Death by Market Power and the Production-Safety Tradeoff in the Coal Mining Industry*

Honghao Zheng†

October 17, 2024

Job Market Paper

Click here for the latest version

Abstract

I examine the effects of buyer power on the organization of production in the Chinese coal mining industry. I show how buyer power emanating from downstream coal-fired power plants ultimately affects the production-safety tradeoff of upstream coal mines. I estimate a structural model of coal mines in imperfect output markets featuring (i) joint production of coal and worker safety as outputs, and (ii) choosing where to locate on the production-safety frontier via endogenous safety choices and factor-augmenting productivity. To identify causal effects, I employ a shift-share instrumental variable, leveraging exogenous variations stemming from an electricity sector restructuring and other demand-side shocks. I find an unintended but life-and-death consequence associated with market power—buyer power exposure leads to higher provincial death rates, which are corroborated by lower composite coal outputs and safety at the mine level. Further evidence indicates exposure of coal mines to buyer power prompts a shift toward less capital-intensive, more traditional, and less safe mining technologies. Back-of-the-envelope calculations suggest that the decline in buyer power explains 53% of the improvement in coal mining death rates.

Keywords: Market power, Buyer power, Production function, Productivity, Technological changes, Coal mining industry.

JEL Codes: D24, L13, L23, O14, O33

*I am grateful to Johannes Van Biesebroeck, Jan De Loecker, Frank Verboven, and Allan Collard-Wexler for their constructive advice. I also thank Daniel Ackerberg, Gianluca Antonecchia, Emily Cuddy, Etienne Guigue, Nadine Hahn, Matthew Johnson, Leslie Marx, Michael Rubens, Paul Scott, Leonard Treuren, Daniel Yi Xu, and participants at various workshops and conferences for their valuable comments and suggestions. Special thanks to Bo Chen for kindly sharing the coal mine accident data with me.

†KU Leuven, Naamsestraat 69, 3000 Leuven, Belgium. E-mail: honghao.zheng@kuleuven.be

1 Introduction

Market power is pervasive across industries and countries in both input and output markets.[1] A growing literature studies the consequence of market power, or competition, on welfare distribution (De Loecker et al., 2021; Edmond et al., 2023), allocative efficiency (Cicala, 2022; Rubens, 2023), innovation (Aghion et al., 2005; Bloom et al., 2016), and productivity (Olley and Pakes, 1996; Fabrizio et al., 2007; Backus, 2020), among others. However, though Backus (2020) directs that within-firm adjustment mechanisms are of first-order standing for revealing the effects of market power, less is known about how market power affects the within-firm organization of production.

In this paper, I speak to this question by empirically examining the effects of downstream buyer power on the within-firm organization of production in the Chinese coal mining industry, through which coal mines choose where to be on the production-safety frontier and determine the corresponding tradeoff. The Chinese coal mining industry is exposed to substantial buyer power from downstream coal-fired power plants while concurrently contending with record-breaking high coal mining death rates—#death per million tons of coal—a consequent outcome of coal and safety decisions. Meanwhile, exogenous demand-side shocks in the buying industry, specifically the electricity sector, bring rich variation in the coal mining industry's exposure to buyer power, providing a unique causal inference framework to trace its impacts on the joint production of coal and worker safety.

There are at least three reasons why answering this question empirically is challenging. First, the measurement of market power frequently encounters some measurement issues. Either by relying on concentration ratio, e.g., Herfindahl-Hirschman Index (HHI), which sensitively depends on market definitions (Syverson, 2019; Berry et al., 2019), or by using the production approach when physical input/output data is commonly unavailable, incurring challenges to estimate output elasticities consistently (De Loecker and Goldberg, 2014; De Loecker and Syverson, 2021).[2] Second, within-firm multi-output data, let alone input allocation information by output (Orr, 2022), is of limited availability in widely used firm-level datasets. Third, market power is endogenous to a firm's production decisions, inducing simultaneity bias to estimates of the market power impacts.

I address the above-mentioned concerns by first employing the production approach with a non-substitutable technology following De Loecker and Scott (2022) in the coal-fired power

[1] Market power has been quantified using different methods and data in the product markets (De Loecker et al., 2020; Benkard et al., 2021; Döpper et al., 2024); in the input markets (Berger et al., 2022; Yeh et al., 2022; Rubens, 2023); across countries (De Loecker and Eeckhout, 2018); across industries (Asker et al., 2019; De Loecker and Scott, 2022; Gentzkow et al., 2024; Grieco et al., 2024); among others.

[2] One could also employ an underlying demand system and assume firm conduct to estimate demand primitives and infer market power information. See Berry et al. (1995) and Nevo (2001).

generation sector, i.e., the vertically downstream industry to coal mines. I take advantage of a novel and comprehensive plant-level power generation dataset, providing both physical input and output information for each power plant. These benefits leave my specification free from the input/output price bias (Klette and Griliches, 1996; De Loecker and Goldberg, 2014) and facilitate a consistent measure for the buyer power of power plants without assuming any model of bilateral competition and conduct.

Second, I construct a structural model for coal mines to examine the production-safety tradeoff in the spirit of Grieco and McDevitt (2017).[3] I exploit coal mine-level accident records to construct a measure of the safety output at the mine level, i.e., the coal mine's predicted accident probability, further instrumenting with focal depth-weighted earthquake magnitudes in estimation. I incorporate a transformation function governing the production process from multi-inputs to multi-outputs of coal quantity and safety level without imputing the output-specific input allocation scheme. In particular, I consider a general constant elasticity of substitution technology with factor-augmenting productivity incorporated following Doraszelski and Jaumandreu (2018). I first utilize the ratio of the coal mine's optimal first-order conditions for labor and material to infer and control for the unobserved labor-augmenting productivity. The identifying variation comes from different effects of labor-augmenting productivity on the optimal factor demand for labor and material. In contrast, I use a control function with the previously measured buyer power exposure and other necessary factors to control for the unobserved Hicks-neutral productivity under imperfect competition and unobserved demand shocks. In addition, I employ an endogenous productivity process (De Loecker, 2011, 2013), which nests the classical exogenous law of motion of productivity as a special case, for both Hicks-neutral and labor-augmenting productivity to allow the buyer power to affect coal mines' future productivity.

Third, I exploit spatial-temporal variation in buyer power exposure to the coal mining industry caused by exogenous changes from an electricity sector restructuring and other demand-side shocks to construct a shift-share instrumental variable (IV) to estimate causal effects.[4] I follow Borusyak et al. (2022) to assume that the shifters (provincial demand-side shocks in coal consumption from the coal-fired power generation sector) are conditionally orthogonal to the coal mining industry outcomes, allowing the shares (historical interprovincial coal selling networks) to be endogenous. By leveraging the IV, I conduct causal inferences

[3] Grieco and McDevitt (2017) propose an estimating framework with neutral productivity to measure the magnitude of the quality-quantity tradeoff in the dialysis industry context.

[4] Recent literature formalizes identifying properties for shift-share (or "Bartik") instruments (Goldsmith-Pinkham et al., 2020; Borusyak et al., 2022) and implements them in different contexts (Autor et al., 2013; Hummels et al., 2014; Imbert et al., 2022). The causal identification of shift-share instruments can come from either conditional exogenous shares (Goldsmith-Pinkham et al., 2020) or shifters (Borusyak et al., 2022), allowing the other component to be endogenous.

to study how buyer power affects provincial death rates in the coal mining industry. The death rate measure, i.e., #death per million tons of coal, straightforwardly correlates coal production with safety.

Finally, I further zoom in to the coal mine level to investigate the within-firm adjustments in multi-output production, input decisions, cost of capital, and technological changes in response to variations in the exposed buyer power. Crucially, although the specific technology adoption information is commonly unavailable at firm-level datasets, I can infer technological changes for coal mines from input intensity patterns, exploiting the distinct engineering characteristics across different mining technologies.

The estimates reveal that power plants procure coal on average at a price 12% below their marginal revenue product of coal. The time-series variations of buyer power from power plants well capture the shock of electricity sector restructuring, slumping on average by 7.7 percentage points in 2004 and onwards. This corroborates that the restructuring brings exogenous shocks to coal mines' buyer power exposure when constructing the shift-share IV. Regarding coal mines' production, I find a substantial production-safety tradeoff for coal mining: a coal mine aiming to reduce the probability of accidents by 1 percentage point would need to decrease its coal production by 3.8%. Put differently, increasing coal output by 1% would necessitate a reduction in safety level, leading to a 0.26 percentage point rise in accident probability, holding others fixed. The significant tradeoff between safety and production in coal mining aligns with Gowrisankaran et al. (2015).

In addition, I find a 1% increase in the buyer power exposure can lead coal mines' future Hicks-neutral and labor-augmenting productivity to decrease by 0.22% and 0.19%, respectively, which both reject an exogenous productivity process. However, the way how buyer power affects future productivity is heterogenous: the effects of buyer power on future Hicks-neutral productivity vary substantially depending on the coal mine's productivity, while it affects labor-augmenting productivity directly and (log-)linearly regardless of the distinct productivity levels.

Using all the estimates and the constructed IV, I examine how buyer power affects provincial death rates in the coal mining industry and other within-firm adjustment margins. The results indicate that the exposed buyer power in the coal mining industry increases by 1% would incur the death rate increase by 4.55%, mapping to a rise in deaths by 11 people per province a year, holding the coal production quantity fixed. Coal mine-level evidence presents qualitatively consistent findings that, on average, a 1% increase in buyer power of power plants leads to a 0.93% drop in coal output if the safety level is held fixed, or equivalently, 0.24 percentage points reduction in safety level, keeping the coal output constant. This empirical evidence corroborates and aligns well with theoretical findings.

To investigate the mechanism behind this, I examine how buyer power exposure affects technological changes, which naturally affect coal mines' efficiency and safety performance. I find that a 1% increase in buyer power induces a decline in capital adoption by 1.67% and reduces the capital-to-labor ratio by 1.23%. In contrast, the adjustments in capital-to-material and material-to-labor ratios in response to higher buyer power exhibit reductions in point estimates but are not statistically significant. Combining these patterns with distinct engineering details in input intensity across different mining technologies, the underlying mechanism emerges: higher buyer power could reduce the coal mine's capital adoption and shift the mining technology from conventionally-mechanized mining to blasting mining, which tends to be more dangerous by nature, leading to a higher death rate. Further evidence shows that higher buyer power raises the cost of capital, partially explaining the adjustments in capital adoption and technical upgrades.

To support the above findings, I conduct a battery of robustness checks and show that all results are robust. In particular, as an extension of my model, I allow for the heterogeneity in the production-function tradeoff, by which the slope of the production-safety frontier could vary across mines depending on their technology or capital-to-labor ratio. I find that coal mines utilizing more advanced modern technology tend to flatten the production-safety tradeoff, expanding the production frontier toward safety compared to coal mines employing relatively traditional technology. This evidence aligns well with the technological characteristics associated with distinct mining methods in practice.

Using estimates obtained from the model, I impute that the opportunity cost of saving one miner's life is roughly 9 million RMB Yuan, while the "national benchmark" of the compensation payments in post-accident settlements for each worker killed was only 2% of the former. This tends to drive coal mine owners to have much higher incentives to sacrifice safety rather than forgo coal production. On top of this, the back-of-the-envelope calculations indicate substantial unintended consequences of reducing buyer power on coal mining death rates. Declined buyer power explains 53% of the improved coal mining death rate performance.

Related Literature. I contribute to three strands of literature. First, I examine how market power affects the within-firm organization of production. Whereas a large body of empirical literature exists on the consequence of market power and competition in the fields of industrial organization (Olley and Pakes, 1996; Aghion et al., 2005; Asker et al., 2019; Backus, 2020; Rubens, 2023), macroeconomics (De Loecker et al., 2021; Edmond et al., 2023), international economics (Pavcnik, 2002; Bloom et al., 2016; Brandt et al., 2017), and energy economics (Fabrizio et al., 2007; Gao and Van Biesebroeck, 2014; Cicala, 2022), among others, less is documented about the within-firm impacts of market power on multi-output production and multi-dimensional productivity. In contrast, I estimate the effects of buyer power on these margins and find significant within-firm adjustment mechanisms involved, which is crucial to be considered when attempting to understand the holistic impacts of market power.

Second, this paper adds to the literature on production function estimation with factor-biased productivity (Van Biesebroeck, 2003; Doraszelski and Jaumandreu, 2018; Zhang, 2019; Demirer, 2022; Rubens, 2024; Rubens et al., 2024). I develop a framework for estimating transformation functions when production is multi-output and productivity is multi-dimensional, without imputing input allocation scheme by output as in the multi-product production function estimation literature (De Loecker et al., 2016; Orr, 2022; Valmari, 2023). The framework provides valuable insights for analyzing production tradeoffs beyond the scope of the production-safety context discussed in the paper, such as identifying tradeoffs between production and emissions as in Shu et al. (2024).

Third, this study shows an unintended but literally life-and-death consequence associated with market power. I find that buyer power from the buying industry could ultimately affect technological adoptions and death outcomes of firms in their input market, complementary to Gaynor et al. (2013) that indicates competition could save lives in the output market in a healthcare context. The findings suggest significant implications for competition policy that extend beyond safety outcomes, which further contribute to the recent literature on the competitive effects of mergers in markets with buyer power (Loertscher and Marx, 2019a,b, 2021). As buyer power becomes prominent in antitrust litigation and debates, competition authorities should consider unintended externalities, in addition to concerns about consumer welfare in the product market, when exercising their discretion in merger and antitrust evaluations, particularly within upstream industries.

The remainder of the paper is structured as follows. Section 2 presents the background of the coal mining industry in China. Section 2.3 introduces different datasets used in the empirical analysis. In Section 3, I develop a theoretical framework to formalize the intuition that power plants' buyer power downstream affects coal mines' tradeoff decisions upstream. Section 4 presents the empirical framework for identifying how buyer power affects the organization of coal mine production, and Section 5 provides the estimation results from the empirical models. Finally, Section 6 concludes.

2 The Coal Mining Industry in China and Data

2.1 Production and Safety

Production and Safety Generating Process. China accounts for the most annual coal production worldwide.[5] Substantial variations in geological conditions of coal seams in China lead to heterogeneous adoption of coal extraction methods and technology.[6] Depending on the thickness, dip angle, and other factors of the coal seams, Chinese coal mines can adopt more than 20 different underground coal mining methods for extraction (Peng, 2010).[7] Regardless of which specific technology is adopted, a typical coal extraction face consists of five main steps that cycle in sequence: coal cutting, loading, transporting, roof supporting, and goaf stowing, as seen in Figure 1.[8] See Appendix B for operation details about each step of the coal production process. The degree of mechanization and corresponding labor participation in different operations, therefore, matter for ultimate coal output, as well as production safety.

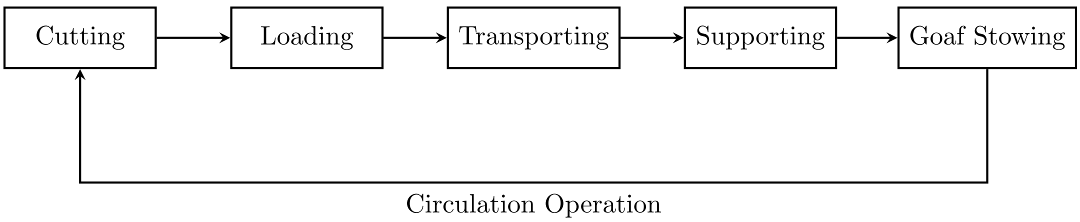


Figure 1: A Typical Coal Mining Process

On the other hand, coal mines also need to adopt auxiliary production equipment and relevant workers for operation and maintenance to support the regular coal mining process. The coal production process cannot be carried out smoothly without the normal operation of the auxiliary system. For example, an indispensable auxiliary system, among others, is mine ventilation. The benefit of installing the mine ventilation system is straightforward to ensure fresh air and, most importantly, guarantee the dilution and discharge of toxic and dangerous gases underground. Hence, these safety-enhancing auxiliary production systems and associated labor inputs, though they do not directly contribute to coal production, are

[5] China accounted for 31% and 51% of world coal production in 2002 and 2022, respectively. Data source: International Energy Agency (IEA).

[6] For example, the thickness of the coal seam matters a lot for adopting coal-cutting machines. Too narrow coal veins are unsuitable for introducing mechanical cutting techniques (Delabastita and Rubens, 2024).

[7] In general, these methods are all variant forms of the longwall extraction method but differ in specific types of specialized equipment required in different steps of the production process.

[8] In the context of coal mining, "face" refers to the area of a coal seam from which coal is being extracted. It is the working zone where miners and machinery cut, drill, blast, load, and transport coal.

essential to support ultimate coal production in the background.

Technological Change. The mechanization rate in China's coal mining industry was relatively low between 2000 and 2010. Specifically, mechanized coal mining accounted for less than 40% of the coal mines in China in 2000, while the state-owned key coal mines—the most advanced coal mines in China at the time—only had 75% mechanized coal extraction (NEA, 2001). *The 10th five-Year Plan for the Coal Industry* set mechanization targets in 2005 for large- and medium-scale coal mines as 90% and 60%, while small-scale coal mines were planned to "begin to take off". Ultimately, the national mechanization degree of coal mining and excavation in 2010 turned to be 65% and 52%, respectively (NEA, 2016). The reasons for the low mechanization rate in China were partially due to the difficult geological conditions for adopting mechanical extraction techniques (Peng, 2010). But a more straightforward reason, among others, would be the expensive fixed cost of mechanizing the whole coal production line (Wright, 2004). In 2000, most coal mines in China were commonly not profitable and financially constrained to do so (Wang, 2007; Shi, 2013; Guo et al., 2018).

Nevertheless, as coal mines gradually mechanized, more and more mines began to upgrade their technology from blasting mining to conventionally-mechanized mining and finally to fully-mechanized coal mining technology (Wright, 2012, 2022). See Appendix B for a detailed introduction to different coal mining technologies. Table 1 summarizes distinct input intensity characteristics concerning the mining process for different types of mining technology.

Table 1: Input Intensity Characteristics of Mining Technology

Mining Process	Type of Technology		
	Blasting	Conventionally-mechanized	Fully-mechanized
Cutting	L	L or K	K*
Loading	L	L or K	K*
Transporting	K	K	K*
Supporting	M and L	M and L	K*
Goaf Stowing	L	L	K*

Notes: I refer to Yan et al. (2009) for a comprehensive introduction and technical details of different coal mining technologies adopted in China. I highlight and summarize the differences in input intensity characteristics for different coal mining technologies. Herein, for the exposition, "L" refers to labor-intensive, "M" denotes material-intensive, "K" represents capital-intensive, while "K*" reflects even higher capital intensity than "K".

The distinct characteristics of input intensity across mining technologies, as shown in Table 1, provide a unique lens for understanding technological changes in the coal mining industry. Although I do not directly observe the specific technology information for coal mines, one can infer technological adoption from the input intensity.[9]

In summary, fully-mechanized coal mining technology is the most capital-intensive. Any technological changes from other mining methods to fully-mechanized coal mining require substantial capital inputs, hence a significant increase in capital-related input intensities, e.g., $\frac{K}{L}$ and $\frac{K}{M}$. In contrast, technological changes from blasting mining to conventionally-mechanized mining also incur a rise in $\frac{K}{L}$ and $\frac{K}{M}$, though the extent wouldn't be that substantial compared to the transition to fully-mechanized mining. While changes in $\frac{M}{L}$ from technological switching are unclear as blasting and conventionally-mechanized mining demand both labor and material much, and fully-mechanized coal mining requires little for both inputs, leading to ambiguous realization in $\frac{M}{L}$.

Coal Mining Accidents. The coal mining industry is considered one of the most hazardous industries worldwide due to its sordid working conditions and complicated production system (ILO, 2015).[10] China is the world's largest coal producer, with 97% of coal mines underground, experiencing the highest coal mining death rate (NMSA, 2016). For producing every one million tons of coal, five people died in China, compared with 0.5 people dead in India and 0.04 people dead in the United States during the same period from 1992 to 2001.[11] Difficult geological conditions played some roles in the high fatality mining rates—gassy mines are naturally pervasive in China—but the primary reasons should come from lacking safety equipment, e.g., gas-detection equipment, as other advanced coal-producing countries' experience has shown that, in principle, these problems have technological solutions (Wright, 2004; ILO, 2006; Wright, 2012; Murray and Silvestre, 2015).[12] Certainly, some of the disasters were caused by less-skilled workers' misconduct or mistaken operations (Liu et al., 2018, 2021). Random factors, such as earthquakes, also correlate with many coal mining accidents in China (Chen, 2020).[13]

[9] Unless one directly observes technological adoption information, e.g., Collard-Wexler and De Loecker (2015) for the US steel industry and Rubens (2022) for the US coal mining industry, most wildly-used firm-level datasets have little information about technology choices.

[10] ILO (2015) indicates that the mining industry, although only accounts for 1% of the global workforce, is responsible for 8% of fatal accidents at work. While coal mining fatalities in China comprise around 70% of the worldwide coal fatalities (Chen et al., 2013).

[11] The detailed data for cross-national comparisons for fatality rates come from Wright (2004). In its comprehensive description of China's coal mining industry, Tim Wright notes that the fatality rates in China's township and village mines at the end of the 20th century were similar to that of Belgium in the early 20th or Britain in the third quarter of the 19th.

[12] Among all coal mining accidents in China from 1994-1999, compared with respective proportions in Belgium (18% and 39% in 1881-1913) and the UK (13% and 51% in 1873-1932), gas explosions and roof falls accidents accounted for 49% and 29% of all accidents, respectively. Gas explosions accounted for an abnormally high proportion in China.

[13] Some political economy literature attributes China's high coal mining death rates to the difficulty in enforcing the law on coal mining safety due to the disincentive from the economic growth lost for implementing safety regulation (Wright, 2004; Li and Zhou, 2005), or collusion between regulators and firms (Jia and Nie, 2017). In contrast, Shi and Xi (2018) find that coal mine safety played a salient role in local officials' performance evaluations.

In addition to affecting worker safety, mining disasters tend to incur substantial economic costs for coal mines per se, including suspension of coal production, compensation of injured or dying workers, direct machinery capital loss due to the disaster, administrative penalties due to the safety regulation, etc. See Charles et al. (2022) for a detailed discussion about underlying accident costs.

Nevertheless, from 2000 to 2020, though the death rate remained high compared with other advanced coal-producing countries, China's coal mining safety improved significantly, from 5.71 deaths per million tons of coal production in 2000 to 0.06 deaths per million tons of coal production in 2020, around 100 times lower than two decades ago.[14] Figure 2 shows the evolution of coal production and mining death rates from 1995 to 2007. Increased coal production since the 21st century has witnessed a drastically declining death rate, defined by total casualties divided by total coal output.[15] Throughout my sample period, each province produced about 70 million tons of coal per year, for which coal mining accidents caused 243 fatalities.

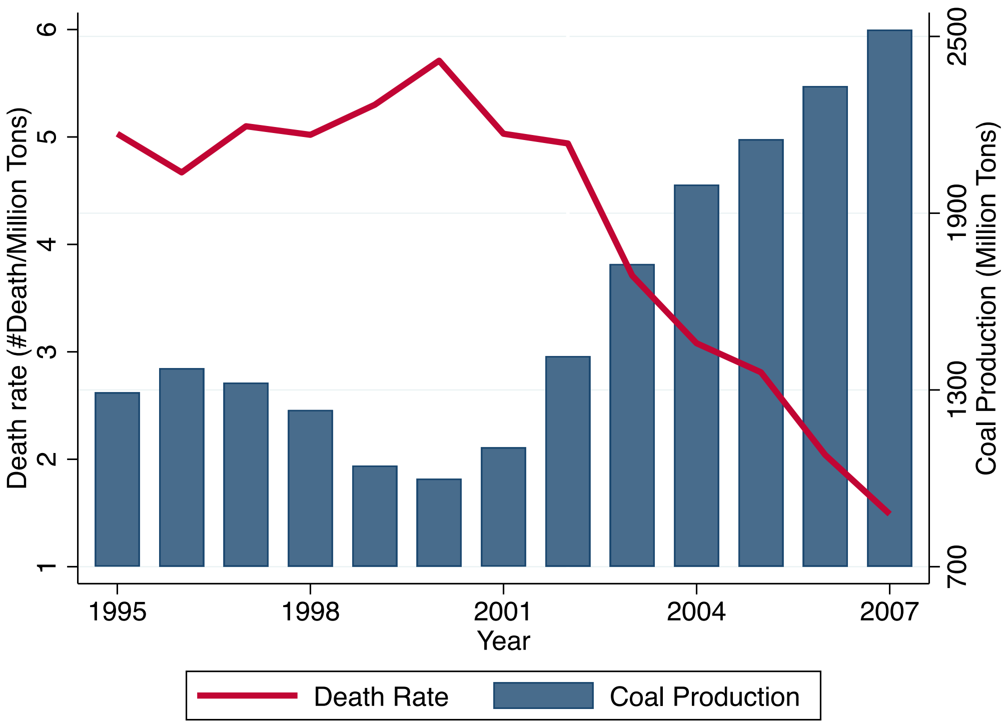


Figure 2: Coal Production and Safety in China: 1995-2007

Sources: Compilation of Coal Statistics of China's Coal Industry.

[14]The death rate data in China's coal mining industry in 2020 comes from Wright (2022).

[15]#death per million tons of coal is a widely adopted safety performance measure in the mining industry worldwide. Alternative death rate measures, e.g., #death per worker, indicate similar evolution patterns, as seen in Appendix A. However, since #death per million tons of coal straightforwardly correlates coal production with safety, it is more informative as it can be regarded as a rough measure of the production-safety tradeoff.

2.2 Demand and Source of Buyer Power

Demand. Coal resources are abundant in China but with substantial variation in geographical distribution. In 2000, the top 10 coal-producing provinces produced 75% of the national output, while the southeast coastal regions, the main coal consumers and GDP contributors, could only produce 18% of coal for their own use.[16] Hence, most coal was transported inter-provincially from north to south, from inland provinces to coastal provinces, by railway mostly.[17] At the time, the railway transportation sector was wholly state-owned with regulated uniform freight rates by product (MRC, 2000). The coal-fired power generation sector was the largest coal-buying sector (54% in 2000, even higher in the following years), consuming more coal than all other industries combined. To ensure the security of energy provision, the Ministry of Railways prioritized and guaranteed most of the coal transportation for coal-fired power plants (Wang, 2007).

However, the coal mining production was unconcentrated in China. Yang et al. (2017) shows that the top four and eight coal mining companies accounted for around 10% and 20% of the national market share in 2000, respectively, and even lower before. To increase their bargaining power over price, coal mining companies frequently attempted to form price cartels during the annual coal ordering fair organized by the government. More details about the annual coal ordering fair are in the following subsection.

Coal Pricing. To transition from the planned economy to the market economy, China adopted a unique dual-track pricing approach to market liberalization in various factor markets (Sicular, 1988; Li, 1999; Lau et al., 2000; Che and Facchini, 2007).[18] The market of thermal coal, also called "electricity coal," the type of coal that coal-fired power plants use as input to generate electricity, was one of the last markets still under the dual-track system during the sample period (1999-2006).[19]

To implement the dual-track thermal coal system, at the beginning of the year, an annual

[16] The top 10 coal-producing provinces were Shanxi, Shandong, Henan, Inner Mongolia, Hebei, Heilongjiang, Anhui, Liaoning, Sichuan, and Guizhou, and the southeast coastal regions consist of Jiangsu, Shanghai, Zhejiang, Fujian, Guangdong, and Hainan. The numbers are calculated using the year 2000 data from the *China Energy Statistical Yearbook*. Other numbers regarding provincial coal consumption and production in the same paragraph come from the same data source.

[17] In 2000, coal import only accounted for 0.2% of total coal consumption.

[18] Under a dual-track system, the planned economy and the market economy coexisted, while the market track progressively expanded with the gradual weakening of the planned track until it achieved market liberalization. Under the planned track, economic agents were supposed to buy or sell fixed quantities of goods at fixed plan prices. Without adjustment cost, the dual-track approach has been proved efficient Pareto-improving economic reform. See Lau et al. (2000) for a detailed introduction and proof.

[19] The coal market, including both thermal and non-thermal coal, was entirely under the market track in 1994 for a while, but the thermal coal market went back to embracing the planned track after price chaos soon happened between the coal mining and power generation industries—the thermal coal price went too high for power plants which threaten the power provision security (Guo et al., 2018). The dual-track thermal coal pricing system kept operating with few essential changes until it ultimately ended in 2012.

coal ordering fair was supposed to be organized, in which both the supply and demand sides of coal would participate, i.e., coal mining companies and coal-fired power plants (Wang, 2007). Coal prices were nominally decided via bilateral bargaining between both sides. However, the government always intervened in coal prices with a so-called "government guidance price," which was a linear per-unit coal tariff and more favorable to power plants when both sides had a hard time achieving agreements.[20]

Drivers of Buyer Power. Therefore, the ability of coal-fired power plants to exert buyer power over upstream coal mines primarily comes from the historical and institutional regime. Relative to the coal mining industry, rich policy and structure changes happened downstream of the power market, hence incurring substantial variation in buyer power.[21] Besides, the incentives mentioned above of coal mining companies to form price cartels during the coal ordering fair tend to reduce buyer power downstream. However, it was unclear whether the price cartels were effective and practical (Yang et al., 2012). Finally, coal safety regulations related to restricting production or enforcing closedowns could also affect upstream bargaining power and ultimate equilibrium downstream buyer power. Disentangling specific sources of buyer power is not the main interest of the paper, but rich variations in buyer power do help identify how buyer power affects variables of interest in the coal mining industry. I elaborate more about this in Section 4.1.

2.3 Data

I combine multiple data sources to study the comprehensive impact of buyer power on various outcomes. The datasets include production and cost information for coal mines and coal-fired power plants, as well as coal mine accidents and fatalities. The data sources are supplemented by several additional datasets covering separate information on state-owned key coal mines, railway coal transportation, and earthquakes, etc. More details on dataset summary statistics are provided in Appendix C.

Annual Survey of Industrial Firms Data. I obtain production and cost data for coal mines and power plants between 1999 and 2006 primarily from the *Annual Survey of In-*

[20]The guidance price for electricity coal has been announced as canceled, but due to the severe price conflicts between coal mines and power plants at the time, the practice of intervening by using a so-called "reference price," which is essentially the same thing but with a different name, has been adopted again.

[21]For example, after the 2002 electricity sector restructuring, the monopoly power generation company was split into five independent companies, and private power plants were also allowed to enter the power market. Each power company needed to bargain with upstream coal mines to make thermal coal contracts during the annual coal trade fair; before the 2002 electricity sector restructuring, it was the China Fuel Corporation as the sole representative for the whole power generation sector. Hence, the restructuring tends to reduce power plants' buyer power. I refer to Gao and Van Biesebroeck (2014) for a detailed description of the deregulation and vertical unbundling of the electricity sector restructuring in 2002.

dustrial firms (ASIF), which is collected by the National Bureau of Statistics (NBS).[22] This dataset contains detailed information on the firm's industry classification, total production (sale), intermediate expenditure (material), total employment and wage (labor), and real capital stock (capital) for all state-owned firms and non-state-owned firms with sales above 5 million RMB.[23] I retain all coal mines and coal-fired power plants under the 4-digit Chinese Industry Classification (CIC) codes of "0600" and "4011", which refer to the coal mining sector and coal-fired power generation sector.[24]

I complement the ASIF dataset with separate product-firm-month-level total production (quantity) data for the same sample period from the NBS, though I aggregate the production quantity information to the firm-year level for consistency. Due to the industry features of coal mining and coal-fired power generation, most firms produce homogenous coal/electricity as their single and ultimate output.

Coal Mine Accident Data. Coal mine accident records (2000-2006) are from the State Administration of Workplace Safety (SAWS). The SAWS requires local regulatory bureaus to report information on each coal mine disaster about the coal mine's name, accident type, occurrence time, location, number of fatalities, and direct causes (Shi and Xi, 2018; Chen, 2020). I merge the coal mine accident records with the coal mines in the ASIF dataset by fuzzy matching coal mines' names. See Appendix C for a credibility evaluation of the coal mine accident data.

Coal-fired Power Plants Production Data. I supplement coal-fired power plants' production and cost information in the ASIF data with a novel plant-level power generation dataset digitized by myself, which was originally compiled and examined by the *China Electricity Council* (CEC), for all power plants in China with a capacity above 6 MW that operated during 1999 and 2006.[25] This dataset provides comprehensive information on a power plant's name, nameplate capacity, operating hours, power generated (quantity), and coal used (quantity).[26] Using detailed quantity information of power plants for both input and output enables estimating the physical production function and, hence, a proper

[22]The unit of observation in the ASIF dataset is a firm (or an establishment), which is defined as a legal unit (faren danwei). Most of the firms in the ASIF dataset are single-plant firms, e.g., the share of single-plant firms was 96.6% in 2007 (Brandt et al., 2014). Hence, I refer to the consistent firm-level units when I mention "coal mines" or "power plants" in different contexts for the rest of the paper.

[23]See Brandt et al. (2012, 2014, 2017) for a comprehensive discussion about the data composition and variable construction of the ASIF dataset.

[24]To obtain a consistent classification across the sample period, I dealt with changes in the CIC codes by adjusting and unifying industry codes following Brandt et al. (2014).

[25]The threshold of 6 MW for being included in the dataset is fairly small for coal-fired power plants. For comparison, a new offshore wind turbine has a capacity of 8-12 MW nowadays (IRENA, 2024). Hence, the CEC dataset basically covers all coal-fired power plants in China during the sample period.

[26]The CEC dataset reports power plant-level coal heat rate annually, which I convert into coal used in quantity given the plant's power generation quantity.

measurement of buyer power afterward. More details will be introduced in Section 4.1.

Additional Datasets. I access various additional datasets over the same sample period to augment main data sources. First, I obtain bilateral railway transportation data for inter-provincial coal transportation from the *China Railway Yearbook*, provincial coal mining fatality and production information, and state-owned key coal mines' inter-provincial and sectoral sales from the *Compilation of Coal Statistics of China's Coal Industry*. Provincial transport characteristics come from the NBS. Second, I access the provincial coal demand information from the *China Energy Statistics Yearbook*. Third, I collect earthquake data from the China Earthquake Networks Center, which provides detailed information on the occurrence of longitude and latitude, time, magnitudes of the earthquake, and focal depth.

3 Theoretical Model

I develop a theoretical framework to formalize the intuition that firms' buyer power downstream affects firms' tradeoff decisions upstream. The specification is introduced in the context of upstream coal mines and downstream coal-fired power plants. Coal mines and coal-fired power plants contract on prices, taking into account expected coal mines' production and safety decisions.

3.1 Linking Bargaining to Buyer Power

A Nash Bargaining Example. I start by considering a simple complete-information bargaining case in which an upstream coal mine bargains with a downstream power plant over price, following a linear coal contract.[27] The upstream coal mine bargains with the downstream power plant over the coal price, P, given the expected amount of coal Q:

$$\max_{P} \left[PQ - C(Q)\right]^{1-b} \times \left[R(Q) - PQ\right]^{b}, \tag{1}$$

where the parameters b and $1-b$ indicate the bargaining power of the coal-fired power plant and the coal mine, respectively. $C(Q)$ is the total variable cost of the coal mine for producing Q unit of coal, which I will elaborate more about its cost function in Section 3.2. $R(Q)$ is the total revenue of the power plant for generating $\widetilde{Q}$ unit of electricity.[28]

[27] Linear contracts are commonly adopted for commodity trading, such as coal (Joskow, 1985, 1988).

[28] Assuming the expected dispatched order the power plant would be commanded is $\widetilde{Q}$, and the power plant's heat rate of coal is η, and hence $Q := \widetilde{Q}/\eta$ is the corresponding determined amount of coal needed as input for power generation.

The bargaining solution of price is characterized by the first-order condition of Eq.(1):

$$P = b \times \frac{C(Q)}{Q} + (1-b) \times \frac{R(Q)}{Q}, \tag{2}$$

indicating that, given Q, the negotiated coal price is a weighted average of the average variable cost of the coal mine ($\frac{C(Q)}{Q}$) and the average revenue of the power plant ($\frac{R(Q)}{Q}$), using the bargaining power of both sides as the weight.

Bargaining and Buyer Power. As noted in Section 2.2, the negotiation procedure between the coal mine and the power plant in practice could be distorted and vague. Both bargainers are incentivized to hide private information to ensure themselves a better bargaining position. Bargaining models under incomplete information in the case would be more realistic, in which each bargainer is uncertain about her opponent's information, evaluating the latent values that her adversary might hold under a subjective probability distribution (Chatterjee and Samuelson, 1983; Loertscher and Marx, 2021, 2022). However, parametric assumptions of the subjective probability about private valuations need discretion in practice; even though imposing given parametric assumptions, two-sided bargaining games with two-sided incomplete information likely have no known analytical solutions (Larsen and Zhang, 2018, 2021).[29]

Hence, I do not take a stance on building a specific bargaining protocol to model the underlying bilateral bargaining structure. Instead, I model the underlying bilateral bargaining structure and capture equilibrium bargaining outcomes in a reduced-form way. Still, suppose the upstream coal mine and downstream power plant bargain over price under certain bargaining rules given the quantity of coal. I capture the outcome of the bilateral bargaining game by an input price schedule from the downstream power plant's perspective, which maps the power plant's demand for coal input, Q, to the negotiated coal price, P:

$$P = P(Q; \mathbf{A}, \mathbf{b}), \tag{3}$$

where $\mathbf{A}$ is a vector governing supply and demand conditions between the buyer-seller network, and $\mathbf{b}$ is a vector capturing exogenous buyer power-related primitives. One can find that Eq.(2) nests well into Eq.(3), where $\mathbf{A}$ absorbs $C(\cdot)$ and $R(\cdot)$, and $\mathbf{b}$ incorporates b, respectively.

From the coal mine's perspective, given the quantity of coal, Q, optimality conditions induce it to minimize its cost to obtain optimal payoffs. I will elaborate more on this in the

[29] Larsen and Zhang (2018, 2021) propose an approach to estimate players' valuations and expected gains from trade under asymmetric information in the wholesale used-car market without imposing a particular extensive form.

following section.

3.2 Joint Production of Coal and Safety

Production. I now discuss the coal mine's cost function and corresponding input decisions for producing the given unit of coal Q. To explicitly model the coal mine's production and safety decisions, I assume the coal mine adopts K^q and K^s, for production capital and safety capital, and L^q and L^s, for production and safety workers, respectively, to produce outputs of coal, $Q = Q(K^q, L^q, \Omega^q)$, and safety, $S = S(K^s, L^s, \Omega^s)$, where Ω^v ($v \in \{q, s\}$) are exogenous productivity for different production processes respectively. For the exposition, I define $A(\cdot)$ as the measure of accidents negatively correlate with $S(\cdot)$. Without loss of generality, let $A(\cdot) = -S(\cdot)$.[30]

Assume that $Q(\cdot)$ and $A(\cdot)$ are continuous and twice differentiable w.r.t to their arguments, and the coal mine determines its input choices by minimizing the variable cost, $C(Q)$. Therefore, the associated Lagrangian function is

$$\mathcal{L}(K^q, K^s, L^q, L^s, \lambda) = \kappa \times A(\cdot) + w \times (L^q + L^s) + r(\mathbf{b}) \times (K^q + K^s) + \lambda \left(Q - Q(\cdot)\right), \quad (4)$$

where λ is the marginal cost at the given level of Q, and κ, w, and $r(\mathbf{b})$ denote corresponding prices for accident (or negative safety), labor, and capital stock, respectively. I model an increase in $\mathbf{b}$ leads to higher $r(\mathbf{b})$, i.e., $\frac{\partial r(\mathbf{b})}{\partial \mathbf{b}} > 0$. The economic intuition behind this is that financial liquidity, the strength of the balance sheet, asset structure, and/or creditworthiness are all negatively correlated with downstream buyer power (Charles et al., 2022). I directly verify this assumption in Section 5.3.[31]

The first-order conditions for any input is $\frac{\partial \mathcal{L}}{\partial V} = 0$ with $V \in \{K^q, K^s, L^q, L^s\}$. By rearranging the equilibrium terms, I have

$$r(\mathbf{b}) = \lambda \times \frac{\partial Q(\cdot)}{\partial K^q} = -\kappa \times \frac{\partial A(\cdot)}{\partial K^s}, \quad (5)$$

$$w = \lambda \times \frac{\partial Q(\cdot)}{\partial L^q} = -\kappa \times \frac{\partial A(\cdot)}{\partial L^s}. \quad (6)$$

The optimizing mine employs different types of capital and labor until the benefit derived from the last unit spent equals the corresponding increase in cost. The relative prices of

[30] Clearly, $A(\cdot)$ is an isomorphic function of $S(\cdot)$ that each of them can be reserved by an inverse mapping from each other.

[31] One can further model wage as a function of accident risk, as in Sider (1983), to capture risk premium effects. Nevertheless, I follow Gowrisankaran et al. (2015) to account for the direct cost of the accident but not only for the increased wage costs from accident risk. In Section 4.2, I utilize high-order polynomials to control for risk premiums in estimation.

inputs determine the ratio of different inputs adopted for the same production process. Buyer power downstream heterogeneously changes the input prices for different production processes upstream, leading to distinct impacts on input adoption. I discuss this more in detail in Section 3.3.

3.3 The Effects of Buyer Power

I now examine the effects of change in downstream power plants' buyer power, i.e., markdown, on upstream coal mines' production and safety decisions.

Assumption. *$Q(\cdot)$ and $A(\cdot)$ are continuous and twice differentiable. For $X^q \in \{K^q, L^q\}$ and $X^s \in \{K^s, L^s\}$, I assume $\frac{\partial Q}{\partial X^q} > 0$, $\frac{\partial\left(\frac{\partial Q}{\partial X^q}\right)}{\partial X^q} < 0$, $\frac{\partial A}{\partial X^s} < 0$, and $\frac{\partial\left(\frac{\partial A}{\partial X^s}\right)}{\partial X^s} > 0$.*

This assumption is straightforwardly due to explicit economic intuitions in the coal mining industry, as seen in Section 2.1, implying more safety inputs generate more safety output and, hence, fewer accidents. In contrast, more production inputs lead to more coal output. In the coal mining industry, producing more coal induces coal mines to work and operate at deeper underground or thicker coal layers, leading to diminishing marginal product of different inputs for producing both coal and safety (or negative accidents), which boils down to the second-order partial derivatives assumptions above.

Proposition 1. *The buyer power of downstream power plants induces upstream coal mines to adopt less capital, ultimately leading to lower capital-to-labor ratios for both production processes.*

Proof: See Appendix D.

The intuition of the impact of downstream buyer power on upstream production decisions in **Proposition 1** is that buyer power affects upstream input prices heterogeneously. An increased cost of capital induces upstream coal mines to adopt less capital input for both safety and production but does not directly affect labor input. The relative capital-to-labor ratio upstream, determined by the relative input price of capital and labor, holding others fixed, unambiguously decreases due to the increased capital input price when buyer power downstream increases.

Corollary 1. *The buyer power of downstream power plants can induce more coal mining accidents upstream via increased cost of capital adoption.*

Proof: See Appendix D.

The intuition of **Corollary 1** is self-contained by **Proposition 1** in the sense that less safety capital, resulting from increased capital costs, induces more accidents and more deaths, holding others fixed.

Overall, the theoretical model suggests that buyer power downstream from power plants can affect coal mines' production and safety decisions and, hence, safety outcomes. Nevertheless, the model is silent about efficiency and technological changes in response to buyer power variations. Instead of pre-imposing assumptions about the directions of technological changes and the endogenous evolution process of productivity, I leave these margins to be examined directly in empirics in Section 4.

4 Empirical Framework and Estimation

This section presents the empirical framework for identifying how buyer power affects the organization of coal mine production. I first employ a production approach combined with unique power plant-level generation data to consistently measure the buyer power of power plants in Section 4.1. Second, I develop a multiple-output production framework with endogenous safety choice and labor-augmenting technology to identify the production-safety tradeoff in Section 4.2. Finally, in Section 4.3, I elaborate on the empirical strategy, relying on an instrument variable methodology.

4.1 Buyer Power Measure

Production Approach. As illustrated in Section 2.2, coal mines and power plants made contracts under incomplete information, opaque rules, and arbitrary intervention. Therefore, specifying any behavioral assumption on competition and conduct between upstream and downstream tends to be restrictive in measuring buyer power consistently. Instead, I employ the production approach to measure the buyer power of power plants without assuming any model of competition and conduct (De Loecker and Warzynski, 2012; Rubens, 2023; De Loecker and Scott, 2022).

Specifically, consider a cost-minimizing power plant f generates Q_{ft} units of electricity using variable inputs such as coal, M_{ft}, labor, L_{ft}, and a fixed input of capital, K_{ft}. I specify a Leontief production technology for electricity generation:

$$Q_{ft} = \min\left[\varrho_{ft} M_{ft}, F(L_{ft}, K_{ft}; \boldsymbol{\beta}) \exp\left(\omega_{ft}\right)\right], \tag{7}$$

where ϱ_{ft} is the heat rate depicting the inverse of a fixed per-unit materials input requirement, $F(\cdot)$ is a common technology across labor and capital with parameters $\boldsymbol{\beta}$, and ω_{ft} is the unobserved productivity term.[32,33] Leontief (i.e. fixed-proportion) technology governs an essential feature of the power generation process, where coal cannot be simply substituted with more labor or capital.

I assume that coal and labor are static inputs that can be adjusted at time t, while capital is dynamic and can only be changed through investment at time $t-1$. Given coal and labor are not substitutable, determining labor immediately implies choosing a quantity of coal. Due to the institutional background shown in Section 2.2, I allow power plants to have input market power over coal, which means once coal's input quantity choice is made, power plants set equilibrium coal input prices if the input supply is upward-sloping (Rubens, 2023).[34] I therefore consider the associated Lagrangian function for power plant f:

$$\mathcal{L}_{ft} = P_{ft}^{M}\left(M_{ft}\right) M_{ft} + P_{ft}^{L} L_{ft} + P_{ft}^{K} K_{ft} + \lambda_{ft}\left(\overline{Q}_{ft} - Q_{ft}\left(M_{ft}, L_{ft}, K_{ft}, \omega_{ft}\right)\right), \tag{8}$$

where P_{ft}^{X} with $X = \{M, L, K\}$ are input price for different inputs, and λ_{ft} is the marginal cost. Define power plant f's output price, i.e., on-grid power price, is P_{ft}. Taking the first-order condition with respect to coal, rearranging terms, and defining the coal markdown as $\psi_{ft}^{M} \equiv \frac{\partial P_{ft}^{M}}{\partial M_{ft}} \frac{M_{ft}}{P_{ft}^{M}} + 1$ and the markup as $\mu_{ft} \equiv \frac{P_{ft}}{\lambda_{ft}}$, the markdown can be expressed as:[35]

$$\psi_{ft}^{M} = \frac{1}{\alpha_{ft}^{M}}\left(\frac{1}{\mu_{ft}} - \frac{\alpha_{ft}^{L}}{\beta^{L}}\right), \tag{9}$$

where $\alpha_{ft}^{M} = \frac{P_{ft}^{M} M_{ft}}{P_{ft} Q_{ft}}$ and $\alpha_{ft}^{L} = \frac{P_{ft}^{L} L_{ft}}{P_{ft} Q_{ft}}$ are the revenue shares of coal and labor, and $\beta^{L} = \frac{\partial Q_{ft}}{\partial L_{ft}} \frac{L_{ft}}{Q_{ft}}$ is the output elasticity of labor. The coal markdown, ψ_{ft}^{M}, measures the extent of buyer power over the coal input market that plant f exerts at time t.

Hence, to calculate power plant-level coal markdowns, I need data on coal and labor revenue shares (α_{ft}^{M} and α_{ft}^{L}), plant-level markups (μ_{ft}), and the estimate of labor output elasticity (β^{L}), where the first is readily observed in the data.[36] I set the markup to be

[32]The Leontief technology specification and cost minimization assumption for coal-fired power plants are well fitted to the practical power generation and dispatching feature. Similar specifications are well adopted in the literature; see Fabrizio et al. (2007), Gao and Van Biesebroeck (2014), Atkinson and Luo (2024), Demirer and Karaduman (2024), among others.

[33]In estimation, an unobserved term of measurement error or unanticipated shocks in output will also be specified.

[34]In this way, P_{ft}^{M} can be expressed as a function of M_{ft}, i.e., $P_{ft}^{M}\left(M_{ft}\right)$.

[35]Equivalently, I can express the markup as $\mu_{ft} = \left(\alpha_{ft}^{M}\psi_{ft}^{M} + \frac{\alpha_{ft}^{L}}{\beta^{L}}\right)^{-1}$.

[36]To be precise, considering the measurement error or unanticipated shocks in output, ϵ_{ft}, what I observe in the data is $\widetilde{Q}_{ft} := Q_{ft}\exp(\epsilon_{ft})$, which leads to wrong measure of revenue shares. Fortunately, my following production function estimation procedure will provide an estimate for ϵ_{ft}, enabling me to correct coal and labor revenue shares. Therefore, the corrected revenue shares should be $\hat{\alpha}_{ft}^{M} = \frac{P_{ft}^{M} M_{ft}}{P_{ft}\widetilde{Q}_{ft}/\exp(\epsilon_{ft})}$ and $\hat{\alpha}_{ft}^{L} = \frac{P_{ft}^{L} L_{ft}}{P_{ft}\widetilde{Q}_{ft}/\exp(\epsilon_{ft})}$. I use the corrected input revenue shares to compute coal markdowns. De Loecker

constant as in Rubens (2023), utilizing the institutional feature of the Chinese power market regulation during the sample period, by which the government sets on-grid power prices plant by plant based on their cost information (Lam, 2004; Ma, 2011).[37] I let $\mu_{ft} = 1$ in the main text.[38] Finally, the output elasticity of labor requires estimating the production function.

After I obtain power plant-level time-varying coal markdown measures (ψ_{ft}^M), I first aggregate it to provincial-level using the plant's market share as weight: $\psi_{dt} := \sum_{f \in \mathcal{D}_d} s_{ft}\psi_{ft}^M$, where $\mathcal{D}_d$ is the set of power plants in province d and s_{ft} is power plant f's market share in the same province. Then, I construct the provincial buyer power exposure ($\widetilde{\psi_{ot}}$) using interprovincial coal railway freight volumes as weight. Specifically, $\widetilde{\psi_{ot}} = \sum_{d \in \Theta} s_{od}\psi_{dt}$.

Production Function Estimation. Now I show how to estimate the production function of power plants to obtain the output elasticity of labor (β^L) in a two-stage estimation procedure.

Taking the logarithm of the production function, Eq.(7), where lowercases refer to the corresponding logarithmic variables, and allowing for an unobserved term ϵ_{ft}, which captures measurement error in output or unanticipated shocks to production,[39] the estimating production function turns out to be:

$$q_{ft} = f(l_{ft}, k_{ft}; \boldsymbol{\beta}) + \omega_{ft} + \epsilon_{ft}. \tag{10}$$

As coal is specified to be non-substitutable, it does not enter the above-estimating equation directly; however, the Leontief specification implies a fixed-proportion rule in equilibrium between coal and other inputs, i.e., $\varrho_{ft}M_{ft} = F(L_{ft}, K_{ft}; \boldsymbol{\beta}) \exp(\omega_{ft})$, enabling a simple control for unobserved productivity: $\omega_{ft} = \ln(\varrho_{ft}) + m_{ft} - f(l_{ft}, k_{ft}; \boldsymbol{\beta})$, as opposed to relying on inverting optimal input demand function in classicial control function literature.[40] Substituting the control function into Eq.(10) helps to extract the measurement error or unanticipated shocks in output, ϵ_{ft}, and purge it out in the first stage of the estimation

and Warzynski (2012) indicates that eliminating measurement error when computing markups (markdowns in my case) is crucial for measuring the wedges consistently. Treuren (2022) also highlights similar insights when measuring input wedges.

[37]It is called "cost-plus" or "rate-of-return" regulation, which was common in the electricity generation market, e.g., Cicala (2015) and Gowrisankaran et al. (2024), among others.

[38]See Appendix A for alternative calibrations.

[39]More precisely, I observe logged output $q_{ft} = \ln(Q_{ft}) + \epsilon_{ft}$ in the data, where ϵ_{ft} are i.i.d. shocks unobserved to firms when making their optimal input decisions.

[40]Similar technological control functions for unobserved productivity have been used in the beer (De Loecker and Scott, 2022) and car industries (Hahn, 2024). This control helps depart from the standard working assumptions of perfectly competitive input and output markets (Olley and Pakes, 1996; Levinsohn and Petrin, 2003; Ackerberg et al., 2015), allowing *any* form of competition in either market, which is essential to my setup as I'm interested in consistently measuring the input market power of power plants.

procedure.[41]

Specifically, I first project the quantity of electricity output on the quantity of coal input and a plant-time specific heat rate factor, ϱ_{ft}. Following De Loecker and Scott (2022), I approximate ϱ_{ft} by a nonparametric function of $(m_{ft}, l_{ft}, k_{ft}, w_{ft}, p^M_{ft}, \boldsymbol{D}_{ft})$, where w_{ft} and p^M_{ft} are plant-level wage and coal input price, and $\boldsymbol{D}_{ft}$ captures ownership dummies, year dummies, region dummies, and interactions.[42,43] The first stage specification, therefore, is:

$$q_{ft} = \varrho_t\big(m_{ft}, l_{ft}, k_{ft}, w_{ft}, p^M_{ft}, \boldsymbol{D}_{ft}\big) + \epsilon_{ft}, \tag{11}$$

where $\varrho_t(\cdot)$ is a nonparametric function. I run Eq.(11) to obtain estimates of expected output $(\hat{\varrho}_{ft})$ and an estimate for ϵ_{ft}.

In the second stage, I can compute productivity using $\omega_{ft}(\boldsymbol{\beta}) = \hat{\varrho}_{ft} - f(l_{ft}, k_{ft}; \boldsymbol{\beta})$ given any value of $\boldsymbol{\beta}$. I adopt a Cobb-Douglas specification for $f(\cdot)$ in the main text, which leads to $f(l_{ft}, k_{ft}; \boldsymbol{\beta}) = \beta^L l_{ft} + \beta^K k_{ft}$, where $\boldsymbol{\beta} = (\beta^L, \beta^K)$. To identify and obtain estimates of production function coefficients, I consider a standard first-order Markov process, $g(\cdot)$, to depict the law of motion for productivity:

$$\omega_{ft} = g(\omega_{ft-1}) + \xi_{ft}, \tag{12}$$

where ξ_{ft} denotes innovation in the productivity process. Given $\boldsymbol{\beta}$, therefore, I can recover the innovation to productivity, $\xi_{ft}(\boldsymbol{\beta})$, by nonparametrically regressing $\omega_{ft}(\boldsymbol{\beta})$ on $\omega_{ft-1}(\boldsymbol{\beta})$.

I can now use the innovation, $\xi_{ft}(\boldsymbol{\beta})$, to construct moment conditions and estimate the production function parameters $\boldsymbol{\beta}$ using standard generalized method of moments (GMM) techniques:

$$\mathbb{E}\left(\xi_{it}(\boldsymbol{\beta})\begin{bmatrix} l_{it-1} \\ k_{it} \end{bmatrix}\right) = 0. \tag{13}$$

The instruments' exclusion restrictions come from the timing assumptions about labor and capital inputs I set before, by which productivity innovation is orthogonal to the current capital and lagged labor.

[41] After the substitution, the estimating production function turns out to be: $q_{ft} = \ln(\varrho_{ft}) + m_{ft} + \epsilon_{ft}$.

[42] The economic intuitions of the control function for heat rate are straightforward since many factors can affect a plant's heat rate: different amount and types of inputs (m_{ft}, l_{ft}, and k_{ft}) due to the start-up costs and economies of scale, quality or proficiency of workers (w_{ft}), quality differentiation of coal input (p^M_{ft} and dummy variables in $\boldsymbol{D}_{ft}$), among others (interactions in $\boldsymbol{D}_{ft}$).

[43] Specifically, I use a third-order polynomial in the inputs, and interactions between the third-order input terms and wage or coal input price. The dummy variables do not interact.

4.2 A Model of the Production-Safety Tradeoff in Coal Mining

The Production Technology. Unlike in the theoretical model of the coal mine's production-safety tradeoff as in Section 3.2, where I model the coal mine adopts different production and safety inputs distinguishably, such output-specific input allocation schedule is rarely observed in commonly available firm-level datasets in empirical studies, as noted by De Loecker and Syverson (2021).[44]

Specifically, I do not have data on how coal mines allocate safety- and production-related workers separately or how much they mechanize different parts of coal mining operations, e.g., adopting coal-cutting machines for coal extraction or mine ventilation systems for safety. Employing a widely used single-output production function for coal and ignoring safety output would undoubtedly incur bias because not whole input bundles are used for producing coal directly, while some inputs are for supporting the production in the background, as demonstrated in Section 2.1 about the coal mining feature.

Instead of imputing the output-specific input allocation scheme for coal and safety, I consider a transformation function, $T(\cdot)$, governing how the coal mine j assigns its productive capacity between the two targeted outputs, i.e., the quantity of coal produced, Q_{jt}, and the safety level during the coal production process, S_{jt}, in the spirit of Grieco and McDevitt (2017):[45]

$$T(Q_{jt}, S_{jt}) = F\left(K_{jt}, L_{jt}, M_{jt}, \Omega_{jt}\right), \tag{14}$$

where production function $F(\cdot)$ describes how the firm-level capital, K_{jt}, labor, L_{jt}, material, M_{jt}, and unobserved productivity, Ω_{jt}, determine the overall productive capacity.

An explicit advantage of adopting the transformation function approach is that I can allow joint or public inputs; that is, some inputs may be used in both coal and safety generating processes simultaneously, which is essential for depicting coal mine production behavior accurately, as described and modeled in Section 2.1.[46]

Model Setup. Specifically, I adopt a parsimonious Cobb-Douglas transformation function

[44] This echoes with a strand of literature on multi-product production function estimation (De Loecker et al., 2016; Orr, 2022; Valmari, 2023). As the allocation of firm-level inputs across products is not available, existing approaches in literature adopt different assumptions to impute the input allocation information. I refer to De Loecker and Syverson (2021) for a detailed review.

[45] As coal is a relatively homogenous product, considering a transformation function in my case is free from the production differentiation concern drawn by De Loecker and Syverson (2021).

[46] Existing multi-product production function estimation literature on imputing input allocation information across products usually rules out public or joint inputs and physical synergies across products. See De Loecker and Syverson (2021) and Orr (2022) for a detailed discussion and comparison.

(in logs) as a benchmark specification[47]

$$\ln\left(T(Q_{jt}, S_{jt})\right) = \tilde{q}_{jt} + \alpha_s \tilde{s}_{jt} \tag{15}$$

and a constant elasticity of substitution (CES) production function (in logs) with labor-augmenting technology following Doraszelski and Jaumandreu (2018) as[48]

$$\ln\left(F\left(K_{jt}, L_{jt}, M_{jt}, \Omega_{jt}\right)\right) = \frac{\sigma\kappa}{\sigma - 1} \ln\left(\beta_K K_{jt}^{-\frac{1-\sigma}{\sigma}} + [\exp(\omega_{Ljt})L_{jt}]^{-\frac{1-\sigma}{\sigma}} + \beta_M M_{jt}^{-\frac{1-\sigma}{\sigma}}\right) + \omega_{Hjt}. \tag{16}$$

Therein, $\tilde{q}_{jt}$ and $\tilde{s}_{jt}$ are logged expected coal output and safety level, α_s represents the key technological parameter governing the production-safety tradeoff, describing the coal mine's incentive to increase coal output by forgoing safety, κ is the returns to scale, σ denotes the elasticity of substitution, and β_K and β_M are distributional parameters. I allow for both Hicks-neutral productivity, ω_{Hjt}, as in the literature, and labor-augmenting productivity, ω_{Ljt}, due to the fact of coal mines' increased capital-to-labor ratio shown in Figure A2 in Appendix A, for the coal mine j at time t.[49]

To estimate the production frontier empirically, I substitute Eq.(15) and Eq.(16) into Eq.(14), and allow for measurement error (or unanticipated shocks) in both quantity and safety as I only observe realized quantity, $q_{jt} := \tilde{q}_{jt} + \epsilon_{jt}^q$, and realized safety, $s_{jt} := \tilde{s}_{jt} + \epsilon_{jt}^s$, in the data, where ϵ_{jt}^q and ϵ_{jt}^s are i.i.d. shocks, rearranging safety-related terms to the right-hand side, the estimating equation then is:

$$q_{jt} = -\alpha_s s_{jt} + \frac{\sigma\kappa}{\sigma - 1} \ln\left(\beta_K K_{jt}^{-\frac{1-\sigma}{\sigma}} + [\exp(\omega_{Ljt})L_{jt}]^{-\frac{1-\sigma}{\sigma}} + \beta_M M_{jt}^{-\frac{1-\sigma}{\sigma}}\right) + \omega_{Hjt} + \alpha_s \epsilon_{jt}^s + \epsilon_{jt}^q. \tag{17}$$

However, estimating Eq.(17) directly with the coal mine's input and output data using GMM techniques or other nonlinear methods leads to inconsistent estimates. There are two unobserved productivity terms, ω_{Hjt} and ω_{Ljt}, and measurement error of ϵ_{jt}^s is also correlated with s_{jt}; all of them induce endogeneity issues. In addition, the identification of α_s could also be problematic if there is no independent variation of $\tilde{s}_{jt}$ from other variables in $F\left(K_{jt}, L_{jt}, M_{jt}, \Omega_{jt}\right)$ as the coal mine makes safety choice with the initial states of

[47] In Section 5.4, I further allow for the heterogeneity in the production-safety tradeoff, by which the slope of the production-safety frontier could vary across mines depending on their technology or capital-to-labor ratio.

[48] $F\left(K_{jt}, L_{jt}, M_{jt}, \Omega_{jt}\right) = \{\beta_K K_{jt}^{-\frac{1-\sigma}{\sigma}} + [\exp(\omega_{Ljt})L_{jt}]^{-\frac{1-\sigma}{\sigma}} + \beta_M M_{jt}^{-\frac{1-\sigma}{\sigma}}\}^{\frac{\sigma\kappa}{\sigma-1}} \exp(\omega_{Hjt})$ in level specification. The CES production function nests the Leontief ($\sigma \to 0$), Cobb-Douglas ($\sigma = 1$), and linear ($\sigma \to \infty$) production functions as special cases depending on the elasticity of substitution, σ.

[49] In theory, I can further allow for output-specific Hicks-neutral productivity following Dhyne et al. (2022), which proposes a general two-equation system for estimating the transformation function of multi-product production. As understanding differentiated output-specific efficiency is not the first-order interest of the paper, I leave the extension to future research.

productivity being aware.

I address all these issues in detail as follows.

Timing and Decisions. I first illustrate the timing and decision-making for the coal mine, which is necessary to be exploited to infer unobserved productivity differences, ω_{Hjt} and ω_{Ljt}, and provide independent variation of $\tilde{s}_{jt}$. Figure 3 shows the timing of decisions and how states evolve.

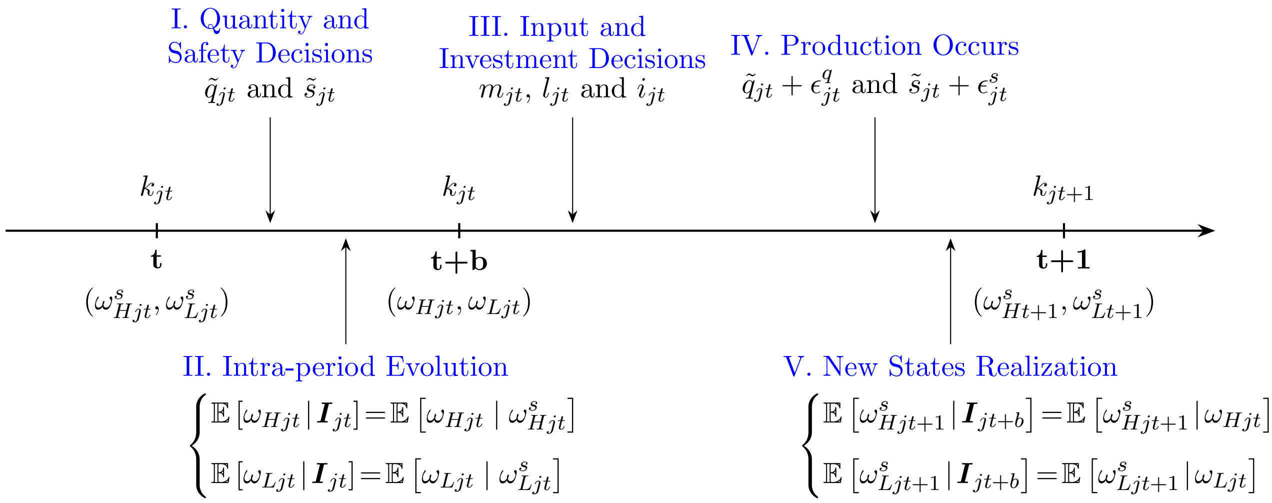


Figure 3: Timing of Decisions

At the beginning of time **t**, the initial states, $(k_{jt}, \omega^s_{Hjt}, \omega^s_{Ljt})$, are known to the coal mine j, before any input decision is made. I denote the initial productivities, ω^s_{Hjt} and ω^s_{Ljt}, with a superscript s as it represents the productivities the mine observes when making the safety decision (s for safety). Based on the initial states, the coal mine sets targeted output level for coal and safety, $(\tilde{q}_{jt}, \tilde{s}_{jt})$,[50] to optimize its objective function as in Section 3.2, while capital (k_{jt}), decided at time **t-1**, remains fixed. One can think of this stage as the coal mine making contracts with power plants based on its expected productivity, while actual productivity will be revealed before production.[51]

During period **t+b**, which is in-between time **t** and **t+1**, the mine updates its productivity beliefs to the actual productivities, ω_{Hjt} and ω_{Ljt}, following an exogenous Markov process, $\mathbb{E}\left[\omega_{Hjt} \,|\, \boldsymbol{I}_{jt}\right] = \mathbb{E}\left[\omega_{Hjt} \mid \omega^s_{Hjt}\right]$ and $\mathbb{E}\left[\omega_{Ljt} \,|\, \boldsymbol{I}_{jt}\right] = \mathbb{E}\left[\omega_{Ljt} \mid \omega^s_{Ljt}\right]$, where $\boldsymbol{I}_{jt}$ represents the mine's information set at the beginning of time **t**. With this updated information, the

[50]The choices of outputs are not modeled explicitly as I take them as given, which is innocuous to following empirical estimation. Though, the intuitions of output choices embedded are consistent with what I illustrated in Section 3.2, where taking coal output as given, and safety choice can be captured by a reduced-form function as $\tilde{s}_{jt} = s_t(\tilde{q}_{jt}, k_{jt}, \omega^s_{Hjt}, \omega^s_{Ljt})$.

[51]The specification aligns with China's coal mining industry, as coal mines and power plants sign contracts during the coal trade fair at the beginning of the year while producing and delivering coal in batches afterward. During this, productivity evolves over time.

mine decides labor (l_{jt}) and material (m_{jt}) inputs, as well as investment (i_{jt}), with new capital (k_{jt+1}) becoming available at the start of the next period, **t+1**. Production occurs afterward, which reveals the actual production outcomes, (q_{jt}, s_{jt}), where $q_{jt} = \tilde{q}_{jt} + \epsilon^q_{jt}$ and $s_{jt} = \tilde{s}_{jt} + \epsilon^s_{jt}$.

Finally, assuming productivity expectations follow a first-order Markov process, the expected productivity difference for the next period are $\mathbb{E}\left[\omega^s_{Hjt+1} | \boldsymbol{I}_{jt+b}\right] = \mathbb{E}\left[\omega^s_{Hjt+1} | \omega_{Hjt}\right]$ and $\mathbb{E}\left[\omega^s_{Ljt+1} | \boldsymbol{I}_{jt+b}\right] = \mathbb{E}\left[\omega^s_{Ljt+1} | \omega_{Ljt}\right]$, where $\boldsymbol{I}_{jt+b}$ captures the mine's information set at the intra-period **t+b**.

Unobserved Productivity Differences. There exists two unobserved productivity differences, ω_{Hjt} and ω_{Ljt}. I first discuss how to control the unobserved labor-augmenting productivity, ω_{Ljt}, by exploiting the coal mine's first-order conditions for labor and material.[52] Let $Y_{jt} := Q_{jt} S^{\alpha_s}_{jt}$ denote the composite outputs for coal and safety. Suppose the coal mine j minimizes variable costs to generate the given amount of outputs, $\bar{Y}_{jt}$. The mine's cost minimization problem is:

$$\min_{L_{jt}, M_{jt}} P^L_{jt} L_{jt} + P^M_{jt} M_{jt} \quad \text{s.t.} \quad \{\beta_K K_{jt}^{-\frac{1-\sigma}{\sigma}} + [\exp(\omega_{Ljt}) L_{jt}]^{-\frac{1-\sigma}{\sigma}} + \beta_M M_{jt}^{-\frac{1-\sigma}{\sigma}}\}^{\frac{\sigma\kappa}{\sigma-1}} \exp(\omega_{Hjt}) \geqslant \bar{Y}_{jt},$$

where P^L_{jt} and P^M_{jt} are input prices for labor and material. The ratio of associated first-order conditions with respect to labor and material gives $\frac{\partial Y_{jt}(\cdot)}{\partial L_{jt}} / \frac{\partial Y_{jt}(\cdot)}{\partial M_{jt}} = \frac{P^L_{jt}}{P^M_{jt}}$, rearranging terms, finally leads to an explicit expression for the labor-augmenting productivity with observed variables and estimable parameters:[53]

$$[\exp(\omega_{Ljt})]^{-\frac{1-\sigma}{\sigma}} = \beta_M \frac{P_{Ljt} L_{jt}}{P_{Mjt} M_{jt}} \left(\frac{M_{jt}}{L_{jt}}\right)^{-\frac{1-\sigma}{\sigma}}, \tag{18}$$

using which I can substitute and control the unobserved ω_{Ljt} in Eq.(17).

Now, I introduce how to control the unobserved Hicks-neutral productivity, ω_{Hjt}, using intermediate input (i.e. materials) to construct a proxy (Olley and Pakes, 1996; Levinsohn and Petrin, 2003; Ackerberg et al., 2015). Specifically, based on the updated intra-period states, the coal mine makes material (m_{jt}) input decision, which is given by the conditional material demand equation

$$m_{jt} = m_t(k_{jt}, l_{jt}, \omega_{Hjt}, \omega_{Ljt}, \mathbf{z}_{jt}), \tag{19}$$

[52]Using similar first-order conditions to control factor-augmenting productivity is widely adopted in the literature on non-neutral production function estimation. See, e.g., Van Biesebroeck (2003), Doraszelski and Jaumandreu (2018), Zhang (2019), and Demirer (2022), among others.

[53]See Appendix E for detailed derivations.

where the vector $\mathbf{z}_{jt}$ includes all other variables that affecting the material demand (De Loecker, 2011; De Loecker and Warzynski, 2012; De Ridder et al., 2024). In my application, I collect the exposed buyer power ($\widetilde{\psi_{ot}}$), output price, average wage, time-fixed effects, and dummy variables of ownership and geography in the vector $\mathbf{z}_{jt}$.

Disscussion about $\boldsymbol{z}_{jt}$ controls. I discuss about the inclusion of $\mathbf{z}_{jt}$ here specifically. De Ridder et al. (2024) propose to include prices, time-fixed effects, and market shares, controlling for the conduct and demand conditions, under imperfect competition in the first stage of the production function estimation, though they highlight that market share is not a perfect control for markup and demand conditions in every case. Alternatively, I control the coal mine's exposed buyer power ($\widetilde{\psi_{ot}}$), which I measured in Section 4.1 without utilizing any information from the coal mine, into the demand equation to directly control imperfect competition and unobserved demand conditions in output markets, on top of controlling output price variation. The direct inclusion of the buyer power exposure helps me depart from assuming the classical monopolistic competition (De Loecker, 2011) or other widely used competition models in output markets (Treuren, 2022; Ackerberg and De Loecker, 2024), while still satisfying the scalar unobserved assumption for the demand equation.[54]

In addition, the serially correlated average wage (labor price) is also included in the vector $\mathbf{z}_{jt}$ to exploit lagged variable inputs as valid instruments later in estimation (De Loecker and Warzynski, 2012; Gandhi et al., 2020). Conditional on the labor usage (in quantity), the included labor price further controls the coal mine's heterogeneity in labor quality. Finally, the time-fixed effects represent disembodied technical changes in the spirit of Olley and Pakes (1996),[55] and other remaining variables are different demand shifters of material input to be controlled.

Inverting the static material demand equation gives the control function for the Hicks-neutral productivity:[56]

$$\omega_{Hjt} = h_t(k_{jt}, l_{jt}, m_{jt}, \omega_{Ljt}, \mathbf{z}_{jt}), \tag{20}$$

where $h_t(\cdot)$ is a nonparametric function. Substituting Eq.(18) into Eq.(20) yields $\omega_{Hjt} =$

[54] An alternative option one can adopt in theory to accommodate imperfect competition in the context of the control function approach is to directly include the entire "state vector" of market participants in the demand input equation. Nevertheless, the problems are explicit in at least two folds, especially in my application. First, not all the participants are observed, and even observed, inverting all firms' productivity shocks as a function of all other firms' choices is tricky. Second, too many market participants easily lead to the curse of dimensionality and undermine the operation in practice. I refer to Ackerberg and De Loecker (2024) for a detailed discussion about this and other existing approaches in the literature estimating production function in the presence of imperfect competition. On top of these, they propose a novel sufficient statistic approach to identify production functions under imperfect competition.

[55] Olley and Pakes (1996) employ a time trend term to capture the technical changes.

[56] As I use a static material control to proxy for productivity, I do not need to solve the complicated dynamic programming problem for any dynamic input, which frees me from proving the invertibility property for the input demand (De Loecker, 2011; Ackerberg et al., 2015; De Loecker et al., 2016).

$h_t(k_{jt}, l_{jt}, m_{jt}, \frac{P_{Ljt}L_{jt}}{P_{Mjt}M_{jt}}, \mathbf{z}_{jt})$, which I exploit to replace ω_{Hjt} in Eq.(17).

Measurement Errors. Two measurement error terms (or unanticipated shocks) for coal quantity (ϵ^q_{jt}) and safety level (ϵ^s_{jt}) in Eq.(17) will bias the consistent estimation of the production frontier. To resolve the concerns, I purge out the measurement error of coal quantity (ϵ^q_{jt}) using the control function approach, as introduced before, in the first stage of the production function estimation procedure. In contrast, I address the measurement error issue of safety (ϵ^s_{jt}) by instrumenting for the mine-level accident probability, which is the proxy for the coal mine's safety level, with focal depth-weighted earthquake magnitudes. I will elaborate more in detail later in the following subsection of estimation.

Safety Measure. I now turn to illustrate how to construct the mine-level safety measure, the predicted coal mine's accident probability.[57,58] Specifically, as the coal mine makes its safety choice $\tilde{s}_{jt}$ given the targeted coal output $\tilde{q}_{jt}$, based on the initial states, $(k_{jt}, \omega^s_{Hjt}, \omega^s_{Ljt})$, at the beginning of time $\mathbf{t}$, one can capture the mine's safety choice by $\tilde{s}_{jt} = s_t(\tilde{q}_{jt}, k_{jt}, \omega^s_{Hjt}, \omega^s_{Ljt})$. Define the indicator function χ_{jt} to be equal to 1 if the mine happens an accident and 0 otherwise. Combining $\tilde{s}_{jt}$ and the control functions for ω_{Hjt} and ω_{Ljt} derived before, the accident probability can be given by

$$\begin{aligned}
\Pr(\chi_{jt} = 1) &= \Pr\left[\varsigma_{jt} \geq \bar{\varsigma}_{jt}(\tilde{s}_{jt}, \tilde{\mathbf{z}}_{jt})\right] \\
&= \varsigma(\tilde{s}_{jt}, \tilde{\mathbf{z}}_{jt}) \\
&= \varsigma\left(\tilde{q}_{jt}, k_{jt}, \omega^s_{Hjt}, \omega^s_{Ljt}, \tilde{\mathbf{z}}_{jt}\right) \\
&= \varsigma\left(k_{jt}, l_{jt}, m_{jt}, \frac{P_{Ljt}L_{jt}}{P_{Mjt}M_{jt}}, \mathbf{e}_{jt}, \tilde{\mathbf{z}}_{jt}\right) \\
&\equiv 1 - s_{jt},
\end{aligned} \tag{21}$$

where ς_{jt} represents the mine-specific time-varying accident risk, $\bar{\varsigma}_{jt}$ denotes the maximum risk resistance level conditional on its safety choice ($\tilde{s}_{jt}$), and relevant state variables ($\tilde{\mathbf{z}}_{jt}$), which consists of $\mathbf{z}_{jt}$ and other safety-related variables, e.g., the coal mine's age and year dummy. $\varsigma(\cdot)$ is a nonparametric function, and $\mathbf{e}_{jt}$ is a vector consisting of the intra-temporal i.i.d. shocks in productivity.[59]

[57] Note that "safety" can refer to many aspects in the coal mining context, e.g., the risk of coal mining accidents, the health outcome of employees, or whether having a sound accident rescue system. I focus on a specific safety dimension related to reducing coal mines' accident risk, which is arguably the most crucial concern for coal mine workers, owners, and policymakers.

[58] Since I can observe coal mine accident events, which contain casualty information, an alternative thought is to exploit the casualty information directly to construct the mine-level death rate measure as the proxy of safety level. However, as coal mining accidents are rare events overall, many observations are zero in most years; however, zero death rates don't necessarily mean high safety levels, as it can just because, for example, the coal mine accidents didn't cause death.

[59] More explicitly, one can write the intra-evolution of productivity as $\omega_{Hjt} = \mathbb{E}\left[\omega_{Hjt} \mid \omega^s_{Hjt}\right] + \epsilon^{\omega_H}_{jt}$ and

If the mine-specific accident risk is higher than the maximum risk resistance level, a coal mining accident would happen. Hence, s_{jt} measures the negative accident probability. In estimation, I use a flexible fourth-order polynomial in $(k_{jt}, l_{jt}, m_{jt}, \frac{P_{Ljt}L_{jt}}{P_{Mjt}M_{jt}})$ with a second-order polynomial in all elements of $\tilde{\mathbf{z}}_{jt}$ as regressors in a probit estimation in the spirit of Olley and Pakes (1996).[60]

Instrumenting for the Safety. I rely on the focal depth-weighted earthquake magnitudes to instrument for s_{jt} to address the measurement error issue in safety. Chen (2020) finds almost half of coal mine accidents in China were accompanied by earthquakes nearby that caused stress disturbances. Earthquakes are natural and exogenous events that are nearly unaffected by other economic activities, enabling them to satisfy exclusion restriction requirements and be a valid IV for instrumenting the coal mine's accident probability when controlling all other production-related variables in Eq.(17).

Estimation Procedures. I now turn to the details of estimating the production frontier in a two-step estimation procedure.

Substituting the control functions of the labor-augmenting and Hicks-neutral productivity, Eq.(18) and Eq.(20), respectively, into Eq.(17), denoting $\mathbf{x}_{jt} = \left\{k_{jt}, l_{jt}, m_{jt}, \frac{P_{Ljt}L_{jt}}{P_{Mjt}M_{jt}}\right\}$ to save notations, the first-stage estimating equation is:

$$q_{jt} = -\alpha_s s_{jt} + \phi_t(\mathbf{x}_{jt}, \mathbf{z}_{jt}) + \boldsymbol{\epsilon}_{jt}, \tag{22}$$

where $\phi_t(\mathbf{x}_{jt}, \mathbf{z}_{jt}) = \kappa \ln(M_{jt}) + \frac{\sigma\kappa}{\sigma-1} \ln\left(\frac{\beta_K}{\beta_M}\left(\frac{K_{jt}}{M_{jt}}\right)^{-\frac{1-\sigma}{\sigma}} + \left(1+\frac{P_{Ljt}L_{jt}}{P_{Mjt}M_{jt}}\right)\right) + \frac{\sigma\kappa}{\sigma-1}\ln\beta_M + h_t(\mathbf{x}_{jt}, \mathbf{z}_{jt})$ and $\boldsymbol{\epsilon}_{jt} = \alpha_s \epsilon^s_{jt} + \epsilon^q_{jt}$. See Appendix E for detailed derivations. Note that the safety choice is $s_{jt} = s_t(\cdot, \omega^s_{Hjt}, \omega^s_{Ljt})$, while the optimal intermediate input demand is $m_{jt} = m_t(\cdot, \omega_{Hjt}, \omega_{Ljt})$. The intra-period evolutions of Hicks-neutral and labor-augmenting productivity, following the exogenous Markov process, update the information sets and provide independent variation between $s_t(\cdot)$ and $\phi_t(\cdot)$ to identify α_s.

First-stage Estimation: Identifying the Production-Safety Tradeoff. To estimate Eq.(22) consistently, I first approximate the function $\phi_t(\cdot)$ with a flexible polynomial in all its variables in $(\mathbf{x}_{jt}, \mathbf{z}_{jt})$.[61] I run a two-stage least squares (2SLS) regression afterward using the exogenous provincial focal depth-weighted earthquake magnitudes as the IV for s_{jt}. The validity of earthquake IV is natural in the sense that earthquakes are relevant to coal mines' safety levels in China, as indicated by the findings of (Chen, 2020), and serve as a natural

$\omega_{Ljt} = \mathbb{E}\left[\omega_{Ljt} \mid \omega^s_{Ljt}\right] + \epsilon^{\omega_L}_{jt}$, where $\epsilon^{\omega_H}_{jt}$ and $\epsilon^{\omega_L}_{jt}$ are i.i.d. shocks. Hence, $\mathbf{e}_{jt} = (\epsilon^{\omega_H}_{jt}, \epsilon^{\omega_L}_{jt})$.

[60] The dummy variables enter the estimating regression linearly.

[61] Specifically, I use a fourth-order polynomial in all elements in $\mathbf{x}_{jt}$ and their interactions with a second-order polynomial in $\mathbf{z}_{jt}$. All dummy variables enter linearly.

exogenous supply shifter.

The first-stage estimation provides a consistent estimate of α_s and purges the composite measurement errors or unanticipated shocks, $\boldsymbol{\epsilon}_{jt}$, and ultimately, yields an estimate of predicted composite output, $\hat{\phi}_{jt}$. I can, therefore, express the Hicks-neutral productivity as:

$$\omega_{Hjt}(\boldsymbol{\theta}) = \hat{\phi}_{jt} - \kappa \ln(M_{jt}) - \frac{\sigma\kappa}{\sigma-1} \ln\left(\frac{\beta_K}{\beta_M}\left(\frac{K_{jt}}{M_{jt}}\right)^{-\frac{1-\sigma}{\sigma}} + \left(1 + \frac{P_{Ljt}L_{jt}}{P_{Mjt}M_{jt}}\right)\right) - \frac{\sigma\kappa}{\sigma-1}\ln\beta_M, \tag{23}$$

where $\boldsymbol{\theta} = (\kappa, \sigma, \frac{\beta_K}{\beta_M})$. Without loss of generality, I let $\beta_K = 1 - \beta_M$ following Doraszelski and Jaumandreu (2018).[62]

Second-stage Estimation: Identifying the Production Parameters. Relying on the law of motion of the Hicks-neutral productivity, which follows

$$\omega_{Hjt} = g\Big(\omega_{Hjt-1}, \widetilde{\psi_{ot-1}}; \boldsymbol{\beta}_g\Big) + \xi_{jt}, \tag{24}$$

where $\boldsymbol{\beta}_g$ is the vector of parameters in the productivity process $g(\cdot)$, which are estimated alongside $\boldsymbol{\theta}$, and ξ_{jt} refers to the innovation in the productivity shock that I use to form moments for estimating production function coefficients. I explicitly include the lagged exposed buyer power ($\widetilde{\psi_{ot-1}}$) in the productivity process following De Loecker (2011, 2013) to allow for market power effects on productivity to take place, if any.[63] In particular, $\frac{\partial g(\cdot)}{\partial \widetilde{\psi_{ot-1}}}$ informs whether and how would past buyer power impacts a firm's future productivity. The identification of the buyer power effects comes from the timing assumption, as illustrated in Figure 3, in which the buyer power information was realized prior to the coal mine receiving the updated productivity shock.[64]

Specifically, I obtain the productivity innovation by nonparametrically regressing $\omega_{Hjt}(\boldsymbol{\theta})$

[62]Note that β_K and β_M are not separately identified without such parametric restrictions. However, the non-identification result of β_K and β_M would not hurt the identification of output elasticities of the production frontier because the ratio of β_K and β_M, i.e., $\frac{\beta_K}{\beta_M}$, is still identified. I obtain indifferent output elasticities regardless of whether I impose such parametric restrictions. See Demirer (2022) for a detailed discussion and proof about different parameters' (non-)identification results in a general nonparametric production function setup, to which my CES specification is a special case. Nevertheless, the cost of such generality is the non-identification of the elasticity of substitutions, which is one of my key interests to be investigated in empirics.

[63]Note that the endogenous productivity process $g(\cdot)$ doesn't presuppose the existence of buyer power effects; oppositely, it nests the exogenous productivity process as a special case. One would expect non-significant estimates for the added endogenous variables (i.e., $\widetilde{\psi_{ot-1}}$ in my case) if the true date-generating process for productivity is exogenous to it. I report relevant estimates in Section 5.3.

[64]This implies that $\mathbb{E}(\xi_{jt} \cdot \widetilde{\psi_{ot-1}}) = 0$. Note that ξ_{jt} is a composite innovation term that consists of two i.i.d. shocks from stages of "new states realization" and "intra-period evolution", respectively, as in Figure 3, where both innovations are orthogonal to $\widetilde{\psi_{ot-1}}$.

on $\omega_{Hjt-1}(\boldsymbol{\theta})$, and form moments that identify the parameters, using a GMM estimator:

$$\mathbb{E}\left(\xi_{jt}(\boldsymbol{\theta})\begin{bmatrix} l_{jt-1} \\ m_{jt-1} \\ k_{jt} \end{bmatrix}\right) = 0. \tag{25}$$

As the timing assumptions elaborated, the firm makes labor and material input decisions after the actual productivity is realized at time **t+b**, while capital investment is chosen at time **t−1**. Hence, the productivity shock is orthogonal to lagged labor and material usage and current capital stock.

4.3 Empirical Strategy

Estimating Equation. My baseline empirical specification starts with regressing coal mining death rates in province o at year t, Death_{ot}, on the buyer power exposure variable, $\widetilde{\psi_{ot}}$, controlling for province-fixed effects γ_o and year-fixed effects γ_t, at the provincial level:

$$\ln \text{Death}_{ot} = \beta \ln \widetilde{\psi_{ot}} + \mathbf{x}_{ot}'\gamma + \gamma_o + \gamma_t + \varepsilon_{ot}, \tag{26}$$

where $\mathbf{x}_{ot}$ captures time-varying provincial characteristics, including provincial coal output and transport capacity. $\widetilde{\psi_{ot}}$ measures the aggregate exposed buyer power from power plants, which I construct from the power plant-level buyer power measure (ψ_{ft}^{M}), weighted by interprovincial coal railway freight volumes. More details will be elaborated in Section 4.1. By controlling the province-fixed effects, Eq.(26) isolates within-province variations in the death rates and buyer power exposure to identify the coefficients of interest.

I further use a variation of Eq.(26) with coal mine-fixed effects (γ_j) at the coal-mine level

$$\ln Y_{jot} = \delta \ln \widetilde{\psi_{ot}} + \mathbf{x}_{jot}'\gamma + \gamma_j + \gamma_t + \varepsilon_{jot}, \tag{27}$$

where Y_{jot} are coal mine-level outcome variables, and $\mathbf{x}_{jot}$ includes $\mathbf{x}_{ot}$ and other coal-mine level characteristics depending on different specifications. The coefficient of δ is identified from the correlation between the average growth rate of Y_{jot} across all mines in province o and the change in exposed buyer power with respect to province o.

Instrumenting for the Buyer Power Exposure. However, estimating these equations with ordinary least squares (OLS) leads to biased estimates if the buyer power exposure is endogenous. For example, increased electricity demand due to positive demand shocks (e.g., WTO accession) could incur more coal demand and supply, causing higher coal mining death rates and affecting buyer power. An alternative source of endogeneity could be the measurement of buyer power exposure is subject to error, resulting in attenuation bias.

To address the endogeneity issues, I construct a shift-share IV, z_{ot}, for province o's buyer power exposure by combining residual national coal demand of province d (excluding purchases from province o), $G_{dt}^{(o)}$ (the shifters), with interprovincial transportation patterns of past coal freight volumes, s_{od} (the shares).[65] Concretely,

$$z_{ot} = \sum_{d \in \Theta} s_{od} G_{dt}^{(o)}, \tag{28}$$

where Θ denotes the set of provinces, s_{od} is the export share of coal sold to province d in total coal sales for province o in the initial-sample year (Borusyak et al., 2022), and

$$G_{dt}^{(o)} = \sum_{o' \in \Theta \backslash \{o\}} Q_{o'dt} \tag{29}$$

with $Q_{o'dt}$ capturing the amount coal that province d bought from province $o' \in \Theta \backslash \{o\}$ at year t. I further adjust $Q_{o'dt}$ to be the coal sold specifically to the power generation sectors using key state-owned coal mines' inter-provincial and sectoral sales (in quantity) to construct weight. The adjustment explicitly attempts to absorb power market-related demand-side exogenous shocks. Nevertheless, the results remain robust regardless of adjustments, as seen in Section 5.4.

The shift-share IV exploits variations from two sources. First, provincial demand-side shocks in coal consumption from the coal-fired power generation sector (the shifters) predict variations in buyer power from coal-fired power plants in these provinces. Second, coal mines in coal-selling provinces tend to be exposed to more buyer power from power plants in coal-buying provinces with which they have sold more coal in the past (the share). I follow Borusyak et al. (2022) to assume that the shifters (demand-side shocks from power plants) are conditionally orthogonal to coal mining industry outcomes, allowing the shares (historical coal selling patterns) to be endogenous.

Threats to Identification. For an instrument to be considered valid, it must shift coal demand without influencing coal mines other than through the buyer power channel. Suppose a rise in $G_{dt}^{(o)}$ for province d is not only due to changes in province d's demand shocks but also reflects shocks to supply in province o. For example, safety regulations in the coal mining industry of province o, such as shutting down small coal mines or restricting coal production, could reduce coal mining death rates and increase $G_{dt}^{(o)}$. I deal with this concern

[65] o for "origin" (coal-selling province) and d for "destination" (coal-buying province).

by incorporating the provincial total coal output in $\mathbf{x}_{ot}$ to control for time-varying supply shocks in province o.

In addition, shocks to transportation patterns might also affect the coal supply from province o.[66] A newly constructed railway or oppositely insufficient transport capacity due to imperfect railway scheduling could also cause either an increase or decrease in $G_{dt}^{(o)}$ and ambiguous change in the coal supply. Thus, I control the provincial transport capacity for different types of transportation, e.g., railways, waterways, and roads, in $\mathbf{x}_{ot}$. Note that s_{od} also incorporates information on transportation patterns, but it would not affect consistency as the exposure shares are allowed to be endogenous (Borusyak et al., 2022). Nevertheless, I use data from the initial sample year to calculate s_{od} to fix transportation patterns and thus focus on buyer power's role in my empirics.

5 Empirical Results

This section provides the estimation results from the empirical models. I start by illustrating the estimated coal price markdowns and the evolution of buyer power of power plants. Second, I present the production function estimates of coal mines and nonparametric estimates of buyer power on productivity. Finally, I use these estimates to implement the empirical strategy. The aim is to identify and understand how buyer power affects the organization of coal mine production.

5.1 The Evolution of Buyer Power

Panel A of Table 2 presents the estimated production function coefficients of power plants for both specifications using OLS and the control function approach, i.e., GMM. The second column shows the output elasticities of labor (0.409) and capital (0.847) estimated using the control function approach. In contrast, the OLS method in the first column overestimates the output elasticity of labor (0.458) and underestimates the output elasticity of capital (0.637). Both specifications imply a capital-intensive nature and increasing returns of scale for power plants (1.095 and 1.256), which aligns with findings in the broad literature on electric utilities.[67]

Based on the different production function estimates, the coal price markdowns are in Panel B of Table 2. The second column is the preferred specification. One can note that, at the median power plant, the coal price markdown is 1.077, while the average coal price

[66] Note that the transport costs would not affect the coal supply due to the context of the regulated railway industry and the priority of coal transportation as mentioned in Section 2.2.

[67] See Nerlove et al. (1961), Atkinson and Primont (2002), Atkinson (2019), and Atkinson and Luo (2024), among others.

Table 2: Model Estimates of Power Plants

	Estimation Method	
	OLS (1)	GMM (2)
Panel A: Production function		
Labor	0.458***	0.409***
	(0.029)	(0.118)
Capital	0.637***	0.847***
	(0.026)	(0.054)
Returns to scale	1.095***	1.256***
	(0.018)	(0.080)
Observations	5,315	5,315
Panel B: Coal price markdown		
Median	1.096***	1.077***
	(0.016)	(0.076)
Average	1.159***	1.136***
	(0.018)	(0.088)

Notes: The returns to scale coefficients sum up the corresponding capital, labor, and material coefficients. Standard errors are computed by bootstrapping 200 times. ***: p<0.01; **: p<0.05; *: p<0.1.

markdown is 1.136, which means that power plants, on average, procure coal at a price 12% below its marginal revenue product of coal.[68] Though it's not very comparable for markdown estimates across industries, the average markdown estimate is in line with other studies that also measure buyer power in material input markets, e.g., Treuren (2022) and Avignon and Guigue (2022), and is sensibly milder than that in Rubens (2023) which studies on more restricted oligopsony markets.

To illustrate the evolution of buyer power, I regress the coal price markdown on power plant- and year-fixed effects, where standard errors are clustered at the prefectural level. Figure A3 plots the coefficients of different year dummies, where 2001, the year prior to the electricity sector restructuring announced, is normalized to zero. Two key patterns emerge. First, there were rich time-series variations in the coal price markdowns. Second, the coal price markdown was relatively stable before 2003 and became more volatile afterward, slumping on average by 7.7 percentage points in 2004 and onwards. The time-series variations pick up well the shock of electricity sector restructuring, which corresponds with the literature and anecdotal evidence that the restructuring did take several years to materialize (Gao and Van Biesebroeck, 2014).[69] The evolution of coal price markdowns is very robust to different

[68] $12\% = 1 - \frac{1}{1.136}$.

[69] Gao and Van Biesebroeck (2014) indicates that the restructuring and vertical unbundling in the electricity sector finished around 2004, as power plants' ownership types kept changing until then.

markup calibrations, as seen in Appendix A.

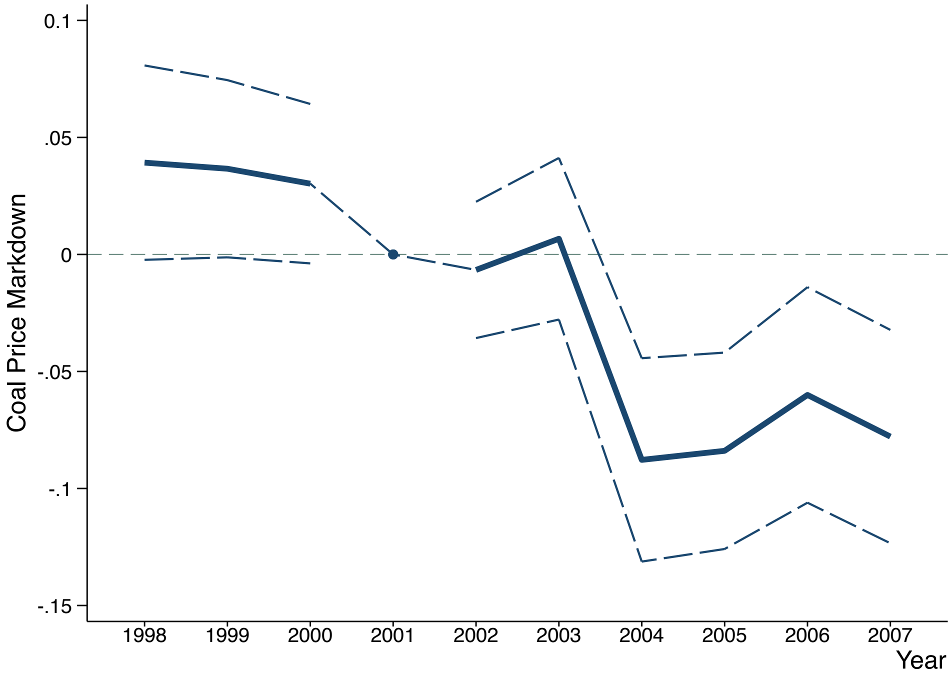


Figure 4: The Evolution of Coal Price Markdowns

Notes: I normalize the year before the announcement of the electricity sector restructuring to zero with omitted confidence intervals. Standard errors are clustered at the prefectural level. 95% confidence intervals in dashed lines are shown.

Though I do not attempt to disentangle the source of the change in buyer power, as I only rely on rich variation in the coal price markdowns to enable identifying the role of buyer power, it seems that the national electricity sector restructuring is a driving force of buyer power that cannot be ignored. Therefore, I preserve the time-varying shocks from the electricity sector restructuring by exploiting the complete time-series variations in my main specification to better understand the buyer power effects, as shown in Section 4.3.

5.2 Production Function Estimates

Table 3 reports the estimates of the transformation function of coal mines. The second column is the preferred specification, while the first column presents the estimated parameters using OLS for comparison.

Both specifications provide evidence of a statistically significant production-safety tradeoff, showcasing that coal mines seriously consider safety when making production decisions. However, the magnitudes of the tradeoff coefficient differ substantially between the structural model (-0.038) and naive OLS (-0.007), highlighting the importance of addressing the endogeneity problems when estimating the production function for coal mines. Specifically,

Table 3: Model Estimates of Coal Mines

		Estimation Method	
		OLS	Model
		(1)	(2)
Safety	$-\alpha_s$	-0.007***	-0.038***
		(0.001)	(0.011)
Output elasticity of capital	θ_{Kjt}	0.051***	0.254***
		(0.004)	(0.001)
Output elasticity of labor	θ_{Ljt}	0.152***	0.126***
		(0.017)	(0.001)
Output elasticity of material	θ_{Mjt}	0.749***	0.391***
		(0.021)	(0.001)
Returns to scale	κ	0.952***	0.771***
		(0.004)	(0.025)
Elasticity of substitution	σ		2.270***
			(0.474)
Observations		20,144	20,007

Notes: The returns to scale coefficient for the OLS specification is summing up over its capital, labor, and material coefficients, while its elasticity of substitution is 1, per the Cobb-Douglas specification. Column (2) reports the model's mean output elasticities of capital, labor, and material. Standard errors are computed by bootstrapping 200 times. ***: p<0.01; **: p<0.05; *: p<0.1.

the estimate of the production-safety tradeoff, α_s, from the structural model, is -0.038, which means a coal mine that targets to reduce accident probability by 1 percentage point would need to forgo its coal production by 3.8%. This corresponds to an elasticity of coal output with respect to safety of -3.76 at the median and -3.72 at the mean. Put differently, a coal mine could increase its coal output by 1% by reducing the safety level such that its accident probability rises by 0.26 percentage points, holding others fixed.[70] The significant tradeoff between safety and production in coal mining is consistent with Gowrisankaran et al. (2015).

The mean output elasticities of capital, labor, and material are 0.254, 0.126, and 0.391, respectively, from the structural model. The OLS method, as often found in the production function literature, underestimates the capital coefficient (0.051) and overestimates the labor coefficient (0.152), while the material coefficient is 0.749. Both specifications indicate that coal mines are relatively material-intensive, which aligns with the coal mining industry's feature that consumables such as pitwood and gunpowder account for a substantial proportion of input expenditure.

The returns to scale coefficients are 0.952 and 0.771, respectively, for specifications of OLS and model, reflecting decreasing returns to scale in the coal mining industry. A decreas-

[70] 0.26 p.p. $= \frac{1}{3.8}$.

ing returns to scale coefficient is familiar in the (single-output) production function literature (Collard-Wexler and De Loecker, 2015; Allcott et al., 2016), while also fits well with the nature of a multi-output production process, as the extra complexity and coordination costs embedded could render decreasing returns to scale. Finally, the elasticity of substitution is 2.270, greater than unity, primarily showcasing that material and capital are gross substitutes with each other, which matches well with the mining technological characteristics as mentioned in Section 2.1, and also corresponds to recent findings about a higher than unity elasticity of substitution in the context of Chinese industries using different methods (Berkowitz et al., 2014; Grieco et al., 2022; Li and Zhang, 2022; Meng et al., 2023).

Nonparametric Estimates of Buyer Power on Productivity. Alongside estimating production function coefficients, I obtain the nonparametric estimates of buyer power on Hicks-neutral productivity at the same time, as shown in Eq.(24). Regarding labor-augmenting productivity, note that I do not impose any restrictions on its evolution other than the classical first-order Markov process assumption, as the unobserved labor-augmenting productivity is replaced by corresponding first-order conditions, i.e., Eq.(18), and disappeared in the estimating equation. It is, therefore, flexible to any complex productivity dynamics on top of the first-order Markov process.[71] For comparison with Hicks-neutral productivity, I start by assuming that labor-augmenting productivity follows an analogous process as $\omega_{Hjt} = g\left(\omega_{Hjt-1}, \widetilde{\psi_{ot-1}}; \boldsymbol{\beta}_g\right) + \xi_{jt}$. I approximate both productivity processes by a third-order polynomial in productivity and their interactions with buyer power exposure and test whether the endogenous productivity process is plausible.

Table 4 presents the buyer power effects on productivity. Three interesting results emerge. First, the F-tests on the joint significance of all parameters for buyer power-related variables show that it is essential to incorporate buyer power into account as it can affect future Hicks-neutral and labor-augmenting productivity, which both reject an exogenous productivity process.

Second, buyer power exposure affects future productivity in heterogeneous ways with respect to different types of productivity. For Hicks-neutral productivity, as the interaction terms of buyer power and productivity are significant, the effects of buyer power vary substantially depending on the coal mine's productivity. Coal mines with high productivity would suffer more from the imperfect competition, which corresponds to Backus (2020) that finds firms in the upper productivity quantiles have higher productivity gains when facing competition. In contrast, I do not find significant estimates for any interactions with

[71] A recent paper by Caselli et al. (2024) proposes a novel method to study productivity and quality for multi-product firms without imputing intra-firm input allocations and allows for flexibility in exploring complex productivity dynamics after estimation, which shares similar insights.

Table 4: Buyer Power Effects on Productivity

	ω_{Hjt} (1)	ω_{Ljt} (2)	ω_{Ljt} (3)
Productivity	0.271***	0.758***	0.729***
	(0.038)	(0.017)	(0.007)
Productivity^2	0.061***	0.023***	0.022***
	(0.018)	(0.005)	(0.002)
Productivity^3	0.009**	-0.017***	-0.014***
	(0.004)	(0.001)	(0.001)
Buyer power exposure, $\ln \widetilde{\psi_{ot-1}}$	-1.982***	-0.138*	-0.194***
	(0.186)	(0.083)	(0.058)
$\ln \widetilde{\psi_{ot-1}} \times$ Productivity	0.976***	-0.113*	
	(0.132)	(0.063)	
$\ln \widetilde{\psi_{ot-1}} \times$ Productivity^2	0.195***	0.003	
	(0.056)	(0.018)	
$\ln \widetilde{\psi_{ot-1}} \times$ Productivity^3	-0.107***	0.012**	
	(0.013)	(0.005)	
Average effect (*in percent*)	-0.216	-0.195	-0.194
Observations	13,393	24,711	24,711
F-test	$F(4, 13385) = 36.71$	$F(3, 24703) = 2.27$	$F(1, 24706) = 11.07$
p-value	0.000	0.079	0.001

Notes: The F-test is on the joint significance of the coefficients from all buyer power-related variables. The corresponding p-value of the test is reported below. Column (2) reports the F-test on the joint significance of all interaction terms with buyer power, while including buyer power per se leads the F-test result to be $F(4, 24703) = 4.47$. ***: p<0.01; **: p<0.05; *: p<0.1.

labor-augmenting productivity at 1% significance level, as seen in column 2. The F-test of the joint significance of the interactions is also barely significant at 10% significance level. That said, buyer power affects labor-augmenting productivity directly and (log-)linearly regardless of the heterogenous productivity levels. The findings of heterogeneous buyer power effects on productivity are novel in the literature but sensible as they imply heterogeneous technological changes.

Third, holding fixed productivity, columns 1 and 3 present that a 1% increase in the buyer power exposure leads to, on average, a decrease in future Hicks-neutral and labor-augmenting productivity by 0.22% and 0.19%, respectively. This finding is consistent with frequently observed results in the productivity literature, which finds that competition boosts firms' efficiency.[72]

[72] I refer to Backus (2020) for a detailed review.

5.3 Death by Market Power and Beyond

Death Rates. Table 5 presents estimates of Eq.(26) for the death rates in the coal mining industry. The dependent variable is logged death rates due to coal mining accidents, and I control the provincial coal physical output on the right-hand side so that the estimates correspond to absolute casualty effects.

Table 5: Buyer Power Effects on Coal Mining Death Rates

$\ln \text{Death}_{ot}$	(1)	(2)	(3)
Buyer power exposure, $\ln \widetilde{\psi_{ot}}$	0.798***	3.865**	4.549**
	(0.287)	(1.555)	(2.077)
Method	OLS	2SLS	2SLS
First stage F-statistic		22.33	11.61
Observations	176	176	176

Notes: Column (1) reports the OLS results, while columns (2)-(3) present the 2SLS results. Therein, provincial coal physical output and province- and year-fixed effects are controlled for all columns. Column (3) further controls transportation patterns. Robust standard errors are reported in the parentheses. ***: p<0.01; **: p<0.05; *: p<0.1.

Specifically, the 2SLS IV estimate in column 2 indicates that a 1% increase in the buyer power exerted by power plants causes coal mines to encounter 3.87% more deaths, conditional on coal production. In contrast, the OLS point estimate (0.8%) strongly underestimates the effects of buyer power on death. It is consistent with the expected biased directions due to either measurement error or omitted positive demand shocks/negative supply shocks. By further controlling transport characteristics, column 3 shows a slightly larger IV estimate than column 2, meaning the exposed buyer power in the coal mining industry increases by 1% would incur the death rate increase by 4.55%. The larger estimate in column 3 suggests that transport capacity negatively correlates with buyer power exposure. In Section 5.4, I conduct a robustness check using different definitions of death rates, and it presents qualitatively consistent estimates.

Using the fact that the mean casualties caused by coal mining accidents in each province are 243, the point estimate suggests that a 1% increase in buyer power exposure increases the number of deaths by $4.55\% \times 243 \approx 11$ people per province a year, holding the coal production quantity fixed. Put differently, given that the mean coal production output is around 70 million tons per province, a coal-producing province exposed to one more percent buyer power increases its death rate by $4.55\% \times 243/70 \approx 0.16$ deaths per million tons of coal production.

Outputs of Coal and Safety: Coal Mine Level. I now directly detect the buyer power effects on the coal mine's outputs at the coal mine level. To do so, I first construct the coal mine's composite output as $y_{jt} := q_{jt} + \hat{\alpha}_s s_{jt}$, where $\hat{\alpha}_s = 0.039$ as in Table 3. I then plug y_{jt} into Eq.(27) and control a flexible fourth-order polynomial in all variables in $(\mathbf{x}_{jt}, \mathbf{z}_{jt})$ other than the exposed buyer power $(\widetilde{\psi_{ot}})$, adding provincial characteristics $\mathbf{x}_{ot}$ on top of the firm- and year-fixed effects. Standard errors are clustered at the prefectural level, which is the highest level I can cluster.[73]

Table 6 shows the estimated buyer power effects on coal and safety outputs. The coefficient of $\widetilde{\psi_{ot}}$ is -0.926, meaning that, on average, a 1% increase in buyer power of power plants leads to a 0.93% drop in the composite outputs y_{jt}, or the same magnitude impact on coal output q_{jt} if the safety level is held fixed. Equivalently, holding the coal output q_{jt} constant, a coal mine exposed to one extra percent of buyer power reduces its safety level by 0.24 percentage points.[74] This finding aligns well with the classical industrial organization theory, as buyers can exert their market power by withholding input purchases. The impact of buyer power on output (or input, proportionally) is also in line with some empirical studies in magnitudes, e.g., Rubens (2023), though in different industries.[75]

Table 6: Buyer Power Effects on the Coal Mine's Outputs

	$q_{jt} + \hat{\alpha}_s s_{jt}$
Buyer power exposure, $\ln \widetilde{\psi_{ot}}$	-0.926***
	(0.353)
Method	2SLS
First stage F-statistic	51.90
Observations	18,390

Notes: Firm- and year-fixed effects are controlled, while all control variables are omitted from the reported results. Standard errors are clustered at the prefectural level. ***: p<0.01; **: p<0.05; *: p<0.1.

Overall, both coal mine-level and provincial evidence consistently indicate that higher buyer power results in lower safety levels and more deaths, whereas lower buyer power exposure contributes to better safety outcomes and fewer deaths.

I now turn to disentangle the underlying mechanism behind it.

[73]Note that the groups of provinces (#province = 22) are too few to be clustered to obtain consistent standard errors.

[74]0.24 p.p. = $\frac{0.926}{3.8}$.

[75]Rubens (2023) studies a consolidation policy in China's tobacco industry and finds that, on average, consolidation increases manufacturers' markdowns by 37%, and aggregate cigarette production falls by 38% accordingly.

Technological Changes. The buyer power induced more coal mining deaths and lower productivity, both Hicks-neutral and labor-augmenting productivity, reflecting embedded technological changes. To check this, I first investigate how buyer power affects the coal mine's input adoptions, as seen in Table 7. Each input column controls the other two inputs in the input bundle $\{k_{jt}, l_{jt}, m_{jt}\}$. I also control $\frac{P_{Ljt}L_{jt}}{P_{Mjt}M_{jt}}$ in the spirits of control functions for Hicks-neutral and labor-augmenting productivity, on top of provincial characteristics $\mathbf{x}_{ot}$ and firm- and year-fixed effects like other specifications. Column 1 shows that increasing buyer power exposure by 1% significantly induces a decline in capital adoption by 1.67%, holding other inputs fixed. In contrast, both labor and material inputs are mute in response to the buyer power exposure, as seen in Columns 2 and 3. The result is intuitive as capital adoption, such as better excavation machines or enhanced ventilation systems, requires substantial upfront investments, especially disincentivized once profit margins are squeezed when coal mines face higher buyer power exposure. This empirical finding is consistent well with Proposition 1 in Section 3.3. Of course, cost of capital may straightforwardly speak to capital adoption, but I will leave that for a detailed discussion in the following subsection.

Table 7: Buyer Power Effects on Input

	ln K	ln L	ln M
	(1)	(2)	(3)
Buyer power exposure, $\ln \widetilde{\psi_{ot}}$	-1.663***	-0.317	-0.425
	(0.625)	(0.344)	(0.843)
Method	2SLS	2SLS	2SLS
First stage F-statistic	47.63	47.28	47.14
Observations	18,431	18,431	18,431

Notes: Each input column controls the other two inputs in the input bundle $\{k_{jt}, l_{jt}, m_{jt}\}$. $\frac{P_{Ljt}L_{jt}}{P_{Mjt}M_{jt}}$, provincial characteristics $\mathbf{x}_{ot}$, and firm- and year-fixed effects are controlled for all columns, but are omitted from the reported results. Standard errors are clustered at the prefectural level. ***: p<0.01; **: p<0.05; *: p<0.1.

I also check the coal mine's adjustments in input intensity margin. The relative input intensity reflects the composition of technological changes immediately, as seen in Section 2.1 that different types of technology have distinct input intensity features, which can help to infer the potential technology adopted by coal mines. Table 8 presents the buyer power effects on the coal mine's input intensity. Column 1 demonstrates that a 1% increase in buyer power exposure leads to a significant 1.23% reduction in the capital-to-labor ratio, which aligns with Proposition 1 in Section 3.3. Conversely, the capital-to-material and material-to-labor ratios, as shown in Columns 2 and 3, exhibit a reduction in point estimate but no significant response to changes in buyer power.

Note that the capital adoption is found to be reduced in response to increased buyer power exposure in Table 7, but only the capital-labor intensity is adjusted significantly but not capital-material intensity, as seen in Table 8. Combining the specific technological features of different mining technologies from Table 1, the underlying mechanism appears: buyer power reduces the coal mine's capital adoption and shifts the mining technology to be less capital-intensive with respect to labor but not material. The reduced capital-to-labor ratio reflects the transition from relatively modern to traditional technology. In contrast, no significant reduction in capital-to-material ratio is observed, indicating the technical switching could only be from the conventionally-mechanized mining method to the blasting mining method, which tends to be more dangerous by nature, suggesting higher death rates. In Section 5.4, I allow for the heterogeneity in the production-safety tradeoff concerning technology and capital-to-labor ratio, which further corroborates what I find here.

Table 8: Buyer Power Effects on Input Intensity

	ln K/L (1)	ln K/M (2)	ln M/L (3)
Buyer power exposure, $\ln \widetilde{\psi_{ot}}$	-1.234**	-0.990	-0.182
	(0.518)	(0.904)	(0.804)
Method	2SLS	2SLS	2SLS
First stage F-statistic	47.87	47.59	47.25
Observations	18,431	18,431	18,431

Notes: Each input intensity column controls the remaining input in the input bundle $\{k_{jt}, l_{jt}, m_{jt}\}$. $\frac{P_{Ljt}L_{jt}}{P_{Mjt}M_{jt}}$, provincial characteristics $\mathbf{x}_{ot}$, and firm- and year-fixed effects are controlled for all columns, but are omitted from the reported results. Standard errors are clustered at the prefectural level. ***: p<0.01; **: p<0.05; *: p<0.1.

Cost of Capital. In the theoretical model in Section 3.2, I assume that buyer power can affect the cost of capital with $\frac{\partial r(\mathbf{b})}{\partial \mathbf{b}} > 0$, i.e., an increase in buyer power exposure leads to higher cost of capital, resulting in Corollary 1 in Section 3.3. Though the cost of capital is unobserved, the assumption is verifiable indirectly in empirics by inferring from the relative output elasticity of inputs.

To see this, one can derive that, under the cost minimization, a firm's output elasticity of a flexible input, e.g., labor L, without adjustment cost and input market power can be expressed as $\theta^L := \mu \frac{P^L L}{PQ}$, where μ is the firm-level markup, P^L and P are input and output prices, respectively, and Q is output. See De Loecker and Warzynski (2012) for a detailed derivation. I omitted all subscripts for succinct. Denote a quasi-fixed input K. One can treat it as if it's a flexible input to write out the first-order condition and express the corresponding output elasticity as $\theta^K := \mu\nu\frac{P^K K}{PQ}$ with an additional term ν to capture all capital-relevant distortions, e.g., adjustment cost, searching cost, capital input market power, etc. Dividing θ^K by θ^L, one can obtain $\frac{\theta^K}{\theta^L} = \frac{\nu P^K}{P^L}\frac{K}{L}$. Therefore, conditional on capital-to-labor ratio $\frac{K}{L}$ and labor input price P^L, one can infer the distortion-absorbed cost of capital $\widetilde{P^K} := \nu P^K$ by checking the relative output elasticity $\frac{\theta^K}{\theta^L}$.[76] Table 9 presents the buyer power effects on the coal mine's relative output elasticity for capital and labor.

Table 9: Buyer Power Effects on Relative Output Elasticity

	$\ln \theta^K_{jt}/\theta^L_{jt}$ (1)	$\ln \theta^K_{jt}/\theta^L_{jt}$ (2)	$\ln \theta^K_{jt}/\theta^L_{jt}$ (3)
Buyer power exposure, $\ln \widetilde{\psi_{ot}}$	0.572	1.269**	0.478**
	(0.588)	(0.590)	(0.204)
Capital-to-labor ratio, ln K/L		0.568***	0.700***
		(0.016)	(0.007)
Wage per worker, $\ln P^L_{jt}$			-0.944***
			(0.004)
Method	2SLS	2SLS	2SLS
First stage F-statistic	47.94	47.80	47.67
Observations	18,447	18,431	18,427

Notes: m_{jt}, $\frac{P_{Ljt}L_{jt}}{P_{Mjt}M_{jt}}$, provincial characteristics $\mathbf{x}_{ot}$, and firm- and year-fixed effects are controlled for all columns but are omitted from the reported results. Standard errors are clustered at the prefectural level. ***: p<0.01; **: p<0.05; *: p<0.1.

Column 1 shows that the capital and labor's relative output elasticity as a whole doesn't react significantly to changes in buyer power exposure, while once the capital-to-labor ratio is controlled, as in Column 2, one can find a significantly positive estimate of buyer power on relative output elasticity. This is unsurprising as Column 1 in Table 8 indicates a significantly negative point estimate for capital-to-labor ratio, which is a component of $\theta^K_{jt}/\theta^L_{jt}$ as seen before. Hence, what Column 2 captures is that a 1% increase in buyer power raises the relative price of capital and labor by 1.27%. Conditioning the labor price in Column 3, one can further isolate the buyer power effect on the cost of capital, which is 0.48%. This corresponds to the reduced capital adoption found in Table 7, which corroborates that at least a portion of reduced capital was due to the increased cost of capital.

[76] Similar insights have been adopted to infer input market wedges when comparing relative output elasticity and cost expenditure of two variable inputs by Morlacco (2019), Wong (2021) and Delabastita and Rubens (2024), among others.

5.4 Robustness Checks

Table 10 presents robustness checks on alternative measures for dependent, endogenous, and instrumental variables when studying the buyer power effects on coal mining death rates. I directly report the 2SLS results, controlling provincial coal physical output, transportation patterns, and province- and year-fixed effects. Estimates are very similar and consistent using alternative measures.

Specifically, the first column in Table 10 indicates that changing the death rate measure to death per worker, a 1% increase in the exposed buyer power leads to increased death per worker by 4.18%. In contrast, in column 2, I do not adjust the interprovincial transportation patterns of coal freight volumes by the key state-owned coal mines' inter-provincial and sectoral sales. This allows the IV to directly capture demand-side shocks to other coal-consuming industries, e.g., the steel sector, in the buying provinces. The result shows a slightly higher estimate than that in the main text that the exposed buyer power in the coal mining industry increases by 1%, which would induce the death rate to increase by 5.52%. Ultimately, I use alternative weight, i.e., interprovincial input-output flows of coal (in monetary value) in 1997, to construct the provincial buyer power exposure. A concern about using the input-output data is that it only reports monetary values, which implicitly absorbs price and market power effects. Still, column 3 presents a qualitatively consistent and robust estimate of buyer power effects on death rates.

Table 10: Robustness Checks: Alternative Measures

	Death per worker	Alternative IV	Alternative $\ln \widetilde{\psi_{ot}}$
	(1)	(2)	(3)
Buyer power exposure, $\ln \widetilde{\psi_{ot}}$	4.177**	5.521**	5.516*
	(2.034)	(2.273)	(3.182)
Specification	2SLS	2SLS	2SLS
First stage F-statistic	11.61	11.17	15.49
Observations	176	176	215

Notes: All columns control provincial coal physical output, transportation patterns, and province- and year-fixed effects. Robust standard errors are reported in the parentheses. ***: p<0.01; **: p<0.05; *: p<0.1.

I further allow for the heterogeneity in the production-safety tradeoff, by which the slope of the production-safety frontier could vary across mines depending on their technology or capital-to-labor ratio. I report the tradeoff estimates in Table 11. I first proxy modern technology by checking whether a coal mine is state-owned. As noted in Section 2.1, state-owned coal mines tend to utilize the most advanced mining technology with the highest mechanization rate among all other counterparts. Column 1 indicates that the slope of the production frontier is related to the technology type. Coal mines with more advanced modern technology flatten the production-safety tradeoff and expand the production frontier toward safety compared to coal mines using relatively traditional technology. This evidence is in line well with the technological features of different mining methods, as introduced in Section 2.1. The average production-safety tradeoff estimate is -0.051, which aligns with that found in the main text, though slightly larger in magnitude. The estimate corresponds to an average elasticity of coal output with respect to the safety of -5.01.

Table 11: Robustness Checks: Heterogeneity in the Production-Safety Tradeoff

	Modern-Tech (1)	Capital-to-Labor Ratio (2)
Safety, s_{jt}	-0.065***	-0.038
	(0.013)	(0.100)
Safety, $s_{jt} \times$ Tech	0.050***	
	(0.014)	
Safety, $s_{jt} \times \ln \mathrm{K/L}$		0.004
		(0.036)
Specification	Model	Model
Observations	20,007	20,007

Notes: Tech dummy is proxied by state-owned ownership status. Standard errors are computed by bootstrapping 200 times. ***: p<0.01; **: p<0.05; *: p<0.1.

Nevertheless, using coal mines' ownership to proxy technological differences is not free from defects, as state-owned coal mines may also have different regulatory incentives other than technological features affecting their production and safety decisions. In column 2, I further allow for the interaction of the capital-to-labor ratio with the safety measure (s_{jt}) to isolate heterogeneity concerning technological differences in the tradeoff. The results are qualitatively consistent. In particular, coal mines with higher capital-to-labor ratios, like adopting more advanced modern technology in column 1, make the production-safety tradeoff flatter, shifting the production frontier towards safety.

However, the coefficients are less precisely estimated, though the point estimate of the production-safety tradeoff is almost identical to the model in the main text. The high correlation between the capital-to-labor ratio and other nonparametric terms is further exacerbated when it interacts with the safety variable, inflating the standard errors of the estimated coefficients and making them less precise.[77] I, therefore, take the model in the main text as

[77]Similar results of inflated standard errors when considering heterogeneity in the tradeoff have also been

the primary estimate of the production-safety tradeoff in the coal mining industry.

6 Conclusion

In this paper, I estimate the impacts of buyer power on the within-firm organization of production in the Chinese coal mining industry, instrumenting buyer power exposure with a shift-share IV exploiting exogenous changes from a electricity sector restructuring and other demand-side shocks.

Four main conclusions are to be drawn. First, I construct a structural model for coal mines with multiple outputs of coal quantity and safety level without imputing the output-specific input allocation scheme, incorporating endogenous safety choices and factor-augmenting productivity. The results show that increased buyer power reduces coal mines' future labor-augmenting productivity and negatively impacts Hicks-neutral productivity heterogeneously.

Second, using estimates obtained from the model, I find that a 1% increase in buyer power exposure increases the number of deaths by 11 people per province a year, holding the coal production quantity fixed. In contrast, coal mine-level evidence presents that a 1% increase in buyer power of power plants leads to a 0.93% drop in coal output if the safety level is held fixed. Taking the average coal price of 151 RMB Yuan per ton in 2001 (Wang and Horii, 2008) and the average provincial coal production output of 70 million tons as the benchmark, one can attribute the economic loss due to a 1% increase in buyer power, holding safety level fixed, is around 98 million RMB Yuan per province.[78] This suggests the opportunity cost of saving one miner's life is roughly 9 million RMB Yuan. However, the "national benchmark" of the compensation payments in post-accident settlements for each worker killed was 200,000 RMB Yuan (CLB, 2008), 2% of the opportunity cost. Apparently, coal mine owners would have much higher incentives to sacrifice safety to produce more coal output, which partially explains the record-breaking alarming coal mining death rate performance at the time.

Third, back-of-the-envelope analysis indicates substantial unintended consequences of buyer power on coal mining death rates. On average, a 7.7 percentage points (equivalently 6.8%) reduction in buyer power after 2004 caused coal mining death rates to decrease by 27%,[79] while the death rates overall decreased by 51% during the same period. Back-of-the-envelope calculations demonstrate that the declined buyer power explains 53% of the improved coal mining death rate performance.

observed in Grieco and McDevitt (2017).

[78] $98 = 0.93\% \times 70 \times 151$.

[79] $27\% = 1 - \exp(4.549 \times \ln(1 - 6.8\%))$.

Fourth, I show that increased buyer power reduces capital adoption, shifting mining technology to be less capital-intensive and more traditional. This corresponds to the technological switch from conventionally-mechanized to blasting mining, which is more dangerous by nature, leading to higher death rates. All the findings mentioned above highlight the profound effects of buyer power on the within-firm organization of production, underscoring the need for careful consideration of market power and its holistic consequences in regulatory and policy frameworks, with implications that extend beyond the context of the coal mining industry in the paper.

Appendix

A Additional Figures

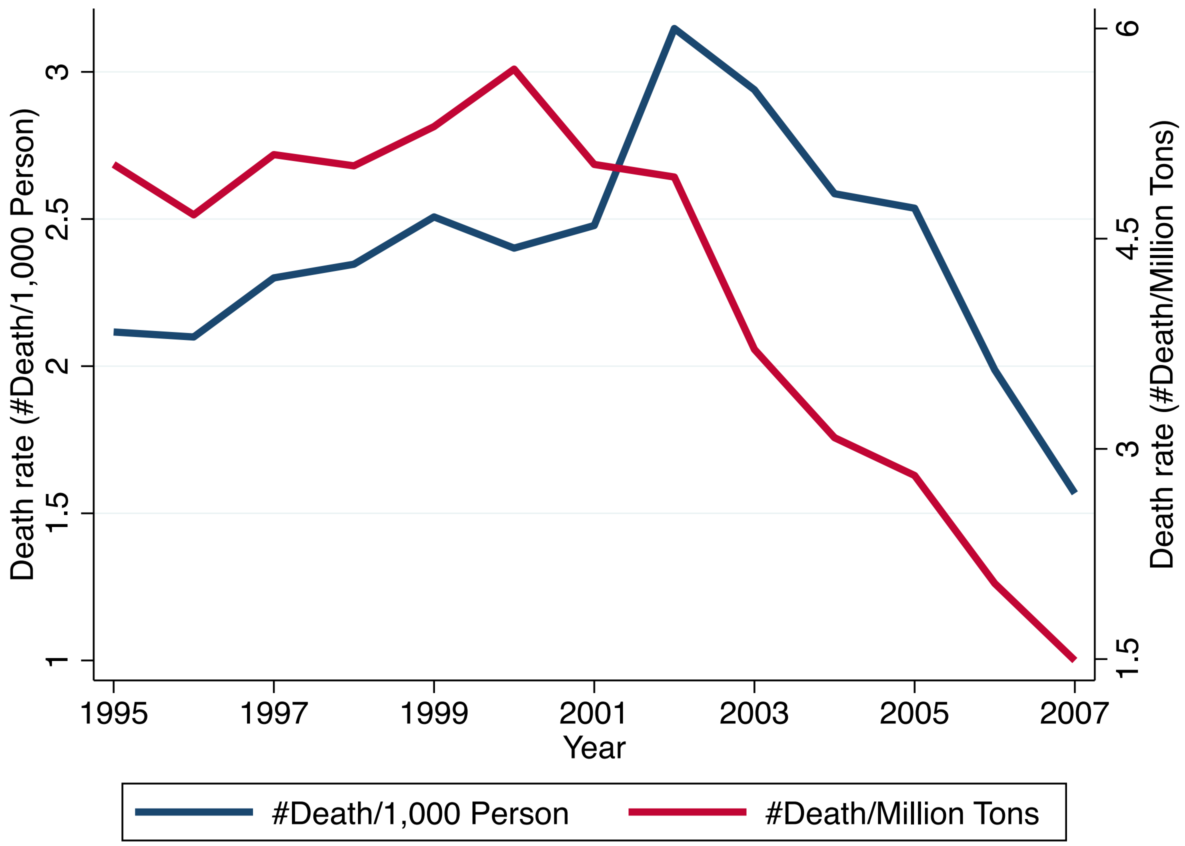


Figure A1: Evolution of Death Rates in the Coal Mining Industry in China: 1995-2007

Sources: The total number of workers in the coal mining industry is obtained from the ASIF database. Other information comes from the Compilation of Coal Statistics of China's Coal Industry.

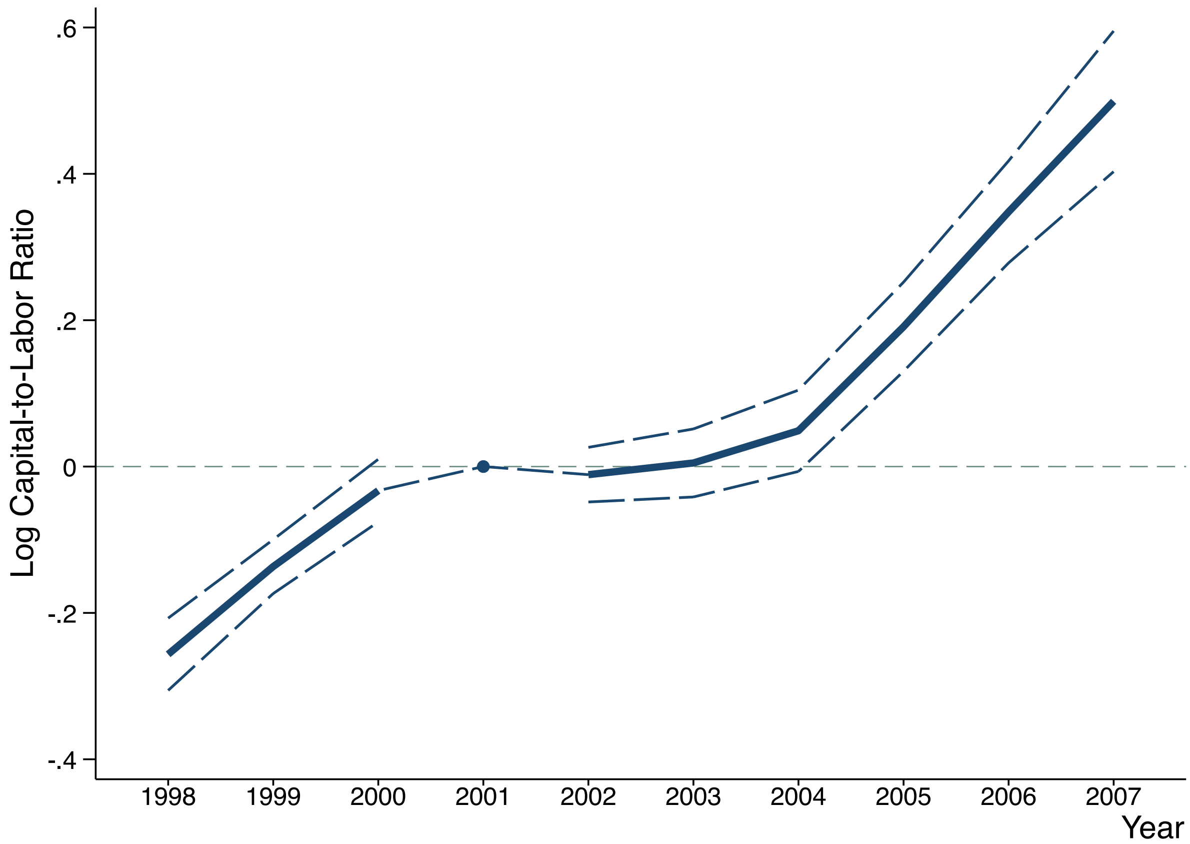
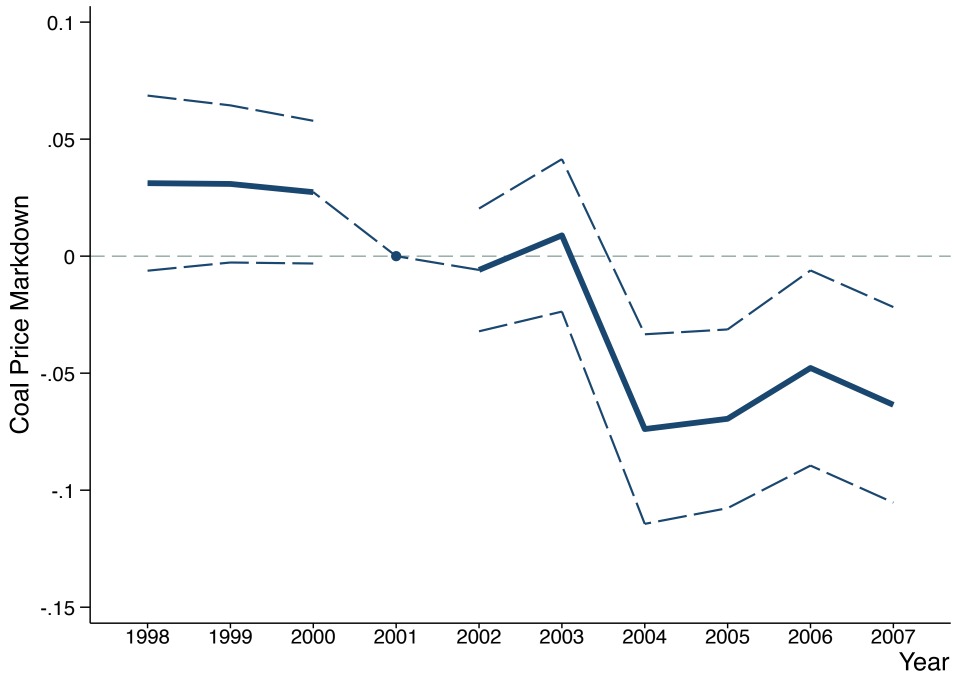
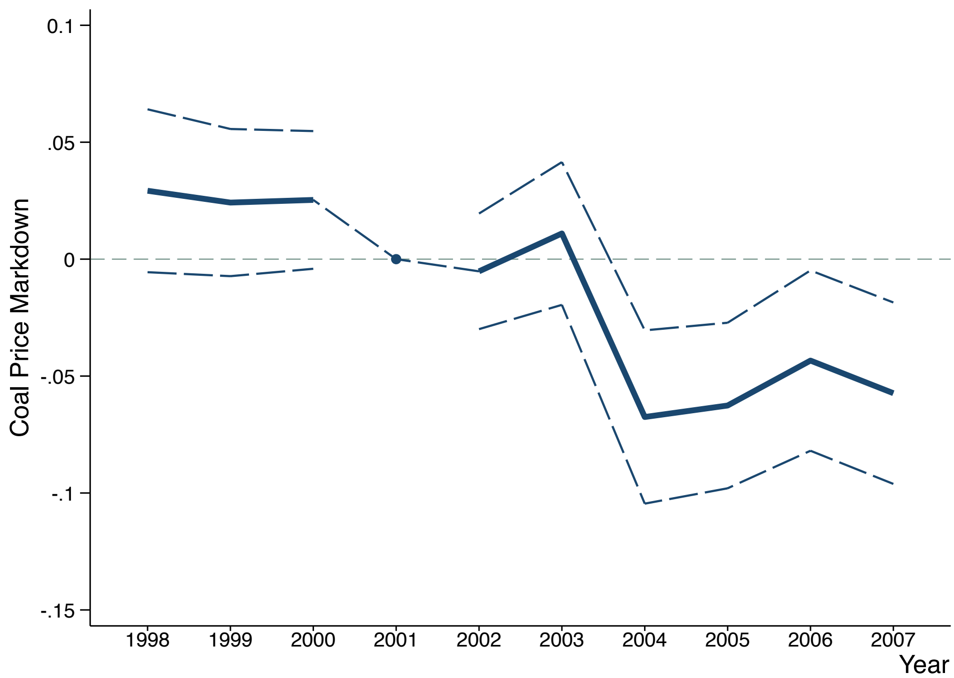


Figure A2: Evolution of Capital-to-Labor Ratio

Notes: I regress the capital-to-labor ratio on coal mine- and year-fixed effects, where standard errors are clustered at the prefectural level. The coefficients of different year dummies are plotted, where 2001, the year prior to the electricity sector restructuring announced, is normalized to zero.



(a) $\mu_{ft} = 1.1$



(b) $\mu_{ft} = 1.2$

Figure A3: Evolution of Coal Price Markdowns for Different Markup Calibrations

Notes: I normalize the year before the announcement of the electricity sector restructuring to zero with omitted confidence intervals. Standard errors are clustered at the prefectural level. 95% confidence intervals in dashed lines are shown. The evolution of coal price markdowns is very robust to different markup calibrations.

B Details of the Coal Production Process Across Mining Methods

Regardless of the specific technology employed, a typical coal extraction face involves a cyclical sequence of five primary steps: coal cutting, loading, transporting, roof supporting, and goaf stowing. Each step plays a crucial role in ensuring both coal production and safety. I now describe and summarize the details of each step conditional on different coal mining technologies following Yan et al. (2009).

Coal Cutting. This is the first phase of coal mining, where the coal seam is mechanically cut or blasted, depending on the technology adopted. Specialized equipment, such as shearers or continuous miners under conventional or fully-mechanized mining methods, is utilized to remove coal from the seam.

Blasting Mining. Coal is broken up using controlled explosives. Holes are drilled into the coal seam, and explosives are inserted and detonated, fragmenting the coal.

Conventionally-mechanized Mining. Coal is typically cut using semi-mechanized tools or shearers. This method involves more manual labor than fully-mechanized mining methods.

Fully-mechanized Mining. Continuous miners mechanically cut the coal from the seam. These large machines operate efficiently, and the coal is cut directly without explosives, making the process more controlled and safer.

Coal Loading. Once the coal is cut or blasted, it needs to be loaded onto conveyors or shuttle cars for transportation.

Blasting Mining. After blasting, the fragmented coal is manually loaded onto shuttle cars, conveyors, or haul trucks with shovels.

Conventionally-mechanized Mining. Using Conventionally-mechanized Mining methods, coal is loaded manually or with small loaders onto carts, conveyors, or trucks. This process is slower and more labor-intensive than in fully-mechanized mining.

Fully-mechanized Mining. Continuous loaders are often integrated with the mining machines, and coal is loaded directly onto conveyor belts without requiring separate loading equipment.

Coal Transporting. The coal is then transported out of the mining face via conveyor belts, rail systems, or haul trucks, depending on the mine's layout and scale.

Blasting Mining. The coal is transported using shuttle cars or haul trucks, and later conveyed to the surface via conveyor belts or rail systems.

Conventionally-mechanized Mining. After loading onto shuttle cars or manually operated carts, the coal is transported along fixed rail systems or conveyor belts.

Fully-mechanized Mining. The transport system is usually fully integrated with the cutting and loading machines. Conveyor belts directly transport coal from the face to the surface, allowing for continuous operation and higher efficiency.

Roof Supporting. After the coal is extracted, roof support becomes essential to prevent collapses. Pitwoods, hydraulic props, or self-advancing hydraulic supports are used to keep the roof stable and ensure miner safety.

Blasting Mining. After each blast, roof supports are installed. Pitwoods or hydraulic props are set up manually, and the roof support process may take longer as workers must enter the mining area after each blast.

Conventionally-mechanized Mining. Roof support is also typically manual or semi-mechanized. Workers install hydraulic props to support the roof after coal removal.

Fully-mechanized Mining. Automated and self-advancing hydraulic roof supports are installed as the shearer or continuous miner advances. These supports are more effective and provide real-time protection for workers and equipment. The roof support system can advance with the machinery, ensuring continuous protection.

Goaf Stowing. This involves filling the void left by coal extraction with material such as crushed stone or waste from the mining process. This step prevents subsidence and improves the stability of the remaining structure.

Blasting Mining. Goaf stowing is typically done manually, using waste rock or crushed stone brought in after each blasting cycle. This can be labor-intensive and time-consuming.

Conventionally-mechanized Mining. Goaf stowing is often performed by manually bringing in waste material to fill voids or with small mechanical aids.

Fully-mechanized Mining. Fully-mechanized mining often employs the caving method, which allows the roof to naturally cave in after coal has been extracted, gradually filling the goaf. This method is cost-effective, as it doesn't require additional materials for filling.

Overall, distinct engineering characteristics of different mining technologies result in heterogeneous input intensity patterns for coal mines. The blasting mining technology involves substantial material usage and high labor intensity in most mining processes, but it is easy to utilize, especially in regions with complex geological conditions. In contrast, the conventionally-mechanized coal mining technology improves efficiency by mechanizing coal cutting and loading. However, significant workers and materials are still needed to support the roof and stow the goaf manually, the same as employing blasting mining technology. Regarding the fully-mechanized coal mining technology, it mechanizes all relevant mining processes and hence requires minor labor inputs and is more efficient, though it demands substantial fixed costs.

C Data Appendix

Summary Statistics. Table C1 presents the summary statistics for the primary datasets I use in the paper. The full sample consists of 10,538 and 42,535 observations for power plants and coal mines during 1999-2007, respectively. Specifically, around 1,000 power plants and 4,000 coal mines are active annually. Nevertheless, not all firms report each variable in each year, especially for quantity-related input and output variables. Thus, I use all observations that report all required variables to estimate corresponding production functions and apply the estimates of output elasticities to the remaining observations.

Table C1: Summary Statistics of Power Plants and Coal Mines

	Unit	Mean	Std. dev.	p10	p50	p90
Panel A: Power plants						
Electricity Output	GWh	1,307	2,247	35	246	4,152
Real capital	1,000 RMB Yuan	764,340	1,955,757	17,711	151,161	2,210,997
Employment	person	758	1,736	97	408	1,708
Coal input	ton	696,325	1,020,771	34,194	212,866	2,094,626
Nominal wage	1,000 RMB Yuan	23,567	62,783	1,392	7,444	57,492
Nominal intermediate input	1,000 RMB Yuan	227,156	337,583	10,336	73,760	699,012
Nominal revenue	1,000 RMB Yuan	408,014	961,567	13,570	97,194	1,121,152
Panel B: Coal mines						
Coal Output	ton	188,544	415,207	21,900	77,623	359,234
Accident probability	%	98.52	2.42	95.36	99.63	99.99
Real capital	1,000 RMB Yuan	28,208	112,900	1,324	5,220	41,576
Employment	person	522	1,412	45	180	970
Nominal wage	1,000 RMB Yuan	6,138	17,802	360	2,031	10,744
Real wage	1,000 RMB Yuan	4,771	14,203	286	1,549	8,245
Nominal intermediate input	1,000 RMB Yuan	25,612	63,265	2,339	8,467	51,834
Real intermediate input	1,000 RMB Yuan	22,161	55,412	2,080	7,440	44,345

Notes: Summary statistics for main variables based on the full sample of 10,538 and 42,535 observations for power plants and coal mines, respectively, during the 1999-2007 period.

Credibility Evaluation of the Coal Mine Accident Data. Table C2 presents the summary statistics of coal mine accidents. After dropping duplicated observations, I retain 7,415 coal mine accidents from 2000 to 2007. I identify accident types by checking accident descriptions manually. Ultimately, it indicates that roof accidents accounted for China's most frequent coal mine accidents (49%), followed by gas accidents (21%), composing the top two accident types between 2000 and 2007. The snapshot of the coal mine accidents aligns well with public information and what I introduced in Section 1.

Table C2: Summary of Coal Mine Accidents

Accident Type	Number	Share (%)
Roof	3,651	49
Gas	1,552	21
Transport	802	11
Safety operation	500	7
Water inrush	361	5
Other	549	7
Total	7,415	100

Sources: State Administration of Workplace Safety.

In addition, Fisman and Wang (2017) document a "death ceiling" effect in reported deaths in China, where a sharp discontinuity in reported deaths at the ceiling was observed, suggesting the local bureaucrats' data manipulation. The incentive behind this was that if the accidental deaths exceeded certain ceilings, it would hinder government officials' promotion. Similar incentives in political promotion competition regarding safety and accidental deaths have also been corroborated in Shi and Xi (2018). Specifically, the death ceiling was assigned by the severity of the accident: "severe" for 3 or more deaths and "very severe" for 10 or more deaths. Hence, if there was data manipulation in #death in coal mine accidents, one would expect a sharp discontinuity in reported deaths around the ceilings of 2 and 9. Figure C1 presents the density curve of #death in coal mine accidents. As we can see, the density curve of #death in coal mine accidents is smooth in trend, and no abnormal discontinuity is observed around the ceilings of 2 and 9, suggesting a regular data reporting pattern.

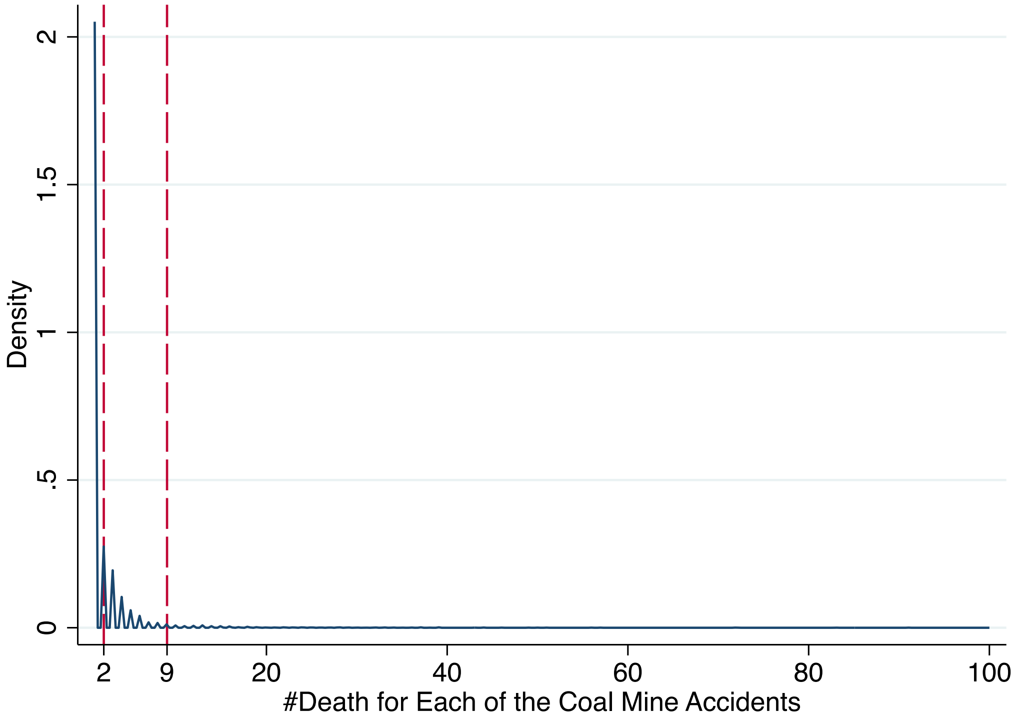


Figure C1: Density Curve of #Death in Coal Mine Accidents

Notes: Data source comes from the State Administration of Workplace Safety. I set the higher range bound to 100 to enable clearer visualization of the left tail. There are only four observations that report #death higher than 100.

D Proof of Proposition 1 and Corollary 1

Proposition 1. *The buyer power of downstream power plants induces upstream coal mines to adopt less capital, ultimately leading to lower capital-to-labor ratios for both production processes.*

Proof: The proof consists of two parts.

i) Comparative statics of an exogenous buyer power increase. By rewriting Eq.(5) and Eq.(6), one can see that the optimal levels of capital stocks, K^{q*} and K^{s*}, and labor, L^{q*} and L^{s*}, are given by

$$\frac{\partial Q}{\partial K^q} = \frac{r(\mathbf{b})}{\lambda}, \qquad -\frac{\partial A}{\partial K^s} = \frac{r(\mathbf{b})}{\kappa}, \qquad \frac{\partial Q}{\partial L^q} = \frac{w}{\lambda}, \qquad -\frac{\partial A}{\partial L^s} = \frac{w}{\kappa}. \tag{30}$$

Given $\frac{\partial r(\mathbf{b})}{\partial \mathbf{b}} > 0$, an increase in $\mathbf{b}$ leads to higher $r(\mathbf{b})$. Combining $\frac{\partial Q}{\partial X^q} > 0$, $\frac{\partial\left(\frac{\partial Q}{\partial X^q}\right)}{\partial X^q} < 0$, $\frac{\partial A}{\partial X^s} < 0$, and $\frac{\partial\left(\frac{\partial A}{\partial X^s}\right)}{\partial X^s} > 0$ ($X^q \in \{K^q, L^q\}$ and $X^s \in \{K^s, L^s\}$), an increased $r(\mathbf{b})$ unambiguously induces lower K^{q*} and K^{s*}. Hence, higher buyer power downstream leads to less adoption of capital upstream.

ii) The impact of buyer power downstream on capital-to-labor ratios upstream. Dividing first-order conditions of capital by that of labor in Eq.(30), one can infer capital-to-labor ratios for production and safety input from $\frac{r(\mathbf{b})}{w} = \frac{\partial Q}{\partial K^q} / \frac{\partial Q}{\partial L^q} = \frac{\partial A}{\partial K^s} / \frac{\partial A}{\partial L^s}$. Buyer power downstream, $\mathbf{b}$, only affects capital-to-labor ratios via influencing $r(\mathbf{b})$, where an increase in $\mathbf{b}$ leads to higher $r(\mathbf{b})$. For either $v \in \{q, s\}$, a higher $r(\mathbf{b})$ induce a higher $\frac{\partial Q}{\partial K^v}$, hence a lower K^v. Note that an increase in $\mathbf{b}$ leads to L^{s*} and L^{q*} unchanged. Hence, both conditions lead to a decrease in the capital-to-labor ratio of K^v/L^v.□

Corollary 1. *The buyer power of downstream power plants can induce more coal mining accidents upstream via increased cost of capital adoption.*

Proof: The effect of the downstream buyer power on the amount of accidents upstream is given by

$$\frac{\partial A}{\partial \mathbf{b}} = \underbrace{\underbrace{\frac{\partial A^*}{\partial K^s}}_{<0} \underbrace{\frac{\partial K^{s*}}{\partial r(\mathbf{b})}}_{<0} \underbrace{\frac{\partial r(\mathbf{b})}{\partial \mathbf{b}}}_{>0}}_{>0}.$$

The proofs are straightforward using **Proposition 1**.□

E Derivations

Expression for the Labor-Augmenting Productivity. Given the coal mine's (variable) cost minimization problem:

$$\min_{L_{jt},M_{jt}} P^L_{jt}L_{jt}+P^M_{jt}M_{jt} \quad \text{s.t.} \quad \{\beta_K K_{jt}^{-\frac{1-\sigma}{\sigma}}+[\exp(\omega_{Ljt})L_{jt}]^{-\frac{1-\sigma}{\sigma}}+\beta_M M_{jt}^{-\frac{1-\sigma}{\sigma}}\}^{\frac{\sigma\kappa}{\sigma-1}}\exp(\omega_{Hjt}) \geqslant \bar{Y}_{jt}.$$

One can consider the associated Lagrangian function for the coal mine:

$$\mathcal{L}_{jt} = P^L_{jt}L_{jt} + P^M_{jt}M_{jt} + \lambda_{jt}\left(\overline{Y}_{jt} - Y_{jt}\left(M_{jt}, L_{jt}, K_{jt}, \boldsymbol{\omega}_{jt}\right)\right), \tag{31}$$

where $\boldsymbol{\omega}_{jt} = (\omega_{Hjt}, \omega_{Ljt})$. The first-order conditions with respect to labor and material give to

$$\frac{\partial \mathcal{L}_{jt}}{\partial L_{jt}} = P^L_{jt} - \lambda_{jt}\frac{\partial Y_{jt}(\cdot)}{\partial L_{jt}} = 0, \tag{32}$$

$$\frac{\partial \mathcal{L}_{jt}}{\partial M_{jt}} = P^M_{jt} - \lambda_{jt}\frac{\partial Y_{jt}(\cdot)}{\partial M_{jt}} = 0. \tag{33}$$

Dividing Eq.(32) by Eq.(33), one can get

$$\frac{\partial Y_{jt}(\cdot)}{\partial L_{jt}} \Big/ \frac{\partial Y_{jt}(\cdot)}{\partial M_{jt}} = \frac{P^L_{jt}}{P^M_{jt}}. \tag{34}$$

Given $Y_{jt}(M_{jt}, L_{jt}, K_{jt}, \boldsymbol{\omega}_{jt}) = \left\{\beta_K K_{jt}^{-\frac{1-\sigma}{\sigma}} + [\exp(\omega_{Ljt})L_{jt}]^{-\frac{1-\sigma}{\sigma}} + \beta_M M_{jt}^{-\frac{1-\sigma}{\sigma}}\right\}^{\frac{\sigma\kappa}{\sigma-1}}\exp(\omega_{Hjt})$, I can derive out

$$\begin{aligned}\frac{\partial Y_{jt}(\cdot)}{\partial L_{jt}} &= \frac{\sigma\kappa}{\sigma-1}\left\{\beta_K K_{jt}^{-\frac{1-\sigma}{\sigma}} + [\exp(\omega_{Ljt})L_{jt}]^{-\frac{1-\sigma}{\sigma}} + \beta_M M_{jt}^{-\frac{1-\sigma}{\sigma}}\right\}^{\frac{\sigma\kappa}{\sigma-1}-1}\exp(\omega_{Hjt}) \\ &\qquad \times [\exp(\omega_{Ljt})]^{-\frac{1-\sigma}{\sigma}}\left(-\frac{1-\sigma}{\sigma}\right)(L_{jt})^{-\frac{1-\sigma}{\sigma}-1}, \end{aligned} \tag{35}$$

$$\begin{aligned}\frac{\partial Y_{jt}(\cdot)}{\partial M_{jt}} &= \frac{\sigma\kappa}{\sigma-1}\left\{\beta_K K_{jt}^{-\frac{1-\sigma}{\sigma}} + [\exp(\omega_{Ljt})L_{jt}]^{-\frac{1-\sigma}{\sigma}} + \beta_M M_{jt}^{-\frac{1-\sigma}{\sigma}}\right\}^{\frac{\sigma\kappa}{\sigma-1}-1}\exp(\omega_{Hjt}) \\ &\qquad \times \beta_M\left(-\frac{1-\sigma}{\sigma}\right)(M_{jt})^{-\frac{1-\sigma}{\sigma}-1}. \end{aligned} \tag{36}$$

Substituting Eq.(35) and Eq.(36) into Eq.(34), rearranging terms, the expression for the labor-augmenting productivity with observed variables and estimable parameters is given by

$$[\exp(\omega_{Ljt})]^{-\frac{1-\sigma}{\sigma}} = \beta_M\frac{P_{Ljt}L_{jt}}{P_{Mjt}M_{jt}}\left(\frac{M_{jt}}{L_{jt}}\right)^{-\frac{1-\sigma}{\sigma}}. \tag{37}$$

Deriving the First-Stage Estimating Equation. Plugging Eq.(18), the control function of the labor-augmenting productivity, into Eq.(17), rearranging terms, the production frontier to be estimated turns to be:

$$q_{jt} = -\alpha_s s_{jt} + \kappa \ln(M_{jt}) + \frac{\sigma\kappa}{\sigma-1} \ln\left(\frac{\beta_K}{\beta_M}\left(\frac{K_{jt}}{M_{jt}}\right)^{-\frac{1-\sigma}{\sigma}} + \left(1+\frac{P_{Ljt}L_{jt}}{P_{Mjt}M_{jt}}\right)\right) + \frac{\sigma\kappa}{\sigma-1}\ln\beta_M$$
$$+\omega_{Hjt} + \alpha_s\epsilon^s_{jt} + \epsilon^q_{jt}. \tag{38}$$

Substituting the control function of the Hicks-neutral productivity, $\omega_{Hjt} = h_t(k_{jt}, l_{jt}, m_{jt}, \frac{P_{Ljt}L_{jt}}{P_{Mjt}M_{jt}}, \mathbf{z}_{jt})$, into Eq.(38), one can rewrite the production frontier as:

$$q_{jt} = -\alpha_s s_{jt} + \kappa \ln(M_{jt}) + \frac{\sigma\kappa}{\sigma-1} \ln\left(\frac{\beta_K}{\beta_M}\left(\frac{K_{jt}}{M_{jt}}\right)^{-\frac{1-\sigma}{\sigma}} + \left(1+\frac{P_{Ljt}L_{jt}}{P_{Mjt}M_{jt}}\right)\right) + \frac{\sigma\kappa}{\sigma-1}\ln\beta_M$$
$$+h_t\left(k_{jt}, l_{jt}, m_{jt}, \frac{P_{Ljt}L_{jt}}{P_{Mjt}M_{jt}}, \mathbf{z}_{jt}\right) + \alpha_s\epsilon^s_{jt} + \epsilon^q_{jt}. \tag{39}$$

Substituting $\mathbf{x}_{jt} = \left\{k_{jt}, l_{jt}, m_{jt}, \frac{P_{Ljt}L_{jt}}{P_{Mjt}M_{jt}}\right\}$ into the above equation to save notations, the first-stage estimating equation as in the main text is:

$$q_{jt} = -\alpha_s s_{jt} + \phi_t(\mathbf{x}_{jt}, \mathbf{z}_{jt}) + \boldsymbol{\epsilon}_{jt},$$

where $\phi_t(\mathbf{x}_{jt}, \mathbf{z}_{jt}) = \kappa\ln(M_{jt}) + \frac{\sigma\kappa}{\sigma-1}\ln\left(\frac{\beta_K}{\beta_M}\left(\frac{K_{jt}}{M_{jt}}\right)^{-\frac{1-\sigma}{\sigma}} + \left(1+\frac{P_{Ljt}L_{jt}}{P_{Mjt}M_{jt}}\right)\right) + \frac{\sigma\kappa}{\sigma-1}\ln\beta_M + h_t(\mathbf{x}_{jt}, \mathbf{z}_{jt})$ and $\boldsymbol{\epsilon}_{jt} = \alpha_s\epsilon^s_{jt} + \epsilon^q_{jt}$.

References

Ackerberg, Daniel A and Jan De Loecker, "Production function identification under imperfect competition," 2024.

_, **Kevin Caves, and Garth Frazer**, "Identification properties of recent production function estimators," *Econometrica*, 2015, *83* (6), 2411–2451.

Aghion, Philippe, Nick Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt, "Competition and innovation: An inverted-U relationship," *The quarterly journal of economics*, 2005, *120* (2), 701–728.

Allcott, Hunt, Allan Collard-Wexler, and Stephen D O'Connell, "How do electricity shortages affect industry? Evidence from India," *American Economic Review*, 2016, *106* (3), 587–624.

Asker, John, Allan Collard-Wexler, and Jan De Loecker, "(Mis) allocation, market power, and global oil extraction," *American Economic Review*, 2019, *109* (4), 1568–1615.

Atkinson, Scott E, "Panel Data in Energy Economics," in "Panel Data Econometrics," Elsevier, 2019, pp. 495–519.

_ **and Daniel Primont**, "Stochastic estimation of firm technology, inefficiency, and productivity growth using shadow cost and distance functions," *Journal of Econometrics*, 2002, *108* (2), 203–225.

_ **and Rong Luo**, "Estimation of Production Technologies with Output and Environmental Constraints," *International Economic Review*, 2024, *65* (2), 755–780.

Autor, David H., David Dorn, and Gordon H. Hanson, "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American Economic Review*, October 2013, *103* (6), 2121–68.

Avignon, Rémi and Etienne Guigue, "Markups and markdowns in the french dairy market," Technical Report, Technical report, mimeo 2022.

Backus, Matthew, "Why is productivity correlated with competition?," *Econometrica*, 2020, *88* (6), 2415–2444.

Benkard, C Lanier, Ali Yurukoglu, and Anthony Lee Zhang, "Concentration in product markets," Technical Report, National Bureau of Economic Research 2021.

Berger, David, Kyle Herkenhoff, and Simon Mongey, "Labor market power," *American Economic Review*, 2022, *112* (4), 1147–1193.

Berkowitz, Daniel, Hong Ma, and Shuichiro Nishioka, "Recasting the Iron Rice Bowl: the evolution of China's state owned enterprises," *Review of Economics and Statistics*, 2014.

Berry, Steven, Martin Gaynor, and Fiona Scott Morton, "Do increasing markups matter? Lessons from empirical industrial organization," *Journal of Economic Perspectives*, 2019, *33* (3), 44–68.

Berry, Steven T, James A Levinsohn, and Ariel Pakes, "Automobile prices in market equilibrium," *Econometrica*, 1995, *63* (4), 841–890.

Biesebroeck, Johannes Van, "Productivity dynamics with technology choice: An application to automobile assembly," *The review of economic studies*, 2003, *70* (1), 167–198.

Bloom, Nicholas, Mirko Draca, and John Van Reenen, "Trade induced technical change? The impact of Chinese imports on innovation, IT and productivity," *The review of economic studies*, 2016, *83* (1), 87–117.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel, "Quasi-experimental shift-share research designs," *The Review of economic studies*, 2022, *89* (1), 181–213.

Brandt, Loren, Johannes Van Biesebroeck, and Yifan Zhang, "Creative accounting or creative destruction? Firm-level productivity growth in Chinese manufacturing," *Journal of development economics*, 2012, *97* (2), 339–351.

_ , _ , and _ , "Challenges of working with the Chinese NBS firm-level data," *China Economic Review*, 2014, *30*, 339–352.

_ , _ , Luhang Wang, and Yifan Zhang, "WTO accession and performance of Chinese manufacturing firms," *American Economic Review*, 2017, *107* (9), 2784–2820.

Caselli, Mauro, Arpita Chatterjee, and Shengyu Li, "Productivity and Quality of Multi-product Firms," *Unpublished paper*, 2024.

Charles, Kerwin Kofi, Matthew S Johnson, Melvin Stephens Jr, and Do Q Lee, "Demand conditions and worker safety: Evidence from price shocks in mining," *Journal of Labor Economics*, 2022, *40* (1), 47–94.

Chatterjee, Kalyan and William Samuelson, "Bargaining under incomplete information," *Operations research*, 1983, *31* (5), 835–851.

Che, Jiahua and Giovanni Facchini, "Dual track reforms: With and without losers," *Journal of Public Economics*, 2007, *91* (11-12), 2291–2306.

Chen, Bo, "Stress-induced trend: the clustering feature of coal mine disasters and earthquakes in China," *International Journal of Coal Science & Technology*, 2020, *7* (4), 676–692.

Chen, Hong, Qun Feng, Ruyin Long, and Hui Qi, "Focusing on coal miners' occupational disease issues: A comparative analysis between China and the United States," *Safety Science*, 2013, *51* (1), 217–222.

Cicala, Steve, "When Does Regulation Distort Costs? Lessons from Fuel Procurement in US Electricity Generation," *American Economic Review*, January 2015, *105* (1), 411–44.

_ , "Imperfect markets versus imperfect regulation in US electricity generation," *American Economic Review*, 2022, *112* (2), 409–441.

CLB, "Bone and Blood: The Price of Coal in China," Technical Report, China Labour Bulletin 2008. Accessed: September 22, 2024.

Collard-Wexler, Allan and Jan De Loecker, "Reallocation and technology: Evidence from the US steel industry," *American Economic Review*, 2015, *105* (1), 131–171.

Delabastita, Vincent and Michael Rubens, "Colluding against workers," *Available at SSRN 4208173*, 2024.

Demirer, Mert, "Production function estimation with factor-augmenting technology: An application to markups," *Job Market Paper*, 2022.

_ **and Ömer Karaduman**, "Do Mergers and Acquisitions Improve Efficiency? Evidence from Power Plants," Technical Report, National Bureau of Economic Research 2024.

Dhyne, Emmanuel, Amil Petrin, and Frederic Warzynski, "Deregulation and Investment Spillovers in Multi-Product Production Settings," 2022.

Döpper, Hendrik, Alexander MacKay, Nathan Miller, and Joel Stiebale, "Rising markups and the role of consumer preferences," *Harvard Business School Strategy Unit Working Paper*, 2024, (22-025).

Doraszelski, Ulrich and Jordi Jaumandreu, "Measuring the bias of technological change," *Journal of Political Economy*, 2018, *126* (3), 1027–1084.

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu, "How costly are markups?," *Journal of Political Economy*, 2023, *131* (7), 1619–1675.

Fabrizio, Kira R, Nancy L Rose, and Catherine D Wolfram, "Do markets reduce costs? Assessing the impact of regulatory restructuring on US electric generation efficiency," *American Economic Review*, 2007, *97* (4), 1250–1277.

Fisman, Raymond and Yongxiang Wang, "The Distortionary effects of incentives in government: evidence from china's "death ceiling" program," *American Economic Journal: Applied Economics*, 2017, *9* (2), 202–218.

Gandhi, Amit, Salvador Navarro, and David A Rivers, "On the identification of gross output production functions," *Journal of Political Economy*, 2020, *128* (8), 2973–3016.

Gao, Hang and Johannes Van Biesebroeck, "Effects of Deregulation and Vertical Unbundling on the Performance of C hina's Electricity Generation Sector," *The Journal of industrial economics*, 2014, *62* (1), 41–76.

Gaynor, Martin, Rodrigo Moreno-Serra, and Carol Propper, "Death by market power: reform, competition, and patient outcomes in the National Health Service," *American Economic Journal: Economic Policy*, 2013, *5* (4), 134–166.

Gentzkow, Matthew, Jesse M Shapiro, Frank Yang, and Ali Yurukoglu, "Pricing power in advertising markets: Theory and evidence," *American Economic Review*, 2024, *114* (2), 500–533.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift, "Bartik instruments: What, when, why, and how," *American Economic Review*, 2020, *110* (8), 2586–2624.

Gowrisankaran, Gautam, Ashley Langer, and Mar Reguant, "Energy transitions in regulated markets," Technical Report, National Bureau of Economic Research 2024.

_ , Charles He, Eric A Lutz, and Jefferey L Burgess, "Productivity, safety, and regulation in underground coal mining: Evidence from disasters and fatalities," Technical Report, National Bureau of Economic Research 2015.

Grieco, Paul LE and Ryan C McDevitt, "Productivity and quality in health care: Evidence from the dialysis industry," *The Review of Economic Studies*, 2017, *84* (3), 1071–1105.

_ , Charles Murry, and Ali Yurukoglu, "The evolution of market power in the us automobile industry," *The Quarterly Journal of Economics*, 2024, *139* (2), 1201–1253.

— , **Shengyu Li, and Hongsong Zhang**, "Input prices, productivity, and trade dynamics: long-run effects of liberalization on Chinese paint manufacturers," *The RAND Journal of Economics*, 2022, *53* (3), 516–560.

Guo, Jingpu, Qianming Zuo, Jie Zhou, and Zhimin Wang, "The dual-track system of coal prices: a lesson from history (in Chinese)," Technical Report, Cinda Securities Research and Development Center 2018. Accessed: September 14, 2024.

Hahn, Nadine, "Who Is in the Driver's Seat? Markups, Markdowns, and Profit Sharing in the Car Industry," *Markups, Markdowns, and Profit Sharing in the Car Industry (July 08, 2024). ZEW-Centre for European Economic Research Discussion Paper*, 2024, (24-047).

Hummels, David, Rasmus Jørgensen, Jakob Munch, and Chong Xiang, "The wage effects of offshoring: Evidence from Danish matched worker-firm data," *American Economic Review*, 2014, *104* (6), 1597–1629.

ILO, "Safety and health in underground coalmines," Technical Report, International Labor Organization 2006. Accessed: September 7, 2024.

— , "Mining: a hazardous work," Technical Report, International Labor Organization 2015. Accessed: September 7, 2024.

Imbert, Clement, Marlon Seror, Yifan Zhang, and Yanos Zylberberg, "Migrants and Firms: Evidence from China," *American Economic Review*, June 2022, *112* (6), 1885–1914.

IRENA, "Wind energy," Technical Report, International Renewable Energy Agency 2024. Accessed: September 14, 2024.

Jia, Ruixue and Huihua Nie, "Decentralization, collusion, and coal mine deaths," *Review of Economics and Statistics*, 2017, *99* (1), 105–118.

Joskow, Paul L, "Vertical integration and long-term contracts: The case of coal-burning electric generating plants," *The Journal of Law, Economics, and Organization*, 1985, *1* (1), 33–80.

— , "Price adjustment in long-term contracts: the case of coal," *The Journal of Law and Economics*, 1988, *31* (1), 47–83.

Klette, Tor Jakob and Zvi Griliches, "The inconsistency of common scale estimators when output prices are unobserved and endogenous," *Journal of applied econometrics*, 1996, *11* (4), 343–361.

Lam, Pun-Lee, "Pricing of electricity in China," *Energy*, 2004, *29* (2), 287–300.

Larsen, Bradley and Anthony Lee Zhang, "A mechanism design approach to identification and estimation," Technical Report, National Bureau of Economic Research 2018.

_ **and** _ , "Quantifying bargaining power under incomplete information: A supply-side analysis of the used-car industry," *Available at SSRN 3990290*, 2021.

Lau, Lawrence J, Yingyi Qian, and Gerard Roland, "Reform without losers: An interpretation of China's dual-track approach to transition," *Journal of political economy*, 2000, *108* (1), 120–143.

Levinsohn, James and Amil Petrin, "Estimating production functions using inputs to control for unobservables," *The review of economic studies*, 2003, *70* (2), 317–341.

Li, Hongbin and Li-An Zhou, "Political turnover and economic performance: the incentive role of personnel control in China," *Journal of public economics*, 2005, *89* (9-10), 1743–1762.

Li, Shengyu and Hongsong Zhang, "Does external monitoring from the government improve the performance of state-owned enterprises?," *The Economic Journal*, 2022, *132* (642), 675–708.

Li, Wei, "A tale of two reforms," *The RAND Journal of Economics*, 1999, pp. 120–136.

Liu, Quanlong, Xinchun Li, and Maureen Hassall, "Regulatory regime on coal Mine Safety in China and Australia: Comparative analysis and overall findings," *Resources Policy*, 2021, *74*, 101454.

Liu, Rulin, Weimin Cheng, Yanbin Yu, and Qingfeng Xu, "Human factors analysis of major coal mine accidents in China based on the HFACS-CM model and AHP method," *International journal of industrial ergonomics*, 2018, *68*, 270–279.

Loecker, Jan De, "Product differentiation, multiproduct firms, and estimating the impact of trade liberalization on productivity," *Econometrica*, 2011, *79* (5), 1407–1451.

_ , "Detecting learning by exporting," *American Economic Journal: Microeconomics*, 2013, *5* (3), 1–21.

_ **and Chad Syverson**, "An industrial organization perspective on productivity," in "Handbook of industrial organization," Vol. 4, Elsevier, 2021, pp. 141–223.

_ **and Frederic Warzynski**, “Markups and firm-level export status,” *American economic review*, 2012, *102* (6), 2437–2471.

_ **and Jan Eeckhout**, “Global market power,” Technical Report, National Bureau of Economic Research 2018.

_ **and Paul Scott**, “Markup Estimation using Production and Demand Data. An Application to the US Brewing Industry,” *Unpublished paper*, 2022.

_ **and Pinelopi Koujianou Goldberg**, “Firm performance in a global market,” *Annu. Rev. Econ.*, 2014, *6* (1), 201–227.

_ **, Jan Eeckhout, and Gabriel Unger**, “The rise of market power and the macroeconomic implications,” *The Quarterly Journal of Economics*, 2020, *135* (2), 561–644.

_ **,** _ **, and Simon Mongey**, “Quantifying market power and business dynamism in the macroeconomy,” Technical Report, National Bureau of Economic Research 2021.

_ **, Pinelopi K Goldberg, Amit K Khandelwal, and Nina Pavcnik**, “Prices, markups, and trade reform,” *Econometrica*, 2016, *84* (2), 445–510.

Loertscher, Simon and Leslie M Marx, “Merger review for markets with buyer power,” *Journal of Political Economy*, 2019, *127* (6), 2967–3017.

_ **and** _ , “Merger review with intermediate buyer power,” *International Journal of Industrial Organization*, 2019, *67*, 102531.

_ **and** _ , “Incomplete-Information Models for Industrial Organization,” 2021.

_ **and** _ , “Incomplete information bargaining with applications to mergers, investment, and vertical integration,” *American Economic Review*, 2022, *112* (2), 616–649.

Ma, Jinlong, “On-grid electricity tariffs in China: Development, reform and prospects,” *Energy policy*, 2011, *39* (5), 2633–2645.

Meng, Lingsheng, Yifan Zhang, and Yunbin Zhang, “Estimating Capital-Labor Substitution in China: Evidence from Firm-Level Data,” 2023.

Morlacco, Monica, “Market power in input markets: Theory and evidence from french manufacturing,” *Unpublished, March*, 2019, *20*, 2019.

MRC, “Railway Freight Rates Rules of the Ministry of Railways of the People’s Republic of China,” 2000. Tieyun[2000] §7.

Murray, John E and Javier Silvestre, "Small-scale technologies and European coal mine safety, 1850–1900," *The Economic History Review*, 2015, *68* (3), 887–910.

NEA, "The 10th Five-Year Plan for the Coal Industry," Technical Report, National Energy Administration of China 2001. Accessed: September 7, 2024.

_ , "The 13th Five-Year Plan for the Coal Industry," Technical Report, National Energy Administration of China 2016. Accessed: September 7, 2024.

Nerlove, Marc et al., *Returns to scale in electricity supply* 1961.

Nevo, Aviv, "Measuring market power in the ready-to-eat cereal industry," *Econometrica*, 2001, *69* (2), 307–342.

NMSA, "The 13th Five-Year Plan for Coal Mine Safety Production," Technical Report, National Mine Safety Administration of China 2016. Accessed: September 7, 2024.

Olley, G Steven and Ariel Pakes, "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 1996, *64* (6), 1263–1297.

Orr, Scott, "Within-firm productivity dispersion: Estimates and implications," *Journal of Political Economy*, 2022, *130* (11), 2771–2828.

Pavcnik, Nina, "Trade liberalization, exit, and productivity improvements: Evidence from Chilean plants," *The Review of Economic Studies*, 2002, *69* (1), 245–276.

Peng, Syd, "Understanding the Chinese Coal Industry," *Coal Age* August 2010. Accessed: September 7, 2024.

Ridder, Maarten De, Basile Grassi, and Giovanni Morzenti, "The Hitchhiker's Guide to Markup Estimation: Assessing Estimates from Financial Data," 2024.

Rubens, Michael, "Management, productivity, and technology choices: Evidence from US mining schools," *RAND Journal of Economics*, 2022.

_ , "Market structure, oligopsony power, and productivity," *American Economic Review*, 2023, *113* (9), 2382–2410.

_ , "Labor Market Power and Factor-Biased Technology Adoption," 2024.

_ , Yingjie Wu, and Mingzhi Jimmy Xu, "Exploiting or Augmenting Labor?," 2024.

Shi, Xiangyu and Tianyang Xi, "Race to safety: Political competition, neighborhood effects, and coal mine deaths in China," *Journal of Development Economics*, 2018, *131*, 79–95.

Shi, Xunpeng, "China's small coal mine policy in the 2000s: a case study of trusteeship and consolidation," *Resources Policy*, 2013, *38* (4), 598–604.

Shu, Yunyu, Ruozi Song, Bing Zhang, and Honghao Zheng, "Local Knowledge or Misallocation: Efficiency Costs of Discretion in Regulatory Enforcement," 2024.

Sicular, Terry, "Plan and market in China's agricultural commerce," *Journal of Political Economy*, 1988, *96* (2), 283–307.

Sider, Hal, "Safety and productivity in underground coal mining," *The Review of Economics and Statistics*, 1983, pp. 225–233.

Syverson, Chad, "Macroeconomics and market power: Context, implications, and open questions," *Journal of Economic Perspectives*, 2019, *33* (3), 23–43.

Treuren, Leonard, "Wage markups and buyer power in intermediate input markets," 2022.

Valmari, Nelli, "Estimating production functions of multiproduct firms," *Review of Economic Studies*, 2023, *90* (6), 3315–3342.

Wang, Bing, "An imbalanced development of coal and electricity industries in China," *Energy Policy*, 2007, *35* (10), 4959–4968.

Wang, Hongying and Nobuhiro Horii, "China's Energy Market and Pricing System Reform - Case Study of Shanxi Province's Pilot Reform," Technical Report, Institute of Developing Economies 2008. Accessed: September 22, 2024.

Wong, Horng Chern, "Understanding high-wage and low-wage firms," *Available at SSRN 3446088*, 2021.

Wright, Tim, "The political economy of coal mine disasters in China:"your rice bowl or your life"," *The China Quarterly*, 2004, *179*, 629–646.

_ , *The political economy of the Chinese coal industry: black gold and blood-stained coal*, Routledge, 2012.

_ , "The Political Economy of China's Dramatically Improved Coal Safety Record," *The China Quarterly*, 2022, *249*, 91–113.

Yan, Haipeng, Jiangning Huang, and Yi Liu, *Caimei Gongyi [Coal Mining Technology]*, Jiangsu, China: China University of Mining and Technology Press, 2009.

Yang, Chi-Jen, Xiaowei Xuan, and Robert B Jackson, "China's coal price disturbances: Observations, explanations, and implications for global energy economies," *Energy Policy*, 2012, *51*, 720–727.

Yang, Qing, Lei Zhang, and Xin Wang, "Dynamic analysis on market structure of China's coal industry," *Energy Policy*, 2017, *106*, 498–504.

Yeh, Chen, Claudia Macaluso, and Brad Hershbein, "Monopsony in the US labor market," *American Economic Review*, 2022, *112* (7), 2099–2138.

Zhang, Hongsong, "Non-neutral technology, firm heterogeneity, and labor demand," *Journal of Development Economics*, 2019, *140*, 145–168.