Labor Market Power and Factor-Biased Technology Adoption

Michael Rubens*

July 16, 2024

Abstract

Although monopsony power can induce deadweight loss, it can also incentivize employers' technology adoption by reducing holdup. Using a structural model that features these two opposing forces, I quantify the net welfare effects of employer monopsony power in the late-19th-century Illinois coal mining industry. I find that a one-standard-deviation increase in employer power would have increased usage of mechanical coal cutters, a new technology, by 14%, thereby lowering marginal costs. However, output would still have decreased by 23%, meaning that the deadweight-loss effect dominates the marginal-cost reduction due to additional investment. Ignoring endogenous investment leads to overestimating the consumer and total welfare costs of monopsony power by 7% and 15%, respectively.

Keywords: Monopsony, Buyer Power, Investment, Technological Change

JEL codes: L11, L13, J42, N51

---

*UCLA, Department of Economics. Contact: rubens@econ.ucla.edu.
This paper previously circulated under the titles "Monopsony Power and Factor-Biased Technology Adoption" and "Oligopsony Power and Factor-Biased Technology Adoption." I thank Jan De Loecker, Frank Verboven, Jo Van Biesebroeck, Chad Syverson, Otto Toivanen, Mert Demirer, Scott Stern, Penny Goldberg, Steve Berry, Suresh Naidu, and Daniel Haanwinckel, for feedback and discussion, as well as participants at various seminars and conferences. Financial support from the Belgian American Educational Foundation, the Fulbright program, the Research Foundation Flanders (FWO grant 11A4618N), the European Research Council (ERC grant 816638), and the Terry and Susan Kramer Family Foundation is gratefully acknowledged.

1 Introduction

There is growing empirical evidence for the existence of "monopsony" or "buyer power" across various industries, countries, and types of factor markets.[1] Recent research on the welfare consequences of "classical" monopsony or oligopsony power, such as Berger et al. (2022) and Lamadon et al. (2022), has typically assumed firms' technology choices and investments to be exogenous to the degree of monopsony power. However, employers' investment decisions are endogenous if employers need some degree of monopsony power to recover the fixed costs incurred when investing in human and/or physical capital.

In this paper, I examine the welfare effects of monopsony power by disentangling these two opposing forces. On one hand, monopsony power leads to deadweight loss, because it makes employers cut back on input usage in order to push down input prices. On the other hand, monopsony power can incentivize the adoption of new productivity-enhancing technologies because it allows employers to appropriate more of the rents created by this technology adoption. This is a version of the classical holdup problem of Williamson (1971). The net effect of monopsony power on equilibrium output, consumer surplus, and total welfare is ambiguous, as it depends on the relative magnitude of the deadweight-loss and technology-adoption channels. This trade-off between anticompetitive distortions and endogenous human-capital investment plays an important role in various debates around labor market power, such as regulation of noncompete agreements (Starr, Prescott, & Bishara, 2021; Shi, 2023), and the role of buyer power in horizontal merger control (Hemphill & Rose, 2018; Loertscher & Marx, 2019)

I start this paper with a theoretical model of labor demand and supply that allows for both monopsony-induced deadweight loss and endogenous investment. Although the model is written in terms of employers and employees, it applies to vertical relationships between buyers and sellers more broadly. The model features employers that face upward-sloping

[1] See literature reviews by Ashenfelter et al. (2010) and Manning (2011), and recent papers by, among many others, Naidu et al. (2016); Berger et al. (2022); Morlacco (2017); Lamadon et al. (2022); Kroft et al. (2020); Rubens (2023); Chambolle et al. (2023).

labor supply curves and negotiate over wages with workers using a linear wage contract in a "Nash-in-Nash" bargaining model (Horn & Wolinsky, 1988). The model collapses to the classical monopsony model, in which employers unilaterally post wages, if bargaining power lies fully on the employers' side. Employers combine different labor types to produce output using a constant elasticity of substitution (CES) production function, and they choose whether to adopt a new technology. This technology decision determines the employers' Hicks-neutral and factor-specific productivity levels and is incurred at a fixed cost. Employers sell their output in product markets where they face a downward-sloping goods demand curve. Simulating this model, I find that the relationship between equilibrium output and employer power is monotonically decreasing when assuming exogenous technology usage, as only the deadweight loss mechanism is at play. However, when allowing for endogenous technology adoption, the output-employer power relationship assumes an inverted U-shape, as the deadweight-loss and endogenous-investment mechanisms counteract each other. The output-maximizing bargaining weight depends on the relative size of the curvature of the labor supply curve and of the productivity effects of the technology.

Given that the equilibrium effect of employer power on output is ambiguous, I perform an empirical application to quantify the relative importance of the deadweight loss and endogenous investment mechanisms. I study how the mechanization of the Illinois coal mining industry between 1884 and 1902 was affected by employers' market power over their employees. There are two main reasons why this industry and period provides a unique setting to study the relationship between buyer power and innovation. First, the introduction of coal-cutting machines in the United States in 1882 provides a large observed technological shock. Second, 19th-century Illinois coal mining towns are a textbook example of monopsonistic labor markets, with geographically isolated local labor markets, which makes it likely that employers had some market power over their workers.

I implement an empirical version of the model to fit the coal-industry setting; I estimate it using a novel, uniquely rich archival dataset on mine-level production, coal prices, input quantities and prices, technology usage, and geological data. I rely on observed variation in the thickness of coal veins as cost shifters to estimate coal demand, and on international coal-price shocks and exporting data to estimate labor supply. Identifying the production function relies, as usual, on timing assumptions on input choices in function of both Hicks-neutral and labor-augmenting productivity shocks. In line with anecdotal historical evidence, I find that cutting machines were unskill-biased, similar to many other technologies that were developed throughout the 19th century (Mokyr, 1990; Goldin & Katz, 2009).

Using the estimated production, labor supply, and coal demand models, I estimate the relative bargaining weights of the employers and the coal miners. I find that employers and coal miners had roughly equal bargaining weights on average, although their relative bargaining power fluctuated over time. I find that a series of large strikes in 1898 led to a relative increase of union bargaining power at striking mines. Finally, I estimate the fixed costs of cutting-machine usage by comparing the variable profit gains from machine adoption to the observed machine usage rates.

Using the estimated model, I numerically solve for output and input quantities, coal prices, miner wages, and machine usage in the observed equilibrium, and in two counterfactual equilibria. In the first, I examine how all equilibrium outcomes and welfare would have changed if employer power would have increased by one standard deviation. In the second, I examine the same counterfactual exercise, but I hold cutting-machine usage fixed. This second counterfactual is informative about the relative importance of the endogenous machine-usage mechanism compared to the deadweight-loss effect. I find that increased employer power would increase cutting-machine usage by 14% by the end of the panel in 1902. Hence, both total factor productivity (TFP) and skill-augmenting productivity would have increased in response to an increase in employer power. However, equilibrium output would have decreased by 23%, because the deadweight-loss effect dominates the marginal cost decrease of increased cutting-machine adoption. Endogenous capital investment alters the welfare effects of increased labor market power. If capital investment would be exogenous, the increase in employer power would reduce consumer, worker, and total welfare by 21.2%, 37.3 %, and 12.5%, respectively. Allowing for endogenous capital investment reduces these welfare losses to 19.8%, 35.6%, and 10.9%. Hence, assuming exogenous capital investment leads to overestimating the total welfare losses from monopsony power by 15.4%.

This paper contributes to three distinct literatures. First, I contribute to the literature on the welfare consequences of monopsony power by considering endogenous technological change. Existing work on classical monopsony/oligopsony power usually focuses on deadweight loss and/or on misallocation (Berger et al., 2022; Lamadon et al., 2022; Jarosch, Nimczik, & Sorkin, 2024; Morlacco, 2017; Rubens, 2023) but keeps technology choices fixed. In contrast, I show that endogenous technology choices present an additional channel through which input market power affects welfare. On the other hand, there is work on holdup in a class of labor search models (Acemoglu & Shimer, 1999; Shi, 2023) and in efficient bargaining models (Abowd & Lemieux, 1993; Van Reenen, 1996; Menezes-Filho & Van Reenen, 2003), but these models do not feature monopsony-induced deadweight loss. I contribute by empirically investigating and comparing monopsony-induced deadweight loss to the inefficiency created by holdup, in order to obtain a net effect of employer power on equilibrium output and welfare.

Second, this paper builds on the vertical relations literature. In contrast to existing work on holdup (Williamson, 1971; Joskow, 1987; Zahur, 2022), I allow for monopsony distortions by including upward-sloping marginal-cost curves of the suppliers, and I also use a model with multiple substitutable inputs, rather than a single input. In contrast to the literature that studies the effects of buyer power on technology choices of suppliers (Just & Chern, 1980; Huang & Sexton, 1996; Köhler & Rammer, 2012; Parra & Marshall, 2024), I focus on its effects on the technology choices of the buyers.

Third, this paper relates to the literature on labor market power during the late 19th century, such as Boal (1995), Naidu and Yuchtman (2017), and Delabastita and Rubens (2024), and on technological change during the same period (Goldin & Katz, 1998; Atack, Bateman, & Margo, 2008; Katz & Margo, 2014; Buyst, Goos, & Salomons, 2018; Hornbeck, Hsu, Humlum, & Rotemberg, 2024). I bring these two literatures together by studying how employer power during the late 19th century affected the adoption of directed technologies, and by quantifying the resulting welfare effects on consumers, producers, and workers.

The rest of this paper is structured as follows. Section 2 contains the theoretical model. Section 3 discusses the data, the industry background, and the empirical model. Section 4 covers the estimation of the model and the counterfactual simulations. Section 5 concludes.

2 Labor Market Power and Investment

I start with a theoretical model to examine the welfare effects of labor market power when allowing for both deadweight losses and endogenous investment.

2.1 Primitives

Firms f produce output Q_f using two variable inputs, H_f and L_f. I consider a firm to be an employer, and the inputs to be high- and low-skilled labor, but the model can equivalently be interpreted as wholesalers f that purchase various inputs from manufacturers. I rely on the CES production function (1), in which both inputs are substitutable at a constant elasticity σ. For simplicity, I assume constant returns to scale, but I relax this in Appendix C.1. Skill-augmenting productivity is denoted as A_f, the low-skilled labor coefficient as β^l, and Hicks-neutral productivity as Ω_f.

$$Q_f = \left((A_f(K_f)H_f)^{\frac{\sigma-1}{\sigma}} + \beta^l L_f^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} \Omega_f(K_f) \tag{1}$$

Firms choose their capital stock K, which enters the production function in two ways. First, capital changes the skill-augmenting productivity term $A_f(K_f)$. Second, capital affects Hicks-neutral productivity $\Omega_f(K_f)$. Firms sell their product at a price P_f. Consumer demand for the good is given by a standard horizontal differentiation demand system, with an average industry-level price P_0, unobserved firm characteristics ξ_f, and a constant demand elasticity η:

$$Q_f = \left(\frac{P_f}{P_0}\right)^{\eta} \xi_f \tag{2}$$

High-skilled workers have an outside option Z_f, which they can earn when choosing not to be employed at firm f. I allow this outside option Z_f to be an upward-sloping curve, with constant inverse elasticity ψ, as shown in Equation (3). Firms are differentiated from the worker's perspective through an amenity term ζ_f. The average industry wage is equal to W_0. In contrast to models with a constant outside-option value, such as in Abowd and Lemieux (1993), an increasing outside option generates an upward-sloping labor supply curve, which allows for the possibility of monopsonistic behavior by employers. An increasing outside-option curve can be rationalized by the fact that workers are heterogeneous in terms of their outside options, and that firms cannot wage-discriminate in function of these worker-specific outside options, as in Berger et al. (2022). Hence, the labor supply curve to each firm is upward-sloping: hiring an extra worker requires a higher wage to compensate the higher outside option of the marginal worker compared to the outside options of the inframarginal workers:

$$Z_f = \frac{W_0}{1+\psi}\left(\frac{H_f}{\zeta_f}\right)^{\psi} \tag{3}$$

In contrast, the outside option of low-skilled labor is assumed to be equal to a constant V. This implies that firms pay low-skilled workers a uniform wage V and that low-skilled labor supply is perfectly elastic.

High-skilled workers are unionized at the firm level. The utility of the labor union at firm f is denoted as Π^u_f, which is defined as the difference between high-skilled earnings and the outside option to high-skilled workers:

$$\Pi^u_f = (W_f - Z_f)H_f$$

Employer profits are denoted as Π_f^d:

$$\Pi_f^d = P_f Q_f - W_f H_f - V L_f - \phi K_f$$

2.2 Strongly Efficient Bargaining

Behavior and Equilibrium

For illustration purposes, I start with a version of the "strongly efficient" bargaining model Abowd and Lemieux (1993), in which a labor union and employers collectively bargain over both employment and wages, before moving to the actual model in which they bargain only over wages. The strongly efficient bargaining model is a useful benchmark to compare the results from the full model against because the strongly efficient model does not feature any monopsony distortions, only endogenous technology choices.

In this model, employers and unions bargain over both employment and wages in a Nash bargaining protocol, with γ_f indicating union bargaining power:

$$\max_{H_f, L_f, W_f} (\Pi_f^u)^{\gamma_f} (\Pi_f^d)^{1-\gamma_f}$$

The model implies that the union and employers jointly optimize joint profits and split the surplus according to the bargaining parameters γ_f. Taking the first-order condition for the high-skilled wage results in:

$$W_f = (1 - \gamma_f) Z_f + \gamma_f \Big(\frac{P_f Q_f - V L_f}{H_f} \Big) \tag{4}$$

The first-order conditions for the labor inputs are given by:

$$P_0 \Big(\frac{1+\eta}{\eta} \Big) Q_f^{\frac{1}{\eta}} \Big(\frac{Q_f}{H_f} \Big)^{\frac{1}{\sigma}} (\Omega_f A_f)^{\frac{\sigma-1}{\sigma}} = W_0 \Big(\frac{H_f}{\zeta_f} \Big)^{\psi} \tag{5}$$

$$P_0 \Big(\frac{1+\eta}{\eta} \Big) Q_f^{\frac{1}{\eta}} \Big(\frac{Q_f}{L_f} \Big)^{\frac{1}{\sigma}} (\Omega_f)^{\frac{\sigma-1}{\sigma}} \beta_l = V \tag{6}$$

Equilibrium $(P_f^*, Q_f^*, H_f^*, L_f^*)$ is the solution to the system of equations (1), (2), (5), and (6): the production function, the goods demand curve, and the two input demand equations. Wages are determined in function of the bargaining parameter, as described in Equation (4), and do not have any effect on equilibrium output, inputs, and goods prices as long as capital is held fixed.

An important difference between the model highlighted above and the model of Abowd and Lemieux (1993) is that the latter assume the outside option of workers to be a scalar, whereas I allow the outside option to be increasing. The analog of this feature in models of vertical relations would be that sellers face increasing marginal costs.

Effects of Employer Power: Endogenous Investment

Although employer power $(1-\gamma_f)$ does not affect equilibrium output when holding the capital stock K_f fixed, employer power optimal investment, which in turn affects marginal costs and, hence, equilibrium output. Suppose firms need to pay a capital cost ϕ per unit of capital K_f, which is a fixed cost because it does not vary with production. We assume that capital increases employer variable profits, implying that it increases skill-augmenting and/or Hicks-neutral productivity. This is an uncontroversial assumption: if capital would not increase buyer variable profits, firms would never invest unless subsidized to do so. Proposition 1 says that under strongly efficient bargaining, employer power increases firms' technology adoption.

Proposition 1 *Under strongly efficient bargaining, buyer power increases capital investment:*

$$\frac{\partial K_f}{\partial(1-\gamma_f)} > 0$$

The proof of this theorem is straightforward. Denoting joint profits as $\Pi^j \equiv \Pi^d + \Pi^u$, the effect of capital on employer profits Π^d is given by:

$$\frac{\partial \Pi_f^d}{\partial K_f} = \frac{\partial \Pi_f^d}{\partial A_f}\frac{\partial A_f}{\partial K_f} - \Phi = (1-\gamma_f)\frac{\partial \Pi_f^j}{\partial A_f}\frac{\partial A_f}{\partial K_f} - \Phi$$

Taking the derivative with respect to employer power $(1-\gamma)$ gives:

$$\frac{\partial}{\partial(1-\gamma_f)}\Big(\frac{\partial\Pi_f^d}{\partial K_f}\Big) = \frac{\partial\Pi_f^j}{\partial A_f}\frac{\partial A_f}{\partial K_f}$$

This last term is positive under the assumption that the technology is variable profit-enhancing.

The intuition behind Proposition 1 is that buyer power increases the share of the rents created by capital investment that flows to the buyer. Hence, this increases the incentive for the buyer to invest. This a reformulation of the well-known holdup mechanism from Williamson (1971), which hinges on the assumption that workers and firms can only write incomplete contracts that do not condition on investments by the employer. The wage contracts used in the Illinois coal mining industry are an example of such an incomplete contract.

Corrolary 1 *Under strongly efficient bargaining, buyer power increases equilibrium output.*

It follows immediately from Proposition 1 that employer power increases equilibrium output in the strongly efficient bargaining model. Given the strong efficiency assumption, employer power does not affect output conditional on technology adoption K_f. However, employer power increases technology adoption, hence, decreases marginal costs. This marginal-cost reduction results in increased equilibrium output.

2.3 Weakly Efficient Bargaining

In reality, employees and employers rarely bargain over both wages and employment, usually only over wages. Assuming linear wage contracts, two types of bargaining models are possible. First, employers could choose employment and bargain with workers, this is the "weakly efficient" bargaining model in Abowd and Lemieux (1993).[2] Such a model

[2]In the vertical relations literature, this is similar to models where buyers choose output prices in their product market and simultaneously bargain wholesale prices with their supplier, such as in Crawford and Yurukoglu (2012) and Grennan (2013).

does not allow for monopsony-induced deadweight loss: when all bargaining power goes to the employer, this model converges to the strongly efficient bargaining model, not to the monopsony model with distortions due to wage markdowns.

Instead, in this paper I consider a different bargaining protocol: I assume that workers decide how much to work, and that they simultaneously bargain over wages with the employers. There are two reasons for making this modeling assumption. One, this model has the benefit of collapsing to the classical monopsony model when assuming perfect buyer power ($\gamma = 0$), with the ensuing monopsony distortion. Two, as I will discuss in detail in Section 3, the empirical evidence shows that a large-scale strike in 1898, which resulted in great bargaining power for the miners union, was followed by increased output, rather than decreased output. This finding aligns with the model in which employers exert monopsony power, whereas it opposes a sequential monopoly model in which miners exert market power.[3]

The labor union decides how much labor it is willing to supply for any given wage level, in order to maximize union profits Π^u, which leads to the following upward-sloping high-skilled labor supply curve:

$$W_f = W_0 \Big(\frac{H_f}{\zeta_f} \Big)^{\psi} \tag{7}$$

Wages are bargained over between the labor union and the employers according to their relative bargaining power γ_f:

$$\max_{W_f} (\Pi_f^u)^{\gamma_f} (\Pi_f^d)^{1-\gamma_f}$$

[3]In the latter case, the labor union's increased bargaining power should result in a decrease in equilibrium output, rather than an increase.

The resulting wage equation is given by Equation (8):

$$\gamma_f\Big((1-\frac{\partial Z_f}{\partial W_f})H_f+\frac{\partial H_f}{\partial W_f}(W_f-Z_f)\Big)(P_fQ_f-W_fH_f-VL_f) \\ +(1-\gamma_f)(W_f-Z_f)H_f(P_f\frac{\partial Q_f}{\partial H_f}\frac{\partial H_f}{\partial W_f}-H_f-W_f\frac{\partial H_f}{\partial W_f})=0 \quad (8)$$

Finally, low-skilled workers are chosen by the employers to maximize their profits:

$$\max_{L_f}(\Pi_f^d)$$

The low-skilled labor demand function is the same as in the strongly efficient bargaining case, Equation (6).

I rely on the "Nash-in-Nash" equilibrium concept of Horn and Wolinsky (1988). Equilibrium $(Q_f^*, P_f^*, W_f^*, H_f^*, L_f^*)$ is given by Equations (1), (2), (6), (7), and (8), which are the production, goods demand, low-skilled labor demand, high-skilled labor supply, and wage bargaining equations.

Effects of Employer Power: Endogenous Investment vs. Deadweight Loss

In contrast to the strongly efficient bargaining case, an increase in buyer power $(1-\gamma_f)$ now decreases equilibrium output when holding capital K_f fixed. This is because employer power allows employers to exercise monopsony power, which creates deadweight loss. In the extreme case when $\gamma_f = 0$, employers set wages at a fixed markdown in function of the labor supply elasticity, which negatively distorts equilibrium labor usage and output.

On the other hand, buyer power still affects capital investment. The effect of capital investment on buyer profits is given by:

$$\frac{\partial \Pi_f^d}{\partial K_f}=\frac{\partial \Pi_f^d}{\partial A_f}\frac{\partial A_f}{\partial K_f}=(1-\gamma_f)\frac{\partial \Pi_f^j}{\partial A_f}\frac{\partial A_f}{\partial K_f}-\Phi$$

Taking the derivative with respect to union power γ gives:

$$\frac{\partial}{\partial \gamma_f}\Big(\frac{\partial \Pi_f^d}{\partial K_f}\Big) = -\frac{\partial \Pi_f^j}{\partial A_f}\frac{\partial A_f}{\partial K_f} + (1-\gamma_f)\frac{\partial}{\partial \gamma_f}(\frac{\partial \Pi_f^j}{\partial A_f}\frac{\partial A_f}{\partial K_f})$$

The first term is still negative: buyer power increases the buyer's share of the profit increase from technology adoption. The second term is the effect of buyer power on the joint profit effect of capital investment and is no longer zero, given that union power affects equilibrium output and input quantities. Numerically simulating the model reveals that the sign of the second term is ambiguous, as I show below.

As a result, the net effect of buyer power on equilibrium output is ambiguous. On one hand, increased buyer power decreases output through the monopsony distortion. On the other hand, buyer power can increase technology usage, which reduces marginal costs, and hence increases output. Which of these effects dominates depends on the relative magnitude of the deadweight loss and the endogenous investment mechanism. In the empirical application, I will quantify the relative size of these effects to examine how counterfactual changes in employer power affect equilibrium output, producer surplus, consumer surplus, and worker welfare.

Simulation

To illustrate the importance of endogenous technology usage for the welfare effects of labor market power, I simulate a parametrized version of the model.[4] I calibrate the goods demand elasticity at $\eta = -7$ and the inverse labor supply elasticity at $\psi = 0.25$, following the estimates for U.S. construction workers in Kroft et al. (2020). I consider a new technology that increases H-augmenting productivity A by 5%, and increases TFP Ω by 20%.

Figure 2a plots equilibrium technology usage K against employer power $(1-\gamma_f)$. The solid red line depicts the model in which technology usage is allowed to change in function of employee bargaining power. As I explained in the theoretical model, technology usage increases with the level of employer power. By comparison, the dashed blue line depicts

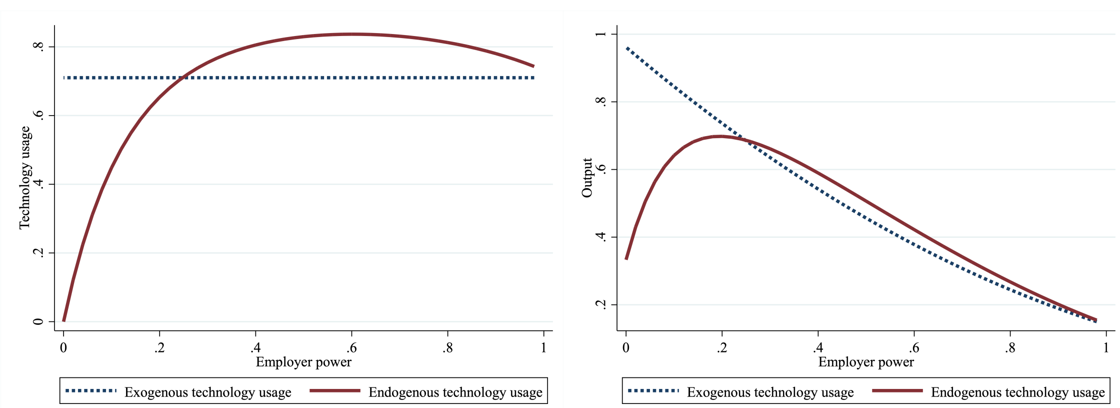
[4]The parametrization is specified in Appendix B.1.

the model in which technology usage is exogenous to the degree of employer power. In this model, technology usage is fixed equal to average technology usage in the endogenous adoption model.

Figure 2 shows equilibrium output Q as a function of employer power $(1 - \gamma_f)$. Under the assumption of exogenous technology usage, the blue solid line, output monotonically decreases with employer power. This is due to deadweight loss induced by the employer's monopsony power. The wage markdown set by the employer shrinks to zero as employer bargaining power goes to zero, reducing deadweight loss to zero. However, allowing for endogenous capital usage turns the output-bargaining power relationship into an inverted U-shape. At low levels of employer power, the output decrease due to monopsony power is countered by the reduction in marginal costs due to increased technology usage. Under the parametrization of the model, the positive output effect of increased technology usage outweighs deadweight loss until the bargaining weight of the employer is 0.20. Hence, equilibrium output is maximized at this level of employer power. In Appendix B.1, I show that this pattern is robust to various alternative parametrizations.

Figure 1: Bargaining Power, Capital Investment, and Output

(a) Capital Investment **(b) Output**



Notes: Panel (a) shows how capital investment changes with the degree of employer power in the simulated model. Panel (b) shows output in function of the employer's bargaining parameter, both when assuming exogenous investment and when letting investment vary with employer power.

3 Empirical Model: Illinois Coal Extraction

Simulating the model in the previous section reveals that the relationship between employer power and output is ambiguous, as it depends on the relative magnitude of monopsony-induced deadweight loss and endogenous technology adoption mechanisms. In this section, I quantify the relative magnitude of these forces by estimating the model in the context of the 19th-century Illinois coal mining industry.

3.1 Data

The main dataset is derived from the *Biennial Report of the Inspector of Mines of Illinois*, which I digitized for the purpose of this project. I observe every bituminous coal mine in Illinois between 1884 and 1902 at two-year intervals, which results in 7,996 observations. The dataset records the name of the mine, the mine owner's name, yearly coal extraction, average employee counts for both skilled and unskilled workers, days worked, and a dummy for cutting-machine usage per year. Materials are measured as the total number of powder kegs used in a given year. Other technical characteristics are observed for a subset of years, such as dummies for the usage of various other technologies (locomotives, ventilators, longwall machines), and technical characteristics such as mine depth and the mine entrance type (shaft, drift, slope, surface). Not all these variables are used in the analysis, given that some are observed in a small subset of years.[5]

I observe the average piece rate for skilled labor throughout the year and the daily wage for unskilled labor from 1888 to 1896. At some of the mines, "wage screens" were used, which means that skilled workers were paid only for their output of large coal pieces, rather than for their total output. This introduces some measurement error in labor costs. However, according to the dataset, wage screens were used for merely 2% of total employment in 1898. Skilled wages and employment are separately reported for the summer and winter months from 1884 to 1894. For some years, I observe additional variables such as mine

[5] Appendix Table A6 contains a full list of observed variables and the years in which they are observed.

capacities, the value of the total capital stock, and a breakdown of coal sales by destination. Wages and employee skill types are not observed in 1896. I deflate all monetary variables using historical CPI estimates from Hoover (1960). The reported monetary values are all in 1884 U.S. dollars.

In addition to the main biennial dataset, I utilize other datasets. First, the Inspection Report of 1890, which contains monthly data on wages and employment for both types of workers, and monthly production quantities for a sample of 11 mines that covers 15% of skilled and 9% of unskilled workers. Second, town- and county-level information are derived from the 1880 and 1900 population census and the censuses of agriculture and manufacturing. Third coal cutting machine costs are obtained from Brown (1889). Appendix A contains more details on my data sources and cleaning procedures.

3.2 Industry Background

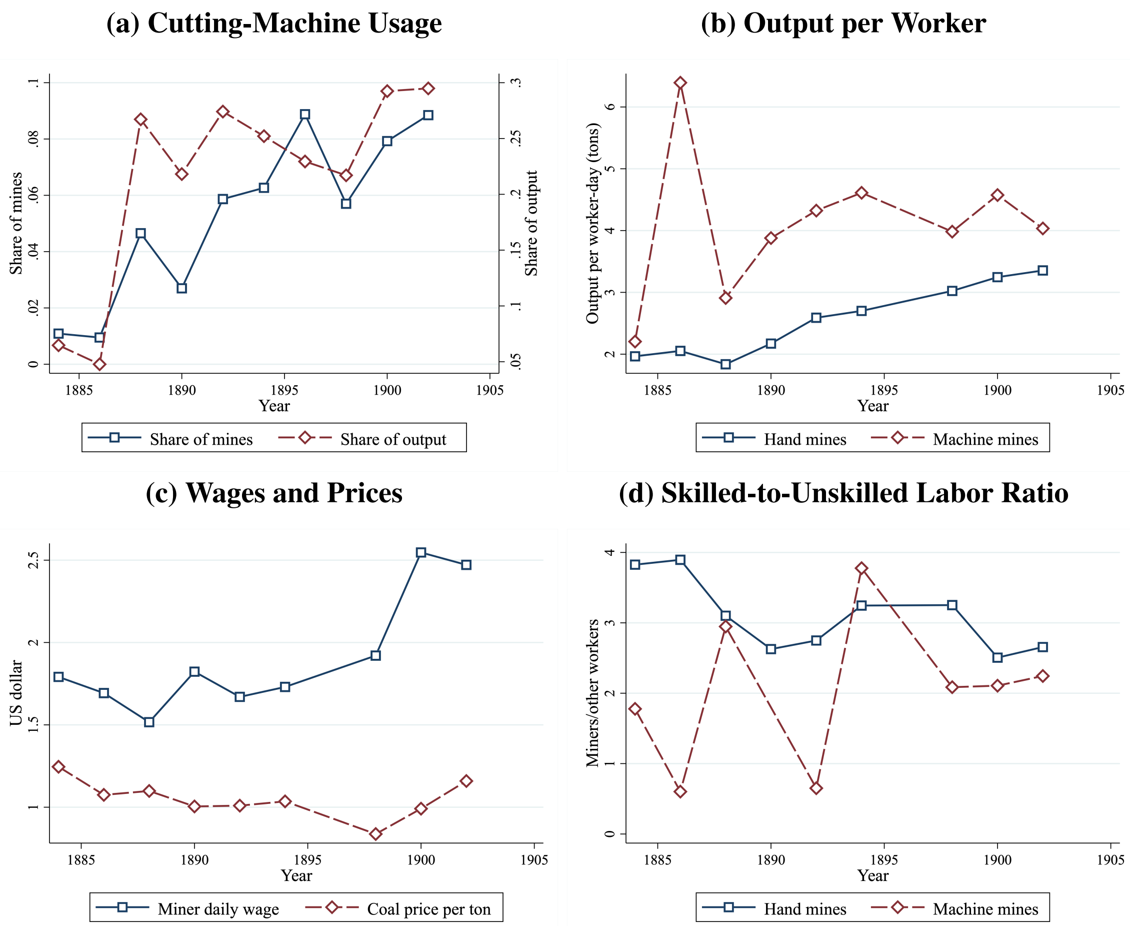
My empirical setting is the Illinois coal mining industry between 1884 and 1902, throughout which the industry grew rapidly: annual output tripled from 8.7 to 28.1 megatons. This was due to both an increase in the average mine size and because the number of mines grew from 643 to 859 units.

Coal-Extraction Process

The coal-extraction process consisted of three main steps. First, the coal vein had to be accessed, as it lay below the surface for 98.75% of the mines and 99.3% of output. Second, after miners reached the vein, the coal wall was "undercut", traditionally by hand, but from 1882 onward also with coal-cutting machines. Mechanization of the cutting process is considered to be the most significant technological advancement during this period (Fishback, 1992). Third, the coal had to be transported to the surface and separated from impurities. The hauling was done using mules or underground locomotives. Mines used two types of intermediate inputs. First, black powder was used to blast the coal wall. This powder and other materials, such as picks, were purchased and brought to the mine by the miners. Second, coal itself was used to power steam engines, electricity generators, and air compressors.

Figure 2(b) plots the ratio of total output over total days worked at mines that used cutting machines ("machine mines") and mines that did not ("hand mines"). Daily output per worker increased from 2.0 to 3.4 tons for hand mines, and from 2.2 to 4.0 tons for machine mines.[6]

Figure 2: Output, Inputs, and Prices



Notes: Panel (a) plots the evolution of cutting-machine usage, both as a share of firms and weighted by output. Panel (b) documents the evolution of output per worker at mines where coal was cut manually, and at mines where cutting machines were used. Panel (c) shows the evolution of daily skilled wages and of the coal price per ton in Illinois, weighted by employment and output, respectively. Panel (d) shows the evolution of the aggregate ratio of skilled to unskilled workers in Illinois for both hand and machine mines.

[6] This series is adjusted for the reduction of hours per working day in 1898, as explained in Appendix A.

Although different coal types exist, the mines in the dataset all extracted bituminous coal. There might have been minor quality differences even within this coal type due to variation in sulfur content, ash yield, and calorific value (Affolter & Hatch, 2002). Most of this variation is, however, dependent on the mine's geographical location and, hence, not a choice variable of the firms conditional on operating in a certain location.

Occupations

Coal mining involved numerous occupational tasks. The inspector report from 1890 reports wages at the occupation level, I report this subdivision in Appendix Table A1 for the 20 occupations with the highest employment shares, together covering 97% of employment. Three of five workers were miners; they did the actual coal cutting. This required significant skill: to determine the thickness of the pillars, miners had to trade off lower output with the risk of collapse. The other 40% of workers did various tasks such as clearing the mine of debris ("laborers"), hauling coal to the surface using locomotives or mules ("drivers" and "mule tenders"), loading coal onto the mine carts ("loaders"), opening doors and elevators ("trappers"), etc. The skills required to carry out these tasks were usually less complex than those of the miners, and moreover they were not specific to coal mining: tending mules and loading carts were general-purpose tasks, in contrast to mining-specific tasks such as undercutting coal walls.

The difference in industry-specific skills is reflected in daily wages: miners earned higher daily wages than any other mining employee type, except for "pit bosses" (middle managers), and "roadmen", who maintained and repaired mine tracks, but these two categories of workers represent barely 2% of the workforce. The higher wages of miners cannot be explained as a risk premium, because most other occupations were performed below the surface as well, and were hence subject to the same risks of mine collapse or flooding. From this point onward, I classify workers into two types: miners, which I denote as "skilled labor", and all other employees, which I denote as "unskilled labor". This follows the categorization of labor skill levels provided in the dataset.

Technological Change

The first prototype of a mechanical coal cutter in the United States was invented by J.W. Harrison in 1877.[7] The Harrison patent was acquired and adapted by Chicago industrialist George Whitcomb, whose "Improved Harrison Cutting Machine" was released on the market in 1882.[8] As shown in Figure 2a, the share of Illinois coal mines using a cutting machine increased from 1% to 9% between 1884 and 1902. Mechanized mines were larger: their share of output increased from 7% to 30% over the same time period. Mechanization of the hauling process, which replaced mules with underground locomotives, was another source of technical advancement that started during the 1870s. By the start of the panel, in 1884, mining locomotives were already widely used in Illinois: the share of output of mines that operated locomotives was around 80% in 1884 and 90% in 1886.

As shown in Figure 2(b), output per worker was higher in cutting machine mines. The composition of labor was also different: in Figure 2(d), I plot the ratio of the total number of skilled-labor days over the number of unskilled-labor days.[9] Mines without cutting machines used between three and four skilled labor-days per unskilled labor-day throughout the sample period, compared to two to three skilled labor-days per unskilled labor-day for machine mines. In every year except 1894, machine mines had a lower skilled-to-unskilled labor ratio than the other mines. On average, the skilled-to-unskilled labor ratio was 13% lower for machine mines compared to hand mines, and this difference is statistically significant.[10] However, this difference is not necessarily a causal effect of cutting machines on skill-augmenting productivity: mines with higher productivity levels were probably more likely to adopt cutting machines. For estimates of the causal effect of cutting machines on TFP and factor-augmenting productivity levels, I refer to the empirical model in Section 4.

[7] Simultaneously, prototypes of mechanical coal-cutting machines were invented in Northern England in the late 1870s (Reid, 1876; Ackermann, 1902).

[8] Appendix Figure A5 depicts the patent. The spatial diffusion of cutting machines is shown in Appendix Figure A3.

[9] 1890 is omitted for machine mines due to employment being unobserved for most machine mines in that year.

[10] Regressing the log skilled-to-total-labor ratio on a cutting-machine usage indicator results in a coefficient of -0.145 with a standard error of 0.011.

Anecdotal evidence suggests that cutting machines led to the substitution of unskilled for skilled workers. In his 1888 report, the Illinois Coal Mines Inspector asserts:

> "Herein lies the chief value of the [cutting] machine to the mine owner. It relieves him for the most part of skilled labor [...] it opens to him the whole labor market from which to recruit his forces [...] The mining machine is in fact the natural enemy of the coal miner; it destroys the value of his skill and experience, and reduces him to the rank of a common laborer" (Lord, 1892).

Labor Markets

Skilled workers received a piece rate per ton of coal mined, whereas unskilled workers were paid a daily wage.[11] Converting the piece rates to daily wages, the net salary of skilled labor was on average 16% higher compared to unskilled labor. "Net salary" means net of material costs and other work-related expenses. Rural Illinois was (and still is) sparsely populated: the median and average populations of the towns in the dataset were 872 and 1,697 inhabitants. In the average town, 13.6% of the population was employed in a coal mine. Women and children under the age of 12 did not work in the mines, which implies that a large share of the local working population was employed in coal mining. Of all the towns, 43% had just one coal firm, and 75% had three or fewer. On average 62% of mining employment was performed in towns with three or fewer coal firms. Although most of the towns in the dataset were connected by railroad, these were exclusively used for freight: passenger lines operated only between major cities (Fishback, 1992). Given that the average village was 6.6 miles from the next village, and that skilled workers had to bring their own supplies to the mine, commuting between towns was not an option, and the mining towns can be considered as isolated local labor markets. Most roads were unpaved and automobiles had not yet been introduced. To switch employers, miners had to migrate to another town.

The first attempts to unionize the Illinois coal miners started around 1860, but without

[11] Piece rates were an incentive scheme in a setting with moral hazard, as permanent miner supervision would have been very costly.

much success (Boal, 2017). Although Illinois coal miners were unionized, for instance through the United Mine Workers of America and the Knights of Labor, union power was contrained by the use of “yellow-dog” labor contracts, which forced employees not to join a trade union.[12] A major strike in 1897–1898 led to a modest increase in wages, to a reduction of working hours, and to the introduction of annual wage negotiations, which took place in January (Bloch, 1922). Nevertheless, industrial relations remained tense for the ensuing years (Bloch, 1922).

Wages were bargained over in a tiered negotiation procedure: first, a general agreement was made at the state-industry level; afterwards, mine owners individually negotiated wages with miner representatives (Bloch, 1922). There was no minimum wage law. In contrast to other states, the mines in the dataset did not pay for company housing of the miners (Lord, 1883, 75), which would otherwise have been a labor cost in addition to miner wages.

Figure 2(c) reports the aggregate skilled-labor daily wage, defined as the total wage bill spent on skilled labor over the total number of skilled labor-days. The fast growth in labor productivity did not translate into higher wages until 1898; daily miner wages remained around \$1.80. After the strikes, wages rose sharply to around \$2.50 per day.

Coal Markets

Coal was sold at the mine gate, and there was no vertical integration with postsales coal treatment, such as coking. On average, 92% of the mines’ coal output was either sold to railroad firms or transported by train to final markets. The remaining 8% was sold to local consumers. The main coal-destination markets for Illinois mines were St. Louis and Chicago. Railway firms acted as an intermediary between coal firms and consumers, and were also major coal consumers themselves. Historical evidence points to intense competition in coal markets during the last two decades of the 19th century, before the large consolidation wave in the early 1900s (Graebner, 1974). Nevertheless, coal was still costly to transport, which means that coal markets were likely not perfectly integrated: coal prices

[12]These contracts were criminalized in Illinois in 1893, with fines of \$100, which was equivalent to on average six months of a miner’s wage. (Fishback, Holmes, & Allen, 2009).

varied considerably across Illinois. In 1886, for instance, the coal price varied between 91 cents/short ton at the 10th percentile of the price distribution to 1.75 dollars/short ton at the 90th percentile, and this price dispersion slightly increased over time. Figure 2(c) shows that the mine-gate coal price per ton, weighted by output shares, fell from $1.25 to $0.84 between 1884-1898, after which it increased again.

3.3 Stylized Facts

In this section, I present two sets of stylized facts to motivate the assumptions made in the structural model. First, I show that skilled wages covaried with seasonal labor demand shocks, whereas unskilled wages did not. Second, I show that output increased in response to the 1897–1898 coal strikes at striking mines.

Seasonal Wage Variation

Coal demand was seasonal: demand for energy was higher in winter than in summer. Coal storage costs meant that firms could not fully arbitrage between winters and summers, and, hence, needed to hire more workers during winter. Joyce (2009) mentions that miners were (partially) unemployed during the summer months. This cyclical pattern provides useful variation to compare wage responses of skilled and unskilled workers to coal demand shocks. In Figure 3(a), I show that skilled wages followed this coal demand cycle: they were higher during winters than during summers. However, this pattern held only for skilled wages, not for unskilled wages. Although the seasonal demand shocks increased demand for both skilled and unskilled labor, only skilled wages increased during winter. This is also shown in Figure 3(c), which plots how average daily wages for both skilled and unskilled workers in 1890 changed with the monthly number of worker-days of each type at the mine-month level throughout 1890.[13] Skilled wages were positively correlated with monthly skilled employment, whereas the unskilled worker wage-employment schedule

[13]Unlike skilled wages and employment, unskilled wages and employment are not broken down by season in the entire dataset. However, monthly wage and employment data are available for a sample of mines selected by the Illinois Bureau of Labor Statistics across five counties in 1890, which cover 16% of skilled employment and 9% of unskilled employment.

was flat. Moreover, skilled wages varied greatly across mines and months, but there was very little cross-sectional and intertemporal variation in unskilled wages.

The fact that skilled wages increased in response to coal demand shocks whereas unskilled wages did not, and the fact that unskilled wages were nearly uniform across Illinois whereas skilled wages were not, suggests an inelastic high-skilled labor supply curve and a fully elastic low-skilled labor supply curve. This implies that firms had the potential to exert monopsony power on the high-skilled labor market, but not on the low-skilled labor market. However, it could be that seasonal employment movements reflected not only labor demand variation but also labor supply shocks, for instance due to the harvesting season. Moreover, within-year demand shocks trace out a short-run supply curve, whereas labor supply could be more elastic in the longer run. Hence, in the structural model, I will rely on an instrumental-variable strategy that relies on international coal-price shocks to identify the high-skilled labor supply elasticity.

Figure 3: Wage-Employment Profile by Skill Type

(a) Wages **(b) Wage-Employment Profile**

Notes: Panel (a) shows howthe wages of skilled miners and other mine employees evolved monthly during 1890. Panel (b) plots mine-month-level wages for both types of workers against monthly employment, again for both types of workers.

Output and Investment Response to the 1897–1898 Coal Strikes

During 1897–1898, a large strike broke out in the Illinois coal basin, which became known as the "Illinois coal war". Miners went on strike at one-quarter of the coal mines in Illinois;

this resulted in a wage increase at 92% of the striking coal mines. Given that the strike was successful at increasing wages, this provides a useful shock to the relative bargaining power of the miners union. Using a difference-in-differences model, I compare the evolution of log coal output between mines at which miners went on strike in 1898, $I(strike)_f = 1$, to mines where miners did not strike, before and after 1898. I define the strike indicator as mines where miners struck for at least a week during 1898, and I include both mine fixed effects and a linear time trend. I exclude the year 1898 from the analysis, to avoid taking into account the reduction in output during the strike.

$$q_{ft} = a_0 + a_1 I(strike)_f + a_2 I(strike)_f I(t \geq 1898) + a_3 I(t \geq 1898) + a_4 t + \delta_f + e_{ft}$$

I also estimate the same difference-in-differences model, but with cutting-machine usage on the left-hand side instead of log output. Rather than a linear model with firm fixed effects, I use a conditional logit model with firm fixed effects because cutting machine usage is a binary variable. In Table 1a, I report the coefficient on the interaction term, a_2. At mines that went on strike, output increased by 33% after 1898 relative to the mines that did not go on strike. In contrast, machine usage is estimated to decrease relatively more at the striking mines, although this effect is not statistically significant.

In Table 1b, I compare pretrends by estimating the interaction effect between the strike indicators and a linear time trend prior to 1898. Although the pretrend in output was not signficantly different between mines that went on strike and mines that did not, cutting-machine usage grew less prior to the strike at mines that went on strike compared to mines that did not. Hence, the difference-in-difference estimates for cutting-machine usage are potentially invalid. To examine the mechanism behind the output and mechanization effects of changes in employer power, and to conduct welfare counterfactuals, a structural model is necessary.

Table 1: Strikes, Output, and Investment

(a) Diff-in-diff results	Log(Output)		Machine Usage	
	Est.	S.E.	Est.	S.E.
1(Strike)*1(year$\geq$ 1898)	0.288	0.096	-0.048	0.505
Model	Linear		Logit	
Firm fixed effects	Yes		Yes	
R-squared	.944		.	
Observations	7154		598	
Firm fixed effects:	No		No	

(b) Pre-trends	Log(Output)		Machine Usage	
	Est.	S.E.	Est.	S.E.
1(Strike)*yr	0.004	0.018	-0.170	0.038
Model	Linear		Logit	
R-squared	.205		.	
Observations	5437		5437	

Notes: Panel (a) reports the difference-in-differences estimates of how log output and cutting-machine usage changed differently after 1898 between striking and nonstriking mines. Panel (b) compares the pretrend in output and cutting machine usage between the mines that went on strike and those that did not.

3.4 Empirical Model

Production Function

I implement an empirical model of the Illinois coal industry based on the general model outlined in Section 2. Let f index coal firms per town and let t index all even years between 1884 and 1902. The model is set up at the firm-town-year level: it is plausible that employers optimize at the firm level, rather than at each mine independently. However, I let firms optimize on a labor-market-by-labor-market basis: firms with mines in different labor markets do not internalize the cross-labor-market effects of their decisions. This is consistent with the model, given that it does not feature strategic interaction between firms on the labor market. Annual coal extraction in short tons is denoted as Q_{ft}, the number of days worked by high-skilled labor is denoted as H_{ft}, and the number of low-skilled labor-days is denoted as L_{ft}. In contrast to the theoretical model, capital investment is modeled as a binary variable: firms choose whether to use cutting machines or not, with usage being denoted as $K_{ft} \in \{0, 1\}$. I abstract from other technologies, such as mining locomotives, because they were widely adopted already by the start of the panel, and because they are not observed in all years of the sample.

I maintain the CES production function from Equation (1), with an elasticity of input substitution σ and low-skilled-labor coefficient β^l. Firms differ in terms of their skill-augmenting productivity A_{ft} and in their Hicks-neutral productivity Ω_{ft}. In Appendix C.1, I estimate various extensions of the production model to allow for nonconstant returns to scale, the existence of intermediate inputs, and the possibility that cutting machines change scale returns. All these extensions lead to very similar production-function estimates. Taking the logs of Equation (1) and adding the time index leads to Equation 9, which is the production function I will estimate:

$$q_{ft} = \left(\frac{\sigma - 1}{\sigma}\right) \log\left(\left(H_{ft}A_{ft}(K_{ft})\right)^{\frac{\sigma-1}{\sigma}} + \beta^l L_{ft}\right)^{\frac{\sigma}{\sigma-1}} + \omega_{ft}(K_{ft}) \tag{9}$$

Cutting-machine usage K_{ft} is allowed to affect both productivity terms Ω_{ft} and A_{ft}. The logarithms of both these productivity terms, a_{ft} and ω_{ft}, are assumed to evolve as AR(1) processes, as specified in Equations (10) and (11). The productivity terms have serial correlations ρ^a and ρ^ω and are assumed to be affected linearly by cutting-machine usage, as parametrized by the coefficients α^k and β^k for labor-augmenting and Hicks-neutral productivity, respectively.[14] Skill-augmenting and Hicks-neutral productivity shocks are denoted

[14]Although these AR(1) specifications do not allow for richer models of cost dynamics in which current productivity is a function of the total amount of output produced in the past, they do have the benefit of not requiring inversion of the production function, thereby allowing for rich heterogeneity in both productivity terms, markdowns, and markups. See Appendix C.2 for a motivation and discussion of this assumption.

as e^a_{ft} and e^ω_{ft}:

$$a_{ft} = \alpha^k K_{ft} + \rho^a a_{ft-1} + e^a_{ft} \tag{10}$$

$$\omega_{ft} = \beta^k K_{ft} + \rho^\omega \omega_{ft-1} + e^\omega_{ft} \tag{11}$$

I assume that mines do not face a binding capacity constraint. This is consistent with the data: in 1898, the only year for which capacities are observed, merely 1.9% of the mines operated at full capacity, and they were responsible for just 2.7% of industry sales.[15] I also abstract from stockpiling of coal, and I assume that coal must be sold immediately after extraction: coal storage usualy led to deteriorating coal quality; moreover it was expensive and dangerous (Stoek, Hippard, & Langtry, 1920). As Williams (1901) asserted:

> "The product of a mine can be stored with economy only in the mine itself [...] Coal must be sold, therefore, as fast as it is mined"

Labor Supply

Adding time subscripts to the inverse labor supply function (7), then inverting it, results in the labor supply equation (12a). The daily wage of high-skilled workers W_{ft} is computed as the piece rate multiplied by daily tonnage per worker. I measure W_{0t} as the average daily wage in year t.

I include observed firm characteristics $\mathbf{X}^h_{ft}$ next to the latent amenity term ζ_{ft}. Specifically, I include a linear time trend, county fixed effects, and the logarithm of the minimal distance of the firm to Chicago and St. Louis as observed characteristics, to account for proximity to the large regional population centers.

$$H_{ft} = \left(\frac{W_{ft}}{W_{0t}}\right)^{\frac{1}{\psi}} \exp(\mathbf{X}^h_{ft})^{\psi^x} \zeta_{ft} \tag{12a}$$

[15] Figure A4 depicts the entire distribution of capacity utilization rates.

I estimate the labor supply equation in logs, which is given by Equation (12b):

$$h_{ft} = \frac{1}{\psi}(w_{ft} - w_{0t}) + \boldsymbol{\psi}^x \mathbf{X}^h_{ft} + \log(\zeta_{ft}) \tag{12b}$$

The amenity term ζ_{ft} captures firm differentiation from the miner's perspective. In contrast to Delabastita and Rubens (2024), who rely on a homogeneous employers model, I do allow for firm differentiation because skilled wages varied substantially across mines, even within the same labor markets.[16]

Similarly to the theoretical model, the market for low-skilled labor is assumed to be perfectly competitive, and low-skilled workers are paid a uniform wage V, which is equal to their outside option. The main reason for this assumption lies in the fact that unskilled wages barely varied across Illinois, nor did they react to seasonal weather shocks, as shown in Figure 3.

Coal Demand

Coal produced in Illinois mines was a nearly homogeneous product. However, coal firms were differentiated by their locations, which resulted in price differences between coal firms. I again include the shortest distance to either Chicago or St. Louis, county fixed effects, and a linear time trend, because these variables likely affected coal demand:

$$Q_{ft} = \left(\frac{P_{ft}}{P_{0t}}\right)^{\eta} \exp(\mathbf{X}^q_{ft})^{\boldsymbol{\eta}^x} \xi_{ft} \tag{13}$$

Taking logarithms of Equation (13) results in Equation (14), which is the demand model I estimate:

$$q_{ft} = \eta(p_{ft} - p_{0t}) + \boldsymbol{\eta}^x \mathbf{X}^q_{ft} + \ln(\xi_{ft}) \tag{14}$$

[16] In Appendix Table A5, I report the R^2 of regressing log daily miner wages on subsequently year, county, town, and firm dummies. Town and year dummies explain only 29% of the variation in skilled miner wages.

Type of Bargaining

I follow the weakly efficient bargaining model, in which workers provide labor through an upward-sloping labor supply curve, for three reasons. First, the stylized facts in Section 3.3 showed that output increased in response to strikes. These strikes were a negative shock to employer bargaining power, as I show in Appendix C.2. In the strongly efficient bargaining environment of Section 2.2, output should not change in response to a shock to bargaining power; this would be a mere transfer. In a weakly efficient bargaining environment with a fully elastic labor supply curve and an inelastic product demand curve, a decrease in employer power should decrease output, due to double marginalization. In contrast, as I show in Section 2.3, in a weakly efficient bargaining model with inelastic labor supply, a drop in employer power increases output, as the wage markdown falls. Second, the description of miner union bargaining in Bloch (1922) shows that employers and unions bargained over wages, not employment, which is in line with the weakly efficient bargaining model. Third, the structural model will find evidence for an upward-sloping labor supply curve, which calls for considering monopsony power.

Variable Input Decisions

I assume firms make decisions in two stages. First, they choose whether or not to use cutting machines. Second, conditional on this choice, they choose high-skilled wages and low-skilled employment. I will discuss these two stages in reverse order to match the order in which the model is estimated.

In year t, employers negotiate a high-skilled wage rate with the union according to the bargaining protocol specified in Equation (15), which is equivalent to the weakly efficient bargaining model described in Equation (8).

$$\gamma_{ft}\Big((1-\frac{\partial Z_{ft}}{\partial W_{ft}})H_{ft}+\frac{\partial H_{ft}}{\partial W_{ft}}(W_{ft}-Z_{ft})\Big)(P_{ft}Q_{ft}-W_{ft}H_{ft}-VL_{ft})$$
$$+(1-\gamma_{ft})(W_{ft}-Z_{ft})H_{ft}(P_{ft}\frac{\partial Q_{ft}}{\partial H_{ft}}\frac{\partial H_{ft}}{\partial W_{ft}}-H_{ft}-W_{ft}\frac{\partial H_{ft}}{\partial W_{ft}})=0 \quad (15)$$

Following the labor supply curve, Equation (12b), this negotiated wage rate results in a certain amount of high-skilled labor being supplied at each coal firm. Coal firms simultaneously choose low-skilled labor, as specified in Equation (16). I assume that these variable input decisions happen after the productivity shocks e^{ω} and e^{a} are observed by the firm. The combination of low-skilled and high-skilled labor and capital, of which the decisions are specified below, results in coal output Q_f according to the production function, Equation (9). The coal demand curve, Equation (14), determines the price every firm can charge at that level of coal output.

$$P_{0t}\Big(\frac{1+\eta}{\eta}\Big)Q_{ft}^{\frac{1}{\eta}}\Big(\frac{Q_{ft}}{L_{ft}}\Big)^{\frac{1}{\sigma}}(\Omega_{ft})^{\frac{\sigma-1}{\sigma}}\beta_l = V \tag{16}$$

Fixed Input Choices

In every year t, the equilibrium quantities and prices at every firm can be found by solving the system of Equations (9), (12a), (13), (15), and (16): the production function, high-skilled labor supply, coal demand, high-skilled labor demand, and low-skilled labor demand. This delivers certain equilibrium outcomes $(Q^1_{ft}, P^1_{ft}, H^1_{ft}, L^1_{ft}, W^1_{ft})$ if the firm uses cutting machines, and different equilibrium outcomes $(Q^0_{ft}, P^0_{ft}, H^0_{ft}, L^0_{ft}, W^0_{ft})$ if the firm does not use cutting machines. The variable profit gain of the employer from using cutting machines is denoted as $\Delta\Pi^d_{ft}$:

$$\Delta\Pi^d_{ft} \equiv (P^1_{ft}Q^1_{ft} - W^1_{ft}H^1_{ft} - V_tL^1_{ft}) - (P^0_{ft}Q^0_{ft} - W^0_{ft}H^0_{ft} - V_tL^0_{ft}) \tag{17}$$

The fixed costs of technology usage are denoted Φ_t, so total employer profits are equal to $\Pi^d_{ft} - \Phi_t K_{ft}$. As in Peters, Roberts, Vuong, and Fryges (2017), I parametrize fixed costs as an exponential distribution. I let the rate parameters (ϕ^0, ϕ^1) evolve over time, with ϕ^0 measuring the time-invariant fixed cost of technology usage and ϕ^1 its time trend:

$$\Phi \sim \exp(\phi^0 + \phi^1 t)$$

I assume that prior to observing the productivity shocks e^{ω} and e^{a}, firms independently and simultaneously choose whether they will use cutting machines or not. They make this decision by trading off the costs of machine adoption Φ_t with the expected variable profit return $\Delta\Pi^d$. I assume that firms do not choose their cutting machines in a dynamic manner, but rather that they optimize their technology mix period by period. The main reason for this assumption is that observed entry *and* exit of machine usage is frequent: 106 instances of cutting machine installation, which were scrapped in 60 instances. This suggests the existence of an aftermarket for capital.

Using the exponential form of fixed costs, the probability that a firm uses cutting machines $p^k_{ft}(\phi)$ is equal to:

$$p^k_{ft}(\phi^0, \phi^1) = 1 - \exp\left(\frac{-\Delta\Pi^d_{ft}}{\phi^0 + \phi^1 t}\right) \tag{18}$$

4 Identification, Estimation, and Counterfactuals

4.1 Labor Supply Estimation

Although the model is specified at the firm level, the dataset is observed at the mine level. Given that firms are assumed to have optimized at the firm-town level, I aggregate all the relevant variables to the firm-town-year level, as detailed in Appendix A.2.

I start with the identification of the skilled labor supply curve, Equation (12b). The labor supply elasticity $\frac{1}{\psi}$ cannot be recovered by simply regressing high-skilled labor wages on employment, because of the latent firm characteristics ζ_{ft}. Firms with a high ζ_{ft} knew they were attractive to miners, which permitted them to offer a lower wage to attract the same number of miners. To identify the slope of the skilled labor supply curve, a shock to labor demand that is excluded from skilled labor utility is necessary.

I rely on international coal-market price shocks for identification. I obtain the average annual coal price on international coal markets in Europe from Degrève (1982). I use as instruments both this coal price and its interaction term with an indicator for whether or not a mine shipped coal over the railroad network. These instruments imply three assumptions. One, individual Illinois coal mines were too small to affect equilibrium coal prices on the European market. This makes sense given that Illinois produced only around 10% of the total U.S. output, and that U.S. bituminous coal mines exported only around 1.2% of their coal in 1898 (Graebner, 1974). Two, international coal-price shocks affected the demand for Illinois coal. Given that Chicago was one of the destination markets for Illinois coal, and that Chicago also sourced coal from both the East coast and other coal fields by lake steamers, changes in nonlocal coal prices affected demand for Illinois coal. Three, international coal price shocks affected coal demand more if coal mines were shipping their coal over the railroad network compared to coal mines that sold their coal only locally. This third assumption makes sense given that mines that sold only locally did not compete with coal fields outside of Illinois; neither these mines not their consumers were connected to the railroad network.

I compute the baseline wage level $w_{0i(f)t}$ as the average miner wage in Illinois. I estimate Equation (12b) with a two-stage least squares estimator using the European coal price and an interaction term of the European coal price and a shipping dummy as instruments for the log relative wage at each firm. I control for whether the firm was a shipping mine or a mine that sold only locally, and I include county fixed effects and a linear time trend.

For unskilled wages, I rely on the average daily wage for unskilled labor in the Illinois coal industry in every year. Given that I only observe this wage from 1884-1894, I linearly interpolate for the time period 1896-1902 using a loglinear time trend.

The skilled labor supply elasticity is estimated to be 3.324 with a standard error of 1.589, as reported in Table 2(a). This means that in the monopsony case, which corresponds to full employer power $\gamma_{ft} = 0$, skilled-labor wages would be set 23.13% below the marginal revenue product of labor.[17]

[17] The wage markdown is equal to $\frac{MRPL-w}{MRPL} = \frac{\psi}{\psi+1}$.

Table 2: Model Estimates

(a) Labor Supply		Est.	S.E.
Labor supply elasticity	$\frac{1}{\psi}$	3.324	1.589
1(Shipping mine)		1.631	0.306
Year		-0.035	0.004
First-stage F-stat		12.4	
Observations		6315	
(b) Coal Demand			
Coal demand elasticity	η	-5.341	0.431
Log(min. distance to big city)		-0.155	0.103
No. railroads		0.136	0.043
Year		-0.001	0.005
First-stage F-stat		773.1	
Observations		3127	
(c) Production Function			
Input substitution elasticity	σ	0.361	0.093
Skill-augmenting technology effect	α^k	0.194	0.105
Hicks-neutral technology effect	β^k	0.107	0.146
Low-skilled labor coefficient	β^l	0.006	0.008
Serial correlation Hicks-neutral productivity	ρ^ω	0.288	0.128
Serial correlation skill-augmenting productivity	ρ^ω	0.829	0.120
Observations			
(d) Fixed Machine Costs			
Fixed cost in 1882 ('000 USD)	ϕ^0	23.680	5.334
Fixed cost time trend ('000 USD)	ϕ^1	-2.358	0.606

Notes: Panel (a) reports the skilled-labor supply estimates, Panel (b) reports the estimates of the coal demand function, Panel (c) contains the estimates of the production function, with block-bootstrapped standard errors over 200 iterations, and Panel (d) reports the cutting-machine fixed-cost estimates.

4.2 Coal Demand Estimation

I estimate the coal demand function in logarithms, Equation (14), using firm-level quantities and prices. I rely on the thickness of the coal vein as a cost shifter: whereas the vein thickness affects the marginal cost of mining, it does not enter consumer utility conditional on the coal price, because vein thickness does not affect coal quality (Affolter & Hatch, 2002). The thickness of the coal vein was the result of geological variation, and hence not a choice variable.

I estimate Equation (14) using a two-stage least squares estimator, with the log average vein thickness in the town as the instrument for coal output. In the observed covariates vector $\mathbf{X}^q_{ft}$, I include the following coal demand shifters: the logarithm of the minimal distance to either Chicago or St. Louis, the number of railroads passing through the mine's town, whether the mine was a shipping mine or not, and a linear time trend. I compute p_{0t} as the average coal price in each year.

Table 2(b) contains the coal demand elasticity. The number of observations, 3,127, is lower than when estimating labor supply because the thickness of the coal veins is not observed in 1888 and 1890. The demand elasticity is estimated at -5.341 with a standard error of 0.431. This means that firms set coal prices at a markup of 23% above marginal costs.[18] The minimal distance from either Chicago or St. Louis had a negative but statistically insignificant effect on demand. A more important demand shifter seems to be the number of railroad lines passing through the mine's town. Coal demand was roughly stable throughout the sample period.

4.3 Production Function Estimation

I estimate the production function in two steps. First, I estimate the elasticity of input substitution σ and the skill-augmenting effects of cutting machines, α^k. Second, I estimate all other production-function coefficients, β^l and β^k.

[18]The coal-price markup above marginal costs is $\frac{P_{ft}}{mc_{ft}} = \frac{\eta}{\eta+1}$.

Elasticity of Substitution

The elasticity of substitution is usually estimated by taking the ratio of the input demand functions from the employer's profit-maximization first-order conditions, e.g., in Doraszelski and Jaumandreu (2018). In the bargaining model, however, the marginal-revenue product of high-skilled labor is not equal to its wage as long as $\gamma < 1$. Setting γ to zero in Equation (8), which implies perfect monopsony power, gives:

$$\frac{\partial R_{ft}}{\partial H_{ft}} = W_{ft}(1+\psi)$$

Conversely, if γ becomes one, which implies that the labor union had all the bargaining power, the wage of high-skilled workers is equated to their marginal revenue product:

$$\frac{\partial R_{ft}}{\partial H_{ft}} = W_{ft}$$

These two first-order conditions for extremes of the bargaining parameter γ_f can be linearly interpolated using the bargaining parameter γ_{ft}, which results in a linear approximation of the first order conditions:

$$\frac{\partial R_{ft}}{\partial H_{ft}} = W_{ft}(1+(1-\gamma_{ft})\psi) \tag{19}$$

Working out the first-order conditions (6) and (19), then dividing (6) by (19), results in Equation (20). This equation is a variant of the first-stage regression from Doraszelski and Jaumandreu (2018), except that the labor supply elasticity enters into the first-order conditions, as in Rubens, Wu, and Xu (2024):

$$l_{ft} - h_{ft} = \sigma(w_{ft} - v + \ln(1+(1-\gamma_{ft})\psi)) + \underbrace{\sigma\ln(\beta^l) + (1-\sigma)a_{ft}}_{\equiv \tilde{a}_{ft}} \tag{20}$$

Given that Equation (10) specifies an AR(1) process for the factor-augmenting productivity term a_{ft}, the residual $\tilde{a}_{ft}$ also evolves as an AR(1). Hence, taking ρ^a differences of

Equation (20) isolates the productivity shock e^a_{ft} as a function of the coefficients ρ^a, σ, and α^k. Using the previously stated assumptions that capital is chosen prior to observing the skill-augmenting productivity shock e^a_{ft}, but that variable inputs are chosen afterwards, the moment conditions, Equation (21), can be specified to estimate the elasticity of input substitution σ, the skill-augmenting productivity effect of cutting machines α^k, and the serial correlation in skill-augmenting productivity ρ^a:

$$\mathbb{E}\Big[e^a_{ft}(\rho^a, \alpha^k, \sigma)| \begin{pmatrix} K_{ft-r} \\ L_{ft-r-1} \\ H_{ft-r-1} \end{pmatrix} \Big]_{r=0}^{T-1} = 0 \tag{21}$$

Second-Stage Production-Function Coefficients

From Equation (20), the log factor-augmenting productivity residual a_{ft} can be written as a function of the estimated parameters σ and ψ, and the yet-to-be-estimated parameters β^l and β^k:

$$a_{ft} = \Big(\frac{l_{ft} - h_{ft}}{1-\sigma}\Big) - \frac{\sigma}{1-\sigma}\big(\ln(\beta^l)\big) - \frac{\sigma}{1-\sigma}\big(w_{ft} - v_t + \ln(1 + (1-\gamma_{ft})\psi)\big)$$

Substituting this factor-augmenting productivity term into the log production function gives:

$$q_{ft} = \frac{\sigma}{\sigma-1}\ln\left(\Big(H_{ft}\exp\big(\frac{l_{ft}-h_{ft}}{1-\sigma} - \frac{\sigma}{1-\sigma}\ln(\beta^l) - \frac{\sigma}{1-\sigma}(w_{ft} - v_t + \ln(1+(1-\gamma_{ft})\psi))\big)\Big)^{\frac{\sigma-1}{\sigma}} + \beta^l L_{ft}^{\frac{\sigma}{\sigma-1}}\right) + \omega_{ft} \tag{22}$$

I define the first linear term in the log production function as $B_{ft}(.)$:

$$B_{ft} \equiv \frac{\sigma}{\sigma-1}\ln\Bigg(\Big(H_{ft}\exp\big(\frac{l_{ft}-h_{ft}}{1-\sigma} - \frac{\sigma}{1-\sigma}\ln(\beta^l) - \frac{\sigma}{1-\sigma}(w_{ft} - v_t + \ln(1+(1-\gamma_{ft})\psi))\big)\Big)^{\frac{\sigma-1}{\sigma}} + \beta^l L_{ft}^{\frac{\sigma}{\sigma-1}}\Bigg)$$

Using the productivity transition in Equation (11), taking ρ^ω differences isolates the Hicks-neutral productivity shock e^ω_{ft} as a function of the parameters $(\rho^\omega, \beta^k, \beta^l)$:

$$e^\omega_{ft} = q_{ft} - \rho^\omega q_{ft-1} - \beta^k(k_{ft} - \rho^\omega k_{ft-1}) - (B_{ft} - \rho^\omega B_{ft-1})$$

Using the timing assumption that employers chose capital prior to the realization of the Hicks-neutral productivity shock but chose low-skilled labor and bargained over wages *after* the realization of this shock leads to the moment conditions in Equation (23). I estimate this model using lags of up to one time period.

$$\mathbb{E}\Big[e^\omega_{ft}(\rho^\omega, \beta^k, \beta^l)| \begin{pmatrix} K_{ft-r} \\ L_{ft-r-1} \\ H_{ft-r-1} \end{pmatrix}\Big]_{r=0}^{T-1} = 0 \tag{23}$$

4.4 Bargaining Weights

Adding time subscripts to the wage-bargaining first-order condition, Equation (15), and rearranging terms in function of the union bargaining weights γ_{ft} leads to Equation (24). I estimate the bargaining parameters using this equation, whose variables are all either observed or have been estimated in the production, labor supply, and goods demand models.

$$\gamma_{ft} = \frac{(W_{ft} - Z_{ft})(\frac{\partial R_{ft}}{\partial H_{ft}}\frac{\partial H_{ft}}{\partial W_{ft}} - H_{ft} - W_{ft}\frac{\partial H_{ft}}{\partial W_{ft}})}{(W_{ft} - Z_{ft})(\frac{\partial R_{ft}}{\partial H_{ft}}\frac{\partial H_{ft}}{\partial W_{ft}} - H_{ft} - W_{ft}\frac{\partial H_{ft}}{\partial W_{ft}}) - (\frac{\psi H_{ft}}{1+\psi} + \frac{\partial H_{ft}}{\partial W_{ft}}(W_{ft} - Z_{ft}))\frac{\Pi^d_{ft}}{H_{ft}}} \tag{24}$$

Fixed-Point Algorithm

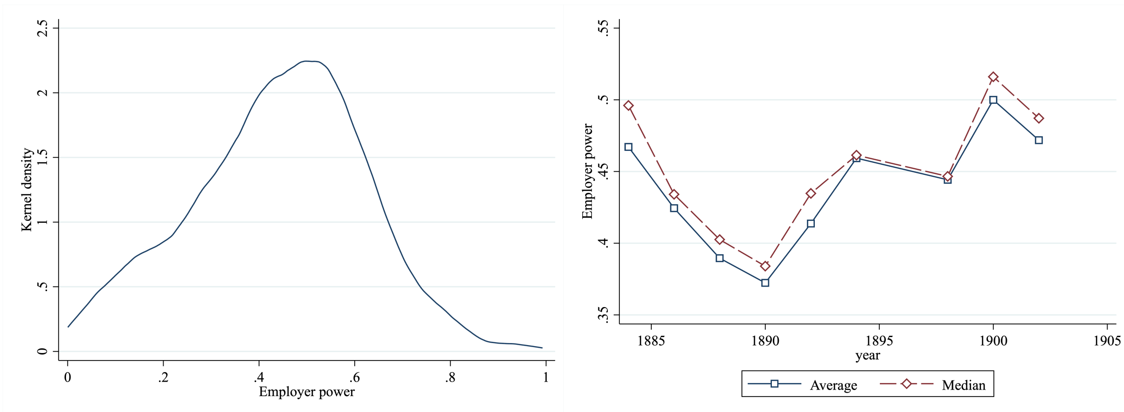
To estimate both the first and second stages of the production-function estimation, Equations (20) and (22), the bargaining parameters γ_{ft} need to be known, as they enter into the first-order condition for high-skilled labor demand. However, estimating the bargaining parameters γ_{ft} requires knowing the production-function coefficients, as is clear from Equation (24). I proceed by estimating the production function and the bargaining parameter using a fixed-point algorithm. I start with an initial value of $\gamma_{ft} = 0.5$ to estimate the production function and the bargaining parameter. Then, I use the resulting bargaining parameter to re-estimate the production function and the bargaining parameter. I continue this estimation loop until the production function coefficients converge, up to a sensitivity level of 0.001. I find that the estimates quickly converge to a fixed point.

Results

Table 2(c) reports the production function estimates. The elasticity of substitution between skilled and unskilled miners is estimated at 0.361, which implies that these two types of workers are gross complements. It is not surprising that this elasticity is relatively low, given that skilled miners were used for cutting coal whereas unskilled miners were used mainly for hauling coal, two tasks that are complementary. Cutting machines are estimated to increase skill-augmenting productivity by 21.4%, so cutting machines are a skill-augmenting technology. Given that skilled and unskilled labor are gross complements, this makes cutting machines an unskill-biased technology (Acemoglu, 2002), similar to many other technologies developed throughout the 19th century, that were also unskill-biased (Mokyr, 1990; Goldin & Katz, 2009). The finding that cutting machines were unskill-biased is consistent with the fact that cutting machines automated the cutting process, which was reliant on skilled miners, in contrast to the hauling process, which was mainly reliant on unskilled workers. Besides increasing skill-augmenting productivity, cutting machines also increased Hicks-neutral productivity by 11.3%, although this effect is not statistically significant. The low-skilled labor parameter β^l is estimated, imprecisely, at 0.006. Easier

Figure 4: Bargaining Parameter Estimates

(a) Employer Power: Distribution **(b) Employer Power: Evolution**



to interpret are the corresponding output elasticities of low- and high-skilled labor, which are estimated at 0.680 and 0.320, respectively. Finally, skill-augmenting and Hicks-neutral productivity have serial correlations of 0.288 and 0.829.

The employer's bargaining weight was 0.452 and 0.440 at the average and median firm, respectively, so bargaining power was roughly split equally between mine owners and the miners' union. I keep only the bargaining parameter values that range between zero and one, as values outside of this range are meaningless in the context of the bargaining model. This reduces the number of observations by 11%. Figure 4a shows the entire distribution of employer power, while Figure 4b shows the evolution of employer power on average and at the median firm. After an initial decrease in the first years of the sample, employer power increased by more than one-third between 1890 and 1902. This coincided with an aggregate decline in both nominal and real wages in the United States during the 1890s (Douglas & Lamberson, 1921).

Table 3 regresses the logarithm of the employer bargaining parameter over high-skilled workers ($\ln(1 - \gamma_f)$) on firm and market characteristics. The first two columns do not include firm fixed effects, the last two do. In both specifications, employers with higher market shares on the high-skilled worker market had more bargaining power. Moreover, employer power was increasing at a rate of 0.7%-1.0% per year. Employers in markets with more immigrants and a higher population share of African Americans had more bargaining power. This is consistent with the worse bargaining power faced by both immigrants and African Americans, and with the use of "strike-breakers" from the southern U.S. to decrease the bargaining ability of the labor unions. Finally, firms in counties with higher manufacturing wages had less bargaining power, as this provided a more favorable negotiation position for miners. Although the 1897-1898 strikes did not succeed in countering the overall rise in employer power, employer power did fall at the striking mines relatively to the mines that did not strike.[19]

Table 3: Employer Power: Covariates

	Log(Employer Bargaining Power)			
	Est.	S.E.	Est.	S.E.
High-skilled-employment market share	0.261	0.021	0.173	0.037
Low-skilled-employment market share	-0.288	0.021	-0.156	0.043
Year	0.010	0.002	0.007	0.003
Pop. share foreigners	0.863	0.281		
Pop. share Afro-Americans	2.425	0.663		
Log(manufacturing wage)	-0.177	0.066		
Firm FE:	No		Yes	
R-squared	.066		.670	
Observations	3459		3459	

Notes: In this table, I regress $\ln(1-\gamma_f)$ on firm and market characteristics.

4.5 Cutting-Machine Fixed Costs

Solving for Market Equilibrium Conditional on Machine Usage

Using the estimated model, I solve the system of Equations (9), (12a), (13), (15), and (16) for every firm in every year, both if using cutting machines and if not using cutting machines. Given that this system of equations is nonlinear and cannot be solved analytically,

[19]I document this in Appendix C.2.

I solve for equilibrium numerically.[20] For every outcome variable $Y \in \{Q, P, H, L, W\}$, this yields an equilibrium outcome if the firms used cutting machines, denoted as Y_{ft}^1, and if not, denoted as Y_{ft}^0.

Estimation of Fixed Costs

Using the estimated equilibrium, the cutting-machine probabilities at each firm in each year, $p_{ft}^k(\phi)$, can be computed using Equation (18), up to the unknown fixed-cost parameters (ϕ^0, ϕ^1). I estimate these fixed-cost parameters using a maximum likelihood estimator. Using Equation (18), the log likelihood function of using cutting machines $\ln(\mathcal{L}_{ft}(\phi)$ can be written as:

$$\ln(\mathcal{L}_{ft}(\phi)) = \sum_{f,t} [K_{ft} \ln(p^k(\phi)) + (1 - K_{ft}) \ln(1 - p_{ft}^k(\phi))]$$

I estimate the fixed cost parameters (ϕ^0, ϕ^1) by maximizing the log likelihood function $\ln(\mathcal{L}_{ft}(\phi))$. Because the number of observed capital-usage decisions is sparse, I do not rely on the observed capital usages K_{ft} in the raw data; I rather estimate a conditional choice probability $\tilde{K}_{ft}$ first by running a probit model of cutting-machine usage on log Hicks-neutral and labor-augmenting productivity, the labor supply shifter, and the coal demand shifter. I estimate this model on the entire sample and obtain predicted usage rates of cutting machines for every firm in every year. Next, I use these predicted usage rates in the log likelihood function to estimate cutting-machine fixed costs. The resulting estimates are in Table 2(d). The fixed cost of using a cutting machine is estimated to be \$23,680 in 1882, and is estimated to fall by \$2,358 every two years throughout the sample period. This decline in machine fixed cost is in line with the falling costs of many new technologies. I obtained external cost information for cutting machines in 1889 from Brown (1889), who reported a purchasing cost of \$8,000 for eight cutting machines. The average firm in the dataset also used around eight cutting machines, so the estimated fixed cost in 1888 of

[20] I use the Matlab optimizer `fsolve` with function tolerance 10^{-3}, 10^5 maximum iterations, and 600 maximum function evaluations.

\$4,816 is lower than the external cost estimate, although in the same order of magnitude.

Solving for Market Equilibrium and Equilibrium Machine Usage

Using Equation (18), I estimate equilibrium cutting machine usage for every firm in every year. I compute the equilibrium values $\hat{Y}_{ft}$ for variables $Y \in \{Q, P, H, L, W\}$ as the weighted average of the value when the firm used machine usage and when if did not, weighted by the probability of using cutting machines:

$$\hat{Y}_{ft} = Y^0_{ft} Pr(K_{ft} = 0) + Y^1_{ft} Pr(K_{ft} = 1) \tag{25}$$

In Appendix B.2, I discuss how the predicted equilibrium values $\hat{\mathbf{Y}}_{ft}$ compare to the observed variables $\mathbf{Y}_{ft}$. As shown in Appendix Figure A2, the model closely tracks the observed evolution of all equilibrium variables, though the estimation procedure does not explicitly target any of these moments, except for cutting-machine usage rates.

4.6 Counterfactual: Welfare Effects of Labor Market Power

To examine the effects of employer power, I compute how all equilibrium outcomes and welfare would have changed if the bargaining power of employers would increase by one standard deviation, which I denote as a counterfactual bargaining parameter $\tilde{\gamma}_{ft}$.

Welfare

In both the actual and counterfactual equilibria, I compute consumer surplus CS_{ft} as the area in between the demand curve and the equilibrium price:

$$CS_{ft} \equiv \int_0^{\hat{Q}_{ft}} (P_0 \Big(\frac{Q_{ft}}{\xi_{ft}}\Big)^{\frac{1}{\eta}} - \hat{P}_{ft}) dQ_{ft} = \Big(\frac{-1}{\eta + 1}\Big) \frac{P_0}{\xi^{\frac{1}{\eta}}} (\hat{Q}_{ft})^{\frac{\eta+1}{\eta}}$$

Similarly, I compute worker surplus WS_{ft} as the area between the labor supply curve and the equilibrium wage $\hat{W}_{ft}$:

$$WS_{ft} \equiv \int_0^{\hat{H}_{ft}} (\hat{W}_{ft} - W_{0,i(f)t}\Big(\frac{H_{ft}}{\zeta_{ft}}\Big)^{\psi})dH_{ft} = \Big(\frac{\psi}{\psi+1}\Big)\frac{W_0}{\zeta_{ft}^{\psi}}(\hat{H}_{ft})^{\psi+1}$$

Finally, producer surplus is equal to variable employer profits:

$$PS_{ft} \equiv \hat{P}_{ft}\hat{Q}_{ft} - \hat{H}_{ft}\hat{W}_{ft} - \hat{L}_{ft}V_t$$

Counterfactual Results

Table 4a reports the equilibrium effects of a one-standard-deviation increase in employer power. If capital investment would be exogenous, coal output would decrease by 25.1%, because the exertion of monopsony power induces deadweight loss. However, capital investment does not remain fixed: the increase in employer power results in an increase in cutting-machine usage of 14.1%.[21] As a result, the output decline due to monopsony power is lower, at 23.2%, when taking into account endogenous cutting-machine usage, given that increased investment lowers marginal costs. The net effect on equilibrium output is still negative because the deadweight loss effect dominates the reduction in marginal costs due to higher technology adoption. Similarly, the exogenous-investment model overestimates both the decline in employment and wages and the increase of the coal price. Both Hicks-neutral and skill-augmenting productivity increase on average when employer power increases, due to higher capital investment.

Table 4b. reports the welfare counterfactuals. Consumer surplus is estimated to fall by 21.2% on average in the exogenous-investment model, which reduces to a 19.8% drop in the endogenous-investment model. Similarly, the reduction in labor surplus changes from 37.3% to 35.6% once accounting for endogenous capital usage.

Under increased employer power, consumers and workers experience a drop in sur-

[21] The usage rate of 10% in the observed equilibrium is slightly lower than the usage rate of 9% in Figure A2e because the comparison table conditions on the equilibrium being solved in both the actual and the counterfactual equilibrium. This leads to a different sample selection with slightly higher capital-usage rates.

plus, although consumer surplus could in theory have increased if the technology-adoption mechanism would dominate the deadweight-loss effect. In contrast, producer surplus obviously increases when employers gain bargaining power. Assuming exogenous investment, producer surplus would increase by 10.5%, and by 12.6% under endogenous investment. The increased producer surplus is insufficient to compensate the losses to consumer and labor surplus, as total surplus falls by 10.9%. However, assuming exogenous investment would predict a total surplus drop of 12.5%, which overestimates the true welfare loss by 15.4%.

Table 4: Counterfactual: Increased Employer Power

(a) Equilibrium		Exogenous K		Endogenous K	
	Actual	C.F.	%diff	C.F.	%diff
Output	1599.327	1198.244	-25.078	1228.505	-23.186
Price	1.727	1.763	2.086	1.760	1.923
High-skilled labor	672.052	499.661	-25.651	505.802	-24.738
Low-skilled labor	232.268	178.024	-23.354	181.353	-21.921
High-skilled wage	1.873	1.717	-8.324	1.720	-8.133
Cutting machine usage	0.104	0.104	0.000	0.119	14.075
Hicks-neutral productivity	1.689	1.689	0.000	1.691	0.077
Skill-augmenting productivity	5.309	5.309	0.000	5.323	0.275
(b) Welfare		Exogenous K		Endogenous K	
	Actual	C.F.	%diff	C.F.	%diff
Consumer surplus	508.631	400.995	-21.162	407.728	-19.838
Producer surplus	460.257	508.660	10.517	518.191	12.587
Worker surplus	251.425	157.687	-37.283	161.859	-35.623
Total surplus	1220.313	1067.342	-12.535	1087.778	-10.861

Notes: Panel (a) reports averages for all equilibrium outcomes in 1902 in the observed equilibrium (the "Actual" column) and in the counterfactual equilibrium where employer power does not increase (the "C.F." columns). The first C.F. column keeps cutting machine usage exogenous, whereas the second C.F. column allows cutting machine usage to be endogenous to the degree of bargaining power held by employers.

Factor Intensity and Monopsony Power

The equilibrium employment changes in Table 4a reveal that increased employer power leads to a reduction of the high-skilled-to-low-skilled labor ratio, from 2.89 to 2.79. This reduction in relative high-skilled labor usage is due to a combination of two effects. First, the increase in employer power affects employer bargaining power only over high-skilled workers, not over low-skilled workers, who are supplied on a perfectly competitive market. Hence, increased monopsony power over high-skilled workers increases the marginal cost for these workers and induces employers to substitute low-skilled workers for high-skilled workers. This is the mechanism documented in Goolsbee and Syverson (2023).

However, a second mechanism is at play. The increase in employer power results in higher cutting-machine adoption, which is an unskill-biased technology. This leads to an additional substitution of low-skilled workers for high-skilled workers. Comparing the exogenous-investment to the endogenous-investment counterfactual shows that 83% of the reduction in the high-skilled-to-low-skilled labor ratio was due to the monopsony-induced substitution effect, whereas the remaining 17% was due to increased cutting-machine adoption.

5 Conclusion

In this paper, I investigate the welfare effects of employer power by studying the trade-off between monopsony distortions and endogenous investment. Using a model of production and labor supply that allows for monopsony power, wage bargaining, and imperfectly competitive goods markets, I find that an increase in employer power could either increase or decrease output and total welfare, depending on the relative size of the monopsony distortion, of the marginal-cost reduction due to endogenous investment, and on the initial level of employer power. In the empirical context of the mechanization of the late-19th-century Illinois coal mining industry, I find that an increase in employer power lowered equilibrium output because the monopsony distortion dominated the marginal-cost reduction that was due to the adoption of additional coal-cutting machines. Although total welfare declined when employer power increases, this decline is 15% smaller than one would find when holding capital investment fixed. Hence, the model and the results show that taking into account endogenous capital quantitatively matters for assessing the welfare effects of labor market power.

References

Abowd, J. A., & Lemieux, T. (1993). The Effects of Product Market Competition on Collective Bargaining Agreements: The Case of Foreign Competition in Canada. *The Quarterly Journal of Economics*, *108*(4), 983–1014.

Acemoglu, D. (2002). Directed Technical Change. *The Review of Economic Studies*, *69*(4), 781–809.

Acemoglu, D., & Shimer, R. (1999). Holdups and efficiency with search frictions. *International Economic Review*, *40*(4), 827–849.

Ackermann, A. (1902). *Coal-cutting by machinery in America* (No. 16). The Colliery Guardian Company Limited.

Affolter, R. H., & Hatch, J. R. (2002). Characterization of Quality of Coals from the Illinois Basin. In R. H. Affolter & J. R. Hatch (Eds.), (Vol. 1625, p. 1-227). U.S. Geological Survey.

Ashenfelter, O. C., Farber, H., & Ransom, M. R. (2010). Labor Market Monopsony. *Journal of Labor Economics*, *28*(2), 203-210.

Atack, J., Bateman, F., & Margo, R. A. (2008). Steam power, establishment size, and labor productivity growth in nineteenth century American manufacturing. *Explorations in Economic History*, *45*(2), 185–198.

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

Berger, D., Herkenhoff, K., & Mongey, S. (2022). Labor Market Power. *American Eco-*

nomic Review, *112*(4), 1147–93.

Bloch, L. (1922). *The Coal Miners' Insecurity: Facts about Irregularity of Employment in the Bituminous Coal Industry in the United States* (No. 7). Russell Sage Foundation.

Boal, W. M. (1995). Testing for Employer Monopsony in Turn-of-the-Century Coal Mining. *RAND Journal of Economics*, *26*(3), 519-536.

Boal, W. M. (2017). What did Unions do? The case of Illinois Coal Mining in the 1880s. *Journal of Labor Research*, *38*(4).

Brown, W. (1889). *Transactions of the federated institution of mining engineers* (Tech. Rep. No. 1). Federated Institution of Mining Engineers.

Buyst, E., Goos, M., & Salomons, A. (2018). Job polarization: an historical perspective. *Oxford Review of Economic Policy*, *34*(3), 461–474.

Chambolle, C., Christin, C., & Molina, H. (2023). Buyer Power and Exclusion: A Progress Report. *International Journal of Industrial Organization*, 102969.

Crawford, G. S., & Yurukoglu, A. (2012). The welfare effects of bundling in multichannel television markets. *American Economic Review*, *102*(2), 643–685.

Degrève, D. (1982). *Le commerce extérieur de la Belgique, 1830-1913-1939: Présentation critique des données statistiques* (Vol. 1b). Brussels: Palais des Académies.

Delabastita, V., & Rubens, M. (2024). Colluding Against Workers. *Available at SSRN*.

Doraszelski, U., & Jaumandreu, J. (2018). Measuring the Bias of Technological Change. *Journal of Political Economy*, *126*(3), 1027-1084.

Douglas, P. H., & Lamberson, F. (1921). The movement of real wages, 1890-1918. *The American Economic Review*, *11*(3), 409–426.

Fishback, P. V. (1992). *Soft coal, hard choices: The economic welfare of bituminous coal miners, 1890-1930*. Oxford University Press.

Fishback, P. V., Holmes, R., & Allen, S. (2009). Lifting the curse of dimensionality: measures of states' labor legislation climate in the United States during the progressive era. *Labor History*, *50*(3), 313-346.

Goldin, C., & Katz, L. F. (1998). The origins of technology-skill complementarity. *The*

Quarterly Journal of Economics, *113*(3), 693–732.

Goldin, C., & Katz, L. F. (2009). *The race between education and technology*. Harvard University Press.

Goolsbee, A., & Syverson, C. (2023). Monopsony power in higher education: A tale of two tracks. *Journal of Labor Economics*, *41*(S1), S257-S290.

Graebner, W. (1974). Great expectations: The search for order in bituminous coal, 1890-1917. *The Business History Review*, 49–72.

Grennan, M. (2013). Price discrimination and bargaining: Empirical evidence from medical devices. *American Economic Review*, *103*(1), 145–177.

Hemphill, C. S., & Rose, N. L. (2018). Mergers that harm sellers. *Yale Law Journal*, *127*, 2078.

Hoover, E. D. (1960). Retail prices after 1850. In *Trends in the American economy in the nineteenth century* (pp. 141–190). Princeton University Press.

Horn, H., & Wolinsky, A. (1988). Bilateral monopolies and incentives for merger. *The RAND Journal of Economics*, 408–419.

Hornbeck, R., Hsu, S. H.-M., Humlum, A., & Rotemberg, M. (2024). *Gaining Steam: Incumbent Lock-in and Entrant Leapfrogging* (Tech. Rep.). National Bureau of Economic Research.

Huang, S.-Y., & Sexton, R. J. (1996). Measuring returns to an innovation in an imperfectly competitive market: Application to mechanical harvesting of processing tomatoes in Taiwan. *American Journal of Agricultural Economics*, *78*(3), 558–571.

Jarosch, G., Nimczik, J. S., & Sorkin, I. (2024). Granular search, market structure, and wages. *The Review of Economic Studies*.

Joskow, P. L. (1987). Contract duration and relationship-specific investments: Empirical evidence from coal markets. *The American Economic Review*, 168–185.

Joyce, R. (2009). *The Early Days of Coal Mining in Northern Illinois.* Illinois Labor History Society. (Accessed through http://www.illinoislaborhistory.org/labor-history-articles/early-days-of-coal-mining-in-northern-illinois)

Just, R. E., & Chern, W. S. (1980). Tomatoes, technology, and oligopsony. *The Bell Journal of Economics*, 584–602.

Katz, L. F., & Margo, R. A. (2014). Technical change and the relative demand for skilled labor: The United States in historical perspective. In *Human capital in history: The American record* (pp. 15–57). University of Chicago Press.

Köhler, C., & Rammer, C. (2012). Buyer power and suppliers' incentives to innovate. *ZEW-Centre for European Economic Research Discussion Paper*.

Kroft, K., Luo, Y., Mogstad, M., & Setzler, B. (2020). *Imperfect Competition and Rents in Labor and Product Markets* (Tech. Rep.). National Bureau of Economic Research.

Lamadon, T., Mogstad, M., & Setzler, B. (2022). Imperfect competition, compensating differentials, and rent sharing in the US labor market. *American Economic Review*, *112*(1), 169–212.

Loertscher, S., & Marx, L. M. (2019). Merger review for markets with buyer power. *Journal of Political Economy*, *127*(6), 2967–3017.

Lord, J. S. (1883). *Statistics of Coal Production in Illinois, 1883: A Supplemental Report of the State Bureau of Labor Statistics*.

Lord, J. S. (1892). *Coal in Illinois, 1884-1892* (Biennial Reports No. 2-7).

Manning, A. (2011). Imperfect competition in the labor market. In *Handbook of labor economics* (Vol. 4, pp. 973–1041). Elsevier.

Menezes-Filho, N., & Van Reenen, J. (2003). Unions and innovation: a survey of the theory and empirical evidence. *International handbook of trade unions*, 293–334.

Mokyr, J. (1990). *Twenty five centuries of technological change: An historical survey* (Vol. 35). Taylor & Francis.

Morlacco, M. (2017). *Market power in input markets: theory and evidence from French manufacturing* (Tech. Rep.).

Naidu, S., Nyarko, Y., & Wang, S.-Y. (2016). Monopsony Power in Migrant Labor Markets. *Journal of Political Economy*, *124*(6), 1735-1792.

Naidu, S., & Yuchtman, N. (2017). Labor Market Institutions in the Gilded Age of American History. *The Oxford Handbook of American Economic History*.

Parra, A., & Marshall, G. (2024). Monopsony power and upstream innovation. *Journal of Industrial Economics*, *72*(2), 1005-1020.

Peters, B., Roberts, M., Vuong, V. A., & Fryges, H. (2017). Estimating Dynamic R&D Demand: An Analysis of Costs and Long-Run Benefits. *RAND Journal of Economics*, *48*(2), 409-437.

Reid, A. (1876). *Transactions of the North of England Institute of Mining and Mechanical Engineers*. North of England Institute of Mining and Mechanical Engineers.

Rubens, M. (2023). Market Structure, Oligopsony Power, and Productivity. *American Economic Review*, *113*(9), 2382–2410.

Rubens, M., Wu, Y., & Xu, M. (2024). *Exploiting or Augmenting Labor?* (Tech. Rep.).

Shi, L. (2023). Optimal regulation of noncompete contracts. *Econometrica*, *91*(2), 425–463.

Starr, E., Prescott, J. J., & Bishara, N. (2021). Noncompetes in the US labor force. *Journal of Law and Economics*.

Stoek, H., Hippard, C., & Langtry, W. (1920). Bituminous Coal Storage Practice. *Bulletin of the University of Illinois Engineering Experiment Station*, *116*, 1-142.

Van Reenen, J. (1996). The creation and capture of rents: wages and innovation in a panel of UK companies. *The Quarterly Journal of Economics*, *111*(1), 195–226.

Whitcomb, G. D. (1882, 08 8). *Coal mining machine* (No. 262225).

Williams, T. (1901). The Anthracite Coal Crisis. *The Atlantic*, *April*.

Williamson, O. E. (1971). The vertical integration of production: market failure considerations. *The American Economic Review*, *61*(2), 112–123.

Zahur, N. B. (2022). Long-term contracts and efficiency in the liquefied natural gas industry. *Available at SSRN 4222408*.

Appendices

A Appendix: Data

A.1 Sources

Mine Inspector Reports

My main data source is the biennial report of the Bureau of Labor Statistics of Illinois, of which I collected the volumes between 1884 and 1902. Each report contains a list of all mines in each county and reports the name of the mine owner, the town in which the mine is located, and a selection of variables that varies across the volumes. An overview of all the variables (including unused ones), and the years in which they are observed, is in Tables A6 and A7. Output quantities, the number of miners and other employees, mine-gate coal prices, and information on the usage of cutting machines are reported in every volume. Miner wages and the number of days worked are reported in every volume except 1896. The other variables, which include information about the mine type, hauling technology, other technical characteristics, and other inputs, are reported in a subset of years.

Census of Population, Agriculture, and Manufacturing

I digitized the 1880 Illinois population census census of manufacturing, and census of agriculture by accessing copies of the original prints from the HathiTrust Digital Library[22]. I rely on the population census to gather information on county population sizes, demographic compositions, and areas. I also observe the county-level capital stock and employment in manufacturing industries from the 1880 Illinois census of manufacturing, and the number of farms and improved farmland area from the 1880 Illinois census of agriculture.

[22] https://www.hathitrust.org/

Monthly Data

The 1888 report contains monthly production data for a selection of 11 mines in Illinois, across six counties. I observe the monthly number of days worked and the number of skilled and unskilled workers. I also observe the net earnings for all skilled and unskilled workers per mine per month, and the number of tons mined per worker per month. This allows me to compute the daily earnings of skilled and unskilled workers per month.

A.2 Data Cleaning

Employment

In every year except 1896, workers are divided into two categories, "miners" and "other employees." In 1896, a different distinction is made, between "underground workers" and "above-ground workers." This does not correspond to the miner-others categorization, because all miners were underground workers but some underground workers were not miners (e.g., doorboys, mule drivers). Hence, I do not use the 1896 data.

From 1888 to 1896, boys are reported as a separate working category. Given that miners (cutters) were adults, I include these boys in the "other employees" category. The number of days worked is observed for all years. The average number of other employees per mine throughout the year is observed in every year except 1896; in 1898, it is subdivided into underground other workers and above-ground other workers, which I add up into a single category.

The quantity of skilled and unskilled labor is calculated by multiplying the number of days worked with the average number of workers in each category throughout the year. Up to and including 1890, the average number of miners is reported separately for winters and summers. I calculate the average number of workers during the year by taking the simple average of summers and winters. If mines closed down during winters or, more likely, summers, I calculate the annual amount of labor-days by multiplying the average number of workers during the observed season with the total number of days worked during the year.

Wages

Only miner wages are consistently reported over time at the mine level. The piece rate for miners is reported. Up to 1894, miner wages per ton of coal are reported separately for summers and winters. I weight these seasonal piece-rate wages using the number of workers employed in each season for the years 1884 to 1890. In 1892 and 1894, seasonal employment is not reported, so I take simple averages of the seasonal wage rates. In 1896, wages are unobserved. From 1898 onward, wages are no longer reported seasonally, because wages were negotiated biennially from that year onwards. For these years, wages are reported separately for hand and machine miners. In the mines that employed both hand and machine miners, I take the average of these two piece rates, weighted by the amount of coal cut by hand and cutting machines.

Output Quantity and Price

The total amount of coal mined is reported in every year, in short tons (2,000 lbs). Up to and including 1890, the total quantity of coal extraction is reported, without distinguishing different sizes of coal pieces. After 1890, coal output is reported separately between "lump" coal (large pieces) and smaller pieces, which I sum in order to ensure consistency in the output definition. Mine-gate prices are normally given on average for all coal sizes, except in 1894 and 1896, where they are only given for lump coal (the larger chunks of coal). I take the lump price to be the average coal price for all coal sizes in these two years. There does not seem to be any discontinuity in the time series of average or median prices between 1892 and 1894 or 1896 and 1898 after doing this, which I see as motivating evidence for this assumption.

Cutting Machine Usage

Between 1884 and 1890, the number of cutting machines used in each mine is observed. Between 1892 and 1896, a dummy is observed for whether coal was mined by hand, using cutting machines, or both. I categorize mines using both hand mining and cutting machines as mines using cutting machines. In 1898, I infer cutting-machine usage by looking at which mines paid “machine’ wages” and “hand wages” (or both). In 1888, the number of cutting machines is reported by type of cutting machine as well. Finally, in 1900 and 1902, the output cut by machines and by hand is reported separately for each mine, on the basis of which I again know which mines used cutting machines, and which did not.

Deflators

I deflate all monetary variables using the Consumer Price Index from the *Handbook of Labor Statistics* of the U.S. Department of Labor, as reported by the Minneapolis Federal Reserve Bank website.[23]

Hours Worked

In 1898, eight-hour days were enforced by law for the first time, which means that the “number of days” measure changes in unit between 1898 and 1900. Because the Inspector Report from 1886 shows that ten-hour days were the standard, I multiply the number of working days after 1898 by 80% in order to ensure consistency in the meaning of a “work-day”, i.e., to ensure that in terms of the total number of hours worked, the labor-quantity definition does not change after 1898. Given that the model is estimated on the pre-1898 period, this does not affect the model estimates, only the descriptive evidence.

Mine and Firm Identifiers

The raw dataset reports mine names, which are not necessarily consistent over time. Based on the mine names, it is often possible to infer the firm name as well, in the case of multi-mine firms. For instance, the Illinois Valley Coal Company No. 1 and Illinois Valley Coal Company No. 2 mines clearly belong to the same company. For single-mine firms, the operator is usually mentioned as the mine name, (e.g. “Floyd Bussard”). For the multi-mine firms, I made mine names consistent over time as much as possible.

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

Town Identifiers and Labor Market Definitions

The dataset contains town names. I link these names to geographical coordinates using Google Maps. I calculate the shortest distance between every town in the data. For towns that are located less than two miles from each other, I merge them and assign them randomly the coordinates of either of the two mines. This reduces the number of towns in the dataset from 480 to 391. The resulting labor markets lie at least two miles from the nearest labor market.

Coal-Market Definitions

Using the 1883 Inspector Report, I link every coal mining town to a railroad line, if any. Some towns are located at the intersection of multiple lines, in which case I assign the town to the first line mentioned. I make a dummy variable that indicates whether a railroad is located at the crossroads of multiple railroad lines. Given that data from 1883 is used, expansion of the railroad network after 1883 is not taken into account. However, the Illinois railroad network was already very dense by 1883.

Aggregation From Mine to Firm Level

I aggregate labor from the mine-bi-year- to firm-bi-year level by taking sums of the total days worked and labor expenses for both types of workers, both per year and per season. I calculate the wage rates for both types per worker by dividing firm-level labor expenditure by the firm-level number of labor-days. I also sum powder usage, coal output, and revenue to the firm level and calculate the firm-level coal price by dividing total firm revenue by total firm output. I aggregate mine depth and vein thickness by taking averages across the different mines of the same firm. I define the cutting-machine dummy at the firm level as the presence of at least one cutting machine in one of the mines owned by the firm. I define a “firm” as the combination of the firm name in the dataset and its town (the merged towns that are used to define labor markets), as firms are assumed to optimize input usage on a town-by-town basis.

B Appendix: Simulations and Model Fit

B.1 Simulating the Theoretical Model

Baseline Parametrization

In Section 2.3, I simulate the theoretical model with the following parameter values. I use the estimates from Kroft et al. (2020) for the U.S. construction industry to set the product-demand elasticity to $\eta = -7$ and the inverse labor-supply elasticity to $\psi = 0.25$. I calibrate the elasticity of substitution between high- and low-skilled labor at $\sigma = 0.7$. I normalize most parameters at one: $\xi = 1$, $\zeta = 1$, $w_0 = p_0 = v = 1$, $\omega = 1$, $a = 1$. I set the low-skilled production coefficient at 0.2: $\beta^l = 0.1$. I simulate a dataset with 50 observations, in which the bargaining parameter γ_f is distributed uniformly between 0 and 1. I let fixed technology costs be distributed as an exponential distribution with a mean of 0.05.

Under these parametrizations, I solve the system of equations (1), (2), (6), (7), (8) for equilibrium (Q, P, W, H, L).

Alternative Parametrizations

In Figure A1, I compare the baseline calibration of the structural model to various alternative parametrizations. First, I let labor supply be more inelastic. Second, I increase the productivity effects of the new technology.

B.2 Model Fit

Figure A2 compares the model-predicted equilibrium outcomes against the observed outcomes in the data. The model is not estimated to target any of these outcomes, except for capital investment, through the maximum-likelihood estimation of fixed technology costs. Nevertheless, the model generates a very similar evolution of average wages, prices, employment, output, and investment between the predicted and observed outcomes. Although the model performs well in terms of generating the observed evolution of these variables over time, it performs slightly less well in terms of absolule magnitudes. The model-predicted output and employment levels are underestimated compared to the truth, and coal prices are overestimated.

C Appendix: Extensions and Robustness Checks

C.1 Alternative Production-Function Specifications

Nonconstant Returns to Scale

In the main text, the production function (1) relied on constant returns to scale. In contrast, Equation (28) allows for nonconstant returns to scale, as parametrized by ν:

$$Q_f = \left((A_f(K_f)H_f)^{\frac{\sigma-1}{\sigma}} + \beta^l L_f^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\nu\sigma}{\sigma-1}} \Omega_f(K_f) \tag{26}$$

The first step of the production function estimation procedure, the estimation of Equation (20), remains the same. However, the second step of the estimation procedure needs to estimate the scale parameter ν in addition to the other production function coefficients ρ^ω, β^l, and β^k. Given that we have four instruments (lagged employment for both labor types, current and lagged capital), the model is still identified.

$$q_{ft} = \frac{\nu\sigma}{\sigma-1} \ln\left(\left(\exp\left((\frac{l_{ft}-h_{ft}}{1-\sigma}) - \frac{\sigma}{1-\sigma}(\ln(\beta^l)) - \frac{\sigma}{1-\sigma}(w_{ft} - v_t + \ln(1+(1-\gamma_{ft})\psi))\right)H_{ft}\right)^{\frac{\sigma-1}{\sigma}} \right.$$
$$\left. + \beta^l L_{ft}^{\frac{\sigma}{\sigma-1}}\right) + \omega_{ft}$$

The results are in the first column of Table A4. The scale parameter is estimated at 1.032, which indicates modestly increasing returns to scale, but is not significantly different from 1. Hence, the assumption of constant returns to scale cannot be rejected. The other production coefficients look very similar to the estimates in the main model, which assumes constant returns to scale.

Adding Materials

As a second robustness check, I add the materials to the production function as a third production input. I use the number of kegs of black powder to measure materials, as this is the main intermediate input that is measured in the dataset. This implies that a fifth coefficient, β^m, needs to be estimated. I assume that changing the stock of black powder requires adjustment costs: black powder is a durable good but needs to be safely stored. Hence, it is conceivable that there was an adjustment cost when increasing the stock of black powder, as additional storage space was needed. Conforming with this assumption, I include current and lagged materials as an additional instrument when estimating the production function:

$$Q_f = \left((A_f(K_f)H_f)^{\frac{\sigma-1}{\sigma}} + \beta^l L_f^{\frac{\sigma-1}{\sigma}} + \beta^m M_f^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\nu\sigma}{\sigma-1}} \Omega_f(K_f) \tag{27}$$

The estimates are in the second column of Table A4. The material coefficient is estimated to be very close to zero, which means that ignoring materials in the main production model does not matter much. The remaining production coefficient look very similar to the previous ones, with the exception of the serial correlation in TFP, which increases to 0.516.

Capital and Returns to Scale

The degree of returns to scale may have changed when firms adopted cutting machines. To test this, I interact the returns-to-scale parameter with the cutting-machine indicator variable, thereby allowing returns to scale to differ between firms that do and do not use cutting machines. Now, an additional instrument is needed to identify all six parameters in the production function. I rely on nonfatal accident rates as shifters of labor supply, which should affect input usage but not productivity directly. I measure the probability of non-fatal accidents as the ratio of the number of such accidents over total employment at the mine, in days worked:

$$Q_f = \left((A_f(K_f)H_f)^{\frac{\sigma-1}{\sigma}} + \beta^l L_f^{\frac{\sigma-1}{\sigma}} + \beta^m M_f^{\frac{\sigma-1}{\sigma}}\right)^{\frac{(\nu_0+\nu_1 K_f)\sigma}{\sigma-1}} \Omega_f(K_f) \tag{28}$$

The estimates are in the third column of Table A4. The interaction effect between returns to scale and cutting machines is close to zero and not statistically significant. Hence, the null hypothesis that returns to scale are invariant to cutting-machine usage cannot be rejected.

C.2 Additional Results

Strikes and Employer Power

In this appendix, I repeat the difference-in-differences analysis from Section 3.3, but now use the log of the estimated bargaining parameter γ_{ft} as the left-hand side variable, in order to examine how employer power changed in response to the 1897 strike. In the left column of Table A3, I compare all firms with strikes to the non-striking firms. The bargaining power of the labor union is estimated to increase by 10.0% at striking firms compared to non-striking firms, although this effect is not statistically significant. When only considering the striking firms at which wages increased after the strike, the increase in union bargaining power is higher, at 19.8%, and becomes statistically significant.

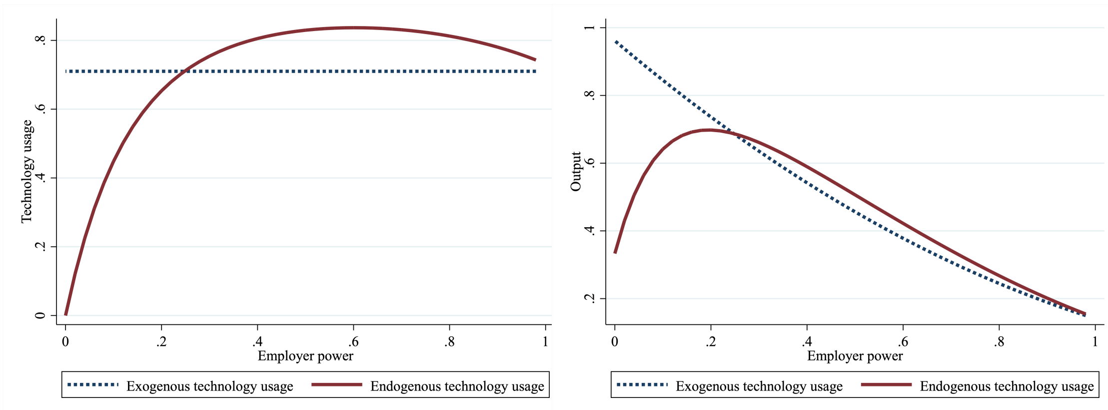
Cost Dynamics

In Table A2, in the spirit of Benkard (2000), I test for cost dynamics by regressing labor productivity, measured as output per labor-day, on log cumulative output. I find that when not taking mine fixed effects, cumulative past output correlates with higher productivity. However, this is likely due to a selection effect: more productive mines exist longer and produce more. As soon as I include mine fixed effects and look at time-series variation in productivity within mines, the relationship between log cumulative output and labor productivity vanishes. This suggests that cost dynamics are not a key feature to be included in the model.

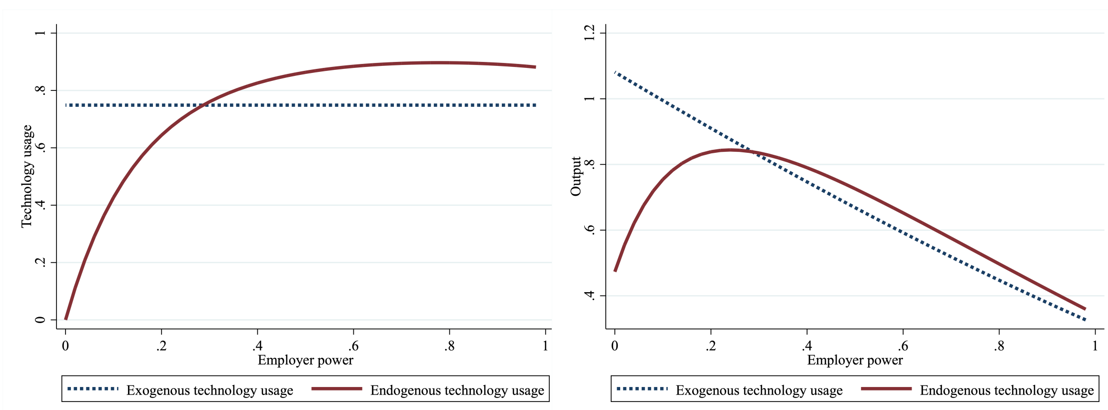
C.3 Appendix Tables and Figures

Figure A1: Simulations: Alternative Parametrization

(a) Baseline: $\psi = 0.25, \beta^k = 0.2$



(b) $\psi = 0.5, \beta^k = 0.2$



(c) $\psi = 0.25, \beta^k = 0.05$

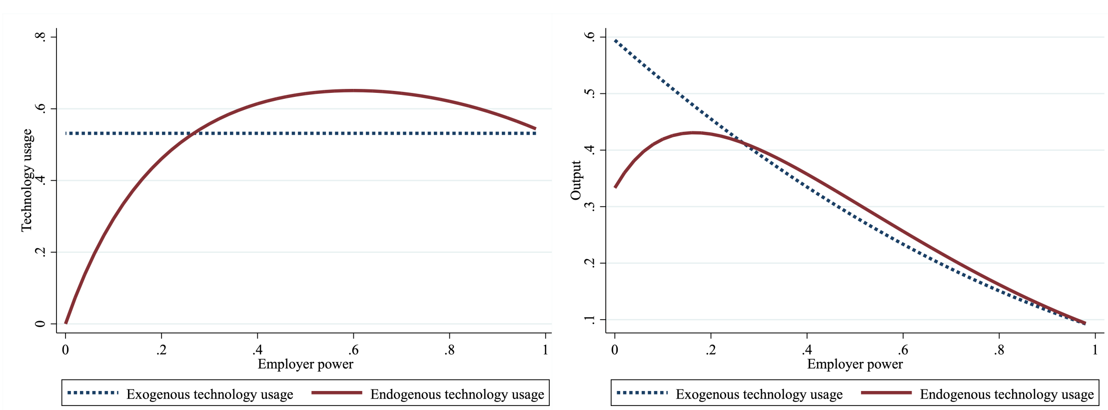
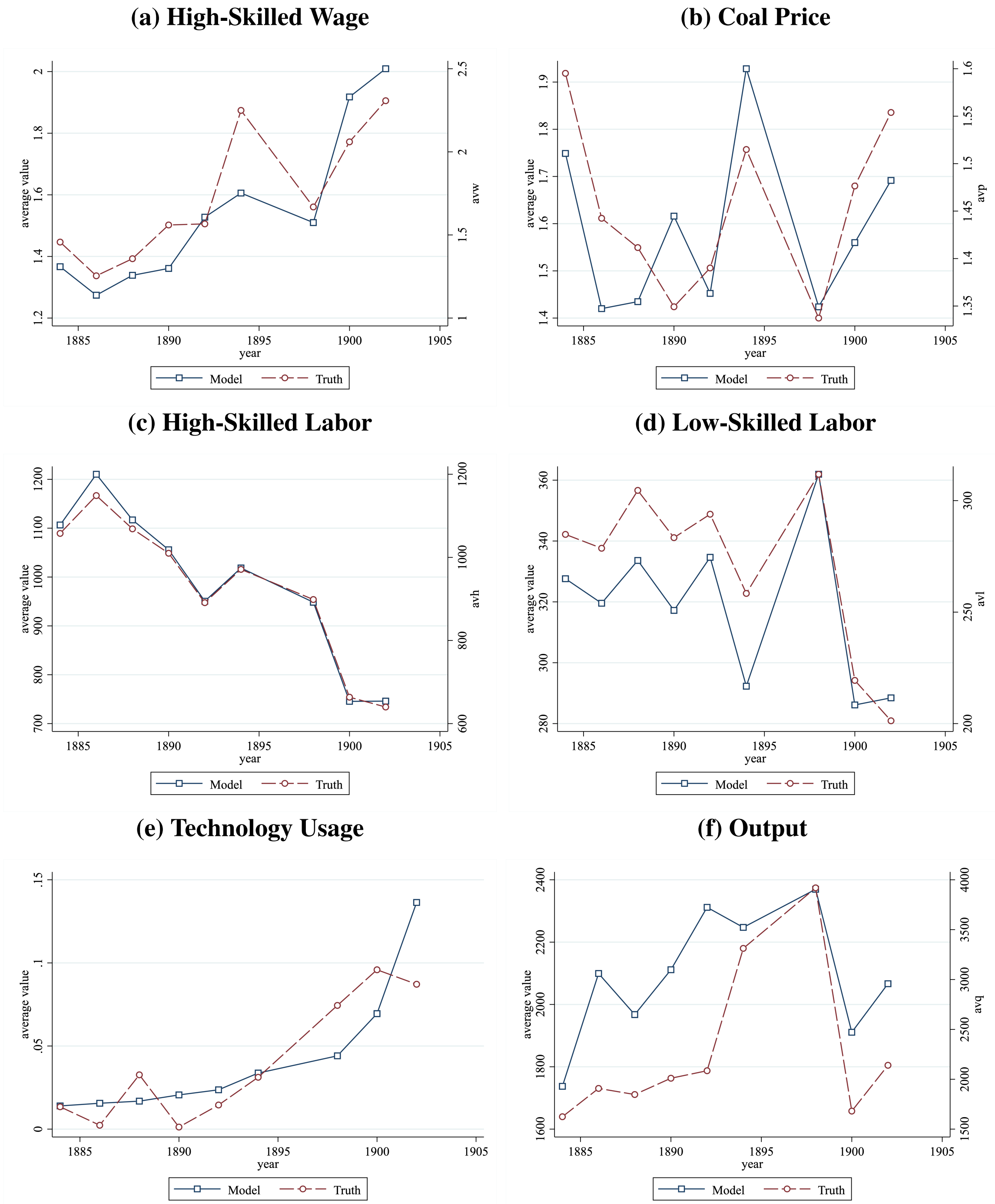
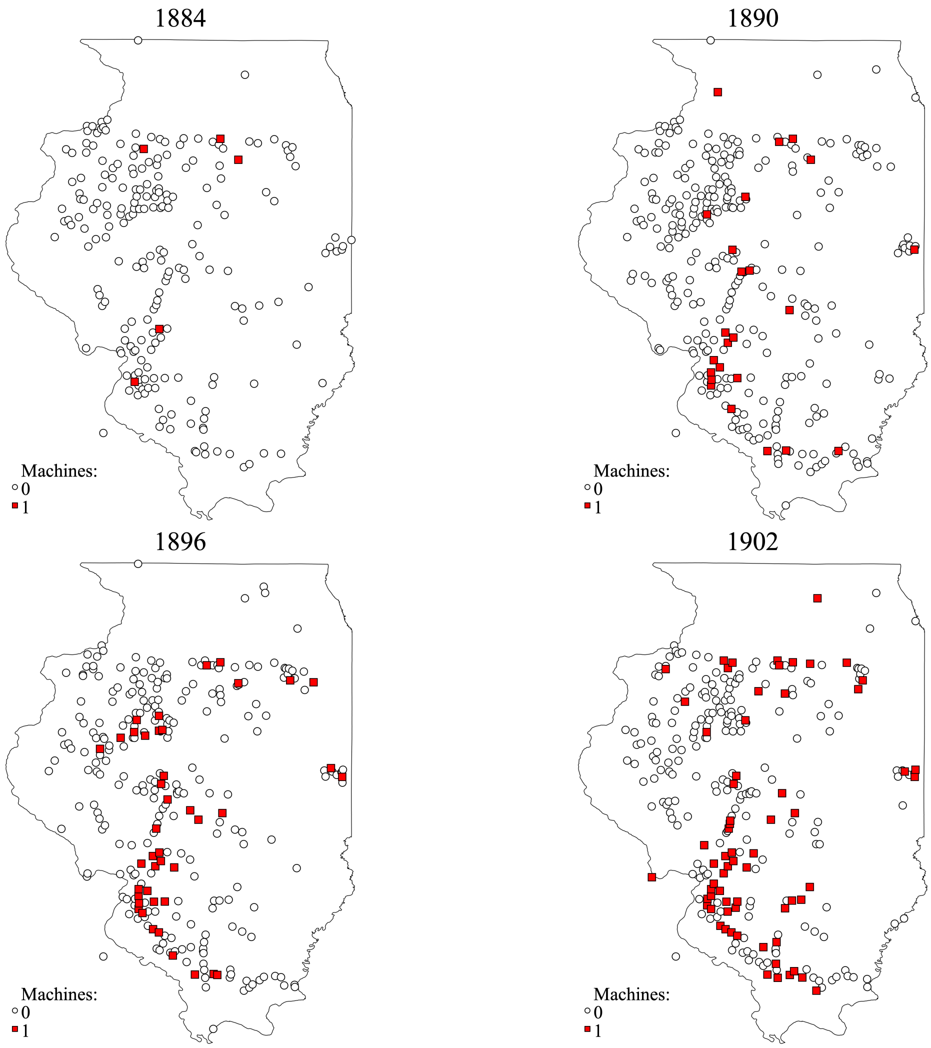


Figure A2: Model Fit



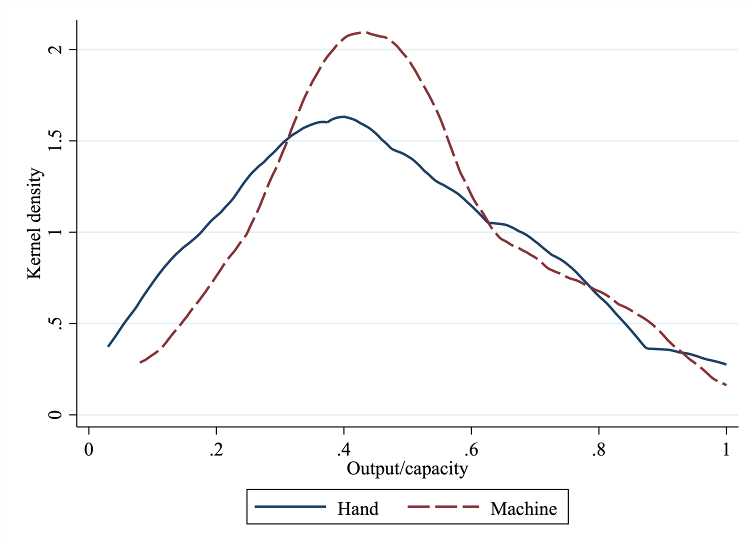
Notes: Figures compare the average equilibrium variables between their observed values and the predicted values from the model, for each year.

Figure A3: Geographical Spread of Cutting Machines



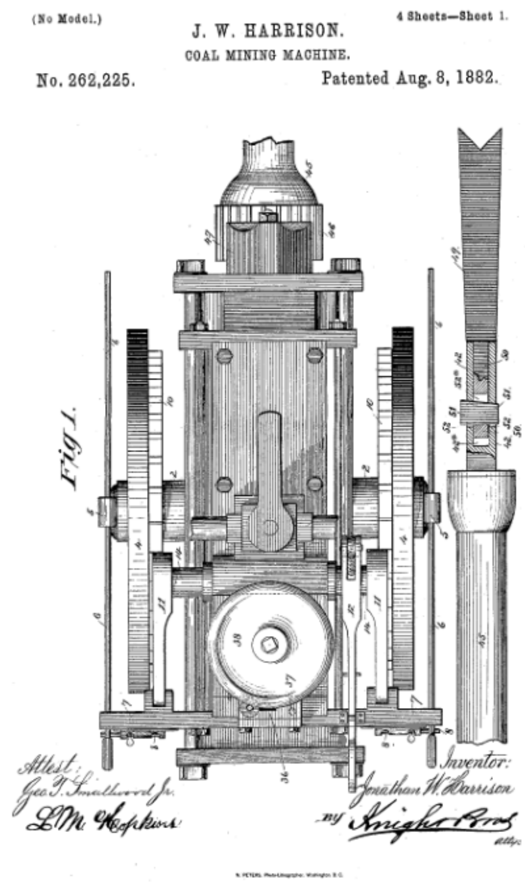
Notes: The dots indicate mining towns, each of which can contain multiple mines. Towns with squares contain at least one machine mine.

Figure A4: Capacity Utilization



Notes: This graph plots the distribution of capacity utilization, defined as annual mine output over annual mine capacity, across mines in Illinois in 1898. I distinguish hand mines, which did not use cutting machines, from machine mines, which did.

Figure A5: Patent for the Harrison Coal Mining Machine



Notes: U.S. patent for the 1882 Harrison Coal Mining Machine (Whitcomb, 1882). This was the most frequently used coal-cutting machine in the dataset.

Table A1: Occupations and Wages

	Daily wage (USD)	Employment share (%)
Miner	2.267	61.5
Laborers	1.76	14.30
Drivers	1.83	5.91
Loaders	1.74	3.63
Trappers	0.80	1.86
Timbermen	2.02	1.68
Roadmen	2.36	1.46
Helpers	1.70	0.92
Brusher	2.06	0.75
Cagers	1.87	0.70
Engineer	2.11	0.61
Firemen	1.60	0.57
Entrymen	2.01	0.56
Pit boss	2.70	0.56
Carpenter	2.09	0.53
Blacksmith	2.08	0.46
Trimmers	1.50	0.36
Dumper	1.68	0.36
Mule tender	1.65	0.31
Weighmen	1.95	0.29

Notes: Occupation-level data for the top-20 occupations by employment share in the 1890 sample of 11 mines in Illinois. The 20 occupations with the highest employment shares together cover 97% of coal mining workers in the sample.

Table A2: Cost Dynamics

	Log(Output/(Labor-Days))			
	Est.	S.E.	Est.	S.E.
Log(Cumulative Output)	0.126	0.004	-0.010	0.017
Mine FE	No		Yes	
R-squared	.336		.818	
Observations	3614		3614	

Notes: Regression of log output per worker-day against log cumulative output (lagged by one time period) at the mine-year level. Sample includes only mines for which lagged output is observed.

Table A3: Union Power and Strikes

	Log(Union Bargaining Power)			
	Est.	S.E.	Est.	S.E.
1(strike)*1(year $\geq$ 1898)	0.095	0.082	0.181	0.069
Strike indicator	Any		Successful	
R-squared	.716		.670	
Observations	3347		3780	

Notes: This table re-estimates the difference-in-differences model for the 1897–1898 strikes, but using the log of the labor union's bargaining power, $\ln(\gamma_{ft})$, as the left-hand-side variable instead of log output. The left column compares all mines that went on strike, the right column only the mines at which strikes resulted in wage increases.

Table A4: Production Function: Extensions

	Nonconstant RTS		Adding Materials		Capital and RTS	
	Est.	S.E.	Est.	S.E.	Est.	S.E.
Returns to scale	1.032	0.041	1.051	0.090	0.980	0.120
Labor coefficient	0.010	0.011	0.014	0.029	0.003	1.681
Capital coefficient	0.050	0.162	-0.056	0.190	0.863	1.464
Serial corr. TFP	0.337	0.114	0.516	0.161	0.370	0.265
Materials coefficient		.	0.000	0.048	0.000	0.027
Returns to scale * K		.		.	-0.012	0.024
Observations	664		296		296	

Notes: This table reports the estimates for the various extensions of the production function. Standard errors are block-bootstrapped with 200 iterations.

Table A5: Wage Variation

	R^2	R^2	R^2	R^2
Log(Daily Skilled Miner Wage)	0.099	0.186	0.285	0.734
Year FE	X	X	X	X
County FE		X	X	X
Town FE			X	X
Firm FE				X

Notes: The four columns report the R^2 of regressing log wages on, alternatively, year, county, town, and firm fixed effects.

Table A6: All Variables per Year

Year	1884	’86	’88	’90	’92	’94	’96	’98	1900	’02
Output Quantities										
Total	X	X	X	X	X	X	X	X	X	X
Lump					X	X	X	X	X	X
Mine run									X	X
Egg									X	X
Pea									X	X
Slack									X	X
Shipping or local mine					X	X	X			
Shipping quantities										X
Input Quantities										
Miners, winter	X	X	X	X						
Miners, summer	X	X	X	X						
Miners, avg entire year					X	X		X	X	X
Miners, max entire year					X	X				
Other employees	X	X	X	X	X	X		X	X	X
Other employees, underground								X		
Other employees, above ground								X		
Other employees winter							X			
Other employees summer							X			
Boys employed underground			X	X	X	X	X			
Mules		X								
Days worked	X	X	X	X	X	X		X	X	X
Kegs powder	X	X	X	X	X	X		X		X
Men killed	X	X	X	X	X	X		X		X
Men injured	X	X	X	X	X	X		X		X
Capital (in dollars)	X									

Table A7: All Variables per Year (cont.)

Year	1884	’86	’88	’90	’92	’94	’96	’98	1900	’02
Output Price										
Price/ton at mine	X	X	X	X	X			X	X	X
Price/ton at mine, lump						X	X	X		
Input Prices										
Miner piece rate (summer)	X	X	X	X	X	X				
Miner piece rate (winter)	X	X	X	X	X	X				
Miner piece rate (hand)								X	X	X
Miner piece rate (machines)								X	X	
Piece rate dummy					X					
Payment frequency						X	X	X	X	
Net/gross wage							X			
Oil price							X			
Mine Characteristics										
Type (drift, shaft, slope)	X	X			X	X	X	X		
Hauling technology	X	X			X	X		X		
Depth	X	X			X	X	X	X	X	
Thickness	X	X			X	X	X	X	X	
Geological vein type	X	X			X	X		X		
Longwall or PR method	X	X			X	X	X		X	
Number of egress places	X	X								
Ventilation type	X	X								
New/old mine					X	X				
# Acres					X	X	X			
Mine capacity								X		
Mined or blasted								X		
Cutting Machines										
Cutting machine dummy					X	X	X	X		
# Cutting machines	X	X	X	X						
# Tons cut by machines									X	X
# Cutting machines, by type			X							