Path Dependence in the Development of 20th Century U.S. Coal-fired Electricity Capacity

Kyle C. Meng*

PRELIMINARY AND INCOMPLETE

Abstract

Understanding the determinants and dynamics of coal-fired electricity is crucial for designing policies that reduce its social costs. This paper examines the role of path dependence - the ongoing effects of obsolete past determinants - on the development of U.S. coal-fired electricity capacity over the 20th century. I exploit a spatial reconfiguration in county-level coal prices driven by new coal extraction techniques which made earlier mines obsolete for some counties. I find that distance to an obsolete coal mine has an increasing effect on later relative coal capacity up to six decades after initial obsolescence. By the 1990s, path dependence explains 60% of total coal-fired capacity over sample counties. Analysis of electricity generating efficiency suggests path dependence may be occurring through coal-specific technological change. I expand on Acemoglu et al. (2012a)'s model of directed technical change to link my reduced-form estimates with a formal definition of technology-driven path dependence and a structural parameter which dictates policy design. The implied parameter value suggests that a future permanent transition away from coal could be induced by a temporary research subsidy towards non-coal electricity. I discuss implications for the current U.S. transition towards natural gas-fired capacity.

*Princeton University, Environmental Defense Fund, and University of California, Santa Barbara, email: kmeng@bren.ucsb.edu. I thank Jonathan Dingel, Anna Thompsett, Solomon Hsiang, Per Krusell, Gary Libecap, Suresh Naidu, Trevor O’Grady, Maxim Pinkovskiy, Bernard Saliñé, Inge van den Bijgaart, and Tom Vogl for helpful comments. Kayleigh Campbell Bierman provided excellent research assistance. Support from the EDF/Princeton High Meadows Post-Doctoral Fellowship is gratefully acknowledged. Errors are my own. This version: July, 2014.
1 Introduction

The use of coal for electricity production is associated with large social costs.\(^1\) At the global level, 50% of cumulative human-induced carbon dioxide emissions since the pre-industrial era is attributed to coal burning (Boden and Andres 2013), making coal the primary driver of climate change. Within the U.S., coal is the input fuel for roughly 40% of electricity capacity\(^2\) since the 1960s. Pollution due to the persistent use of coal for electricity production has been indirectly (Chay and Greenstone 2003, 2005) and directly (Barreca, Clay and Tarr 2014) linked with a variety of health risks. Reducing these social costs requires climate and energy policies that induce a structural transition away from coal-fired electricity. However, the effective design of such policies relies fundamentally on our understanding of the determinants and dynamics of coal-fired electricity production.

There are two common arguments for why coal has historically dominated the U.S. electricity sector, each with different implications for future energy transition policies. The locational fundamentals argument observes that the U.S. has the world’s largest endowment of coal reserves. As such, coal was favored over other inputs during each stage of electricity expansion because of the constant abundance of coal supplies relative to that of other input fuels. Implied by this view is that, absent sustained changes in relative input supply, policies that induce a permanent transition away from coal-fired electricity will require a permanent input price adjustment to offset coal’s time-invariant supply advantage.

Alternatively, the path dependence argument poses that the decision to build coal-fired capacity reflects the accumulation of historical circumstances, which, over time may dominate locational fundamentals. In this view, historical determinants of past coal-fired capacity triggered a dynamic process which affected the development of later coal-fired capacity even as certain historical determinants themselves became obsolete. One possible dynamic process driving path dependence may be the accumulation of coal-specific technologies which over time increasingly favored coal-fired capacity. If path dependence due to technological change were the dominant driver of modern U.S. coal use, policies aimed at a permanent transition away from coal-fired electricity may only require a temporary research subsidy directed towards elevating non-coal technologies until it is on par with that of coal (Acemoglu et al. 2012a, b). Once coal’s technological advantage is eliminated, growth in non-coal capacity will occur endogenously without policy intervention.

This paper tests these competing hypotheses by estimating the role of path dependence in the development of U.S. electricity capacity over the 20th century. Using subsurface spatial data characterizing coal mine depth, location, and operating dates over the U.S. midwest, I exploit a spatial reconfiguration in county-level

\(^{1}\)Local pollution from domestic coal use was one of the examples of externalities explored in Coase (1960).

\(^{2}\)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 generator and, as discussed below, provides a measure of electricity producing machine stock or infrastructure.
coal prices at the start of the 20th century driven by the new availability of deeper coal reserves following the introduction of mechanized mining. The resulting opening of deeper coal mines rearranged county-level coal prices by making existing shallow mines obsolete for certain counties. My identification strategy amounts to asking whether, conditional on contemporaneous and time-invariant determinants, counties that were located closer to an obsolete coal mine continued to have more relative coal-fired capacity than counties further away in the decades following obsolescence. Using a county-level panel of input-specific capacity for 1890-2000 reconstructed from modern records, I find that the effects of distance to an obsolete coal mine increases over time. Specifically, relative to the contemporaneous effect of distance to an active coal mine, a county that is 1% closer to an obsolete coal mine experiences 2.5% greater relative coal-fired capacity six decades after obsolescence. This result is robust to a variety of subsamples and choices made in data construction. My estimates imply that by the 1990s, 60% of total coal capacity over sample counties can be explained by the effects of path dependence with the remaining capacity explained by other determinants such as time-invariant locational fundamentals and time-varying contemporaneous factors.

My reduced-form estimates of path dependence may emerge from various dynamic processes. Additional evidence examining the thermal efficiency of existing coal-fired generators suggests that the relevant “first-stage” mechanism may be the accumulation of coal-specific technology stock. To explore the implications of this particular channel, I extend Acemoglu et al. (2012a)’s model of directed technical change with Schumpeterian innovation to provide formal definitions of weak and strong technology-driven path dependence in the development of relative coal capacity. Effects that increase over time, as is the case with my reduced-form estimates, imply the presence of strong path dependence. Importantly, my model also provides a mapping between my reduced-form estimates and a key structural parameter, the elasticity of substitution between coal and non-coal electricity, whose value dictates the duration of policy intervention required in order to induce a permanent transition away from coal-fired electricity. The elasticity of substitution implied by my reduced form estimates suggests that a *temporary* research subsidy towards non-coal electricity is sufficient to induce a *permanent* transition away from coal-fired electricity. As such, this paper provides a counterintuitive policy implication. Whereas coal-fired capacity has historically been driven by technological accumulation, the same dynamics could take hold for non-coal electricity provided that a sufficient “boost” is given to the technology associated with other inputs.

While this paper focuses on the development of coal-fired capacity over the 20th century, the presence of strong path dependence in the energy sector can inform upon the current U.S. transition away from coal and into natural gas-fired capacity. In particular, if similar dynamics of path dependence continue to hold, an endogenous expansion in natural gas-fired electricity capacity may occur with only a temporary relative input price shock provided it is sufficiently strong. However, because natural gas burning also emits carbon...
dioxide, if expected climate damages are large, the ensuing path dependence in natural gas capacity may further delay the necessary amount of decarbonization required from the electricity sector.

Identifying path dependence in relative coal capacity requires finding variation in past coal prices which satisfies two statistical properties. First, this past variation must become obsolete, by not having direct contemporaneous effects on subsequent relative coal capacity. Second, this source of variation must be uncorrelated with unobserved contemporaneous determinants of relative coal capacity prior to obsolescence. These two requirements are similar to the exclusion restriction and exogeneity assumptions needed for an instrumental variables setup. To satisfy the exclusion restriction, I follow the argument in Hotelling’s location model (Hotelling 1929) and consider only the component of county-by-decade coal prices driven by the straight-line distance to the nearest operating coal mine. For a given county, a coal mine becomes obsolete when it is no longer the nearest operating mine. By this definition, a subset of existing coal mines becomes obsolete whenever a new mine opens and the timing of this event may not always be exogenous to county-level relative coal capacity. To address this issue, I exploit only obsolescence events driven by the introduction of mechanized mining which allowed the extraction of deep coal reserves that was previously inaccessible. I also only consider the opening of deep coal mines covering large areas which are less likely to open due to demand from a single county. The resulting spatial reconfiguration of county-level coal prices is arguably uncorrelated with relative coal capacity of nearby counties conditional on controls for contemporaneous and time-invariant determinants. These controls include county and state-by-year fixed effects as well as the time varying effect of distance to the subsequent deep coal mine. Finally, as further checks on my identifying assumption, I demonstrate that distance to an obsolete coal mine is uncorrelated with pre-obsolescence changes in county-level determinants such as population and manufacturing activity that may directly affect changes in relative coal capacity after obsolescence.

This paper is linked with several bodies of literature. Conceptually, this paper ties in with a long tradition in the economics of technological change arguing that dynamic processes may differentially favor certain technologies over time (Schumpeter 1942; Schmookler 1966; Aghion and Howitt 1992) resulting in technology “lock-in” or path dependence (David 1985; Arthur 1994). While path dependence of certain technologies due to productivity accumulation may not pose economic consequences per se during initial periods, significant costs may arise if the path dependence of early technologies lead to a time path of resource allocation that is cumulatively inefficient. Coal-fired electricity is one such example in which initial drivers of coal capacity may lead to the persistent use of coal resulting in suboptimal levels of cumulative carbon dioxide emissions.

The modeling of directed technical change and environmental externalities was most recently explored by Acemoglu et al. (2012a,b). In their two-sector model of directed technical change under climate change,
Acemoglu et al. (2012a) characterize the conditions whereby an optimal policy intervention may require both research subsidies to the clean sector as well as a Pigouvian tax on carbon emissions. To the best of my knowledge, the only other paper which uses the directed technical change framework to empirically estimate path dependence in energy-related technology choice is Aghion et al. (2012) who explore patent differences in clean and dirty automobile technology across 80 countries in response to oil price changes from 1986-2005. This paper has the advantage of exploring path dependence over a much longer time horizon and for a sector which has an arguably greater role for climate change mitigation. At the aggregate level, Hassler, Krusell and Olovsson (2012) explore the potential role of directed technical change in the U.S. energy sector by examining aggregate energy-specific productivity following the 1970s oil shock.

Empirically, this paper connects with an emerging literature exploring the time varying effects of natural endowments on various economic outcomes. This literature can be generally classified by properties of the treatment. The ex-post adjustment literature typically uses panel data methods to explore the dynamics of transition in the aftermath of a major change in natural endowments such as the introduction of the potato into the Old World during the 18th century (Nunn and Qian, 2011), the American Dust Bowl of the 1930s (Hornbeck, 2012), the access to groundwater over the Ogallala Aquifer after World War II (Hornbeck and Keskin, 2014), and the Great Mississippi Flood of 1927 (Hornbeck and Naidu, 2014). The path dependence or obsolescence literature primarily uses cross-sectional methods to examine the persistent effect of historically relevant but subsequently obsolete natural endowments such as the location of U.S. portages sites (Bleakley and Lin, 2012) and that of historically active U.S. mineral mines in 1900 (Glaeser, Kerr and Kerr, 2012). For path dependence following shocks to other factors of production, Davis and Weinstein (2002) explore growth in Japanese cities following World War II bombings while Redding, Sturm and Wolf (2011) study the location of airports following the post-war German division. I follow in this literature by examining path dependence through the ongoing effects of distance to obsolete coal mines. In addition, I employ panel data methods used in many adjustment papers in order to estimate the dynamics of path dependence for decades following initial obsolescence as well as to control for time-invariant and time-varying determinants.

The remainder of the paper is organized as follows. Section 2 provides some motivating facts about the history of electricity capacity development in the U.S. in light of the locational fundamentals and path dependence explanations. Section 3 presents a model of directed technical change which provide formal definitions of weak and strong technology-driven path dependence, policy implications, and an equilibrium condition which informs my empirical specification for estimating path dependence in relative coal capacity. Section 4 covers data sources, construction, and validity checks for my construction procedures. Section 5 discusses my empirical strategy and specification. Section 6 presents the main path dependence results as well as robustness checks. Section 7 explores mechanisms of path dependence and finds evidence consistent
with path dependence being driven by technological change. The paper concludes in Section 8 with a brief discussion relating my results to the transition from coal to natural gas currently underway in the U.S.

2 Motivation: locational fundamentals vs. path dependence in 20th century U.S. coal capacity

The United States is the world’s largest cumulative emitter of man-made carbon dioxide and was, until recently, also the largest annual emitting nation [Boden and Andres 2013]. U.S. carbon dioxide emissions have historically been high even after accounting for the size of its economy. Figure A.1 plots country-level carbon dioxide emissions per capita against GDP per capita in 1960 and 2000 for all non-OPEC countries. Relative to countries of similar per capita income, U.S. carbon dioxide emissions per capita was a positive 2.4 and 1.8 \( \sigma \) outlier in 1960 and 2000 respectively.

Geography is often cited when explaining the U.S.’s persistently high carbon dioxide emissions. In particular, the U.S. has the world’s largest reserves of coal, the most carbon-intensive of primary energy inputs. Figure A.2 ranks countries by modern known coal reserves and shows that the U.S. has roughly double the coal reserves of the next most coal abundant country, China [BP 2014]. As the locational fundamental argument goes, the abundance of coal reserves, in turn, explains why the U.S. electricity sector, which accounts for 38% of U.S. greenhouse gas emissions, uses coal in over 40% of its generating capacity [Energy Information Administration 2013].

County-level evidence suggests a potentially richer explanation. Figure 1 plots the county-level log relative coal capacity\(^5\) in the 1990s against log relative coal capacity in the 1920s. For a given county, the relative coal capacity share during the 1920s predicts relative coal capacity in the 1990s which implies a role for time-invariant determinants. However, relative coal capacity in the 1920s explains only 4% of the variation in relative coal capacity during the 1990s. To explore the potential role of time varying causes, Figure 2 follows the 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 coal reserves in the Illinois Basin. Counties that are closer to coal reserves consistently have higher relative coal capacity compared to counties that are further away as expected based on differences in geography. However, the discrepancy in coal use between these counties grow dramatically over the 20th century suggesting the increasingly important role played by

---

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 coal rank 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=7&t=11](http://www.eia.gov/tools/faqs/faq.cfm?id=7&t=11)

5 For county \( i \) in decade \( t \), the coal-fired electricity capacity is \( X_{ic} \). Likewise, \( X_{int} \) denotes the capacity summed over all other energy inputs. Thus, the log relative coal capacity is \( \ln(\frac{X_{ict}}{X_{int}}) \).
time-varying factors.

The evolution of relative coal capacity for counties close to and far from coal reserves shown in Figure 2 is consistent with the presence of path dependence, defined as the ongoing effect of obsolete past determinants. However, this evidence is suggestive as other explanations could produce Figure 2. For example, it may be that distance to coal reserves never becomes obsolete but instead has an increasing contemporaneous effect on relative coal capacity over the 20th century.\(^6\) Thus, in order to estimate the effects of path dependence, it is crucial that the identifying source of variation affect past relative coal capacity but has no direct contemporaneous effect. Section 5 details an identification strategy with an exclusion restriction designed to address this estimation challenge. I now present a model which provides a formal definition of weak and strong path dependence in relative coal capacity driven by directed technical change.

3 Theory: Path dependence in a model of directed technical change

This section presents a model of directed technical change with Schumpeterian innovation which formalizes path dependence in the use of particular technologies, in this case relative coal-fired electricity capacity, as discussed in Arthur (1994). Specifically, this model demonstrates how past coal prices can have persistent effects on the subsequent development of relative coal capacity through the accumulation of coal-specific technology driven by directed technical change. While directed technical change may not be the only channel that generates path dependence, an analysis of various potential mechanisms in Section 7 provides evidence that is consistent with directed technical change as the “first-stage” in driving path dependence.

The model is largely based on Acemoglu et al. (2012) with a few extensions.\(^7\) It presents an equilibrium condition that formalizes definitions for weak and strong technology-driven path dependence. This expression can then be readily linked to an empirical specification in which reduced form estimates map onto the elasticity of substitution between coal and non-coal electricity, the key structural parameter dictating the strength of path dependence in relative coal capacity. Furthermore, this parameter also informs the form of policy intervention needed in order to initiate future transitions away from coal-fired electricity capacity. Below, I describe the modeling environment and equilibrium conditions. I then provide an equilibrium expression that formalizes definitions for weak and strong technology-driven path dependence.

\(^6\) More formally, consider the following two period equation for log relative coal capacity:

\[
\begin{align*}
\hat{X}_{i1} &= \beta_1 \text{dist}_i \\
\hat{X}_{i2} &= \beta_2 \text{dist}_i + \rho \text{dist}_i
\end{align*}
\]

When only path dependence occurs, \(\text{dist}_i\) has no direct effect on \(\hat{X}_{i2}\), so that \(\beta_2 = 0\), implying \(\hat{X}_{i2} - \hat{X}_{i1} = [\rho - \beta_1] \text{dist}_i\). When only increasing effect occurs, \(\rho = 0\) and \(\beta_2 > \beta_1\), such that \(\hat{X}_{i2} - \hat{X}_{i1} = [\beta_2 - \beta_1] \text{dist}_i\). Both models could generate the patterns shown in Figure 2.

\(^7\) There are two notable differences. First, I replace labor as an input in the production of intermediate goods with sector-specific primary energy inputs, coal and non-coal. Second, I consider only the equilibrium whereby innovation occurs in both coal and non-coal sectors at all periods.
pression for path dependence in relative coal capacity, formal definitions of weak and strong path dependence, and implications for energy transition policies.

### 3.1 Environment

There are two periods, \( t = 1, 2 \). The final good, electricity, denoted by \( Y_t \), is produced by two intermediate inputs \( Y_{ct} \) and \( Y_{nt} \) representing coal-fired electricity and non-coal electricity using all other input fuels respectively.\(^8\) Production of \( Y_t \) takes the following Constant Elasticity of Substitution form:

\[
Y_t = \left( \frac{Y_{ct}^{(\epsilon-1)/\epsilon} + Y_{nt}^{(\epsilon-1)/\epsilon}}{\epsilon} \right)^{\epsilon/(\epsilon-1)}
\]

(1)

where \( \epsilon \) is the elasticity of substitution between electricity produced by the two intermediate inputs and assumed to be gross substitutes such that \( \epsilon > 1 \).\(^9\) I normalize the price of the final good to 1.

The two intermediate inputs, indexed by \( j \in \{c, n\} \), in turn, are each produced using sector-specific energy input, \( E_{jt} \) and a continuum of sector-specific machines \( x_{jmt} \) with quality \( A_{jmt} \), indexed by \( m \in [0, 1] \). Specifically, production has the form:

\[
Y_{jt} = E_{jt}^{1-\alpha} \int_0^1 A_{jmt}^{1-\alpha} x_{jmt}^\alpha dm \quad \text{for} \quad j \in \{c, n\}
\]

(2)

where \( \alpha \in (0, 1) \). As will be discussed in Section 5.2, I treat the price of energy input \( E_{jt} \), \( w_{jt} \), as an exogenous variable. Machine \( x_{jmt} \) depreciates completely after each period. I further normalize the price of the two intermediate inputs at each date:

\[
\left[ p_{ct}^{1-\epsilon} + p_{nt}^{1-\epsilon} \right]^{1/\epsilon} = 1
\]

(3)

The objective of the model is to understand the dynamics of input specific electricity capacity, which within this framework is captured by the sum of all sector-specific machines:

\[
X_{jt} = \int_0^1 x_{jmt} dm \quad \text{for} \quad j \in \{c, n\}
\]

(4)

Machines are supplied by monopolistically competitive firms at price \( p_{jmt} \). To simplify notation, I normalize the unit cost of producing machine regardless of quality or sector to \( \alpha^2 \). Machine ownership in each sector is given each period to scientists drawn from a continuum \( s \) with normalization, \( s_{ct} + s_{nt} \leq 1 \). At the beginning of the period, each scientist decides whether to conduct research in the coal or other input sector based on the relative profitability of the two sectors. Once decided, a scientist is then randomly assigned exclusively

---

\(^8\)For simplicity, this implies an assumption that electricity produced by all other inputs are perfect substitutes.

\(^9\)This is a reasonable assumption as while electricity produced by coal may have different properties than electricity produced by other inputs (such as reliability), in general they may be regarded as substitutes.
to one machine in that sector and with probability \( \eta \in (0, 1) \) is successful in improving the technology, which for simplicity I assume is the same for both sectors. When an innovation is successful, the quality of that machine increases by \( 1 + \gamma \) with \( \gamma > 0 \). The scientist then obtains a one-period patent for that machine and becomes a monopolist for the production of that machine. When innovation is not successful, monopoly rights for the previous technology are allocated to a scientist drawn randomly from the overall pool. Aggregate productivity for each sector are defined as:

\[
A_{jt} = \int_0^1 A_{jmt} dm \quad \text{for} \quad j \in \{c, n\} \tag{5}
\]

while the innovation possibilities frontier specified above implies the following equation of motion for aggregate innovation in each sector:

\[
A_{jt} = (1 + \gamma \eta s_{jt}) A_{jt-1} \quad \text{for} \quad j \in \{c, n\} \tag{6}
\]

### 3.2 Equilibrium

My objective is to solve for relative coal-fired electricity capacity as a function of exogenous current and past relative energy input prices under a competitive equilibrium. In the competitive equilibrium, final good producers, intermediate good producers, capacity owners, and scientists maximize profits. Given Eq. 2 the optimization problem for producer of intermediate input \( j \) at time \( t \) can be written as:

\[
\max_{x_{jmt}, E_{jt}} \left\{ p_{jt} E_{jt}^{1-\alpha} - \alpha \int_0^1 A_{jmt}^{1-\alpha} x_{jmt}^\alpha dm - w_{jt} E_{jt} - \int_0^1 p_{jmt} x_{jmt} dm \right\} \tag{7}
\]

The first order condition with respect to \( x_{jmt} \) leads to the following iso-elastic demand curve:

\[
x_{jmt} = \left( \frac{\alpha p_{jt}}{p_{jmt}} \right)^{\frac{1}{1-\alpha}} A_{jmt} E_{jt} \tag{8}
\]

The profit maximizing problem for machine supplier \( m \) in sector \( j \) at time \( t \) is:

\[
\max_{p_{jmt}} (p_{jmt} - \alpha^2 x_{jmt}) \tag{9}
\]

which given Eq. 8 yields an equilibrium price \( p_{ijt} = \alpha \). Thus, the equilibrium demand for machines \( i \) in sector \( j \) is:

\[
x_{jit} = (p_{jt})^{\frac{1}{1-\alpha}} A_{jit} E_{jt} \tag{10}
\]

Integrating over all machines \( j \), using the definition for aggregate capacity in Eq. 4 and taking the ratio of
coal to other input capacity yields:

\[ \tilde{X}_t = \tilde{p}_t^{1-\alpha} \tilde{E}_t \tilde{A}_t \]  

(11)

where \( \tilde{X}_t = \frac{X_t}{X_{nt}} \) and likewise for all other variables. Next, we can simplify Eq. 11 into a function of exogenous variables using several additional equilibrium relationships. First, optimization by the final good producer results in the following expression for the relative price of the two intermediate inputs:

\[ \tilde{Y}_t = \tilde{p}_t^{-\epsilon} \]  

(12)

Taking the derivative of Eq. 7 with respect to the primary energy input and using Eq. 5, we obtain the following expression for relative output prices:

\[ \bar{p}_t^{-\alpha} = \bar{w}_t \bar{A}_t^{-1} \]  

(13)

Next the equilibrium demand in Eq. 10 inserted into the production function in Eq. 2 and rearranged in terms of relative primary energy input becomes:

\[ \tilde{E}_t = \tilde{A}_t^{-1} \tilde{Y}_t \tilde{p}_t^{1-\alpha} \]
\[ = \tilde{A}_t^{-1} \tilde{p}_t^{1-\alpha} \]
\[ = \tilde{A}_t^{-\psi} \bar{w}_t^{1-\psi} \]  

(14)

where the second line uses Eq. 12 and the definition \( \psi = (1-\alpha)(1-\epsilon) \) and the third line uses Eq. 13. Finally, Eq. 14 together with Eq. 13 is inserted into Eq. 11 and becomes:

\[ \tilde{X}_t = \bar{w}_t^{\psi} \bar{A}_t^{-\psi} \]  

(15)

This equation expresses the contemporaneous relationship between relative coal capacity and relative primary energy prices and technology levels. Assuming \( \epsilon > 1 \) implies that \( \psi < 0 \). Thus the relative coal capacity decreases as coal becomes relatively more expensive. Furthermore, relative coal capacity increases with relative coal-specific aggregate productivity. Because technology is cumulative in the Schumpeterian “standing on the shoulder of giants” sense, it is evident that path dependence of historical relative prices would manifest through this term, which we now explore.
3.3 Defining path dependence

Here, the analysis takes a departure from Acemoglu et al. (2012a). I assume that innovation occurs in both sectors in each period. This is in contrast to the baseline Acemoglu et al. (2012a) model in which initial conditions are selected such that innovation in the competitive equilibrium only occurs in the coal sector.

Scientists allocate into both sectors when the expected profit from both sectors are equal. Plugging the equilibrium demand function Eq. 10 into the monopolist’s problem yields the following equilibrium expected profit function for a scientist conducting research in sector $j$ at time $t$:

$$
\Pi_{jt} = \eta(1 + \gamma)(1 - \alpha)\alpha p_{jt}^{\frac{1}{\gamma}} E_{jt} A_{jt-1}
$$

Substituting in Eqs. 13, 14, and the equation of motion Eq. 6, the expected profit ratio becomes:

$$
\tilde{\Pi}_t = \tilde{w}_t^{\psi} \left( \frac{1 + \gamma \eta s_{ct}}{1 + \gamma \eta s_{nt}} \right)^{-\psi-1} \tilde{A}_{t-1}^{-\psi}
$$

where the second equality follows from assuming innovation occurs in both sectors in both periods. I now present two assumptions:

**ASSUMPTION 1** The elasticity of substitution for coal-fired electricity and that electricity produced by other inputs is limited to the “weak equilibrium relative bias” (Acemoglu, 2003) such that $-1 < \psi < 0$, or equivalently $1 < \epsilon < \frac{2-\alpha}{1-\alpha}$.

**ASSUMPTION 2** Innovation occurs in both sectors during both periods such that $\tilde{\Pi}_t = 1$ for $t = 1, 2$.

Applying Assumption 1\textsuperscript{10} into Eq. 17 rearranging terms, and using Eq. 6 again to replace $\frac{1 + \gamma \eta s_{ct}}{1 + \gamma \eta s_{nt}}$ yields:

$$
\tilde{A}_t = \tilde{w}_t^{-\psi} \tilde{A}_{t-1}^{\psi-1}
$$

Under Assumption 1, Eq. 18 shows that relative coal capacity decreases with contemporaneous relative coal prices and increases with the past technology ratio. Inserting Eq. 18 into Eq. 15 and expressing in terms of specific time periods:

$$
\tilde{X}_2 = \tilde{w}_2^{-\psi} \tilde{A}_{2}^{\psi-1}
$$

$$
\tilde{X}_2 = \tilde{w}_2^{-\psi} \tilde{w}_1^{\psi+1} \tilde{A}_{1}^{-\psi}
$$

$$
\log(\tilde{X}_2) = \frac{\psi}{\psi + 1} \log(\tilde{w}_2) + \frac{-\psi^2}{(\psi + 1)^2} \log(\tilde{w}_1) + \frac{-\psi}{(\psi + 1)^2} \log(\tilde{A}_0)
$$

\textsuperscript{10}Following Eq. 17, the parameter space defined in Assumption 1 implies that the equilibrium relative profit, $\tilde{\Pi}_t$, is declining in $s_{ct}$. Assumption 2 amounts to assuming parameter values of $\eta$, $\gamma$, and $A_0$ which allow for a unique interior solution $s^*_ct \in (0, 1)$ such that $\tilde{\Pi}_t = 1$ for $t = 1, 2$. 

10
where the second line results from applying Eq. [18] from period $t = 1$, and the final line follows from a log transformation. $\tilde{A}_0$ is the initial technology ratio. Eq. [19] is the ultimate expression of interest as it relates current relative coal capacity to current and past relative input prices. In particular, under Assumption I, Eq. [19] captures path dependence through sector-specific technology accumulation by showing that relative coal capacity decreases with both current and past relative coal prices. I now define two forms of path dependence which compares the effects of current and past relative coal prices on current relative coal capacity:

**DEFINITION 1** *Weak technology-driven path dependence:* Suppose Assumptions 2 and 1 hold. The effect of past relative coal prices weaken over time, 
\[ \frac{\partial \log(\tilde{X}_2)}{\partial \log(\tilde{w}_1)} - \frac{\partial \log(\tilde{X}_2)}{\partial \log(\tilde{w}_2)} = \frac{-\psi^2}{(\psi+1)^2} - \frac{\psi}{\psi+1} > 0, \text{ when } -0.5 < \psi < 0, \text{ or equivalently when } 1 < \epsilon < \frac{1.5-\alpha}{1-\alpha}. \]

**DEFINITION 2** *Strong technology-driven path dependence:* Suppose Assumptions 2 and 1 hold. The effect of past relative coal prices strengthen over time, 
\[ \frac{\partial \log(\tilde{X}_2)}{\partial \log(\tilde{w}_1)} - \frac{\partial \log(\tilde{X}_2)}{\partial \log(\tilde{w}_2)} = \frac{-\psi^2}{(\psi+1)^2} - \frac{\psi}{\psi+1} < 0, \text{ when } -1 < \psi < -0.5, \text{ or equivalently when } \frac{1.5-\alpha}{1-\alpha} < \epsilon < \frac{2-\alpha}{1-\alpha}. \]

which follows directly from comparing \( \frac{-\psi^2}{(\psi+1)^2} \) and \( \frac{\psi}{\psi+1} \). Importantly, Definition 2 formalizes the idea noted by Arthur (1994) that past relative determinants could exert increasing influence over time. Within the parameter space specified by Assumption I, there exists a range for $\epsilon$ in which the negative effect of past relative coal prices on current relative coal capacity may be larger in magnitude than the effect of current relative coal prices. Definitions 1 and 2 is tightly connected with the competing forces of the market size and price effects that is typical of models of directed technical change (Acemoglu, 2002; Acemoglu et al., 2012a). I now turn to policy implications of weak and strong technology-driven path dependence.

### 3.4 Policy implications for transitioning away from coal

Suppose policy makers were interested in preventing the future growth of coal-fired electricity, $Y_{ct}$. Would a temporary research subsidy which lowers the technology ratio, $\tilde{A}_t$, such that it would be profitable for scientists thereafter to direct research only in the non-coal sector ensure $Y_{ct} / Y_{ct} \leq 0$ in the long run? In other words, in the long run, would $Y_{ct}$ continue to grow if $A_{nt} = (1 + \gamma \eta)A_{nt-1}$ while $A_{ct}$ remains constant? To examine this, use Eqs. 8 and 2 to rewrite the equilibrium coal-fired electricity:

\[ Y_{ct} = A_{ct} E_{ct} (p_{ct})^{\frac{1}{1-\alpha}} \tag{20} \]

\[ 11 \]This may be due to concerns about climate damages associated with increased use of $Y_{ct}$. See Acemoglu et al. (2012a) for a full analysis of optimal climate policy under climate damages.
Next, rewriting Eq. 3 and substituting in Eq. 13, I can write the price of coal-fired electricity as:

\[ p_{ct} = \left( \frac{(A_{nt}w_{ct})^\psi}{(A_{ct}w_{nt})^\psi + (A_{nt}w_{ct})^\psi} \right)^{\frac{1}{1-\epsilon}} \] (21)

Plugging in Eqs. 14 and 21 into Eq. 20 yields:

\[ Y_{ct} = \frac{w_{nt}}{w_{ct}} A_{ct}(A_{nt}w_{ct})^{\alpha+\psi} \left[ (A_{ct}w_{nt})^\psi + (A_{nt}w_{ct})^\psi \right]^{-\frac{\psi}{\alpha}} \] (22)

If input prices \( w_{jt} \) are stationary in the long run, Eq. 22 implies that \( Y_{ct} \) grows asymptotically at the same rate as \( A_{nt}^{\alpha+\psi} \) or \( (\alpha+\psi)\log(1+\gamma\eta) \).\(^{12}\) Thus when \( \alpha + \psi \leq 0 \), or equivalently when \( \epsilon > 1/(1 - \alpha) \), a temporary subsidy making research in the non-coal sector permanently profitable would cause \( Y_{ct} \) to decline in the long run. In the parlance of directed technical change models, this relative optimistic policy scenario corresponds to the case when the market size effect dominates due to the increased use of non-coal energy inputs. If, however, \( \alpha + \psi > 0 \) or equivalently when the elastically of substitution is sufficiently low, \( 1 < \epsilon < 1/(1 - \alpha) \), the price effect dominates such that an increase in the relative price of coal-fired electricity encourages the further production of \( Y_{ct} \). In such a case, a temporary research subsidy alone will not be sufficient to stop growth in \( Y_{ct} \) and instead some form of permanent intervention is additionally required to lower \( Y_{ct} \).

Interestingly, these thresholds for \( \epsilon \) correspond approximately with those under Definitions 1 and 2 for weak and strong path dependence. In particular, for reasonable values of \( \alpha \),\(^{13}\) the parameter space for \( \epsilon \) defining strong path dependence mostly overlaps the parameter space characterizing the more optimistic policy scenario in which a temporary research subsidy is sufficient to arrest future growth in \( Y_{ct} \). Thus, the dynamics presented in this model suggest a counterintuitive interpretation for the presence of strong path dependence. If coal-fired capacity has historically grown due to coal-specific technological accumulation, the same dynamics may take hold for electricity produced by other fuels. As such, a sufficient temporary “boost” in the technology stock of non-coal capacity may be sufficient in order to induce a permanent decline in coal-fired electricity.

\(^{12}\)Taking the log of Eq 22 and the time derivative yields:

\[ \frac{\dot{Y}_{ct}}{Y_{ct}} = (\alpha + \psi) \frac{\dot{A}_{nt}}{A_{nt}} - \alpha \left[ \frac{(A_{nt}w_{ct})^\psi}{(A_{ct}w_{nt})^\psi + (A_{nt}w_{ct})^\psi} \right] \frac{\dot{A}_{nt}}{A_{nt}} \] (23)

Asymptotically, as \( t \to \infty \), \( A_{nt}^\psi \to 0 \) so that \( \frac{\dot{Y}_{ct}}{Y_{ct}} = (\alpha + \psi)\log(1 + \gamma\eta) \).

\(^{13}\)An \( \alpha = 1/3 \) corresponds roughly to the national capital share in the U.S. (Acemoglu et al., 2012a).
4 Data

4.1 U.S.G.S. assessment of coal reserves and mines in the Illinois Basin

This paper’s identification strategy relies on the spatial rearrangement of operating coal mines following the introduction of mechanized coal extraction which made deeper coal reserves accessible. The ideal dataset, which would include the opening and closing dates of coal mines and the method of extraction used, does not exist. Instead, the most relevant data available is 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 reserves at various depths from the surface for the Appalachian, Illinois, Gulf Coast, Northern Rocky Mountains, and Colorado Plateau coal basins. This three dimensional spatial data provide a rich characterization of the stratigraphy of the coal-bearing subsurface geology and can be used to approximate access to coal at various points in time as coal mining gradually deepened across the U.S. Of the coal reserves in the N.C.R.A, only the Illinois Basin assessment includes additional GIS data on the coal mines with opening and closing dates since 1880 thus limiting my sample to counties in the midwest.14

4.2 Reconstructing county-level input-specific electricity capacity

Electricity is typically produced by extracting heat from an energy input such as coal, oil, or natural gas. This is done within a power plant by a series of machines known together as a generating unit.15 Generating units have two important characteristics relevant for this paper: the type of energy input used and its capacity which describes the maximum electricity that can be produced by the sum of the machines in that generating unit.

My main outcome of interest is the county-level relative coal capacity, defined as the ratio of coal-fired electricity capacity over capacity using all other energy inputs. An ideal dataset would be a county-by-energy input panel of capacity since initial electricity expansion around 1890. However, such data was either never collected or likely no longer exist at or below the county-level today. A summary of the history of U.S. electricity data and a discussion of what is likely still available today is presented in Section Appendix A. Given this data limitation, I reconstruct a historical county-by-energy input capacity dataset from 1890-2000 using generator-level data collected over recent years. The key is that this data include the start and retirement dates of many generating units that are currently retired.

---

14Historical coal mine data for other coal basins are held by state-level agencies. However, they are based on the “final map” for a coal mine which only notes the closing year. My definition of obsolescence requires observing the year in which a mine opens.

15According to the U.S. Department of Energy’s Energy Information Agency, a generating unit is a “combination of physically connected generators, reactors, boilers, combustion turbines, and other prime movers operated together to produce electric power.” A typical power plant could include several generating units which may or may not use the same primary energy input.
I reconstruct the history of electricity capacity across the major energy inputs for each county using public utility electricity generator data collected annually by the U.S. Department of Energy’s Energy Information Agency (EIA) in its EIA-860 form available digitally since 1990. EIA-860 includes a wealth of generating-unit characteristics among which, relevant for this paper, include capacity, opening year, primary type of input fuel, prime-mover,\(^{16}\) thermal efficiency,\(^{17}\) and the county it is situated in. Most critically, EIA-860 reporting is required not just for currently operating generating units but also for all retired generators located on any currently operating power plant and notes the year of retirement. This allows me to use cross-sectional data from the EIA-860 forms to reconstruct a county-by-input-by-decade panel dataset of operating capacity from 1890-2000.\(^{18}\) For a given generator, I take the most recently reported values within all 1990-2012 EIA-860 forms to ensure that I capture generating units located on any power plants that may have retired since 1990. As I show in Section 6.3, my estimates of path dependence are not sensitive to the particular EIA-860 reporting year used in reconstructing my historical panel.

### 4.3 Checking validity of reconstructed capacity data

I make three assumptions in order for my EIA-860 reconstructed county-by-input-by-decade panel from 1890-200 to match actual historical data. First, I assume that nearly all power plants that have operated since 1890 still host operating generating units today. If it is true that power plants rarely retire, this implies that censoring of retired generators is limited in the EIA-860 data. Second, for a generator reported in EIA-860, I assume that the reported capacity today is the same as the historical capacity. Third, I assume that fuel switching rarely occurs for a given generating unit such that the input fuel reported in the EIA-860 form reflects the historical input fuel.

I conduct several checks for these assumptions. As a test of my first assumption, while actual historical fuel-specific capacity at or below the county-level is not available, the historical national capacity is available at the prime-mover level from the U.S. Census Bureau (1975). Panel (A) of Figure A.5 compares the national hydropower capacity reported by the historical census from 1920-1970 against national hydropower capacity based on my reconstructed panel using EIA-860 forms. As expected given concerns over data censoring, my reconstructed national aggregate is slightly below the census data from 1920-1960 but the two time series converging thereafter. Generators can operate for several decades. As such, the censoring issue displayed in Panel (A) may be due to missing generators on retired power plants that opened prior to 1920 or due

\(^{16}\)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.

\(^{17}\)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).

\(^{18}\)My reconstructed panel dataset includes only on generating units owned by utility companies because units owned by non-utilities are not consistently reported over the history of the EIA-860.

14
to those opening after 1920. Panel (B) plots the annual percentage change in hydropower capacity for the census data and the EIA-860 reconstructed data. Percentage changes from the reconstructed data closely track that from the census data suggesting that the EIA-860 forms are including most generating units that opened since 1920. Panels (C) and (D) of Figure A.5 provides similar comparisons between census data on steam-power capacity and data from my EIA-860 reconstruction. As with hydropower capacity, there is under some data censoring from 1920 to 1955 which appears to disappear afterwards. Annual percentage changes in reconstructed steam-powered capacity again closely track the census data suggesting that the EIA-860 forms include most steam-powered generating units which opened after 1920.

Steam-powered capacity can use a variety of inputs, including but limited to coal. While national census data does not report fuel-specific electricity capacity, it does provide fuel-specific electricity generation which is the product of capacity and a capacity factor (see Section Appendix B for details on capacity factor and production data imputation). Panels (A) and (B) of Figure A.6 compare coal-fired electricity production between the census data and my reconstructed data in levels and percentage changes. Again, coal-fired electricity production in my reconstructed data is underreported prior to 1960, though the annual per change change is roughly similar to that shown from the true historical data. In Panel (C) of Figure A.6 shows that my reconstructed data over reports the steam-powered electricity produced by other inputs prior to 1960 while Panel (D) again shows that the annual percentage change closely matches that of the census data. Because Figure A.6 suggests that my reconstructed historical data underreports coal-fired electricity and over reports electricity from other inputs, my EIA-860 reconstructed data at the county-level is likely to under report relative coal capacity.

Turning to my other two assumptions, I now examine whether generator capacity and fuel input has changed over time by exploring the consistency in reported values both across 1990-2012 EIA-860 records and across the 2012 EIA-860 record and the earliest available generator-level dataset. Table A.1 examines the consistency of relevant generator characteristics across the 1990-2012 EIA-860 records. For each variable across the columns of Table A.1 the rows show the percentage of EIA-860 reports from 1990-2011 that differed from the value recorded in 2012. For example, 94% of generators reported using the same primary fuel throughout 1990-2011 as was reported in 2012. Likewise, for reported generator capacity, 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 which includes the earliest available comprehensive generator-level dataset collected in the late 1970s (see Section Appendix A for more details). Figure A.7 plots the generator-level capacity reported in 2012 against the capacity reported in the late 1970s. The relationship between the

---

19Steam engines or turbines can be fed by coal, oil, natural gas, nuclear fuel, and various forms of biomass.
two reported values is nearly one-to-one. Table A.2 shows the distribution of reported primary fuel in 2012 conditional on primary fuel reported in the late 1970s. 92% of generators 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, available checks with national census data and earlier generator-level data suggests that a panel of county-level relative coal capacity reconstructed using modern EIA-860 generator-level data should reasonably approximate the true historical data.

4.4 Historical and modern county characteristics

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). These variables include total population, rural population as well as manufacturing variables such as output, number of establishments, employment, and capital value. 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.) using a method which modifies the procedure developed by Hornbeck (2012). Modern county-level data on highways and railroad networks was obtained from the U.S. Department of Transportation’s National Transportation Atlas Database.

5 Empirical setup

This section details how I estimate reduced-form effects of path dependence. Such estimates will allow me to recover the structural parameter $\psi$ which captures both the strength of path dependence in the development of relative coal capacity and informs whether a temporary research intervention is sufficient for limiting growth in future coal-fired electricity. This can be done by estimating the equilibrium condition Eq. [19] in Section 3.3 which, for county $i$, can be expanded into:

$$\log(\tilde{X}_{i2}) = \frac{\psi}{\psi + 1} \log(w_{i2}) + \frac{-\psi^2}{(\psi + 1)^2} \log(w_{ic1}) - \frac{\psi}{(\psi + 1)^2} \log(w_{in2}) - \frac{-\psi^2}{(\psi + 1)^2} \log(w_{in1}) + \frac{-\psi}{(\psi + 1)^2} \log(\tilde{A}_0)$$ (24)

Identifying path dependence in Eq. (24) requires finding variation in past coal prices, $\log(w_{ic1})$, which

---

20 Available: http://doi.org/10.3886/ICPSR02896.v3
21 Available: https://www.nhgis.org/documentation/gis-data
satisfies two statistical properties. First, this past variation must become obsolete, that is it can not have direct contemporaneous effects on subsequent relative coal capacity. Second, this source of variation must be uncorrelated with unobserved contemporaneous determinants of relative coal capacity. In particular, the technology ratio during initial electricity expansion, $\tilde{A}_0$, is not observed and may be correlated with $\log(w_{i1c1})$. I also do not observe the composite price of all other primary energy inputs, $\log(w_{in1})$ and $\log(w_{in2})$, which may also be correlated with $\log(w_{i1c1})$. These two requirements are similar to the exclusion restriction and exogeneity assumptions needed for an instrumental variables setup.

To satisfy the exclusion restriction, I use Hotelling’s location model (Hotelling, 1929) to define coal prices for a given county at any period as the straight-line distance to the nearest operating coal mine. For a given county, a coal mine becomes obsolete when it is no longer the nearest mine. The obsolete coal price faced by a county is therefore that county’s distance to the previously closest coal mine. By this definition, a subset of existing coal mines becomes obsolete whenever a new mine opens. For a given county, I call this the obsolescence event. In order to satisfy exogeneity, I need to ensure that both distance to an obsolete coal mine and the timing of obsolescence is uncorrelated with unobserved determinants of contemporaneous relative coal capacity. By using straight-line distance between a county’s centroid and the nearest coal mine, I am likely exploiting the component of transport cost that is driven by subsurface geographical features and not surface transport infrastructure which may be correlated with a county’s relative coal capacity. To ensure that the timing of obsolescence is exogenous to relative coal capacity, I exploit only obsolescence events driven by the introduction of mechanized mining which opened production of deep coal reserves which were not accessible previously. This spatial reconfiguration of operating mines due to an advance in mining technology is arguably exogenous to the relative coal capacity of nearby counties. I also only consider the opening of deep coal mines covering large areas which are less likely to open due to coal demand from any single county. Because the timing of obsolescence events varies across counties, I estimate Eq. 24 using a panel data structure which includes county-level fixed effects and state-by-year year effects. Finally, as further checks on exogeneity, I demonstrate that distance to an obsolete coal mine is uncorrelated with pre-obsolescence changes in county-level determinants such as population and manufacturing activity that may directly affect changes in relative coal capacity after obsolescence. I now discuss the empirical context in which I define coal mine obsolescence.

---

\[ \frac{\partial \log(X_{i2})}{\partial \log(w_{i1})} = \frac{-\psi^2}{(\psi+1)^3} \] and not \[ \frac{\partial \log(X_{i2})}{\partial \log(w_{i1})} = \frac{-\psi^2}{(\psi+1)^2} + \frac{\psi}{(\psi+1)} \]

---

23In the context of Equation 24, it must be that
24This is equivalent to a first difference of Eq. 24
5.1 Spatial reconfiguration of coal prices following mechanized mining

The location of coal mines depends on several factors including the location of demand, transport costs, subsurface geology of coal reserves, and extraction technology. Prior to the 20th century, most coal in the U.S. was manually extracted which limited the location of coal mines primarily over near-surface coal reserves 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. Indeed, Figure A.3 shows that nearly all the increase 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 reserves which was previously inaccessible (Speight [1994]). The opening of deeper coal mines following the introduction of mechanization altered the spatial arrangement of coal prices. Some counties situated near deep coal reserves became closer to coal mines while others became further away as older shallow mines closed.

To locate mines that were manually and mechanically extract, I use the N.C.R.A. data for the Illinois basin to create a spatial union for all shallow coal reserves in the Illinois basin that is situated less than 200 feet from the surface and all deep reserves that are situated greater than 200 feet from the surface. Because the Illinois coal basin is composed largely of one sloping coal layer, there are very few locations where both shallow and deep coal reserves coexist at different depths. I then characterize coal mines over the Illinois basin as either accessible by manual mining or accessible by mechanized mining by whether the mine is situated over a shallow or deep coal reserve.

Figure 3 maps the location of coal reserves by depth and coal mines in the Illinois Basin, the only coal reserve in the U.S. for which GIS data on coal reserves and historical mine opening and closing dates are available (see Section 4.1). The lighter shaded areas show bituminous coal reserves that are less than 200 feet below the surface and approximate reserves that are recoverable using manual extraction. After the introduction of mechanized mining, the darker shaded areas in Figure 3, which indicate reserves that are situated at depths greater than 200 feet deep, became accessible opening up the interior of the Illinois coal basin to mining.

Figure 3 also locates large coal mines in operation 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

25For a few small areas where there is overlap, I consider that area to only have a shallow coal reserve. This is unlike the Appalachian coal basin which has many overlapping layers. For such subsurface formations, one may be concerned that mining of a shallow coal mine could lead to mining of deeper coal mines.

26This corresponds to a spatial area more than 0.01 square mile. This definition of “large” mines is not essential to the main result.
to demand from any individual county and thus their obsolescence is unlikely to be correlated with county-level determinants of relative coal capacity. To show the timing of large mine openings due to mechanized extraction, Figure 4 plots the opening of large mines by decade from 1890 to 1990 by the depth of the underlying coal reserve. While mines over shallow coal reserves continued to be opened throughout the 20th century, coal mines over deeper reserves was an increasing share of new coal mines after the 1900s consistent with the increased use of mechanized extraction shown in Figure A.3.

5.2 Defining obsolescence

For a given county, I consider a coal mine to be obsolete when it is no longer the nearest operating mine. To limit the number of obsolescent events to those that are plausibly exogenous to county-level relative coal capacity, I consider only mine openings that have likely occur due to the introduction of mechanized extraction as approximated by the depth of the coal reserve located under a new mine. I now formalize my definition of obsolescence based on the logic of Hotelling’s location model (Hotelling, 1929).

There are two time periods, \( t = 1, 2 \). Consider a continuum of potential mines that lie over coal reserves characterized by a location index \( H \in \mathbb{R}^1 \) and coal reserve depth \( z_h \in \{z^S, z^D\} \) with \( z^S < z^D \). \( z^S \) is the depth over shallow coal that can be extracted manually and \( z^D \) is the depth over deep coal that can only be recovered with mechanized extraction. For potential coal mine \( h \), the fixed cost of extraction is \( F_{ht} \). In period \( t = 1 \), when only manual mining is available \( F_{h1} = 0 \) if \( z_h = z^S \) and \( F_{h1} = F \) if \( z_h = z^D \) with \( F > 0 \).

At period \( t = 2 \) when mechanized mining is available, \( F_{h2} = 0 \) for all potential mines \( h \).

In addition, there is a continuum of counties buying coal indexed by location \( I \in \mathbb{R}^1 \). Mine \( z \) charges county \( i \) price for coal that is, for simplicity, captured by their distance, \( \text{dist}_{ih} = |i - h| \). A mine will enter into production when revenue exceeds fixed cost and thus whenever \( F_{ht} = 0 \). Finally, define the set of mines that are operating at each time period as \( S_t = \{h \in H : F_{ht} = 0\} \). As in the standard Hotelling setup, mines engage in Bertrand price competition such that the coal price faced by county \( i \) is determined entirely by that of the closest mine.\(^{27} \)

Thus the coal price faced by county \( i \) in period \( t \) is:

\[
w_{ict} = \text{dist}_{it}^* = \min_{h \in S_t} \{\text{dist}_{ih}\} \tag{25}
\]

Because \( w_{ict} \) is driven by aggregate chances in mining technology, it is possibly uncorrelated with unobserved contemporaneous determinants of county-level relative coal capacity which would satisfy my exogeneity assumption. For a given county \( i \), a mine becomes obsolete from period 1 to 2 when it is no longer the

\(^{27}\)One can think of mining firms as engaging in backwards induction of a two-stage problem. In the first stage, the fixed cost of extraction, \( F_{ht} \), is observed for all mines \( h \). In the second stage, coal prices for each county are set based on distance to the nearest operating coal mine. Thus, mining firms decide whether to enter into production if revenue from the second stage exceeds the fixed cost from the first stage.
closest operating mine due to mine opens driving by the introduction of mechanized extraction. Thus, distance to obsolete coal mines due to mechanized mining provides a source of variation for recovering the reduced-form effect of obsolete coal prices on later relative coal capacity.

5.3 Empirical specification

Turning now to my empirical specification, the model presented in Section 3 assumes full machine depreciation between each period. However, in reality electric generators are built to last several decades. Thus, to capture path dependence I must estimate over a sufficiently long time horizon such that most of the electricity capacity during the initial period has since retired.

I construct a panel dataset in which each county-decade observation is assigned the distance to the nearest operating mine which captures the contemporaneous coal price following Eq. 25. My sample includes only counties that are within 250 miles of the Illinois coal basin and furthermore are closer to the Illinois Basin than the Appalachian Basin (see Figure A.4). For a given county, an obsolescence events occurs when the nearest mine switches from a shallow mine to a deep large mine. The obsolete coal price is then the distance to the now obsolete shallow coal mine, $dist_i^O$, with $dist_i^A$ denoting the distance to the nearest deep coal mine following the obsolescence event. Because opening of deep coal mines occurred throughout the first half of the 20th century (see Figure 4), the timing of the event varies across counties as shown in Figure 5. Most counties experiences the obsolescence event in the 1900s. I code additional years prior to the obsolescence event as $\tau = 0$ if the obsolete mine was the nearest mine for several decades, though doing so is inessential to my results. I also add a value of 1 to all input-specific capacity values, $X_{ijst}$, to ensure that relative coal capacity, $\bar{X}_{ist} = \frac{X_{ijst}}{X_{inst}}$, is never undefined when $X_{inst} = 0$. Because generator capacities are typically in large discrete values, this data transformation has little effect on relative coal capacity whenever $X_{inst} \neq 0$.

This data environment allows me to estimate the time-varying effect of distance to an obsolete coal mine on a county’s relative coal capacity decades after the obsolescence event. Specifically, I estimate the following linear panel model for relative log coal capacity $\log(\bar{X}_{ist})$ for county $i$ in state $s$ during decade $t$:

$$\log(\bar{X}_{ist}) = \sum_{\tau=1}^{10} \beta^O_{\tau} \log(dist_i^O) \ast 1[t = \tau] + \sum_{\tau=1}^{10} \beta^A_{\tau} \log(dist_i^A) \ast 1[t = \tau] + \sum_{\tau=1}^{10} \gamma_{\tau} \ast 1[t = \tau] + \mu_i + \nu_{st} + \epsilon_{ist}$$

where $\tau$ captures event time and $\beta^O$, my coefficient of interest, captures the time-varying elasticity of distance.

---

28 According to the EIA-860 forms in 2012, the average retired U.S. electric generator recorded operated for 43 years. The average operating U.S. electric generator in 2012 has been operating for 38 years.

29 I consider only counties that are closer to coal reserves in the Illinois Basin than to reserves in the Appalachian Basin to avoid the influence of coal mining over the Appalachian Basin for which there is not adequate mining data.

30 Of the counties in the main sample, 40% faced obsolescence during the 1900s while another 37% faced obsolescence in the 1960s.
to the obsolete mine relative to elasticity prior to the obsolescence event. In particular, \( \beta^O \) corresponds to \( \frac{-\psi^2}{(\psi+1)^2} - \frac{\psi}{\psi+1} \) from Eq. 19. An estimate of \( \hat{\beta}^O > 0 \) implies weak path dependence while \( \hat{\beta}^O < 0 \) implies strong path dependence.

Equation 26 includes a number of controls to address exogeneity concerns. Figure A.8 plots the relationship between distance to obsolete coal mine, \( dist^O_i \), and distance to the deep coal mine after obsolescence, \( dist^A_i \), showing that the two variables are strongly correlated. To ensure that \( \beta^O \) isn’t capturing the effects of contemporaneous coal prices, Eq. 26 includes the time-varying effects of \( dist^A_i \) captured by \( \beta^A \). Furthermore, \( \gamma_T \) captures the event time effect unrelated to distance to mine, \( \mu_i \) controls for unobserved time-invariant county characteristics, and \( \nu_{st} \) controls for unobserved characteristics that vary at the state-by-year level. Standard errors are clustered at the county-level to allow for heteroscedasticity and serial correlation of arbitrary form for a given county.

6 Results

This section presents my main results for path dependence in relative coal capacity. First, I demonstrate that distance to an obsolete coal mine is not correlated with changes in county characteristics prior to the obsolescence event that may influence subsequent relative coal-fired electricity capacity. I then turn to my main results finding evidence of strong path dependence. This is followed by a series of robustness checks examining functional form and whether results are sensitive to estimation sample and data construction. Finally, using my estimates, I examine the relative role of path dependence in explaining total coal capacity over my sample counties during the 20th century.

6.1 Changes in county characteristics before obsolescence

My exogeneity assumption requires that changes in relative coal capacity after the obsolescence event is driven only by distance to the obsolete coal mine. To examine whether other plausible drivers of relative coal capacity prior to the obsolescence event may be correlated with distance to the obsolete coal mine, I run the following two regressions when \( \tau = 1 \):

\[
\log(Z_{ist}) = \alpha_1 \log(dist^O_i) + \alpha_2 \log(dist^A_i) + \nu_{st} + \epsilon_{ist} \quad \text{if } \tau = 1 \tag{27}
\]

\[
\log(Z_{ist}) - \log(Z_{ist-1}) = \alpha_1 \log(dist^O_i) + \alpha_2 \log(dist^A_i) + \nu_{st} + \epsilon_{ist} \quad \text{if } \tau = 1 \tag{28}
\]

where \( Z_{ist} \) is a characteristic of county \( i \) in state \( s \) at the start of decade \( t \) from the historical censuses. For example, if the obsolescence event for a particular county occurs during the 1900s, Eq. 27 examines whether the level in characteristic \( Z_{ist} \) recorded in the year 1900, when the coal mine has yet to become
obsolete, is correlated with distance to that mine. Similarly, the first difference specification in Eq. 28 which corresponds to a fixed effects model, examines whether changes in characteristic $Z_{ist}$ from the years 1890 to 1900 is correlated with distance to that mine. I choose county characteristics which may affect relative coal capacity as well as overall demand for electricity. These includes total population, urban population. It also includes total output, number of establishments, average number of wage earners per establishment, and capital value in the manufacturing sector.

Columns (1) and (2) of Table 1 presents $\alpha_1$ from Eq. 27 for models with and without state-by-year fixed effects. Columns (3) and (4) presents similar results for estimates of Eq. 28. The years covered for each characteristic is also shown in Table 1. The outcome is not available during each sample county’s obsolescence event because some of the chosen county characteristics are not consistently reported throughout the ICPSR historical census database. Distance to the obsolete coal mine is correlated with several of the selected county characteristics across Columns (1)-(3). However, none of the correlations are statistically significant when characteristics are estimated in pre-obsolescence changes and when state-by-year fixed effects are included as shown in Column (4). This suggests that the inclusion of county fixed effects and state-by-year fixed effects is important for identifying the effects of path dependence.

6.2 Effects of path dependence

Table 2 shows my path dependence effects using the specification in Eq. 26. The outcome in Column (1) is the log relative coal capacity, $\log(\tilde{X}_{ist})$, with coefficients corresponding to the elasticity $\beta^O_\tau$ in Eq. 26. I find effects which increase over time which is consistent with strong path dependence. Relative to the effect of contemporaneous coal prices, a county that is 1% closer to an obsolete coal mine experiences 2.5% greater relative coal-fired capacity six decades after obsolescence. Relating to my model in Section 3, this corresponds to an elasticity of substitution of $\psi = -0.68$.\footnote{68} Column (2) estimates Eq. 26 for the relative coal capacity of new generators during each decade and demonstrates that distance to an obsolete coal mine affects the building of new coal capacity relative to new capacity using other inputs. Columns (3)-(6) explores path dependence effects for the components of the outcomes in Columns (1) and (2). Columns (3) and (4) estimates Eq. 26 for log coal capacity and log new coal capacity respectively, the numerator for the outcomes in Columns (1) and (2). Similarly, Columns (5) and (6) examines log non-coal capacity and log new non-coal capacity, the denominator for the outcomes in Columns (1) and (2). It appears that distance to obsolete coal mines has a persistent effect primarily on coal capacity without much of a net effect for all non-coal capacity suggesting that there is limited spillover onto non-coal electricity capacity.

\footnote{68}$\beta^O_\tau = \frac{\psi^2}{(\psi+1)^2} - \frac{\psi}{\psi+1}$ takes a quadratic form. When $\beta^O_\tau = -2.5$, the root which satisfies Assumption 1 is $\psi = -0.68$.}
The log-log specification in Eq. 26, which is directly informed by my model, assumes that relative coal capacity and distance to an obsolete coal mine has a power function relationship. Figure 7 examines whether the my data supports this functional form assumption by estimating a variant of Eq. 26 which breaks distance to obsolete coal mine into 50 mile distance bins. This allows me to plot the general function $\gamma_\tau + f_\tau(dist_{oi})$ semi-parametrically for each period $\tau$ after the obsolescence event. The shape of the plots shown in Figure 7 generally takes on the shape of a power function.

My results imply both strong path dependence and an optimistic policy scenario for future transitions away from coal-fired capacity. Returning to my initial motivation, I am also interested in the relative role of path dependence in explaining the development of coal-fired capacity over my sample counties throughout the 20th century. To answer this question, I first extract the component of $\log(X_{icst})$ that is driven by path dependence from Column (3) of Table 2:

$$\hat{X}_{icst}^P = \exp \left( \sum_{\tau=1}^{10} \beta_{i\tau} \cdot \log(dist_{oi}) \cdot 1[t = \tau] + \sum_{\tau=1}^{10} \gamma_{i\tau} \cdot 1[t = \tau] \right)$$

such that the component of historical total coal capacity over my sample counties that is unrelated to path dependence is:

$$\sum_i X_{icst}^{NP} = \sum_i \left( X_{icst} - \hat{X}_{icst}^P \right)$$

Figure 8 plots $X_{ct}$ and $X_{ct}^{NP}$ summed over the sample counties from the 1900s to the 1990s. The difference between the two lines, shown in dark shading is the amount of coal-fired capacity explained by path dependence. This component is small initially but grows dramatically after the 1960s. By the 1990s, 60% of coal capacity over my sample counties is driven by path dependence with the rest explained by either time-invariant locational fundamentals or time-varying contemporaneously determinants.

### 6.3 Robustness to subsamples and data construction

I now examine the robustness of my main result to several estimation and data construction decisions. In Table 3, I examine subsamples of my primary dataset. Column (1) reproduces my benchmark result in Column (1) of Table 2. Column (2) shrinks the sample from counties less than 250 miles from the Illinois coal basin to counties that are less than 200 miles and shows that my result do not vary with distance cutoff. Column (3) shows that my estimates are similar when I drop extreme values in log relative coal capacity that are above the 99th percentile and below the 1st percentile in the unconditional distribution. Recall from 32 $\hat{\gamma}_\tau$ is the time-varying effect of the obsolescence event that is unrelated to both distance to obsolete coal mine and distance to subsequent deep coal mine. However, $\hat{\gamma}_\tau$ is relatively insensitive to whether or not distance to subsequent deep coal mine is included in Eq. 26.
Section 5.3 that my sample includes all years prior to an obsolescence event in which a shallow coal mine is active. In Column (4), I keep the year that is immediately before the obsolescence event, showing that my result does not depend on this decision. Finally, beginning in the 1970s, deregulation of public utilities in the U.S. led to more competitive electricity pricing for various parts of the country. To examine whether deregulation altered my full-sample path dependence estimates, in Column (5) I exclude counties that faced obsolescence after the 1970s with estimates that are similar to my benchmark results.

Table 4 conducts further robustness checks by examining whether my results are sensitive to how I’ve constructed both my treatment and outcome data. Again, Column (1) replicates my benchmark result. Recall that my definition of a “large” deep coal mine in the Illinois basin uses a cutoff area of 0.01 square miles. In Column (2) I demonstrate that strong path dependence is detected even as I use a smaller cutoff threshold of 0.005 square miles while Column (3) shows that strong path dependence is also present when a larger cutoff of 0.015 square miles is used. Hotelling’s location model, discussed in Section 5.2, argues that coal prices are dictated by distance to the nearest coal mine. In reality, full price differentiation may be an overly strong assumption. In Column (4), I construct my distance measure based on that of the 3 nearest coal mines rather than just the single nearest mine. The resulting point estimates are very close to that of my benchmark result though the standard errors increase.

In Section 4.3, I discussed that reported generator characteristics are remarkably consistent across EIA-860 forms from 1990 to 2012. As such, my primary county-by-input-by-decade 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 form while in Column (6) I use only the 2005 EIA-860 form. I find evidence of strong path dependence using either particular years of the EIA-860 forms.

7 Mechanisms

My model of directed technical change presented in Section 3 suggests that path dependence in the development of U.S. coal-fired capacity over the 20th century is driven by an accumulation in coal-specific technology stock. However, I have yet to examine whether technological change is the specific mechanism through which path dependence occurs. In this section, I provide additional evidence which supports this particular mechanism. Specifically, I examine the interacted effect of distance to an obsolete coal mine and decade since an obsolescence event on the thermal efficiency of a generator operating today, a standard engineering measure which captures a generator’s technological level. This amounts to a “first-stage” for
the reduced form estimate shown in Eq. 20. Conversely, I do not find evidence in support of other possible mechanisms.

7.1 Testing for technological change

From 1990 to 1995, the EIA-860 forms included data on annual thermal efficiency for steam-based generating units, measured as the ratio of the heat content of electricity produced (in BTUs) over the heat content of input fuel consumed (in BTUs). Unlike physical characteristics of a generating unit such as capacity and type of fuel consumed which are unlikely to vary over the lifetime of that unit, thermal efficiency is more likely to fluctuate annually due to variation in input fuel and electricity prices. Thus, I am unable to reproduce an historical generator-level panel of thermal efficiency as was done for county-level capacity used for my path dependence estimates. Regardless, one could estimate the persistent effects of distance to an obsolete coal mine since obsolescence on thermal efficiency at the generator level in the cross-section. The primary drawback of this approach is the inability to control for county and time effects which as shown in Section 6.1 was important for identifying path dependence effects.

For this cross-sectional specification, log thermal efficiency, $\log(\text{eff}_{gis})$, averaged across the 1990-1995 EIA-860 forms for generator $g$, in county $i$, and state $s$, is modeled as:

$$
\log(\text{eff}_{gis}) = \kappa_1 \log(\text{dist}^O_i) * \text{sinceEvent}_i + \kappa_2 \log(\text{dist}^O_i) + \kappa_3 \log(\text{dist}^A_i) * \text{sinceEvent}_i + \kappa_4 \log(\text{dist}^A_i) + \kappa_5 \text{sinceEvent}_i + \omega \text{Z}_g + \eta_s + \mu_{gis}
$$

(31)

where $\text{sinceEvent}$ is the number of decades between the 1990s when thermal efficiency is observed and the initial decade of obsolescence for county $i$. $\kappa_1$ is my coefficient of interest and captures the change in the effect of distance to an obsolete coal mine as time since obsolescence increases. $\kappa_2$ captures the effect of distance to the obsolete coal mine prior to obsolescence. In order to control for the effect of coal prices after obsolescence, I also include the interacted and uninteracted effects of distance to the nearest deep coal mine following obsolescence with $\kappa_3$ and $\kappa_4$ having similar interpretations as $\kappa_1$ and $\kappa_2$. Finally, because Eq. (31) uses only cross sectional variation, I add controls $\text{Z}_g$ indicating the generator’s latitude, longitude, and state-fixed effects, $\eta_s$, that serve as proxies for time-invariant characteristics. $\text{Z}_g$ also includes the generator’s age and the number of generators on the given power plant site which may affect the efficiency of a generating unit.

Column (1) of Table 5 displays estimates for Eq. (31) for my main sample of generators located in counties less than 250 miles away from the Illinois Basin. $\kappa_1$ is statistically significant and negative implying that the path dependence effect increases in magnitude as more time passes since the obsolescence event. Column (2)
shows that result are nearly identical if I only consider counties less than 200 miles from the Illinois Basin. To test whether thermal efficiency is the mechanism driving path dependence, I include the generator’s capacity in Column (3). Because capacity is itself a function of technology stock, this creates a bad or proxy control problem (see p. 66 of Angrist and Pischke (2008)) and thus would not lead to the identification of $\kappa_1$. However, if estimates of $\kappa_1$ changes dramatically after capacity is included, this may suggest that technological change is a possible driver for coal capacity. Indeed Column (3) shows that $\kappa_1$ gets smaller in magnitude and is no longer statistically significant when generator capacity is included in Eq. 31. This occurs as well for the sample of generators located in counties that are less than 200 miles away from the Illinois Basin as shown in Column (4).

7.2 Alternative mechanisms

Could path dependence in relative coal capacity be driven by other mechanisms? Table 6 tests a few competing mechanisms. One alternative explanation involves the presence of other durable capital with high fixed costs and low depreciation rates that may be complementary to input-specific electricity capacity. For example, electricity production requires other capital investments, in particular transport infrastructure for delivering energy inputs. As such, counties that used to be close to active coal mines may have historically invested in transport infrastructure such that coal-fired capacity was preferred even after nearby coal mines became obsolete due to the high fixed costs of building transport infrastructure for delivering other energy inputs.

I offer two lines of evidence to suggest this explanation is less plausible. First, complementary capital with high fixed costs and low depreciation rates such as transport infrastructure may lead to path dependence in relative coal capacity, but it should not result in path dependence which increases over time as I find in Section 6.2. Second, if transport infrastructure is driven path dependence in coal capacity, one should observe that counties that are closer to obsolete coal mines have both more transport infrastructure today and perhaps more infrastructure as a function of time since obsolescence.

I conduct this test in Column (1) of Table 6 by examining railroad density, measured as miles of railroad tracks per square mile in 2010, using a county-level version of the specification in Eq. 31. I find that distance to an obsolete coal mine does not have a statistically significant effect on modern railroad density regardless of time since obsolescence. Column (2) shows this is also the case for counties located less than 200 miles from the Illinois Basin. Columns (3) and (4) examine highway density for the two county samples and also cannot reject a null effect.

The remainder of Table 6 examines county characteristics in 2000 related to electricity demand similar
to those examined prior to obsolescence in Section 6.1. I do not detect a statistically significant relationship between distance to the obsolete coal mine and modern county-level population, number of manufacturing establishments, or average number of wage earners per establishment. While the results in Table 6 does not allow me to entirely rule out path dependence driven by other explanations, this evidence combined with that shown in Table 5 suggests that path dependence is more likely arising due to technological change.

8 Discussion

This paper examines the relative role of path dependence, defined as the ongoing effect of obsolete past determinants, on the development of U.S. electricity capacity over the 20th century. Specifically, I find that counties closer to an obsolete coal mine continued to develop more coal-fired capacity than counties further away for up to six decades after the coal mine becomes obsolete. By the 1990s, 60% of total coal capacity over my sample counties can be explained by path dependence. This suggests that path dependence played an increasingly important role in the expansion of U.S. coal-fired capacity over the course of the 20th century, possibly dominating the contributions of time-invariant locational fundamentals and time-varying contemporaneous factors in determining the energy mix of the modern U.S. electricity sector. Auxiliary evidence suggests that path dependence is being driven by accumulations over time in coal-specific technology.

The presence of increasing path dependence effects has important implications for the long-run transition away from coal. Under such circumstances, the same dynamics of technological accumulation which has historically favored coal-fired capacity can also be exploited to induce technological change for capacity using other inputs. The implied policy intervention therefore only requires a temporary subsidy in non-coal research until non-coal technology catches up with that of coal. Once this “boost” is completed, the increased market size for non-coal capacity would initiate similar dynamics of path dependence towards non-coal capacity that has characterized coal-fired capacity over the 20th century. This contrasts with the policy intervention implied by the locational fundamentals view in which the time-invariant relative abundance of coal can only be offset by a permanent adjustment to relative input prices. Thus, by finding strong path dependence in the historical development of coal-fired electricity, this paper lends an optimistic view towards future climate and energy policies aimed at transitions in energy input use.

This result is particularly timely given the energy transition currently underway within the U.S. The recent expansion in natural gas production due to hydraulic fracturing has led both to the large-scale retirement of existing coal-fired capacity and the proposed construction of new natural gas-fired capacity. Thus, the requirements for transitioning away from coal-fired capacity suggested by this paper’s results may already
be underway due not to policy interventions but rather to dramatic changes in relative input prices.

Regardless, by examining the role of path dependence in the expansion of U.S. coal-fired electricity capacity during the 20th century, this paper provide two important insights into the current transition towards natural gas-fired electricity. First, the presence of strong path dependence within electricity sector suggests that a permanent transition from coal to natural gas-fired electricity can be induced with a temporary natural gas relative price advantage provided it is sufficiently strong. Under these circumstances, the dynamics of path dependence and directed technical change would ensure that at some point, research in natural gas technology would becomes more profitable than research in coal technology allowing relative natural gas capacity to grow endogenously thereafter. Second, burning of natural gas, while less carbon intensive than coal nonetheless contributes to climate change. Thus, if damages from climate change are sufficiently large to require dramatic cuts in greenhouse gas emissions and technological spillovers are limited, a structural transition towards natural gas today may only further delay the eventually needed transition towards zero-carbon electricity. This is because at some point, research must be directed towards zero-carbon electricity in order to offset the technological advantage now accumulated towards natural gas use. Thus, if the history of path dependence in coal-fired capacity is repeated, it may be nearly another century before electricity production becomes fully decarbonized.
References


Figures

Figure 1: County-level coal-fired capacity in the 1920s and 1990s

Notes: Scatter shows county-level log relative coal capacity in the 1990s against log relative coal capacity during 1920s. Data reconstructed using modern EIA-860 forms (see Section 4).

Figure 2: U.S. electricity capacity by input (1900-2000)

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 (see Section 4).
Figure 3: Map of Illinois coal basin by reserve depth and mine location

Notes: Green/lighter shaded area indicates the location of coal reserves less than 200 feet deep. Black/darker shaded area indicates the location of coal reserves 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 reserve. Light bars shows mines located over coal reserves less than 200 feet deep. Dark bars shows mines located over coal reserves greater than 200 feet deep.
**Figure 5:** 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.4). Light shading indicates pre-obsolescence event. Dark shading indicates decades during and after obsolescence event.

**Figure 6:** Path dependence result: effect of distance to obsolete coal mine

*Notes:* Plot shows coefficients $\beta^O_{\tau}$ and 90% confidence intervals from Eq. 26. Coefficients correspond to the change in relative coal capacity for a county that is 1% closer to an obsolete coal mine $\tau$ decades after obsolescence. Estimates are relative to the pre-obsolescence period $\tau = 0$. 
Figure 7: Nonlinearity of path dependence effect

Notes: Plot examines the nonlinearity of path dependence effect and shows distance bin dummies for the path dependence effect from 1 to 4 decades after obsolescence.

Figure 8: Total coal-fired capacity with and without path dependence

Notes: Plot breaks down total coal-fired capacity summed across sample counties into components predicted by path dependence effects and by all other determinants (see Eqs. 29 and 30).
Table 1: County characteristics by distance to obsolete mine before obsolescence

<table>
<thead>
<tr>
<th>Outcome</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-obsolescence levels</td>
<td>Pre-obsolescence changes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log population</td>
<td>-0.42***</td>
<td>-0.36**</td>
<td>0.094***</td>
<td>-0.024</td>
</tr>
<tr>
<td>1890-1990</td>
<td>[0.16]</td>
<td>[0.18]</td>
<td>[0.029]</td>
<td>[0.024]</td>
</tr>
<tr>
<td>Obs</td>
<td>261</td>
<td>261</td>
<td>261</td>
<td>261</td>
</tr>
<tr>
<td>Log urban population</td>
<td>-0.81***</td>
<td>-0.4</td>
<td>0.075</td>
<td>-0.026</td>
</tr>
<tr>
<td>1890-1980</td>
<td>[0.26]</td>
<td>[0.42]</td>
<td>[0.056]</td>
<td>[0.076]</td>
</tr>
<tr>
<td></td>
<td>183</td>
<td>183</td>
<td>171</td>
<td>171</td>
</tr>
<tr>
<td>log mfg output</td>
<td>-1.71***</td>
<td>-0.5</td>
<td>0.52***</td>
<td>0.13</td>
</tr>
<tr>
<td>1890-1940</td>
<td>[0.56]</td>
<td>[0.65]</td>
<td>[0.17]</td>
<td>[0.25]</td>
</tr>
<tr>
<td></td>
<td>114</td>
<td>114</td>
<td>113</td>
<td>113</td>
</tr>
<tr>
<td>log mfg establishments</td>
<td>0.61**</td>
<td>-0.52**</td>
<td>0.57***</td>
<td>0.11</td>
</tr>
<tr>
<td>1890-1990</td>
<td>[0.26]</td>
<td>[0.22]</td>
<td>[0.15]</td>
<td>[0.10]</td>
</tr>
<tr>
<td></td>
<td>260</td>
<td>260</td>
<td>260</td>
<td>260</td>
</tr>
<tr>
<td>log mfg employment</td>
<td>-1.34***</td>
<td>-1.02***</td>
<td>0.081</td>
<td>-0.014</td>
</tr>
<tr>
<td>1890-1990</td>
<td>[0.32]</td>
<td>[0.37]</td>
<td>[0.12]</td>
<td>[0.11]</td>
</tr>
<tr>
<td></td>
<td>258</td>
<td>258</td>
<td>255</td>
<td>255</td>
</tr>
<tr>
<td>log mfg capital</td>
<td>-0.85*</td>
<td>-0.44</td>
<td>0.29*</td>
<td>0.23</td>
</tr>
<tr>
<td>1890-1900</td>
<td>[0.50]</td>
<td>[0.62]</td>
<td>[0.16]</td>
<td>[0.18]</td>
</tr>
<tr>
<td></td>
<td>106</td>
<td>106</td>
<td>105</td>
<td>105</td>
</tr>
</tbody>
</table>

State-by-year FEs NO YES NO YES

Notes: Each coefficient from a separate regression. Outcomes are county-level characteristics at the beginning of the decade in which obsolescence occurs. Data coverage for outcome variables indicated below variable name. All models include log distance to obsolete mine (reported coefficient) and log distance to subsequent mine (coefficient now shown). Columns (1) and (2) estimates outcome log levels. Columns (3) and (4) estimates outcome log first differences. Columns (2) and (4) include state-by-year fixed effects. Robust standard errors clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1
### Table 2: Path dependence result: Effect of distance to obsolete coal mine

<table>
<thead>
<tr>
<th>Persistence effects</th>
<th>(1) coal/other capacity</th>
<th>(2) new coal/other capacity</th>
<th>(3) coal capacity</th>
<th>(4) new coal capacity</th>
<th>(5) other capacity</th>
<th>(6) new other capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 decade later</td>
<td>-0.19</td>
<td>-0.065</td>
<td>-0.47*</td>
<td>-0.37</td>
<td>-0.28</td>
<td>-0.31</td>
</tr>
<tr>
<td></td>
<td>[0.39]</td>
<td>[0.77]</td>
<td>[0.26]</td>
<td>[0.36]</td>
<td>[0.23]</td>
<td>[0.49]</td>
</tr>
<tr>
<td>2 decades later</td>
<td>-1.13***</td>
<td>-0.73*</td>
<td>-1.28***</td>
<td>-1.19***</td>
<td>-0.15</td>
<td>-0.46</td>
</tr>
<tr>
<td></td>
<td>[0.36]</td>
<td>[0.39]</td>
<td>[0.32]</td>
<td>[0.29]</td>
<td>[0.28]</td>
<td>[0.32]</td>
</tr>
<tr>
<td>3 decades later</td>
<td>-1.63***</td>
<td>-0.4</td>
<td>-1.73***</td>
<td>-0.54</td>
<td>-0.1</td>
<td>-0.14</td>
</tr>
<tr>
<td></td>
<td>[0.48]</td>
<td>[0.40]</td>
<td>[0.43]</td>
<td>[0.56]</td>
<td>[0.34]</td>
<td>[0.40]</td>
</tr>
<tr>
<td>4 decades later</td>
<td>-1.73***</td>
<td>0.73</td>
<td>-1.90***</td>
<td>-0.02</td>
<td>-0.17</td>
<td>-0.75**</td>
</tr>
<tr>
<td></td>
<td>[0.57]</td>
<td>[0.56]</td>
<td>[0.49]</td>
<td>[0.27]</td>
<td>[0.35]</td>
<td>[0.37]</td>
</tr>
<tr>
<td>5 decades later</td>
<td>-2.43***</td>
<td>-0.75</td>
<td>-2.40***</td>
<td>-0.31</td>
<td>0.034</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>[0.68]</td>
<td>[0.48]</td>
<td>[0.49]</td>
<td>[0.28]</td>
<td>[0.47]</td>
<td>[0.35]</td>
</tr>
<tr>
<td>6 decades later</td>
<td>-2.51**</td>
<td>-0.47</td>
<td>-2.20***</td>
<td>-0.38</td>
<td>0.31</td>
<td>0.088</td>
</tr>
<tr>
<td></td>
<td>[1.00]</td>
<td>[0.81]</td>
<td>[0.77]</td>
<td>[0.61]</td>
<td>[0.66]</td>
<td>[0.58]</td>
</tr>
<tr>
<td>7 decades later</td>
<td>-0.39</td>
<td>1.72</td>
<td>-0.76</td>
<td>0.38</td>
<td>-0.38</td>
<td>-1.34*</td>
</tr>
<tr>
<td></td>
<td>[1.52]</td>
<td>[1.20]</td>
<td>[1.29]</td>
<td>[0.75]</td>
<td>[1.02]</td>
<td>[0.75]</td>
</tr>
<tr>
<td>8 decades later</td>
<td>-0.98</td>
<td>0.37</td>
<td>-0.81</td>
<td>0.64</td>
<td>0.16</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>[1.48]</td>
<td>[1.26]</td>
<td>[1.36]</td>
<td>[0.95]</td>
<td>[1.10]</td>
<td>[0.74]</td>
</tr>
<tr>
<td>9 decades later</td>
<td>-0.43</td>
<td>0.023</td>
<td>-0.48</td>
<td>-0.21</td>
<td>-0.044</td>
<td>-0.23</td>
</tr>
<tr>
<td></td>
<td>[1.56]</td>
<td>[0.64]</td>
<td>[1.38]</td>
<td>[0.60]</td>
<td>[1.14]</td>
<td>[0.42]</td>
</tr>
<tr>
<td>10 decades later</td>
<td>-0.31</td>
<td>0.53</td>
<td>-0.0022</td>
<td>-0.71</td>
<td>0.31</td>
<td>-1.24*</td>
</tr>
<tr>
<td></td>
<td>[1.74]</td>
<td>[0.81]</td>
<td>[1.36]</td>
<td>[0.44]</td>
<td>[1.11]</td>
<td>[0.65]</td>
</tr>
</tbody>
</table>

**Notes:** Estimates of distance to mine elasticity using Equation 26. Sample includes counties that are less than 250 miles from nearest Illinois Basin coal reserve and are closer to nearest Illinois Basin coal reserve than to nearest Appalachian Basin coal reserve. Each model includes event time dummies, time-varying effects of distance to subsequent deep mine, county fixed effects, and year-by-state fixed effects. Outcome in Columns (1) and (2) is relative coal-fired capacity. Outcome in Columns (3) and (4) is coal-fired capacity. Outcome in Columns (5) and (6) are other input capacity. Columns (1), (3), and (5) estimates operating capacity while Columns (2), (4), and (6) estimates new capacity. Robust standard errors clustered at the county-level. *** p<0.01, ** p<0.05, * p<0.1
Table 3: Robustness of path dependence result: subsamples

<table>
<thead>
<tr>
<th>Persistence effects</th>
<th>Benchmark</th>
<th>Distance &lt;200 miles</th>
<th>Drop outliers &lt;1%, &gt;99%</th>
<th>Keep 1 active year</th>
<th>Event before 1970s</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 decade later</td>
<td>-0.19</td>
<td>-0.17</td>
<td>-0.11</td>
<td>-0.13</td>
<td>-0.47*</td>
</tr>
<tr>
<td></td>
<td>[0.39]</td>
<td>[0.40]</td>
<td>[0.42]</td>
<td>[0.42]</td>
<td>[0.26]</td>
</tr>
<tr>
<td>2 decades later</td>
<td>-1.13***</td>
<td>-1.22***</td>
<td>-0.91***</td>
<td>-1.07***</td>
<td>-1.28***</td>
</tr>
<tr>
<td></td>
<td>[0.36]</td>
<td>[0.38]</td>
<td>[0.34]</td>
<td>[0.38]</td>
<td>[0.32]</td>
</tr>
<tr>
<td>3 decades later</td>
<td>-1.63***</td>
<td>-1.74***</td>
<td>-1.13***</td>
<td>-1.49***</td>
<td>-1.73***</td>
</tr>
<tr>
<td></td>
<td>[0.48]</td>
<td>[0.52]</td>
<td>[0.38]</td>
<td>[0.50]</td>
<td>[0.43]</td>
</tr>
<tr>
<td>4 decades later</td>
<td>-1.73***</td>
<td>-1.78***</td>
<td>-1.19***</td>
<td>-1.46**</td>
<td>-1.90**</td>
</tr>
<tr>
<td></td>
<td>[0.57]</td>
<td>[0.61]</td>
<td>[0.45]</td>
<td>[0.61]</td>
<td>[0.49]</td>
</tr>
<tr>
<td>5 decades later</td>
<td>-2.43***</td>
<td>-2.55***</td>
<td>-1.68***</td>
<td>-2.23***</td>
<td>-2.40***</td>
</tr>
<tr>
<td></td>
<td>[0.68]</td>
<td>[0.70]</td>
<td>[0.50]</td>
<td>[0.74]</td>
<td>[0.49]</td>
</tr>
<tr>
<td>6 decades later</td>
<td>-2.51**</td>
<td>-2.58**</td>
<td>-2.22***</td>
<td>-2.25**</td>
<td>-2.20**</td>
</tr>
<tr>
<td></td>
<td>[1.00]</td>
<td>[1.01]</td>
<td>[0.79]</td>
<td>[1.03]</td>
<td>[0.77]</td>
</tr>
<tr>
<td>7 decades later</td>
<td>-0.39</td>
<td>-0.3</td>
<td>-0.96</td>
<td>-0.21</td>
<td>-0.76</td>
</tr>
<tr>
<td></td>
<td>[1.52]</td>
<td>[1.53]</td>
<td>[1.57]</td>
<td>[1.60]</td>
<td>[1.29]</td>
</tr>
<tr>
<td>8 decades later</td>
<td>-0.98</td>
<td>-0.98</td>
<td>-0.67</td>
<td>-0.78</td>
<td>-0.81</td>
</tr>
<tr>
<td></td>
<td>[1.48]</td>
<td>[1.49]</td>
<td>[1.69]</td>
<td>[1.55]</td>
<td>[1.36]</td>
</tr>
<tr>
<td>9 decades later</td>
<td>-0.43</td>
<td>-0.32</td>
<td>-0.36</td>
<td>-0.23</td>
<td>-0.48</td>
</tr>
<tr>
<td></td>
<td>[1.56]</td>
<td>[1.56]</td>
<td>[1.72]</td>
<td>[1.63]</td>
<td>[1.38]</td>
</tr>
<tr>
<td>10 decades later</td>
<td>-0.31</td>
<td>-0.27</td>
<td>-0.23</td>
<td>-0.11</td>
<td>-0.0022</td>
</tr>
<tr>
<td></td>
<td>[1.74]</td>
<td>[1.73]</td>
<td>[1.87]</td>
<td>[1.79]</td>
<td>[1.36]</td>
</tr>
</tbody>
</table>

Observations: 2113, 1756, 2051, 1966, 1957
Counties: 261, 208, 261, 261, 229

Notes: Estimates of distance to mine elasticity using Equation 26. Sample includes counties closer to nearest Illinois Basin coal reserve than to nearest Appalachian Basin coal reserve. Each model includes event time dummies, time-varying effects of distance to subsequent closest mine, county fixed effects, and year-by-state fixed effects. Column (1) shows benchmark model. Column (2) includes only counties less than 200 miles away from nearest Illinois Basin coal reserve. Column (3) drops outcome outliers less than 1st percentile and greater than the 99th percentile. Column (4) keeps only the year that is immediately before obsolescence event. Column (5) includes only counties with obsolescence event occurring before 1970s. Robust standard errors clustered at the county-level. *** p<0.01, ** p<0.05, * p<0.1
Table 4: Robustness of path dependence result: data construction

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>&gt;0.005 sq. mile</td>
<td>&gt;0.015 sq. mile</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 decade later</td>
<td>-0.19</td>
<td>-0.81**</td>
<td>-0.84</td>
<td>-0.25</td>
<td>-0.057</td>
<td>-0.3</td>
</tr>
<tr>
<td></td>
<td>[0.39]</td>
<td>[0.37]</td>
<td>[0.56]</td>
<td>[0.76]</td>
<td>[0.50]</td>
<td>[0.82]</td>
</tr>
<tr>
<td>2 decades later</td>
<td>-1.13***</td>
<td>-1.21**</td>
<td>-2.00***</td>
<td>-0.85</td>
<td>-1.24**</td>
<td>-0.98**</td>
</tr>
<tr>
<td></td>
<td>[0.36]</td>
<td>[0.53]</td>
<td>[0.73]</td>
<td>[0.95]</td>
<td>[0.50]</td>
<td>[0.44]</td>
</tr>
<tr>
<td>3 decades later</td>
<td>-1.63***</td>
<td>-1.34*</td>
<td>-2.58***</td>
<td>-1.77</td>
<td>-1.82***</td>
<td>-1.35**</td>
</tr>
<tr>
<td></td>
<td>[0.48]</td>
<td>[0.74]</td>
<td>[0.80]</td>
<td>[1.13]</td>
<td>[0.61]</td>
<td>[0.57]</td>
</tr>
<tr>
<td>4 decades later</td>
<td>-1.73***</td>
<td>-1.87***</td>
<td>-2.82***</td>
<td>-2.09*</td>
<td>-1.55**</td>
<td>-1.48**</td>
</tr>
<tr>
<td></td>
<td>[0.57]</td>
<td>[0.67]</td>
<td>[0.80]</td>
<td>[1.26]</td>
<td>[0.62]</td>
<td>[0.61]</td>
</tr>
<tr>
<td>5 decades later</td>
<td>-2.43***</td>
<td>-1.5</td>
<td>-2.97***</td>
<td>-2.84*</td>
<td>-2.60***</td>
<td>-1.76**</td>
</tr>
<tr>
<td></td>
<td>[0.68]</td>
<td>[0.93]</td>
<td>[1.03]</td>
<td>[1.63]</td>
<td>[0.83]</td>
<td>[0.83]</td>
</tr>
<tr>
<td>6 decades later</td>
<td>-2.51**</td>
<td>-0.74</td>
<td>-2.95**</td>
<td>-0.81</td>
<td>-0.59</td>
<td>-2.05*</td>
</tr>
<tr>
<td></td>
<td>[1.00]</td>
<td>[1.04]</td>
<td>[1.39]</td>
<td>[3.09]</td>
<td>[1.31]</td>
<td>[1.15]</td>
</tr>
<tr>
<td>7 decades later</td>
<td>-0.39</td>
<td>1.79</td>
<td>-4.06***</td>
<td>0.51</td>
<td>1.24</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>[1.52]</td>
<td>[1.79]</td>
<td>[1.49]</td>
<td>[3.62]</td>
<td>[1.38]</td>
<td>[1.64]</td>
</tr>
<tr>
<td>8 decades later</td>
<td>-0.98</td>
<td>0.77</td>
<td>-3.61**</td>
<td>-0.27</td>
<td>0.44</td>
<td>-0.12</td>
</tr>
<tr>
<td></td>
<td>[1.48]</td>
<td>[1.90]</td>
<td>[1.78]</td>
<td>[3.74]</td>
<td>[1.41]</td>
<td>[1.63]</td>
</tr>
<tr>
<td>9 decades later</td>
<td>-0.43</td>
<td>1.7</td>
<td>-2.96</td>
<td>0.77</td>
<td>1.31</td>
<td>-0.094</td>
</tr>
<tr>
<td></td>
<td>[1.56]</td>
<td>[1.87]</td>
<td>[1.86]</td>
<td>[3.74]</td>
<td>[1.41]</td>
<td>[1.63]</td>
</tr>
<tr>
<td>10 decades later</td>
<td>-0.31</td>
<td>0.44</td>
<td>-3.29</td>
<td>-0.44</td>
<td>1.49</td>
<td>-0.19</td>
</tr>
<tr>
<td></td>
<td>[1.74]</td>
<td>[2.28]</td>
<td>[2.14]</td>
<td>[4.43]</td>
<td>[1.82]</td>
<td>[1.87]</td>
</tr>
</tbody>
</table>

Observations: 2113, 862, 2505, 1830, 1796, 1856
Counties: 261, 187, 303, 230, 227, 228

Notes: Estimates of distance to mine elasticity using Equation 26. Sample includes counties closer to nearest Illinois Basin coal reserve than to nearest Appalachian Basin coal reserve. Each model includes event time dummies, time-varying effects of distance to subsequent closest mine, county fixed effects, and year-by-state fixed effects. 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 3 mines with area greater than 0.01 sq. miles. Outcome in Column (5) built from only the 2000 EIA-860 form. Outcome in Column (6) 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 5: Effect of distance to obsolete coal mine on modern generator efficiency

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dep. var. is log generator thermal efficiency</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>log(distO) x sinceTreat</td>
<td>-0.032**</td>
<td>-0.033**</td>
<td>-0.021</td>
<td>-0.019</td>
</tr>
<tr>
<td></td>
<td>[0.014]</td>
<td>[0.016]</td>
<td>[0.016]</td>
<td>[0.017]</td>
</tr>
<tr>
<td>log(distA) x sinceTreat</td>
<td>0.041</td>
<td>0.045</td>
<td>0.032</td>
<td>0.042</td>
</tr>
<tr>
<td></td>
<td>[0.030]</td>
<td>[0.054]</td>
<td>[0.032]</td>
<td>[0.054]</td>
</tr>
<tr>
<td>log(distO)</td>
<td>0.24</td>
<td>0.25</td>
<td>0.16</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>[0.16]</td>
<td>[0.24]</td>
<td>[0.18]</td>
<td>[0.25]</td>
</tr>
<tr>
<td>log(distA)</td>
<td>-0.27</td>
<td>-0.32</td>
<td>-0.19</td>
<td>-0.3</td>
</tr>
<tr>
<td></td>
<td>[0.30]</td>
<td>[0.54]</td>
<td>[0.31]</td>
<td>[0.53]</td>
</tr>
<tr>
<td>sinceTreat</td>
<td>-0.041</td>
<td>-0.051</td>
<td>-0.046</td>
<td>-0.09</td>
</tr>
<tr>
<td></td>
<td>[0.11]</td>
<td>[0.21]</td>
<td>[0.11]</td>
<td>[0.20]</td>
</tr>
<tr>
<td>age</td>
<td>-0.0078***</td>
<td>-0.0077***</td>
<td>-0.0055**</td>
<td>-0.0051*</td>
</tr>
<tr>
<td></td>
<td>[0.0026]</td>
<td>[0.0029]</td>
<td>[0.0024]</td>
<td>[0.0027]</td>
</tr>
<tr>
<td># of generators on plant</td>
<td>0.017</td>
<td>0.019</td>
<td>0.019*</td>
<td>0.022*</td>
</tr>
<tr>
<td></td>
<td>[0.011]</td>
<td>[0.013]</td>
<td>[0.011]</td>
<td>[0.013]</td>
</tr>
<tr>
<td>capacity</td>
<td>0.00018</td>
<td>0.00022*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[0.00011]</td>
<td>[0.00012]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>267</td>
<td>231</td>
<td>267</td>
<td>231</td>
</tr>
<tr>
<td>County sample</td>
<td>&lt;250 mi.</td>
<td>&lt;200 mi.</td>
<td>&lt;250 mi.</td>
<td>&lt;200 mi.</td>
</tr>
</tbody>
</table>

Notes: Estimates from Eq. 31. Sample includes counties closer to nearest Illinois Basin coal reserve than to nearest Appalachian Basin coal reserve. Each model includes generator latitude and longitude and state-fixed effects. Odd columns include generators in counties located less than 250 miles from Illinois coal basin. Even columns include generators in counties located less than 200 miles from Illinois coal basin. Columns (1) and (2) show benchmark model. Columns (3) and (4) adds generator capacity. Robust standard errors clustered at the county-level. *** p<0.01, ** p<0.05, * p<0.1
Table 6: Effect of distance to obsolete coal mine on modern county characteristics

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
<th>(10)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rail density</td>
<td>Highway density</td>
<td>Population</td>
<td>Mgf establishments</td>
<td>Mgf employment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>log(distO) x sinceTreat</td>
<td>0.029</td>
<td>0.0064</td>
<td>-0.015</td>
<td>-0.021</td>
<td>-0.12</td>
<td>-0.14</td>
<td>-0.12</td>
<td>-0.13</td>
<td>-0.07</td>
<td>-0.05</td>
</tr>
<tr>
<td></td>
<td>[0.055]</td>
<td>[0.053]</td>
<td>[0.030]</td>
<td>[0.028]</td>
<td>[0.12]</td>
<td>[0.11]</td>
<td>[0.10]</td>
<td>[0.096]</td>
<td>[0.062]</td>
<td>[0.063]</td>
</tr>
<tr>
<td>log(distA) x sinceTreat</td>
<td>-0.074</td>
<td>-0.023</td>
<td>-0.037</td>
<td>-0.018</td>
<td>0.0051</td>
<td>0.059</td>
<td>0.03</td>
<td>0.054</td>
<td>0.11*</td>
<td>0.097</td>
</tr>
<tr>
<td></td>
<td>[0.075]</td>
<td>[0.073]</td>
<td>[0.046]</td>
<td>[0.036]</td>
<td>[0.16]</td>
<td>[0.14]</td>
<td>[0.13]</td>
<td>[0.11]</td>
<td>[0.063]</td>
<td>[0.065]</td>
</tr>
<tr>
<td>log(distO)</td>
<td>-0.24</td>
<td>0.0023</td>
<td>-0.13</td>
<td>-0.035</td>
<td>0.25</td>
<td>0.47</td>
<td>0.45</td>
<td>0.55</td>
<td>0.53</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>[0.43]</td>
<td>[0.40]</td>
<td>[0.26]</td>
<td>[0.22]</td>
<td>[1.09]</td>
<td>[0.98]</td>
<td>[0.84]</td>
<td>[0.76]</td>
<td>[0.47]</td>
<td>[0.48]</td>
</tr>
<tr>
<td>log(distA)</td>
<td>0.24</td>
<td>-0.24</td>
<td>0.38</td>
<td>0.16</td>
<td>0.038</td>
<td>-0.51</td>
<td>-0.19</td>
<td>-0.46</td>
<td>-0.9</td>
<td>-0.78</td>
</tr>
<tr>
<td></td>
<td>[0.66]</td>
<td>[0.64]</td>
<td>[0.42]</td>
<td>[0.32]</td>
<td>[1.47]</td>
<td>[1.25]</td>
<td>[1.16]</td>
<td>[0.99]</td>
<td>[0.56]</td>
<td>[0.57]</td>
</tr>
<tr>
<td>sinceTreat</td>
<td>0.16</td>
<td>0.043</td>
<td>0.16</td>
<td>0.12</td>
<td>0.34</td>
<td>0.2</td>
<td>0.21</td>
<td>0.18</td>
<td>-0.19</td>
<td>-0.24</td>
</tr>
<tr>
<td></td>
<td>[0.20]</td>
<td>[0.20]</td>
<td>[0.12]</td>
<td>[0.084]</td>
<td>[0.33]</td>
<td>[0.28]</td>
<td>[0.32]</td>
<td>[0.29]</td>
<td>[0.16]</td>
<td>[0.18]</td>
</tr>
</tbody>
</table>

Observations: 458 378 458 378 458 378 376 313 354 292
County sample: <250 mi. <200 mi. <250 mi. <200 mi. <250 mi. <200 mi. <250 mi. <200 mi. <200 mi.

Notes: Outcomes in logs. Sample includes counties closer to nearest Illinois Basin coal reserve than to nearest Appalachian Basin coal reserve. Each model includes county latitude and longitude and state-fixed effects. Odd columns include counties located less than 250 miles from Illinois coal basin. Even columns include counties located less than 200 miles from Illinois coal basin. Robust standard errors clustered at the county-level. *** p<0.01, ** p<0.05, * p<0.1
Appendix A  The history of U.S. electricity data collection

Section 4.2 notes that actual historical data on electricity capacity by energy input at or below the county-level from 1890-200 were either never collected or likely to no longer exist. This section summarizes what relevant data were collected and the gap in data availability today which prompts the need to reconstruct a historical panel using modern EIA-860 records.

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) and not energy input which is necessary for determining input-specific capacity. Input costs are provided by fossil-fuel but fuel price would be needed in order to recover input fuel consumption which could then be converted to input-specific capacity using a capacity factor.

**Availability**  Summaries of these censuses with aggregate statistics are available digitally.\(^{33}\) However, the original power plant-level data could not be located following private conversation with archivists at the National Archives and Records Administration (NARA).\(^{34}\)

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 capacity and built year at the generating-unit. However, it only shows fuel use at the power plant level which makes assignment of generating-unit fuel use difficult.

**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”\(^{35}\) which was digitized and used for my data validation exercises in Section 4.3. The ”Steam-Electric Plant Construction Cost and Annual Production Expenses” from 1948 to 1974 has plant-level values from Form 1 and Form 12.\(^{36}\) Unfortunately, in addition to the limitations mentioned above for Form 1 and 12, power plant coverage is low. In 1948, this report includes only 219 power plants covering 67% of total U.S. steam capacity. Finally, private conversation with NARA archivists suggests that the original Form 4 and Form 12 documents may no longer be held by NARA.

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

\(^{33}\) Available: http://hdl.handle.net/2027/mdp.39015028113663
\(^{34}\) 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.
\(^{35}\) Available: http://hdl.handle.net/2027/mdp.39015023906806
\(^{36}\) Available: http://catalog.hathitrust.org/Record/000904499
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 include 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\(^{37}\) and was digitized for my data validation exercises discussed in Section 4.3. 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.\(^{38}\)
Appendix B  Calculating capacity factors

Section 4.3 compares national aggregates based on my reconstruction of historical county-level capacity using EIA-860 forms against true historical census data. The “Historical Statistics of the United States from Colonial Times to 1970” provides national electricity production (gigawatt-hours) by input fuel, $Y_{jt}$. To compare national input-specific capacity, $X_{jt}$ to national input-specific production, $Y_{jt}$, one must use an annual capacity factor, $F_{jt}$, defined for a generating unit as the ratio of electricity produced over the electricity produced if that generating unit were running continuously over the year. Observe that a capacity factor falls outside the theoretical considerations in Section 3 as my model assumes that capacity is fully used within a given year. The relationship is:

$$Y_{jt} = 8760 \times X_{jt} \times F_{jt} \quad (A.1)$$

where 8760 is the number of hours in the year. I want to compare my EIA-860 reconstructed coal and non-coal steam (i.e. non-hydro) capacity, $X_{ct}^{EIA}$ and $X_{nt}^{EIA}$, against aggregate census coal and non-coal electricity output, $Y_{ct}^{census}$ and $Y_{nt}^{census}$. However, I only have the capacity factor for coal, $F_{ct}$, obtained from McNerney, Farmer and Trancik (2011) and capacity-weighted aggregate capacity-factor for all steam capacity obtained from U.S. Census Bureau (1975), defined as $\tilde{F}_{t} = \frac{Y_{ct}^{census} + Y_{nt}^{census}}{8760(X_{ct}^{EIA} + X_{nt}^{EIA})}$. This allows me to recover the following implied input-specific production using my EIA-860 reconstructed data:

$$Y_{ct}^{EIA} = 8760 \times X_{ct}^{EIA} \times F_{ct} \quad (A.2)$$
$$Y_{nt}^{EIA} = 8760 \times \tilde{F}(X_{ct}^{EIA} + X_{nt}^{EIA}) - Y_{ct}^{EIA} \quad (A.3)$$

Figure A.6 compares my EIA-860 implied national coal and non-coal electricity production, $Y_{ct}^{EIA}$ and $Y_{nt}^{EIA}$, against that provided by the historical census.
Appendix Figures

Figure A.1: U.S. CO₂ emissions intensity in 1960 and 2000

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

Figure A.2: Country-level coal reserves

Figure A.3: Bituminous coal production and mechanization over time

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

Figure A.4: Sample counties and coal reserves

Notes: Counties are included in the sample (in yellow) if county centroid is 1) < 250 miles from nearest Illinois coal reserve, 2) closer to nearest Illinois coal reserve than to nearest Appalachian coal reserve.
Figure A.5: Comparing census and reconstructed steam and hydro capacity

Notes: Panel (A) compares aggregate U.S. steam capacity from 1920-1970 reported by historical census (solid blue) against that reported by reconstructed panel using EIA-860 (dashed red). Panel (B) compares percentage change in aggregate U.S. steam-capacity reported by historical census (solid blue) against that reported by reconstructed panel using EIA-860 (dashed red). Panels (C) and (D) the same as Panels (A) and (B) respectively for aggregate U.S. hydro capacity.

Figure A.6: Comparing census and reconstructed coal-fed and other-fed production

Notes: Panel (A) compares aggregate U.S. coal-fired steam electricity production from 1920-1970 reported by historical census (solid blue) against that reported by reconstructed panel using EIA-860 (dashed red). Panel (B) compares percentage change in aggregate U.S. coal-fired electricity production reported by historical census (solid blue) against that reported by reconstructed panel using EIA-860 (dashed red). Panels (C) and (D) the same as Panels (A) and (B) respectively for aggregate U.S. steam electricity production from other inputs. See calculation of capacity factors in Section Appendix B.
Figure A.7: 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 1970s in the “Inventory of Power Plants in the United States”

Figure A.8: County distance to obsolete and subsequent nearest mine

Notes: Scatter shows the county distance to the nearest deep coal mine after obsolescence against the distance to the obsolete coal mine.
Appendix Tables

Table A.1: Data consistency across EIA-860 forms (1990-2012)

<table>
<thead>
<tr>
<th>Number of different values</th>
<th>Percent of generators with different reported values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary fuel</td>
<td>Capacity</td>
</tr>
<tr>
<td>0</td>
<td>94.25</td>
</tr>
<tr>
<td>1</td>
<td>1.49</td>
</tr>
<tr>
<td>2</td>
<td>0.57</td>
</tr>
<tr>
<td>3</td>
<td>0.23</td>
</tr>
<tr>
<td>4</td>
<td>0.2</td>
</tr>
<tr>
<td>5</td>
<td>0.51</td>
</tr>
<tr>
<td>6</td>
<td>0.46</td>
</tr>
<tr>
<td>7</td>
<td>0.29</td>
</tr>
<tr>
<td>8</td>
<td>0.27</td>
</tr>
<tr>
<td>9</td>
<td>0.2</td>
</tr>
<tr>
<td>10</td>
<td>0.25</td>
</tr>
</tbody>
</table>

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.2: Consistency in reported primary fuel between late 1970s and 2012 EIA-860

<table>
<thead>
<tr>
<th>Primary fuel in 1970s</th>
<th>Primary fuel in 2012</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coal</td>
</tr>
<tr>
<td>Coal</td>
<td>92.2</td>
</tr>
<tr>
<td>Hydro</td>
<td>0.0</td>
</tr>
<tr>
<td>Nat. gas</td>
<td>0.8</td>
</tr>
<tr>
<td>Nuclear</td>
<td>0.0</td>
</tr>
<tr>
<td>Oil</td>
<td>1.0</td>
</tr>
</tbody>
</table>

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.