Proximity to coal mines and mortality rates in the Appalachian Region of the United States: a spatial econometric analysis

Sriparna Ghosh and Richard J. Cebula

ABSTRACT
Over the last several years, a body of the scholarly literature has emerged analysing the effects of proximity to coal mines and various human health indicators in the United States. These studies have found evidence of an increased poor health status associated with the production of coal and with close proximity to coal mines. The present study contributes to this literature by analysing the spatial effects of coal production on neighbouring counties within the entire 420-county, 13-state Appalachian Region in terms of increased rates of death attributable to respiratory diseases.

ARTICLE HISTORY
Received 22 June 2020; Accepted 5 March 2021

KEYWORD
spatial panel data model; regional analysis of coal mine production; analysis of health and proximity to coal mines

JEL
C12; I19; Q53; R12

INTRODUCTION
The emergence of market capitalism has been interpreted as being a robust cause of the Industrial Revolution. The latter, in turn, has been shown to have created historically unprecedented economic growth and prosperity (Cebula et al., 2015, ch. 1; Landes, 1998; North & Thomas, 1973). However, the face of capitalism is not, and indeed has certainly not always been, a pretty one, even in contemporary societies, despite various and seemingly extensive forms of well-intended regulation and legislation.

Consider, for instance, the case of coal mining. Although leading to economic profits and dividends for owners and stockholders and presumably helping to restrain the growth rate of energy prices, the environmental externalities associated with coal mining, in all its various forms, generate an increase in morbidity as well as a host of serious forms of human health costs along with, given the low-wage structure typically found within the industry, serious inter-regional income inequalities and poverty (Richmond-Bryant et al., 2020). These human health...
costs, low-wage/poverty per se considerations aside, involve myriad profoundly health-compro-
mising dimensions for humans with a history of working in the mines as well as for those resid-
ning in closer proximity to coal-mining activities.

Based effectively upon considerations such as those referred to above, a body of scholarly lit-
erature has emerged in recent years seeking to analyse the effect of proximity to coal mines upon
human health indicators, including such conditions as chronic obstructive pulmonary disease
(COPD), disabling pneumoconiosis, forced vital lung capacity (FVC), cardiac disease and a var-
iety of birth defects, as well as elevated morbidity per se (e.g., Artfield, 1985, 1992; Borak et al.,
2012; Carta et al., 1996; Coggon & Taylor, 1998; Hendryx & Ahern, 2008; Hendryx & Ahern,
2012; Lapp et al., 1994; Marine et al., 1988). These studies have found evidence of reduced life
expectancy and poor health status being associated with the production of coal. Most of these
studies use regression analysis to measure the impacts of coal mines on health outcomes (e.g.,
Zullig & Hendryx, 2011).

However, to date, the existing literature, with only one apparent exception to date (Ahern
et al., 2011, where only a very limited number of contiguous counties in central Appalachia
are considered), overlooks the possibility that coal production within a given county may
adversely affect the health status or mortality of the population residing in neighbouring coun-
ties as well. This shortcoming of the related literature is problematic precisely because it over-
looks these spillovers associated with coal production. For example, as a practical matter, the coal
dust resulting from coal production and related activities is carried in varying degrees and in
varying ways through the atmosphere, water tables, and the like, for example, by the wind
and the rain as well as on trucks and other vehicles and on clothing, as well as on coal miners
and other employees themselves, from coal-mining operation sites to other surrounding geo-
graphical areas. Hence, it is logical to infer that adverse health impacts imposed by coal dust
can be similarly transferred and be classified as spillovers. A failure to allow for such circum-
stances (spillovers) leads to miscalculations, misleading inferences and poorly informed policy-
making decisions.

Accordingly, the present study seeks to fill this lacuna in the literature in terms of capturing the
spatial effects of coal production on the health outcomes of neighbouring counties and to do so for
the entirety of the Appalachian Region of the United States. As observed above, the study by
Ahern et al. (2011) is the one of the very few to investigate the spatial correlation between moun-
taintop coal mining and health, in this case in the form of birth defects. However, that study
focuses solely on the ‘central’ Appalachian states, that is, Kentucky, Tennessee, Virginia and
West Virginia. The present study contributes to this literature by analysing the spatial effects
of coal production on neighbouring counties within the entire 420-county, 13-state Appalachian
Region (see Figures A1 and A2 in in Appendix A in the supplemental data online) in terms of
increased rates of death attributable to respiratory diseases. The motivation of this study is in
part associated with the fact that the undertaking of coal-mining activity generates environmental
externalities that are not fully reflected in market prices but are reflected in augmented human
health costs, expressed in the present study in terms of increased mortality rates per se. We
seek to provide more accurate estimates of the latter phenomenon than has heretofore been avail-
able. Arguably, our findings can be relevant to the formulation of coal-mining policies.

LITERATURE REVIEW

Over the years, numerous studies have empirically investigated the linkage of various measures
of compromised health and coal mining. It is commonplace for such studies to find that coal-
mining activities and coal dust exposure adversely have impacted upon health among the popu-
lation (e.g., Artfield, 1985, 1992; Carta et al., 1996; Coggon & Taylor, 1998; Hendryx & Ahern,
2008; Lapp et al., 1994; Marine et al., 1988; Oxman et al., 1993; Zullig & Hendryx,
More specifically, numerous early studies documented significantly poorer health conditions across multiple health indicators among people living in mining versus those living in non-mining areas.

For instance, Carta et al. (1996) study the role of coal dust exposure on the incidence of respiratory symptoms and the decline of lung function in young coal miners. Their results show that even a pattern of moderate exposure to mixed coal dust significantly affects the lung function and incidence of symptoms of underground miners deleteriously. Moreover, in logistic models, they determined that coal dust exposure was also found to be a significant predictor of the onset of respiratory symptoms, along with age and smoking.

Indeed, such effects go beyond the general population of miners and ex-miners and even extend to newborn infants. For example, in Ahern et al. (2011), birth defects are examined in mountaintop coal-mining areas and compared with other coal-mining areas and non-coal-mining areas at the county level for four states in Central Appalachia, namely, Kentucky, Tennessee, Virginia and West Virginia. The hypothesis tested by Ahern et al. is that higher birth-defect rates are present in mountaintop mining areas. National Center for Health Statistics natality files were used to analyse 1996–2003 live births in four Central Appalachian states (N = 1,889,071). Poisson regression models that control for a number of covariates compare birth defect prevalence rates associated with maternal residence in mountaintop mining areas, other mining areas and non-mining areas. In a finding that is compatible in principle with a majority of previous related published research, Ahern et al. infer that the prevalence rate ratio for nearly every form categorized birth defect was found to be significantly higher in mountaintop mining areas as compared with non-mining areas, even after controlling for a host of covariates. In still greater detail, it was found that prevalence rate ratios were significantly higher in mountaintop mining areas for six out of seven types of defects: circulatory/respiratory, central nervous system, musculoskeletal, gastrointestinal, urogenital and ‘other’. Moreover, there was also preliminary evidence suggesting that spatial correlation between mountaintop mining and birth defects was present, implying in turn that there can be impacts (spillovers) from mountaintop mining in a focal county on birth defects in neighbouring counties.

Shifting focus from health issues such as those identified above to mortality rates/probabilities per se, there is a published literature linking higher morbidity rates to coal dust exposure, albeit a literature accompanied by at least some degree of controversy. To illustrate the latter phenomenon, consider, for example, Borak et al. (2012), who seek to determine the predictive value of coal mining and other risk factors for explaining disproportionately high mortality rates within Appalachia. Mortality and covariate data were reported to have been obtained from publicly available databases for 2000–04. The study undertook ordinary least squares (OLS) regression with age-adjusted mortality as the dependent variable. Among other things, this study found that age-adjusted all-cause mortality was related to the percentage poverty rate (positively), median household income (negatively), the percentage of the population with a high-school diploma (negatively) and obesity (positively), but not with the unemployment rate, the per cent of the population having health insurance, the per cent of the population with a four-year degree, the availability of physicians (per capita), smoking, diabetes or even coal mining. The lattermost finding runs against the findings in effectively all the related literature, whereas Borak et al. emphasize instead the ramifications of substantial economic and cultural disadvantages that adversely impact health in Appalachia, especially in the more coal-mining-intensive areas of Central Appalachia.

Moreover, the findings of Borak et al. (2012) seem very much at odds not only with previous studies (e.g., Landen et al., 2011; Liddell, 1973; Miller & Jacobsen, 1985; Rockette, 1977) but also with subsequent studies (e.g., Laney and Weissman, 2014; Reynolds et al., 2017). Interestingly, the findings of Borak et al. (2012) were seriously questioned in a ‘Reply’ crafted by Hendryx
and Ahern (2012). Hendryx and Ahern (2012) state that they find significant effects of coal mining on population mortality that are robust to multiple model specifications, including the specification adopted by and estimated by Borak et al. (2012). Indeed, this disparity prompted a challenge by Hendryx and Ahern (2012) to Borak et al. (2012) to openly demonstrate the validity of their (arguably very ‘atypical’) findings in terms of coal-mining activities per se.

To appreciate more fully the position/perspective expressed by Hendryx and Ahern (2012), one could also consider Landen et al. (2011), who begin with the premise that background particulate exposure from air pollution has been found to elevate the risk of ischaemic heart disease (IHD) mortality. Landen et al. (2011, p. 727) find that after having adjusted for age, smoking and body mass index, the risk of IHD mortality increased at higher levels of coal dust exposure. Other research yields very similar, that is, parallel results to those in Ahern et al. (2011) and Landen et al. (2011), as evidenced by, for example, by Brabin et al.’s (1994) study of school children who have been exposed to coal dust and Reynolds et al. (2017), who look at coal workers’ pneumoconiosis (CWP) across the nation. The study by Reynolds et al. (2017, p. 513) is pertinent if not compelling because it observes that ‘Coal is mined in approximately half of all U.S. states …’. Reynolds et al. emphasize the magnitude of the miner population, and hence of the population at risk in the coal-mining industry when observing that ‘miners working outside central Appalachia account for 57.1% of the country’s 65,000 coal miners’ (p. 513).

Moreover, the concern over the health effects and ultimately, the morbidity effects (end-stage disease effects) of coal dust exposure continues on many fronts. As observed by Laney and Weissman (2014, p. S18):

coal mine dust causes a spectrum of lung diseases collectively termed coal mine dust lung disease (CMDLD). These include Coal Workers’ Pneumoconiosis [CWP], silicosis, mixed dust pneumoconiosis, dust-related diffuse fibrosis … and chronic obstructive pulmonary disease. CMDLD continues to be a problem in the United States, particularly in the central Appalachian region. Treatment of CMDLD is symptomatic. Those with end-stage disease are candidates for lung transplantation. Because CMDLD cannot be cured, prevention is critical.

The present study is ultimately based in the aforementioned scholarly literature. However, it seeks to proceed beyond the scope and method of that literature, in part reflecting the preliminary evidence referred to above by Ahern et al. (2011) that suggests the possible presence of spatial correlation between coal mining and human health. The continuing serious concern regarding coal dust effects in terms not only of several health markers but also morbidity as such provides motivation for this study. Moreover, the nation’s continued reliance on coal mining reinforces that motivation.

The current literature focuses on the detrimental effect of coal dust on health outcomes (Laney & Weissman, 2014; Ahern et al., 2011; Landen et al., 2011; Reynolds et al., 2017). However, according to our knowledge, none of the existing literature has examined in a rigorous econometric context the spatial interactions of presence of coal mines on the deaths in neighbouring states. The toxins and impurities resulted from the coal by-products create air pollution, often causing cardiovascular and pulmonary diseases. The health hazards are not created by occupational exposure to coal, but simply living close to coal mines might result in sever health damage and deaths. Thus, this paper’s primary contribution is to examine whether spatial dependence in the deaths is due to proximity to coal mines. Our hypothesis is that once spatial dependence is taken into account, we will find the stronger effect of coal mines on mortality rates. Indeed, our results show that while taking into consideration the spillover effect of proximity to coal mines, the effect on the mortality rate is much higher. This research indicates that the spatial correlation between the coal mines and mortality rate is effective from a policy perspective to limit coal dust in the Appalachian Region.
BACKGROUND

The Appalachian Region, as defined by the Appalachian Regional Commission (ARC), consists of 420 contiguous counties in 13 states across the East from New York to Mississippi. According to a recent report by the ARC, approximately 25 million people reside within the Appalachian Region. The residents of the Appalachian Region not only face poor socioeconomic conditions but also (as summarized in the literature review provided above) suffer disproportionately from poor health and potentially increased risks of adverse health outcomes, including: (1) an increase in the mortality rate among persons aged 35–64 years; (2) a potentially greater likelihood of heart disease; and (3) a greater incidence of chronic pulmonary disease, as compared with the most of the rest of the nation (Laney & Weissman, 2014; Reynolds et al., 2017). Additionally, the population of the Appalachian Region suffers higher rates of total and premature mortality due to obesity and diabetes. In the view of many epidemiologists and public health researchers, Appalachia is characterized by an ‘increased chronic disease burden, limited access to health care, and elevated rates of behavioural risks’ (Borak et al., 2012, p. 1). Zullig and Hendryx (2011) find that people residing in the mountaintop mining areas of the Appalachian Regions suffer from poor mental and physical health conditions (measured by self-reported indicators). As stressed earlier, Ahern et al. (2011) find that there is a spatial correlation between mountaintop coal mining in counties and birth defects in the neighbouring counties. The present study contributes to the literature in two ways. First, it investigates the spatial spillover of coal production on death rates in the neighbouring countries in the Appalachian Region. Second, it also considers all the counties within the entire 13-state Appalachian Region.

DATA

In terms of health outcomes, several datasets are used to examine the effect of coal mine dust on human health. County-level deaths from respiratory disease have been obtained from the Centers for Disease Control and Prevention (CDC). The data for the total number of deaths per county per year were based on death certificates filed in the 50 states and the District of Columbia. County-level covariates such as total population, income and unemployment rates were obtained from the US Census and the Area Health Resources Files. Data for coal production were obtained from the Annual Coal Report and reported in thousands of tons. Descriptive statistics for these variables are shown in Table 1. The number of observations (N) across the entire 420 contiguous counties across 13 states is 2616. The average annual number of deaths in a county is 70, with a minimum of 1 and a maximum of 1472. The average amount of coal production on a county basis is 5,147,080 tons. The mean unemployment rate at the county level is 6.97%; the mean income per capita is US$35,624.14. For perspective, the national mean unemployment rate over the same period of the present study was only 5.63% (Council of Economic Advisors, 2017, tab. B-12), whereas the mean for per capita income nationally was a much higher at US$50,151 (Council of Economic Advisors, 2013, tab. B-31).

Table 1. Descriptive statistics, 2000–12.

| Variable              | Observations | Mean   | SD        | Minimum | Maximum |
|-----------------------|--------------|--------|-----------|---------|---------|
| Deaths                | 2616         | 70     | 116.5028  | 1       | 1472    |
| Population            | 2616         | 61,570 | 107,502   | 4743    | 1,281,666 |
| Income                | 2616         | 35,624.14 | 8443.92 | 960     | 77,830 |
| Total coal production | 2616         | 5147.08 | 25,258.75 | 170     | 415,924 |
| Unemployment rate     | 2616         | 6.97   | 2.79      | .1      | 23.3    |
EMPIRICAL ANALYSIS

The empirical analysis in this study involves the adoption of the spatial econometric approach. By using the spatial Durbin error model (SDEM), we are able to capture the local spillovers of coal production in a county on its neighbouring counties. The advantage of using the SDEM for this purpose is that it contains spatially lagged explanatory variables measuring the exogenous interactions between the explanatory variables and the error terms. In this case, the SDEM helps to measure the local spillover of coal production in any county on the mortality and health outcomes of the neighbouring counties. The SDEM, unlike other models, focuses on the spatial dependence among the explanatory variables. This particular feature of the SDEM is very useful in the present framework because it is arguable that the health hazards resulting from the operation of coal mines are most likely to be restricted to the closest neighbouring counties as compared with more distant ones. This empirical study contributes to the existing literature by considering the effect of spatial dependence of coal production on mortality and health outcomes.

It is hypothesized that independent and error variables have spatial effects. Accordingly, we employ the SDEM for this study. We begin by explaining a family of related spatial models and then the actual model choice, that is, the SDEM. The family of related spatial models can be represented, as follows:

\[
y_{it} = \rho \sum_{j=1}^{N} w_{ij} y_{jt} + x_{it} \beta + \theta \sum_{j=1}^{N} w_{ij} x_{jt} + \mu_i + \lambda_t + u_i \\
\]

\[
u_i = \delta \sum_{j=1}^{N} w_{ij} \nu_{jt} + \epsilon_{it}
\]

where \(i\) represents cross-sectional units. For the purposes of this study, \(i\) represents U. states and ranges from \(i = 1\) to \(N\); \(t\) represents the time dimension, that is, year, and ranges from \(t = 1\) to \(T\). Therefore, \(y_{it}\) represents an observation for the dependent variable in state \(i\) in year \(t\). \(x_{it}\), the explanatory variables, is a row vector of observations with dimension \((1 \times K)\). \(\beta\) is a \((K \times 1)\) vector of parameters associated with \(x_{it}\) variables, and is fixed and unknown. The terms \(\mu_i\) and \(\lambda_t\) represent spatial fixed effects and time fixed effects, respectively. The spatially lagged dependent variable and spatially lagged explanatory variables are represented by \(\rho\) and \(\theta\), whereas the spatially lagged error term is represented by \(\delta\). The addition of these terms in association with \(w_{ij}\) makes the above model a spatial econometric model. In the model, \(w_{ij}\) is an element of a spatial weighting matrix, \(W\). \(W^3\) symbolizes ‘neighbour-to-neighbour’ relationships and is of dimension \((N \times N)\).

For example, if \(i\) and \(j\) are identified as neighbouring counties, the \(w_{ij}\) element is assigned a value of 1, and 0 otherwise. In creating weighted matrices, \(W\) is designed to be row-stochastic, meaning that rows of \(W\) sum to 1. Hence, the term \(W_y\) represents the weighted average of the surrounding \(y\)’s. Similarly, \(W_x\) represents the weighted average of the surrounding explanatory variables; and \(W_u\) represents the surrounding error terms. Since equation (1) characterizes a family of spatial models, restricting parameters in the equation generates various specific spatial econometric models. For example, by restricting \(\theta\) and \(\delta\) to 0, we obtain spatial dependence only in the dependent variable. This type of model is a spatial autoregressive (SAR) model. By setting \(\rho = 0\) and \(\delta = 0\), we observe spatial dependence only with the independent variable. Such a model is named the spatial lag of \(X\) (SLX) model. Similarly, setting \(\rho = 0\) and \(\theta = 0\), we observe spatial dependence only on the error term; this model is named spatial error model (SEM). The spatial Durbin model (SDM) is obtained by setting \(\delta = 0\), and the spatial Durbin error model (SDEM)
is obtained by restricting \( \rho \) to 0. Spatial fixed effects (\( \mu_i \)) and time fixed effects (\( \lambda_t \)) in equation (1) may be included in all the models described above.

The meaning of average direct effects can be interpreted in a straightforward fashion. For example, if Coal Production of state \( i \) is changed, how does it affect the Mortality in state \( i \)? This is called \( i \)'s ‘own effects’. However, average direct effect also captures spillover effects, in which case the change in Coal Production in state \( i \) affects Coal Production in its neighbouring state \( j \), which, again, ‘feeds back’ to state \( i \). This is in fact called the ‘feedback effect’. The average of the diagonal elements of the \( S_r(W) \) matrix captures both own effect and the feedback effect and is referred to as the ‘average direct effect’.

In Table 2, Elhorst test findings are provided for determining the presence of space- and time-fixed effects (Elhorst, 2009). The Lagrange multiplier (LM) tests check whether there is spatial correlation in the data. The null hypothesis is given by \( H_0: \) No spatial dependence in the dependent variable, is revealed by the LM lag test, whereas, the null hypothesis, \( H_0: \) No spatial dependence in the error term, is revealed by the LM error test for each specification. Standard likelihood ratio (LR) tests are performed to determine the joint significance of space- and time-fixed effects (Elhorst, 2014). The null hypotheses for such tests regarding the presence of state fixed effects and year fixed effects are represented by the following:

\[
H_0: \mu_1, \mu_2, \mu_3, \mu_4, \ldots, \mu_n = 0
\]

(2)

\[
H_0: \lambda_1, \lambda_2, \lambda_3, \lambda_4, \ldots, \lambda_n = 0
\]

(3)

In view of the results shown in Table 2, the most appropriate model to adopt is one that includes only the time fixed effects.

In this case SDEM is the parsimonious model. (We have provided the details in the Appendix in the supplemental data online how we chose SDEM as the parsimonious model.) However, as there are no robust ways to estimate SDEM model, we use both the SDM and SEM to capture direct and indirect effects under the SDEM. In order to fully understand the effects of the explanatory variables in the SDM, it is imperative to understand how the \( \beta \)-coefficient is interpreted. In the SDM, unlike the case of a standard linear model, \( \beta \) not only represents the marginal effects, meaning an increase in \( \beta \) captures the explanatory variable changes and how it affects the dependent variable, but now also captures the average direct, average indirect and average total effects, which are referred to as ‘term effects estimates’. Since models with a spatially lagged dependent variable yield estimates that are difficult to interpret, the data-generating process in reduced form for such models (in this case, the SDM) can be mathematically

|                  | OLS  | Spatial FE | Time-period FE | Spatial and time-period FE | Spatial (joint significance) | Time-period (joint significance) |
|------------------|------|------------|----------------|---------------------------|-------------------------------|---------------------------------|
| LM lag (robust)  | 0.0050 | 0.0930     | 0.0166         | 0.6060                    |                               |                                 |
| LM error (robust)| 0.0010 | 0.0787     | 0.0062         | 0.7270                    |                               |                                 |
| LR test          |       | 0.000      | 0.0825         |                           |                               |                                 |

Note: FE, fixed effects; OLS, ordinary least squares.
written as:

\[ y = \rho W y + X \beta + WX \theta + \epsilon \]  

\[ y = (I_n - \rho W)^{-1}(X \beta + WX \theta) + (I_n - \rho W)^{-1} \epsilon \] 

\[ (I_n - \rho W)^{-1} = I_n + \rho W + \rho^2 W^2 + \cdots + \rho^q W^q \] 

\[ S_r(W) = \frac{\partial y}{\partial x} = (In - W)^{-1} (\beta + W \theta) \] 

where the \( r \) subscript in the term \( S_r(W) \) represents individual explanatory variables in the \( X \) matrix.

In the SDM framework, \( S_r(W) \) is intended to capture the average direct, average indirect and average total effects of the change in a variable on the dependent variable and can be mathematically shown as follows:

Average Direct Effect: \[ \frac{\partial E(y_i)}{\partial x_{ir}} = S_r(W)_{ii} \]  

Average Indirect Effect: \[ \frac{\partial E(y_i)}{\partial x_{ir}} = S_r(W)_{ij} \] 

These equations follow directly from equation (1). Average total effects are the sum of the average direct and average indirect effects.

The meaning of average direct effects can be interpreted in a straightforward fashion. For example, if \( \text{Coal Production} \) of state \( i \) is changed, how does it affect the \( \text{Mortality} \) in state \( i \)? This is called \( i \)'s 'own effects'. However, the average direct effect also captures spillover effects, in which the change in \( \text{Coal Production} \) in state \( i \) affects \( \text{Coal Production} \) in its neighbouring state \( j \), which, again, 'feeds back' to state \( i \). This is the 'feedback effect' alluded to above. The average of the diagonal elements of the \( S_r(W) \) matrix captures both own effect and the feedback effect and is called the average direct effect.

The average indirect effect captures spillovers effects of a change in an explanatory variable in state \( i \) and how that affects observations in its neighbouring state \( j \), where state \( i \) and \( j \) are not same. This effect is captured by the average of the off-diagonal elements of the \( S_r(W) \) matrix. Since the indirect effects are accumulated over all neighbours, its magnitude can actually be larger than the direct effects. Finally, the average total effect is the summation of the average direct effect and the average indirect effect.

**EMPIRICAL RESULTS**

The objective of this study is to investigate the spatial effects of coal production on the number of deaths in neighbouring counties in the Appalachian Region of the United States. The focus on deaths reflects the argument that the death rate encompasses the negative health effects of coal in the most comprehensive way, that is, more inclusively than any of the other health measures available, such as COPD, disabling pneumoconiosis, FVC, cardiac disease and a variety of birth defects. This is the underlying reasoning for previous studies' emphasis on deaths (Borak et al., 2012; Hendryx & Ahern, 2012; Landen et al., 2011; Liddell, 1973; Rockette, 1977). It would be potentially very useful for future research to address the issue at hand in term of these specific forms of health problems. We also include population size (Population), Unemployment Rate, and Per Capita Income as control variables. The analysis models coal-production related deaths due to respiratory diseases (Deaths) as a negative function of population.
size \((\text{Population})\) and \(\text{Income}\), and a positive function of \(\text{Total Coal Production}\) and the unemployment rate.

Table 3 shows the estimated coefficients from the SEM regression, Table 4 reports the estimated coefficients from the panel regression including county fixed effects and year fixed effects. Table 5 reports the coefficients from SDM panel estimation.

As shown in Table 5, the panel shows that with a 1 unit increase in coal production, the number of deaths increases by 0.02%. However, this does not include the spatial spillover effect from the neighbouring counties. Table 6 reveals at the 1% significance level that with a 1 (thousand) unit increase in coal production, the number of deaths increases by 1.24% (total effect) significantly.

After including the spatial correlation, the direct effect of coal production on death is higher as well. The SDM model shown in Table 5 reports that with a 1 (thousand) unit increase in coal production the direct effect on death is 1.1%. Similarly, with a 1 (thousand) unit increase in coal production in neighbouring counties, the spatial spillover on number of deaths is 0.12%, meaning with a 1 (thousand) unit increase in coal production in neighbouring counties, the death rate in a county will rise significantly by 0.12%. Comparing with the panel regression, we find that coal production (both direct and indirect) has higher adverse effect on number of deaths due to respiratory diseases. Furthermore, if income in a county increases by US$1, the death rate in that particular county falls significantly by 0.04%, whereas if income rises by US$1 in the neighbouring counties, the death rates significantly falls by 0.03% in a county. If population in a county increases by 1 unit, the death rate decreases by 0.05% significantly, whereas if the population increases in the neighbouring counties the death rate falls by 0.2% significantly. If unemployment increases by 1% the death increases significantly by 1.552% significantly whereas if unemployment increases in the neighbouring counties the death rates does not get affected significantly. Furthermore, if the death rate increases by 1%, the death rates in the neighbouring counties also significantly increases by 0.225% significantly. For robustness check, we repeat the regressions with a \(W\) matrix with five neighbouring counties (see the Appendix in the supplemental data online) and find similar results.

Table 6 reports the results of coal production on mortality rates. It consists of the average direct, average indirect and average total effect estimates from the SDM panel regression. One of the noteworthy findings is that \(\rho\) is statistically significant at the 1% level, indicating that

**Table 3. Spatial error model (SEM) estimation results.**

| Variable         | SEM          |
|------------------|--------------|
| Unemployment rate| 1.591***     |
|                  | (0.001)      |
| Coal production  | 0.0011***    |
|                  | (0.002)      |
| Income           | 0.0043***    |
|                  | (0.031)      |
| Total population | -0.0003**    |
|                  | (0.001)      |
| Year fixed effects| Yes        |
| County fixed effects| Yes     |
| Observations     | 2616         |
| \(R^2\)          | 0.9503       |
| Prob > Wald Chi-square | 0.00       |

Notes: The dependent variable is death due to respiratory diseases. Robust standard errors are shown in parentheses. 
***\(p < 0.01\), **\(p < 0.05\), *\(p < 0.1\).
### Table 4. Panel estimation results.

| Variable               | Panel regression |
|------------------------|------------------|
| Total coal production  | 0.002*           |
|                        | (0.085)          |
| Unemployment rate      | 1.14***          |
|                        | (0.001)          |
| Income                 | 1.100***         |
|                        | (0.001)          |
| Total population       | − 0.0001*        |
|                        | (0.011)          |
| Year fixed effects     | Yes              |
| County fixed effects   | Yes              |
| Observations           | 2616             |
| $R^2$                  | 0.6              |
| Prob > F               | 0.00             |

Notes: The dependent variable is death due to respiratory diseases. Robust standard errors are shown in parentheses.

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.***

### Table 5. Spatial Durbin model (SDM) estimation results.

| Variable               | SDM               |
|------------------------|-------------------|
| Unemployment rate      | 1.552**           |
|                        | (0.008)           |
| Coal production        | 0.011***          |
|                        | (0.012)           |
| Income                 | −0.0004**         |
|                        | (0.011)           |
| Total population       | −0.0005**         |
|                        | (0.001)           |
| W*Unemployment rate    | 2.061             |
|                        | (0.003)           |
| W*Coal production      | 0.0012***         |
|                        | (0.006)           |
| W*Income               | −0.0003***        |
|                        | (0.002)           |
| W*Total population     | −0.002***         |
|                        | (0.112)           |
| W*Deaths               | 0.225**           |
|                        | (0.012)           |
| Year fixed effects     | Yes               |
| County fixed effects   | Yes               |
| Observations           | 2616              |
| $R^2$                  | 0.06              |
| Prob > F               | 0.00              |

Notes: The dependent variable is death due to respiratory diseases. The W matrix is constructed with four neighbouring counties. Robust standard errors are shown in parentheses.

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.***
the mortality rates in the Appalachian Region are spatially correlated. Table 6 shows that increasing coal production exercises both positive average direct and indirect effects on the mortality rate. Table 6 also shows that with a 1% increase in the unemployment rate, the number of deaths directly increases by 1.553%. However, the indirect effect of unemployment is insignificant implying that increases in unemployment do not affect the mortality rates in neighbouring counties. The unemployment rate has positive direct effect, suggesting that higher unemployment is related to higher mortality in a state. A 1 unit increase in income in a county reduces the number of deaths by 0.04% (direct effect), while increases in income in a county also decrease the mortality by 0.03% in the neighbouring counties (significance is at the 1% level). Higher income might lead to healthy lifestyle choices and better development of the county, which might affect the living standards of the neighbouring counties due to a greater availability of jobs and associated spending multiplier effects. We find that as the population size in a county increases, the number of deaths falls significantly not only in the county but also in the neighbouring counties. This might be due to the fact that a higher population implies a greater availability of medical care facilities, which would tend to reduce the mortality rates. Hendryx and Ahern (2008) found that proximity to coal mines is associated with poor health status in West Virginia, measured by higher rates of cardiopulmonary disease, COPD, hypertension, lung disease and kidney disease. However, this paper is an improvement over the aforementioned study as it considers the spillover effect of coal production on mortality rates in the entire Appalachian Region.

### CONCLUSIONS

The existing body of scholarly studies regarding the impact of coal production on deaths are very useful insofar as they funnel attention to a critical public health issue. For the most part, such studies for the United States take into consideration only the simple correlation between coal production and death rates. Ahern et al. (2011) is one of the very few studies that considers the spatial correlation between coal production and the human health effects thereof, albeit for solely the case of birth defects in certain (although not all) counties of the Appalachian Region. The present study contributes to the literature in terms of analysing the spatial effects of coal production on deaths due to respiratory diseases throughout the counties of the entire Appalachian Region. In addition, the main findings of this study address the fact that full impact in a given county of coal production on mortality rates is not limited to the particular coal-producing county but rather the fact that the amount of coal produced also has measurable spillover effects on the mortality rates of the neighbouring counties. Our findings imply that the hazardous health effects of coal production do lead to an increased death rate among the residents of the neighbouring counties as well. Furthermore, for reasons provided above, it may well be that the residents of the counties producing coal in Appalachian Region also exhibit poor lifestyle choices of residents leading to higher death rates.

| Variable                | Direct effect | Indirect effect | Total effects |
|-------------------------|---------------|-----------------|---------------|
| Unemployment rate       | 1.553**       | 2.057           | 3.61          |
| Coal production         | 0.011***      | 0.0013***       | 0.0124***     |
| Income                  | −0.0004**     | −0.0003***      | −0.0008**     |
| Total population        | −0.0007**     | −0.003***       | −0.0037**     |

Notes: Robust standard errors are shown in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1.
One of the principal concerns of this study is the set of endogenous regressors. Future research should focus on identifying possible endogeneity and accordingly including appropriate instruments to address that endogeneity should it be necessary to do. Moreover, future research on the topic of this study might focus upon expanding the model to include additional explanatory variables as well. Finally, the issues considered in this study may provide motivation for future research efforts to be directed at circumstances in other nations (i.e., nations other than the United States) that have extensive coal mining (or even other types of mining) industries that generate externalities that jeopardize human health. Arguably, in any event, our findings can be relevant to the formulation of public policies towards coal mining by adding depth to the cost–benefit tools of policymakers dealing with the environment, public health, and regulatory efficiency.

NOTES

1 The Appalachian Region, covering 205,000 square miles, was initially divided into ARC 360 counties as a result of periodic acts of Congress. It was increased to 399 counties in 1991, 406 counties in 1998, 410 counties in 2002 and, finally, 420 counties in 2008.

2 Data are from the Multiple Cause of Death Files, 1999–2017, as compiled from data provided by the 57 vital statistics jurisdictions through the Vital Statistics Cooperative Program (http://wonder.cdc.gov/ucd-icd10.html, accessed June 1, 2019).

3 We use the $k$-nearest-neighbour $W$ matrix in this paper. We used the four neighbouring counties in this study. The neighbouring counties were picked such that the counties touch borders.

4 The results reflect the spatial correlation of coal production on mortality rates across all US counties with a special emphasis on the Appalachian Region.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the authors.

ORCID

Sriparna Ghosh © http://orcid.org/0000-0001-8059-3210

REFERENCES

Ahern, M. M., Hendryx, M., Conley, J., Fedorko, E., Ducatman, A., & Zullig, K. J. (2011). The association between mountaintop mining and birth defects among live births in central Appalachia, 1996–2003. *Environmental Research, 111*(6), 838–846. https://doi.org/10.1016/j.envres.2011.05.019

Artfield, M. D. (1985). Longitudinal decline in FEV in United States coal miners. *Thorax, 40*(1), 182–189. https://doi.org/10.1136/thx.40.2.132

Artfield, M. D. (1992). Pulmonary function of U.S. Coal miners related to dust exposure estimates. *American Review of Respiratory Disease, 145*(3), 605–609. https://doi.org/10.1164/ujrccm/145.3.605

Borak, J., Salipante-Zaidel, C., Slade, M. D., & Fields, C. A. (2012). Mortality disparities in Appalachia: Reassessment of major risk factors. *Journal of Occupational and Environmental Medicine, 54*(2), 146–156. https://doi.org/10.1097/JOM.0b013e318246f395

Brabin, B., Smith, M., Milligan, P., Benjamin, C., Dunne, E., & Pearson, M. (1994). Respiratory morbidity in Merseyside schoolchildren exposed to coal dust and air pollution. *Archives of Disease in Childhood, 70*(4), 305–312. https://doi.org/10.1136/adc.70.4.305
Carta, P., Aru, G., Barbieri, M. T., Avataneo, G., & Casula, D. (1996). Dust exposure, respiratory symptoms, and longitudinal decline of lung function in young coal miners. *Occupational and Environmental Medicine, 53*(2), 312–319. https://doi.org/10.1136/oem.53.5.312

Cebula, R. J., Hall, J. C., Mixon, F. G., & Payne, J. E. (2015). Economic behavior, economic freedom, and entrepreneurship. Edward Elgar Publishing.

Coggon, D., & Taylor, A. N. (1998). Coal mining and chronic obstructive pulmonary disease: A review of the evidence. *Thorax, 53*(4), 398–407. https://doi.org/10.1136/thx.53.5.398

Council of Economic Advisors. (2013). Economic report of the president, 2013. U.S. Government Printing Office.

Council of Economic Advisors. (2017). Economic report of the president, 2017. U.S. Government Printing Office.

Elhorst, J. P. (2009). Spatial panel data models. In M. M. Fischer & A. Getis (Eds.), *Handbook of applied spatial analysis*, Chapter 2 (pp. 81–103). Springer.

Elhorst, J. P. (2014). Spatial panel data models. In *Spatial econometrics: From cross-sectional data to spatial panels*, Chapter 3 (pp. 37–93). Springer.

Hendryx, M., & Ahern, M. (2012). Reply to Borak et al ‘mortality disparities in Appalachia: Reassessment of major risk factors. *Journal of Occupational and Environmental Medicine, 54*(7), 768–769. https://doi.org/10.1097/JOM.0b013e318254622f

Hendryx, M., & Ahern, M. M. (2008). Relations between health indicators and residential proximity to coal mining in West Virginia. *American Journal of Public Health, 98*(4), 669–671. https://doi.org/10.2105/AJPH.2007.113472

Landen, D. D., McWilliams, L., Patel, A., & Wassell, J. T. (2011). Coal dust exposure and mortality from ischemic heart disease among a cohort of U.S. coal miners. *American Journal of Industrial Medicine, 54*(10), 727–733. https://doi.org/10.1002/ajim.20986

Landes, D. S. (1998). *The wealth and poverty of nations: Why some are so rich and some so poor*. W.W. Norton.

Laney, A. S., & Weissman, D. N. (2014). Respiratory diseases caused by coal mine dust. *Journal of Occupational and Environmental Medicine, 56*(10), S18–S22. https://doi.org/10.1097/JOM.0b013e318254622f

Lapp, N. L., Morgan, W. K., & Zaldivar, G. (1994). Airways obstruction, coal mining, and disability. *Occupational and Environmental Medicine, 51*(4), 234–238. https://doi.org/10.1136/oem.51.4.234

Liddell, F. D. (1973). Mortality of British coal miners in 1961. *British Journal of Industrial Medicine, 40*(11), 25–24. https://doi.org/10.1136/oem.40.11.15

Marine, W. M., Gurr, D., & Jacobsen, M. (1988). Clinically important respiratory effects of dust exposure and smoking in British coal miners. *American Review of Respiratory Disease, 137*(1), 106–112. https://doi.org/10.1164/ajrccm/137.1.106

Miller, B. G., & Jacobsen, M. (1985). Dust exposure, pneumoconiosis, and mortality of coal miners. *British Journal of Industrial Medicine, 42*(11), 723–733. https://doi.org/10.1136/oem.42.11.723

North, D. C., & Thomas, R. (1973). *The rise of the western world: A new economic history*. Cambridge University Press.

Oxman, A. D., Muir, D. C., Shannon, H. S., Stock, S. R., Hnizdo, E., & Lange, H. J. (1993). Occupational dust exposure and chronic obstructive pulmonary disease: A systematic overview of the evidence. *American Review of Respiratory Disease, 148*(1), 38–48. https://doi.org/10.1164/ajrccm/148.1.38

Reynolds, L. E., Blackley, D. J., Laney, A. S., & Halldin, C. N. (2017). Respirator morbidity among U.S. coal miners in states outside of central Appalachia. *American Journal of Industrial Medicine, 60*(6), 513–517. https://doi.org/10.1002/ajim.22727

Richmond-Bryant, J., Mikati, I., Benson, A. F., Luben, T. J., & Sacks, J. D. (2020). Disparities in distribution of particulate matter emissions from US coal-fired power plants by race and poverty status after accounting for reductions in operations between 2015 and 2017. *American Journal of Public Health, 110*(5), 655–661

Rockette, H. E. (1977). Cause specific mortality of coal miners. *Journal of Occupational and Environmental Medicine, 19*(12), 795–801. https://doi.org/10.1097/00043764-197712000-00004

Zullig, K. J., & Hendryx, M. (2011). Health-related quality of life among central Appalachian residents in mountaintop mining counties. *American Journal of Public Health, 101*(5), 848–853. https://doi.org/10.2105/AJPH.2010.300073