates with an *Engineer of Mines (E.M.)* title[8] quadrupled from 80 to 320 between 1898 and 1914, as shown in figure 3. As the supply of mining engineers grew, these graduates started to enter the coal mining industry. The solid red line in figure 3 shows that the share of mines managed by a college-level mining engineer increased from none in 1898 to 6% in 1914. The fraction of mines managed by a graduate with other college degrees grew from 2 to 6% as well. They slowly replaced an older generation of non-educated managers who had usually entered the mines around the age of twelve.[9] Managers with an E.M. degree were on average responsible for 6 different mines, twice as many as other managers. Their mines also employed nearly twice as many workers.

[Figure 3 here]

What did mining colleges teach? Appendix table A8 contains summary statistics on the curricula of 7 important mining colleges. They were heavily dominated by the natural sciences and various engineering branches, which made up for around 95% of credits. The remaining credits were used for non-science subjects such as language courses and administrative topics such as accounting and law. 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.[10]

Mining engineering graduates in Pennsylvania mainly graduated from local schools such as Lafayette College and Lehigh University. Managers with college degrees in other fields hailed, in contrast, predominantly from elite universities: two thirds of these managers graduated from Ivy League colleges such as Yale, Princeton, Cornell and Penn. Managers with college degrees were on average in their early thirties, or 15 years younger than the average non-educated manager. Summary statistics on managers of different backgrounds are in appendix table A7.

Figure 4 compares output per worker between mines managed by mining college graduates and other mines. I calculate labor productivity by dividing annual output by the number of worker-days.[11] Labor productivity grew on average by 3% per year between 1900 and 1914 in mines managed by a mining college graduate, compared to merely 1.7% for the other mines. The productivity gap between both groups was largest around 1910 and closed again by 1914.

[Figure 4 here]

[8]To fix terminology, when I mention 'college-level mining programs', I mean four-year undergraduate degrees with a specific focus on mining engineering, delivering an *Engineer of Mines (E.M.)* degree. This definition also excludes non-college-level technical schools and military academies.

[9]This becomes clear from reading biographies of important managers and state inspectors in the *Reports*: most of these men were born in England or Wales, had entered the mines at an early age, and had then migrated to the USA.

[10]I observe this, for instance, in the catalogs of the University of Utah and New Mexico School of Mines, where electricity was introduced in the mining curriculum during the late 1890s.

[11]I include mines that are not managed by a mining engineer, but were in the past three years in the 'mining engineering' category: the effects of mining engineers outlasts their tenure at the mine.

The visual correlation between managerial education and productivity could reflect that more productive firms chose better managers, or vice-versa. Besides, it is more meaningful to compare total factor productivity between these mines, rather than just output per worker. A more comprehensive production model is hence necessary.

2.2 Data sources

Production

Mine output and input data are obtained from the *Report of the Bureau of Mines* by the Department of Internal Affairs of Pennsylvania. It includes 615 Pennsylvania coal mines between 1898-1914. I observe annual coal extraction in tons and the share of output that is shipped, sold locally or re-used as inputs. Labor is measured in employee counts and days worked. Intermediate inputs include black powder and dynamite, which are measured in quantities (kegs and pounds, respectively). I also observe the number of mules used at each mine.

Managers

The given, middle and surnames of all firm and mine managers and of their deputies are observed. I only focus on mine superintendents because firm managers and foremen were almost never college-educated. I match full manager names with population census records and with college alumni records using *Ancestry.com*. I also cross-checked all managers with alumni records from US mining schools between 1870 and 1914, just to be sure I really observe all college-level engineers.

Technology

A complication is that the number of locomotives of each type are observed at the county-firm-year level, while all other variables are observed at the mine-year level. The average firm operated 2.6 mines, but four out of five firms operated just one mine. The average county in the dataset 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 calculations. I choose to bring the entire dataset to the mine-level in the baseline model 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.[12] Finally, geographical coordinates of the mining towns are obtained using Google

[12] As a robustness check, I re-estimate the main regressions at the firm-county-year level in the appendix, and it does not change the main conclusions.

Maps. A map with mining village locations and mining engineers is in appendix figure A1, and further details concerning the data sources are in appendix A.

3 Managers, productivity and technology choice

In this section, I set up a production and cost model of the coal mining industry in which managers both enter the production function directly, and choose the technology.

3.1 Production and costs

Mines i extract Q_{it} tons of coal in year t. Variable inputs are denoted $\mathbf{V}_{it}$ and includes labor, materials and mules. There are three types of locomotives: steam, electrical and/or compressed air locomotives. These types are denoted as $\tau \in \{\text{st, el, ca}\}$. Dummies indicating usage of each of these technologies are denoted $K_{it}^{\tau} \in \{0, 1\}$. The vector $\mathbf{K}_{it}$ collects these three dummies. It is possible to operate mines without any locomotive, in which case $\mathbf{K}_{it} = (0, 0, 0)$. This was the case for a third of all mines. Multiple locomotive types could also be operated simultaneously. Some capital was, finally, required to dig the tunnels to reach the coal seam, but these costs are considered sunk and is not formally modeled.

Mines are managed by superintendents, with a dummy $X_{it} \in \{0, 1\}$ indicating whether they obtained mining college degree. Managers often managed multiple mines. I will, however, consider input decisions to be independently made for at each mine separately, as markets are assumed perfectly competitive and costs assumed to evolve independently across mines. Let the production function be given by equation (1), with parametrization $\boldsymbol{\beta}$. In line with most of the literature, I rule out unobserved heterogeneity in the production parameters $\boldsymbol{\beta}$ across mines or over time. The educational background of managers enters the production function as an input.

$$Q_{it} = F\Big(\mathbf{V_{it}}, \mathbf{K_{it}}, X_{it}; \boldsymbol{\beta}\Big) \exp(\omega_{it}) \tag{1}$$

As is usual in the productivity literature, the residual ω is assumed to be a scalar which evolves following an AR(1) which is determined by a function $g(.)$:

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

I implement a functional form for $F(.)$ that allows each locomotive type to have different factor-biased effects. Adopting locomotives can change the marginal product of mules. In logarithms, the concrete production function used is given by equation (2):

$$q_{it} = \beta_v \mathbf{v}_{it} + \boldsymbol{\beta}_{vk} \mathbf{v}_{it} \circ \mathbf{K}_{it} + \boldsymbol{\beta}_k \mathbf{K}_{it} + \boldsymbol{\beta}_x \mathbf{X}_{it} + \omega_{it} \tag{2}$$

The output elasticity of a locomotive type is denoted θ^τ_{it}, and is a function of variable input usage and of the coefficient vectors $\boldsymbol{\beta}_{vk}$ and $\boldsymbol{\beta}_k$. The output elasticity of a locomotive can hence differ across mines depending on how many variable inputs they are using, due to the factor-biased technology effects.

$$\theta^\tau_{it} \equiv \frac{\partial Q_{it}}{\partial K^\tau_{it}} \frac{K^\tau_{it}}{Q_{it}} = \boldsymbol{\beta}^\tau_{vk} \mathbf{v}_{it} + \beta^\tau_k$$

The formulation of the production function in equation (2) implies different assumptions on coal extraction. First, locomotive output elasticities do not depend on whether the mine is managed by a mining engineer or not. I verify this assumption in appendix B.1. Next, I impose perfect substitutability between all three locomotive types: operating both an air and electrical locomotive does, for instance, not generate a higher return from either locomotive types. I again verify this assumption empirically in appendix B.1. Third, the locomotive coefficients do not change over time, so their is no technological progress within each locomotive type.

Variable vs. fixed inputs

Labor, materials, mules and managers can be flexibly adjusted every year can be freely adjusted every year. This makes sense as black powder and mules were sold on spot markets. Labor markets were very flexible and unregulated in the USA at the time (Naidu & Yuchtman, 2017). In appendix B.1, I extend the model to allow for adjustment costs on managerial labor, which barely changes the results.

Locomotives $\mathbf{K}_{it}$ evolve, however, dynamically. Adoption of type-τ capital is denoted $A^\tau_{it} \in \{0, 1\}$. I abstract from scrapping decisions, and only focus on adoption. Hence, mines need to decide one period ahead on which types of capital, if any, to use in the next period. I assume there is no depreciation. Due to the discreteness of the capital stock, the capital transition is given by equation (3).

$$K^\tau_{it} = K^\tau_{it-1} + A^\tau_{it-1}(1 - K^\tau_{it-1}) \tag{3}$$

Variable input demand

Mine managers are assumed to choose variable inputs annually in order to minimize costs. Let both product and input markets be perfectly competitive. I hence assume all mines make input decisions independently from each other. Denote input prices of inputs $\mathbf{V}$ as $\mathbf{W}^{\mathbf{V}}$. Variable inputs are chosen by minimizing static costs in each period, taking output Q^* as given:

$$\mathbf{V_{it}} = \arg\min \Big[\mathbf{W^V_{it}} \mathbf{V_{it}} - \lambda\Big(Q^*_{it} - F(.) \exp(\omega_{it})\Big)\Big] \tag{4}$$

As managers are assumed to be variable inputs as well in the baseline specification, demand for managers follows the same minimization problem as above, with the only difference that X is a binary variable.

Caveats

The model so far assumes there are no cost dynamics: past production does not affect current productivity. Cost dynamics may, however, be important in mining because 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 & Luengo, 2017; Asker, Collard-Wexler, & De Loecker, 2019). On the other hand, there could be learning by doing as in Benkard (2000). As an extension, I use dynamic specification in which lagged cumulative output enters flexibly in the productivity transition in appendix B.2. This delivers very similar results to the baseline approach.

I also assume there are no capacity constraints. Coal mine data from Illinois from the same period containing capacity data show that 95% of mines operated under 90% of their maximum extraction capacity, which suggests that capacity constraints were mostly non-binding.

3.2 Technology choice

The model so far assumed perfect information about prices and output elasticities of all variable inputs. All managers hence choose variable inputs in the same way, by solving the cost minimization problem in equation (4).

As different locomotive types were newly invented, it is more likely that there was uncertainty about the returns to these technologies for at least some managers. I therefore specify a technology adoption model in which managers learn about locomotive returns through different channels.

Information and learning

Managers of mine i have perfect information about all production function coefficients, except for the locomotive coefficients $\boldsymbol{\beta_{vk}}$ and $\boldsymbol{\beta_k}$. Together with variable inputs usage, these coefficients determine the output elasticity of each locomotive type, θ^{τ}_{it}. In time $t = 0$, just before the first year of observation, the manager of mine i has a private prior expectation $\hat{\theta}^{\tau}_{i0}$ about the output elasticity of locomotive type τ. The distribution of this prior depends on whether managers obtained a mining degree or not. The standard deviation of the prior is given by $\sigma^{\tau}(X_{i0})$, and is lower for mines managed by a manager with a mining degree: $\sigma^{\tau}(1) \leq \sigma^{\tau}(0)$. Otherwise said, priors about each technology by mining engineers are more precise than priors of other managers.

$$\hat{\theta}^{\tau}_{i0}(X_{it})z \sim \mathcal{D}\Big(\theta^{\tau}_{it}, \sigma^{\tau}(X_{i0})\Big)$$

Each period, a noisy signal u^{τ}_{it} arrives about the output elasticity of every locomotive type. As more locomotives get adopted in other mines in the same town ℓ, the precision of this signal increases. Denote the set of all mines in town ℓ by $\mathcal{I}_{\ell}$ and the number of locomotives of type τ which are being used at other mines as $K^{\tau}_{-it} \equiv \sum_{r \in \mathcal{I}_{\ell} \setminus \{i\}}(K^{\tau}_{rt})$. The signal u^{τ}_{it} has an unbiased distribution $\mathcal{E}$, of which the standard deviation $\varsigma^{\tau}(K^{\tau}_{-it})$ decreases with the number of locomotives in the same town:

$$u^{\tau}_{it} \sim \mathcal{E}\Big(\theta^{\tau}_{it}, \varsigma^{\tau}(K^{\tau}_{-it})\Big)$$

Using Bayesian updating, the updated prior in period t, $\hat{\theta}^{\tau}_{it}$, is a weighted average of the previous period's prior and of the signal u^{τ}_{it}, as long as the mine does not operate any locomotives of type τ. As soon as a locomotive of type τ is installed, managers immediately observe its output elasticity θ^{τ}.

$$\hat{\theta}^{\tau}_{it} = \Big[\Big(1 - \alpha^{\tau}_{it}(K^{\tau}_{-it})\Big)\hat{\theta}^{\tau}_{it-1} + \alpha^{\tau}_{it}u^{\tau}_{it}\Big](1 - K^{\tau}_{it}) + \theta^{\tau}_{it}K^{\tau}_{it}$$

The Bayesian weights α^{τ}_{it} depend on the relative standard deviations of the original prior, σ and of the information shocks, ς. Both these standard deviations depends on whether the mine is managed by a mining engineer and on adoption by neighboring rivals:

$$\alpha^{\tau}_{it} = \alpha^{\tau}(X_{it}, K^{\tau}_{-it})$$

In contrast with the prior literature in which managers have common priors, such as in Munshi (2004), I allow the weights α^{τ} do differ across managers and over time. Managers with a mining degree have a perfect signal about each technology, meaning that $\sigma^{\tau}(1,.) = 0$. In this case, managers do not place any weight on the signal u: $\alpha^{\tau}(0, K^{\tau}_{-it}) = 0$. In case the manager does not have a mining degree, the weight on the prior diminishes as more locomotives are adopted by competitors, as this reduces the standard deviation of the signal u^{τ}_{it}:

$$\frac{\partial \alpha^{\tau}(X_{it}, K^{\tau}_{-it})}{\partial K^{\tau}_{-it}} \geq 0$$

Eventually, as many locomotives get adopted, the signal u^{τ} becomes a very precise estimate of the true output elasticity β^{τ} so managers only base their prior on this signal: α_{it} approaches unity.

Fixed costs

Locomotives of type τ have fixed costs Φ^{τ}_{it}. I assume these costs depend on usage of other locomotive types, which I collect in a vector $\tilde{\mathbf{K}}^{\tau}_{it}$ and on mine time-invariant characteristics δ_i, and have an i.i.d. component ν^{τ}_{it}. I assume the managers know the distribution from which the transient fixed cost shocks ν^{τ}_{it} are drawn.

$$\Phi^{\tau} = \Phi(\tilde{\mathbf{K}}^{\tau}_{it}, \delta_i) + \nu^{\tau}_{it}$$

Decision problem

I assume managers are risk neutral. Most existing papers on technology adoption under uncertainty assume risk aversion, but this would not change the model implications of interest. Each year t, managers decide whether to adopt locomotives of each type, A^{τ}_{it}. Normalizing exogenous coal prices to one, annual added value from using technology τ is denoted $\Delta^{\tau}\Pi$ and given by:

$$\begin{aligned}
\Delta^{\tau}\Pi_{it} &\equiv \Pi(., K^{\tau}_{it} = 1) - \Pi_{it}(., K^{\tau}_{it} = 0) \\
&= Q(., K^{\tau}_{it} = 1) - Q(., K^{\tau}_{it} = 1) - \Phi^{\tau}_{it} \\
&= \theta^{\tau}_{it} Q_{it} - \Phi^{\tau}_{it}
\end{aligned}$$

Locomotive resturns are uncertain, so managers base their adoption decision on their expectation of these returns. This expectation is given by equation (5). As shown in the theoretical model above, expectations of the output elasticity $\hat{\theta}^{\tau}_{it}$ are a function of managerial education, adoption by others and the interaction between these variables. Managers also form an expectation over fixed costs Φ and output Q, but these expectations are not mine-specific as the distributions of the underlying transient stochastic variables ε_{it} (productivity shock) and ν_{it} (fixed cost shock) are assumed to be public knowledge.

$$\mathbb{E}_{it}[\Delta^{\tau}\Pi_{it+1}] = \hat{\theta}^{\tau}_{it}(\theta^{\tau}_{it}, X_{it}, K^{\tau}_{-it})\mathbb{E}_t[Q_{it+1}] - \mathbb{E}_t[\Phi^{\tau}_{it+1}] \tag{5}$$

The manager of mine i decides on locomotive adoption A^{τ} which maximizes his expected profit from doing so:

$$A^{\tau}_{it} = 1 \Leftrightarrow \mathbb{E}_{it}[\Delta^{\tau}\Pi_{t+1}] > 0$$

3.3 Model implications

Information about and benefits of various locomotives

Before deriving implications from the model for locomotive usage, I now discuss how locomotives compared in terms of output elasticities and the public/private nature of this information. As will be shown later in the production function results section, electrical engines had on average higher value added than steam locomotives, which had a similar value compared to compressed air locomotives:

$$\mathbb{E}[\beta_{it}^{el}] > \mathbb{E}[\beta_{it}^{st}] = \mathbb{E}[\beta_{it}^{ca}]$$

Both electrical and compressed air locomotives had benefits which were not publicly known in 1900. The value added of steam locomotives was, however, publicly known to all in 1900 as 80% of mines were located in a town with at least one steam engine. In terms of the variances of the knowledge shocks, this implies that:

$$\varsigma^{\tau} \geq 0 \text{ if } \tau \in \{\text{el,ca}\}$$
$$\varsigma^{\tau} = 0 \text{ if } \tau \in \{\text{st}\}$$

Mean fixed operating costs Φ^{τ} were very similar for all three types, as becomes clear from the cost data.

Five hypotheses about technology choice

I now derive five key implications of the model which will be verified empirically.

Hypothesis 1 *For non-engineers, adoption of electrical locomotives by others correlates positively with their own electricity adoption.*

Managers without a mining degree start out with a noisy prior about each technology's benefits. As shown in the model, they hence place a positive weight α^{el} on electrical locomotive adoption by other managers. Such adoption increases the precision of the manager's expectation of electrical locomotive benefits, which increases the likelihood of adopting them.

Hypothesis 2 *For non-engineers, adoption of compressed-air locomotives by others correlates positively with their own compressed-air adoption.*

For compressed-air locomotives, the same argument holds as for hypothesis 1, with the only difference that air locomotives had low benefits. As compressed air adoption by others makes it clear that these locomotives were suboptimal to use, adoption fell.

Hypothesis 3 *Independently of the manager's education, adoption of steam locomotives by others does not correlate with their own steam adoption.*

This directly follows from the fact that information about steam locomotives value added was public. Hence, $\hat{\beta}_{it}^{st} = \beta_{it} \; \forall i, t$ and $\alpha^{st} = 0$.

Hypothesis 4 *For mining engineers, adoption of electrical and compressed air locomotives by others does not correlate with their own adoption of these locomotives.*

Because their prior is much more precise than the signal they obtain from other mines' locomotive usage, mining engineers solely rely on their prior. In the model, this means that the value of α is zero, as already stated in the model.

Hypothesis 5 *Mining engineers will adopt electrical locomotives faster than other managers, but the difference between both manager types decreases over time.*

The probability that a mining engineer believes that electrical engine benefits surpass all other locomotive types is higher compared to non-engineers, as the latter are assumed to have more noisy priors. Managers with a mining degree will hence adopt electrical locomotives faster than other managers. As time elapsed, though, non-engineers updated their beliefs about the benefits of all locomotive types. The difference in electrical locomotive usage between both groups of managers should hence decrease over time.

4 Empirical analysis

4.1 Estimation

Production function

The estimation of the production function mostly follows Ackerberg, Caves, and Frazer (2015). The only difference is that technology adoption by others now needs to enter the input demand function, and hence the first stage regression.

I use the log production function from equation (2). Denoting the vector of all variable inputs except intermediate inputs as $\tilde{\mathbf{v}}_{\mathbf{it}}$, intermediate input demand is given by:

$$m_{it} = m(\omega_{it}, \tilde{\mathbf{v}}_{\mathbf{it}}, \mathbf{k}_{\mathbf{it}}, \mathbf{k}_{-\mathbf{it}}, x_{it}, t)$$

The first stage regression is given by the following equation, with Ψ being a second-order polynomial.

$$q_{it} = \Psi(\omega_{it}, \mathbf{v}_{\mathbf{it}}, \mathbf{k}_{\mathbf{it}}, \mathbf{k}_{-\mathbf{it}}, x_{it})$$

Next, I estimate the production function coefficients using the moment conditions in (6). The timing assumptions are consistent with the variable/fixed inputs classification already made: labor, intermediate inputs and mules react immediately to productivity shocks, while the various capital types can only be adjusted with a one-year lag. In the baseline specification, I consider managers to be just another type of labor, which can hence be adjusted immediately. In a robustness check I allow for managerial adjustment costs, but this does not change the main conclusions.

$$\mathbb{E}\left\{ \xi_{it}(\boldsymbol{\beta}_v, \boldsymbol{\beta}_k, \boldsymbol{\beta}_{vk}, \boldsymbol{\beta}_x) \begin{pmatrix} \mathbf{v}_{\mathbf{it-1}} \\ \mathbf{k}_{\mathbf{it}} \\ \mathbf{v}_{\mathbf{it-1}} \circ \mathbf{k}_{\mathbf{it}} \\ \mathbf{x}_{\mathbf{it-1}} \end{pmatrix} \right\} = 0 \tag{6}$$

Labor is measured as employee counts multiplied by days worked, materials as the number of powder kegs used and animals as the average number of mules used in a given year. The three capital variables are measured as dummies for the usage of at least one electrical locomotive, air locomotive or steam locomotive. As I take logs of both the number of powder kegs used and the number of mules, I omit mines which do not use any powder or any mules. This amounts to 18% of the observations (879 observations), but just 6.8% of total output. The omitted mines are hence much smaller than average. Locomotive usage by others is calculated as the total number of mines in the same town (as reported in the data) who use a certain locomotive type. I define spillovers at the town-level, as most mining towns were small and isolated. I refer to appendix A for further details on measurement. In order to obtain the correct standard errors, I bootstrap using 50 iterations.

Technology choice

Equation (5) showed that usage of technology τ depends on (i) the true output elasticity of that technology, θ_{it}^{τ}, (ii) output Q_{it}, the educational background of the manager, X, (iii) technology usage by others, K_{it}^{τ} and (iv) fixed costs Φ_{it}^{τ}. The output elasticity of each technology depends on variable input usage, which is observed, and on the production function coefficients, which were estimated in the previous section. Fixed costs are latent, but assumed to depend on prior locomotive usage of other types and on time-invariant unobservables, such as location dummies, δ_i.

As introduced before, usage of other locomotive types than type τ is given by a vector $\tilde{\mathbf{K}}^{\tau}_{\mathbf{it}}$. I collect the variables affecting locomotive choice which are not of direct in the paper into a vector $\mathbf{z}^{\tau}_{\mathbf{it}} = (\mathbf{v}_{\mathbf{it}}, \tilde{\mathbf{K}}^{\tau}_{\mathbf{it}}, \omega_{it}, t, \delta_i)$.

Imposing a type-I i.i.d. distribution for the transient fixed costs shocks ν^{τ}_{it} in equation (5) leads to the logit model in (7).

$$Pr(K^{\tau}_{it+1} = 1|.) = \frac{\exp(\gamma^{\tau}_X X_{it} + \gamma^{\tau}_{XK} X_{it} K_{-it} + \gamma^{\tau}_K K_{-it} + \boldsymbol{\gamma}^{\tau}_Z \mathbf{z}^{\tau}_{\mathbf{it}})}{1 + \exp(\gamma^{\tau}_X X_{it} + \gamma^{\tau}_{XK} X_{it} K_{-it} + \gamma^{\tau}_K K_{-it} + \boldsymbol{\gamma}^{\tau}_Z \mathbf{z}^{\tau}_{\mathbf{it}})} \tag{7}$$

In the baseline model, I allow for county-specific fixed costs by defining δ_i as a set of county dummies. In case there are still unobserved mine-level variables which affect either the cost or return from operating each locomotive type and mining college graduate hiring, this results in an endogeneity problem. I use an alternative linear specification with mine fixed effects in equation in appendix B.4. In that same appendix, I also let mining college graduates enter with lags and leads in order to examine whether mining college graduates led to different technology choices, or the other way around.

4.2 Results

Production function

The estimated output elasticities are in the first two columns of table 1. These are not the full set of β coefficients from equation (2), but the partial derivatives of output w.r.t. each input. The manager's educational background did not change output significantly, conditional on all other inputs. The production process relied heavily on labor, with an output elasticity of 0.77. The three locomotive types had very different output elasticities. The coefficient on electrical locomotives is 0.11, meaning that adopting such a machine increased output by 11.3%.[13] The other two locomotive types, with steam and compressed air engines, did not increase output on average. For steam locomotives this can be due to the fact that there is very little variation in its usage, as most mines had already adopted this technology in 1900. Compressed-air locomotives were, on the contrary, a brand-new technology, but seemingly did not lead to any productivity gains. This is consistent with historical evidence which was discussed in section 2.

[Table 1 here]

[13] $\exp(0.107) - 1$

Mules had an average output elasticity of 0.173, but this is an average over mines with and without locomotives, which were a substitute to mules. The interaction terms in table 1 shows that the output elasticity of mules was higher when not using any locomotives. Locomotives hence reduced the usefulness of mules, which is to be expected as they were complementary technologies. Interestingly, not all three locomotive types had the same mule-substituting effects. Electrical and steam locomotives both considerably reduced the output elasticity of mules, but compressed-air locomotives did not. This is yet another indication of the limitations of compressed-air locomotives.

Technology choice

The estimates of the technology choice model from equation (7) are in table 2. I report the marginal effects at the mean and do not report the control variables in $\mathbf{z_{it}}$. Three main results stand out. First, managers with mining degrees are 54 percentage points more likely to adopt electrical locomotives, on an average adoption rate of 42%. This effect is thus very large in size, and significant. Mining college graduates are, however, not more likely to adopt either of the other locomotive types. Managers with other college degrees did not adopt more locomotives of any type. The fact that that mining college graduates were much faster to adopt electrical locomotives compared to other managers is consistent with hypothesis 5.

[Table 2 here]

Second, adoption of the same locomotive type at other mines in the same market is associated with higher adoption of electrical engines, while this correlation is negative for compressed-air locomotives and practically zero for steam engines. An additional mine using an electrical locomotive is associated with an increase in electrical engine usage of 2.1 p.p., which is a relative increase of 5%. [14]. For steam locomotives this increase is 0.7 p.p., which is a relative change of only 0.1%, i.e. close to nothing. For compressed-air locomotives, adoption of those locomotives in other mines are associated with a *fall* in usage of 5.3 p.p., or a relative decline of 18%[15]. These findings are in line with hypotheses 1-3.

Third, locomotive usage by others does not correlate with locomotive adoption of any type when mines are managed by mining college graduates, as predicted in hypothesis 4. This knowledge effect is only present for mining engineers: the interaction terms with other college degree types (not reported) are not significant and small.

[14]=0.021/0.426
[15]-0.05/0.289

Evolution of locomotive usage over time

If imperfect information is the main driver for the higher electrical locomotive adoption by mining engineers, than the difference in this adoption between managers should fall over time, as argued in hypothesis 5. I interact the mining degree effects on electricity adoption with indicator variables of two time periods: the initial period 1900-1903 and a second period 1904-1907. The final period of 1908-1914 serves as the comparison group. The choice of these time periods is inspired by the usage graph in figure 2, which shows that electrical and compressed-air usage grew at similar rates up to 1904, which suggests that the benefits from electrification were unknown to the general public up to that point.

The estimates in table 3 show that the difference between mining engineers and other managers was indeed the largest in the earlier years. The ratio of usage probabilities between both groups fell from 2.6 before 1903 to 1.73 up to 1907, and to 1.38 afterwards. This is again in line with hypothesis 5 which supports the information spillover mechanism.

[Table 3 here]

4.3 Productivity and profitability effects

Productivity

How much did better technology choices by mining college graduates contribute to the productivity of the Pennsylvania coal mining industry? I calculate the expected value of output conditional on all inputs, except electrical locomotives, when mines are headed by a mining college graduate in the numerator of equation (8). The denominator is the expected value of output when having another type of manager. The derivation of this equation is in appendix C.1.

$$\frac{\mathbb{E}(Q|X=1)}{\mathbb{E}(Q|X=0)} = \frac{(1+\theta^e)\mathbb{E}(K=1|X=1)+\mathbb{E}(K=0|X=1)}{(1+\theta^e)\mathbb{E}(K=1|X=0)+\mathbb{E}(K=0|X=0)} \tag{8}$$

Calculating this ratio, I find that mines with mining engineers extracted 5.5% more coal with the same bundle of inputs, except for mining engineers and electrical locomotives. This ratio is, however, merely an average and conceals large heterogeneity over time. As was already shown in table 3, the difference between mining college graduates and other managers in terms of electrical engine adoption, and hence in terms of productivity, was largest initially and declined throughout the observed time period.

[Table 4 here]

Profits

While they arguably increased productivity, both electrical locomotives and managers with mining degrees came at a price, so what were their effects on profits? As the theoretical model assumed perfectly competitive output and input markets, this naturally implies zero profits. Combining, however, both locomotive cost data from archival sources with the locomotive benefit estimates from the model, I find that electrical locomotives earned a higher marginal product compared to their costs. This is not necessarily contradictory to price-taking behavior of firms on locomotive markets, as the information of most managers was imperfect.

I calculate the mean profit gain from having a mining college graduate as manager, $\Delta\Pi \equiv \mathbb{E}(\Pi(X = 1)) - \mathbb{E}(\Pi(X = 0))$. The precise expression for this profit gain is in appendix C.2. In order to know locomotive fixed costs, I use cost data from Randolph (1905). This data reports that installing 4 electrical locomotives cost \$32,100 in 1900, as shown in figure A5. In 2019 US dollars, which I will now express all monetary amounts in, this is around \$1 million. Conditional on operating at least one electrical locomotive, the average mine operated 26 electrical locomotives, which gives a cost of \$6.3 M. In reality, locomotives did not of course not depreciate entirely in one year's time. I use two different linear depreciation rates of 0.1 and 0.2, meaning that electrical locomotives last for 10 and 5 years. It is possible that managers with mining degrees earned higher wages than to other managers[16], but I do not observe manager-level wages. In one scenario, I assume mining college graduates earned the same wage as other managers, which was on average \$940 in 1910, or \$24,391 per year in 2019. In a second scenario, I let managers with a mining degree earn three times as much as the others.

The average profit gains from hiring a mining engineer managers are listed in table 4. The four rows correspond to four scenarios. In the most optimistic case, with an annual depreciation of 10% and no wage premium for mining engineers, profits increase by \$1.23 M in 2019 US dollars, or 8.3% of annual revenue. In the most conservative scenario, with 20% depreciation and a wage premium of 200%, the profit increase was \$0.86 M, which was 6.0% of annual revenue.

[Table 4 here]

How much large was the relative increase in profits? This depends on how large initial variable profits, net of locomotive and manager adoption, were. First, I use a benchmark profit rate (not considering the effects of electrical locomotives) of 0%, as input and product markets are assumed

[16] although it took mine owners some time to realize the potential of electrical locomotives, so the same might hold for their assesment of technically educated managers

to be competitive. Next, I also allow for a baseline profit rate of 10% of annual revenue. In the first case, hiring a mining college graduate increases profits by between 111 and 118$\%$. Note that profits without mining engineers, $\Pi(X = 0)$ are above zero: even mines without mining engineers adopt electrical locomotives, whose benefits exceed their costs. In the second case, which seems much more plausible, profits increased by 14 to 19% when adopting a mining college graduate.

Why not all firms hired mining engineers

It seems surprising that so few mines hired mining college graduates considering that they increased profitability. The supply of mining college graduates was, however, still tiny during the first decade of the 20th century, even if it grew fastly. In 1909, after the large entry wave of mining colleges, still only 4000 persons had ever graduated from a U.S. mining college program in the USA, and some of these could have already died by then. In comparison, there were over 6000 coal mines in the USA and many more in other extractive industries such as copper or gold mining. College yearbooks show that many mining engineers migrated to work in Mexican, Canadian or Australian mines, and many also entered different industries, became civil servants or academic researchers (Ochs, 1992). Coal mining was also more competitive and hence less profitable than some other extractive industries, which made it less attractive for highly solicited mining college graduates.

5 Conclusion

In this paper, I find that the introduction of college-level mining engineering degrees changed technology choices made by managers. Managers without these degrees had to learn about new technologies by observing other mines who adopted such technologies. Mining engineering graduates accessed, however, private knowledge about these technologies and hence did not need to learn from competitors. This led to considerably increased productivity and profits. The findings of this paper show that the differentiation of the US higher education system from the late 19th century onwards had important consequences for innovation and productivity growth of firms, but that these gains were also conditional on the arrival of new technologies.

References

Ackerberg, D., Caves, K., & Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica*, *83*(6), 2411-2451.

Aguirregabiria, V., & Luengo, A. (2017). *A microeconometric dynamic structural model of copper mining decisions.* (Mimeo)

Asker, J., Collard-Wexler, A., & De Loecker, J. (2019). (mis) allocation, market power, and global oil extraction. *American Economic Review*, *109*(4), 1568-1615.

Bender, S., Bloom, N., Card, D., Van Reenen, J., & Wolter, S. (2018). Management practices, workforce selection, and productivity. *Journal of Labor Economics*, *36*(S1), S371–S409.

Benkard, C. L. (2000). Learning and forgetting: the dynamics of aircraft production. *American Economic Review*, *90*(4), 1034-1054.

Benson, A., Li, D., & Shue, K. (2018). *Promotions and the peter principle* (Tech. Rep.). National Bureau of Economic Research.

Bertrand, M., & Schoar, A. (2003). Managing with style: The effect of managers on firm policies. *Quarterly Journal of Economics*, *118*(4), 1169-1208.

Besley, T., & Case, A. (1993). Modeling technology adoption in developing countries. *The American economic review, Papers and Proceedings*, *83*(2), 396–402.

Bianchi, N., & Giorcelli, M. (2017, March). *Scientific education and innovation: From technical diplomas to university STEM degrees.*

Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., & Roberts, J. (2013). Does management matter? evidence from India. *The Quarterly Journal of Economics*, *128*(1), 1-51.

Bloom, N., Sadun, R., & Van Reenen, J. (2017). *Management as technology?* (NBER Working Paper No. 22327). National Bureau of Economic Research.

Bloom, N., & Van Reenen, J. (2007). Measuring and explaining management practices across firms and countries. *The Quarterly Journal of Economics*, *122*(4), 1351-1408.

Braguinsky, S., Ohyama, A., Okazaki, T., & Syverson, C. (2015). Acquisitions, productivity and profitability: evidence from the Japanese cotton spinning industry. *The American Economic Review*, *105*(7), 2086-2119.

Caroli, E., & Van Reenen, J. (2001). Skill-biased organizational change? evidence from a panel of British and French establishments. *Quarterly Journal of Economics*, *116*(4), 1449-1492.

Church, J. A. (1871). Mining schools in the United States. *The North American Review*, *112*(230), 62-81.

Conley, T. G., & Udry, C. R. (2010). Learning about a new technology: Pineapple in Ghana. *American Economic Review*, *100*(1), 35-69.

Foster, A. D., & Rosenzweig, M. R. (1995). Learning by doing from others: Human capital and technical change in agriculture. *Journal of Political Economy*, *103*(6), 1176-1209.

Gairns, J. (1904). Industrial locomotives for mining, factory and allied uses. In (Vol. 26, p. 291-309). Cassier's Magazine.

Hodges, C. (1905). Compressed air locomotives for mine haulage. *Mining Magazine*, *11*(3), 305-310.

Hoffman, M., & Tadelis, S. (2018). *People management skills, employee attrition, and manager rewards: An empirical analysis* (Tech. Rep.). National Bureau of Economic Research.

Jovanovic, B., & Nyarko, Y. (1996). Learning by doing and the choice of technology. *Econometrica*, *64*(4), 1299-1310.

Lazear, E. P., Shaw, K. L., & Stanton, C. T. (2015). The value of bosses. *Journal of Labor Economics*, *33*(4), 823–861.

Levin, R. C., Klevorick, A. K., Nelson, R. R., Winter, S. G., Gilbert, R., & Griliches, Z. (1987). Appropriating the returns from industrial research and development. *Brookings papers on economic activity*, *1987*(3), 783–831.

Lleras-Muney, A., & Lichtenberg, F. R. (2005). Are the more educated more likely to use new drugs? *Annales d'Économie et de Statistique*, 671–696.

Lundgreen, P. (1990). 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*, *47*(1), 33-75.

Munshi, K. (2004). Social learning in a heterogeneous population: technology diffusion in the indian green revolution. *Journal of development Economics*, *73*(1), 185–213.

Naidu, S., & Yuchtman, N. (2017). Labor market institutions in the gilded age of American history. *The Oxford Handbook of American Economic History*, *Forthcoming*.

Ochs, K. H. (1992). The rise of American mining engineers: A case study of the Colorado School of Mines. *Technology and Culture*, *33*(2), 278-301.

Randolph, B. S. (1905). Comparison of electric and compressed-air locomotives in American mines. In M. W. Brown (Ed.), *Transactions of the institution of mining engineers* (Vol. 27). Institution of Mining Engineers.

Rosenzweig, M. R. (1995). Why are there returns to schooling? *The American Economic Review*, *85*(2), 153–158.

Skinner, J., & Staiger, D. (2005). *Technology adoption from hybrid corn to beta blockers* (Tech. Rep.). National Bureau of Economic Research.

Terviö, M. (2009). Superstars and mediocrities: Market failure in the discovery of talent. *The Review of Economic Studies*, *76*(2), 829–850.

Toivanen, O., & Väänänen, L. (2016). Education and invention. *Review of Economics and Statistics*, *98*(2), 382–396.

Van Biesebroeck, J. (2003). Productivity dynamics with technology choice: An application to automobile assembly. *The Review of Economic Studies*, *70*(1), 167-198.

Wozniak, G. D. (1984). The adoption of interrelated innovations: A human capital approach. *The*

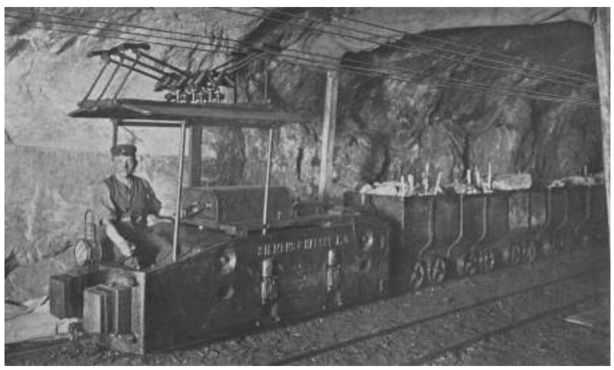
Review of Economics and Statistics, 70–79.

Wozniak, G. D. (1987). Human capital, information, and the early adoption of new technology. *Journal of Human Resources*, 101–112.

Figure 1: Mining locomotive types

(a) Steam

(b) Electrical

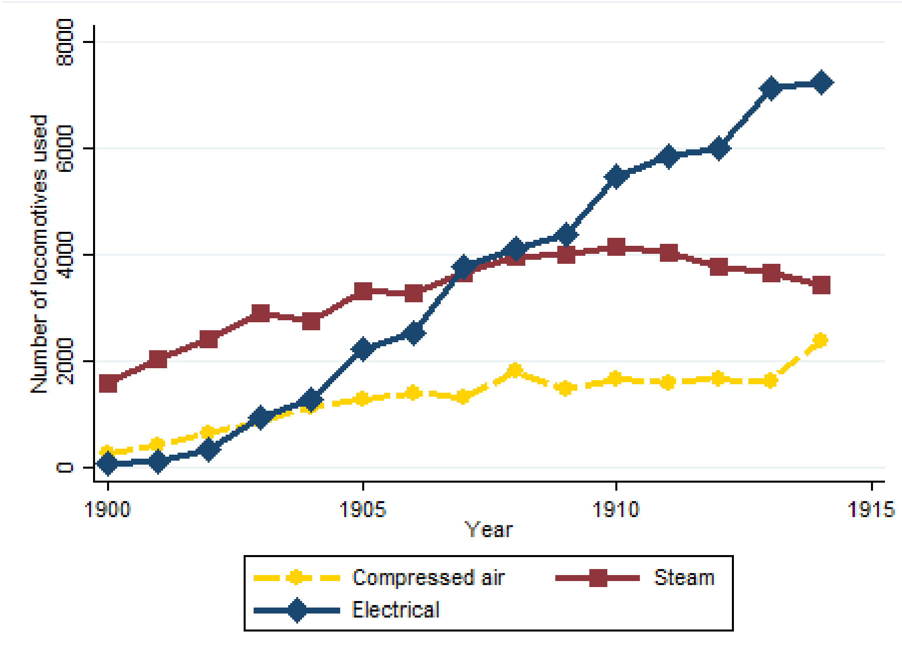


(c) Compressed air

Source: Gairns (1904)

Figure 2: Usage of mining locomotives

(a) Total number in use



(b) Share of mines

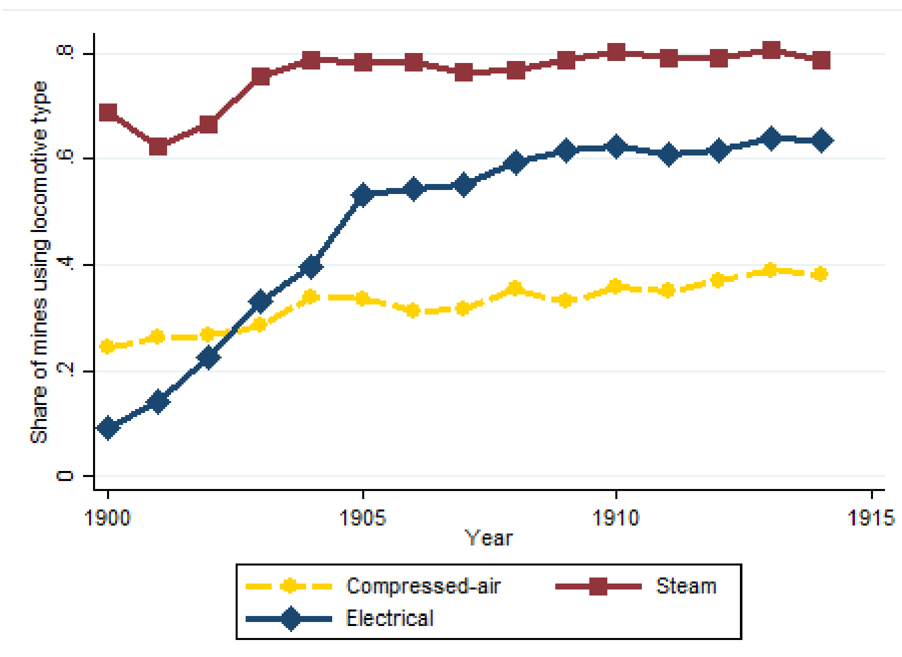
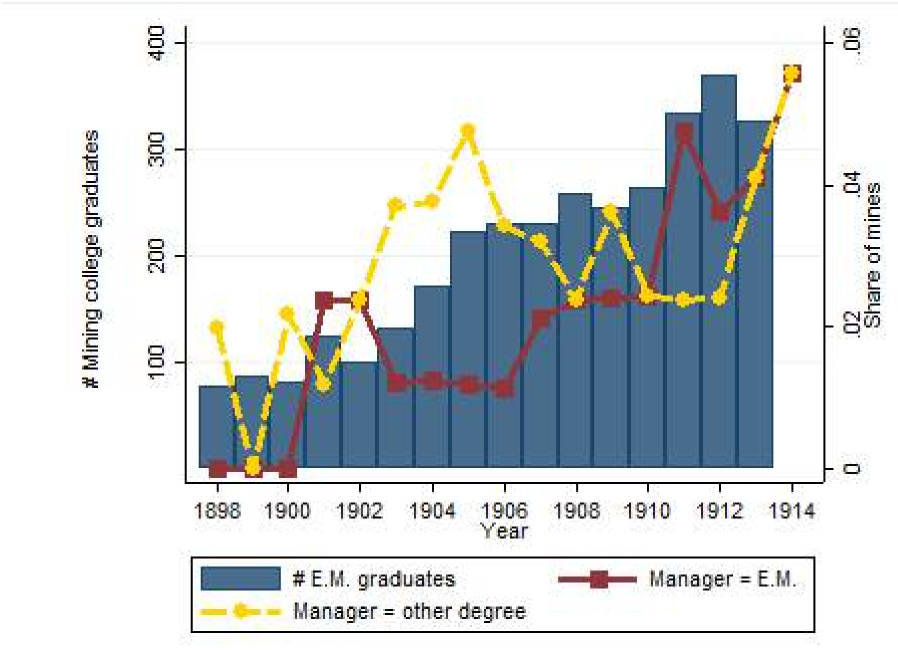
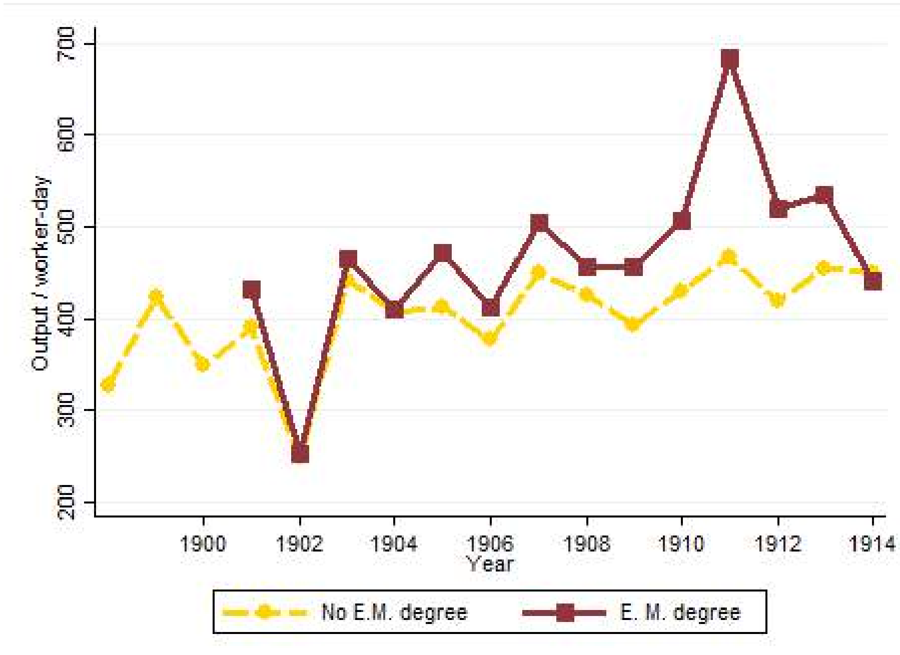


Figure 3: The rise of U.S. mining colleges



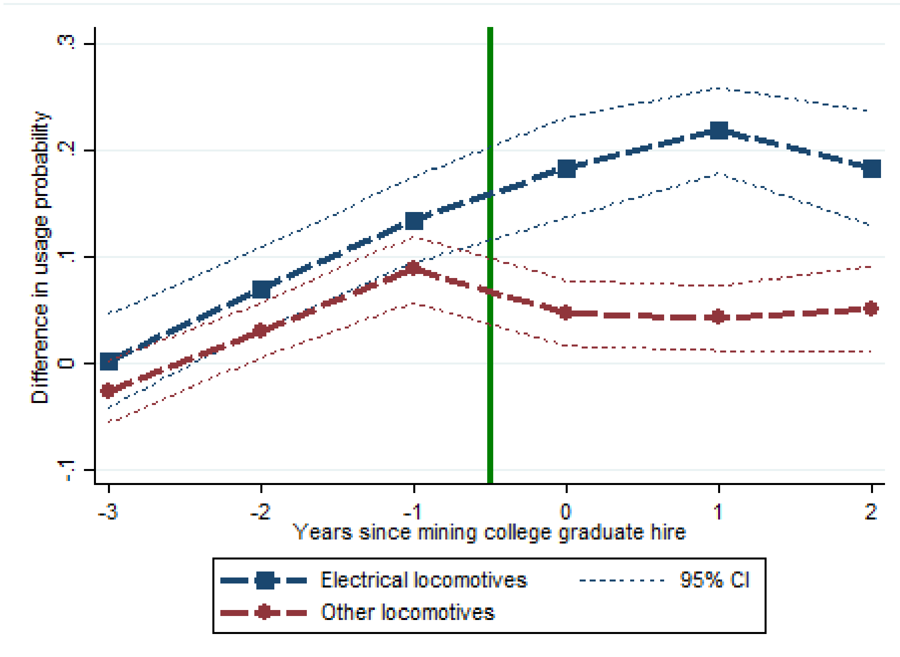
Notes: The blue bars show the annual number of graduates (left axis) with an Engineer of Mines (E.M.) degree from all US mining colleges. The solid red line plots the share of Pennsylvania anthracite mines with a manager with such an E.M. degree, the dotted yellow line does the same for other college degrees.

Figure 4: Labor productivity and managerial education



Notes: The solid red line shows plots the evolution of output per worker per day in mines managed by a mining college graduate. The dotted yellow line does the same for mines without such a manager.

Figure 5: Mining college graduates and electrical locomotive adoption



Note: This graph plots the estimated difference in electrical locomotive usage between mines managed by mining college graduates and other mines three years before and two years after adopting the graduate. Controls: productivity, variable input usage, other locomotive types usage, mine fixed effects, year fixed effects.

Table 1: Production function

Panel (a): Output elasticities	log(Output)	
	Estimate:	SE:
1(Mining degree)	-0.019	(0.183)
1(Elec. loc.)	0.107	(0.046)
1(Air loc.)	0.045	(0.268)
1(Steam loc.)	0.043	(0.042)
log(Labor)	0.774	(0.181)
log(Materials)	0.016	(0.022)
log(Mules)	0.173	(0.148)
Observations	3,221	
R-squared	0.893	
Panel (b): Interaction terms	log(Output)	
	Estimate:	SE:
log(Elec. loc.)*log(Mules)	-0.225	(0.083)
log(Steam loc.)*log(Mules)	-0.173	(0.068)
log(Air loc.)*log(Mules)	0.040	(0.061)

Notes: Dependent variable is log output. All variables in logs , except for mining college and locomotive dummies. Bootstrapped standard errors in parentheses, 50 iterations. Panel (a) contains output elasticities, panel (b) interaction terms.

Table 2: Technology choice

Locomotive types:	1(Elec. loc.)		1(Steam loc.)		1(Air loc.)	
	Estimate:	SE:	Estimate:	SE:	Estimate:	SE:
1(Mining degree)	0.539	(0.068)	-0.084	(0.071)	-0.153	(0.146)
1(Other degree)	-0.041	(0.069)	0.002	(0.062)	-0.278	(0.148)
Adoption of others	0.021	(0.003)	0.007	(0.003)	-0.053	(0.013)
Adoption of others * 1(Mining degree)	-0.099	(0.015)	0.013	(0.009)	0.104	(0.047)
Observations	3,339		3,339		3,339	
Mean usage:	0.426		0.676		0.289	

Notes: Dependent variables are dummies for usage of each locomotive type. Logit model used, marginal effects reported at the mean. Controls: size, variable input usage, other locomotive types usage, time trend, county dummies.

Table 3: Technology usage differences over time

Panel (a): Choice estimates	1(Elec. loc.)	
	Estimate:	SE:
1(Mining degree)	0.153	(0.039)
1(Mining degree)*1(1900 $\leq$ t $\leq$ 1903)	0.185	(0.062)
1(Mining degree)*1(1904 $\leq$ t $\leq$ 1907)	0.116	(0.059)
1(1900 $\leq$ t $\leq$ 1903)	-0.191	(0.041)
1(1904 $\leq$ t $\leq$ 1907)	-0.036	(0.027)
Panel (b): Imputed choice probabilities	$\Pr\big(1(\text{Elec. loc.}) = 1\big)$	
	Mining engineers:	Other managers:
1900-1903	0.216	0.553
1904-1907	0.371	0.640
1908-1914	0.407	0.560

Notes: Panel (a) compares the marginal probability effects at the mean of having a mining degree on electrical locomotive usage across three time periods. Controls include mine size, county dummies, and usage of both air and steam locomotives. Panel (b) compares the probability of electrical locomotive usage between mining school graduates and other managers during three time periods.

Table 4: Profit gains from hiring mining college educated managers

Depreciation	Wage premium[§]	Profit gain $\Delta\Pi$	$\dfrac{\Delta\Pi}{PQ}$		$\dfrac{\Delta\Pi}{\Pi(X=0)}$	
				prof. margin[‡]:	0	0.2
0.1	0	1.23 M	0.083		1.18	0.19
0.1	2	1.18 M	0.082		1.17	0.18
0.2	0	0.91 M	0.062		1.14	0.15
0.2	2	0.86 M	0.060		1.11	0.14

Note: [§]Wage premium $= \dfrac{\text{Mining degree manager wage}}{\text{Normal manager wage}} - 1$, [‡]Profit margin $= \dfrac{\mathbb{E}[\Pi|X=0, K^e=0]}{PQ}$

Appendices

A Data sources

A.1 Production and cost data

Raw 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. Geographical coordinates were obtained from Google Maps. A full list of all variables used, and their characteristics, is in table A5. 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 postoffice location and the railroad to which it is connected. This is shown in table A2. A second table provides production and cost data at the mine level, a sample is in figure A3. Thirdly, technology choices are reported in a third table, at the firm-county-year level, see A4. Fourthly, the occupational breakdown of labor is given in a fourth table, again at the firm-county-year level. Yearly prices of Pennsylvania anthracite coal are obtained from the *Statistical Abstract of the United States*.

Data cleaning

I make unique mine identifiers by tracking mine name changes over time. It happens that mines have multiple sub-units which merge or split over time. I collect these sub-units to the mine-level in order to have a unit which does not change over time. I sum all inputs and outputs of the sub-units to this mine-level. For the number of days worked, I calculate the means across sub-units. Locomotives are given 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.

A.2 Management data

As explained in the main text, I matched the managers in the production data on their full names, years and residences with Federal Census records using *Ancestry.com*. I checked whether the listed occupations were correct (e.g. 'Coal operator' or 'Mine superintendent'). Next, I retrieve birth years and match on full name and birth year with alumni records, both in my own mining college alumni lists and through *Ancestry.com*.

In principle, full manager names are given in the data. Sometimes, however, only the first letter of the second name is given, or a shorter version of the given name (Joe vs. Joseph). I encode unique manager identifiers by looking the managers up in the US census through *Ancestry.com* and comparing their location in the data to the location in the census, and the observed years in the data with their age in the census.

A list of mining schools on which I have curriculum information is in table A10. Six out of fifteen institutions are specialized in mining engineering, offering no other fields of study. Columbia University was the only private elite university offering a mining degree, all other mining schools and universities in our sample are public and generally younger. Only two schools trained mining engineers before 1885. The mean annual cohort does not exceed 30 students, and is on average 17 students. These small class sizes were customary in engineering education these days, and were considered as being beneficial for educational quality by contemporary professors (Church, 1871).

B Robustness checks

B.1 Alternative production function specifications

Complementarity between education and technology

In the production model, I assumed that locomotive output elasticities did not depend on managerial education of the manager. I test this assumption by interacting each locomotive type with the mining college indicator in the production function. The results in panel (a) of table A1 shows that these interaction terms are insignificant, however.

Complementarity between different locomotive types

Another assumption in the production model was that using different types of locomotives together did not change the output elasticity of each locomotive. I test this by interacting all locomotive types in the production function in panel (b) of table A1. All interaction terms are insignificant, and they are close to zero for the two new technologies, compressed-air and electrical locomotives.

Managerial adjustment costs

In the baseline version of the model, I assumed managers are a variable static input, without any adjustment costs. I now allow for them to be a dynamic input, such as mining locomotives. The moment conditions are still given by (6), but managerial education is now a part of the fixed inputs rather than the variable inputs: $x_{it} \in \mathbf{k_{it}}$. The estimates are in the third column of table A2 and are very similar to those in the main text.

Intensive vs. extensive margin

Throughout the model, the focus of the analysis was on the extensive margin of technology usage (which locomotives to use), not on the intensive margin (how many locomotives to use). Panel (a) of appendix table A3 shows that mining college graduates did not adopt more locomotives of any type. As the previous paragraph showed that marginal returns from locomotives did not depend on managerial education, this implies that any unmodeled marginal costs of locomotive usage (electricity prices, for instance), are also orthogonal to managerial education.

B.2 Cost dynamics

I add cost dynamics by letting the productivity transition depend on past cumulative output:

$$\omega_{it} = \tilde{g}(\omega_{it-1}, C_{it-1}) + \xi_{it} \quad \text{with} \quad C_{it} = \sum_{k=1}^{t} Q_{ik}$$

In the first column of table A2, I regress the productivity residuals of this altered model on log cumulative output and its squared term. Productivity initially increases with cumulative output (i.e. marginal costs fall), but the deeper the mine becomes, the smaller this increase becomes. The estimated output elasticities with cost dynamics are in column 2 of table A2. Accounting for cost dynamics changes the estimated output elasticities of all hauling technologies, but the difference with the baseline estimate is only statistically significant for mules. Both mules and locomotives become necessary more as the mine gets deeper.

The technology choice estimates with cost dynamics are in table A3. Again, nothing changes substantially compared to the baseline specification.

B.3 Additional controls

I add various additional controls to the production function which may drive both managerial change, technology choice and productivity. First, I include a dummy for ownership changes, as this potentially affects both management and firm performance (Braguinsky et al., 2015). I define an ownership dummy to be one if the firm listed as the mine owner changes.

Next, I include a dummy which indicates whether a mine ships at least some of its output over the railroad network to other towns, and also include the share of output which is sold in different towns. This distinguishes purely local producers from 'exporters' (even if they did not necessarily export to foreign countries).

Thirdly, I control for whether the mine managers and foremen have the same surname but a different given name, which indicates family businesses.

Finally, I take into account heterogeneity in labor. I aggregate the 18 labor occupations to three types of workers: managerial employees, skilled labor and unskilled labor. The occupational descriptions and classification is in table A9.

The production function estimates with these additional controls in the productivity transition are shown in column 4 of table A2. They are very similar to those in the main text, with no direct productivity effects of mining college graduates, but large productivity gains from using electrical locomotives. The technology choice estimates also do not change much, as shown in panel (e) of table A4.

B.4 Alternative explanations for technology choice estimates

In this section, I discuss a number of alternative explanations for the technology choice estimates which were reported in section (4.2).

Unobserved heterogeneity and reverse causality

Mines with mining engineer managers may differ from others in unobserved ways, which could be the reason why they are also more likely to use electrical locomotives. It may be that electrical locomotive adoption leads to hiring mining engineers, rather than the other way around.

I address these two issues by estimating an alternative linear specification for the technology adoption model, with mine fixed effects. I let the educational indicators enter this equation with lags and leads, in order to trace the dynamics of the decision process. Concretely, I adjust the logit model (9) to a linear probability model (9) with mine fixed effects, which I estimate for different values of $w \in \{-2, -1, 0, 1, 2\}$.

$$K^{\tau}_{it+1} = \tilde{\boldsymbol{\gamma}}^{\tau}_{Z}\mathbf{z}^{\tau}_{\mathbf{it}} + \tilde{\gamma}^{\tau}_{X} X_{it+w} + \gamma^{\tau}_{\delta}\delta_i + \nu^{\tau}_{it} \tag{9}$$

The estimated coefficients θ for $w = 0$ are in panel (b) of appendix table A3. These results are very similar to the logit specification with county fixed effects. This signals that the correlations found are not due to time-invariant latent mine characteristics. The result for each lag and lead, and its 95% confidence intervals, are plotted in figure 5. Mines that hired a mining college graduate were not different from other mines in terms of which technology they used before the hiring decision was taken. These mines were more likely than others to use any locomotive type, but this difference was not locomotive type-specific. After the mining college graduate took over, however, their mines were around 20 p.p. more likely to use electrical locomotives than others, but not just 5% more likely to use electrical or steam locomotives. This is additional evidence for the fact that mining college graduates did have an effect on the choice of which locomotive type to use, rather than that the locomotive type in use affected which type of manager to hire.

[Figure 5 here]

Spatially correlated costs

Spatially correlated unobservable costs or returns to locomotive adoption could explain the correlation between own adoption and adoption in other mines in the same market, even without information spillovers.[17] In order to explain the results in table 2, however, such cost or return spillovers should be positive for electrical locomotives, zero for steam locomotives and negative for air-powered locomotives. It is hard to think of any latent correlated cost or return shocks which are so dependent on the type of locomotive. Moreover, such cost or return shocks should not not be present for managers with mining degrees. These patterns all seem very hard to reconcile with a common cost or return hypothesis, while they are perfectly consistent with the information spillovers hypothesis.

Manager unobservables

Did mining college graduates make different technology choices because of their educational background, or because of some latent characteristics which made them self-select into mining colleges? The definitive answer to this question cannot be given in the context of this paper, as it would require observing the same managers before and after attending college - which is not possible because all mining college graduates only started managing mines after they graduated. Nevertheless, it is striking that the only observable distinction of mining college graduates in terms of actions and performance relates to the adoption of a technology about which they were educated. Moreover, it was harder to enter established Ivy League colleges such as Yale than new institutions such as Lehigh. It is therefore remarkable that only the mining college graduates were better technology choosers, while other college graduates were not. Finally, even if mining schools were merely signaling mechanisms, they still provided the opportunity to coal mines to select managers who would adopt better technologies.

[17] Electrification could seem, for instance, one reason for this. There was, however, no common electrification network at the time, so each mine had its own electricity generator, making this type of cost externality unlikely.

B.5 Firm-level instead of mine-level

In the baseline model, I allocated firm-county-year level locomotive observations to the mine-year level, as all other variables are measured at that level. In this robustness check, I bring the dataset to the firm-county-year level. Managerial education is now no long a vector of two dummy variables, but a vector with the share of mining college and other college graduates. I calculate these shares by taking simple averages over all mines in the same firm within the same county. The results are in panel (c) of table A3 and figure A6, and reach very similar conclusions as the baseline model.

C Productivity and profit gains

C.1 Productivity gains

I compare expected output when mines are headed by a mining engineer with expected output when they are not. The only difference between mining engineers and other managers is their probability of using electrical locomotives, which increase output.

$$\begin{aligned}\frac{\mathbb{E}(Q|X=1)}{\mathbb{E}(Q|X=0)} &= \frac{\mathbb{E}(Q|K=1)\mathbb{E}(K=1|X=1)+\mathbb{E}(Q|K=0)\mathbb{E}(K=0|X=1)}{\mathbb{E}(Q|K=1)\mathbb{E}(K=1|X=0)+\mathbb{E}(Q|K=0)\mathbb{E}(K=0|X=0)} \\ &= \frac{Q(1+\beta^e)\mathbb{E}(K=1|X=1)+Q\mathbb{E}(K=0|X=1)}{(1+\beta^e)Q\mathbb{E}(K=1|X=0)+Q\mathbb{E}(K=0|X=0)}\end{aligned}$$

Denote $\mathbb{E}(K=1|X=1)\equiv p_1^1, \mathbb{E}(K=1|X=0)\equiv p_0^1, \mathbb{E}(K=0|X=1)\equiv p_1^0$ and $\mathbb{E}(K=0|X=0)\equiv p_0^0$

$$\Leftrightarrow \frac{\mathbb{E}(Q|X=1)}{\mathbb{E}(Q|X=0)} = \frac{(1+\beta^e)p_1^1+p_1^0}{(1+\beta^e)p_0^1+p_0^0}$$

C.2 Profit gains

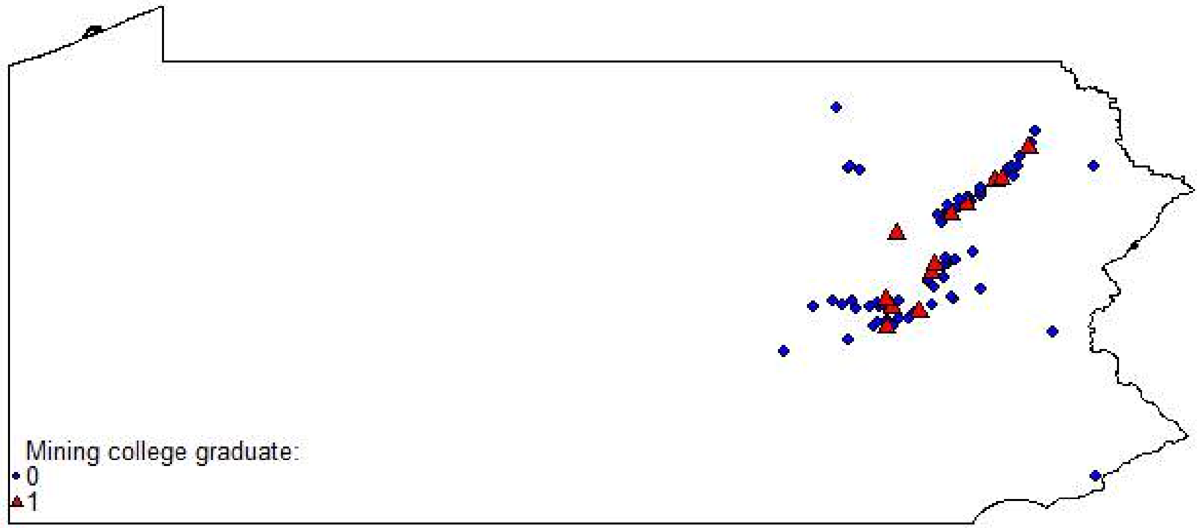
I now do a similar exercise as above, but in monetary terms. I deduct locomotive fixed costs from expected profits in case mines use electrical locomotives, and deduct mining engineering salaries in case a manager with such a degree is used.

$$
\begin{aligned}
\Delta\Pi &\equiv \mathbb{E}(\Pi|X=1) - \mathbb{E}(\Pi|X=0) \\
&= \mathbb{E}(\Pi|K=1, X=1)p_1^1 + \mathbb{E}(\Pi|K=0, X=1)p_1^0 - \mathbb{E}(\Pi|K=1, X=0)p_0^1 + \mathbb{E}(\Pi|K=0, X=0)p_0^0
\end{aligned}
$$

With:

$$
\begin{cases}
\mathbb{E}(\Pi|K=1, X=1) & = PQ\beta^e - \varphi^e - \varphi^x \\
\mathbb{E}(\Pi|K=0, X=1) & = -\varphi^x \\
\mathbb{E}(\Pi|K=1, X=0) & = PQ\beta^e - \varphi^e \\
\mathbb{E}(\Pi|K=0, X=0) & = 0
\end{cases}
$$

Figure A1: Map with mining towns



Note: Red triangles are towns in which there was at least one mining college graduate managing an anthracite mine between 1900 and 1914. Blue circles are towns where this was not the case.

Figure A2: Data example: ownership, management and location

TABLE I—Showing Names of Operators, Railroads, etc., etc., and Location of Collieries in the Third Anthracite District for the year 1901.

Names of Operators and Collieries.	County.	Name of General Superintendent.	P. O. Address.	Name of Superintendent.	P. O. Address.	Railroad to Mine.
Pennsylvania Coal Company.		Sidney Williams,	Dunmore,			Erie and Wyoming.
Barnum No. 1 shaft,	Luzerne,	Sidney Williams,	Dunmore,	John Popling and	Pittston,	Erie and Wyoming.
Barnum No. 2 shaft,	Luzerne,	Sidney Williams,	Dunmore,	John W. Reid	Moosic,	Erie and Wyoming.
Barnum No. 3 shaft,	Luzerne,	Sidney Williams,	Dunmore,	John W. Reid,	Moosic,	Erie and Wyoming.
Laws shaft,	Luzerne,	Sidney Williams,	Dunmore,	John W. Reid,	Moosic,	Erie and Wyoming.
No. 13 shaft,	Lackawanna,	Sidney Williams,	Dunmore,	John W. Reid,	Moosic,	Erie and Wyoming.
No. 9 shaft,	Luzerne,	Sidney Williams,	Dunmore,	John W. Reid,	Moosic,	Erie and Wyoming.
No. 10 shaft,	Luzerne,	Sidney Williams,	Dunmore,	John W. Reid,	Moosic,	Erie and Wyoming.
No. 10 Jr. shaft,	Luzerne,	Sidney Williams,	Dunmore,	John W. Reid,	Moosic,	Erie and Wyoming.
No. 1 shaft,	Luzerne,	Sidney Williams,	Dunmore,	John W. Reid,	Moosic,	Erie and Wyoming.
No. 8 shaft,	Luzerne,	Sidney Williams,	Dunmore,	John W. Reid,	Moosic,	Erie and Wyoming.
No. 7 shaft,	Luzerne,	Sidney Williams,	Dunmore,	John W. Reid,	Moosic,	Erie and Wyoming.
No. 4 shaft,	Luzerne,	Sidney Williams,	Dunmore,	John W. Reid,	Moosic,	Erie and Wyoming.
Hoyte shaft,	Luzerne,	Sidney Williams,	Dunmore,	John W. Reid,	Moosic,	Erie and Wyoming.
No. 6 shaft,	Luzerne,	Sidney Williams,	Dunmore,	John W. Reid,	Moosic,	Erie and Wyoming.
No. 5 shaft,	Luzerne,	Sidney Williams,	Dunmore,	John W. Reid,	Moosic,	Erie and Wyoming.
No. 11 shaft,	Luzerne,	Sidney Williams,	Dunmore,	John W. Reid,	Moosic,	Erie and Wyoming.
No. 14 shaft,	Luzerne,	Sidney Williams,	Dunmore,	John W. Reid,	Moosic,	Erie and Wyoming.
No. 14 tunnel,	Luzerne,	Sidney Williams,	Dunmore,	John W. Reid,	Moosic,	Erie and Wyoming.
No. 6 washery,	Luzerne,	Sidney Williams,	Dunmore,	John W. Reid,	Moosic,	Erie and Wyoming.
No. 8 washery,	Luzerne,	Sidney Williams,	Dunmore,	John W. Reid,	Moosic,	Erie and Wyoming.
Lehigh Valley Coal Company.						
Prospect shaft,	Luzerne,	W. A. Lathrop,	Wilkes-Barre,	Eli P. Conner,	Wilkes-Barre,	Lehigh Valley Railroad.
Oakwood shaft,	Luzerne,	W. A. Lathrop,	Wilkes-Barre,	Eli P. Conner,	Wilkes-Barre,	Lehigh Valley Railroad.
Midvale slope,	Luzerne,	W. A. Lathrop,	Wilkes-Barre,	Eli P. Conner,	Wilkes-Barre,	Lehigh Valley Railroad.
Wyoming Hillman slope,	Luzerne,	W. A. Lathrop,	Wilkes-Barre,	Eli P. Conner,	Wilkes-Barre,	Lehigh Valley Railroad.
Wyoming shaft,	Luzerne,	W. A. Lathrop,	Wilkes-Barre,	Eli P. Conner,	Wilkes-Barre,	Lehigh Valley Railroad.
Henry shaft,	Luzerne,	W. A. Lathrop,	Wilkes-Barre,	Eli P. Conner,	Wilkes-Barre,	Lehigh Valley Railroad.
Maltby shaft,	Luzerne,	W. A. Lathrop,	Wilkes-Barre,	Eli P. Conner,	Wilkes-Barre,	Lehigh Valley Railroad.

Figure A3: Data example: production, sales and inputs

TABLE II—Gives the total number of tons of coal mined in each colliery, number of days worked, number of employes, number of persons killed and injured, number of kegs of powder, etc., used in the Third Anthracite District for the year ending December 31, 1900.

Names of Operators and Collieries.	County.	Shipments of coal in tons by rail or otherwise.	Number of tons used for steam and heat at colliery.	Sold to local trade and used by employes—tons.	Total production of coal in tons.	Number days worked.	Number persons employed.	Number fatal accidents.	Number non-fatal accidents.	Number kegs powder used.	Number pounds of dynamite used.	Number horses and mules.
Pennsylvania Coal Company.												
Barnum No. 1, 2 and 3 shafts,	Luzerne,	252,138.16	7,509.19	....	259,648.15	159.50	769	3	4	9,682	511	55
Laws and No. 13 shafts,	Luzerne,	183,273.15	4,863.04	....	188,136.19	162.75	552	1	2	4,646	1,181	49
Shafts No. 9, 10 and 10 Jr.,	Luzerne,	185,192.04	11,488.14	....	196,680.18	161	800	......	2	5,825	823	71
Shafts No. 1 and 8,	Luzerne,	128,717.18	2,848.00	....	131,565.18	160.50	385	1	......	3,773	819	43
Shafts No. 4, 7 and Hoyte,	Luzerne,	215,247.19	11,805.07	....	227,053.06	141.50	885	2	3	8,087	1,285	72
Shafts No. 5, 6 and 11,	Luzerne,	239,683.16	10,962.01	....	250,645.17	141.50	944	2	10	9,412	6,340	67
No. 14 shaft and tunnels,	Luzerne,	197,369.04	8,509.15	....	205,878.19	154.50	628	3	2	5,537	1,414	57
No. 6 washery,	Luzerne,	55,781.13	2,867.16	....	58,649.09	155	36	......	......	....	....	2
No. 8 washery,	Luzerne,	76,488.02	2,978.07	....	79,466.09	162	60	......	......	....	....	....
Total,	....	1,533,893.07	*63,833.03	....	1,597,726.10	155.33	5,059	12	23	46,912	12,373	416
Lehigh Valley Coal Company.												
Prospect and Oakwood shafts,	Luzerne,											
Wyoming and Midvale slopes,	Luzerne,	242,619.18	23,775.00	6,273.02	272,668.00	146.75	816	2	5	4,839	37,707	90

Figure A4: Data example: technology usage

TABLE II—Continued.

Name of Operators.	County.	Number of Boilers. Cylindrical.	Horse power.	Tubular.	Horse power.	Total horse power.	Locomotives. Steam.	Air.	Electric.	Number steam engines of all classes.	Total horse power.	Number pumps delivering water to surface.	Capacity in gallons per minute.	Quantity delivered to surface per minute—gallons.	Number electric dynamos.	Number air compressors.
Pennsylvania Coal Company, ...	Luzerne, ...	35	1,400	51	7,605	9,005	10	3	......	133	15,041	30	23,792	11,290	......	6
Lehigh Valley Coal Company, ...	Luzerne, ...	27	837	34	6,592	7,429	7	......	......	92	13,661	26	19,033	13,080	3	4
Butler Mine Company, Limited, ...	Luzerne, ...	24	280	6	440	720	2	......	......	16	350	14	3,000	800	......	......
Delaware, Lacka. & West. R. R. Co.	Luzerne, ...	41	1,000	7	905	1,905	......	......	......	31	1,674	16	5,900	2,950	2	......
Temple Iron Company, ...	Luzerne, ...	30	800	20	4,675	5,475	4	......	3	72	3,600	10	11,450	5,450	4	6
Miscellaneous Coal Companies.																
Seneca Coal Company, ...	Luzerne, ...	27	1,135	1	100	1,235	4	......	......	16	716	11	3,500	1,500	......	1
Old Forge Coal Company, ...	Luzerne, ...	6	200	4	450	650	......	......	......	7	386	......	......	800	......	......
Delaware and Hudson Canal Co., ...	Luzerne, ...	15	450	5	750	1,200	......	......	......	19	1,350	3	1,760	1,250	......	1
Raub Coal Company, ...	Luzerne, ...	13	440	5	610	1,050	1	......	......	14	793	1	500	300	......	1
John C. Haddock, ...	Luzerne, ...	15	288	10	1,070	1,358	1	......	......	36	1,902	3	1,580	1,208	......	1
Clear Spring Coal Company, ...	Luzerne, ...	8	275	5	750	1,025	......	......	......	8	800	2	1,200	600	......	......
Florence Coal Company, Limited, ...	Luzerne, ...	14	350	3	150	500	1	......	......	12	366	6	2,600	1,800	......	......
W. G. Payne and Company, ...	Luzerne, ...	24	444	2	200	644	......	......	......	12	1,453	1	2,000	1,200	......	1
Traders' Coal Company, ...	Luzerne, ...	8	160	2	190	350	......	......	......	10	315	2	600	400	......	......

Figure A5: Locomotive costs

A. Compressed-air plant operating two main hauling locomotives, each weighing 30,000 pounds, and five gathering locomotives, each weighing 8,000 pounds, a total weight of 100,000 pounds. The cost was as follows:—

	Dollars.	£
A three-stage air compressor and a compound steam-engine	5,300	1,104
5,600 feet of pipes, 5 inches in diameter ...	5,600	1,167
3,100 ,, ,, 2½ ,, ,, ...	1,700	354
1,000 ,, ,, 1¼ ,, ,, ...	300	63
Two locomotives, each weighing 30,000 pounds	6,000	1,250
Five ,, ,, 8,000 ,,	10,000	2,083
Two boilers, each of 80 horsepower	1,000	208
Installation	4,000	833
Totals	33,900	7,062

B. Electric plant of four locomotives, each weighing 26,000 pounds, a total weight of 104,000 pounds. The cost was as follows:—

	Dollars.	£
Generator, producing 225 kilowatts	3,900	812
Engine, 500 horsepower	4,800	1,000
Boilers, 600 horsepower	4,400	917
Foundations, piping, etc.	1,500	312
Wiring	8,000	1,667
Four locomotives, each weighing 26,000 pounds	9,500	1,979
Totals	32,100	6,687

C. Electric plant, comprising two locomotives, each weighing 26,000 pounds, a total weight of 52,000 pounds. The cost was as follows:—

	Dollars	£
Generators, etc.	24,000	5,000

Source: Randolph (1905)

Figure A6: Education and locomotive adoption: firm-county-year level

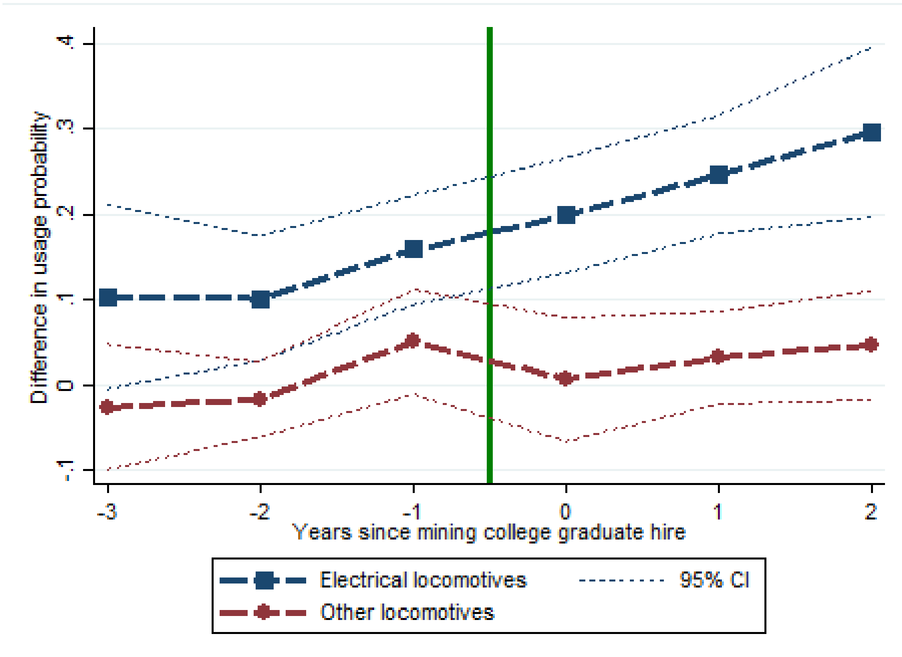


Table A1: Production function: additional interaction effects

Panel (a): Managers and returns from locomotives	log(Output)	
	Estimate:	SE:
1(Air locomotive)*1(Mining degree)	0.060	(0.178)
1(Elec. locomotive)*1(Mining degree)	-0.061	(0.205)
1(Steam locomotive)*1(Mining degree)	-0.189	(0.268)
Observations	3,221	
R-squared	0.895	
Panel (b): Interaction effects between locomotives	log(Output)	
	Estimate:	SE:
1(Air loc.)*1(Elec. loc.)	-0.011	(0.040)
1(Steam. loc.)*1(Elec. loc.)	0.122	(0.092)
1(Steam loc.)*1(Air loc.)	0.100	(0.331)
1(Steam loc.)*1(Air loc.)*1(Elec. loc.)	0.006	(0.001)

Notes: Panel (a) extends the production function by interacting locomotive dummies with the mining college dummy. Panel (b) reports the interaction effects between all locomotive types. Bootstrapped standard errors with 50 iterations are between parentheses.

Table A2: Production function: robustness checks

Specification:	Cost dynamics log(Output)		Dyn. managers log(Output)		Additional controls log(Output)	
	Estimate:	SE:	Estimate:	SE:	Estimate:	SE:
1(Mining degree)	-0.093	(0.500)	-0.019	(0.183)	-0.020	(0.188)
log(Labor)	0.481	(0.249)	0.756	(0.198)	0.462	(0.197)
log(Materials)	0.033	(0.034)	0.016	(0.022)	0.130	(0.069)
log(Mules)	0.500	(0.213)	0.185	(0.159)	0.260	(0.137)
1(Elec. loc.)	0.154	(0.076)	0.118	(0.050)	0.166	(0.070)
1(Steam loc.)	0.100	(0.054)	0.041	(0.042)	0.073	(0.057)
1(Air loc.)	0.067	(0.118)	0.037	(0.067)	0.086	(0.057)
Scale parameter	1.241	(0.577)	1.134	(0.176)	1.156	(0.244)
Observations	2,991		3,221		2,316	
R-squared	0.875		0.893		0.890	

Notes: Bootstrapped standard errors in parentheses, 50 iterations

Table A3: Technology choice: robustness checks

Panel (a) Intensive margin	log(Elec. loc.)		log(Steam loc.)		log(Air loc.)	
	Estimate:	SE:	Estimate:	SE:	Estimate:	SE:
1(Mining degree)	-0.246	(0.171)	0.309	(0.106)	-0.676	(0.188)
1(Other degree)	-0.418	(0.337)	0.046	(0.210)	-0.709	(0.831)
Observations	374		891		164	
R-squared	0.205		0.326		0.315	
Panel (b) Linear probability model	1(Elec. loc.)		1(Steam loc.)		1(Air loc.)	
	Estimate:	SE:	Estimate:	SE:	Estimate:	SE:
1(Mining degree)	0.428	(0.040)	0.358	(0.283)	0.071	(0.096)
1(Other degree)	-0.010	(0.174)	-0.238	(0.098)	0.094	(0.097)
Observations	3,345		3,345		3,345	
R-squared	0.364		0.345		0.255	
Panel (c) Firm-level model	1(Elec. loc.)		1(Steam loc.)		1(Air loc.)	
	Estimate:	SE:	Estimate:	SE:	Estimate:	SE:
1(Mining degree)	0.395	(0.040)	-0.016	(0.134)	0.102	(0.071)
1(Other degree)	0.072	(0.054)	0.077	(0.078)	0.114	(0.127)
Average usage	0.206		0.478		0.090	
Observations	3,339		3,339		3,339	
R-squared	0.364		0.345		0.255	

Notes: Controls include productivity and time trend

Table A4: Technology choice: robustness checks (continued)

Panel (d) Cost dynamics	log(Elec. loc.)		log(Steam loc.)		log(Air loc.)	
	Estimate:	SE:	Estimate:	SE:	Estimate:	SE:
1(Mining degree)	0.306	(0.037)	0.154	(0.025)	0.075	(0.046)
1(Other degree)	-0.082	(0.187)	-0.282	(0.104	0.004	(0.092)
Observations	3,345		3,345		3,345	
R-squared	0.317		0.157		0.135	

Panel (e) Additional controls	log(Elec. loc.)		log(Steam loc.)		log(Air loc.)	
	Estimate:	SE:	Estimate:	SE:	Estimate:	SE:
1(Mining degree)	0.297	(0.044)	0.082	(0.045)	0.125	(0.072)
1(Other degree)	0.327	(0.043)	-0.281	(0.030)	-0.280	(0.038)
Observations	1,694		1,694		1,694	
R-squared	0.330		0.163		0.215	

Notes: Controls include productivity and time trend

Table A5: List of variables

	Unit of measurement	Level	Frequency	Flow/stock
(a) Output and sales				
Coal extracted	short tons (2000 lbs)	Mine	Annual	Flow
Exports	short tons	Mine	Annual	Flow
Local sales	short tons	Mine	Annual	Flow
Output reused as input	short tons	Mine	Annual	Flow
(b) Inputs				
Employees	average counts	Mine	Annual	Stock
Days worked	counts	Mine	Annual	Flow
Powder	kegs of 25lbs	Mine	Annual	Flow
Mules	counts	Mine	Annual	Stock
(c) Management				
Superintendent name	string	Mine	Annual	N/A
Foreman name	string	Mine	Annual	N/A
E. M. degree	binary	Manager	Annual	N/A
College degree	binary	Manager	Annual	N/A
Superintendent Age	string	Manager	Annual	N/A
(d) Technology				
Locomotives	counts	Firm-county	Annual	Stock
Compressed-air locomotives	counts	Firm-county	Annual	Stock
Electrical locomotives	counts	Firm-county	Annual	Stock
Steam locomotives	counts	Firm-county	Annual	Stock
Location	coordinates	Village	Time-invariant	N/A

Table A6: Mine summary statistics

	Mean	Std. Dev.	Min.	Max.	Observations
(a) Output:					
Coal extracted, Mtons	0.21	0.19	0	3.52	4730
Output shipped, share	0.87	0.15	0	1	4572
(b) Variable inputs:					
Employees	512.71	427.8	0	6595	4730
Powder, 1000 kegs	52.42	138.24	0	1748.43	4730
Coal inputs, Ktons	20.92	21.85	0	494.48	4730
Dynamite, 1000 pounds	26.96	47.83	0	654.89	4730
(c) Capital inputs:					
Mules	51.03	43.01	0	276	4730
Mining locomotives	25.37	36.15	0	172	4730
Machinery horsepower x 1000	29.67	44.18	0	212.92	4730
(d) Managerial inputs:					
Manager has college mining degree	0.07	0.25	0	1	4730
Manager has other college degree	0.01	0.1	0	1	4730

Table A7: Manager summary statistics

	Mean	Std. Dev.	Min.	Max.	Observations
(a) Age	46.02	11.34	19.5	77.40	309
if mining degree	33.53	6.98	25.87	43.5	7
if other degree	31.84	5.76	20	43	11
if no degree	46.85	11.06	19.5	77.40	291
(b) Years in firm	1.56	1.97	0	8.54	308
if mining degree	1.33	1.56	0	4.48	7
if other degree	1.11	1.67	0	4.47	11
if no degree	1.58	1.99	0	8.54	290
(c) # Mines managed	2.71	5.17	1	37.32	309
if mining degree	5.72	9.62	1	26.85	7
if other degree	1.2	0.43	1	2.33	11
if no degree	2.7	5.12	1	37.32	291
(d) Output, Mtons	0.12	0.12	0	0.66	310
if mining degree	0.22	0.18	0.06	0.63	7
if other degree	0.13	0.22	0	0.66	11
if no degree	0.12	0.11	0	0.63	292

Table A8: Mining school curricula

Subject	Course examples	% Credits	Usual phase
Science	Mathematics, Chemistry, ...	33.7	Freshman / sophomore
Mining engineering	Drilling, Mine construction, Geology, ...	34.3	Junior / senior
Other engineering	Electricity, Mechanics, ...	24.3	Sophomore / junior
Languages	Foreign languages, writing, retorics	4.7	Freshman
Thesis	Master project	2.0	Senior
Management	Mining economics, mining law, contracts, ...	1.0	Senior

Table A9: Occupation descriptions

Occupation	Location	Type	Description
Slatepicker	Above-ground	Unskilled	Breaking and sorting coal
Mining laborer	Underground	Unskilled	Assisting miner by cleaning and hauling
Driver	Underground	Unskilled	Driving mules
Doorboy	Underground	Unskilled	Opening and closing doors
Blacksmith	Above-ground	Skilled	Blacksmithing
Engineer	Above-ground	Skilled	Machine maintenance
Miner	Underground	Skilled[18]	Cutting coal
Superintendent	Above-ground	Manager	Managing entire mine
Outside Foreman	Above-ground	Manager	Managing above-ground operations
Inside Foreman	Underground	Manager	Managing underground operations
Fire boss	Underground	Manager	Responsible for managing explosions and fire risk

Table A10: School data sources

School name	Document name	Type	Years
Arizona School of Mines	Alumnal Record of the University of Arizona	Alumni record	1916
Colorado School of Mines	Quarterly of the Colorado School of Mines	Bulletin	1908, 1912-1914
Columbia College of Mines	Catalogue of Columbia University	Catalogue	1867-1914
Michigan College of Mines	Graduates of the Michigan College of Mines	Alumni record	1910
	Year Book of the Michigan College of Mines	Catalogue	1910-1914
Missouri School of Mines	School of Mines and Metallurgy Bulletin	Catalogue	1914
Montana School of Mines	Annual Catalogue	Catalogue	1908-1914
Nevada Mackay School of Mines	Register of the University of Nevada	Catalogue	1908-1914
New Mexico School of Mines	Register of the New Mexico School of Mines	Catalogue	1909-1914
Penn State School of Mines	Alumni Directory	Alumni record	1913
South Dakota School of Mines	Annual Catalogue	Catalogue	1912
University of Minnesota	School of Mines Announcement	Catalogue	1897-1914
West Virginia University	Register of Faculty, Alumni and Students	Alumni record	1920