

# Labor Market Power and Factor-Biased Technology Adoption

Michael Rubens\*

March 4, 2024

## Abstract

Although monopsony power can induce deadweight loss, it can also incentivize technology adoption by employers or buyers through reduced hold-up risk. Using an empirical model of labor supply and demand that features both these opposing forces, I quantify the net welfare effects of employer power in the late 19th century Illinois coal mining industry. I find that if employer power, which increased during the 1890s, would have remained constant, the usage of mechanical coal cutters, a new technology, would have decreased by 23%, thereby increasing marginal costs. However, output would still have increased by 9.5%, meaning that the deadweight loss channel dominates the hold-up channel. Ignoring endogenous investment leads to overestimating the consumer welfare gains and the producer welfare losses of unionization.

**Keywords:** Monopsony, Buyer Power, Investment, Technological change, Productivity

**JEL codes:** L11, L13, J42, N51

---

\*UCLA, Department of Economics. Contact: [rubens@econ.ucla.edu](mailto: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, and participants at various seminars and conferences. Financial support from the Belgian American Educational Foundation, the Fulbright program, the Research Foundation Flanders (FWO grant 11A4618N), and the European Research Council (ERC grant 816638) is gratefully acknowledged.

# 1 Introduction

There is increasing empirical evidence for the existence of ‘monopsony’ or ‘buyer power’ across various industries, countries, and types of factor markets.<sup>1</sup> Prior 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 investment to be exogenous. However, investment of employers, or of buyers more generally, is endogenous to the extent of monopsony power through reduced hold-up. The more monopsony power employers have, the larger the share of the increased surplus from technology adoption they can appropriate, which increases the incentives to adopt new technologies.

In this paper, I examine the welfare effects of monopsony power by disentangling two key mechanisms. On one hand, under linear wage contracts, monopsony power leads to deadweight loss because it makes employers cut back on inputs 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 rents created by this technology adoption. Although sellers would want to incentivize buyers to adopt new technologies nevertheless by writing contracts conditional on technology usage, the inability of writing such contracts leads to a hold-up problem. Whereas monopsony power lowers equilibrium output by inducing deadweight loss, it can increase equilibrium output through the second mechanism, increased technology adoption, as technology adoption lowers marginal costs.

I start the paper with a theoretical model of labor demand and supply that allows for both monopsony-induced deadweight loss and hold-up. Although the model is written in terms of employers and employees, it applies to vertical relationships between buyers and sellers more generally. The model features four three elements. First, firms face

---

<sup>1</sup>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).

upward-sloping labor supply curves. They negotiate over wages with workers using a linear wage contract according to their relative bargaining power. In the extreme case where all bargaining power goes to the employer, the model collapses to the classical monopsony model. Second, firms combine different labor types to produce output using a constant elasticity of substitution (C.E.S.) production function, and choose between a menu of technologies. These technology decisions determine the employers' Hicks-neutral and factor-specific productivity levels. Third, employers sell their output on a product markets, thereby facing a downward-sloping goods demand curve. Simulating this model, I find that the relationship between equilibrium output and employee bargaining power is monotonically increasing when assuming exogenous technology usage, as only the deadweight loss mechanism is at play. However, when allowing for endogenous technology adoption, the output-employee power relationship becomes an inverted U-shape, as the deadweight loss and hold-up 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 with endogenous technology adoption, I carry out an empirical application to quantify the relative importance of the deadweight loss and hold-up mechanisms. I study how the mechanization of the Illinois coal mining industry between 1884 and 1902 was affected by market power held by firms over their miners. There are three reasons why this provides a unique and interesting setting to study the relationship between buyer power and innovation. First, the introduction of coal cutting machines in the U.S. in 1882 provides a large observed technological shock. A novel and rich historical data set tracks the usage of these cutting machines and other technologies over time, together with input and output quantities, wages and coal prices, all at the mine level. 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. Third, throughout the sample period, unionized coal miners went on strike on different occasions,

which provides useful shocks to the relative bargaining power of employers and employees.

I implement an empirical version of the model to fit the coal industry setting, and estimate it using the rich mine-level data on production, coal prices, input quantities and prices, and technology usage. I rely on observed geological 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. The identification of 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, similarly to many other technologies that were developed throughout the 19th century (Mokyr, 1990; C. D. Goldin & Katz, 2009). Using the estimated production, labor supply, and coal demand models, I estimate the relative bargaining weights of the employers and labor union. I find that employer power over miners increased substantially throughout the 1890s: in 1890, employers extracted one third of the total rents created in the industry, whereas they extracted almost half of the rents by 1902. 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 equilibrium output, input quantities, coal prices, wages, and machine usage in the observed equilibrium, and in two counterfactual equilibria. In a first counterfactual, I examine how all equilibrium outcomes and welfare would have changed if employer power would have remained constant at its 1890 average, rather than increasing throughout the 1890s. Second, I examine the same counterfactual exercise, but holding cutting machine usage fixed. This second counterfactual is informative of the relative importance of the endogenous machine usage mechanism compared to the deadweight loss effect. I find that in the absence of the employer power increase, cutting machine usage would have been on average 23% lower by the end of the panel in 1902. Hence, both total factor productivity and skill-augmenting productivity would have been lower without the increase in employer power. However, equilibrium output would have increased by 9.5% in the counterfactual, because the markdown elimination

effect was larger than the technology adoption effect. This reduction in capital investment has important implications for the welfare effects of increased union power. If capital investment would be exogenous, the decrease in employer power would have resulted in an increase of consumer surplus of 10.6%. However, when taking into account reduced capital investment, this welfare gain drops to 8.8%. On the producer side, the exogenous capital model predicts a loss in producer surplus of 36.4%, which reduces to merely 14.3% when accounting for changing technology usage. When considering total welfare of workers, firms, and consumers, the exogenous investment model finds that unionization increased total welfare by only 1.9%, whereas the endogenous investment model finds a considerably higher welfare gain of 5.3%.

I examine two additional counterfactual exercises. First, I re-estimate the same counterfactual, but calibrating fixed machine costs to be 10% of their observed level. With lower fixed costs, the hold-up mechanism dominates the deadweight loss mechanism: a decrease in employer power now leads to lower equilibrium output and a decrease in total welfare of 1.5%. Second, I re-estimate the counterfactual but for a skill-biased technology, rather than an unskill-biased technology. I find that the usage rate of an skill-biased technology would be considerably lower, at 2.7% in the observed equilibrium compared to 8.6% for the unskill-biased technology. Second, the skilled labor welfare comparison between the endogenous and exogenous investment case flips. For the unskill-biased technology, additional investment leads to lower skilled labor surplus, whereas for a skill-biased technology, the opposite holds. Hence, a reduction in investment due to decreased employer power over skilled workers increases skilled worker welfare in the endogenous investment case compared to the exogenous investment model, whereas the opposite holds for a counterfactual skill-biased technology.

This paper relates 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, 2019; Morlacco, 2017; Rubens, 2023), but keep technology choices fixed. In contrast, I show that endogenous technology choices present an additional channel through which input market power shapes aggregate outcomes. On the other hand, there is work on hold-up 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 to these literatures by empirically investigating and comparing monopsony-induced deadweight loss to the inefficiency created by hold-up, in order to obtain a net effect of employer power on equilibrium output and welfare. In contrast to Shi (2023), the main research question in this paper is about employer power in general, rather than the effects of noncompete agreements in particular.

Second, this paper builds on the vertical relations literature. In contrast to existing work on hold-up (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, 2021), I focus on its effects on the technology choices of the buyers.

Finally, 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 (2022), and on technological change during this same time period (Atack, Bateman, & Margo, 2008; C. Goldin & Katz, 1998; Katz & Margo, 2014). I bring these two literatures together by studying how increasing employer power during the late 19th century affected the adoption of new skill-biased technologies, and by quantifying the resulting welfare effects on both consumers producers, and workers.

The remainder of this paper is structured as follows. Section 2 contains the theoretical model. Section 3 discusses the data, industry background, and the empirical model, which is estimated in Section 4. Section 5 contains the counterfactual exercises.

## 2 Theory

### 2.1 Primitives

Firms  $f$  produce output  $Q_f$  using two variable inputs  $H_f$  and  $L_f$ . I denote these as high- and low-skilled labor and let firms be employers, but the model can be interpreted as wholesalers  $f$  buying inputs from manufacturers as well. I rely on a C.E.S. production function with two inputs that are substitutable at a constant elasticity  $\sigma$ . For simplicity, I assume constant returns to scale, but relax this in Appendix C.1. Skill-augmenting productivity is denoted  $A_f$ , the low-skilled labor coefficient is  $\beta^l$ , and Hicks-neutral productivity is  $\Omega_f$ .

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

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 can change 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  $P_0$ , demand shocks  $\xi_f$ , and a constant demand elasticity  $\eta$ .

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

Let the outside option of high-skilled labor be an upward-sloping function with constant inverse elasticity  $\psi$ . Firms are differentiated by an amenity term  $\zeta_f$ , the average industry wage is equal to  $W_0$ . In contrast to models with a constant outside option value, such as 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.<sup>2</sup> In contrast, the outside option of low-skilled labor is assumed to be equal to a constant

---

<sup>2</sup>There are various reasons for an increasing outside options curve on aggregate, one of which is that workers

$V$ . Hence, firms pay low-skilled wages  $V$ , and the low-skilled labor market is perfectly competitive.

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

High-skilled workers are unionized at the firm level. The utility of the labor union is at firm  $f$  is denoted  $\Pi_f^u$ , and is defined as the difference between high-skilled earnings and the outside option to high-skilled workers.

$$\Pi_f^u = (W_f - Z_f)Q_f$$

Employer profits are denoted as  $\Pi_f^d$ :

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

## 2.2 Strongly efficient bargaining

### Behavior and equilibrium

For the purpose of illustration, I start with a Nash bargaining model in which a labor union and employers collectively bargain over both employment and wages, before moving to model where they only bargain over wages. Employers and unions bargain over both employment and wages in a Nash bargaining protocol, with  $\gamma$  indicating union bargaining power.

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

---

are heterogeneous in terms of their outside options, and that firms cannot observe these worker-specific outside options, and hence not wage-discriminate in function of these individual outside options.



The model implies that the union and employers jointly optimize joint profits, and split the surplus according to the bargaining parameters  $\gamma$ . This model is equivalent to the ‘strongly efficient’ bargaining model in Abowd and Lemieux (1993), but with the important distinction that the outside option of the workers is increasing, which implies an upward-sloping labor supply curve. The analog of this feature in vertical industry models would be that sellers face increasing marginal costs. The usefulness of the strongly efficient bargaining model is that it does not feature any monopsony distortions, only endogenous technology choices. This makes it a useful benchmark against which the full model can be compared.

Taking the first order condition for the high-skilled wage results in:

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

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

$$P_0 \left( \frac{1 + \eta}{\eta} \right) Q_f^{\frac{1}{\eta}} \left( \frac{Q_f}{H_f} \right)^{\frac{1}{\sigma}} (\Omega_f A_f)^{\frac{\sigma-1}{\sigma}} = W_0 H_f^\psi \quad (5)$$

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

Equilibrium  $(P_f^*, Q_f^*, H_f^*, L_f^*)$  is the solution 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.

### Effects of employer power: hold-up

Holding the capital  $K_f$  fixed, employer power  $(1 - \gamma_f)$  does not affect equilibrium output, consumer prices, or input quantities, only the wage  $W_f$ . However, employer power affects investment. Suppose firms need to pay per-unit capital costs  $\phi$ . We assume that capital increases employer variable profits, which implies 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 receiving a subsidy. Propo-

sition 1 says that under efficient bargaining, employer power increases technology adoption of firms.

**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  $\Pi^d$  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:

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

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

The intuition behind Proposition 1 is that the higher buyer power is, the larger the share of the rents created by capital investment get appropriated by the buyer. Hence, this increases the incentive for the buyer to invest. This is not a new finding, but a reformulation of the well-known hold-up 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, unions and employers usually do not contract on both wages and employment, but only bargain over wages. Assuming they bargain over a linear wage contract, two types of bargaining models are possible. First, it could be that employers choose employment and bargain with workers who supply labor perfectly elastically, as in Abowd and Lemieux (1993). Such a model does not allow for monopsony distortions, as the labor supply curve is not upward-sloping. 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. First, this model has the benefit of collapsing to the classical monopsony model when assuming perfect buyer power ( $\gamma = 0$ ), with the ensuing monopsony distortion. Second, in the empirical background, it will become clear that double marginalization in the empirical application took the form of a markdown-markup marginalization due to monopsony power of the employers, rather than from a sequential markup set by both unions and employers, due to market power of the union. In other words, deadweight loss was caused by employers who exercised market power over workers, resulting in excess labor supply, rather than by unions who exercised market power over firms, which would result in labor shortages in equilibrium.

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

$$W_f = W_0 \left( \frac{H_f}{\zeta_f} \right)^\psi \quad (7)$$

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

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

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

$$\begin{aligned} \gamma_f \left( \left(1 - \frac{\partial Z_f}{\partial W_f}\right) H_f + \frac{\partial H_f}{\partial W_f} (W_f - Z_f) \right) (P_f Q_f - W_f H_f - V L_f) \\ + (1 - \gamma_f) (W_f - Z_f) H_f \left( P_f \frac{\partial Q_f}{\partial H_f} \frac{\partial H_f}{\partial W_f} - H_f - W_f \frac{\partial H_f}{\partial W_f} \right) = 0 \end{aligned} \quad (8)$$

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

$$\max_{L_f} (\Pi_f^d) \quad (9)$$

The low-skilled labor demand function is the same as in the strongly efficient bargaining case, equation (6). Equilibrium  $(Q_f^*, P_f^*, W_f^*, H_f^*, L_f^*)$  is given by equations (1), (2), (7), (6), (8), which are the production, goods demand, high-skilled labor supply, low-skilled labor demand, and wage bargaining equations.

### Effects of employer power: hold-up 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. The reason for this is that 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 . The effect of capital investment on buyer

profits is still 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  $\gamma$  gives:

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

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. The sign of the second term is ambiguous in sign, as we show below using simulated data.

**Conjecture 1** *Under weakly efficient bargaining, buyer power can either increase or decrease equilibrium output.*

As a result, the net effect of buyer power on equilibrium output is ambiguous, as stated in conjecture 1. On one hand, increased buyer power decreases output through the monopsony distortion. On the other hand, buyer power increases 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 hold-up mechanism. Hence, it is a function of, among others, the labor supply elasticity, which determines the size of the deadweight loss, the productivity effects of the technology, which determine the marginal cost effects of the technology. In the empirical application, I will quantify the relative size of these effects in order to examine how counterfactual changes in employer power affect equilibrium output, producer surplus, consumer surplus, and worker welfare.

## Simulation

To illustrate the importance of exogenous technology usage for the welfare effects of labor market power, I simulate a parametrized version of the model. The parametrization

is specified in Appendix B.1. I consider a new technology that increases H-augmenting productivity  $A$  by 5%, and increases TFP  $\Omega$  by 20%.

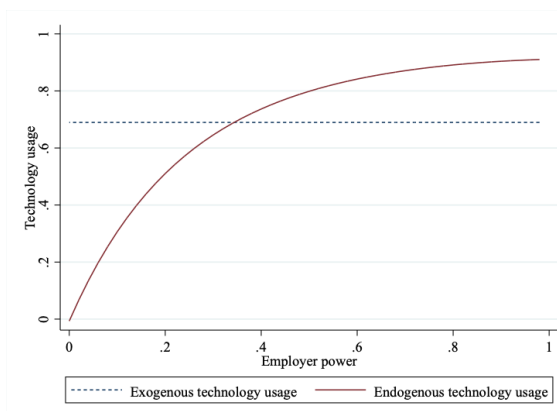
Figure 2a plots equilibrium technology usage  $K$  against employee bargaining power  $\gamma_f$ . The red solid line depicts the model in which technology usage is allowed to change in function of employee bargaining power. As was explained in the theoretical model, technology usage decreases with the level of employee bargaining power. As a comparison, the blue dashed line depicts the model in which technology usage is exogenous to the degree of bargaining 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  $(1 - \gamma)$ . Under the assumption of exogenous technology usage, the blue solid line, output monotonically decreases with employer power. As was explained earlier, 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 0.35. Hence, equilibrium output is maximized at this level of employer power.

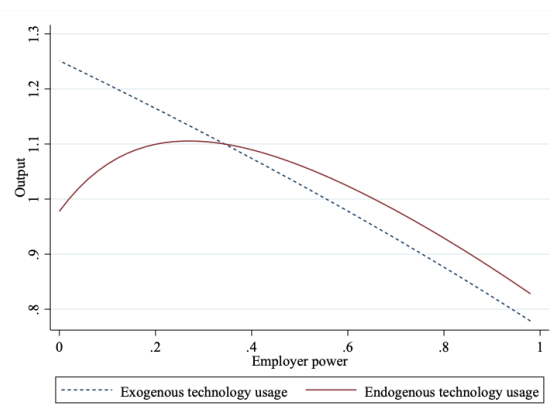
The output-maximizing degree of employer power depends on the labor supply elasticity and on the productivity effects of the technology. If labor supply becomes more inelastic, deadweight loss becomes larger, so the relative importance of the technology usage mechanism compared to deadweight loss falls. As shown in Appendix B.1, the output-maximizing employer weight falls from 0.35 to 0.18 if the labor supply elasticity is  $\psi = 0.5$  rather than  $\psi = 1.5$ . Second, the larger the productivity effects of the technology, the higher employer power should be, because higher productivity effects increase the relative importance of the technology usage mechanism compared to deadweight loss.

**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.

The simulations in Appendix B.1 show that a Hicks-neutral productivity effect of 10%, compared to 20% in the baseline, decrease the output-maximizing bargaining weight from 0.35 to 0.10. Moreover, the output gain between a scenario of full absence of employer power and the output-maximizing employer weight of 0.10 is almost zero. Equilibrium output becomes almost a monotonically decreasing function of employer power, meaning that deadweight loss dominates the technology usage mechanism at any level of employer power.

### 3 Empirical application

The simulated theoretical model in the previous section shows that the relationship between employer power and output is ambiguous, depending 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 used is derived from the *Biennial Report of the Inspector of Mines of Illinois*, which was digitized in the context of this project. I observe every bituminous coal mine in Illinois between 1884 and 1902 at two-year intervals, which results in 8356 observations. The dataset records the name of the mine, the mine owner, yearly coal extraction, average employee counts for both skilled and unskilled workers, days worked, and a dummy for cutting machine usage in every two-year period. 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 of these variables are used in the analysis, given that some are observed in a small subset of years.

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 based on their output of large coal pieces, rather than on their total output. This introduces some measurement error in labor costs. However, the data set reports the usage of wage screens in 1898, and shows that they were used in mines representing merely 2% of total employment. Skilled wages and employment are separately reported for the summer and winter months between 1884 and 1894. For some years I observe additional variables such as mine capacities, the value of the total capital stock and a break-up 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 different other datasets. First, the inspection report from 1890 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 from the



1880 and 1900 population census and the censuses of agriculture and manufacturing are collected as well. Third I collect information on coal cutting machine costs from Brown (1889). I refer to Appendix A for more details regarding the data sources and cleaning procedures.

## **3.2 Industry background**

The setting of the empirical application is the Illinois coal mining industry between 1884 and 1902. Throughout this time period, this industry grew rapidly: annual output tripled from 10 to 30 megatons between 1884 and 1902. This was both due to an increase in the average mine size and to an increase in the number of mines from 680 to 898.

### **Extraction process**

The coal extraction process consisted of three consecutive steps. First, the coal vein had to be accessed, as it lied below the surface for 98.0% of the mines and 99.4% of output. Second, upon reaching the vein, the coal wall was ‘undercut’, traditionally by hand, but from 1882 onward also with coal cutting machines. The mechanization of the cutting process is considered to be the most significant technological change during this time period (Fishback, 1992). Third, coal had to be transported back to the surface and sorted 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, was purchased and brought by the miners. Second, coal itself was used to power steam engines, electricity generators, and air compressors.

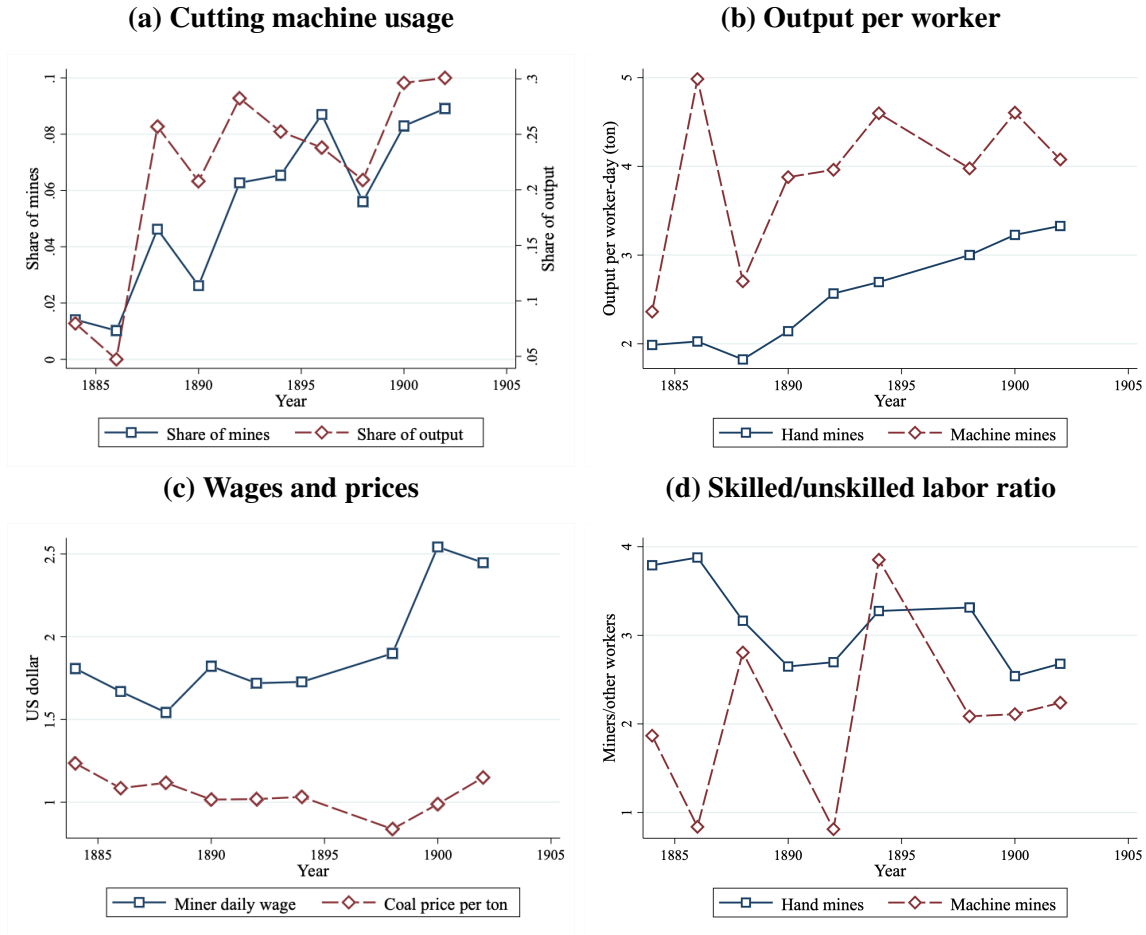
Figure 2(b) plots the ratio of total output over total days worked at mines that used cutting machines (‘machine mines’) and mines that did not (‘hand mines’). Daily output per worker increased from 2 to 3.3 tons for hand mines, and from 2.3 to 4.1 tons for machine mines.<sup>3</sup>

Although different coal types exist, the mines in the data set all extract bituminous coal. There might be minor quality differences even within this coal type due to variation in sul-

---

<sup>3</sup>This series is adjusted for the reduction of hours per working day in 1898, as explained in Appendix A.

**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 is cut manually, and at mines where cutting machines are 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.

fur content, ash yield, and calorific value (Affolter & Hatch, 2002). Most of this variation is, however, dependent on the mine's geographical location.

## Occupations

Coal mining involved a variety of occupational tasks. The inspector report from 1890 reports wages at the occupation-level, and this subdivision is reported in Appendix Table A1 for the 20 occupations with the highest employment shares, together covering 97% of

employment. Three out of five workers were miners, who did the actual coal cutting. This required a significant amount of skill: in order 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 a variety of 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 were moreover not specific to coal mining: tending mules or loading carts are general-purpose tasks, in contrast to undercutting coal walls.

The difference in industry-specific skills are 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 nearly all other occupations worked 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 will denote as 'skilled labor', and all other employees, which are called 'unskilled labor'. This follows the categorization of labor provided in the data set.

### **Technological change**

The first mechanical coal cutter in the U.S.A. was invented by J.W. Harrison in 1877, but it was merely a prototype.<sup>4</sup> The Harrison patent was acquired and adapted by Chicago industrialist George Whitcomb, whose 'Improved Harrison Cutting Machine' was released in 1882.<sup>5</sup> As shown in Figure 2a, the share of Illinois coal mines using a cutting machine increased from below 2% to 9% between 1884 and 1902. Mechanized mines were larger: their share of output increased from 7 to 30% over the same time period. The mechanization

---

<sup>4</sup>Simultaneously, prototypes of mechanical coal cutting machines were invented in Northern England in the late 1870s (Reid, 1876; Ackermann, 1902).

<sup>5</sup>A picture of the patent is in Appendix Figure A5. The spatial diffusion of cutting machines is shown in Appendix Figure A3.

of the hauling process, which replaced mules with underground locomotives, was another source of technical change, and started during the 1870s. By the start of the panel in 1884, mining locomotives were already widely used in Illinois: the share of output mined in locomotive mines was around 90%.

As was 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 worker-days per year.<sup>6</sup> Mines without cutting machines used between 3 and 4 skilled labor-days per unskilled labor-days throughout the sample period, compared to 2 to 3 skilled labor-day per unskilled worker-day for machine mines. In every year, except 1894, machine mines used less skilled per unskilled worker. The skilled-unskilled labor ratio was on average 16.5% lower for machine mines compared to hand mines, and this difference is statistically significant. 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 total factor and factor-augmenting productivity levels, I refer to the empirical model in the next section. 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).

In contrast, underground mining locomotives had very different effects: rather than saving on the skilled miners, locomotives replaced mules and some of the unskilled workers involved in the hauling process, such as mule tenders.

---

<sup>6</sup>1890 is omitted for machine mines in 1890 due to employment being unobserved for most machine mines in that year.

## Labor markets

Skilled workers received a piece rate per ton of coal mined, whereas unskilled workers were paid a daily wage.<sup>7</sup> Converting the piece rates to daily wages, the net salary of skilled labor was on average 22% 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 population sizes of the towns in the dataset were 845 and 1706 inhabitants. In the average town, 16% of the population was employed in a coal mine. Considering that women and children under the age of 12 did not work in the mines, this implies that a large share of the local working population was employed in coal mining. Of all the villages, 42% had just one coal firm, and 75% had three or less coal firms. Two-thirds of all employees worked in a village with three or less coal mines. Although most of the villages in the data set were connected by railroad, these were exclusively used for freight: passenger lines only operated between major cities (Fishback, 1992). Given that the average village was 7.4 miles apart from the next closest village, and that skilled workers had to bring their own supplies to the mine, commuting between villages was not an option, and the mining towns can be considered as isolated local labor markets. Most roads were unpaved and automobiles were not yet introduced. In order to switch employers, miners had to migrate to another town.

First attempts to unionize the Illinois coal miners started around 1860, without 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 constrained by the use of ‘yellow-dog’ labor contracts that forced employees not to join a trade union.<sup>8</sup> 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

---

<sup>7</sup>Piece rates were an incentive scheme in a setting with moral hazard, as permanent miner supervision would be very costly.

<sup>8</sup>These contracts were criminalized in Illinois in 1893, with fines of \$100 USD, which was equivalent to on average six months of a miner’s wage. (Fishback, Holmes, & Allen, 2009).

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 data set did not pay for company housing of the miners (Lord, 1883, 75), which would otherwise be a labor cost in addition to miner wages.

Figure 2(c) reports the aggregate skilled labor daily wage, defined as the total wage bill spend 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, as daily miner wages remained around \$1.8. After the subsequent introduction of wage bargaining, wages rose.

### **Coal markets**

Coal was sold at the mine gate, and there was no vertical integration with post-sales 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, to a lesser extent, Chicago.<sup>9</sup> Railway firms acted as an intermediary between coal firms and consumers, and were also major coal consumers themselves. Historical evidence points to intense competition on coal markets during the last two decades of the 19th century, before the large consolidation wave in the early 1900s (Graebner, 1974). Nevertheless, there was still a considerable transportation cost of coal, which makes that coal markets were likely not entirely integrated. There are large differences in the coal price across Illinois: in 1886, for instance, it varied between 80 cents/short ton at the 10th percentile of the price distribution to 2 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 \$1.2 to \$0.9 between 1884-1898, after which it increased again.

---

<sup>9</sup>Chicago mainly sourced its coal from fields in Ohio, Pennsylvania, and West Virginia using lake steamers (Graebner, 1974).

### 3.3 Stylized facts

In this section, I present two sets of stylized facts to motivate the assumptions made in the structure model. First, I show that skilled wages co-moved 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 very seasonal: during winters, there was more demand for energy than during summers. As was explained earlier, 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 both skilled and unskilled labor demand, only skilled wages increased during winter. This is also shown in Figure 3(c), that plots how average daily wages for both skilled and unskilled workers in 1890 change with the monthly number of worker-days of each type at the mine-month level throughout 1890.<sup>10</sup> Skilled wages were positively correlated with monthly skilled employment, whereas the unskilled worker wage-employment schedule is flat. Moreover, there was a lot of variation in skilled wages across mines and months, but very little cross-sectional and intertemporal variation in unskilled wages.

The fact that unskilled wages were dispersed very little, and did not react to coal demand shocks, whereas skilled wages were very dispersed and reactive to demand shocks, sug-

---

<sup>10</sup>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 is available for a sample of mines selected by the Illinois Bureau of Labor Statistics across 5 counties in 1890, which covers 16% of skilled employment and 9% of unskilled employment.

gests that skilled labor was supplied inelastically and unskilled labor elastically. However, it could be that labor supply also changed seasonally, 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 on the longer term. Hence, in the structural model, I will rely on a different instrumental variables strategy, which relies on international coal price shocks, to identify the labor supply elasticity.

**Figure 3: Wage-employment profile by skill type**



**Notes:** Panel (a) shows how the wages of skilled miners and other mine employees evolved monthly during the year 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'. At 28% of coal mines, miners went on strike, and this resulted in a wage increase at 90% 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 miner's union. Using a difference-in-differences model, I compare the evolution of log coal output between mines at which miners went on strike,  $I(strike)_{ft} = 1$  to mines where miners did not strike, before and after 1898. I define the strike indicator as mines where miners striked 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, in order to not take into



account the reduction in output during the strike.

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

I also estimate the same difference-in-difference 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 27% after 1898 relatively to the mines that did not go on strike. In contrast, machine usage is estimate to decrease relatively at the striking mines, although this effect is measured with much error.

In Table 1b, I compare pre-trends by estimating the interaction effect between the strike indicators and a linear time trend prior to 1898. I find that output was already growing faster at striking mines, although this effect is not significantly positive when using the successful strike indicator. Nevertheless, it is important to keep in mind that the miner's decision to go on strike could have been endogenous to underlying changes in productivity, mechanization, and other determinants of output. This is one of the reasons why a structural model will be useful to examine how changes in relative bargaining power of employers and unions affected equilibrium output.

### 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

**Table 1: Strikes, output, and investment**

<i>(a) Diff-in-diff results</i>	log(Output)		Machine usage	
	Est.	S.E.	Est.	S.E.
1(Strike)*1(year $\geq$ 1898)	0.237	0.099	-0.193	0.472
Model:	Linear		Logit	
Firm fixed effects:	Yes		Yes	
R-squared	.941		.	
Observations	7415		693	
Firm fixed effects:	No		No	
<i>(b) Pre-trends</i>	log(Output)		Machine usage	
	Est.	S.E.	Est.	S.E.
1(Strike)*yr	0.018	0.018	-0.144	0.036
Model:	Linear		Logit	
R-squared	.194		.	
Observations	5637		5703	

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

consistent with the model, given that it does not feature strategic interaction between firms on the labor market. Annual coal extraction is  $Q_{ft}$  tons, the number of days worked by high-skilled labor is denoted  $H_{ft}$  whereas the number of low-skilled labor-days is  $L_{ft}$ . 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  $K_{ft} \in \{0, 1\}$ . 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 C.E.S. production function from Equation (1), with an elasticity of input substitution  $\sigma$  and low-skilled labor coefficient  $\beta^l$ . Firms differ in terms of their skill-augmenting productivity  $A_{ft}$  and in their Hicks-neutral productivity  $\Omega_{ft}$ . In Appendix C.1, I estimate various extensions of the production model to allow for non-constant 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 10, which is the production function I will estimate.

$$q_{ft} = \left( \frac{\sigma - 1}{\sigma} \right) \log \left( (H_{ft} A_{ft}(K_{ft}))^{\frac{\sigma-1}{\sigma}} + \beta^l L_{ft} \right)^{\frac{\sigma}{\sigma-1}} + \omega_{ft}(K_{ft}) \quad (10)$$

Cutting machine usage  $K_{ft}$  is allowed to affect both productivity terms  $\Omega_{ft}$  and  $A_{ft}$ . The logarithms of both these productivity terms,  $a_{ft}$  and  $\omega_{ft}$ , are assumed to evolve as AR(1) processes, as specified in Equations (11) and (12). The productivity terms have serial correlations  $\rho^a$  and  $\rho^\omega$  and are assumed to be affected linearly by cutting machine usage, as parametrized by the coefficients  $\alpha^k$  and  $\beta^k$  for labor-augmenting and Hicks-neutral productivity, respectively.<sup>11</sup> Skill-augmenting and Hicks-neutral productivity shocks are denoted  $e_{ft}^a$  and  $e_{ft}^\omega$ .

$$a_{ft} = \alpha^k K_{ft} + \rho^a a_{ft-1} + e_{ft}^a \quad (11)$$

$$\omega_{ft} = \beta^k K_{ft} + \rho^\omega \omega_{ft-1} + e_{ft}^\omega \quad (12)$$

I assume 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 1.4% of the mines operated at full capacity, and they were responsible for 1.1% of industry sales.<sup>12</sup> I also abstract from stockpiling of coal, and assume that coal must be sold immediately after extraction: coal storage usually led to deteriorating coal quality, moreover it was expensive and dangerous (Stoek, Hippard, & Langtry, 1920). As Williams (1901) asserts:

---

<sup>11</sup>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. I refer to Appendix C.2 for a motivation and discussion of this assumption.

<sup>12</sup>The entire distribution of capacity utilization rates is shown in Figure A4.

“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’ (Williams, 1901)

### Labor supply

Adding time subscripts to the inverse labor supply function (7) and inverting it results in the labor supply equation (13a). The daily wage of high-skilled workers  $W_{ft}$  is computed as the piece rate multiplied by daily tonnage per worker. I diverge from equation (7) in two ways. First, I let the average wage  $W^0$  be labor market-specific, in order to take into account differences in local labor market conditions across Illinois. I define labor markets  $i$  by grouping villages into isolated labor markets, as detailed below in Section 4 and hence let  $W_{0i(f)t}$  be the average daily wage in the labor market  $i$  of firm  $f$  in year  $t$ .

Second, I include observed firm characteristics  $\mathbf{X}_{ft}$  next to the latent amenity term  $\zeta_{ft}$ . Specifically, I include a linear time trend 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 population centers in the area.

$$H_{ft} = \left( \frac{W_{ft}}{W_{0i(f)t}} \right)^{\frac{1}{\psi}} \exp(\mathbf{X}_{ft})^{\psi^x} \zeta_{ft} \quad (13a)$$

I estimate the labor supply equation in logs, which is given by equation (13b).

$$h_{ft} = \frac{1}{\psi} (\ln w_{ft} - \ln w_{0i(f)t}) + \psi^x \mathbf{X}_{ft} + \ln(\zeta_{ft}) \quad (13b)$$

The amenity term  $\zeta_{ft}$  captures firm differentiation from the miner’s perspective. In contrast to Delabastita and Rubens (2022), which relies on a homogeneous employers model, I do allow for firm differentiation because there is substantial variation in skilled wages across mines, even within the same labor markets.<sup>13</sup>

Similarly to the theoretical model, the market for low-skilled labor is perfectly competitive, and low-skilled workers are paid their constant outside option wage  $V$ . The main

<sup>13</sup>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.

reason for this assumption lies in the fact that unskilled wages were the same everywhere in Illinois, as shown in Figure 3, and did not respond to labor demand shocks.

### **Coal demand**

As was explained in the background section, the coal produced in Illinois mines was a nearly homogeneous product. However, coal firms are differentiated due to their locations, which results in price differences between coal firms. I again include the lowest distance to either Chicago or St. Louis and a linear time trend, as these variables could also affect coal demand at each firm.

$$Q_{ft} = \left( \frac{P_{ft}}{P_{0t}} \right)^\eta \exp(\mathbf{X}_{ft})^{\eta^x} \xi_{ft} \quad (14)$$

Taking logarithms of Equation (14) results in Equation (15), which is the demand model I estimate.

$$q_{ft} = \eta(p_{ft} - p_{0t} + \boldsymbol{\eta}^x \mathbf{X}_{ft} + \ln(\xi_{ft})) \quad (15)$$

### **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 inelastic product demand curve, a decrease in employer power should decrease output, as deadweight loss increases. In contrast, as was shown 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 bargaining in Bloch (1922) shows that employers 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 taking monopsony power into consideration.

### Variable input decisions

I assume firms make decisions in two stages. First, they choose whether to use cutting machines, or not. Second, conditional on this technology choice, they choose high-skilled wages, low-skilled employment, and materials. I will discuss these two stages in reverse order, as the estimation of the model will also be done in that order.

In year  $t$ , employers negotiate a high-skilled wage rate with the union according to the bargaining protocol specified in Equation (8). Following the labor supply curve, (13b), this negotiated wage rate results in a certain amount of high-skilled labor supplied at each coal firm. Coal firms simultaneously choose low-skilled labor as specified in Equation (9). I assume that these variable input decisions happen after the productivity shocks  $e^\omega$  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 (10). The coal demand curve (15) determines the price every firm can charge at that level of coal output.

### Fixed input choices

In every year  $t$ , the equilibrium quantities and prices at every firm can be found by solving the system of equations (13a), (15), (10); (8), and (6). This delivers equilibrium outcomes  $(Q_{ft}^1, P_{ft}^1, H_{ft}^1, L_{ft}^1, W_{ft}^1)$  if the firm uses cutting machines, and different equilibrium outcomes  $(Q_{ft}^0, P_{ft}^0, H_{ft}^0, L_{ft}^0, W_{ft}^0)$  if the firm does not use cutting machines. The variable profit gain of the employer from using cutting machines is denoted  $\Delta\Pi_{ft}^d$ :

$$\Delta\Pi_{ft}^d \equiv (P_{ft}^1 Q_{ft}^1 - W_{ft}^1 H_{ft}^1 - V_{ft} L_{ft}^1) - (P_{ft}^0 Q_{ft}^0 - W_{ft}^0 H_{ft}^0 - V_{ft} L_{ft}^0) \quad (16)$$

The fixed costs of technology usage are denoted  $\Phi_t$ , so total employer profits are equal to  $\Pi_{ft}^d - \Phi_t K_{ft}$ . As in Roberts, Peters, Vuong, and Fryges (2017), I parametrize fixed costs as an exponential distribution. I let the rate parameters  $(\phi_0, \phi_1)$ , evolve over time, with  $\phi_0$  measuring the time-invariant fixed cost of technology usage and  $\phi_1$  its time trend.

$$\Phi \sim \exp(\phi_0 + \phi_1 t)$$

I assume that prior to observing the productivity shocks  $e^\omega$  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  $\Phi_t$  with the expected variable profit return  $\Delta\Pi^d$ . I assume that firms do not choose their cutting machines in a dynamic manner, but optimize their technology mix period-by-period. The main reason for this assumption is that observed entry *and* exit of machine usage is frequent: whereas there are 109 observed instances of cutting machine installation, there are 62 observed instances in which an installed cutting machine is scrapped again. This suggests the existence of an aftermarket for capital.

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

$$p_{ft}^k(\phi_0, \phi_1) = 1 - \exp\left(\frac{-\Delta\Pi_{ft}^d}{\phi_0 + \phi_1 t}\right) \quad (17)$$

## 4 Identification and estimation

I now turn to the identification and estimation of the model. Although the model is specified at the firm-bi-year level, the dataset comes at the mine-year level. Given that firms are assumed to optimize at the firm-town level, I aggregate all the relevant variables to the firm-town-year-level, as detailed in Appendix A.2.

## 4.1 Labor supply estimation

I start with the identification of the skilled labor supply curve, Equation (13b). The labor supply elasticity  $\frac{1}{\psi}$  cannot be recovered by simply regressing skilled labor wages on employment because of the latent firm characteristics  $\zeta_{ft}$ . Firms with a high  $\zeta_{ft}$  know they are attractive to miners, and can hence offer a lower wage to attract the same number of miners. In order 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 coal price on international coal markets in Europe in every year from Degreè (1982). I use both this coal price and its interaction term with indicator for whether a mine shipped coal over the railroad network or not as instruments. These instruments imply three assumptions. First, individual Illinois coal mines were too small to have an effect on equilibrium coal prices on international trade markets. This makes sense given that Illinois produced only around 10% of the total U.S. output, and U.S. bituminous coal mines exported only around 1.2% of their coal in 1898 Graebner (1974). Second, international coal price shocks did affect the demand for Illinois coal. Given that Chicago was a main market for Illinois coal mines, and that Chicago also sourced coal by lake steamers from both the East coast and from other coal fields, changes in non-local coal prices affected demand for Illinois coal mines. Third, international coal price shocks affected coal mine demand more if coal mines were shipping their coal over the railroad network compared to coal mines who sold their coal only locally in their own town. This third assumption makes sense given that mines that sold only locally did not compete with coal fields outside of Illinois, given that both these mines and their consumers were not connected to the railroad network, in contrast to mines that did ship coal over the railroad network.

I compute the baseline wage level  $w_{0i(ft)t}$  as the average miner wage in Illinois. I estimate Equation (13b) 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 is a shipping mine or a mine that only sells locally, and include county fixed effects and a linear time trend.

For unskilled wages, I rely on the observed 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.

## Results

The skilled labor supply elasticity is estimated to be 3.799 with a standard error of 1.696, 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 20.84% below the marginal revenue product of labor.<sup>14</sup> Labor supply decreased over time.

## 4.2 Coal demand estimation

### Identification

I estimate the demand function in logarithms, (15), using firm-level quantity and prices. As firms in attractive locations, with a high  $\xi_{ft}$ , will set higher coal prices, this equation cannot be identified by simply regressing coal prices on quantities. I rely on the thickness of the coal vein as a cost shifter: whereas the vein thickness affects the marginal cost of mining, consumers do not care about it, as it does not affect coal quality (Affolter & Hatch, 2002). Vein thickness was the result of geological variation, and hence plausibly exogenous to coal firms, conditional on their location.

### Estimation

I estimate Equation (15) 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  $\mathbf{X}_{ft}$ , I include the following coal demand shifters: the log distance to Chicago and St. Louis, the number of railroads passing through the mine's town, and a linear time trend.

<sup>14</sup>The wage markdown is equal to  $\frac{MRPL-w}{MRPL} = \frac{1}{\psi+1}$ .

**Table 2: Model estimates**

<i>(a) Labor supply</i>		Est.	S.E.
Labor supply elasticity	$\frac{1}{\psi}$	3.799	1.696
1(Shipping mine)		1.522	0.324
Year		-0.038	0.005
First-stage F-stat		12.7	
Observations		6497	
<i>(b) Coal demand</i>			
Coal demand elasticity	$\eta$	-4.406	0.198
Log(min. distance to big city)		-0.255	0.062
No. railroads		0.478	0.041
Year		0.018	0.006
First-stage F-stat		802.9	
Observations		3192	
<i>(c) Production function</i>			
Input substitution elasticity	$\sigma$	0.259	0.066
Skill-augmenting technology effect	$\alpha^k$	0.125	0.073
Hicks-neutral technology effect	$\beta^k$	0.190	0.135
Low-skilled labor coefficient	$\beta^l$	0.001	0.072
Serial correlation Hicks-neutral productivity	$\rho^\omega$	0.356	0.239
Serial correlation skill-augmenting productivity	$\rho^\omega$	0.772	0.272
Observations		1786	
<i>(d) Fixed machine costs</i>			
Time-invariant fixed cost ('000 USD)	$\phi_{-0}$	35.339	7.661
Linear time trend in fixed costs ('000 USD)	$\phi_{-1}$	-3.506	0.872

**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. Panel (d) reports the average and median of the bargaining power distribution. Panel (e) reports the cutting machine fixed cost estimates.

I compute  $p_{0t}$  as the average coal price in each year.

## Results

The coal demand elasticity is in Table 2(b). The number of observations is lower, 3192, because the thickness of the coal veins is not observed in 1888 and 1890. The demand elasticity is estimated at -4.406 with a standard error of 0.198. This means that firms set coal prices at a markup of 29% above marginal costs. Coal demand is lower in markets that are further away from Chicago or St. Louis, and higher the more railroads pass through the mine's town. Coal demand grows at an average rate of 1.8% per two years throughout the sample period.

### 4.3 Production function estimation

I estimate the production function in two steps. In a first step, I estimate the elasticity of input substitution  $\sigma$  and the skill-augmenting effects of cutting machines,  $\alpha^k$ . Second, I estimate all other production function coefficients,  $\beta^l$  and  $\beta^k$ .

#### 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  $\gamma$  to zero in equation (8), which implies perfect monopsony power, gives:

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

Conversely, if  $\gamma$  approximates one, which implies full union power, the wage of high-

skilled workers goes towards their marginal revenue product:

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

In general, we can approximate between these two first order conditions by linearly interpolating these two extremes by the profit sharing parameter  $\gamma$ :

$$\frac{\partial R_{ft}}{\partial H_{ft}} = w_{ft}(1 + (1 - \gamma)\psi) \quad (18)$$

Working out both first order conditions (6) and (18), and dividing (6) by (18), results in a variant of the first stage regression from Doraszelski and Jaumandreu (2018):

$$l_{ft} - h_{ft} = \sigma(w_{ft}^h - v + \ln(1 + (1 - \gamma)\psi^h)) + \underbrace{\sigma \ln(\beta^l) + (1 - \sigma)a_{ft}}_{\equiv \tilde{a}_{ft}} \quad (19)$$

Given that Equation (11) 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  $\rho^a$  differences of (19) isolates the productivity shock  $e_{ft}^a$  as a function of the coefficients  $\rho^a$ ,  $\sigma$ , and  $\alpha^k$ . Using the previously stated assumptions that capital is chosen prior to observing the skill-augmenting productivity shock  $e_{ft}^a$ , but variable inputs are chosen afterwards, the moment conditions (20) can be specified to estimate the elasticity of input substitution  $\sigma$ , the skill-augmenting productivity effect of cutting machines  $\alpha^k$ , and the serial correlation in skill-augmenting productivity  $\rho^a$ :

$$\mathbb{E} \left[ e_{ft}^a(\rho^a, \alpha^k) \middle| \begin{pmatrix} K_{ft-r} \\ L_{ft-r-1} \\ H_{ft-r-1} \end{pmatrix} \right]_{r=0}^{T-1} = 0 \quad (20)$$

## Second-stage production function coefficients

From equation (19), the log factor-augmenting productivity residual  $a_{ft}$  can be written as a function of the estimated parameters  $\sigma$  and  $\psi$ , and the yet to estimate parameters  $\beta^l$  and  $\beta^k$ :

$$a_{ft} = \left( \frac{l_{ft} - h_{ft}}{1 - \sigma} \right) - \frac{\sigma}{1 - \sigma} (\ln(\beta^l)) - \frac{\sigma}{1 - \sigma} (w_{ft} - v_{ft} + \ln(1 + \psi))$$

Substituting this factor-augmenting productivity term into the log production function and assuming a linear function for the price control  $b(p_{ft})$  gives:

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

We define the first linear term in the log production function as  $B_{ft}(\cdot)$ :

$$B_{ft} \equiv \frac{\sigma}{\sigma - 1} \ln \left( \left( \exp \left( \left( \frac{l_{ft} - h_{ft}}{1 - \sigma} \right) - \frac{\sigma}{1 - \sigma} (\ln(\beta^l)) - \frac{\sigma}{1 - \sigma} (w_{ft} - v_{ft} + \ln(1 + \psi)) \right) H_{ft} \right)^{\frac{\sigma-1}{\sigma}} + \beta^l L_{ft}^{\frac{\sigma}{\sigma-1}} \right)$$

Using the productivity transition in Equation (12), taking  $\rho^\omega$  differences isolates the Hicks-neutral productivity shock  $e_{ft}^\omega$  as a function of the parameters  $(\rho^\omega, \beta^k, \beta^l)$ :

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

Making use of the timing assumptions that employers choose capital prior to the realization of the Hicks-neutral productivity shock, but choose low-skilled labor and bargain over wages after the realization of this shock leads to the moment conditions in equation (21). I

estimate this model using lags of up to one time period.

$$\mathbb{E} \left[ e_{ft}^{\omega}(\rho^{\omega}, \beta^k, \beta^l) \middle| \begin{pmatrix} K_{ft-r} \\ L_{ft-r-1} \\ H_{ft-r-1} \end{pmatrix} \right]_{r=0}^{T-1} = 0 \quad (21)$$

## Results

The production function estimates are reported in Table 2(c). The elasticity of substitution between skilled and unskilled miners is estimated at 0.259, hence, 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 complements rather than substitutes. Cutting machines are estimated to increase skill-augmenting productivity by 13.3%, 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), similarly to many other technologies that were developed throughout the 19th century, which were also unskill-biased (Mokyr, 1990; C. D. 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 21%. Finally, the low-skilled labor parameter  $\beta^l$  is estimated at 0.001, but is estimated imprecisely. More informatively, the corresponding output elasticities of low- and high-skilled labor are estimated at 0.56 and 0.44, respectively. Finally, both skill-augmenting and Hicks-neutral productivity are serially correlated, with serial correlations of 0.356 and 0.772.

## 4.4 Bargaining weights

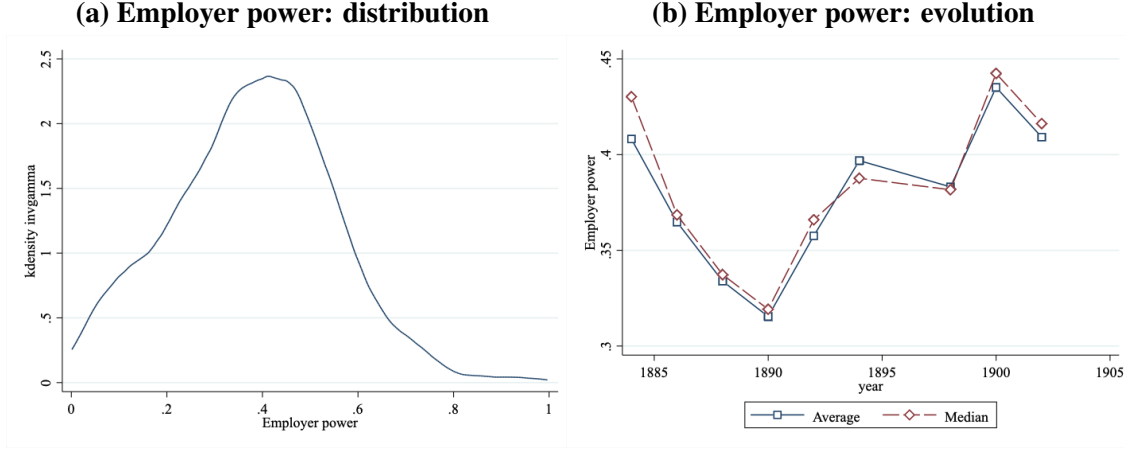
### Estimating the bargaining parameters

Rewriting Equation (8) in function of the union bargaining weights  $\gamma_{ft}$  leads to Equation (22). I estimate the bargaining parameters using this equation, of which all variables are either observed or have been estimated in the production, labor supply, and goods demand models.

$$\gamma_{ft} = \frac{(W_{ft} - Z_{ft})\left(\frac{\partial R_{ft}}{\partial H_{ft}} \frac{\partial H_{ft}}{\partial W_{ft}} - H_{ft} - W_{ft} \frac{\partial H_{ft}}{\partial W_{ft}}\right)}{(W_{ft} - Z_{ft})\left(\frac{\partial R_{ft}}{\partial H_{ft}} \frac{\partial H_{ft}}{\partial W_{ft}} - H_{ft} - W_{ft} \frac{\partial H_{ft}}{\partial W_{ft}}\right) - \left(\frac{\psi H_{ft}}{1+\psi} + \frac{\partial H_{ft}}{\partial W_{ft}}(W_{ft} - Z_{ft})\right)\left(\frac{\Pi_{ft}^d}{H_{ft}}\right)} \quad (22)$$

### Results

The employer's bargaining weight was on average 0.380, and 0.384 at the median firm. I keep only the bargaining parameter values that range in 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 13%. Figure 4a shows the entire distribution of employer power, 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). There are various explanations for this relative increase in employer power. First, there was an economic recession during the 1890s, which led to unemployment and depressed the workers' outside options. Second, as documented in Naidu and Yuchtman (2017), labor market institutions were biased against workers. For instance, the use of strike-breakers that were brought in from the Southern United States decreased the relative bargaining power of the labor unions. Although the 1897-1898 strikes did not succeed in



countering the overall rise in employer power, employer power did fall relatively at the striking mines compared to the mines that did not go on strike.<sup>15</sup>

## 4.5 Cutting machine fixed costs

### Estimation of fixed costs

Using the cutting machine usage probability from Equation (17), the log likelihood function of using cutting machines  $\ln(\mathcal{L}_{ft}(\phi))$  is equal to:

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

I estimate the fixed cost parameters  $(\phi_0, \phi_1)$  by maximizing the log likelihood function  $\ln(\mathcal{L}_{ft}(\phi))$ . Because the number of observed capital adoptions in the reduced-size panel on which the equilibrium model is estimated becomes sparse, I do not rely on the observed capital usages  $K_{ft}$  in the raw data, but 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 func-

<sup>15</sup>I document this in Appendix C.2.



tion to estimate cutting machine fixed costs. The resulting estimates are in Table 2(e). The fixed cost of using a cutting machine is estimated to be \$35,339 in 1882, and is estimated to fall by \$3,506 every two years throughout the sample period. Hence, the average machine fixed costs fell from \$31,883 in 1884, the first year of the sample, to \$ 2790 in 1902, its last year. Fixed costs are estimated to fall over time because machine uses increases whereas the variable profit returns to usage do not increase by much. This falling cost is in line with the falling costs of many new technologies. External cost information for cutting machines in 1889 is obtained from Brown (1889), which reports a purchasing cost of \$8000 for eight cutting machines. The average firm in the data set also used 8 cutting machines, so the estimated fixed cost in 1888 of \$7291 is slightly below the external cost estimate.

## 5 Counterfactuals

### 5.1 Observed equilibrium

Using the estimated model, I solve the system of equations (13a), (15), (10); (8), and (6) for every firm in every year. Given that this system of equations is nonlinear, and cannot be solved analytically, I solve for equilibrium numerically.<sup>16</sup> For every outcome variable  $Y \in \{Q, P, H, L, W\}$ , this yields an equilibrium outcomes if the firms use cutting machines,  $Y_{ft}^1$ , and if they do not,  $Y_{ft}^0$ . Using Equation (17), 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 uses machine usage and when it does not, weighted by the probability of using cutting machines:

$$\hat{Y}_{ft} = Y_{ft}^0 Pr(K_{ft} = 0) + Y_{ft}^1 Pr(K_{ft} = 1) \quad (23)$$

---

<sup>16</sup>I use the Matlab optimizer `fsolve` with function tolerance  $10^{-3}$ ,  $10^5$  maximum iterations, and 600 maximum function evaluations.

Appendix B.2 discusses how the predicted equilibrium values  $\hat{Y}_{ft}$  compare to the observed variables  $Y_{ft}$ . The model closely tracks the observed evolution of all equilibrium variables, despite the fact that the estimation procedure does not explicitly target any of these moments, except for capital investment.

## 5.2 Counterfactual equilibria

I conduct a counterfactual exercises that relate to the wage bargaining process between labor unions and employers. I examine how all equilibrium outcomes and welfare would have changed if employer power would have remained constant at its 1890 level, rather than increasing sharply during the 1890s. I compute counterfactual bargaining parameters by regressing the log bargaining parameter on year dummies with 1890 as base year, and then deducting the year fixed effects from the bargaining parameter to obtain the counterfactual bargaining parameter  $\tilde{\gamma}_{ft}$ .

$$\gamma_{ft} = \sum_{t \setminus \{1890\}} \delta_t + u_{ft}$$

$$\tilde{\gamma}_{ft} = \gamma_{ft} - \sum_{t \setminus \{1890\}} \hat{\delta}_t$$

## 5.3 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  $P_{ft}^*$ :

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

Similarly, I compute worker surplus  $WS_{ft}$  as the area between the labor supply curve

and the equilibrium wage  $\hat{W}_{ft}$ :

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

Finally, producer surplus is equal to total employer profits:

$$PS_{ft} \equiv \hat{P}_{ft} \hat{Q}_{ft} - \hat{H}_{ft} \hat{W}_{ft} - \hat{L}_{ft} V_{ft} - \phi_t \hat{K}_{ft}$$

## 5.4 Results

### No increasing employer power

Table 5a reports the equilibrium effects of keeping employer power constant after 1890, rather than letting it increase, as it did in reality. If capital investment would be exogenous, coal output would increase by 13%, which is due to a reduction in monopsony-induced deadweight loss. However, in reality, reducing employer power results in a decrease in cutting machine usage of 23%.<sup>17</sup> Hence, output would in reality increase by merely 9.5% due to reduced capital investment, as this investment reduction increases marginal costs. Similarly, the exogenous investment model overestimates the increases in employment and wages, and the reduction in prices. Both Hicks-neutral and skill-augmenting productivity fall on average when employer power is reduced, due to the fall in capital investment.

The welfare counterfactuals are reported in Table 5b. Consumer surplus is estimated to increase by 10.6% on average in the exogenous investment model, which reduces to 8.8% in the endogenous investment model. The increase in labor surplus is higher in the endogenous investment model, at 21.3%, compared to the exogenous investment model, at 20.9%. The reason for this is that cutting machines are unskill-biased, so reduced usage leads to a higher relative marginal revenue product for skilled workers compared to unskilled work-

---

<sup>17</sup>The usage rate of 8.6% in the observed equilibrium is lower as the usage rate of 11% in Figure A2e because the comparison table conditions on the equilibrium being solved in both the actual and counterfactual equilibrium. This leads to a different sample selection with slightly lower capital usage rates.

ers. Hence, reduced technology usage leads firms to substitute unskilled workers for skilled workers, although the magnitude of this effect is modest given the small elasticity of substitution between skilled and unskilled workers. The finding that employer power over skilled workers leads firms to substitute towards skilled workers is the opposite finding as in Goolsbee and Syverson (2019), which consider substitution between worker types conditional on investment, in a wage-setting monopsony model and with exogenous technology usage. However, as I show below, inverting the direction of technological change flips the sign of this effect.

The starkest difference in welfare effects lies in producer surplus. Assuming exogenous capital investment, producer surplus would fall by 36.4%. In reality, producer surplus falls by only 14.3%. The main reason for this difference is that under the exogenous capital usage model, firms continue to use too much capital, considering the rents they obtain after bargaining with the workers. In the endogenous investment model, firms respond to their lower bargaining power by using less capital, thereby saving on fixed costs. Finally, total surplus is estimated to increase by merely 1.9% when assuming exogenous investment, whereas it increases to 5.3% under endogenous capital investment. Hence, an increase in worker bargaining power does not merely redistribute surplus between employers, consumers, and workers, but also increases total surplus.

### **Lower fixed costs**

In the counterfactual exercise above, the hold-up mechanism is dominated by the deadweight loss mechanism. One reason for this is that fixed costs are estimated to be high, in order to reconcile the low adoption rate of cutting machines in the data. This low usage rate mechanically reduces the aggregate importance of hold-up. In order to gauge the relative effects of hold-up and deadweight loss in settings with higher capital usage rates, I re-estimate the same counterfactual exercise under a lower calibrated fixed cost of 10% of the estimated value. Given the fast decrease in estimated fixed costs over time, this value of fixed costs would be reached by 1905, three years after the sample ends. The results of this counterfactual are reported in Table 4. Under the lower fixed costs level, the hold-up

**Table 3: Counterfactual: no increase in employer power**

<i>(a) Equilibrium quantities and prices</i>	Actual	C.F. (exog. K)	C.F. (endog. K)
Output	2467.204	2787.442	2701.744
Price	1.570	1.538	1.542
High-skilled labor	873.680	989.929	974.549
Low-skilled labor	378.667	423.527	414.339
High-skilled wage	1.885	1.951	1.945
Cutting machine usage	0.086	0.086	0.066
Hicks-neutral productivity	0.856	0.856	0.853
Skill-augmenting productivity	4.953	4.953	4.942
<i>(b) Welfare</i>	Actual	C.F. (exog. K)	C.F. (endog. K)
Consumer surplus	860.481	951.754	936.565
Producer surplus	254.017	161.537	217.761
Worker surplus	117.170	141.657	142.142
Total surplus	1231.669	1254.947	1296.468

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

mechanism dominates the deadweight loss channel: equilibrium output now decreases in the counterfactual with lower employer power. Although the relative decline in capital usage is the same as in the main counterfactual, this has much larger absolute effects given the higher level of capital usage. Similarly to output, total surplus also falls in the counterfactual of reduced employer power, whereas assuming exogenous investment would lead to a predicted increase in total welfare.

**Table 4: Counterfactual with lower fixed costs**

<i>(a) Equilibrium quantities and prices</i>	Actual	C.F. (exog. K)	C.F. (endog. K)
Output	3678.564	4101.861	3860.979
Price	1.475	1.447	1.465
High-skilled labor	1119.959	1243.215	1197.826
Low-skilled labor	511.320	563.515	534.372
High-skilled wage	1.981	2.052	2.025
Cutting machine usage	0.466	0.466	0.375
Hicks-neutral productivity	0.919	0.919	0.907
Skill-augmenting productivity	5.162	5.162	5.105
<i>(b) Welfare</i>	Actual	C.F. (exog. K)	C.F. (endog. K)
Consumer surplus	1147.547	1276.203	1218.959
Producer surplus	626.281	483.014	503.038
Worker surplus	163.270	185.675	186.059
Total surplus	1937.098	1944.892	1908.056

**Notes:** In this table, I repeat the same counterfactual exercise from Table 5, but calibrating fixed cutting machine costs at 10% of their actual level.

### Inverse direction of technological change

Finally, I re-compute the same counterfactual exercise under the true level of fixed costs, but for a counterfactual technology that is skill-biased, rather than an unskill-biased technology, as cutting machines were. I keep the same level of fixed costs and assume an identical parametrization of the production function, with the only exception that the technology now decreases skill-augmenting productivity by 13.3%, rather than increasing it by 13.3%.<sup>18</sup> First, I find that the usage rate of an skill-biased technology would be considerably lower, at 2.7% in the observed equilibrium compared to 8.6% for the unskill-biased technology. Cutting machines increase skill-augmenting productivity, which in-

<sup>18</sup>Hence,  $\alpha^k = -0.125$  rather than  $\alpha^k = 0.125$ .

creases profits as firms have employer power over skilled workers, and thus capture part of the skill-augmenting productivity increase. The counterfactual technology does the opposite, and increases unskill-augmenting productivity instead. However, the market for unskilled workers is perfectly competitive, so employers do not capture any of the rents that are created by increasing the marginal revenue product of unskilled workers. Second, the skilled labor welfare comparison between the endogenous and exogenous investment case flips. In the cases of cutting machines, a reduction in investment led to higher skilled labor surplus, because cutting machine usage leads firms to substitute away from skilled workers. In contrast, the counterfactual skill-biased technology leads firms to substitute away from unskilled workers. Hence, a reduction in investment in the skill-biased technology due to lower employer power leads to lower skilled labor surplus compared to the exogenous investment case.

## 6 Conclusion

In this paper, I investigate the welfare effects of employer power by studying the trade-off between monopsony distortions and investment hold-up. Using a model of weakly efficient bargaining and linear wage contracts, 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 and of the productivity effects of new technologies, 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 the monopsony distortion dominated the hold-up effects on coal cutting machine adoption. Increasing the relative bargaining position of labor unions would, hence, have resulted in total welfare gains. However, under lower fixed costs of cutting machines, which corresponds to the time period just after the observed sample, the hold-up effect would have started to dominate the deadweight loss channel.

**Table 5: Counterfactual: skill-biased technology**

<i>(a) Equilibrium quantities and prices</i>	Actual	C.F. (exog. K)	C.F. (endog. K)
Output	2123.869	2405.223	2399.789
Price	1.588	1.555	1.556
High-skilled labor	808.750	920.057	919.792
Low-skilled labor	342.732	384.033	384.010
High-skilled wage	1.866	1.930	1.930
Cutting machine usage	0.027	0.027	0.020
Hicks-neutral productivity	0.844	0.844	0.843
Skill-augmenting productivity	4.892	4.892	4.896
<i>(b) Welfare</i>	Actual	C.F. (exog. K)	C.F. (endog. K)
Consumer surplus	799.119	882.572	880.475
Producer surplus	350.903	265.596	289.673
Worker surplus	107.231	131.600	130.395
Total surplus	1257.253	1279.768	1300.543

**Notes:** In this table, I repeat the same counterfactual exercise from Table 5, but setting the cutting machine effects on skill-augmenting productivity to -0.125, rather than 0.125.

## 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 Economic 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.
- Chambolle, C., Christin, C., & Molina, H. (2023). Buyer power and exclusion: A progress report. *International Journal of Industrial Organization*, 102969.
- 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. (2022). Monopsony and industrialization. *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. D., & Katz, L. F. (2009). *The Race between Education and Technology*. Harvard University Press.
- Goolsbee, A., & Syverson, C. (2019). *Monopsony power in higher education: a tale of two tracks* (Tech. Rep.). National Bureau of Economic Research.
- Graebner, W. (1974). Great expectations: The search for order in bituminous coal, 1890-1917. *The Business History Review*, 49-72.
- Hoover, E. D. (1960). Retail prices after 1850. In *Trends in the American economy in the nineteenth century* (pp. 141-190). Princeton University Press.
- 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. (2019). *Granular search, market structure, and wages* (Tech. Rep.). National Bureau of Economic Research.
- 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.
- 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*, Forthcoming.
- Parra, A., & Marshall, G. (2021). *Monopsony power and upstream innovation* (Tech.

Rep.).

- Reid, A. (1876). *Transactions of the North of England Institute of Mining and Mechanical Engineers*. North of England Institute of Mining and Mechanical Engineers.
- Roberts, M., Peters, B., Vuong, V. A., & Fryges, H. (2017). Estimating Dynamic R&D Demand: An Analysis of Costs and Long-Run Benefits. *Rand Journal of Economics*, *Forthcoming*.
- Rubens, M. (2023). Market structure, oligopsony power, and productivity. *American Economic Review*, *113*(9), 2382–2410.
- Shi, L. (2023). Optimal regulation of noncompete contracts. *Econometrica*, *91*(2), 425–463.
- 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**

The 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 includes 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 use the 1880 population census to have 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 census of manufacturing, and the number of farms and improved farmland area from the 1880 census of agriculture.

#### **Monthly data**

The 1888 report contains monthly production data for a selection of 11 mines in Illinois, across 6 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, etc.). 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 employee’ 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 rates wages using the number of workers employed in each season for the years 1884-1890. 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 onwards, 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 (2000 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 1892-1894 or 1896-1898 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. In 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.<sup>19</sup>

## **Hours worked**

In 1898, eight-hour days were enforced by law, which means that the ‘number of days’ measure changes in unit between 1898 and 1900. As 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 ‘workday’, 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, mine names were made consistent over time as much as possible.

## **Town identifiers and labor market definitions**

The data set 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 3 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 448 to 374. The resulting labor markets lie at least 3 miles from the nearest labor

---

<sup>19</sup><https://www.minneapolisfed.org/about-us/monetary-policy/inflation-calculator/consumer-price-index-1800->



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 on a crossroad of multiple railroad lines. Towns not located on railroads are assumed to be isolated coal markets. For the connected towns, the market is defined as the railroad line on which they are located, of which there are 26. 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 number of labor-days 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 on 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 ‘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: Model

### B.1 Simulating the theoretical model

#### Baseline parametrization

In Section 2.3, I simulate the theoretical model with the following parameters being scalars:  $\eta = -10$ ,  $\psi = 1.5$ ,  $\sigma = 0.7$ ,  $\beta^l = 0.02$ ,  $v = 0.1$ ,  $\xi = 0.5$ ,  $\zeta = 0.5$ ,  $w_0 = p_0 = 1$ ,  $\omega = 1$ ,  $a = 1$ . I simulate a dataset with 50 observations, in which the bargaining parameter  $\gamma_f$  is distributed uniformly between 0 and 1. I let fixed technology costs be distributed as an exponential distribution with its mean being equal to the average variable profit gain of technology adoption.

Under these parametrizations, I solve the system of equations (1), (2), (7), (6), (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 absolute 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

#### Non-constant returns to scale

In the main text, the production function (1) relied on constant returns to scale. In contrast, Equation (26) allows for non-constant returns to scale, as parametrized by  $\nu$ .

$$Q_f = \left( (A_f(K_f)H_f)^{\frac{\sigma-1}{\sigma}} + \beta^l L_f^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\nu\sigma}{\sigma-1}} \Omega_f(K_f) \quad (24)$$

The first step of the production function estimation procedure, the estimation of Equation (19) remains the same. However, the second step of the estimation procedure needs to estimate the scale parameter  $\nu$  in addition to the other production function coefficients  $\rho^\omega$ ,  $\beta^l$ , and  $\beta^k$ . Given that we have four instruments (lagged employment for both labor types, current and lagged capital), the model is still identified.

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

The results are in the first column of Table A4. The scale parameter is estimated at 1.038, 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 that 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,  $\beta^m$  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 needed to be added. Conform with this assumption, I include current and lagged materials as an additional instrument when estimating the production function.

$$Q_f = \left( (A_f(K_f)H_f)^{\frac{\sigma-1}{\sigma}} + \beta^l L_f^{\frac{\sigma-1}{\sigma}} + \beta^m M_f^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\nu\sigma}{\sigma-1}} \Omega_f(K_f) \quad (25)$$

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 0.551.

### Capital and returns to scale

It could be that the degree of returns to scale 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 non-fatal 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.

$$Q_f = \left( (A_f(K_f)H_f)^{\frac{\sigma-1}{\sigma}} + \beta^l L_f^{\frac{\sigma-1}{\sigma}} + \beta^m M_f^{\frac{\sigma-1}{\sigma}} \right)^{\frac{(\nu_0 + \nu_1 K_f)\sigma}{\sigma-1}} \Omega_f(K_f) \quad (26)$$

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. The remaining parameters again stay similar, except for the capital coefficient that is now estimated at 0.433, albeit very imprecisely.

## **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  $\gamma_{ft}$  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. Union power is estimated to increase by 3%, 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 8.5%, and this change is statistically significant.

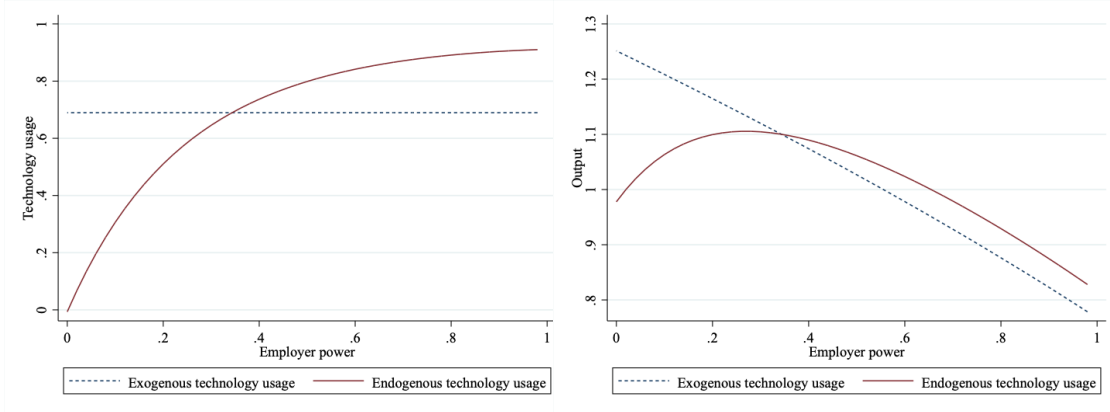
### **Cost dynamics**

In Table A2, I test for cost dynamics by regressing labor productivity, measured as output per labor day, on log cumulative output. This is in the same spirit of Benkard (2000). 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.

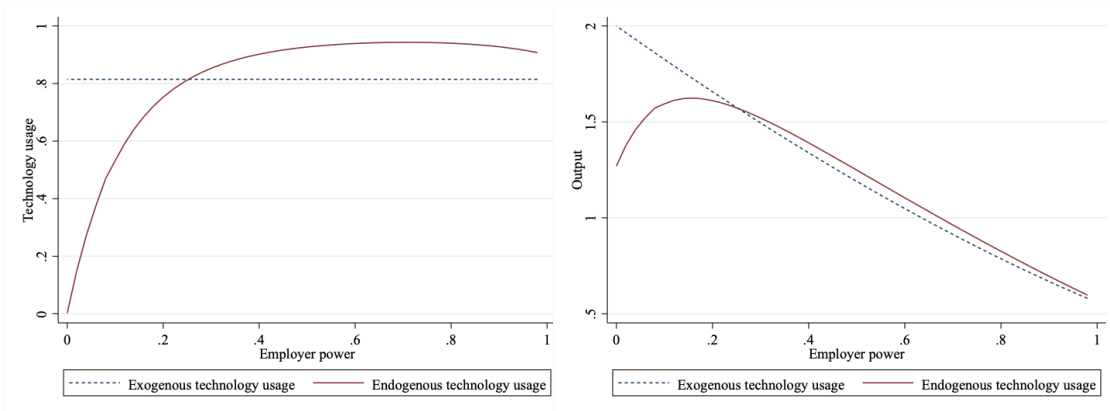
### C.3 Appendix tables and figures

**Figure A1: Simulations: alternative parametrization**

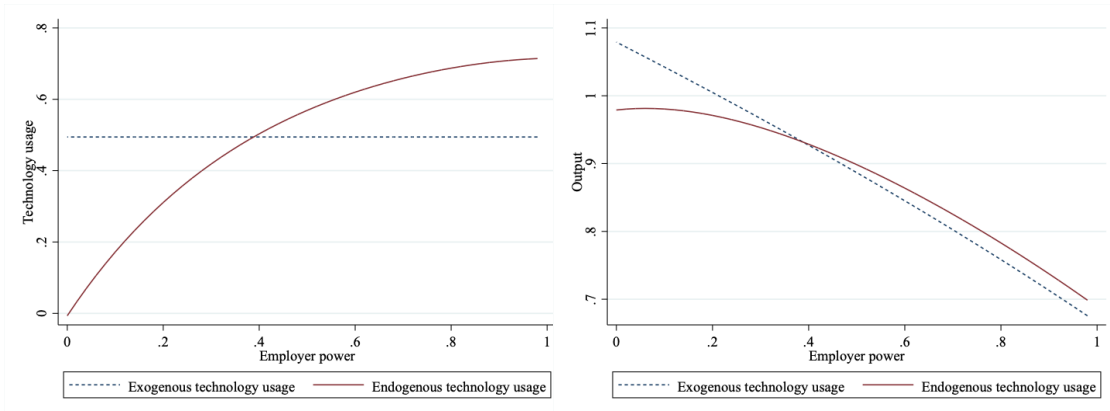
**(a) Baseline:  $\psi = 1.5, \beta^k = 0.2$**



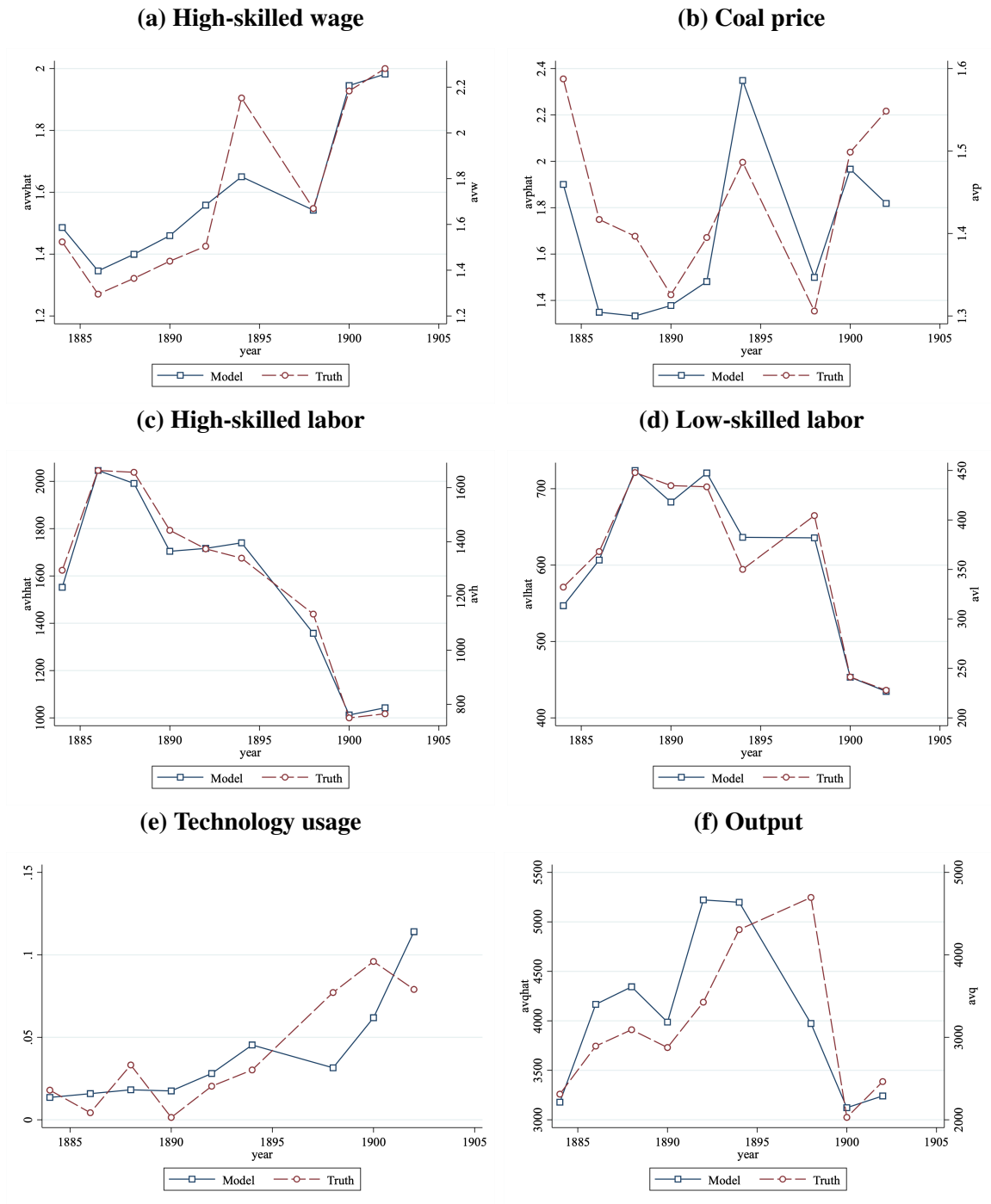
**(b) More inelastic labor supply:  $\psi = 0.5, \beta^k = 0.2$**



**(c) Smaller productivity effect:  $\psi = 1.5, \beta^k = 0.1$**



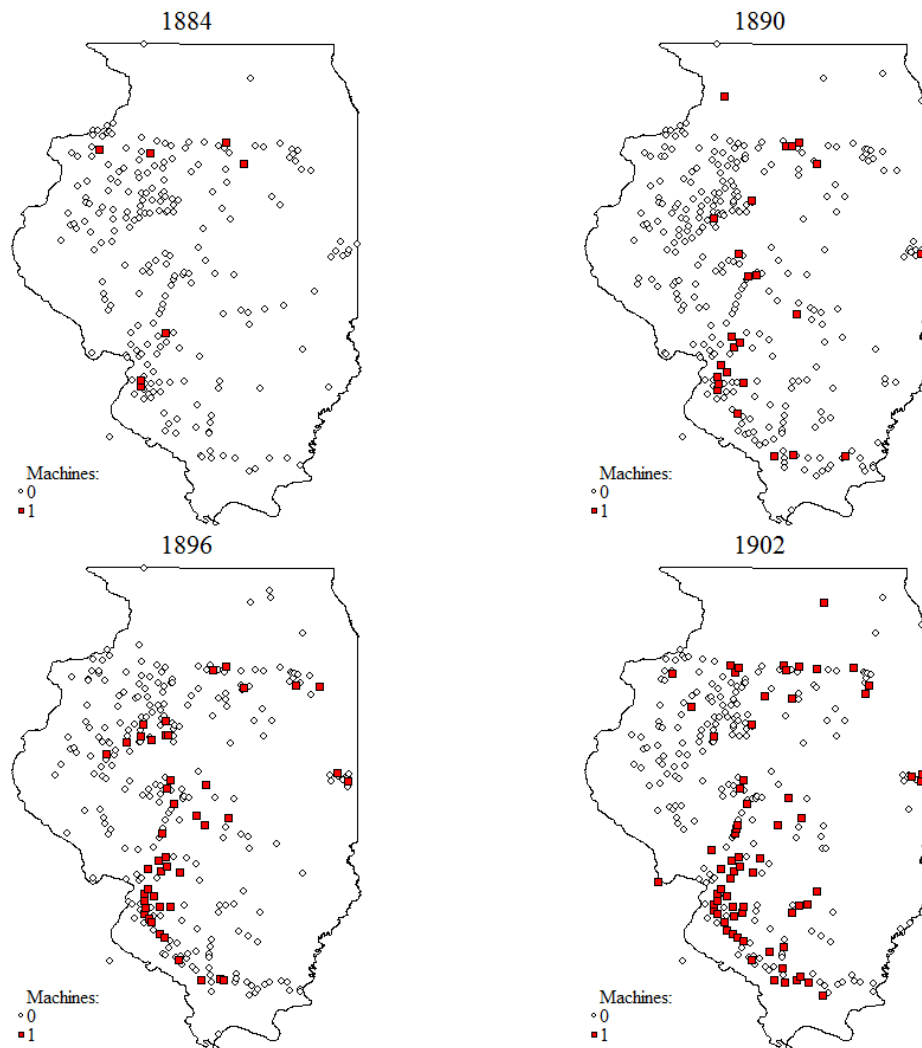
**Figure A2: Model fit**



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

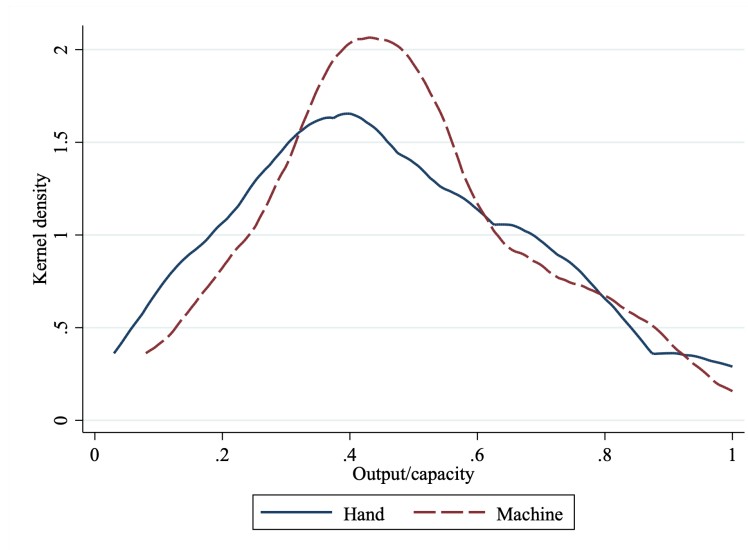


**Figure A3: Geographical spread of cutting machines**



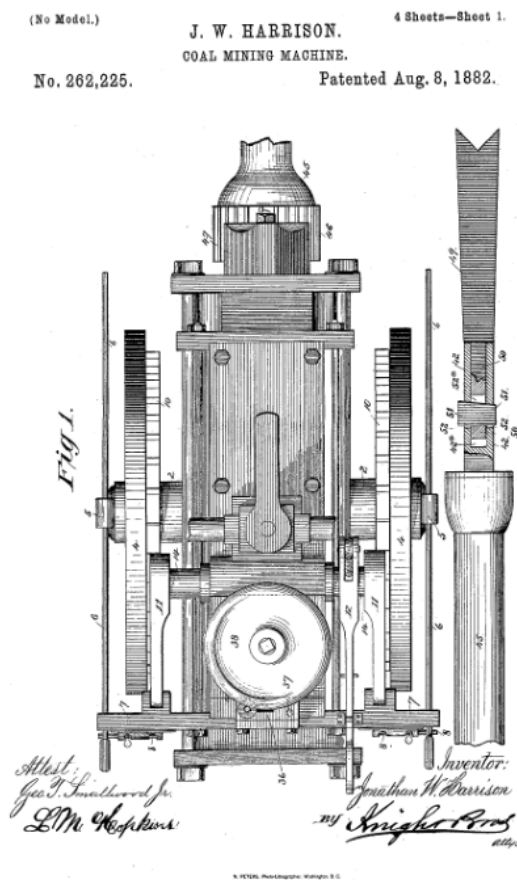
**Notes:** The dots indicate mining towns, each of which can contain multiple mines. Villages with squares contain at least one machine mine.

**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 1898. A distinction is made between hand mines, which did not use cutting machines, and machine mines, which did so.

**Figure A5: Patent of the Harrison Cutting Machine**



**Notes:** U.S. patent of the 1882 Improved Harrison Coal Cutting Machine (Whitcomb, 1882). This was the most frequently used coal cutting machine in the data set.

**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 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(Cum. output)	0.124	0.004	-0.009	0.017
Mine FE	No		Yes	
Observations	3717		3717	
R-squared	.327		.811	

**Notes:** Regression of log output per worker-day against log cumulative output (lagged by one time period) at the mine-year level. Sample only includes 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)	0.117	0.072	0.175	0.063
Strike indicator:	Any		Successful	
R-squared	.703		.664	
Observations	3469		3918	

**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**

	Non-constant RTS		Adding materials		Capital and RTS	
	Est.	S.E.	Est.	S.E.	Est.	S.E.
Returns to scale	1.038	0.030	1.051	0.094	0.974	0.141
Labor coefficient	0.001	0.002	0.002	0.009	0.000	4.242
Capital coefficient	0.130	0.132	0.114	0.552	0.885	0.988
Serial corr. TFP	0.390	0.106	0.562	0.224	0.400	0.306
Materials coefficient		.	0.000	0.046	0.000	0.413
Returns to scale * K		.		.	-0.000	0.050
Observations	732		332		332	

**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	0.185	0.290	0.736
Year F.E.	X	X	X	X
County F.E.		X	X	X
Town F.E.			X	X
Firm F.E.				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	'00	'02
<b>Output quantities</b>										
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
<b>Input quantities</b>										
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 dollar)	X									

**Table A7: All variables per year (cont.)**

Year	1884	'86	'88	'90	'92	'94	'96	'98	'00	'02
<b>Output price</b>										
Price/ton at mine	X	X	X	X	X			X	X	X
Price/ton at mine, lump						X	X	X		
<b>Input prices</b>										
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			
<b>Technicals</b>										
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 egress places	X	X								
Ventilation type	X	X								
New/old mine					X	X				
# Acres					X	X	X			
Mine capacity								X		
Mined or blasted								X		
<b>Cutting machine usage</b>										
Cutting machine dummy					X	X	X	X		
# Cutting machines	X	X	X	X						
# Tons cut by machines									X	X
# Cutting machines, by type			X							