Path Dependence and Structural Change in U.S. Electricity Production

Kyle C. Meng*

Abstract

Understanding the dynamics of long-run fuel switching in the electricity sector is crucial for climate policy. This paper quantifies the role of path dependence - the lagged effects of transitory price shocks - on the development of midwestern U.S. coal-fired electricity capacity over the 20th century. I exploit variation in county coal prices following the introduction of mechanized mining which allowed extraction of previously inaccessible coal resources. I find that transitory coal price shocks have increasing effects on relative coal capacity over the subsequent ten decades, influencing two generations of new electricity infrastructure. I develop and test a model of electricity production nesting competing drivers of path dependence and find empirical support for scale over productivity effects. Interpreted through this model, my reduced-form estimates imply a long-run elasticity of substitution between coal and non-coal electricity of 4. Simulations show a future long-term decline in U.S. electricity CO_2 emissions is possible if recent natural gas prices are sustained for at least two decades.

*Bren School and Dept. of Economics, University of California, Santa Barbara, email: kmeng@bren.ucsb.edu. I thank Dan Benjamin, Javier Birchenall, Hoyt Bleakley, Jonathan Dingel, Walker Hanlon, Akshaya Jha, Per Krusell, Gary Libecap, Suresh Naidu, Trevor O'Grady, Maxim Pinkovskiy, Bernard Salanié, Steve Salant, Anna Thompsett, Inge van den Bijgaart, Tom Vogl, and seminar participants at the NBER DAE Summer Institute, PERC, University of Gothenburg, and University of San Francisco for helpful comments. I am grateful to Kira Fabrizio, Josh Lewis, Jessika Trancik, and Catherine Wolfram for sharing data. Kayleigh Campbell Bierman and Andy Bilich provided excellent research assistance. Errors are my own. This version: March, 2016.

1 Introduction

The regulation of greenhouse gas emissions causing anthropogenic climate change is among the most important policy challenges today. To address climate change, it is widely recognized that the global economy must undergo a structural transition away from electricity produced by coal, the most carbon intensive fossil fuel. Economic analyses of climate policy offer two views on how such a transition could be induced.

Traditional models of climate policy typically require *permanent* externality pricing to counter the relative abundance of coal (Nordhaus, 1994). Recent models featuring structural change demonstrate that in the presence of strong transitional dynamics, a large but *temporary* policy intervention may be sufficient to trigger a long-term emissions decline, overcoming coal's supply advantage (Acemoglu et al., 2012, 2014). While such dynamics can result from different mechanisms, they are broadly defined by whether transitory shocks have lasting impact, leading to effects that exhibit path dependence. Unfortunately, to date there is limited empirical evidence on long-run transitional dynamics in the electricity sector which would lend support to either view.

This paper estimates the strength of path dependence in coal-fired electricity across the 20th century over mid-western U.S. counties. I find that county-level coal price shocks have increasing effects on relative coal-fired electricity capacity[1] over the subsequent ten decades, covering two generations of new electricity infrastructure. To interpret my results structurally, I develop a model of electricity production nesting two potential drivers of path dependence featured prominently in the literature, returns to scale and productivity effects. Additional empirical tests refine the model by detecting scale effects and allow a mapping of my reduced-form estimates onto the long-run elasticity of substitution between coal and non-coal electricity, a common structural parameter across a class of multi-sector models of structural change. Recovery of this parameter allows this paper to inform upon settings where fuel switching may occur due to alternative drivers of path dependence.

Understanding electricity sector transition dynamics is important for three reasons. First, at the global level, the electricity sector contributes a large portion of greenhouse gas emissions, emitting 30% of global emissions since 1970 (International Energy Agency, 2014). Thus, any effective climate policy must directly address the electricity sector, particularly in countries that are currently coal-dependent.[2] Second, it is commonly argued that the U.S. economy is carbon intensive because of its world leading coal resources. Detection of path dependence offers an additional explanation based on past circumstances that dictate historical coal access. Third, the U.S. electricity sector is currently undergoing a potential structural transition from coal to natural gas following the introduction of hydraulic fracturing. The dynamics behind the initial expansion into coal-fired electricity at the start of the 20th century may help inform upon this upcoming transition.

Causal estimates of path dependence in coal-fired electricity capacity requires variation in coal prices satisfying three statistical properties. First, coal prices must be uncorrelated with unobserved contemporaneous and subsequent determinants of relative coal capacity. Second, price variation must be transitory or obsolete such that it does not directly affect subsequent relative coal capacity (Bleakley and Lin, 2012). This second

[1] Capacity is the maximum amount of electricity that can be produced (in watts) by that generator. It is typically used to describe the size of a electricity generator and provides broadly a measure of electricity producing machine or capital stock. I focus on electricity capacity rather than generation due to data limitation issues described in Section 4. Relative capacity is the ratio of coal to non-coal capacity.

[2] Outside of climate change concerns, coal burning has been long associated with social costs. Local pollution from domestic coal use was one of the examples of externalities discussed in Coase (1960). More recently, studies have shown that conventional air pollution from coal burning has been indirectly (Chay and Greenstone, 2003, 2005) and directly (Barreca, Clay and Tarr, 2014) linked with a variety of health risks.

condition ensures that lagged effects can be interpreted as path dependence as distinguished more generally from persistence, or the contemporaneous effects of an ongoing treatment. Third, because electricity capital is highly durable, detecting lagged effects on new capital investments require data covering many decades. To meet these empirical requirements, I use recently digitalized 3-dimensional spatial data characterizing the local and depth of U.S. coal resources. In particular, I exploit spatial changes in county-level delivered coal prices following the introduction of mechanized mining technology which opened mines over previously inaccessible deep coal resources around the start of the 20th century.

The construction of a county-by-decade panel of delivered coal prices from 1890-1990 follows the logic of price determination in Gaudet, Moreaux and Salant (2001) which extends the canonical Hotelling location model (Hotelling, 1929) for a setting where county and coal resources are located arbitrarily in a 2-dimensional space. In this setting, the delivered coal price faced by a county is defined as the straight-line distance to the nearest mine. For a given county, the opening of a deep mine due to mechanization alters the delivered coal price if it becomes the nearest mine to that county. As a consequence, the previous price-setting shallow mine now becomes obsolete to that county. I call this the obsolescence event with my treatment of interest being distance to the obsolete mine. Using this setting, I estimate an event study, asking whether distance to the obsolete mine continues to influence a county's coal capacity in the decades after obsolescence.

This data construction addresses my statistical requirements. First, my measure of delivered coal prices is driven entirely by the interaction between the location and depth of subsurface coal resources and the aggregate expansion of mechanized mining. It is likely orthogonal to local contemporaneous and subsequent determinants of coal demand. Second, my panel data setting allows possible isolation of path dependent effects from time-varying and time-invariant contemporaneous determinants by allowing for state-level time trends, and county fixed effects. Furthermore, because deep coal mines opened at different times, counties experience the obsolescence event at different moments providing assurance, through the use of decade fixed effects, that the timing of obsolescence itself is not endogenous. Third, because my data covers the entirety of the 20th century, I am able to estimate lagged effects over a time horizon relevant for capital investment decisions.

I find that distance to the obsolete mine has a strengthening effect on cumulative relative coal-fired capacity for up to ten decades after the obsolescence event with notable jumps two and seven decades later of -2.3% and -3.7%, respectively. Importantly, these jumps coincide with the expected timing of two generations of new electricity capacity. By comparison, these lagged effects are 3 and 4 times stronger respectively than the contemporaneous price effect. I demonstrate that this evidence of path dependence is robust to various model specifications, functional form assumptions, estimation samples, data construction and other modeling choices. As a check on my identifying assumptions, I further show that distance to an obsolete mine is uncorrelated with pre-event trends in observable county characteristics that may directly affect relative coal capacity. I also validate my constructed measures of county coal prices and fuel-specific capacity against various available historical data over the 20th century.

Path dependence may emerge from various dynamic processes. From a theoretical perspective, path dependence in electricity capacity occurs whenever past relative fuel prices shift the relative marginal product electricity capacity in subsequent periods. Prior literature emphasizes two potential mechanisms: increasing returns to scale (Nerlove, 1963; Christensen and Greene, 1976) and the accumulation of fuel-specific productivity (Acemoglu et al., 2012, 2014). In particular, the thermodynamic properties of power plants provide a physical basis for the presence of increasing returns to scale. To examine these competing mechanisms, I introduce a theoretical framework nesting three features: vintaged fuel-specific electricity capacity with fixed

factor proportions (Komiya, 1962; Solow et al., 1966; Joskow, 1985, 1987; Atkeson and Kehoe, 1999), increasing returns to scale for fuel-specific capacity (Nerlove, 1963; Christensen and Greene, 1976), and imperfect substitutability between electricity from coal and other fuels (Acemoglu et al., 2012). Using predictions from this framework, I test for the presence of scale and productivity effects and find empirical support for scale effects. I also conduct an additional set of tests that fail to detect the role of other mechanisms. These include tests of scale effects in the residential and manufacturing sectors that consume electricity (Kline and Moretti, 2014; Severnini, 2014) and in the coal transport sector (Redding, Sturm and Wolf, 2011).

I further utilize my theoretical framework to formalize definitions of the strength of path dependence due to scale effects. I show that lagged effects which increase over time, or strong path dependence, occurs either when there is sufficiently strong increasing returns to scale or sufficiently high degree of substitutability between coal and non-coal electricity. Using this definition together with my estimated path dependent effects and scale parameter, I recover a value of 4 for the long-run elasticity of substitution.

This parameter appears across a broad class of multi-sector models of structural change driven by supply-side forces, first posited by Baumol (1967). Prominent models in this literature consider structural change driven by sector-level differences in total factor productivity (Ngai and Pissarides, 2007), factor proportions and capital deepening (Acemoglu and Guerrieri, 2008), and directed technical change (Acemoglu, 2002; Acemoglu et al., 2012). As an example of how this paper informs upon these other structural change models, my elasticity value implies that a temporary policy intervention is sufficient to avoid climate change disaster within the context of Acemoglu et al. (2012)'s model of directed technical change.

As such, this paper provides a counterintuitive policy implication. Whereas historical coal access may explain part of the current reliance on coal for U.S. electricity, the same dynamics could take hold for less carbon intensive electricity provided that a sufficient "boost" is provided to electricity produced by other fuels. This particular policy implication for the electricity sector parallels the logic more broadly behind "big-push" theories of development (Rosenstein-Rodan, 1943; Murphy and Shleifer, 1989; Costas Azariadis, 1990). Of particular relevance is the current transition from coal to natural gas-fired electricity in the U.S. To explore these implications, I simulate future U.S. electricity sector CO_2 emissions using my recovered structural parameters. Simulations show that recent natural gas prices can induce a long-run emissions decline provided that they are sustained for at least two decades.

This paper is linked with several bodies of literature. There is a long tradition in the economics of technological change arguing that dynamic processes may differentially "lock-in" certain technologies over time (Schumpeter, 1942; Schmookler, 1966; David, 1985; Aghion and Howitt, 1992; Arthur, 1994). Significant costs may arise if such dynamics lead to a time path of resource allocation that is socially inefficient. Coal-fired electricity is one example in which historical drivers of coal-fired electricity may lead to a socially suboptimal path of greenhouse gas emissions.

An emerging literature examines the empirics of short-run fuel switching. Aghion et al. (2014) explore firm-level patent differences in clean and dirty automobile technology across 80 countries in response to annual oil price changes from 1986-2005, finding evidence of path dependence by demonstrating that a firm's current patenting activity responds to its stock of patents. Knittel, Metaxoglou and Trindade (2015) examine how coal to natural gas fuel-switching from recent contemporaneous price shocks may differ depending on the ownership and market structure of the electricity producer. My analysis has the advantage of exploring transition dynamics over a longer time horizon that is relevant for capital investment decisions. This is particularly important for climate policy given the durability of electricity producing capital.

In terms of identification strategy, this paper connects with a growing empirical literature examining

dynamic effects following plausibly exogenous changes to the supply of a natural resource. This literature is generally classified by the nature of the treatment, whether the resource of interest goes into or out of productive use. The adjustment literature explores lagged effects following a persistent change in resource access (Nunn and Qian, 2011; Hornbeck and Keskin, 2014). The path dependence literature, which this paper follows, explores lagged effects following the obsolescence of a previously productive natural resource (Bleakley and Lin, 2012; Dell, 2012; Hornbeck, 2012; Severnini, 2014; Hornbeck and Naidu, 2014; Glaeser, Kerr and Kerr, 2015).

The remainder of the paper is organized as follows. Section 2 provides some motivating facts about the history of the U.S. electricity sector and introduces several relevant physical attributes of electric power plants. Section 3 details the statistical challenges in identifying path dependence and the proposed solution. Section 4 covers data construction and associated verification checks. Section 5 presents my reduced-form evidence of path dependence along with a variety of robustness checks. Section 6 introduces a theoretical framework to guide an empirical exploration of potential mechanisms. Section 7 uses this framework to formally define weak and strong path dependence, recover an implied elasticity of substitution, and simulate future emission pathways following fuel switching induced by recent natural gas prices. Section 8 concludes the paper.

2 Background

This section serves two purposes. I first provide some prima facie evidence for path dependence in the U.S. electricity sector. I then discuss the physical attributes of electric power plants that will inform the theory considered in Section 6 and provide a basis for the presence of path dependence.

2.1 Prima facie: Why are U.S. greenhouse gas emissions so high?

The United States is the world's largest cumulative emitter of man-made carbon dioxide[3] and was, until recently, also the largest annual emitting nation (Boden and Andres, 2013). U.S. carbon dioxide emissions are high even after accounting for aggregate economic activity. Figure A.1 plots country-level carbon dioxide emissions per capita against GDP per capita in 1960 and 2000 for all non-OPEC countries. U.S. carbon dioxide emissions per capita was a positive 2.4 and 1.8 σ outlier in 1960 and 2000 respectively relative to countries of similar per capita income. The U.S.'s reliance on coal for electricity, the most carbon intensive of energy inputs[4], is the primary reason for its high emissions. Currently, the electricity sector accounts for 38% of U.S. emissions of which 40% of generating capacity primarily burns coal (Energy Information Administration, 2013).

The traditional explanation for the U.S.'s heavily reliance on coal for electricity is that it has the world's largest coal endowment. Figure A.2 ranks countries by modern known coal reserves showing that the U.S. has roughly double the coal reserves of the next most coal abundant country, China (BP, 2014). The important role of endowments is supported by a first-look at county-level data. Figure 1 plots the county-level log relative coal capacity, the ratio of capacity using coal to that using other fuels, in the 1990's against log relative coal capacity in the 1920's. Consistent, with the time-invariant endowment argument, counties that

[3]Carbon dioxide is the most abundant greenhouse gas in the atmosphere and is the primary proximate cause of anthropogenic climate change.

[4]According to the U.S. Energy Information Agency, bituminous coal, the most common type of coal for electricity, yields 206 lbs of CO_2 per million BTU. By contrast, the CO_2 emission factor for gasoline and natural gas is 157 and 117 lbs of CO_2 per million BTU respectively. Available: `http://www.eia.gov/tools/faqs/faq.cfm?id=73&t=11`

have more relative coal capacity in the 1920's also tend to have more relative coal capacity in the 1990's, though relative coal capacity in the 1920's explains only 4% of the variation in relative coal capacity in the 1990's

To explore whether time-varying determinants may play a role, Figure 2 displays the average trajectories of relative coal capacity over the 20th century for counties that are close to (between 0 and 50 miles) and far away (between 200 and 250 miles) from time-invariant coal resources in the Illinois Basin. Counties that are closer to coal resources consistently have higher relative coal capacity compared to counties that are further away. However, the difference in coal use between these sets of counties grow dramatically over the 20th century suggesting that time-varying factors may play an important role.

The time evolution of relative coal capacity for counties close to and far from coal resources shown in Figure 2 is consistent with the presence of path dependence. However, this evidence is suggestive as the pattern displayed in Figure 2 could result from other explanations. For example, it may be that distance to time-invariant coal resources never becomes obsolete but instead exert persistent and increasing contemporaneous effects on relative coal capacity over the 20th century. Section 3.1 will detail the statistical requirements for identifying lagged effects that exhibit path dependence as separate from persistence.

2.2 Features of an electric power plant

A typical fossil fuel power plant converts fuel, such as coal, into electricity through three physical transformations. The fuel is burned in a boiler to generate heat, typically in the form of steam. Next, heat is converted to mechanical work through a steam turbine, which, in the final stage, is converted to electricity via a electricity generator. A generating unit, of which several may exist in a power plant, typically refers to a physically linked boiler, turbine, and generator. Capacity refers to the maximum electricity that can be produced by a single generating unit and is a common measure for the overall capital of a generating unit. I now turn to three features of electricity production that are pertinent for this analysis and the model introduced in Section 6.

Fixed input proportions While some boilers can burn coal, oil, or natural gas, most boilers are designed for a specific fuel and input volume (Avallone, Baumeister and Sadegh (2006), p.875). Burning alternative fuels or the intended fuel at different volumes can result in large efficiency losses (Avallone, Baumeister and Sadegh (2006), p.871). This implies that boiler capital, once built, is generally not substitutable with other inputs: production occurs with fixed input proportions. This "clay"-like nature of boilers has long been recognized in economics (Komiya, 1962; Joskow, 1985, 1987; Atkeson and Kehoe, 1999).

Returns to scale due to thermodynamics and capital modularity Two features of power plants provide a physical basis for increasing returns to scale in electricity production. First, the overall efficiency of a turbine in converting heat to mechanical work is broadly governed by the Second Law of Thermodynamics which states that no system operating in a closed cycle can convert heat entirely into work. Conversion efficiency is dictated by the following relationship from a Carnot heat engine:

$$\eta = 1 - \frac{T_{outside}}{T_{steam}}$$

where $T_{outside}$ is the outdoor temperature and T_{steam} is the steam temperature. This relationship states that the conversion of heat to work increases concavely with steam temperature. Steam temperature from

a boiler, in turn, is determined by the heat content of the fuel and the rate at which the fuel is fed into the boiler. Larger boilers burn more fuel and generate higher steam temperatures. Thus, thermal efficiency is tightly linked with boiler size.[5]

The second relevant feature is the modularity of power plant capital: boilers can provide steam to multiple turbines and generators. As new generators are added, larger boilers are built to serve both new and existing generators, providing efficiency gains to both old and new capital. This spillover effect, together with larger boilers being more thermally efficient, implies the presence of increasing returns to scale.[6]

Substitutability across fuel-specific electricity Electricity produced by different fuels often has distinct properties, making them gross but imperfect substitutes. For example, coal-fired generating units provides a steady supply of base load generation which can meet baseline electricity demand. However, coal-fired boilers cannot be easily ramped up or down in response to fluctuations in short-run electricity demand. In contrast, natural gas or diesel-fired generating units can be adjusted rapidly in response to real-time demand. Similarly, renewable energy sources, may have different reliability profiles given the intermittency of the primary energy supply.

3 Empirical context

This section contains two parts. First, I detail the statistical requirements required for identifying path dependence in coal-fired electricity capacity. Next, I discuss how the introduction of mechanized mining made deep coal resources mineable over time and can serve as a basis for an identification strategy following the logic of Hotelling's location model (Hotelling, 1929) as generalized by Gaudet, Moreaux and Salant (2001).

3.1 Empirical challenges: requirements for identifying path dependence

I present a simple reduced-form statistical model to clarify the statistical requirements needed for identifying path dependence. There are two periods, $t = 1, 2$. For county i during decade t, w_{cit} is the delivered coal price and X_{cit} is coal-fired capacity. Suppose each period has the following data generating process:

$$\begin{aligned} X_{ci1} &= \beta_1 w_{ci1} + \mu_{ci1} \\ X_{ci2} &= \rho X_{ci1} + \beta_1 w_{ci2} + \beta_2 w_{ci1} + \mu_{ci2} \end{aligned}$$

where β_1 and β_2 are the contemporaneous and lagged effects of coal prices on coal capacity respectively. In particular, β_2 captures the persistent effect of past coal prices. ρ is the path dependence parameter of interest, the indirect impact of past coal prices on future coal capacity through past capacity. Inserting the first equation into the second yields:

$$X_{ci2} = (\rho\beta_1 + \beta_2) w_{ci1} + \beta_1 w_{ci2} + \rho\mu_{ci1} + \mu_{ci2}$$

[5] There is also a fixed cost argument for why larger boilers yield higher stem temperatures. Reaching high temperatures require precision engineering and tight process control, which is often cost-prohibitive unless the size of the boiler reaches a certain level.

[6] Formally, suppose electricity Y is produced using capital via the production function, $Y(X) = X$. A doubling of capacity increases the efficiency of all capital units such that $Y(2X) = 2bX$ where $b > 1$ is the efficiency improvement. Thus $Y(2X) > 2Y(X)$, satisfying increasing returns to scale.

Identifying ρ requires several assumptions. First, current and lagged coal prices must be exogenous, or $E[w_{ci1}(\rho\mu_{ci1} + \mu_{ci2})|w_{ci2}] = E[w_{ci2}(\rho\mu_{ci1} + \mu_{ci2})|w_{ci1}] = 0$. Second, past coal prices must be transitory and thus not have a direct effect on current coal capacity. This obsolescence assumption is $\beta_2 = 0$. Without this assumption, it would be impossible to empirically distinguish the effects of path dependence from that of persistence. Additionally, the setup so far assumes every county experiences the obsolescence event after $t = 1$. If this common event is correlated with unobserved common shocks, one would further need a data setting in which the moment of obsolescence differs across counties and is also uncorrelated with the error term.

Finally, there is the difficulty of obtaining historical county-level delivered coal prices. Because I am estimating lagged effects on highly durable electricity capital, detecting path dependent effects requires obtaining price data ideally around the start of the 20th century. Unfortunately, such data only starts in the 1970's (see details in Section 4.1), with earlier data either never collected or no longer available.[7]

3.2 Proposed solution: spatial variation in coal prices from mechanized mining

To overcome these statistical and data collection challenges, I construct a proxy measure county-level delivered coal prices over the 20th century by exploiting spatial changes in county coal prices following the introduction of mechanized mining technology which opened mines over previously inaccessible deep coal resources. Construction for this measure follows the price-determination logic of **?** 2-dimensional spatial extension of the canonical Hotelling location model. This approach allows me to reconstruct a long panel of county-level delivered coal prices over the 20th century using a combination of spatial data indicating the location and depth of coal resources and data on the location and opening date of coal mines

Opening of deeper coal following mechanized mining In geology, a distinction is made between coal resources and reserves. Resources are all known underground coal deposits while reserves are the subset of resources which can be economically mined using existing mining technologies. Prior to the 20th century, most coal in the U.S. was manually extracted which limited the spatial extent of coal reserves primarily over near-surface resources less than 200 feet from the surface (Fisher, 1910; Speight, 1994).

Manual mining made way for mechanized extraction around the turn of the century and eventually came to dominate coal mining. Figure A.3 shows that nearly all increases in the mining of bituminous coal, the type of coal most commonly burned for electricity, from 1890 to 1930 occurred as a result of mechanized extraction. Chief among the benefits of mechanization was the introduction of mechanized cutting drills which allowed for the cost effective excavation of deeper coal resources that was previously inaccessible. In particular, Speight (1994) notes that coal resources deeper than 200 feet from the surface, which was previously unaccessible using manual techniques were, following mechanization, reclassified as reserves leading to many new mine openings. As I discuss next, the opening of mines over deeper coal resources following the introduction of mechanization altered the spatial distribution of delivered coal prices.

Defining county coal prices and obsolescence via Hotelling's location model I follow the logic of price-determination in Gaudet, Moreaux and Salant (2001) which extends Hotelling's location model (Hotelling, 1929) to construct a county-by-decade panel of delivered coal prices using spatial data on the

[7] County coal producer prices are recorded in the U.S. Geological Survey's Bureau of Mine reports "Mineral resources of the United States 1882-1931" and "Minerals yearbook 1932-1970" and are available via `hathitrust.org`. However, the mapping from county-level producer to delivered coal price is not straightforward as a mine in one county can supply multiple counties.

location and depth of coal reserves together with data on the location and opening dates of coal mines. In what follows, I first sketch out a general model for any number of arbitrarily located counties and coal resources on a 2-dimensional plane following Gaudet, Moreaux and Salant (2001).[8] I then demonstrate the basic intuition of the model in a simple 1-dimensional case with 2 counties, 2 coal resources, and 2 periods.

Extending the two-period setting in Section 3.1, county i is now situated on a 2-dimensional plane $\mathbb{D} \in \mathbb{R}^2$. On the same plane are time-invariant coal resources that can, given available mining technology, be converted into a mine. Each resource is characterized by index $k = 1, 2, ..., K$ and resource depth $z_k \in \{z^S, z^D\}$. z^S is the depth of shallow coal resources that can be mined by both manual and mechanized techniques and z^D is the depth of deeper coal resources that can only be mined mechanically, $z^S < z^D$.

I make three simplifying assumptions. First, the size of each coal resource is sufficient large that county coal prices are dominated by transport costs and not resource rents via Hotelling's rule (Hotelling, 1931). Second, transport costs are captured by straight line distance between resource and county. Third, the fixed cost of opening a coal mine, F_{kt}, is determined entirely by the depth of the coal resource and available mining technology. In period $t = 1$ with only manual mining $F_{k1} = 0$ if $z_k = z^S$ and $F_{k1} = \infty$ if $z_k = z^D$ so that the set of operating coal mines in period 1 is $\mathbb{S}_1 = \{\mathbb{D}|z_k = z^S\}$. At period $t = 2$ when mechanized mining is available, $F_{k2} = 0$ for all coal resources and the set of operating coal mines is the entire plane, $\mathbb{S}_2 = \mathbb{D}$.[9]

Following Gaudet, Moreaux and Salant (2001), this setup implies that under a competitive equilibrium[10] a Herfindahl Principle (Herfindahl, 1967) emerges whereby each county i in period t buys coal only from the nearest mine, facing the following delivered coal price:

$$w_{ict} = dist^*_{it} = \min_{k \in \mathbb{S}_t} \{||i - k||\} \tag{1}$$

Observe that in this setup, spatial and temporal variation in coal prices is driven entirely by the opening of coal mines. Coal prices change for a given county if a new mine opens that is closer than the previous closest mine. Mine openings in turn are determined entirely by an interaction between the time-invariant location and depth of coal resources and time varying mining technology. Thus, one can generate a measure of coal prices that is plausibly orthogonal to local determinants of coal demand by interacting an aggregate mining technology shock with the location and depth of coal resources.

Obsolescence is defined within the same context. For a given county i, the closest operating mine over a shallow, manually extracted, resource becomes obsolete when a deeper, mechanically extracted, mine opens in the second period.[11] I call this the obsolescence event. My treatment of interest, the transitory delivered coal price shock, is therefore defined as distance to the obsolete mine:

$$dist^O_{i2} = dist^*_{i1} \quad \text{if} \quad dist^*_{i2} \neq dist^*_{i1} \tag{2}$$

To provide some visual intuition for how this setting defines county coal prices, obsolescence, and path dependence, Figure A.4 considers a simple case for 2 counties i and i', and 2 coal resources, shallow resource k and deep resource k', all located on a line. In the first period as shown in the left panel, both counties buy coal from resource k. Because county i is closer to resource k than county i', coal prices are relatively lower

[8] Kolstad (1994) pioneered in introducing spatial considerations into exhaustible resource theory but required that two extractors be located at the ends of a (Hotelling, 1929) line segment with a continuum of consumers in between.

[9] A fourth assumption is actually required. To avoid indeterminacies, one must assume that the location of counties and coal resources are such that for any county the distance to each coal resource, the norm $dist_{ik} = ||i - k||$, is unique for that county.

[10] Gaudet, Moreaux and Salant (2001) formally solves the social planner problem but because this setting has no externalities, market power, information asymmetries, or strictly positive finite fixed costs, the solution for the planner's problem coincides with the competitive equilibrium. I thank Steve Salant for a clarifying discussion.

[11] Obsolescence is defined for a county-mine pair. A mine may become obsolete for one county but may still be the supplier for another county.

in county i implying also relatively more coal capacity. In period 2, mechanical mining allows both k and k' to be mined. If both counties i and i' are closer to resource k' than resource k, as shown in the right panel, then both counties now buy entirely from resource k' making resource k obsolete for both counties. Because now county i' is relatively closer to resource k' than county i, one would expect that county i' has relatively more coal capacity during period 2 in the absence of path dependence. However, if path dependence were present, the difference in coal capacity between the two counties would be smaller and may even be negative such that county i has relatively more coal capacity in period 2 even though its nearest coal mine is further away than that of county i'.

4 Data construction

Estimating path dependence in electricity infrastructure requires historical records over the 20th century of both county-level delivered coal prices and fuel-specific electricity capacity. Unfortunately, neither is directly available. This section details how spatial data on coal resources and mines are used to constructed a county-by-decade panel of delivered coal prices over the 20th century. I then discuss how a county-by-decade panel of fuel specific electricity capacity over the same period is reconstructed using modern power plant records. I validate both datasets by comparing my constructed variables against historical data that is available. See Appendix A for a detailed summary of all data used.

4.1 Constructing county delivered coal prices

U.S.G.S. assessment of coal resources and mines over the Illinois Basin The construction method for exogenous county delivered coal prices detailed in Section 3.2 interacts the location and depth of time-invariant coal resources with time-varying aggregate mining technology. To implement this approach, I use spatial data from the National Coal Resource Assessment (N.C.R.A.) completed recently by the United States Geological Survey (U.S.G.S.) (East, 2012). The N.C.R.A. includes GIS shape files for coal resources at discrete depths from the surface for the Appalachian, Illinois, Gulf Coast, Northern Rocky Mountains and Great Plains, and Colorado Plateau coal basins (see Figure A.5).

Of these major U.S. coal basins, I consider only coal from the Illinois basin for six reasons. First, compared to other major basins, the Illinois basin has a relatively simple dish-like shape with shallow coal in the outer regions that deepen towards the center of the basin. More geologically complicated basins such as the Appalachian basin that have been subject to processes of folding, tilting, and faulting of subsurface stratigraphic layers over geologic time, often have multiple overlapping coal resources of different depths under the same location. When coal resources are overlapping, one may be concerned that mining of a shallow coal mine could directly lead to mining of deeper coal mines. It is also much harder to define whether mining is occurring over a deep resource that is accessible only via mechanized extraction.

Second, of the eastern U.S. coal basins, coal quality is relatively homogeneous in the Illinois basin, minimizing concerns that path dependence could occur if generating capacity is tied to a particular type of coal (Joskow, 1985, 1987). The three maps shown in Figure A.6 show the heat, sulfur, and ash content of coal produced across the U.S. during 1990-1999 from EIA coal procurement data. Compared with coal found in the Appalachian basin, Illinois basin coal exhibit less variance in heat, sulfur and ash content.

Third, as Figure A.6 shows, the size of counties around the Illinois Basin are relatively homogeneous compared to other parts of the country. This means that measures of straight-line distance between mine location and the area-weighted centroid of a county would be less influenced by variation in county size.

Fourth, the subsurface structure of the Illinois basin has been characterized since the start of the 20th century such that one can credibly interpret the coal resources shown in the N.C.R.A. database as representing resources known throughout the 20th century.[12]. Fifth, of the coal basins surveyed for the N.C.R.A, only the Illinois Basin assessment includes geo-located data on coal mines with opening dates since 1880, which is essential for my construction of county delivered coal prices.[13] Finally, the N.C.R.A. GIS shape files for the Illinois basin separate resources less than and deeper than 200 feet from the surface, the traditional depth cutoff for manual mining (Speight, 1994; Fisher, 1910).

Figure 3 maps the location of coal resources by depth using the N.C.R.A. data. The lighter shaded areas show coal resources that are less than 200 feet below the surface and approximate reserves that were recoverable when only manual mining was available. The darker shaded areas in Figure 3 indicate the location of coal resources situated at depths greater than 200 feet deep. Notice the dish-like shape of the Illinois basin with most of the shallow coal located in the outer ring of the basin.

Figure 3 also locates all large coal mines that operated after 1880, defined as mines that are above the 95% percentile in spatial area. I consider only "large" mines because they are less likely to be depleted due to demand from any individual county, such that Hotelling rents, as discussed in Section 3.2, are less likely to affect the delivered price of coal.[14] In Section 5.4, I demonstrate that this definition is not essential to my findings. To show when mines over shallow and deep resources open, Figure 4 plots the opening of large mines by depth of underlying resource for each decade from 1890 to 1990. While mines over shallow resources continued to be opened throughout the 20th century, coal mines over deeper resources was an increasing share of new coal mines from the 1900s onwards, consistent with the increased national production of mechanically extracted coal shown in Figure A.3.

Construction procedure Construction of county delivered coal prices from 1890-1990 occurs over several steps. Using N.C.R.A. data, I first identify shallow and deep coal mines by spatially overlaying the coordinates of all mines that ever existed since 1890 in the Illinois basin onto shape files of the basin's shallow and deep coal resources. Using the opening and closing years of each mine, I then construct a mine-by-decade panel indicating when each mine was in operation.[15] Next, I restrict attention to all counties for which straight-line distance from the area-weighted county centroid to the nearest Illinois coal resource is i) less than that to the nearest Appalachian coal resource and ii) less than 250 km. Both of these sample restrictions follows from the reasoning in Hotelling's model where local coal prices are set by distance to nearest mine. The first restriction ensures that I only include counties in which delivered prices are unlikely to be determined by the presence of Appalachian coal. The second restriction ensures that I exclude counties that may be affected by other coal basins. In Section 5.4, I show that my results are robust to both restrictions. Figure A.7 highlights the resulting sample of midwestern counties along with the location of Illinois and Appalachian coal basins.

Following Eq. 1, for each county i in decade t, I search through all operating mines to identify the nearest mine according to the straight-line distance between the county's area-weighted centroid and the mine-mouth. Following the Hotelling model reasoning in Section 3.2, that distance is the delivered coal price

[12]See, for example, the Illinois coal basin shown in Campbell (1908)'s map of U.S. coal fields as compared with that shown in Figure A.5.

[13]Historical coal mine data for other coal basins are typically held by state-level agencies. However, they are based on the "final map" for a coal mine which only notes the closing year. Under the Hotelling model, delivered coal prices depends on knowing the year in which a mine opens.

[14]This corresponds to a spatial area more than 0.01 square mile.

[15]Specifically, if the mine was in operation for any year in a decade, I note that it was in operation during that decade.

for county i in decade t. Figure 5 maps county distance to nearest mine from 1890 to 1960. It is evident that the timing and location of mine openings over the 20th century has altered the spatial distribution of delivered coal prices over sample counties with counties becoming both closer and further away from operating mines over the course of the 20th century.

The constructed delivered coal prices shown in Figure 5 are driven by the timing and location of all mine openings. To obtain a more plausibly exogenous measure, I only consider coal price variation occurring when a county first switches from buying coal from a shallow to a deep mine following the introduction of mechanized mining. Because coal depths vary spatially, this obsolescence event occurs at different moments for each county. Figure 4 plots when this switch occurs for the counties in the sample, stacked arbitrarily on the y-axis. The light shaded areas show the period before the obsolescence event while the darker shaded areas show the period during and after the obsolescence event. Because many deep mines opened in the 1900's (see Figure 4), most counties experience the obsolescence event during the 1900's. However, the treatment decade is not uniform for all counties with some facing the obsolescence event before and after the 1900's.[16]

Verifying constructed data against historical data Data limitations discussed in Section 3.1 prohibit one from directly comparing my constructed county delivered coal prices against historical data throughout the 20th century. However, various available historical and modern data can be used to verify both underlying assumptions in the Hotelling model of delivered coal price determination and my constructed coal prices for recent decades. I provide three checks.

First, recall in Section 3.2 that I assumed that delivered coal prices are dominated by transport costs. Previous scholars have noted the high costs of transporting coal, the heaviest per heat content of all energy fuels (Joskow, 1987; McNerney, Farmer and Trancik, 2011). Figure A.8 plots the share of national delivered coal prices due to transport costs over the 20th century. Transport costs were between 40-60% of delivered coal prices for much of the first half of the 20th century, declining thereafter. Thus, assuming that transport costs, as captured by distance traveled between mine and county, serve as a proxy for delivered coal prices is reasonable, particularly during the early decades of the 20th century when most of my sample counties are experiencing the obsolescence event.

Second, the solution to the Hotelling location model states that counties only buy coal from the nearest mine, a version of the Herfindahl Principle. To verify this assumption, I turn to U.S. Energy Information Agency (EIA) data on coal procurement from 1990-1999. The EIA 423 forms document the delivered price and quantity of coal delivered to all power plants, but unfortunately only notes the county of origin of the delivered coal and not the specific mine. Thus, the EIA 423 data allows one to conduct a modified test of the Herfindahl Principle using aggregated data by asking whether coal delivered to a given county varies according to distance between the area-weighted centroids of the destination and origin counties. Because multiple mines exist in a given county, it is unlikely that any county buys coal exclusively from a single origin county, but one would expect more coal to be sourced from nearby counties. For every coal-buying county in the U.S., the left panel of Figure A.9 plots the 1990-1999 average share of total coal purchases and log delivered coal price against the bilateral straight-line distance between area-weighted centroids of the coal destination and origin counties. Notice the steep initial drop in coal shares as distance increases between trading partners which flattens beyond 50 miles. Delivered prices provide a mirrored relationship. The right panel of Figure A.9 shows the same information but according to the rank in bilateral distance

[16] Of the counties in the main sample, 40% faced the obsolescence event during the 1900's while another 37% experience it during the 1960's.

between the trading partners. For a destination country, a rank of 1 means that the coal is sourced within the same county, while a rank of 2 means the coal is sourced by the nearest neighboring county. Consistent with the Herfindahl Principle, there is a steep drop in coal shares during initial increases in distance rank which flattens thereafter. Figure A.10 provide the same information as Figure A.9 but for the subset of sample counties near the Illinois coal basin.

Thus far, I have shown evidence in support of the Hotelling model of coal price determination. As a final check, I compare my constructed county delivered coal prices against available observed data which is available at the county-level using EIA 423 forms over 1972-1999. Table A.1 shows the correlation between observed county log delivered coal prices in the 1970s, 1980s, and 1990s against my proxy measure using log distance to nearest mine. Over all three recent decades, I detect a statistically significant correlation. This relationship holds whether I look at all U.S. counties with coal-fired power plants or just the subset of sample counties near the Illinois coal basin.

4.2 Constructing county-by-fuel electricity capacity

Modern EIA 860 forms on active and retired generating units My main two outcomes of interest are county-level cumulative and new county-level relative coal capacity. Both are defined as the ratio of capacity using coal over capacity using all other fuels. Relative capacity is of interest because it implicitly accounts for the overall scale of the electricity sector in a given county to address concerns about agglomeration forces in electricity demand and because it maps onto theoretical predictions discussed in Section 6.

Ideally, one would use historical data on county-level fuel specific capacity starting around 1890 during the initial introduction of electricity production. However, fuel-specific electricity capacity data at the county-level were either never collected or, if collected, no longer exist today (see Appendix B for a summary of historical electricity data collection and what remains in existence today). Given this data limitation, this paper uses particular features of modern generating unit level data collected by the EIA over 1990-2012 to reconstruct a historical county-by-fuel-by-decade panel of electricity capacity from 1890-1990.

The key feature that makes this reconstruction possible is EIA's collection of both currently operating and retired generating units through the EIA-860 form. EIA-860 includes a wealth of generating-unit characteristics among which, most relevant for this paper, include capacity, opening year, type of primary input fuel, thermal efficiency,[17], prime-mover,[18], and county of location.

Critically, EIA-860 reporting is required not just for currently operating generating units but also for all retired generators located on currently operating power plants, and includes the closing year of the retired generator. This feature of the EIA-860 forms allows one to use cross-sectional EIA-860 data to reconstruct a historical panel.[19] This procedure involves two steps. First, to ensure that I observe as many generating units possible, I create a cross-sectional generating unit database taking the most recently reported data across the 1990-2012 EIA-860 forms. As I show in Section 5.4, my main results are not sensitive to the particular EIA-860 reporting year used. Next, using the opening and closing years of each generating unit, I construct a balanced county-by-fuel-by-decade panel of electricity capacity from 1890-1990.

[17]Thermal efficiency measures how efficiently input energy is converted into electricity. For a given generating unit, the thermal efficiency is the ratio of the heat content from electricity produced (in BTUs) over the heat content of fuel consumed (in BTUs). This was directly collected from 1990-1995.

[18]The prime-mover is the machine which converts fuel into electricity. The most common prime-movers in electricity generation are steam engines, hydropower, and the internal combustion machine.

[19]I only include generating units owned by utility companies because units owned by non-utilities are inconsistently reported over the history of the EIA-860 such that data construction assumptions cannot be verified.

Three data construction assumptions are required in order for my modern EIA-860 reconstructed panel to match actual historical data. First, all power plants since 1890 must still have operating generating units today. That is, power plant sites never retire. Second, a generating unit does not change capacity during its lifetime. Third, a generating unit does not change its primary input fuel. Section 2.2 already provided physical reasons for why the last two assumptions may be reasonable. Next, I turn to the historical data that is available to conduct verification checks of these assumptions.

Verifying constructed data against historical data While actual historical fuel-specific capacity at or below the county level is not available, historical data at more aggregate levels do still exist and can be used to test my first assumption. In particular, capacity at the prime-mover level is available since 1920 at the national level (U.S. Census Bureau, 1975) and power plant level (see (Lewis, 2014) and Appendix B). Panel (A) of Figure A.11 compares the national steam-powered capacity reported by the U.S. historical census from 1920-1970 against national steam-powered capacity aggregated from my panel dataset constructed out of modern EIA-860 forms. There appears to be some underreporting in my reconstructed data from 1920-1955 with convergence between the two datasets afterwards. Because generating units operate over multiple decades, missing capacity may be due to the omission of currently retired power plants which opened either prior to or after 1920. The former reason would imply data inaccuracies in my measure of cumulative electricity capacity while the latter reason would additional imply inaccuracies in my measure of new electricity capacity during each period. Panel (B) of Figure A.11 provides a similar national-level comparison but for new steam-power capacity. New capacity from the modern EIA-860 reconstructed data closely track that from the national census suggesting that I am likely observing power plants that have opened since 1920. Panels (C) and (D) of Figure A.11 provides similar comparisons but for national-level hydropower capacity. Again, Panel (C) modern EIA-860 data appear to be missing some historical hydro power plants between 1920-1955 though the magnitude of the censuring appear less than with steam-power capacity. Panel (D) shows that new hydro capacity across the two datasets are highly similar.

To provide a finer comparison, Figure A.12 replicates Figure A.11 but only for the eleven states covering the estimating sample as shown in Figure A.7 by aggregating power plant-level historical data available from Lewis (2014).[20] Consistently with the national-level data verification, Panel (A) shows that steam-capacity is underreported over 1920-1955 in the EIA-860 reconstructed dataset with comparisons of new capacity shown in Panel (B) indicating that most the the missing power plants are likely those built prior to 1920. A similar pattern of lesser magnitude is shown for hydropower capacity across the eleven sample states in Panels (C) and Panel (D). Again, this suggests that my EIA-860 reconstructed panel may underreport cumulative relative coal capacity but is likely to accurately capture new relative coal capacity since 1920.

Turning to my other two assumptions, I examine whether generating unit capacity and fuel input has changed over time by exploring the consistency in reported values both across 1990-2012 EIA-860 records and between the 2012 EIA-860 dataset and the earliest available generating unit-level dataset. Table A.2 examines the consistency of relevant generator characteristics across the 1990-2012 EIA-860 records. For each variable across the columns of Table A.2, the rows show the percentage of EIA-860 reports from 1990-2011 that differed from the value recorded in 2012. For example, 94% of generating units reported using the same primary fuel throughout 1990-2011 as was reported in 2012. Likewise, for reported generating unit capacity,

[20] While the historical power plant-level dataset constructed by Lewis (2014) includes power plant names, only fuzzing string matching with modern EIA-860 data can be conducted. It is impossible to verify whether any string mismatches is due to power plant name changes, data entry errors, or actual missing data in the EIA-860 forms. As such, comparing capacity at the state level is a more straightforward test of data censuring.

opening year, and retirement year, 75%, 97% and 80% of generators respectively reported the same value in 1990-2011 as was reported in 2012. To examine data consistency going further back, I digitized the E.I.A.'s "Inventory of Power Plants in the United States" in 1980, the earliest available comprehensive generating unit-level dataset collected in the late 1970s (see Section B for more details). Figure A.13 plots the generating unit-level capacity reported in 2012 against the capacity reported in the late 1970s. The relationship between the two reported values is nearly one-to-one. Table A.3 shows the distribution of reported primary fuel in 2012 conditional on primary fuel reported in the late 1970s. 92% of generating units that reported using coal in the late 1970s reported coal use in 2012. Fuel switching appears to occur more frequently between natural gas and oil-fired generators with 77% and 75% of generators which reported use of natural gas and oil respectively in the late 1970s noting the same fuel use in 2012. There is no evidence of any fuel switching since the late 1970s for generators that use nuclear and hydro power.

In summary, data verification checks using available historical data at various levels suggest that my county-by-fuel-by-decade panel of electricity capacity reconstructed using modern EIA-860 data reasonably approximates the true historical data.

5 Results

This section presents reduced-form evidence of path dependence in coal-fired electricity capacity. First, to guide my model selection, I test whether various observed county characteristics are correlated with distance to nearest mine prior to the obsolescence event. Having selected my preferred model, I then test for the presence and strength of path dependence by estimating an event study that examines the effects of distance to the obsolete mine decades obsolescence. Robustness of my estimates to model specification, sample selection, and various data construction choices are then examined.

5.1 Examining pre-obsolescence trends in county characteristics

A key identifying assumption from Section 3.1 requires that conditional on controls, contemporaneous coal prices should be uncorrelated with other contemporaneous determinants of coal capacity. To examine this assumption, in Table 1 I estimate whether distance to the obsolete mine in the period prior to obsolescence is correlated with several county characteristics which may affect coal capacity. These includes total population and urban population as well as the number of establishments, average number of wage earners per establishment, capital value, and total output in the manufacturing sector. These characteristics are indicated in each row of Table 1 along with the outcome data coverage period. For some counties, characteristics were not available prior to the obsolescence event because the ICPSR historical census database does not provide a balanced panel over the 20th century.

I consider models that employ four different identifying assumptions. In each, standard errors are clustered at the county level, allowing for heteroscedasticity and serial correlation of arbitrary form within a county. In Column (1), I estimate the effect of log distance to obsolete mine on each log outcome in levels. Of the six outcomes considered, only the number of manufacturing establishments show a statistically significant correlation. For Column (2), I consider a specification that is similar to a panel estimator with county fixed effects by estimating a model with state fixed effects on first-differenced log outcome variables. None of the outcomes examined are correlated with log distance to obsolete mine in this pre-trend test. In Column (3), I add year fixed effects to the model in Column (2) showing only a statistically significant effect on total population. Finally, Column (4) additionally employs state trends such that effects are identified

relative to county averages, year averages, and state trends. None of the county characteristics are correlated with mine distance before obsolescence. This suggests a benchmark specification that includes county fixed effects, year fixed effects, and state trends.

5.2 Main specification

My two primary outcomes of interest are cumulative and new relative coal-fired electricity capacity. There are two properties of electricity infrastructure that inform key modeling choices. First, the EIA 860 forms omit generating units with capacity less than 1 MW. As a consequence, this data censuring leads to a number of zero or indeterminate values in relative coal capacity when in fact there may be coal and non-coal capacity in that county-decade observation. Second, the distribution of electricity capacity is highly left-skewed. To address the first concern, I add a value of 1 to both coal and non-coal capacity in all observations. This is reasonable since it is possible that all sample counties have some electricity production even if it is not reported in the EIA 860 forms. One could address the second concern by applying a log transformation to the outcome variable. However, a log transformation is sensitive to low values which is now exacerbated by my solution to the first concern. Instead, I elect to estimate a Poisson fixed effects model as my benchmark, showing later in Section 5.4 that results are relatively insensitive to this modeling choice. Specifically, my benchmark specification for relative coal capacity in county i, state s, decade t is:

$$\widetilde{X}_{ist} = exp(\sum_{\tau=1}^{10} \beta_{\tau}^{O} * log(dist_{i}^{O}) * \mathbf{1}[t=\tau] + \beta^{C} log(dist_{it}^{*}) + \sum_{\tau=1}^{10} \gamma_{\tau} * \mathbf{1}[t=\tau] + \nu_t + \phi_s t + \mu_i) + \epsilon_{ist} \quad (3)$$

where as noted earlier $dist_i^O$ is distance to the obsolete mine and τ indexes obsolescence event time in decades. β_τ^O is the parameter of interest and captures the event time-varying effect of distance to the obsolete mine in the decades after obsolescence relative to the effect prior to obsolescence, the omitted variable. Following the evidence from the pre-trend tests in Section 5.1, Equation 3 includes a full set of county fixed effects, μ_i, year fixed effects, ν_t, and state trends, ϕ_s. Additionally, I include set of event time dummies and contemporaneous distance to nearest mine, $dist_{it}^*$. The latter controls for the effect of contemporaneous delivered coal prices. Altogether, the inclusion of controls for contemporaneous determinants of relative coal capacity as well as semi-parametric controls for time invariant and time varying determinants provides reassurance that estimates of β_τ^O can be as capturing path dependence.

In practice, Equation 3 is estimated using conditional maximum likelihood where firm fixed effects are eliminated by conditioning on a sufficient statistic for μ_i, in this case $\sum_t \widetilde{X}_{ist}$. The Poisson distribution also benefits from being a member of the linear exponential family such that even if the density is misspecified, Equation 3 produces consistent point estimates through quasi-MLE provided that the conditional mean function is correctly specified. Poisson models have a general limitation by requiring that the first and second moments be equal which is overly restrictive in most settings. To address this issue, I estimate Eq. 3 with county-level clustered standard errors which relaxes this restriction by allowing for arbitrary forms of heteroskedasticity and serial correlation within a county.

Finally, for reasons due to the location of other U.S. coal basins discussed in Section 4.1, my benchmark sample of counties is restricted to those whose area-weighted centroid is within 250 miles of any resource in the Illinois coal basin and is situated closer to the Illinois basin than to the Appalachian basin. I examine whether estimates are sensitive to this sample restriction in Section 5.4.

5.3 Main result: effects of path dependence

Table 2 shows estimates of Eq. 3. In Column (1), I first examine the event time-varying effects of distance to the obsolete mine on cumulative coal capacity. I detect lagged effects, relative to the contemporaneous effect, that are statistically significant and strengthen over the ten decades following obsolescence. Column (2) shows that path dependence is not simply driven by non-depreciated capital built prior to obsolescence as increasing effects over time can also be detected on new coal capacity built each decade. Columns (3) and (4) estimate Eq. 3 for cumulative and new electricity capacity using all other fuels (i.e. oil, natural gas, nuclear, hydro, etc) respectively. I do not detected a statistically significant lagged effect of distance to obsolete mine on non-coal capacity. For new non-coal capacity, a statistically significant effect is detected 1 and 3 decades later but only at the 10% level, nor do the point estimates strengthen over time like they do for coal capacity. Strangely, distance to current nearest mine appears to have an effect on new non-coal capacity.

Columns (5) and (6) display my main result: path dependence in relative coal capacity. Consistent with the results in Columns (1) and (3), Column (5) shows that distance to the obsolete mine has a strengthening effect on cumulative relative coal capacity that can be detected statistically up to ten decades after obsolescence. Crucially, I detect distinct jumps two and seven decades after obsolescence with a 1% increase in distance to the obsolete mine lowering relative coal capacity by 2.3% and 3.7% two and seven decades later. The timing of these effects coincide approximately with the expected moment of capital replacement for electricity generators.[21] This implies that historic transitory coal price shocks have a detectable effect on two subsequent generations of new electricity capital. Indeed, this is confirmed by the coefficients shown in Column (6) which models effects on new relative coal capacity. To get a sense of magnitude for the path dependence effect, one can compare the lagged effects of distance to the obsolete mine to the contemporaneous effect of distance to the current nearest mine. Doing so shows that the lagged effects two and seven decades later are 3 and 4 times stronger respectively than the contemporaneous effect.[22]

5.4 Robustness checks

This section conducts a series of robustness tests examining the stability of my benchmark path dependence result to different model specifications, estimating sample, data construction choices, and other modeling decisions.

Nonlinearity Eq. 3 implicitly assumes that the relationship between relative coal capacity and distance to the obsolete mine is an isoelastic function and thus has a linear coefficient. While this is suitable for the theory considered in Section 6, it need not be empirically true. Figure 7 examines whether the data supports this functional form assumption by estimating a variant of Eq. 3 that breaks distance to obsolete mine into 50 mile-wide dummy bins. This semi-parametric functional form allows for a flexible relationship between relative coal capacity and distance to obsolete coal mine for each period after obsolescence. Figure 7 shows the dummy coefficients one, three, five, seven, and nine decades after obsolescence. Linearity appears to be a reasonable assumption for each lagged effect.

[21] According to the EIA-860 forms, the average retired generating unit in 2012 had a lifespan of 54 years. The average duration in which a shallow coal served as the price-setter mine before being replaced by a deep mine was 30 years. Thus, the effects observed two and seven decades after obsolescence coincides with the expected moment of electricity capital replacement.

[22] Two decades later: $\frac{2.31+1.15}{1.15} = 3$. Seven decades later: $\frac{3.7+1.15}{1.15} = 4.2$

Model specification Table 3 examines alternative model specifications. In Column (1), I estimate a stripped-down version of Eq. 3 including only county fixed effects and event time dummies. I detect statistically significant lagged effects two to five decades after obsolescence but they do not get stronger over time. Coefficients in Column (1), however, does not control for contemporaneous determinants of relative coal capacity which, if correlated with distance to the obsolete mine, would not empirically separate effects due to persistence versus path dependence. As such, additional controls are added cumulatively across Columns (2) to (5) of Table 3. In Column (2), I include contemporaneous coal prices, as approximated by distance to current nearest mine, which leads to precisely estimated and growing lagged effects up to 10 decades after obsolescence. The magnitude of these effects, however, are lower than my benchmark estimates in Column (5) of Table 2. Columns (3) further includes year fixed effects while Column (4), which replicates my benchmark result, additionally includes state trends. Finally, in Column (6) I additionally control for total population, number of manufacturing establishments, and manufacturing employment, the three county covariates for which county-level data is available over the entire the 20th century. The magnitude of my path dependence effects increase relative to my benchmark model though not enough to be statistically different.

Estimating sample Table 4 examines different county sample definitions. For my benchmark model, reproduced as Column (1), I restrict my sample to counties within 250 miles of the nearest Illinois basin coal resource and situated closer to Illinois basin coal than to Appalachian basin coal. Recall that the rationale for both sample restrictions is to lessen the competing effects of coal resources in neighboring basins. In Column (2), I further weaken the effects of other coal basins by restricting my sample to counties within 200 miles of the nearest Illinois basin coal resource. I find effects that are very similar to my benchmark. In Column (3), I allow more counties into my sample by increasing the distance threshold to within 300 miles of the nearest Illinois basin coal resource. Effects are now slightly smaller though they are within the confidence interval of my benchmark estimates. In Column (4), I allow counties that are situated closer to the Appalachian basin into the sample. As with Column (3), I still detect statistically significant increasing lagged effects but of lower magnitude possibly due to influence of Appalachian basin availability. In Column (5), I consider only counties that faced an obsolescence event prior to the 1960's to test whether there are differences based on the when counties experience the obsolescence event. I obtain estimates that are larger but not statistically very similar to that of my benchmark model.

Data construction choices Table 5 conducts further robustness checks by examining whether results are sensitive to how delivered coal prices and relative coal capacity are constructed. Column (1) replicates the benchmark result. Recall that my definition of a "large" coal mine in the Illinois basin uses a cutoff area of 0.01 square miles, which is the 95% of Illinois Basin mines. In Columns (2) and (3), I demonstrate that using a smaller threshold of 0.005 and a larger threshold of 0.015 square miles still produces lagged effects that strengthen over time, though they are less precisely estimated. Hotelling's location model argues that coal prices are dictated by distance to the nearest coal mine, which finds empirical support in Section 4.1. Nonetheless, assuming that counties buy from a single mine may be an overly strong assumption. In Columns (4) and (5), I construct my distance measure based on the average distance to the nearest two and three mines, respectively. The resulting point estimates are larger though not statistically different from my benchmark.

In Section 4.2, I discussed that reported generator characteristics are remarkably consistent across 1990-

2012 EIA 860 forms. As a consequence, my constructed capacity dataset uses values from the latest year in which a generator is reported across the 1990-2012 EIA-860 forms to ensure that I observe all retired generators that were reported since 1990. To demonstrate that my results are not driven by this particular procedure, in Column (5) I construct my capacity dataset using only the 2000 EIA-860 forms while in Column (6) I use only the 2005 EIA-860 form. Results are very similar to my benchmark estimate.

Other modeling choices Table 6 examines robustness to other modeling choices. Again, for comparison, the benchmark result is reproduce in Column (1). Previously, I coded all years prior to obsolescence in which the obsolete coal mine was the nearest mine as $\tau = 0$. In Column (2), I recode the treatment such that only the single year before obsolescence is assigned $\tau = 0$. As a result, a few observations are dropped but estimates of lagged effects are largely unaffected. In Column (3), I examine what happens when I do not adjust coal and non-coal capacity by add a unit value. I still detect statistically significant lagged effects three to five decades after obsolescence. However, because sample size is now less than one-fifth smaller than my benchmark estimate, I do not detect precise estimates over longer lag periods. In Column (4), I consider a log-linear model of adjusted relative coal capacity, showing effects that are similar to that using a poisson model. Finally, in Column (5) I estimate a log-linear model using the raw unadjusted capacity values. Despite the much smaller sample size, I still detect statistically significant lagged effects two, three, and five decades after obsolescence though they are much smaller in magnitude.

To provide a visual summary of how these robustness checks compare with my benchmark estimate, Figure 8 plots my benchmark point estimates and confidence interval along with point estimates from the 14 models estimates across Tables 4, 5, and 6. Observe that almost all coefficients from these robustness tests fall within the uncertainty of my benchmark estimate.

6 Mechanisms

Section 5 finds robust evidence that past transitory delivered coal price shocks have an increasing effect on subsequent relative coal capacity over 10 decades later. However, different mechanisms can generate such dynamics. Without knowing more about the particular mechanism driving my reduced-form results thus far, it is difficult to know how such historical evidence can inform potential future energy transitions in the U.S. and elsewhere. This section aims to empirically isolate such a mechanism so as to offer a structural interpretation of my reduced-form results.

In the broadest sense, path dependence occurs when past relative fuel prices shift the relative marginal product of capital in subsequent periods. Earlier literature highlights two mechanisms that can generate this effect in the electricity sector: increasing returns to scale and the accumulation of a stock factor such as coal-specific technology or human capital. To examine these competing channels, I first develop a production function for an electric power plant that incorporates both scale and productivity effects. Informed by this theoretical framework, I then conduct a series of nested empirical tests designed to isolate the relevant mechanism. I also test other potential mechanisms that fall outside my theoretical framework. This combination of theory and empirics provides the set up for recovering my ultimate structural parameter of interest, the long-run elasticity between coal and non-coal electricity.

6.1 Theory: scale and productivity effects in electricity production

To guide the ensuing empirical analysis, I first introduce a plausible production function for a power plant. This production function incorporates modeling elements that correspond to the three features of power plants introduced in Section 2.2, each having separately been highlighted in the literature: imperfect substitutability between electricity from coal and other fuels (Acemoglu et al., 2012), increasing returns to scale for fuel-specific capacity (Nerlove, 1963; Christensen and Greene, 1976), and vintaged fuel-specific electricity capacity with fixed factor proportions (Komiya, 1962; Solow et al., 1966; Atkeson and Kehoe, 1999).

These elements are represented by separate tiers of a power plant production function. At the top tier, the power plant produces the final good, electricity, Y_t, in period t using two intermediate goods, Y_{ct} and Y_{nt}, representing electricity from coal and all other fuels respectively.[23] Specifically, it takes the following Constant Elasticity of Substitution form:

$$Y_t = \left(Y_{ct}^{(\epsilon-1)/\epsilon} + Y_{nt}^{(\epsilon-1)/\epsilon}\right)^{\epsilon/(\epsilon-1)} \tag{4}$$

where ϵ is the elasticity of substitution between electricity produced by the two intermediate goods. Following the discussion in Section 2.2, I assume that electricity generated by different energy inputs are gross but imperfect substitutes so that $\epsilon > 1$, arising from differences in reliability and where fuel-specific electricity falls on the dispatch curve. The price of the final good is normalized to 1.[24]

Fuel-specific electricity production occurs over two lower tiers that separately capture scale and technology effects. Fuel-specific electricity is generated by combining output from generating units across two vintages, X_{jt} and X_{jt-1}, in a Cobb-Douglas function with scale parameter r.[25] This scale parameter captures the efficiency gains noted in Section 2.2 from having larger boilers that provide steam to generating units across multiple vintages. At the lowest tier, vintage and fuel-specific generating units produce output by combining capital with fuel, E_{jt} and E_{jt-1}, in fixed proportions as modeled by a Leontief function.[26] These proportions are captured by productivity terms A_{Xjt} and A_{Ejt} representing respectively the productivity of capital and fuel for a generating unit of vintage t using fuel j. The intermediate good production function is:

$$Y_{jt} = (\min[A_{Ejt}E_{jt}, \delta A_{Xjt}X_{jt}])^{\alpha}(\min[A_{Xjt-1}E_{jt-1}, \delta A_{Xjt-1}X_{jt-1}])^{\alpha} \quad \text{for} \quad \text{j}\in\{\text{c,n}\} \tag{6}$$

where δ is capital depreciation, $\alpha \in \{0, 1\}$ is the output elasticity of capacity so that $r = 2\alpha$.[27] Efficient input use under fixed factor proportions imply for each period and intermediate good $A_{Ejt}E_{jt} = A_{Xjt}X_{jt}$ and $A_{Ejt-1}E_{jt-1} = A_{Xjt-1}X_{jt-1}$.

I now explore how scale and productivity effects could generate path dependence for otherwise similar intermediate sectors. Suppose capital and fuel productivities are the same across the two intermediate production functions during period $t-1$, $A_{Xct-1} = A_{Xjt-1}$ and $A_{Ect-1} = A_{Ejt-1}$. The first order condition in period t, expressed in relative terms, and rewriting in terms of of relative current vintage capacity, $\widetilde{X}_t = \frac{X_{ct}}{X_{nt}}$, is (see Appendix C.1 for full derivation):

[23] For simplicity, this implies that electricity produced by all other inputs are perfect substitutes.

[24] Specifically,

$$\left[p_{ct}^{1-\epsilon} + p_{nt}^{1-\epsilon}\right]^{\frac{1}{1-\epsilon}} = 1 \tag{5}$$

[25] Assuming a CES function at this level introduces an extra substitution parameter that complicates the empirical implications but would not alter the main conclusion.

[26] This is also known as "clay-clay" production.

[27] This allows for diminishing marginal product under varying returns to scale. Otherwise, the relative input demand curve becomes upward sloping.

$$\widetilde{X}_t = \widetilde{w}_t^{\frac{\epsilon}{\psi-1}} \widetilde{X}_{t-1}^{\frac{\alpha(1-\epsilon)}{\psi-1}} \widetilde{A}_{Et}^{\frac{-\epsilon}{\psi-1}} \widetilde{A}_{Xt} \tag{7}$$

where $\widetilde{w}_t$ is the relative input price index[28] and $\psi = (1-\alpha)(1-\epsilon) < 0$, from earlier assumptions. Capital depreciation plays no role in Eq. 7 as it is the same in both intermediate sectors. Past relative input prices $\widetilde{w}_{t-1}$ can affect current vintage relative coal capacity via two channels. First, applying Eq. 7 recursively shows that past relative coal prices affect current relative coal capacity via past relative coal capacity, $\widetilde{X}_{t-1}$. This is the scale effect. Second, any form of fuel-specific endogenous technical change (i.e. directed technical change or human capital accumulation) allows past relative input prices to affect current relative coal capacity via current relative fuel productivity, $\widetilde{A}_{Et}$. This is the productivity effect.

Eq. 7 suggests that path dependence at the power plant level can occur through either the scale or productivity effects. To empirically isolate each effect, I exploit the tiered structure of my power plant production function through a nested series of empirical tests. First, I turn to power plant level regressions for plants that only burn coal using standard specification designed to recover the scale parameter. Because such estimates may be biased in the presence of productivity effects, I then turn to generating-unit level regressions to test for productivity effects. Finally, I conduct a series of additional tests to examine the possible role of other mechanisms that are not modeled in my production function.

6.2 Testing for scale effects at the power plant level

To recover the scale parameter, I follow the cost function approach developed initially by Nerlove (1963) and further implemented by Christensen and Greene (1976) to estimate scale economies in the electricity sector. To remove the influence of the elasticity of substitution parameter, I consider a power plant m in county i and state s that primarily uses coal as input fuel. Using cost data from the Utility Data Institute, cost minimization of Eq. 6 suggests the following OLS regression of non-fuel cost (see Appendix C.2 for full derivation):

$$log(\overline{\text{non_fuel_cost}}_{mis}) = \frac{1}{r} log(\overline{Y}_{mis}) + \frac{1}{2} log(\overline{w}_{mis}) + \overline{\boldsymbol{Z}}_{mis}\theta + \epsilon_{mis} \tag{8}$$

where the bar indicates time averaged variables because UDI cost data is only available for the 1981-1999 period such that a panel regression over the 20th century is not possible. My parameter of interest is the scale parameter r. $\boldsymbol{Z}_{mis}$ is a vector of cross-sectional controls intended to absorb cross-sectional differences in generating unit productivity and includes state fixed effects and the latitude and longitude of the county centroid. In addition, I control for differences across transmission girds by including NERC region fixed effects. Standard errors are clustered at the county level.

Table 7 estimates Eq. 8 for three samples of power plants. The first sample shown in Column (1) include only power plants in our sample of midwest counties (see Figure A.7) which only burn coal. I estimate a statistically significant scale parameter of 1.8. The second sample contain midwestern plants which indicate coal as the "primary-fuel".[29] While these plants may have some fuel switching capability, I find in Column (2) that they exhibit similar returns to scale as the coal-only power plants, with a statistically significant scale parameter of 2.1.

[28] Specifically, $w_{jt} = z_{jt} + \frac{A_{Ejt}}{A_{Xjt}} r_t$ where z_{jt} is the primary energy input price and r_t is the rental price of capital.

[29] According to the Platts data, federal forms used in compiling this report make no designation of primary or secondary fuels burned at a plant site. For multi-fuel plants, primary fuel is established by calculating the energy input of each fuel used and then assigning the primary fuel to that with the highest energy input.

Potential simultaneity bias in Eq. 8 has been noted as early as Nerlove (1963). In particular, electricity prices for regulated electric utilities are often set to cover a power plant's average costs which implies that electricity output may be correlated with unobserved determinants of non-fuel costs. To address this concern, Columns (3) and (4) of Table 7 uses past constructed delivered coal prices as an instrument for current electricity output. Specifically, my instrument is the interaction between county distance to the obsolete mine and number of decades since the obsolescence event. For my identification strategy to be valid, my measure of past coal prices must affect current non-fuel costs only through current output. Specifically, my first stage regression is:

$$log\overline{Y}_{mis} = \kappa_1 log(dist_i^O) * sinceEvent_i + \kappa_2 log(dist_i^O) + \kappa_3 sinceEvent_i + \frac{1}{2} log\overline{w}_{mis} + \overline{\boldsymbol{Z}}_{mis}\theta + \mu_{mis} \quad (9)$$

Eq. 8 estimates the event-time varying effects of distance to the obsolete coal mine and is the cross-sectional version of my panel estimator in Eq. 3. In particular, κ_1 captures the event-time varying effect of distance to the obsolete mine after obsolescence and is analogous to the slope of the lagged effects shown in Figure 8. Columns (5) and (6) show IV estimates that are statistically significant, differs little depending on power plant sample, and similar in magnitude to my OLS estimates. Furthermore, these IV estimates are robust to the potential presence of a weak instrument. For both IV estimates, the p-value and confidence interval from a conditional likelihood ratio test (Moreira, 2003) rejects the null that the coefficient on electricity output in the structural equation is zero. Henceforth, my preferred estimate of r is 1.66 from Column (5) of Table 7.

Finally, the theory presented in Section 6.1 assumes that coal and non-coal electricity exhibit similar returns to scale. To test this assumption, I estimate Eq. 8 for power plants that primarily burn natural gas and oil in Columns (3) and (4) of 7. Because non-coal power plants are limited in my sample of midwest counties, I turn to the full sample of U.S. power plants in the UDI dataset. I find scale parameters for plants that primarily burn gas and oil that are very similar to that of coal-fired power plants.

6.3 Testing for productivity effects at the generating unit level

If past coal prices directly affect current power plant productivity, the exclusion restriction assumption for the IV estimator in Section 6.2 fails and the scale parameter cannot be identified. To examine the presence of productivity effects, I turn to generating-unit level regressions where productivity effects are most likely to be isolated. A standard engineering measure of generating unit productivity is thermal efficiency, the ratio of heat from electricity produced to heat from fuel consumed,[30] and best corresponds to $A_{Ect} = \frac{Y_{ct}}{E_{ct}}$ in the theory above. Using cross-sectional data of generating-unit variables from EIA 860 forms averaged for the period 1990-1995[31], I estimate the following regression of thermal efficiency for generating unit g, in power plant m, county i, and state s:

$$log(\frac{\overline{Y}}{\overline{E}})_{gmis} = \omega_1 log(dist_i^O) * sinceEvent_i + \omega_2 log(dist_i^O) + \omega_3 sinceEvent_i + \omega_4 log(dist_i^*) + \overline{\boldsymbol{Z}}_{gmis}\theta + \epsilon_{gmis} \quad (10)$$

where again $\boldsymbol{Z}_{mis}$ includes state and NERC region fixed effects and the latitude and longitude of the county centroid. As with Eq. 9, Eq. 10 examines, through ω_1, whether distance to the obsolete mine has an

[30] Thermal efficiency is also the inverse of the heat rate.

[31] Generating unit heat rates, the inverse of thermal efficiency, was recorded by the EIA 860 form only for the 1990-1995 period.

increasing effect on generator efficiency after obsolescence. If so, this would suggest that historic transitory coal price shocks induce local productivity improvements. Results are shown in Table 8. Columns (1) and (2) include all coal-fired generating units while Columns (3) and (4) include generating units on power plants that only burn coal. Columns (2) and (4) augments the model in Eq. 10 by additionally including the age and capacity of the generating unit. I do not find that historic coal price shocks have a statistically significant effect on generator thermal efficiency. The strongest determinant of thermal efficiency is the age of the generating unit, which likely reflects aggregate technological change and not location-specific changes driven by historic coal price shocks.

It is unsurprising that I am unable to detect local productivity effects. While capital-augmenting technical change, either embodied or disembodied, may respond to aggregate price shocks strong enough to redirect research efforts, it seems unlikely that there is directed technical change in response to local price shocks at the county level. Similarly, if productivity effects occur through coal-specific human capital accumulation, such effects are unlikely to be localized at the county level given that electric utilities typically own power plants in several counties.

6.4 Other mechanisms

The theoretical framework presented in Section 6.1 incorporates two potential drivers of path dependence, scale and productivity effects within a power plant. However, there may be other drivers at play beyond a power plant that may generate path dependence. This section provides examines a few additional mechanisms.

Electricity can be an input for other sectors such as manufacturing and residential. If these other sectors exhibit scale effects or some other driver of path dependence, then historic obsolete coal access may lead to increased total demand for electricity over time. This in turn may drive path dependence in electricity production. Indeed, lagged effects of historic access to electricity on local population density and manufacturing sector employment has previously been detected following the introduction of hydropower dams by Kline and Moretti (2014) and Severnini (2014) respectively.

The first reassurance that my findings cannot be explained by electricity demand-side scale effects alone is that I detect path dependence effects over two generations of not only coal-fired capacity, but relative coal-fired capacity which normalizes for the scale of local electricity demand. To provide additional evidence, Table 9 estimates the distance to obsolete mine effects on various county characteristics that approximate for electricity demand in the residential and manufacturing sectors. Specifically, I separately examine effects on total population, urban population, manufacturing establishments, and manufacturing employment over the eight decades after obsolescence, as indicated across Columns (1)-(8). To be consistent with earlier pre-trend tests of county characteristics, I use first-differenced log outcomes and a specification identical to that shown in Column (4) of Table 1 with a further inclusion of contemporaneous distance to nearest mine to control for current coal prices. A few points are worth noting. First there are a few significant coefficients. Of the total 32 separately estimated coefficients, 6, or 19%, are statistically significant at the 10% level. However, none of the county characteristics exhibit time-varying effects that systematically follow the pattern of increasing lagged effects found for relative coal capacity. This suggests that demand-side drivers are unlikely to be driving path dependence in relative coal capacity.

Infrastructure for coal-fired electricity is not exclusive to a power plant. In order to coal to be delivered to a power plant, rail and highway networks are needed as complementary transport capital. Large fixed costs in transport capital can also generate path dependent dynamics through economies of scale. Unfortunately,

unlike the county outcomes examined in Table 9, panel data for county-level rail and roadway density is not available. Instead, I use measures of rail and highway density (miles per square mile) in 2010 in a county-level cross-sectional regression similar to the thermal efficiency specification in Eq. 10. Results are shown in Table 10. I find that neither forms of transport infrastructure respond to past coal price shocks.

Finally, Joskow (1987) finds that power plants procure coal using long-duration contracts. The presence of such long-term contracts may induce power plants to continue purchasing coal even as contemporaneous economic circumstances change. There are two reasons for why such procurement practices would not fully explain the path dependent effects estimated in Section 5. First, Joskow (1987) showed that the average coal contract length in 1979 lasted 12.8 years with a standard deviation of 10.4 years. Using EIA 423 data form 1983-1997, Jha (2015) finds that this duration has decreased since to 4.4 years with a standard deviation of 5.9 years. Coal contracts of such duration may induce persistence in coal capacity within a decade but is unlikely to generate lagged effects over multiple decades. Furthermore, even if coal contracts were of longer duration, it is unclear why such contracts would cause lagged effects to increase over time.

7 Structural interpretation

Evidence from Section 6 support that path dependence in coal-fired electricity is occurring through scale effects. In this section, I extend the theoretical framework in Section 6.1 to formally define path dependent effects that increase or decrease over time as a function of the scale parameter and the elasticity of substitution between coal and non-coal electricity. Using this definition, I then recover the implied long-run elasticity of substitution as implied by my estimates of path dependence in Section 5 and scale parameter in Section 6.2. This elasticity is of particular interest as it is a common and important parameter across a broad class of structural change models featuring different drivers of path dependence. Finally, I use these structural parameters to simulate how current low natural gas prices could induce a structural transition from coal to natural gas-fired electricity resulting in various future aggregate CO_2 emission pathways for the U.S. electricity sector.

7.1 Strength of path dependence: formal definition

In the absence of productivity effects, one can assume relative productivities are the same in each time period, that is $\widetilde{A}_{Et} = 1$ and $\widetilde{A}_{Xt} = 1$. Under these conditions, applying a log transformation to Eq. 7 and rewriting it recursively:

$$\begin{aligned} \ln \widetilde{X}_t &= \frac{\epsilon}{(\psi-1)} \ln \widetilde{w}_t + \frac{\alpha(1-\epsilon)\epsilon}{(\psi-1)^2} \ln \widetilde{w}_{t-1} + \frac{\alpha^2(1-\epsilon)^2\epsilon}{(\psi-1)^3} \ln \widetilde{w}_{t-2} + ... \\ &= \sum_{s=0}^{\infty} \frac{\epsilon}{(\psi-1)} \left(\frac{\alpha(1-\epsilon)}{(\psi-1)} \right)^s \ln \widetilde{w}_{t-s} \end{aligned} \tag{11}$$

where s is the lag time index and $\psi = (1-\alpha)(1-\epsilon) < 0$, from earlier assumptions. Relative current coal capacity declines with higher relative coal prices, $\frac{\partial \ln \widetilde{X}_t}{\partial \ln \widetilde{w}_t} < 0$. Furthermore, consistent with earlier vintage capital literature, past relative energy price shocks generate lagged effects as both $\frac{\partial \ln \widetilde{X}_t}{\partial \ln \widetilde{w}_{t-1}} < 0$ and $\frac{\partial \ln \widetilde{X}_t}{\partial \ln \widetilde{w}_{t-2}} < 0$. But when do these effects increase or decrease over time? An examination of Eq. 11 provides formal definitions of path dependence strength, defined as the difference between lagged and current effects of relative fuel price shocks:

PROPOSITION 1 ***Weak path dependence:*** *The effect of past relative coal prices weakens over time,* $\frac{\partial \ln(\widetilde{X}_t)}{\partial \ln(\widetilde{w}_{t-1})} - \frac{\partial \ln(\widetilde{X}_t)}{\partial \ln(\widetilde{w}_t)} = \frac{\alpha(1-\epsilon)\epsilon}{(\psi-1)^2} - \frac{\epsilon}{(\psi-1)} > 0$*, or when* $r < \frac{-\epsilon}{1-\epsilon}$.

PROPOSITION 2 ***Strong path dependence:*** *The effect of past relative coal prices strengthens over time,* $\frac{\partial \ln(\widetilde{X}_t)}{\partial \ln(\widetilde{w}_{t-1})} - \frac{\partial \ln(\widetilde{X}_t)}{\partial \ln(\widetilde{w}_t)} = \frac{\alpha(1-\epsilon)\epsilon}{(\psi-1)^2} - \frac{\epsilon}{(\psi-1)} < 0$*, when* $r > \frac{-\epsilon}{1-\epsilon}$.

Strong path dependence occurs whenever an increase in relative coal prices triggers a downward shift in the relative marginal product curve of coal and non-coal electricity that continues in each subsequent period. Two countervailing forces are at play. Consider first just the role of the scale parameter r on coal-fired capacity. When $r > 1$, the cross partial derivative of Y_{ct} exceeds the second partial derivative of Y_{ct} with respect to current coal capacity such that an increase in the coal price has a stronger effect on lagged coal capacity than on current coal capacity.[32] However, $r > 1$ alone does not dictate the strength of path dependence. Coal and non-coal electricity are imperfect substitutes with the degree of substitutabilility, $\epsilon > 1$, determining how much their relative prices favor the less abundant good and thus fuel-specific capital. An increase in the relative coal price induces a contemporaneous decrease in relative coal capacity. However, because vintaged-capital is durable, there is now an excess supply of non-coal capacity in the subsequent period after the price shock has disappeared. The resulting higher prices for coal-fired electricity induces a relative increase in coal capacity such that, over time, relative coal capacity converges back to the pre-shock level. Thus, strong path dependence can only be achieved through a combination of strong increasing returns to scale and a high degree of substitutability between coal and non-coal electricity.[33] In the case where coal and non-coal electricity are weak substitutes, (i.e. a low ϵ or high $\frac{-\epsilon}{1-\epsilon}$), r must be sufficiently large such that the forces of increasing returns offset that of imperfect substitutability. Put in another way, strong path dependence occurs when the alternative fuel to coal produces electricity that can provide baseload power and exhibit strong increasing returns to scale.

7.2 Recovering the long-run elasticity of substitution

For a given $\alpha = r/2$, Eq. 11 provides two ways to recover the long-run elasticity of substitution, ϵ. The first approach uses the difference between the contemporaneous and one-generation lagged effects of relative coal prices on new relative coal capacity, $\frac{\alpha(1-\epsilon)\epsilon}{(\psi-1)^2} - \frac{\epsilon}{(\psi-1)}$. Because of the five decade average lifespan of electricity capital, this corresponds to $\widehat{\beta^O}_{\tau=2}$ in Column (6) of Table 2.[34] The first generation effect implies $\epsilon = 4$. A second approach uses the difference between the contemporaneous and two-generation lagged effects of relative coal prices on new relative coal capacity, $\frac{\alpha^2(1-\epsilon)^2\epsilon}{(\psi-1)^3} - \frac{\epsilon}{(\psi-1)}$, which equals $\widehat{\beta^O}_{\tau=7}$ in Column

[32] This results from a straightforward application of Euler's theorem. Formally, the cross partial derivative of a function $Y(X_t, X_{t-1})$ of homogeneous degree r can be written as:

$$\frac{\partial^2 Y_{ct}}{\partial X_t \partial X_{t-1}} = \left(\frac{r-1}{X_{t-1}}\right)\frac{\partial Y_{ct}}{\partial X_t} - \left(\frac{X_t}{X_{t-1}}\right)\frac{\partial^2 Y_{ct}}{\partial^2 X_t^2}$$

Setting $X_t = X_{t-1}$ so that one can compare the effects of lagged capacity against current capacity, it is evident that $\frac{\partial^2 Y_{ct}}{\partial X_t \partial X_{t-1}} > -\frac{\partial^2 Y_{ct}}{\partial^2 X_t^2}$ only in the presence of increasing returns to scale, or when $r > 1$.

[33] Notice the similarities here with the "price" and "market" effects found in standard models of directed technical change (Acemoglu, 2002; Acemoglu et al., 2012)

[34] To solve for ϵ analytically, one can rewrite $\frac{\alpha(1-\epsilon)\epsilon}{(\psi-1)^2} - \frac{\epsilon}{(\psi-1)} = \widehat{\beta^O}_{\tau=2}$ as the following quadratic form:

$$\epsilon^2[2(1-\widehat{\beta^O}_{\tau=2})\alpha - (1-\widehat{\beta^O}_{\tau=2}) + \widehat{\beta^O}_{\tau=2} * \alpha^2] + \epsilon[-2(1-\widehat{\beta^O}_{\tau=2})\alpha - 2\alpha^2\widehat{\beta^O}_{\tau=2}] + \widehat{\beta^O}_{\tau=2}\alpha^2 = 0$$

where the positive root of the above equation is the implied ϵ.

(6) of Table 2.[35] The second generation effect implies $\epsilon = 3.6$. In the context of Acemoglu et al. (2012), both elasticity values are well within the parameter space for which a temporary policy is sufficient to avoid climate change disaster.

7.3 Simulating future emissions from recent natural gas prices

The recent advent of hydraulic fracturing technology have increased natural gas supplies in the U.S. and elsewhere. Figure A.14 plots log coal to natural gas sales prices in the U.S. from 1970-2011. As a consequence of fracking, relative coal prices since 2009 have increased by 117% relative to the historic trend. To understand the future emission implications of this recent rise in relative coal prices, I use my recovered scale and elasticity of substitution parameters to simulate various future U.S. CO_2 emission pathways for the electricity.

To conduct my simulations, I employ the following simplifying assumptions. First, coal-fired capacity can only be substituted for natural-gas fired capacity.[36] Second, to avoid forecasting trends in fuel combustion efficiency and overall electricity demand, I assume that future fuel use per unit capacity and total coal and natural gas capacity are held at 2000-2009 levels. Third, I assume that in the absence of the relative coal price shock, relative coal prices and electricity depreciation rates are held at 2000-2009 levels. And finally, I assume that the scale and elasticity of substitution parameters are constant over the simulation period.

I generate future emission pathways by varying two simulation features. First, because both scale parameter, r, and the elasticity of substitution, ϵ, are empirically estimated, I conduct independent Monte Carlo draws of each parameter from the estimated empirical distribution. Second, I vary the duration of the current relative coal price shock, examining the implications of a 117% increase in relative coal prices lasting one, two, and three decades. Appendix D provides further details on my simulation procedure.

Simulated emission pathways are displayed in Figure 9. When the relative coal price shock lasts only one decade, as shown in the top panel, the induced substitution into natural gas-fired capacity induces an initial decline in emissions. However, under mean parameter values, the emissions trajectory eventually rises again once the lagged effects of the one period price shock gets overpowered by the coal capacity inducing effect of business-as-usual relative coal prices. Statistical uncertainty in r and ϵ generates path dependence of varying strength, broadly reflecting the bifurcation between cases of weak and strong path dependence. Taking into account statistical uncertainty, only 15% of the emission pathways drawn from the empirical distributions of r and ϵ exhibit declining emissions following a one period shock.

A long-term decline in emissions becomes far more likely following two-decade long price shock. The middle panel of Figure 9 shows a declining emissions pathway under mean parameter values, as path dependent effects of the two-decade shock is able to overcome the effects of business-as-usual relative coal prices thereafter. Drawing from the empirical distribution of the estimated parameters, 87% of simulated pathways now exhibit declining in emissions. The bottom panel of Figure 9 show that if the price shock lasts three decades, the likelihood of declining emissions increases further to 92%. In none of the simulations considered, however, do emissions ever reach zero, but rather they converge over time asymptotically to a steady-state level of emissions. This is because while natural gas is still less carbon intensive than coal, it is still a carbon containing fuel.

[35] The positive root of this 3rd order function is solved for numerically.

[36] In 2009, coal and natural gas constitute 92% of total U.S. fossil fuel-fired electricity capacity.

8 Discussion

This paper examines the role of path dependence - the lagged effects of transitory price shocks - on the development of coal-fired electricity capacity in the U.S. midwest over the 20th century. Exploiting variation in county delivered coal prices driven by the introduction of mechanized mining, I find that transitory price shocks have an increasing effect on the adoption of subsequent coal-fired capacity lasting ten decades, or about two generations of new capacity. Additional evidence find increasing returns to scale in electricity production as the likely mechanism causing path dependence. Interpreted through a multi-sector model of electricity production, my reduced-form estimates map onto a long-run elasticity of substitution between coal and non-coal electricity value of 4.

This historical evidence of path dependence in the electricity sector is particularly timely given the energy transition currently underway in the U.S. The recent spike in relative coal prices due to new sources of natural gas has led to both the large scale retirement of existing coal-fired capacity and the proposed construction of new natural gas-fired capacity. Simulations of future emissions using structural parameters recovered from this analysis suggest that a long-run decline in electricity sector emissions to a steady-state level is possible provided that current relative coal prices last at least two decades.

Whether such a scenario achieves optimal abatement depends on the severity of climate change damages. If climate damages are modest, then the same forces of path dependence that has historically led to increasing reliance on coal for electricity can be leveraged to endogenously expand natural gas-fired electricity even after relative prices return to favoring coal. However, natural gas still contains carbon. If climate damages turn out to be large, a structural transition towards zero-carbon electricity may ultimately be needed. In that case, a transition today towards natural gas may only further delay or increase the cost of the subsequent transition towards a fully decarbonized economy.

References

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

Acemoglu, Daron, and Veronica Guerrieri. 2008. "Capital Deepening and Nonbalanced Economic Growth." *Journal of Political Economy*, 116(3): 467–498.

Acemoglu, Daron., Philippe. Aghion, Leonardo Bursztyn, and David Hemous. 2012. "The Environment and Directed Technical Change." *American Economic Review*, 102(1): 131–66.

Acemoglu, Daron, Ufuk Akcigit, Douglas Hanley, and William Kerr. 2014. "Transition to Clean Technology." *Journal of Political Economy*, forthcoming.

Aghion, Philippe, and Peter Howitt. 1992. "A Model of Growth Through Creative Destruction." *Econometrica*, 60(2): pp. 323–351.

Aghion, Philippe, Antoine Dechezleprtre, David Hemous, Ralf Martin, and John Van Reenen. 2014. "Carbon Taxes, Path Dependency and Directed Technical Change: Evidence from the Auto Industry." *Journal of Political Economy*, forthcoming.

Arthur, W. Brian. 1994. *Increasing Returns and Path Dependence in the Economy.* University of Michigan Press.

Atkeson, Andrew, and Patrick J. Kehoe. 1999. "Models of Energy Use: Putty-Putty versus Putty-Clay." *American Economic Review*, 89(4): 1028–1043.

Avallone, Eugene A, Theodore Baumeister, and Ali Sadegh. 2006. *Marks' Standard Handbook For Mechanical Engineers (Standard Handbook for Mechanical Engineers).* Mcgraw-Hill Professional.

Barreca, Alan, Karen Clay, and Joel Tarr. 2014. "Coal, Smoke, and Death: Bituminous Coal and American Home Heating." National Bureau of Economic Research Working Paper 19881.

Baumol, William J. 1967. "Macroeconomics of unbalanced growth: the anatomy of urban crisis." *The American economic review*, 415–426.

Bleakley, Hoyt, and Jeffrey Lin. 2012. "Portage and Path Dependence." *The Quarterly Journal of Economics*, 127(2): 587–644.

Boden, T.A., G. Marland, and R.J. Andres. 2013. "Global, Regional, and National Fossil-Fuel CO2 Emissions." Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A. Working Paper.

BP. 2014. "Statistical Review of World Energy."

Campbell, Marius R. 1908. "Coal Fields of the United States." : U. S. Geological Survey Map.

Chay, Kenneth Y., and Michael Greenstone. 2003. "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession." *The Quarterly Journal of Economics*, 118(3): 1121–1167.

Chay, Kenneth Y., and Michael Greenstone. 2005. "Does Air Quality Matter? Evidence from the Housing Market." *Journal of Political Economy*, 113(2): pp. 376–424.

Christensen, Laurits R, and William H Greene. 1976. "Economies of scale in US electric power generation." *The Journal of Political Economy*, 655–676.

Coase, R. H. 1960. "The Problem of Social Cost." *Journal of Law and Economics*, 3: pp. 1–44.

Costas Azariadis, Allan Drazen. 1990. "Threshold Externalities in Economic Development." *The Quarterly Journal of Economics*, 105(2): 501–526.

David, Paul A. 1985. "Clio and the Economics of QWERTY." *The American Economic Review*, 75(2): pp. 332–337.

Dell, Melissa. 2012. "Path Dependence in Development: Evidence from the Mexican Revolution." mimeo.

East, J.A. 2012. "Coal fields of the conterminous United States, National Coal Resource Assessment." U.S. Geological Survey 2012-1205.

Energy Information Administration. 1980. "Inventory of Power Plants in the United States."

Energy Information Administration. 2012. "Annual Energy Review, 1982-2011."

Energy Information Administration. 2013. "Annual Energy Outlook."

Federal Power Commission. 1941. "Plant and Ownership List to Accompany Map 'Principal Electric Utility Generating Stations and Transmission Lines'." Government Printing Office, Washington, DC.

Federal Power Commission. 1945. "Plant and Ownership List to Accompany Map 'Principal Electric Utility Generating Stations and Transmission Lines'." Government Printing Office, Washington, DC.

Federal Power Commission. 1951. "Plant and Ownership List to Accompany Map 'Principal Electric Utility Generating Stations and Transmission Lines'." Government Printing Office, Washington, DC.

Federal Power Commission. 1954. "Plant and Ownership List to Accompany Map 'Principal Electric Utility Generating Stations and Transmission Lines'." Government Printing Office, Washington, DC.

Federal Power Commission. 1963. "Principal electric power facilities in the United States (map)." Government Printing Office, Washington, DC.

Fisher, Cassius A. 1910. "Depth and Minimum Thickness of Beds as Limiting Factors in Valuation." United States Geological Survey Bulletin 424.

Gaudet, Grard, Michel Moreaux, and Stephen W. Salant. 2001. "Intertemporal Depletion of Resource Sites by Spatially Distributed Users." *The American Economic Review*, 91(4): pp. 1149–1159.

Glaeser, Edward L, Sari Pekkala Kerr, and William R Kerr. 2015. "Entrepreneurship and urban growth: An empirical assessment with historical mines." *Review of Economics and Statistics*, 97(2): 498–520.

Haines, Michael R, and Inter-university Consortium for Political and Social Research. 2010. "Historical, Demographic, Economic, and Social Data: The United States, 1790-2002." *Inter-university Consortium for Political and Social Research (ICPSR) [distributor].*

Herfindahl, Orris C. 1967. "Depletion and economic theory." *Extractive resources and taxation*, 63–90.

Hornbeck, Richard. 2012. "The Enduring Impact of the American Dust Bowl: Short- and Long-Run Adjustments to Environmental Catastrophe." *American Economic Review*, 102(4): 1477–1507.

Hornbeck, Richard, and Pinar Keskin. 2014. "The Historically Evolving Impact of the Ogallala Aquifer: Agricultural Adaptation to Groundwater and Drought." *American Economic Journal: Applied Economics*, 6(1): 190–219.

Hornbeck, Richard, and Suresh Naidu. 2014. "When the Levee Breaks: Black Migration and Economic Development in the American South." *American Economic Review*, 104(3): 963–90.

Hotelling, Harold. 1929. "Stability in Competition." *The Economic Journal*, 39(153): pp. 41–57.

Hotelling, Harold. 1931. "The Economics of Exhaustible Resources." *The Journal of Political Economy*, 39(2): 137–135.

International Energy Agency. 2014. "CO2 Emissions from Fuel Combustion." OECD/IEA. Paris, France.

Jha, Akshaya. 2015. "Regulatory Induced Risk Aversion: Coal Procurement at U.S Power Plants." mimeo.

Joskow, Paul L. 1985. "Vertical Integration and Long-Term Contracts: The Case of Coal-Burning Electric Generating Plants." *Journal of Law, Economics, & Organization*, 1(1): 33–80.

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

Kline, Patrick, and Enrico Moretti. 2014. "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority." *The Quarterly Journal of Economics*, 129(1): 275–331.

Knittel, Christopher R., Konstantinos Metaxoglou, and Andre Trindade. 2015. "Natural Gas Prices and Coal Displacement: Evidence from Electricity Markets." National Bureau of Economic Research Working Paper 21627.

Kolstad, Charles D. 1994. "Hotelling Rents in Hotelling Space: Product Differentiation in Exhaustible Resource Markets." *Journal of Environmental Economics and Management*, 26(2): 163 – 180.

Komiya, Ryutaro. 1962. "Technological Progress and the Production Function in the United States Steam Power Industry." *The Review of Economics and Statistics*, 44(2): 156–166.

Lewis, Joshua. 2014. "Short-run and Long-run Effects of Household Electrification." mimeo.

McNerney, James, J. Doyne Farmer, and Jessika E. Trancik. 2011. "Historical costs of coal-fired electricity and implications for the future." *Energy Policy*, 39(6): 3042 – 3054.

Moreira, Marcelo J. 2003. "A conditional likelihood ratio test for structural models." *Econometrica*, 71(4): 1027–1048.

Murphy, Kevin M, and Andrei Shleifer. 1989. "Industrialization and the Big Push." *The Journal of Political Economy*, 97(5): 1003–1026.

Nerlove, Marc. 1963. "Returns to Scale in Electricity Supply." In *Measurement in Economics - Studies in Mathematical Economics and Econometrics in Memory of Yehuda Grunfeld.* , ed. Carl F. Christ. Stanford, CA:Stanford Univ. Press.

Ngai, L. Rachel, and Christopher A. Pissarides. 2007. "Structural Change in a Multisector Model of Growth." *American Economic Review*, 97(1): 429–443.

Nordhaus, William D. 1994. *Managing the Global Commons: The Economics of Climate Change.* The MIT Press.

Nunn, Nathan, and Nancy Qian. 2011. "The Potato's Contribution to Population and Urbanization: Evidence From A Historical Experiment." *The Quarterly Journal of Economics.*

Redding, Stephen J, Daniel M Sturm, and Nikolaus Wolf. 2011. "History and Industry Location: Evidence from German Airports." *Review of Economics and statistics*, 93(3): 814–831.

Rosenstein-Rodan, Paul N. 1943. "Problems of industrialisation of eastern and south-eastern Europe." *The Economic Journal*, 202–211.

Schmookler, Jacob. 1966. *Invention and Economic Growth.* Harvard University Press.

Schumpeter, J. 1942. *Capitalism, Socialism and Democracy.* Harper.

Severnini, Edson. 2014. "The Power of Hydroelectric Dams: Agglomeration spillovers." IZA Discussion Paper.

Solow, R.M., J. Tobin, C. C. von Weizscker, and M. Yaari. 1966. "Neoclassical Growth with Fixed Factor Proportions." *The Review of Economic Studies*, 33(2): 79–115.

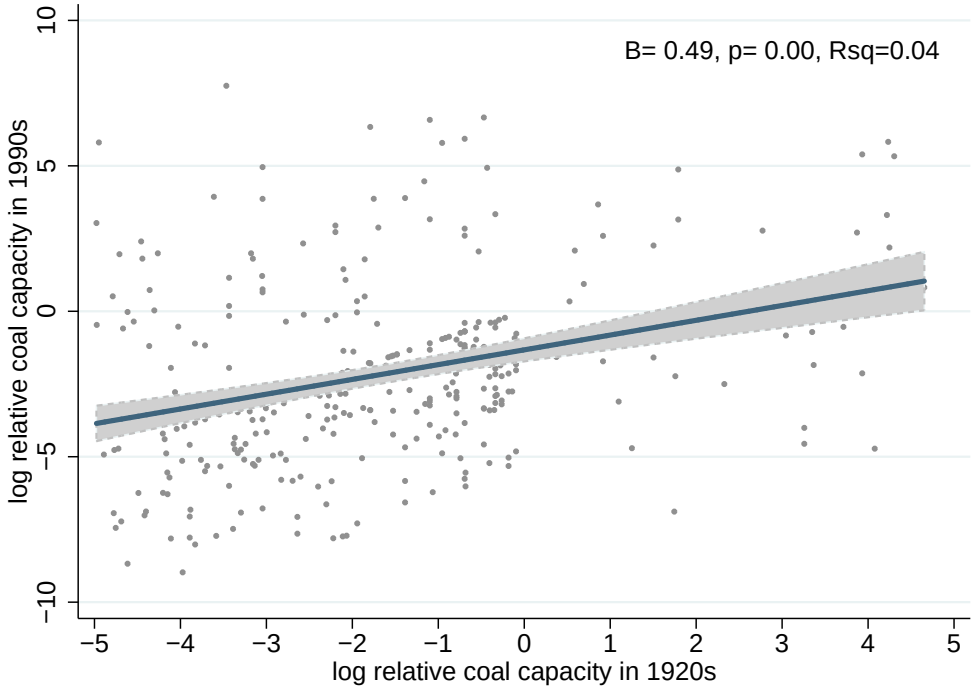
Speight, J. G. 1994. *The Chemistry and Technology of Coal, 2nd edition.* Marcel Dekker, Inc.

U.S. Census Bureau. 1975. "Historical Statistics of the United States, Colonial Times to 1970, Bicentennial Edition." U.S. Department of Commerce U.S. Geological Survey.

World Bank. 2014. "Data retrieved March, 2014, from World Development Indicators Online (WDI) database."

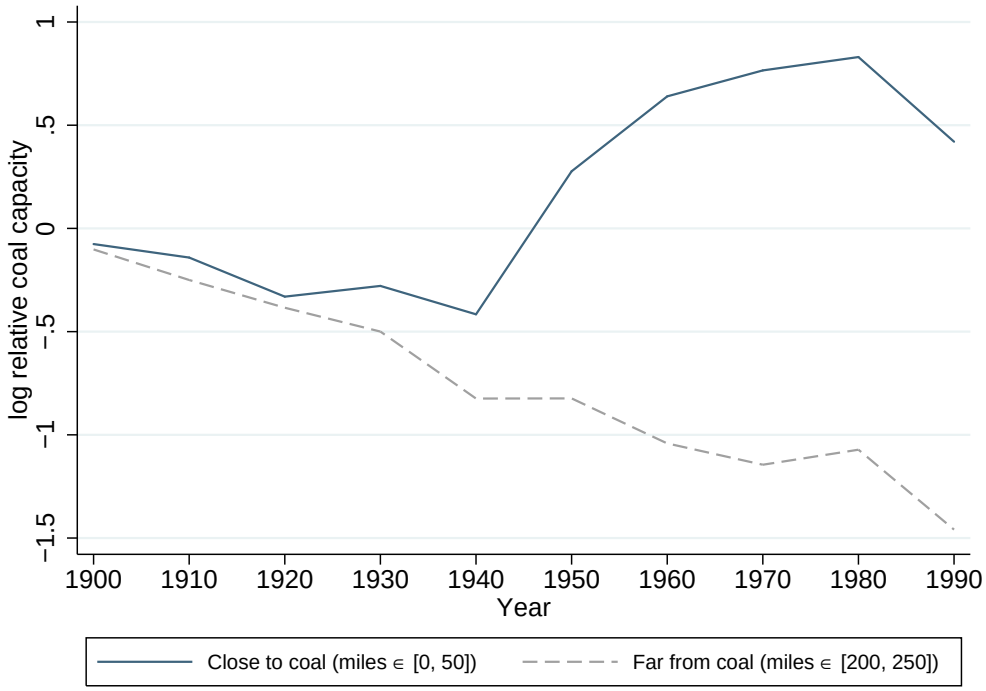
Figures

Figure 1: County relative coal capacity in the 1990's versus the 1920s



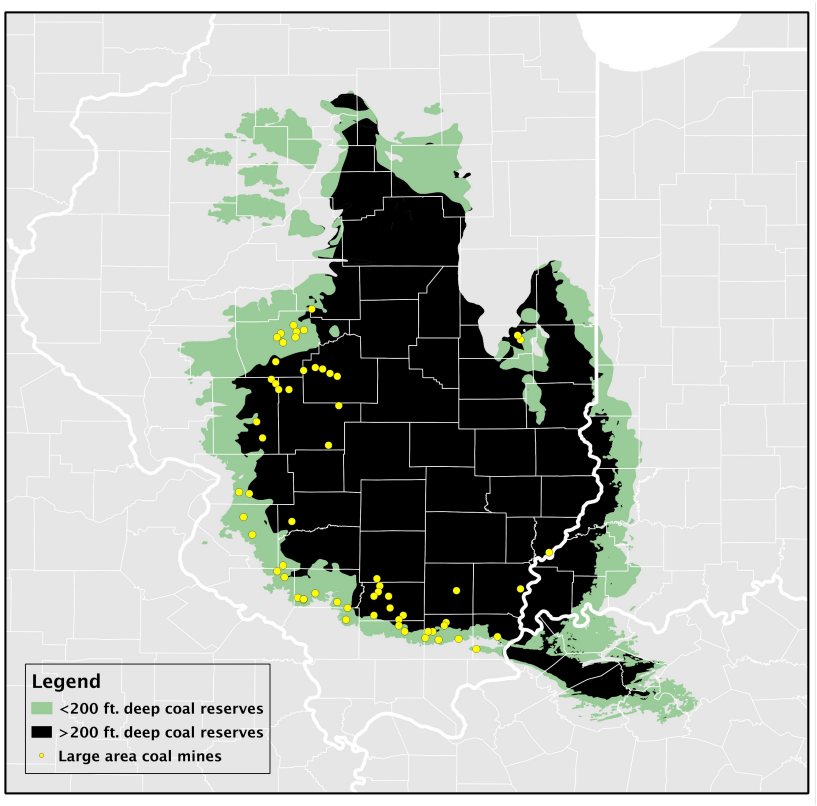
Notes: Scatter shows county-level log relative coal capacity in the 1990s against log relative coal capacity in the 1920s. Data reconstructed using modern EIA-860 forms. Data from modern EIA-860 forms (see Section 4.2 for details).

Figure 2: County relative coal capacity over 20th century for counties close to and far from Illinois basin



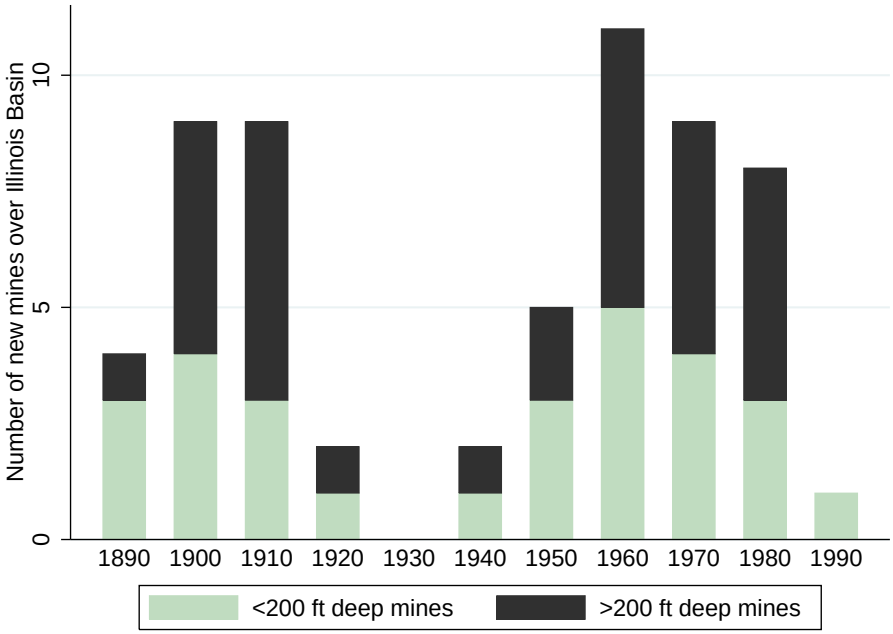
Notes: Average log relative coal capacity over time for counties less than 50 miles (solid blue) and between 200 and 250 miles (dashed gray) from the Illinois coal basin. Data reconstructed using modern EIA-860 forms. Data from modern EIA-860 forms. Illinois coal basin data from East (2012) (see Section 4.1 for details).

Figure 3: Map of Illinois coal basin by resource depth and mine location



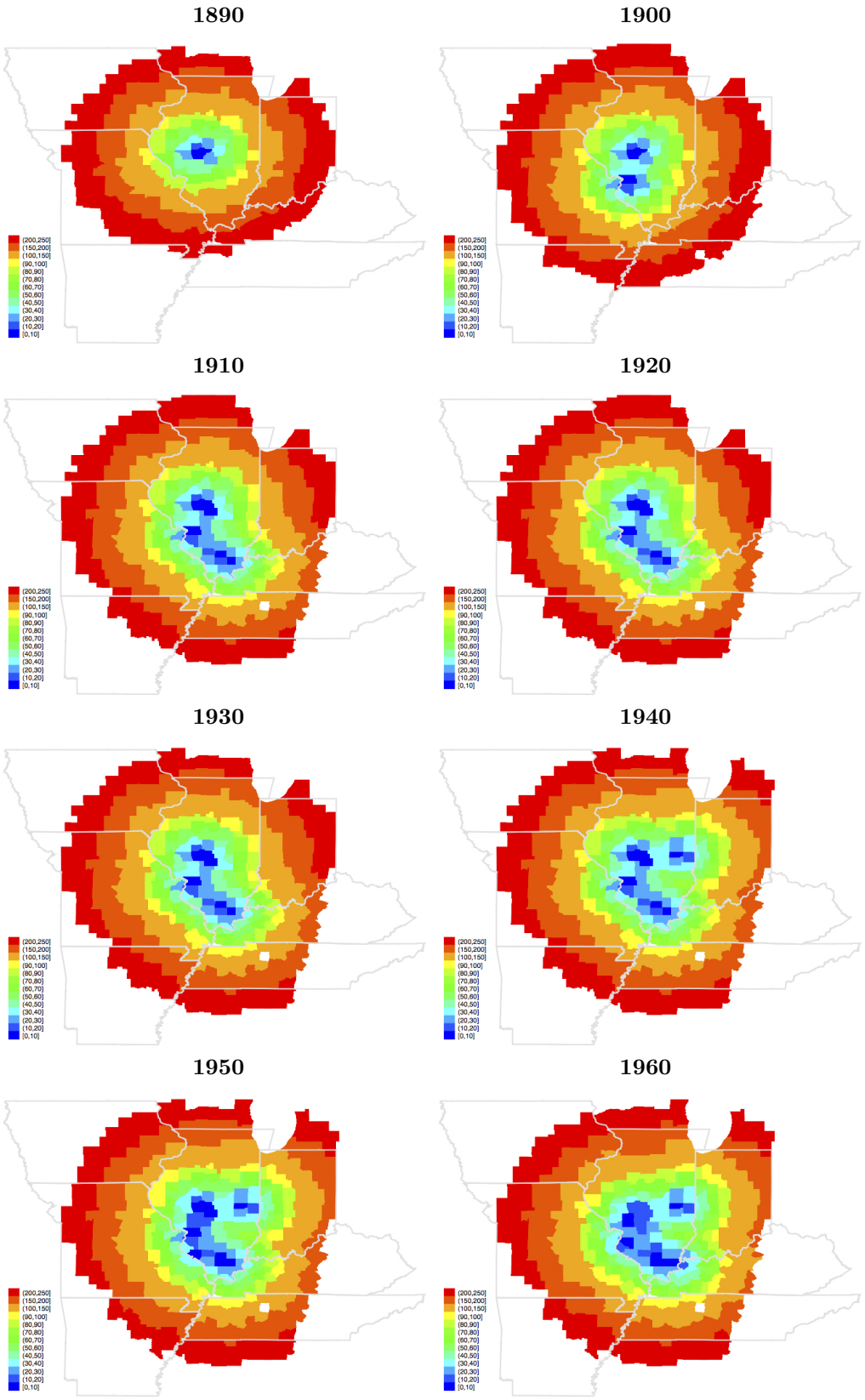
Notes: Green/lighter shaded area indicates the location of coal resources less than 200 feet deep. Black/darker shaded area indicates the location of coal resources more than 200 feet deep. Yellow points show location of large coal mines (>0.01 sq mile) that operated at any point after 1880. County and state boundaries shown. Source: East (2012)

Figure 4: Coal mining openings and depth over the Illinois Basin



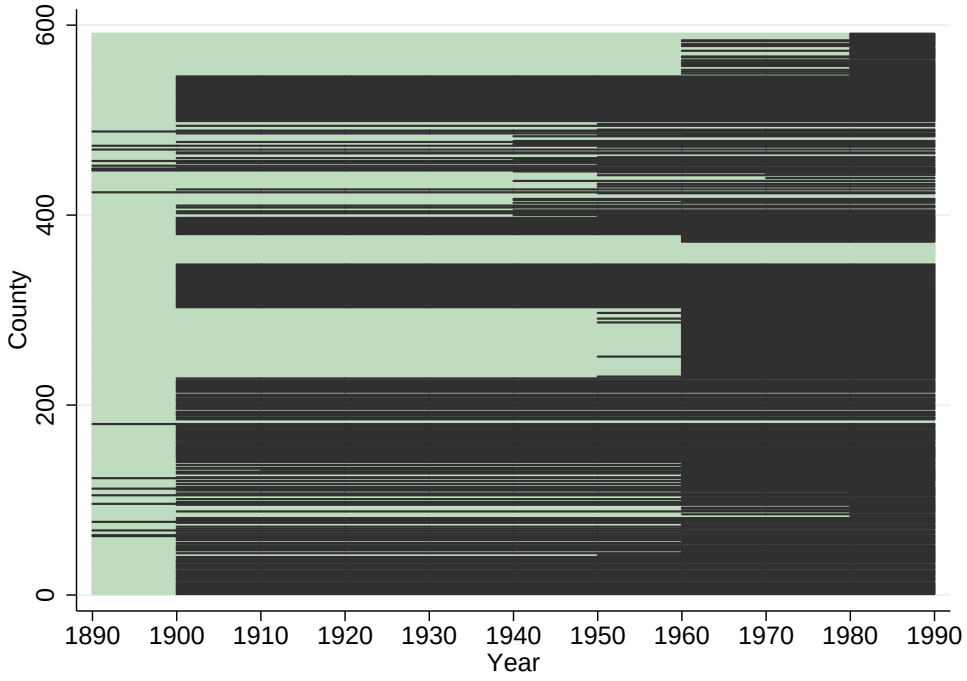
Notes: Plot shows the number of new large coal mines (>0.01 sq mile) opening per decade by depth of coal resource. Light bars shows mines located over coal resources less than 200 feet deep. Dark bars shows mines located over coal resources greater than 200 feet deep. Source: East (2012)

Figure 5: Constructed county coal prices by decade



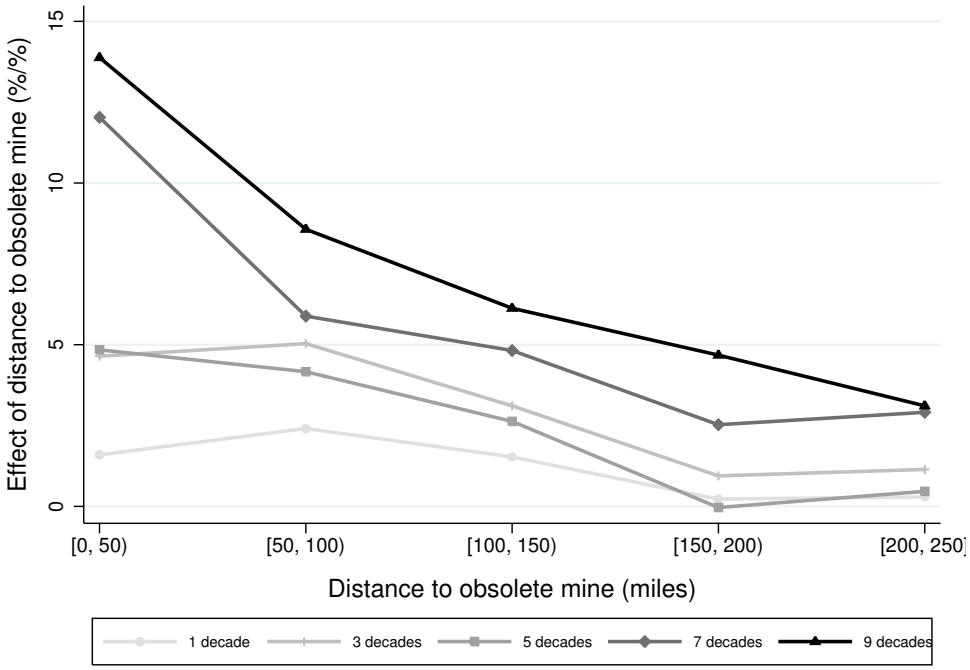
Notes: Constructed county coal prices as approximated by distance to nearest large coal mine from 1890-1960 over sample counties. See Section 4.1 for details.

Figure 6: Timing of obsolescence treatment for each sample county



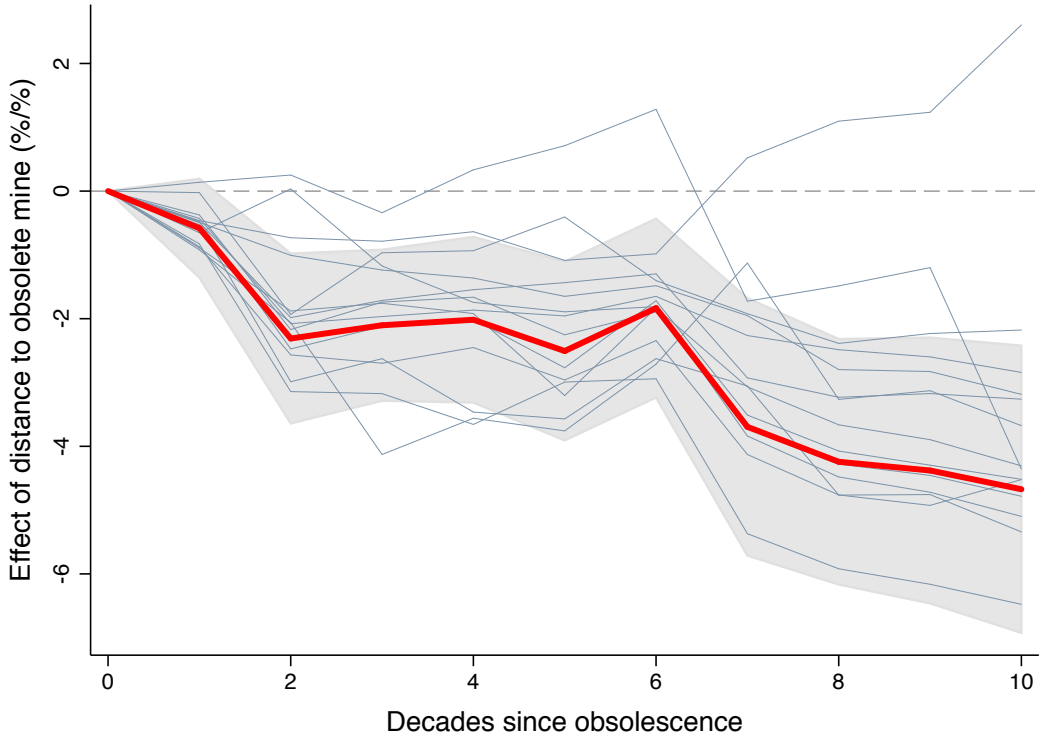
Notes: Plot shows the decade of obsolescence event for counties within the sample area (see Figure A.7). Light shading indicates decades prior to obsolesence event. Dark shading indicates decades during and after obsolescence event.

Figure 7: Nonlinearity in path dependence effects



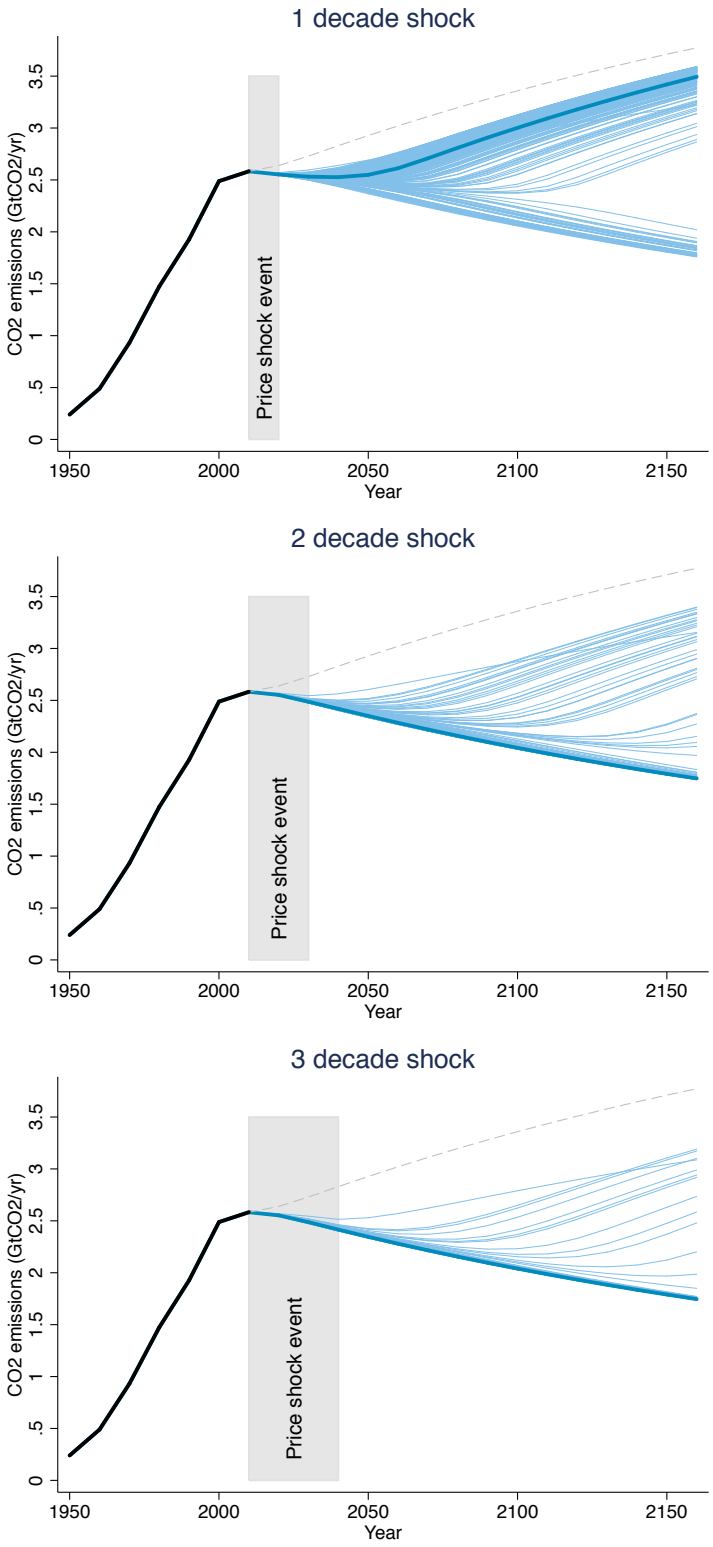
Notes: Plot examines the nonlinearity of path dependence effect. It shows distance bin dummies for the lagged effects of distance to obsolete mine one, three, five, seven, and nine decades after obsolescence.

Figure 8: Robustness in path dependence effects



Notes: Solid red line and shaded gray area shows coefficients β_τ^O and 90% confidence intervals from Eq. 3, corresponding to Column (5) in Table 2. Each narrow blue line shows a coefficient from the 14 robustness tests shown in Tables 4, 5, 6.

Figure 9: Simulating future CO_2 emissions following natural gas price shock



Notes: Top, middle, and bottom panels show simulated U.S. CO_2 electricity sector emissions from coal and natural gas-fired capacity following a relative coal price shock lasting one, two, and three decades, respectively. Shock defined as a 117% increase in relative coal prices compared to baseline levels. Solid black line shows historic emissions. Dashed gray line shows future emissions without price shock. Solid blue line shows projected emissions using mean estimates of scale and elasticity of substitution parameters. Thin blue lines from 250 Monte Carlo draws from the empirically estimated distributions of scale and elasticity of substitution parameters.

Tables

Table 1: Examining pre-obsolescence county characteristics

	(1)	(2)	(3)	(4)
	Pre-obsolescence	Pre-obsolescence		
Outcome	levels	changes		
Log population	-0.13	-0.016	-0.029**	-0.017
1890-1990	[0.086]	[0.013]	[0.012]	[0.011]
Obs	261	261	261	261
Log urban population	-0.085	0.019	0.015	0.011
1890-1980	[0.17]	[0.022]	[0.022]	[0.025]
Obs	183	171	171	171
log mgf establishments	-0.20**	0.046	0.0068	0.02
1890-1990	[0.098]	[0.064]	[0.049]	[0.054]
Obs	260	260	260	260
log mgf employment	-0.16	0.00096	-0.073	-0.086
1890-1990	[0.20]	[0.068]	[0.079]	[0.085]
Obs	258	255	255	255
log mfg capital	0.071	0.093	0.093	0.093
1890-1900	[0.24]	[0.15]	[0.15]	[0.15]
Obs	106	105	105	105
log mfg output	-0.44	0.044	-0.11	-0.016
1890-1940	[0.35]	[0.16]	[0.13]	[0.18]
Obs	114	113	113	113
State fixed effects	No	Yes	Yes	Yes
Year fixed effects	No	No	Yes	Yes
State trends	No	No	No	Yes

Notes: Each coefficient from a separate regression showing effect of distance to obsolete coal mine *before* it becomes obsolete. Outcomes are county-level characteristics at the beginning of the decade during which obsolescence occurs. Time coverage for outcome variables indicated below variable name. Column (1) estimate log outcome while Columns (2)-(4) estimate the difference in current and lagged log outcome. Columns (2)-(4) includes state fixed effects. Columns (3)-(4) includes year fixed effects. Column (4) includes state trends. Robust standard errors clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1

Table 2: Path dependence effects

	(1) cumulative coal	(2) new coal	(3) cumulative other	(4) new other	(5) cumulative relative coal	(6) new relative coal
Lagged effects of log dist. to obsolete mine						
1 decade later	-0.15	-0.59	-0.12	-1.13*	-0.58	-0.49
	[0.32]	[0.66]	[0.17]	[0.60]	[0.47]	[0.92]
2 decades later	-0.68***	-1.69***	0.072	-0.69	-2.31***	-1.72**
	[0.22]	[0.61]	[0.22]	[0.57]	[0.81]	[0.76]
3 decades later	-0.77***	-1.42**	-0.022	-1.34*	-2.10***	-1.04
	[0.23]	[0.60]	[0.31]	[0.69]	[0.72]	[0.71]
4 decades later	-0.87***	-1	0.17	-0.82	-2.02**	-0.33
	[0.25]	[0.91]	[0.34]	[0.72]	[0.79]	[1.11]
5 decades later	-0.98***	-0.41	0.47	0.25	-2.51***	-0.33
	[0.33]	[0.81]	[0.33]	[0.81]	[0.85]	[0.90]
6 decades later	-0.74**	-0.47	0.52	-0.65	-1.83**	-0.45
	[0.36]	[0.91]	[0.34]	[0.78]	[0.85]	[0.93]
7 decades later	-1.78***	-2.62*	0.23	-0.37	-3.70***	-3.15**
	[0.64]	[1.34]	[0.49]	[0.93]	[1.23]	[1.51]
8 decades later	-1.85***	-1.53	0.81	2.32	-4.24***	-1.29
	[0.66]	[1.43]	[0.70]	[1.63]	[1.17]	[1.74]
9 decades later	-1.98***	-1.7	0.79	-1.51	-4.38***	-1.7
	[0.70]	[1.70]	[0.76]	[1.07]	[1.27]	[1.99]
10 decades later	-2.13***	-1.88	0.53	-1.29	-4.67***	-1.99
	[0.74]	[4.08]	[0.90]	[1.35]	[1.37]	[4.63]
Contemp effect of log dist. to nearest mine	-0.80**	-0.86	-0.57	-1.99***	-1.15**	-1.32**
	[0.35]	[0.56]	[0.49]	[0.71]	[0.49]	[0.63]
Observations	2,871	2,871	2,871	2,871	2,871	2,871
Counties	261	261	261	261	261	261

Notes: Estimates of distance to obsolete mine using Equation 3. Sample includes counties that are less than 250 miles from nearest Illinois Basin coal basin and are closer to nearest Illinois Basin coal basin than to nearest Appalachian Basin coal basin. Each model includes event time dummies, contemporaneous distance to nearest mine, county fixed effects, year fixed effects, and state trends. Outcome in Columns (1) and (2) is cumulative and new coal capacity. Outcomes in Columns (3) and (4) is cumulative and new non-coal capacity. Outcome in Columns (5) and (6) are cumulative and new relative coal capacity. Robust standard errors clustered at the county-level. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Robustness in path dependence effects: controls

	(1)	(2)	(3)	(4)	(5)
	Outcome is cumulative relative coal capacity				
Lagged effects of log dist. to obsolete mine					
1 decade later	0.17	0.0085	0.34	-0.58	-0.96**
	[0.27]	[0.29]	[0.34]	[0.47]	[0.47]
2 decades later	-1.22**	-1.33**	-0.86*	-2.31***	-2.32***
	[0.55]	[0.59]	[0.51]	[0.81]	[0.68]
3 decades later	-1.06**	-1.05**	-0.47	-2.10***	-2.43***
	[0.44]	[0.45]	[0.33]	[0.72]	[0.56]
4 decades later	-0.93*	-0.96*	-0.25	-2.02**	-2.52***
	[0.52]	[0.57]	[0.40]	[0.79]	[0.66]
5 decades later	-0.96**	-1.13***	-0.87**	-2.51***	-3.38***
	[0.43]	[0.42]	[0.37]	[0.85]	[0.89]
6 decades later	0.71	0.33	-0.059	-1.83**	-2.77***
	[0.44]	[0.43]	[0.37]	[0.85]	[0.85]
7 decades later	-1.14	-1.91**	-2.08**	-3.70***	-4.39***
	[1.12]	[0.92]	[0.87]	[1.23]	[1.31]
8 decades later	-1.41	-2.17**	-2.36***	-4.24***	-5.21***
	[1.07]	[0.85]	[0.80]	[1.17]	[1.27]
9 decades later	-1.25	-2.00**	-2.19***	-4.38***	-5.32***
	[1.07]	[0.85]	[0.80]	[1.27]	[1.41]
10 decades later	-1.2	-1.96**	-2.14***	-4.67***	-5.73***
	[1.08]	[0.86]	[0.81]	[1.37]	[1.44]
Observations	2,871	2,871	2,871	2,871	2,478
No. of counties	261	261	261	261	261
County fixed effects	Yes	Yes	Yes	Yes	Yes
Obsolescence event-time effects	Yes	Yes	Yes	Yes	Yes
Contemp. dist. to nearest mine	No	Yes	Yes	Yes	Yes
Year fixed effects	No	No	Yes	Yes	Yes
State trends	No	No	No	Yes	Yes
County controls	No	No	No	No	Yes

Notes: Sample includes counties that are less than 250 miles from nearest Illinois Basin coal basin and are closer to nearest Illinois Basin coal basin than to nearest Appalachian Basin coal basin. Each model includes county fixed effects and event time dummies. Additional controls are added cumulatively across columns. Column (2) adds contemporaneous distance to nearest mine. Column (3) adds year fixed effects. Column (4) adds state trends and is same as benchmark model in Column (5) of Table 2. Column (6) adds total population, number of manufacturing establishments, and manufacturing employment. Robust standard errors clustered at the county-level. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Robustness in path dependence effects: estimating sample

	(1)	(2)	(3)	(4)	(5)
	Outcome is cumulative relative coal capacity				
Lagged effects of log dist. to obsolete mine					
1 decade later	-0.58	-0.47	-0.49	-0.43	-0.65**
	[0.47]	[0.51]	[0.44]	[0.45]	[0.29]
2 decades later	-2.31***	-2.47***	-1.98**	-2.08***	0.033
	[0.81]	[0.88]	[0.82]	[0.79]	[0.41]
3 decades later	-2.10***	-2.12***	-1.71**	-1.97***	-1.17*
	[0.72]	[0.77]	[0.77]	[0.70]	[0.62]
4 decades later	-2.02**	-2.01**	-1.55*	-1.87**	-1.74**
	[0.79]	[0.85]	[0.83]	[0.76]	[0.70]
5 decades later	-2.51***	-2.77***	-1.43	-1.96**	-1.89**
	[0.85]	[0.90]	[0.91]	[0.85]	[0.78]
6 decades later	-1.83**	-1.72*	-1.30	-1.65*	-1.81**
	[0.85]	[0.94]	[0.93]	[0.91]	[0.86]
7 decades later	-3.70***	-3.51**	-2.93**	-2.26**	-3.84***
	[1.23]	[1.37]	[1.21]	[1.12]	[1.29]
8 decades later	-4.24***	-4.07***	-3.23***	-2.49**	-4.48***
	[1.17]	[1.28]	[1.21]	[1.22]	[1.29]
9 decades later	-4.38***	-4.30***	-3.17**	-2.60*	-4.72***
	[1.27]	[1.37]	[1.34]	[1.34]	[1.41]
10 decades later	-4.67***	-4.52***	-3.26**	-2.84*	-5.10***
	[1.37]	[1.49]	[1.48]	[1.46]	[1.54]
Observations	2,871	2,288	3,520	3,718	1,452
No. of counties	261	208	320	338	132
Sample	Benchmark	<200 miles from Ill. basin	<300 miles from Ill. basin	Include closer to App. basin	Treated < 1960s

Notes: Estimates of distance to obsolete mine using Equation 3. All models include county fixed effects, event time dummies, contemporaneous distance to nearest mine, year fixed effects, and state trends. Column (1) replicates benchmark model from Column (5) of Table 2. Columns (2) and (3) restricts sample to counties within 200 and 300 miles from the nearest Illinois basin coal resource. Column (4) includes counties that are closer to Appalachian basin coal than to Illinois basin coal. Column (5) includes only counties experiencing obsolescence event prior to the 1960's. Robust standard errors clustered at the county-level. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Robustness in path dependence effects: data construction choices

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Outcome is cumulative relative coal capacity						
Lagged effects of log dist. to obsolete mine							
1 decade later	-0.58	0.14	-0.024	-0.83***	-0.53	-0.87**	-0.89*
	[0.47]	[0.21]	[0.50]	[0.28]	[0.47]	[0.41]	[0.51]
2 decades later	-2.31***	0.25	-1.93**	-3.14***	-2.99***	-2.57***	-1.88*
	[0.81]	[0.46]	[0.87]	[0.84]	[1.00]	[0.76]	[0.97]
3 decades later	-2.10***	-0.34	-0.97	-3.17***	-2.63**	-2.70***	-1.76**
	[0.72]	[0.43]	[0.74]	[0.79]	[1.03]	[0.71]	[0.87]
4 decades later	-2.02**	0.33	-0.93	-3.66***	-3.46***	-2.45***	-1.92*
	[0.79]	[0.52]	[0.84]	[0.96]	[1.14]	[0.73]	[0.99]
5 decades later	-2.51***	0.71	-0.41	-2.99***	-3.57***	-2.96***	-3.20***
	[0.85]	[0.68]	[1.03]	[0.94]	[1.35]	[0.81]	[1.05]
6 decades later	-1.83**	1.28	-1.40	-2.94***	-2.63*	-2.34***	-1.83*
	[0.85]	[1.06]	[1.34]	[1.00]	[1.50]	[0.81]	[1.01]
7 decades later	-3.70***	-1.73*	-1.92	-5.37***	-3.06	-4.13***	-3.07**
	[1.23]	[1.03]	[1.32]	[1.34]	[2.18]	[1.23]	[1.45]
8 decades later	-4.24***	-1.49	-2.39*	-5.92***	-4.76**	-4.76***	-3.66***
	[1.17]	[1.05]	[1.31]	[1.29]	[2.13]	[1.14]	[1.42]
9 decades later	-4.38***	-1.20	-2.23	-6.16***	-4.76**	-4.93***	-3.90***
	[1.27]	[1.14]	[1.42]	[1.41]	[2.37]	[1.23]	[1.47]
10 decades later	-4.67***	-4.36***	-2.18	-6.48***	-5.34**	-4.52***	-4.31***
	[1.37]	[1.56]	[1.52]	[1.51]	[2.46]	[1.28]	[1.53]
Observations	2,871	2,057	3,333	2,315	2,161	2,497	2,508
No. of counties	261	187	303	303	278	227	228
Min mine area (km^2)	.01	.005	.015	.01	.01	.01	.01
Num of nearest mines	1	1	1	2	3	1	1
EIA data year	1990-2012	1990-2012	1990-2012	1990-2012	1990-2012	2000	2005

Notes: All models include county fixed effects, event time dummies, contemporaneous distance to nearest mine, year fixed effects, and state trends. Column (1) shows benchmark model. Column (2) constructs obsolescence event using mines with area greater than 0.005 sq. miles. Column (3) constructs obsolescence event using mines with area greater than 0.015 sq. miles. Column (4) constructs obsolescence event using average distance to the nearest 2 mines with area greater than 0.01 sq. miles. Column (5) constructs obsolescence event using average distance to the nearest 3 mines with area greater than 0.01 sq. miles. Outcome in Column (6) built from only the 2000 EIA-860 form. Outcome in Column (7) built from only the 2005 EIA-860 form. Robust standard errors clustered at the county-level. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Robustness in path dependence effects: other modeling choices

	(1)	(2)	(3)	(4)	(5)
	Outcome is cumulative relative coal capacity				
Lagged effects of log dist. to obsolete mine					
1 decade later	-0.58	-0.37	-0.92	-0.49**	-0.46*
	[0.47]	[0.45]	[0.67]	[0.21]	[0.26]
2 decades later	-2.31***	-2.17***	-2.06	-1.01***	-0.73**
	[0.81]	[0.81]	[1.86]	[0.23]	[0.34]
3 decades later	-2.10***	-1.74***	-4.13**	-1.24***	-0.79*
	[0.72]	[0.67]	[1.67]	[0.32]	[0.45]
4 decades later	-2.02**	-1.66**	-3.56*	-1.36***	-0.64
	[0.79]	[0.75]	[1.89]	[0.36]	[0.58]
5 decades later	-2.51***	-2.25***	-3.76*	-1.65***	-1.09*
	[0.85]	[0.79]	[2.08]	[0.44]	[0.62]
6 decades later	-1.83**	-1.89**	-2.71	-1.48***	-0.98
	[0.85]	[0.84]	[2.51]	[0.52]	[0.70]
7 decades later	-3.70***	-3.70***	-1.13	-1.95**	0.52
	[1.23]	[1.21]	[2.49]	[0.78]	[0.88]
8 decades later	-4.24***	-4.28***	-3.27	-2.80***	1.10
	[1.17]	[1.17]	[2.34]	[0.88]	[0.96]
9 decades later	-4.38***	-4.46***	-3.13	-2.83***	1.23
	[1.27]	[1.28]	[2.36]	[0.99]	[1.20]
10 decades later	-4.67***	-4.78***	-3.68	-3.19***	2.60
	[1.37]	[1.40]	[2.72]	[1.12]	[1.71]
Observations	2,871	2,408	404	2,871	302
No. of counties	261	261	66	261	71
Model	Poisson	Poisson	Poisson	Linear	Linear
Add 1?	Yes	Yes	No	Yes	No
Multiple pre-event years	Yes	No	Yes	Yes	Yes

Notes: All models include county fixed effects, event time dummies, contemporaneous distance to nearest mine, year fixed effects, and state trends. Column (1) replicates benchmark model from Column (5) of Table 2. In Column (2), only the last year before obsolescence is set to $\tau = 0$. Column (3) estimates the effects on unadjusted relative coal capacity. Column (4) estimates a log-linear model on adjusted relative coal capacity. Column (5) estimates a log-linear model on unadjusted relative coal capacity. Robust standard errors clustered at the county-level. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Returns to scale regressions at the power plant level

	(1)	(2)	(3)	(4)	(5)	(6)
	Dep. var. is log non-fuel cost					
	OLS				IV	
log elec output (mwh)	0.56***	0.48***	0.57***	0.51***	0.60***	0.52***
	[0.028]	[0.035]	[0.052]	[0.054]	[0.087]	[0.072]
CLR P-value					0.0081	0.0075
CLR confidence int (90%)					[.54, .89]	[.44, .76]
implied scale parameter, r	1.78***	2.09***	1.75***	1.96***	1.66***	1.93***
	[0.090]	[0.15]	[0.16]	[0.21]	[0.24]	[0.27]
No. of power plants	103	147	32	73	96	139
County sample	Midwest	Midwest	U.S.	U.S.	Midwest	Midwest
Fuel	Only coal	Primarily coal	Primarily gas	Primarily oil	Only coal	Only coal

Notes: Estimates from Eq. 8 and 9. Each model includes average (1981-1999) power plant level fuel price, state and NERC region fixed effects and county centroid longitude and latitude. Columns (1), (2), (5), and (6) includes counties closer to nearest Illinois Basin coal reserve than to nearest Appalachian Basin coal reserve and that are less than 250 miles from Illinois coal basin. OLS regressions for power plants using primarily coal and only coal shown in Columns (1) and (2) respectively. IV regressions for power plants using primarily coal and only coal shown in Columns (5) and (6) respectively. OLS regressions in Columns (3) and (4) include all U.S. power plants burning primarily natural gas or oil. Robust standard errors clustered at the county-level. *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Productivity regressions at the generating-unit level

	(1)	(2)	(3)	(4)
	Dep. var. is log generator thermal efficiency			
Sample	Primarily coal	Primarily coal	Only coal	Only coal
$log(dist_i^O)XsinceEvent_i$	-0.0027	-0.0074	-0.012	-0.0083
	[0.0100]	[0.0070]	[0.014]	[0.010]
$log(dist_i^O)$	0.012	0.015	0.06	0.038
	[0.054]	[0.039]	[0.054]	[0.045]
$sinceEvent_i$	-0.008	0.027	0.022	0.021
	[0.039]	[0.027]	[0.058]	[0.039]
age		-0.0062***		-0.0057**
		[0.0018]		[0.0022]
capacity		0.000082		0.000062
		[0.000068]		[0.000084]
No. of generating units	266	266	160	160
No. of power plants	96	96	63	63

Notes: Estimates from Eq. 10. Each model includes average (1990-1995) county-level distance to nearest coal mine, state and NERC region fixed effects and county centroid longitude and latitude. All samples includes counties closer to nearest Illinois Basin coal reserve than to nearest Appalachian Basin coal reserve and that are less than 250 miles from Illinois coal basin. Columns (1) and (2) include all generators that burn coal as primary fuel. Columns (3) and (4) include only generators on coal-only power plants. Robust standard errors clustered at the county-level. *** p<0.01, ** p<0.05, * p<0.1.

Table 9: Examining electricity demand-side mechanisms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Decade since obsolescence event							
	1	2	3	4	5	6	7	8
Log population	-0.031	-0.049	-0.057**	-0.035	-0.069	-0.03	-0.041	-0.00035
1890-1990	[0.044]	[0.037]	[0.027]	[0.052]	[0.080]	[0.043]	[0.036]	[0.025]
No. of counties	261	231	229	132	119	107	107	107
Log urban population	0.015	-0.049	-0.0072	-0.53*	0.11	0.093*	-0.084	-0.062
1890-1980	[0.17]	[0.084]	[0.12]	[0.29]	[0.073]	[0.051]	[0.10]	[0.055]
No. of counties	153	164	92	89	84	88	93	97
log mgf establishments	0.055	0.24***	0.026	-0.31***	0.054	-0.031	0.021	0.061
1890-1990	[0.14]	[0.051]	[0.078]	[0.075]	[0.098]	[0.11]	[0.088]	[0.075]
No. of counties	154	125	221	123	117	107	107	107
log mgf employment	-0.012	0.27**	0.27	0.1	-0.16	0.071	0.065	-0.16
1890-1990	[0.24]	[0.12]	[0.16]	[0.31]	[0.24]	[0.18]	[0.095]	[0.19]
No. of counties	139	103	193	114	107	104	95	87

Notes: Outcomes in logs. Each coefficient from a separate regression showing effect of distance to obsolete coal mine *after* it becomes obsolete. Outcomes are first differenced log county-level characteristics during decades after the obsolescence event. Data coverage for outcome variables indicated below variable name. All models include contemporaneous log distance to nearest mine (coefficient not shown), state fixed effects, year fixed effects, and state trends. Robust standard errors clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1

Table 10: Effect of distance to obsolete coal mine on modern county characteristics

	(1)	(2)
	Rail density (2010)	Highway density (2010)
$ln(dist_i^O)XsinceEvent_i$	-0.031	-0.027
	[0.031]	[0.018]
$ln(dist_i^O)$	0.039	0.0021
	[0.17]	[0.11]
$sinceEvent_i$	0.13	0.075
	[0.13]	[0.073]
No. of counties	458	458

Notes: Cross-sectional estimates similar to specification in Eq. 10. Outcomes in logs. Sample includes counties closer to nearest Illinois Basin coal reserve than to nearest Appalachian Basin coal reserve. Each model includes county centorid latitude and longitude, state fixed effects, and NERC fixed effects. Robust standard errors clustered at the county-level. *** p<0.01, ** p<0.05, * p<0.1.

Appendix to

Path Dependence and Structural Change in U.S. Electricity Production

FOR ONLINE PUBLICATION

A Data Sources

This section details the data used in the paper.

A.1 Coal resources, mining, and modern delivered coal prices

U.S.G.S. National Coal Resource Assessment (N.C.R.A.) This paper uses two sets of spatial data from the N.C.R.A. (East, 2012). The first dataset contains vector-based shape files of Illinois and Appalachian Basin coal resources that are situated less than and above 200 feet from the surface. These shape files are used to generate the map in Figure A.7 showing each depth layer of coal resources for the two basins. The second dataset contains characteristics of all coal mines over the Illinois Basin that has operated since 1890. Variables include geo-coded location, opening year, closing year, and spatial area of the mine. Because the coal mines database does not include the depth of the coal resource being mined, this dataset is spatially overlaid onto the depth-specific coal resource shape files, as shown Figure 3, and used to generate a county-decade panel of delivered coal prices displayed in Figure 5. The opening of Illinois Basin coal mines since 1890 by depth is shown in Figure 4 and the moment of obsolescence for each sample county is shown in Figure 6.

EIA 423 forms (1972-1999) Verification of my constructed county delivered coal price panel dataset is performed against coal procurement data available from 1972-1999 EIA 423 forms. The raw data is at the destination power plant by county of coal origin level. It contains data on the quantity and price of the purchased coal as well as its heat, sulfur, and ash content. EIA notes that the county of coal origin may be unreliable.[37] A spot check of the data shows this to be particularly the case for EIA 423 forms before 1990 as the county identifier recorded does not match standard county identifiers.

This data is used to serve three objectives. First, using the 1990-1999 period where coal county of origin is more reliable, EIA 423 data on coal heat, sulfur, and heat content is aggregated to the county of origin and averaged across the years to produce the maps in Figure A.6 documenting relative homogeneity in coal quality in the Illinois Basin. Second, data from 1990-1999 EIA 423 forms are used to test underlying assumptions in the construction of delivered coal prices via the Hotelling local model as discussed in Section 4.1 and shown in Figures A.9 and A.10. Finally, because the county of the destination power plant is recorded relatively accurate, I use the entire set of 1972-1999 EIA 423 forms to construct decade-by-county level delivered coal prices to verify against my constructed delivered coal prices. This is used in the regression results in Table A.1.

A.2 Electricity production data

Generating unit-level data: EIA 860 forms (1980, 1990-2012) The EIA 860 generating-unit data is used for three purposes in this paper. First, fuel-specific capacity, county of location, opening year, and closing year is used to construct the county-by-decade-by-fuel capacity panel dataset from 1890-1990. The assumptions needed for this data construction, discussed in Section 4.2 are tested in Figure A.13 and Tables A.2 and A.3 using both 1990-2012 EIA 860 forms as well as the 1980 EIA 860 form, which was digitized

[37] In the EIA 423 readme file, it is noted that "The instructions for the FERC Form 423 require the respondent to report the county in which the coal was mined. However, this data is not always known or reported correctly. Sometimes a utility will report the location of a tipple or a prep plant as the source of the coal. In these cases, the coal usually comes from the surrounding counties. In some instances, a utility may report the county where MOST of the coal came from. It is very difficult to verify county level data. Users of the data should be aware of this and use the data accordingly."

for this paper. This data was also used when comparing with historical national and state-level data on generating capacity as shown in Figures A.11 and A.12. Second, this constructed panel of historical county-by-decade-by-fuel capacity is used to generate the prima-facie evidence in Figures 1 and 2. Third, the 1990-1995 EIA 860 forms include data on generator heat rate, which is the inverse of thermal efficiency.[38] I use log average thermal efficiency during 1990-1995 as the outcome variable in Table 8.

Power plant-level data: Historical FERC map and directories (1941, 1945, 1951, 1954, 1963) To verify my county-by-decade-by-fuel capacity dataset constructed from modern EIA 860 forms, I construct a state-by-decade-by-prime mover panel using the historical power plant dataset constructed by Lewis (2014). This dataset is constructed by first digitizing a map of power plants in 1963 (Federal Power Commission, 1963) which is then merged with power plant the timing of each power plant opening from directories on electric generating plants (Federal Power Commission, 1941, 1945, 1951, 1954). See Lewis (2014) for more details. This data is used in Figure A.12.

Power plant-level data: PLATTS/UDI O&M cost dataset (1981-1999) Power plant-level data on total and fuel costs for 1981-1999 was purchased from PLATTS/UDI. It is used to construct log non-fuel cost, the outcome variable in Table 7.

U.S. time series data (1890-2012) This paper uses two sources of aggregate U.S. time series data. U.S. Census Bureau (1975) provides total and mechanically produced U.S. bituminous coal production from 1890-1950 (see Figure A.3) and total electricity capacity by steam and hydropower from 1920-1970 (see Figure A.11). For the share of national delivered coal prices due to transport costs from 1902-2007), I use data from McNerney, Farmer and Trancik (2011). National coal and natural gas sales price from Energy Information Administration (2012) (see Figure A.14).

Cross country data Country-level per capita CO_2 emissions and GDP per capita in 1960 and 2000 come from Boden and Andres (2013) and World Bank (2014), respectively. Shown in Figure A.1. Country-level coal reserves in 2013, shown in Figure A.2, from BP (2014).

A.3 Other

County-level U.S. census data (1890-1990) County-level historical variables from 1890-2000 are obtained from historical U.S. censuses as collected by Haines and Inter-university Consortium for Political and Social Research (2010).[39] These variables include total population, rural population as well as manufacturing variables such as number of establishments, employment, capital value, and output. Because U.S. county boundaries were evolving until the 1930s, I redraw county-level data from 1890 to 1930 onto 1930 county boundaries using historical GIS county shape files from the U.S. National Historical Geographical Information System (N.H.G.I.S.)[40] using a method which modifies the procedure developed by Hornbeck (2012). This data is used to test for pre-trends in Table 1, as controls in my path dependence specification in Table 3, and in tests of alternative mechanisms in Table 9.

[38] Specifically, efficiency is 3412 divided by the heat rate where 3412 is the equivalent Btu content of a kWh of electricity. See `https://www.eia.gov/tools/faqs/faq.cfm?id=107&t=3` for more details.

[39] Available: `http://doi.org/10.3886/ICPSR02896.v3`

[40] Available: `https://www.nhgis.org/documentation/gis-data`

County-level transportation data (2010) Modern county-level data on highways and railroad networks was obtained from the U.S. Department of Transportation's National Transportation Atlas Database.[41] These variables are used as outcomes in the regressions shown in Table 10.

B Historical U.S. electricity data and its current availability

Section 4.2 mentioned that actual historical data on electricity capacity by input fuel at or below the county-level from 1890-1990 were either never collected or if collected may no longer exist. This section summarizes what relevant data were collected and the gap in data availability today which prompts for using data from modern EIA-860 forms to reconstruct a historical county-fuel-decade panel of electricity capacity.

1902-1917:

The U.S. federal government first collected power plant-level data in 1902 in the inaugural Central Electric Light and Power Station Census administered by the then Department of Commerce and Labor. This was repeated every five years in 1907, 1912, and 1917. Importantly, power plants are classified by prime-mover (i.e. steam, hydro, internal combustion) and not fuel input needed for this paper. Input costs are provided for each fuel. However, in order to recover an input-specific capacity, fuel price and a generating-unit capacity factor would be needed, neither of which were recorded

Availability Summaries of these censuses with aggregate statistics are available digitally.[42] However, the original power plant-level data could not be located following private conversations with archivists at the National Archives and Records Administration (NARA).[43]

1920-1970:

Soon after the creation of the Federal Power Commission (FPC) in 1920, several forms were administered annually to document electricity production and fuel consumption. The most important of these forms were the Annual Financial and Statistical Reports (Form 1) and the Power System Statements (Form 12). Form 1 shows generating-unit capacity and built year but not type of input fuel. Form 1 shows primary fuel usage by fuel type at the power-plant level. This data has two limitations. First, in order to recover plant-by-input fuel capacity, one needs plant-by-input fuel capacity factors, which is not available. Second, power plant coverage is incomplete. In 1948, Forms 1 and 12 included only 219 power plants covering 67% of total U.S. steam capacity.

Availability State-level aggregate statistics for electricity capacity by prime-mover and energy input consumption by fuel is available in the annual "Production of Energy and Capacity of Plants and Fuel Consumption of Electric Power Plants" and compiled in the "Electric Power Statistics, 1920-1940"[44]. The

[41] Available: `http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/national_transportation_atlas_database/2012/index.html`

[42] Available: `http://hdl.handle.net/2027/mdp.39015028113663`

[43] Once a document has exceeded its agency retention period (which pertains to all pre-1970 documents), only 3% of documents are deemed permanently valuable and retained in the public collection of at NARA.

[44] Available: `http://hdl.handle.net/2027/mdp.39015023906806`

"Steam-Electric Plant Construction Cost and Annual Production Expenses" from 1948 to 1974 has plant-level values from Form 1 and Form 12.[45] This data was also printed on a 1963 map called "Principal electric power facilities in the United States" and digitized by Lewis (2014) and used in the data validation exercises in Section 4.2. Private conversations with NARA archivists concluded that NARA may no longer hold the original Form 4 and Form 12 documents.

1977-1990:

The FPC was replaced by the Federal Energy Regulatory Commission (FERC) under the U.S. Department of Energy in 1977. FERC began publishing the "Inventory of Power Plants in the United States" that year combining generating-unit data from the Monthly Power Plant Report (Form 4), Annual Power System Statement (Form 12), and the Supplemental Power Statement (Form 12E-2). This report includes generating-unit capacity, type of energy input, and built year. Unfortunately, retired units documented in the "Inventory of Power Plants in the United States" only included units that were retired during record year and not previously retired generators.

Availability The "Inventory of Power Plants in the United States" for 1980 is available online[46] and was digitized for my data validation exercises discussed in Section 4.2. Reports from other years are kept as microfiche in many research libraries.

1985 - :

The Annual Electric Generator Report, Form EIA-860, was originally implemented in 1985. It replaced the previous Form 4, Form 12 and 12E, Form 67, and Form 411. According to the U.S. Energy Information Agency which administers the form:

> "The Form EIA-860 is a mandatory annual census of all existing and planned electric generating facilities in the United States with a total generator nameplate capacity of 1 or more megawatts. The survey is used to collect data on existing power plants and 10 year plans for constructing new plants, as well as generating unit additions, modifications, and retirements in existing plants. Data on the survey are collected at the individual generator level."

Availability Available digitally online.[47]

[45] Available: `http://catalog.hathitrust.org/Record/000904499`
[46] Available: `http://hdl.handle.net/2027/umn.31951d02987924n`
[47] Available: `http://www.eia.gov/electricity/data/eia860/`

C Electricity producer problem

C.1 Profit maximization

For the final good producer, the full optimization problem is:

$$\max_{X_{ct},X_{nt},E_{ct},E_{nt}} \left(Y_{ct}^{(\epsilon-1)/\epsilon} + Y_{nt}^{(\epsilon-1)/\epsilon}\right)^{\epsilon/(\epsilon-1)} - z_{ct}(E_{ct}+E_{ct-1}) - z_{nt}(E_{nt}+E_{nt-1}) - r_t(X_{ct}+X_{nt})$$

where z_{jt} is the primary energy input price and r_t is the rental rate of capital. Inserting the intermediate good production function from Eq. 6 and using the fixed-factor proportions for generating units, the problem can be simplified in terms of primary energy inputs:

$$\begin{aligned}\max_{E_{ct},E_{nt}} \Big(((A_{Ect}E_{ct})^{\alpha}(\delta A_{Ect-1}E_{ct-1})^{\alpha})^{(\epsilon-1)/\epsilon} + ((A_{Ent}E_{nt})^{\alpha}(\delta A_{Ent-1}E_{nt-1})^{\alpha})^{(\epsilon-1)/\epsilon}\Big)^{\epsilon/(\epsilon-1)} \\ -z_{ct}(E_{ct}+E_{ct-1}) - z_{nt}(E_{nt}+E_{nt-1}) - r_t(\frac{A_{Ect}}{A_{Xct}}E_{ct} + \frac{A_{Ent}}{A_{Xnt}}E_{nt})\end{aligned} \tag{A.1}$$

Suppose input-specific productivities are the same across intermediate goods during period $t-1$, that is $A_{Xct-1} = A_{Xnt-1}$ and $A_{Ect-1} = A_{Ent-1}$. Taking the first order condition of Eq. A.1, obtaining the ratio of first order conditions for coal and non-coal electricity, and rewriting in terms of relative current vintage coal capacity, $\widetilde{X}_t = \frac{X_{ct}}{X_{nt}}$:

$$\widetilde{X}_t = \widetilde{w}_t^{\frac{\epsilon}{\psi-1}} \widetilde{X}_{t-1}^{\frac{\alpha(1-\epsilon)}{\psi-1}} \widetilde{A}_{Et}^{\frac{-\epsilon}{\psi-1}} \widetilde{A}_{Xt} \tag{A.2}$$

where $w_{jt} = z_{jt} + \frac{A_{Ejt}}{A_{Xjt}} r_t$ is the productivity weighted input price index. Eq. A.2 is Eq. 7 in the main text.

C.2 Cost minimization problem to recover scale parameter

Consider a power plant m containing only coal-fired generating units such that one can drop index j and only needs to consider optimizing Eq. 6. Applying the fixed input proportions $E_{mt} = \frac{A_{Xmt}}{A_{Emt}} X_{mt}$ and $E_{mt-1} = \frac{A_{Xmt-1}}{A_{Emt-1}} X_{mt-1}$, the constrained cost minimization problem becomes:

$$\begin{aligned} C(w_{mt}, r_{mt}, Y_{mt}) = \underset{E_{mt}}{\text{minimize}} \quad & z_{mt}\left(E_{mt} + E_{mt-1}\right) + r_t(\frac{A_{Xmt}}{A_{Emt}}E_{mt}) \\ s.t. \quad & Y_{mt} \leq (A_{Emt}E_{mt})^{\alpha}(\delta A_{Xmt-1}E_{mt-1})^{\alpha} \\ & X_{mt} \leq 0 \end{aligned}$$

where coal input for past capital E_{mt-1} can be viewed as a fixed cost. Rewriting the production function as $E_{mt-1} = Y_{mt}^{1/\alpha}(A_{Emt}A_{Emt-1}E_{mt}\delta)^{-1}$, one can write the following equivalent unconstrained minimization problem:

$$\underset{X_{mt}}{\text{minimize}} \quad w_{mt}E_{mt} + z_{mt}Y_{mt}^{1/\alpha}(A_{Emt}A_{Emt-1}E_{mt}\delta)^{-1} \tag{A.3}$$

where again $w_{mt} = z_{mt} - \frac{A_{Xmt}}{A_{Emt}} r$ is the productivity weighted input price index. Taking the first order condition of Eq. A.3 yields the following conditional demand function:

$$E_{mt} = Y_{mt}^{1/2\alpha} w_{mt}^{1/2} z_{mt}^{-1/2} (A_{Emt}A_{Emt-1}\delta)^{-1/2} \tag{A.4}$$

Inserting Eq. A.4 into non-fuel cost at the cost-minimizing input level, $\text{non_fuel_cost}_{mt} = C(w_{mt}, r_{mt}, Y_{mt}) - w_{mt}(E_{mt} + E_{mt-1}) = r_t(\frac{A_{Xmt}}{A_{Emt}}E_t)$, and applying a log transformation yields:

$$\ln(\text{non_fuel_cost}_{mt}) = \frac{1}{2\alpha}\ln Y_{mt} + \frac{1}{2}\ln w_{mt} + \ln r_t - \frac{1}{2}\ln(z_{mt}) - \ln(A_{Xmt}A_{Emt}^{-3/2}A_{Emt-1}^{-1/2}\delta^{-1/2}) \quad \text{(A.5)}$$

which is the structural analog to the OLS specification of Eq. 8 in the main text.

D Simulating future emissions

This section details the procedure for simulating future CO_2 emissions following a relative coal price shock, as shown in Figure 9.

D.1 Parameters

- Estimated returns to scale parameter with mean $\hat{r} = 1.66$ and standard error $\hat{\sigma}_r = 0.24$.
- Estimated new capacity one-generation lagged effect with mean $\widehat{\beta}^0 = -1.72$ and standard error $\hat{\sigma}_{\beta^0} = 0.76$.
- Business-as-usual relative coal prices from 2010 set to 2000's value, $\widetilde{w}_t = 0.7$.
- Relative coal price shock based on 117% increase in relative coal prices during 2009-2011 (see Figure A.14), set at $\widetilde{w}_t = 0.7(1 + 1.17)$.
- Capital depreciation rate set to 2000's value, $\delta = 0.04$.
- Carbon content of coal: $C_c = 4931.3$ lb CO_2/short ton coal. [48]
- Carbon content of natural gas: $C_n = 119.9$ lb CO_2/ thousand cubic feet.[49]

D.2 Historical emissions

After aggregating coal and natural gas capacity to the national level, emissions is calculated by:

$$M_t = E_{ct}C_c + E_{nt}C_n$$

where E_{ct} and E_{nt} is total coal (in short tons) and natural gas (in thousand cubic feet) consumed respectively by the U.S. electricity sector in year t. Data from Energy Information Administration (2012).

D.3 Simulating future emissions

For a relative price shock lasting one, two, or three decades, conduct the following Monte Carlo procedure with $b = 1...250$ draws:

- Draw $r(b) \sim N(\hat{r}, \hat{\sigma_r})$ and $\beta^0(b) \sim N(\widehat{\beta}^0, \hat{\sigma}_{\beta^0})$. Define $\alpha(b) = r(b)/2$.

[48] Available here: `https://www.eia.gov/environment/emissions/co2_vol_mass.cfm`
[49] Available here: `https://www.eia.gov/environment/emissions/co2_vol_mass.cfm`

- Obtain $\epsilon(b)$ by taking the positive root from:

$$\epsilon(b)^2[2(1-\widehat{\beta}(b)^O)\alpha(b)-(1-\widehat{\beta}(b)^O)+\widehat{\beta}(b)^O*\alpha(b)^2]+\epsilon(b)[-2(1-\widehat{\beta}(b)^O)\alpha(b)-2\alpha(b)^2\widehat{\beta}(b)^O]+\widehat{\beta}(b)^O\alpha(b)^2=0$$

  Define $\psi(b)=(1-\alpha(b))(1-\epsilon(b))$.

- For each decade $t = 2010...2150$, implement:

  1. Apply Eq. 11 to obtain new relative coal capacity:

$$\widetilde{\Delta X}(b)_t = exp\left(\sum_{s=0}^{\infty}\frac{\epsilon(b)}{(\psi(b)-1)}\left(\frac{\alpha(b)(1-\epsilon(b))}{(\psi(b)-1)}\right)^s \ln \widetilde{w}_{t-s}\right)$$

  2. Obtain new coal capacity, holding total capacity constant from previous period:

$$\Delta X(b)_{ct} = \left(\frac{1}{\frac{1}{\widetilde{X}(b)_t}+1}\right)(X(b)_{ct-1}+X(b)_{nt-1})\,\delta$$

  3. Obtain new natural gas capacity, holding total capacity constant from previous period:

$$\Delta X(b)_{nt} = \left(1-\frac{1}{\frac{1}{\widetilde{X}(b)_t}+1}\right)(X(b)_{ct-1}+X(b)_{nt-1})\,\delta$$

  4. Obtain cumulative coal capacity:

$$X(b)_{ct} = X(b)_{ct-1}(1-\delta)+\Delta X(b)_{ct}$$

  5. Obtain cumulative natural gas capacity:

$$X(b)_{nt} = X(b)_{nt-1}(1-\delta)+\Delta X(b)_{nt}$$

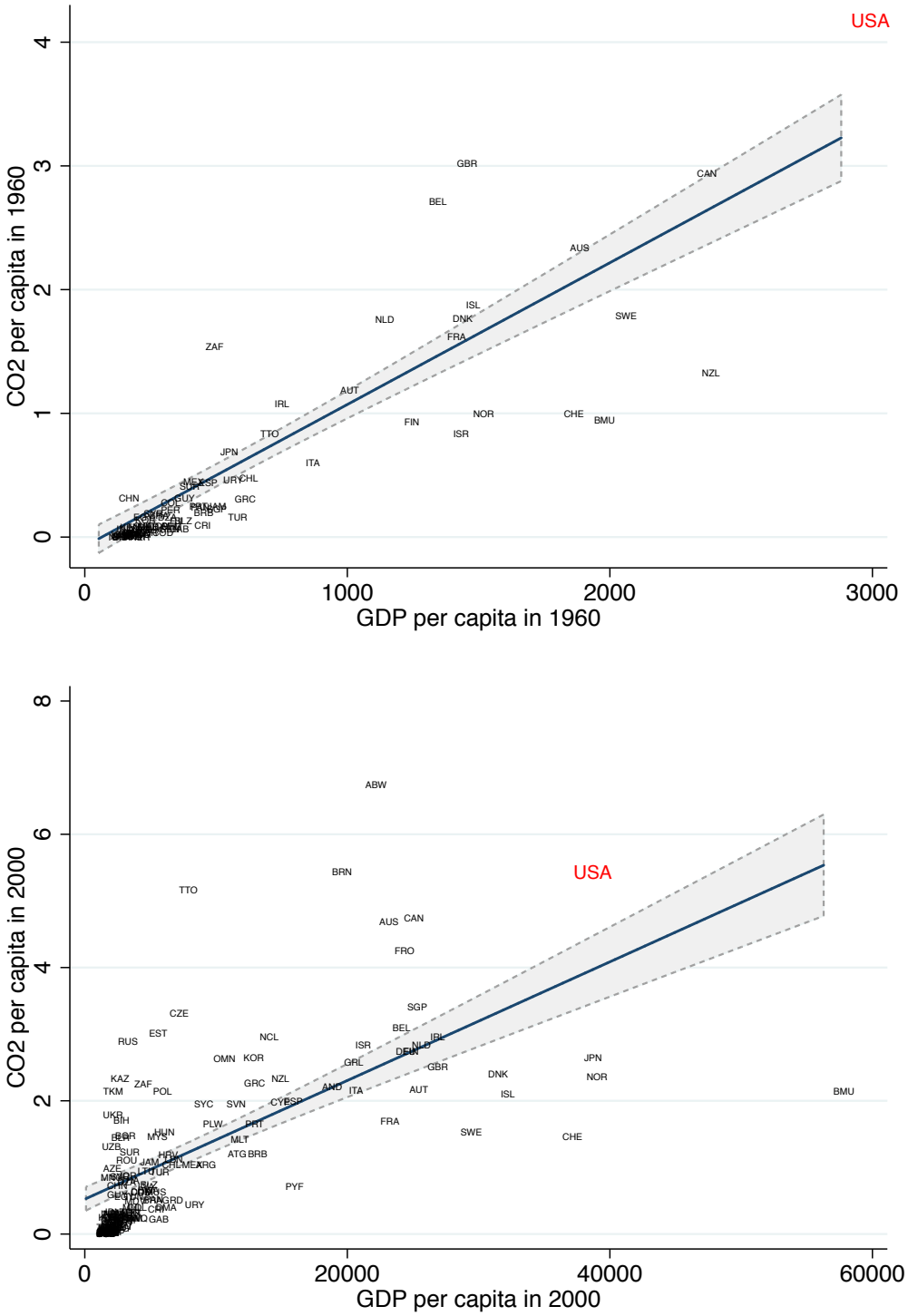
  6. Obtain total CO_2 emissions using 2010 emissions intensity:

$$M(b)_t = X(b)_{ct}\frac{E_{c2010}}{X_{c2000}}C_c + X(b)_{nt}\frac{E_{n2010}}{X_{n2000}}C_n$$

This procedure is implemented once under business-as-usual relative coal prices.

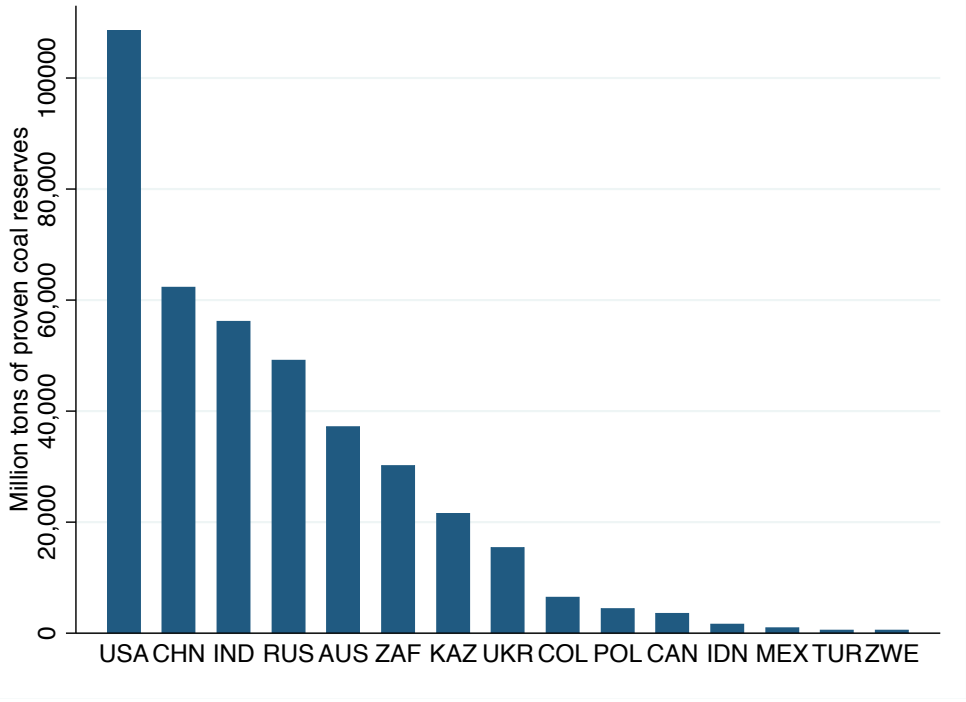
Appendix Figures

Figure A.1: U.S. CO_2 emissions intensity in 1960 and 2000



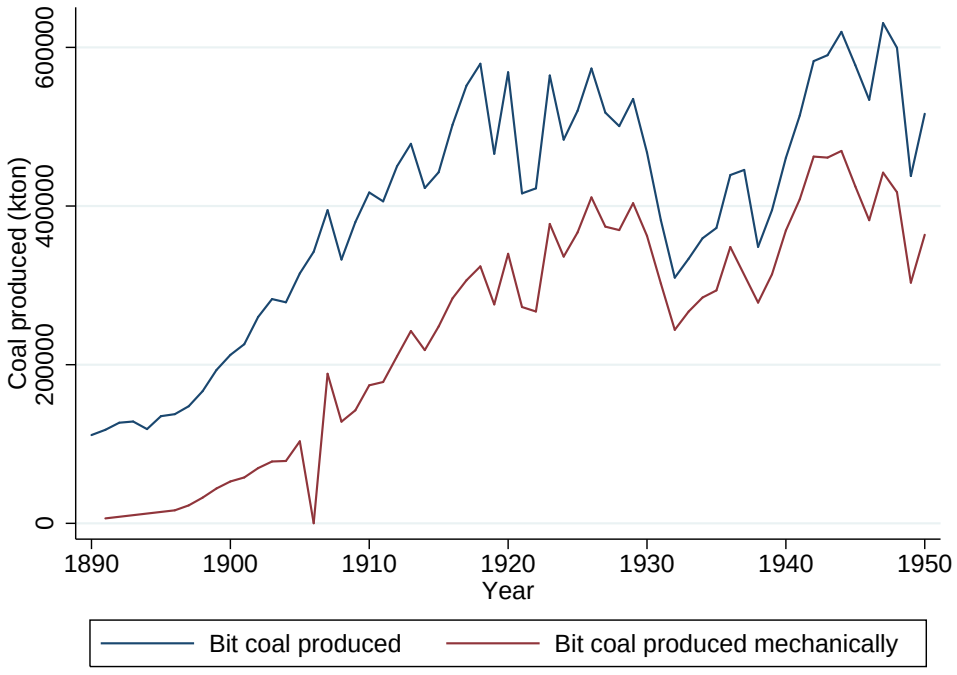
Notes: Scatterplots shows carbon dioxide emissions per capita against income in 1960 (top panel) and 2000 (bottom panel). Linear regression fit shown with 90% confidence interval. OPEC countries excluded. Source: Boden and Andres (2013) and World Bank (2014).

Figure A.2: Country-level coal reserves



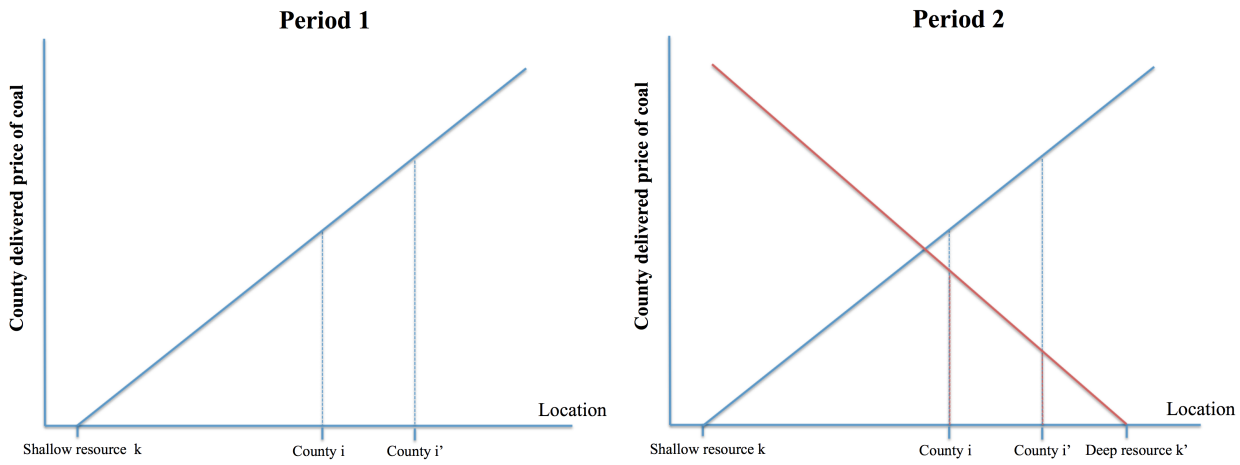
Notes: Known country-level coal reserves in 2013. Source: BP (2014).

Figure A.3: U.S. bituminous coal production and mechanization over time



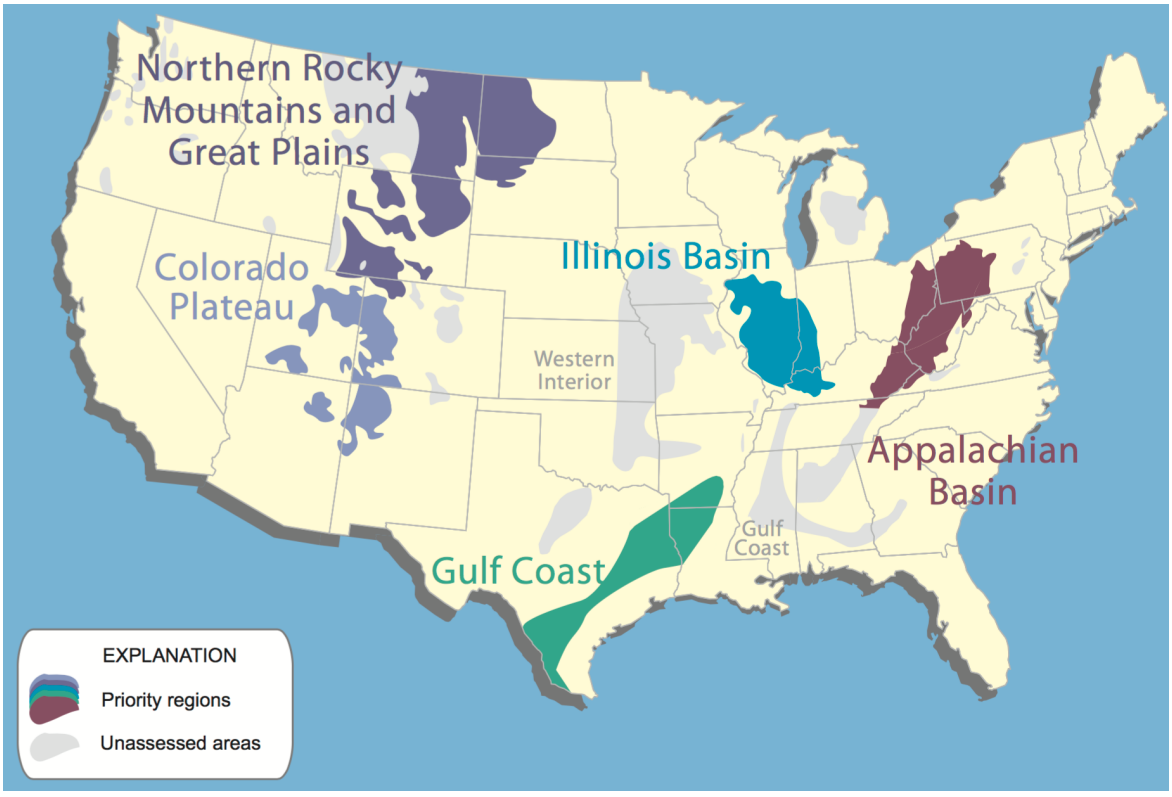
Notes: Bituminous coal production overall (blue) and from mechanized extraction (red). Source: U.S. Census Bureau (1975).

Figure A.4: A 2x2x2 spatial Hotelling location model



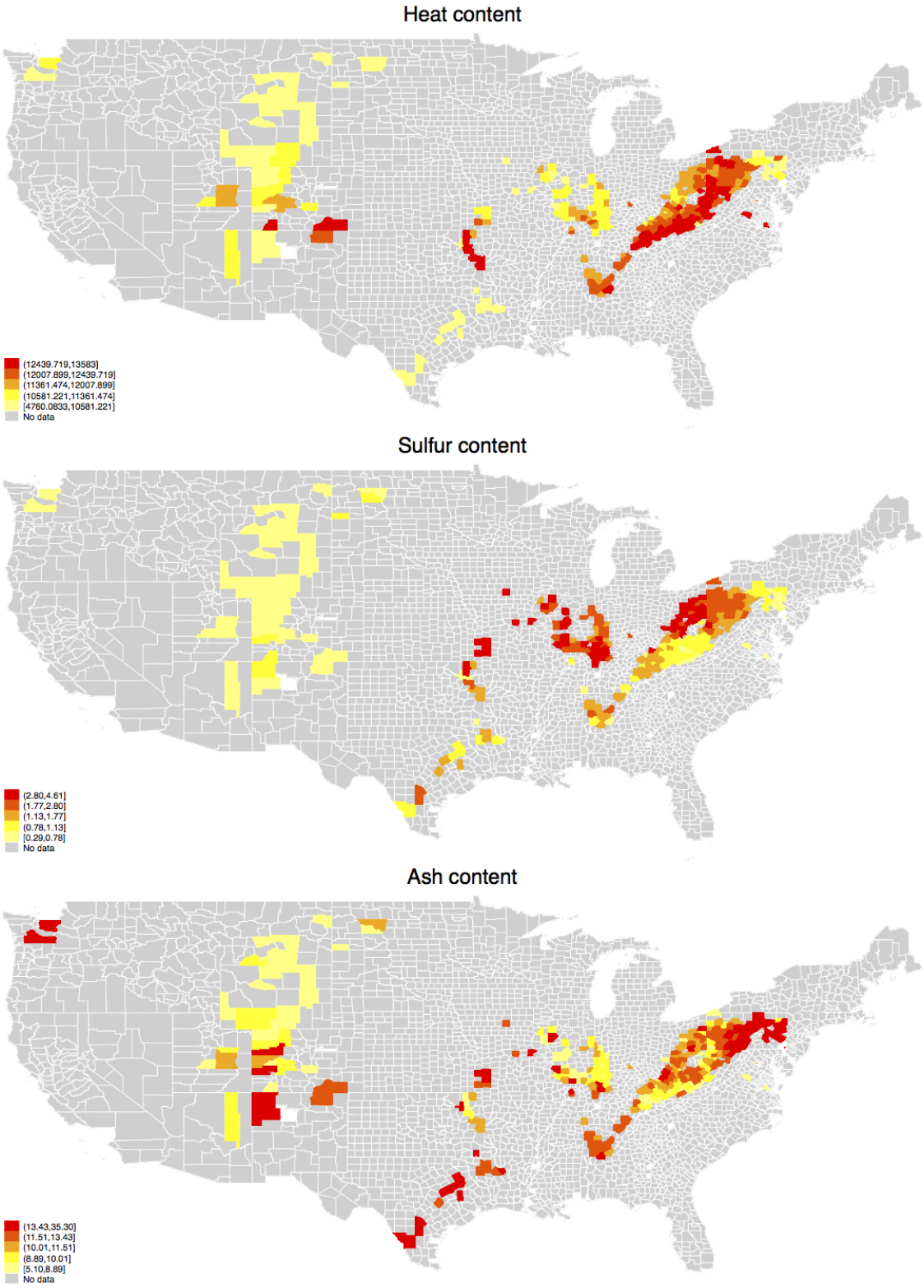
Notes: Schematic for a 2 county, 2 coal resource, 2 period Hotelling location model. See Section 3.2 for discussion.

Figure A.5: U.S. coal basins



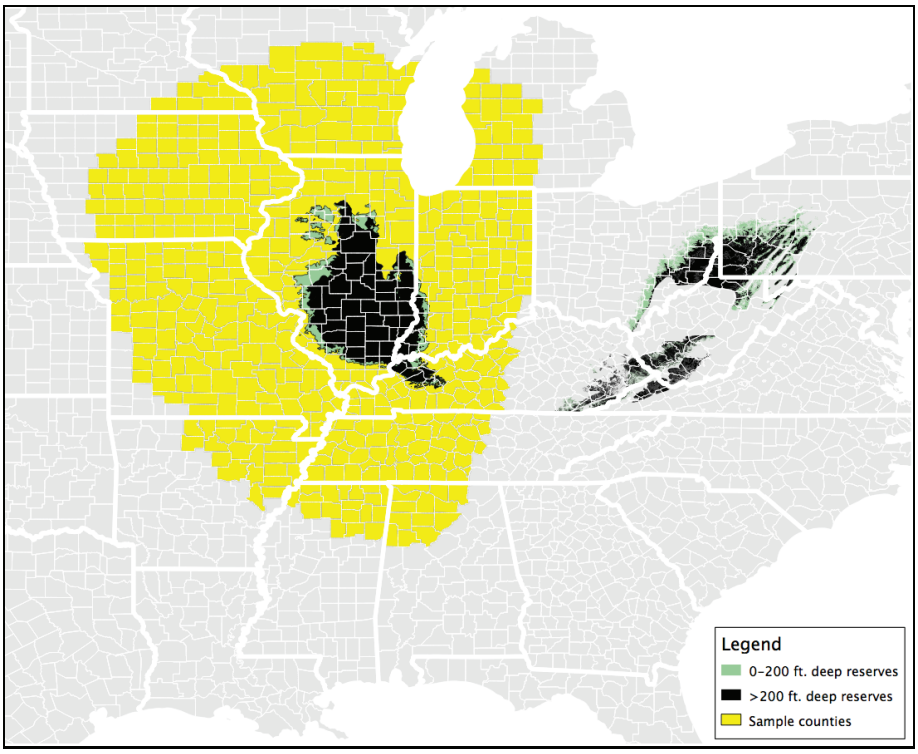
Notes: U.S. coal basins. Reproduced from U.S.G.S. National Coal Resource Assessment (East, 2012)

Figure A.6: Coal quality across the U.S.



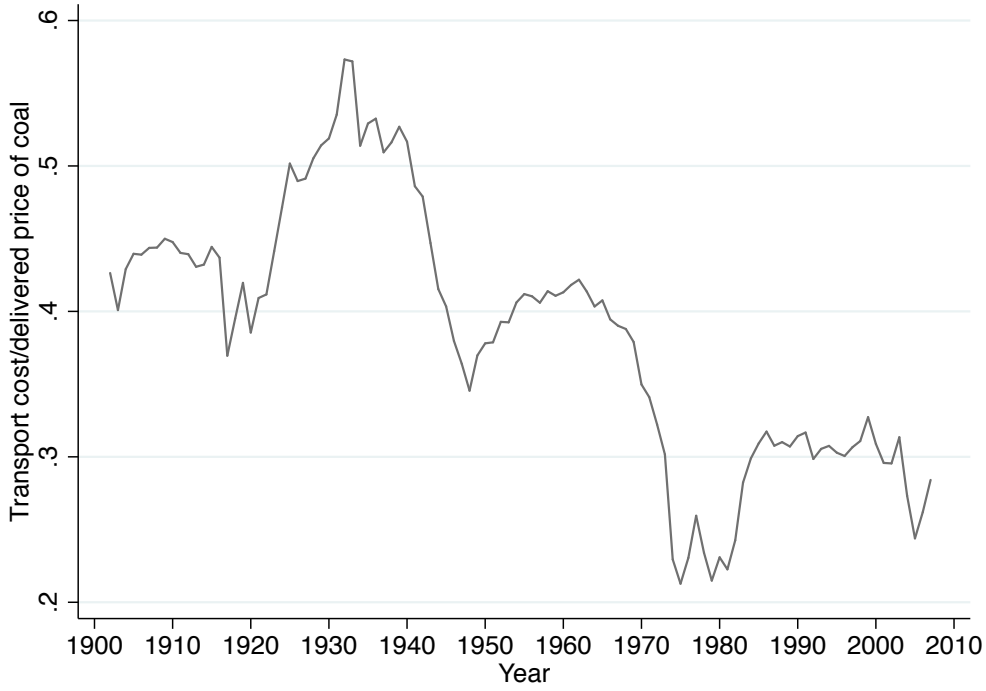
Notes: Maps show county-level average heat (in BTUs per lb, top panel), sulfur (in % of weight, middle panel), and ash (in % of weight, bottom panel) content of coal produced during 1990-1999. Source: EIA 423 forms.

Figure A.7: Sample counties and coal resources



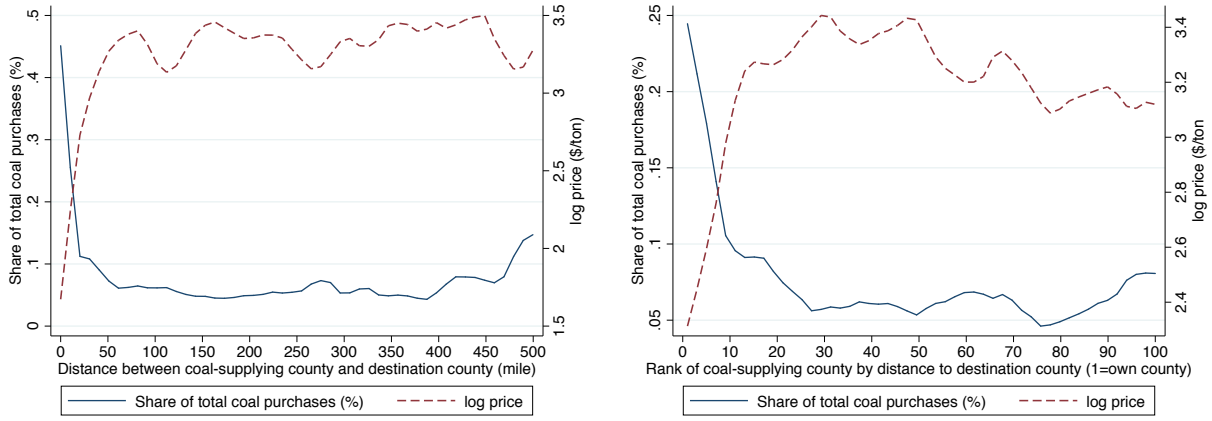
Notes: Counties are included in the sample (in yellow) if area-weighted county centroid is 1) < 250 miles from nearest Illinois coal resource, 2) closer to nearest Illinois coal resource than to nearest Appalachian coal resource.

Figure A.8: Share of delivered coal price due to transport costs at the national level



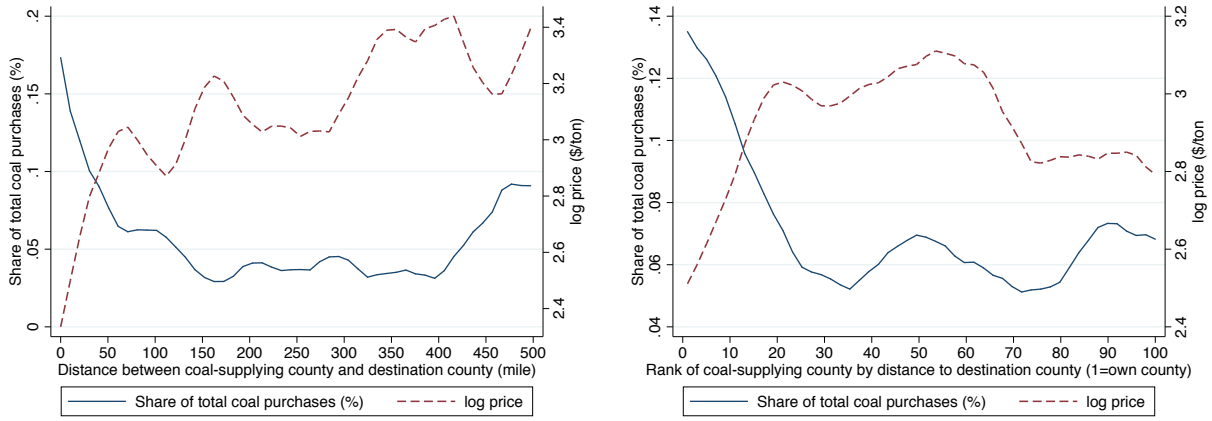
Notes: Historical national share of delivered coal price due to transport costs (1902-2007). Source: McNerney, Farmer and Trancik (2011)

Figure A.9: Testing delivered coal price definition from Hotelling's location model at the national level



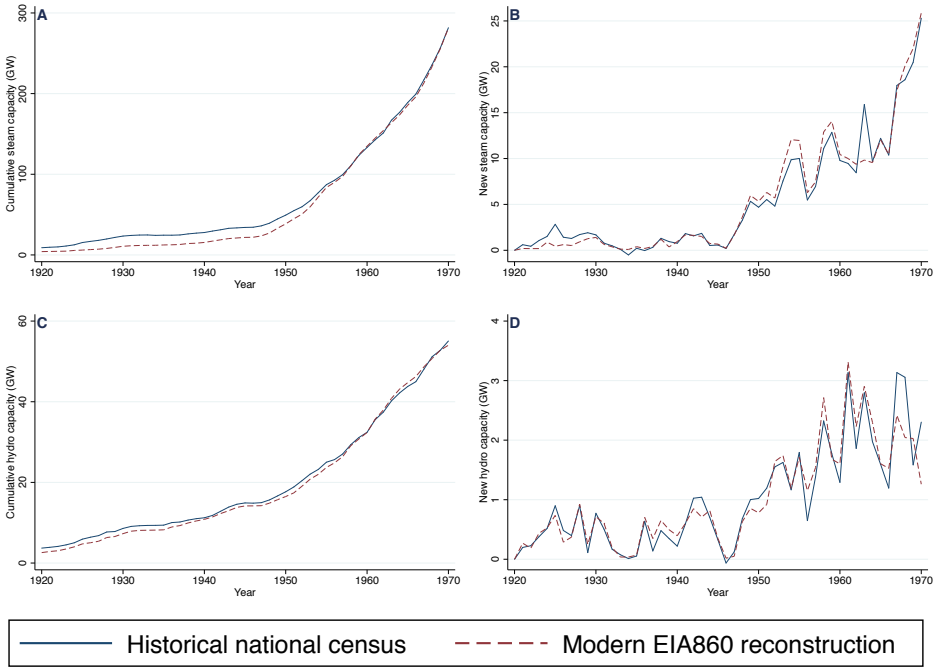
Notes: Share of total coal purchase and log delivered coal price for destination county by distance to origin county (left panel) and rank accordingly to distance (right panel). Includes all U.S. counties averaged over 1990-1999. Source: EIA 423 forms.

Figure A.10: Testing delivered coal price definition from Hotelling's location model for sample states



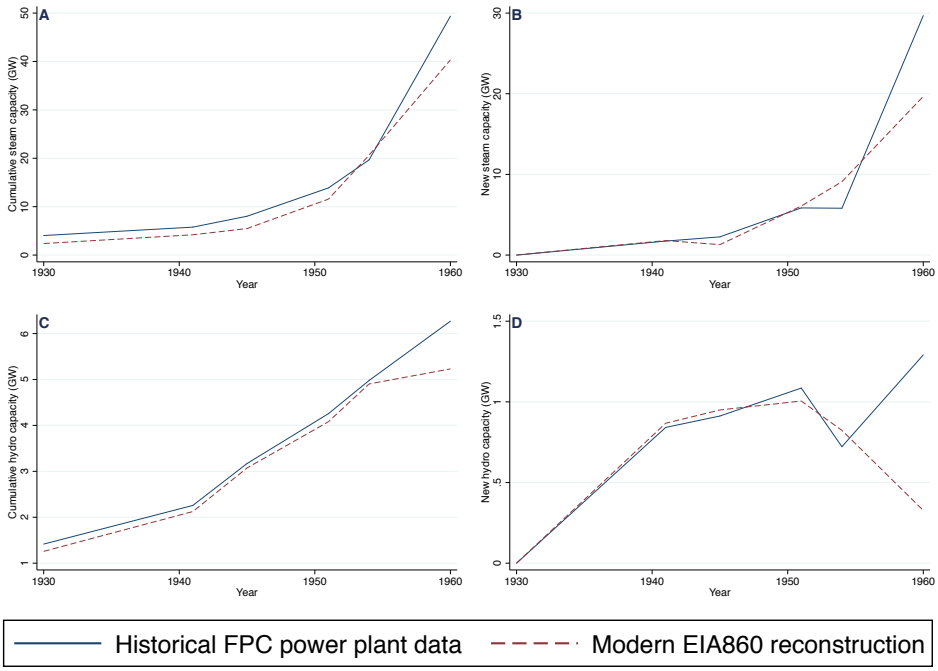
Notes: Share of total coal purchase and log delivered coal price for destination county by distance to origin county (left panel) and rank accordingly to distance (right panel). Includes counties in Alabama, Arkansas, Iowa, Illinois, Indiana, Kentucky, Minnesota, Missouri, Mississippi, Tennessee, and Wisconsin averaged over 1990-1999. Source: EIA 423 forms

Figure A.11: Comparing historical census and reconstructed capacity at the national level



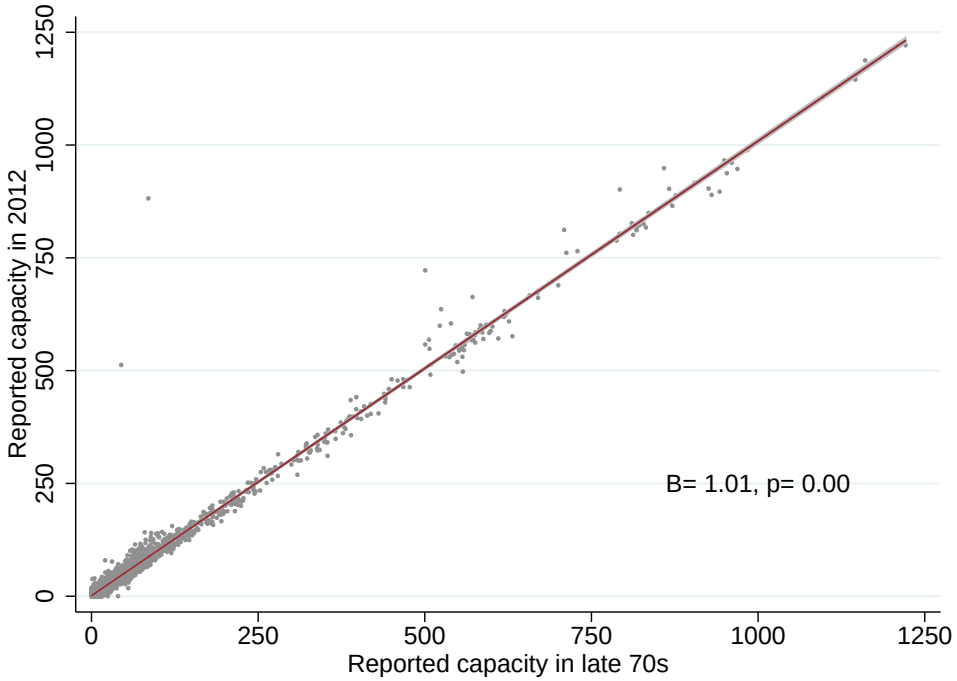
Notes: Panel (A) compares aggregate U.S. steam cumulative capacity from 1920-1970 from the U.S. historical census (U.S. Census Bureau, 1975) (solid blue) against that from modern EIA-860 reconstructed panel (dashed red). Panel (B) compares new capacity. Panels (C) and (D) the same as Panels (A) and (B) respectively for aggregate U.S. hydro capacity.

Figure A.12: Comparing historical census and reconstructed capacity for sample states



Notes: Panel (A) compares aggregate steam cumulative capacity from 1930-1960 within estimating sample states using power-plant level data from historical FPC data (see Lewis (2014)) (solid blue) against that from modern EIA-860 reconstructed panel (dashed red). Panel (B) compares new capacity. Panels (C) and (D) the same as Panels (A) and (B) respectively for aggregate hydro capacity within estimating sample states. States include Alabama, Arkansas, Iowa, Illinois, Indiana, Kentucky, Minnesota, Missouri, Mississippi, Tennessee, and Wisconsin.

Figure A.13: Comparing generator-level capacity in late 1970s and 2012 EIA forms



Notes: Scatter shows the reported generator capacity in the 2012 EIA 860 forms against the reported generator capacity in the late 1970's in the "Inventory of Power Plants in the United States"

Figure A.14: U.S. coal relative to natural gas sales prices



Notes: Log ratio of U.S. sales coal price over natural gas price. Source: Energy Information Administration (2012)

Appendix Tables

Table A.1: Checking observed county-level and constructed delivered coal price

	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome is log delivered price of coal					
log dist. to nearest mine	0.38***	0.38**	0.55***	0.58***	0.33***	0.34**
	[0.14]	[0.14]	[0.16]	[0.16]	[0.11]	[0.11]
No. of counties	377	153	392	153	353	133
Decade	70's	70's	80's	80's	90's	90's
Sample	All	Sample counties	All	Sample counties	All	Sample counties

Notes: Each column is a county-level cross-sectional regression of delivered coal price from EIA 423 forms on distance to nearest mine. All models include state fixed effects. Columns (1) and (2) use 1970s values. Columns (3) and (4) use 1980's values. Columns (5) and (6) use 1990's values. Columns (1), (3), and (5) includes all U.S. counties with coal-fired power plants. Columns (2), (4), and (6) includes only counties in Alabama, Arkansas, Iowa, Illinois, Indiana, Kentucky, Minnesota, Missouri, Mississippi, Tennessee, and Wisconsin. Robust standard errors clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1

Table A.2: Data consistency across EIA-860 forms (1990-2012)

	Percent of generators with different reported values			
Number of different values	Primary fuel	Capacity	Opening year	Retirement year
0	94.25	74.78	96.88	80.01
1	1.49	2.62	0.52	1.98
2	0.57	1.81	0.44	2.92
3	0.23	1.07	0.08	0.87
4	0.2	1.26	0.08	0.87
5	0.51	0.85	0.19	1.5
6	0.46	0.69	0.24	0.67
7	0.29	0.95	0.29	0.72
8	0.27	0.53	0.22	1.08
9	0.2	0.79	0.14	1.74
10	0.25	0.72	0.24	1.04

Notes: Row indicates the number of values from 1990-2011 EIA-860 forms that was different from the 2012 EIA-860 form. Column shows generator-level variables. Each cell shows the percentage of EIA-860 forms from 1990-2011 with a reported value that is different from that reported in 2012. For example, Row 1, Column 1 indicates that 94.25% of generators reported the same primary fuel from 1990-2011 as was reported in 2012.

Table A.3: Data consistency in reported primary fuel between late 1970s and 2012 EIA-860

	Primary fuel in 2012				
Primary fuel in 1970s	Coal	Hydro	Nat. gas	Nuclear	Oil
Coal	92.2	0.0	5.9	0.0	1.1
Hydro	0.0	100.0	0.0	0.0	0.0
Nat. gas	0.8	0.0	77.2	0.0	21.9
Nuclear	0.0	0.0	0.0	100.0	0.0
Oil	1.0	0.0	24.4	0.0	74.6

Notes: Each row shows the distribution of reported primary fuel in the 2012 EIA-860 forms conditional on the primary fuel reported in the 1970s from (Energy Information Administration, 1980). For example, 92.2% of generators which reported to use coal in the 1970s also reported to use coal in 2012.