Opioid Mortality in the US: Quantifying the Direct and Indirect Impact of Sociodemographic and Socioeconomic Factors

Sucharita Gopal[1] , Manfred M. Fischer*[2]

Abstract: This paper employs a spatial Durbin panel data model, an extension of the cross-sectional spatial Durbin model to a panel data framework, to quantify the impact of a set of sociodemographic and socioeconomic factors that influence opioid-related mortality in the US. The empirical model uses a pool of 49 US states over six years from 2014 to 2019, and a nearest-neighbor matrix that represents the topological structure between the states. Calculation of direct (own-state) and indirect (cross-state spillovers) effects estimates is based on Bayesian estimation and inference reflecting a proper interpretation of the marginal effects for the model that involves spatial lags of the dependent and independent variables. The study provides evidence that opioid mortality depends not only on the characteristics of the state itself (direct effects), but also on those of nearby states (indirect effects). Direct effects are important, but externalities (spatial spillovers) are more important. The sociodemographic structure (age and race) of a state is important whereas economic distress of a state is less so, as indicated by the total impact estimates. The methodology and the research findings provide a useful template for future empirical work using other geographic locations or shifting interest to other epidemics.

June 12, 2023

*Corresponding author: Manfred M. Fischer, Vienna University of Economics and Business, Welthandelsplatz 1, A-1020 Vienna, Austria; email: manfred.fischer@wu.ac.at; orcid: 0000-0002-0033-2510
[1] Department of Earth and Environment, Boston University, [2] Department of SocioEconomics, Vienna University of Economics and Business

1 Introduction

Opioid-induced deaths (hereafter, briefly opioid mortality)[1] increased sharply in the United States (US) during the past three decades, and opioid overdose mortality has become an important, yet inadequately understood public health problem (Kopel 2019). Although opioid mortality is a spatial characteristic of a population in a specific area (Cossman et al. 2007), only few studies have used spatial econometric methods to control for spatial dependence[2] in regression models that aim to identify the impact of determinants on the outcome of the opioid mortality process (see Cushing et al. 2020). Sparks and Sparks (2010) apply spatial autoregressive models to the study of US county mortality rates, and Yang et al. (2015) explore geographic variation in US mortality rates using a conventional spatial Durbin model. These studies, however, are cross-sectional in nature on the one side and do not contribute to the opioid mortality literature on the other.

To overcome this shortcoming in mortality research in general and opioid mortality research in particular, we use panel rather than cross-sectional data to study the relationship between opioid mortality rates and explanatory variables from a longitudinal perspective that accounts for temporal and spatial dependence simultaneously. Omitting the spatial dependence issue would increase the risk of obtaining biased estimation results, while the use of panel data results in a more accurate inference of model parameters (Baltagi et al. 2007). Panel data is generally more informative, contains more degrees of freedom and less multicollinearity between variables than cross-sectional data, and improves the efficiency of econometric estimates (see Elhorst 2021). It is more informative since providing information on both the intertemporal dynamics and the individuality of the observational units (regions) and allowing one to control the effects of missing and unobserved variables (for more details, see Hsiao 2007).

The primary objective of this study is to employ a spatial Durbin panel data model,[3] an extension

---

[1]Opioids include natural opioids such as morphine and codeine, semisynthetic opioids such as oxycodone, hydrocodone, hydromorphone and oxymorphone, and synthetic opioids such as methadone, tramadole and fentanyl. Heroine is an illicit opioid synthesized from morphine.

[2]Spatial dependence among observational units in the US is well documented in Rossen et al. (2014), Buchanich et al. (2016), and Stewart et al. (2017), among others.

[3]A spatial Durbin rather than a spatial Durbin error panel data model specification is used because it allows for quantification of spatial spillovers, the focus of interest in this study. Note that a spatial Durbin error specification would rule out spatial spillovers a priori, allowing for spatial impacts arising from dependence in the disturbances but not from situations where changes in a variable from observation unit j impact the outcome of the dependent variable in another unit i, called spatial spillover from j to i.

of the cross-sectional model to a panel data framework, for estimating direct and indirect (cross-state spillover) effects of a set of factors that may influence opioid mortality rates across space and over time. In this model specification, the characteristics of a specific observation unit (in our case: state) and its neighbors are simultaneously considered. To the best of our knowledge, no opioid mortality research has considered spatial spillovers among observational units. Specifically, we consider five variables to capture the racial/ethnic structure of a state, five variables to cover the state's age structure, two variables to represent the educational attainment level of its population, and three covariates that address economic distress. In addition to the sociodemographic and socioeconomic variables, we include the opioid prescription rate to account for one of the opioid supply factors, and health insurance coverage that may have the potential to increase the demand for drugs.

Our choice of using states rather than counties as observation units in this study was forced by data availability.[4] The Annual American Community Survey (ACS) census data is available starting 2014 for the demographic variables included in the study. Hence our study used the period 2014-2019. The model is applied to 49 states over six years, from 2014 to 2019. A k-nearest-neighbor matrix is chosen to represent the topological structure between the states.[5] The calculation of effects estimates is based on Bayesian estimation and inference that reflects a proper interpretation of the marginal effects for our nonlinear model that involves lags of the dependent variable vector. The study provides a rich picture of how sociodemographic and socioeconomic determinants affect opioid mortality. The findings highlight the critical role indirect (spillover) effects play. A neglection of the spillover effects would cause incorrect inferences.

The remainder of the paper is structured as follows. Section 2 presents the econometric framework, outlines the proposed model, discusses prior selection for Bayesian estimation of the model, and shows how to calculate direct and indirect effects estimates necessary for a proper interpretation of the model estimates.[6] Section 3 outlines model specification issues, provides a brief summary

---

[4]Using counties as observation units would be preferable but county-level data is unavailable for all the states in the US for even one year, due to HIPAA (Health Insurance Portability and Accountability Act) privacy rules in some counties. The missing data (especially in mid-western and western states) forced us to select the Census geography of the state as the unit of observation.

[5]This weight matrix constrains the neighbor structure to the k nearest neighbors. For a definition of k-nearest neighbor matrices, see, for example, Fischer and Wang (2011).

[6]It is noteworthy that this econometric framework is so general that it could be also used for quantifying direct and spillover effects of a set of factors that influence mortality caused by other pandemics such as the HIV/Aids pandemic, the Spanish flu or more recently the COVID-19 pandemic.

of the data set used, and presents then the direct and indirect effects estimates of the explanatory variables on opioid mortality rates. The last section summarizes and concludes the paper.

2   Econometric Framework

2.1   The model

The econometric approach we employ in this study is a spatial econometric panel data model that is known as spatial Durbin panel data model in the spatial econometrics literature (see, e.g., LeSage 2014) and can be written in matrix form as

$$y = \rho W y + X\beta + WX\theta + \iota_T \otimes \mu + \nu \otimes \iota_N + \epsilon \qquad (1)$$

where y is the $NT \times 1$ dependent variable vector of opioid mortality rates in state i $(i = 1, ..., N)$ at time t $(t = 1, ..., T)$, organized with t being the slow index for elements y_{it} in the vector y. X is an $NT \times Q$ matrix of Q explanatory variables that is organized in the same way. β is the associated $Q \times 1$ parameter vector.

W is an $NT \times NT$ weight matrix representing some sort of connections among the states. The matrix has a block diagonal form: $I_T \otimes w$, with $\otimes$ being a Kronecker product. w is a $N \times N$ spatial weight matrix reflecting spatial proximity of the N states that make up the panel of states over $t = 1, ..., T$ time periods. The diagonal elements $w_{ii} = 0$ $(i = 1, ..., N)$, and the matrix is row-normalized to have row-sums of unity. The $NT \times 1$ matrix-vector product Wy is referred to as spatial lag and reflects a linear combination of neighboring state values for the dependent variable. The scalar parameter ρ denotes the spatial dependence parameter that indicates the strength of spatial dependence between the vector y and the vector Wy.[7] The $NT \times Q$ matrix-matrix product WX is used to create spatial lags of the Q explanatory variables of the model, and represents a linear combination of characteristics from neighboring states, with the associated parameters θ.

---

[7]Note that the dependence parameter ρ is well defined over a limited interval that safeguards the existence of the matrix inverse $R^{-1} = (I_{NT} - \rho W)^{-1}$. This interval is $(\lambda_{min}, \lambda_{max})$, where $\lambda_{min}, \lambda_{max}$ are the minimum and maximum eigenvalues of the matrix $(I_{NT} - \rho W)$, respectively (LeSage 2021). λ_{min} will be negative, and λ_{max} will be one, provided the matrix W is row-normalized. In this study we use a lower bound of -1 to avoid the need to calculate the minimum and maximum eigenvalues of the large matrix R.

$\iota_T \otimes \mu$ represents an N-vector of state-specific fixed effects μ, with $\otimes$ being a Kronecker product that repeats the vector μ for each time period. $\nu \otimes \iota_N$ reflects a Kronecker product of the T-vector of time-specific effects ν, one for each time period. The $NT \times 1$ vector ϵ is a stochastic disturbance term, assumed to be normally distributed with zero mean and a scalar variance, σ^2.

$$\epsilon \sim \mathcal{N}(0_{NT}, \sigma^2 I_{NT}) \tag{2}$$

2.2   Bayesian estimation

Estimation of the spatial Durbin panel data model outlined in Eq.(1) is based on Markov Chain Monte Carlo (MCMC) estimation, with prior distributions assigned to the parameters ρ, $\delta = (\beta, \theta)'$ and σ^2.[8] Parameter restrictions are imposed on the dependence parameter ρ during MCMC sampling, using methods described in LeSage (2020), and Fischer and LeSage (2020). MCMC estimation involves sequentially sampling each parameter from conditional distributions. One of the main advantages of MCMC sampling is that conditional distributions for each parameter, given values of all other model parameters, take a form that is computationally simple to sample from.

The prior setup we use is standard. We rely on a normal prior for the parameters $\delta = (\beta, \theta)'$, associated with the X- and WX-variables:

$$p(\delta) \sim \mathcal{N}(\bar{\delta}, \bar{\Sigma}_\delta) \tag{3}$$

where $\bar{\delta}$ is a $2Q \times 1$ vector of prior means, and $\bar{\Sigma}_\delta$ a $(2Q) \times (2Q)$ prior variance-covariance matrix. We set the prior means to a value of 0.5 and the prior variance to 0.001.

For the dependence parameter ρ we employ a uniform prior

$$p(\rho) \sim U(-1, 1) \tag{4}$$

because this scalar parameter is constrained to lie in the open interval $(-1, 1)$. The constraint is imposed during MCMC estimation using a Metropolis-Hastings approach with rejection sampling,

---

[8]For μ and ν no priors, just demeaning transformations are used (see LeSage 2021). The reader interested in a recent survey of Bayesian inference methods in regional science may consult Mills and Parent (2021).

using methods described in Fischer and LeSage (2020), and LeSage (2020).

A non-informative prior is placed on the noise variance parameter σ^2, because we use an inverse gamma $(\bar{a}, \bar{b})$ distribution shown in Eq. (5).

$$p(\sigma^2) = \frac{\bar{b}^{\bar{a}}}{gamma(\bar{a})} \; (\sigma^2)^{-(\bar{a}+1)} \; exp(-\bar{b}/\sigma^2) \tag{5}$$

where $\bar{a} \to 0$, $\bar{b} \to 0$ reflect no prior information for this parameter.

As is traditional in the literature, we assume that priors for ρ, δ and σ^2 are independent. Given these priors we use the conditional distributions described in LeSage and Fischer (2020). A set of 4,000 draws were taken for determining posterior estimates of the parameters, with the first 500 omitted for burn-in of the sampler.

2.3 Direct and indirect impact estimates

LeSage and Pace (2009) point out that the estimates for the coefficients β and θ are not meaningful, if one is interested in how changes in the explanatory variables impact the dependent variable outcome, which is the object of interest in this paper. In our model that includes a spatial lag of the dependent and independent variables, a change in a single explanatory variable in state i has a direct impact in i as well as an indirect (spillover) impact from other-states $j \neq i$ to state i. We follow the suggestion of LeSage and Pace (2009) to calculate direct and indirect effects estimates that reflect a proper interpretation of the marginal effects for our nonlinear model which involves spatial lags of the dependent variable.[9]

The partial derivatives used to calculate the direct and indirect effects estimates take the form of an $NT \times NT$ matrix. Using matrix notation, the partial derivative for the qth explanatory variable X^q can be written as shown in Eq. (6).

$$\partial y / \partial X^q = (I_{NT} - \hat{\rho}W)^{-1} \; (\hat{\beta} + W\hat{\theta}) \tag{6}$$

where $\hat{\rho}, \hat{\beta}$ and $\hat{\theta}$ denote parameter estimates. Diagonal elements of this matrix represent own-partial

---

[9]See LeSage and Pace (2021) for a discussion of issues surrounding proper interpretation of the estimates from a variety of spatial regression models.

derivatives, while off-diagonal elements reflect cross-partial derivatives. Equation (6) illustrates that the partial derivatives of y with respect to the qth explanatory variable X^q has two specific properties (Elhorst 2010): *First*, if a particular explanatory variable X^q in a particular state i changes, not only the dependent variable in this state i will change, but also in other states $j \neq i$. *Second*, direct and indirect effects are different for different states in the sample. Direct effects are different, since the own-partial derivatives of the $NT \times NT$ matrix are different. And indirect effects are different, since the cross-partial derivatives of this matrix are different for different states.

If we have N states and Q different explanatory variables, then we get Q different $NT \times NT$ matrices representing direct and indirect effects. To improve the surveyability of the $Q \times NT \times NT$ effects estimates, LeSage and Pace (2009) propose using the mean of the main diagonal elements of the matrix of partial derivatives to produce a scalar summary of the *direct effects*. Using the mean of these NT different values generates a scalar summary that may be interpreted as representing how a change in the qth explanatory variable in the typical state impacts outcome y for this state (LeSage and Pace 2021).[10]

Indirect effects represent the impact of the ith state outcomes y_i from a change in the qth explanatory variable from the jth state $j \neq i$, and capture the off-diagonal elements of the $NT \times NT$ matrix. Specifically, the elements in the ith row of the matrix show

$$\partial y_i / \partial X_j^q \text{ for } i \neq j; \ i = 1, ..., N \tag{7}$$

reflecting how changes in each of the other states' qth explanatory variable impact outcomes in the ith state.

Note that our model can have positive (or negative) direct effects associated with negative (or positive) indirect effects for the qth variable, so that spillover impacts might work in the opposite direction of direct impacts, arising from changes in each explanatory variable. *Total effects* represent the sum of direct and indirect effects. In addition to calculating scalar summary measures of the effects, there is a need to calculate an empirical distribution from which measures of dispersion for

---

[10]Since the panel model includes time-specific fixed effects that capture variation over time periods, we can reasonably average the scalar summary estimates. But these scalar summary measures might ignore impact information over all time periods regarding the impact of changing values X^q at the observation level of direct (indirect) effects of the N states.

the effects can be constructed and used for inference regarding the statistical significance of the effects (LeSage and Pace 2021). This is done using a sequence of say 1,000 MCMC draws for the parameters ρ, β, θ to calculate a sequence of 1,000 scalar summary measures of these effects that reflect the posterior distributions.[11]

Given an estimate of the standard deviation for the scalar summary point estimates, we can test hypotheses regarding the significance of the various types of effects for each of the explanatory variables used in the model. t-statistics, t-probabilities, and lower 0.05 and upper 0.95 credible intervals can be computed using the retained MCMC draws. The mean of the draws for direct effects (for each of the Q explanatory variables), and the standard deviation of the draws is taken to construct the t-statistic, which is then used to find the associated t-probability. The lower 0.05 and upper 0.95 are based on the MCMC draws. Given a set of 1,000 MCMC draws, the lower 0.05 interval would be determined by the 50th value from the lowest set of sorted values, and the upper 0.95 by the 950th value from this sorted set (LeSage 2021).

3   Application of the Model

The empirical model estimated uses a pool of the 49 contiguous US states ($N = 49$) over six years from 2014 to 2019 ($T = 6$). The dependent variable vector y represents the annual opioid mortality rates, and we consider a suit of Q=17 explanatory variables. The choice of the time was based on the availability of data.[12]

3.1   Model specification issues

The definition of a spatial lag in spatial regression models in general, and our spatial Durbin panel data model in particular, depends on the choice of a spatial weight matrix that summarizes the topology of the data set. Clearly, a large number of weight matrices can be derived for the same spatial layout (see Fischer and Wang 2011 for a review). But k-nearest neighbor matrices that

---

[11]Reporting spatial Durbin model estimates based on direct and indirect scalar summaries is widespread in the spatial econometrics literature.

[12]ACS variables are not available for years prior to 2014, and Census data for 2019 was released only in 2021.

constrain the neighbor structure to the k nearest neighbors have received increasing popularity in spatial econometrics in recent years due to several reasons. *First*, there are new algorithms for calculating k nearest neighbors, and these are extremely fast for millions of observations (Matlab implemented these).[13] *Second*, interpretation of spatial spillovers having a constant number of neighbors is more intuitive. The magnitude of spillovers to the k neighbors can easily be calculated and understood. This is not the case where an observation unit (state here) has two neighbors and another has four neighbors as in the case of, for example, a contiguity-based weight matrix.

The number of neighbors, k, is the parameter of this weighting scheme. Its choice is an empirical issue. Our solution to this problem relies on a Bayesian Comparison approach (see LeSage and Fischer 2008).[14] This approach calculates the Bayesian posterior model probabilities of ten different k-nearest neighbor matrices with k=1, ..., 10, given our spatial Durbin panel data model specification. These probabilities are based on the log-marginal likelihood of the model obtained by integrating out all parameters of the model over the entire parameter space on which they are defined. A log-marginal likelihood of -174.86 and an associated model probability close to one indicates that the model specification with the k=8 nearest neighbor matrix is most consistent with the given panel data sample.

W is an 294×294 weight matrix representing connections between the states. W has a block diagonal $I_T \otimes w$, with $\otimes$ denoting a Kronecker product. w is a 49×49 spatial weight matrix constructed based on eight nearest neighboring states and identical during each time period in our model formulation. The diagonal elements $w_{ii} = 0$ $(i = 1, ..., 49)$, and the matrix is row-normalized to have row-sums of unity. The 294×1 matrix-vector product Wy is referred to as spatial lag and reflects a linear combination of neighboring states' values for the dependent variable vector of opioid death rates. The $294 \times Q$ matrix product WX is used to create spatial lags of the $Q = 17$ explanatory variables described in the following subsection.

---

[13]Note that distance is difficult to compute by comparison.

[14]An advantage of this model comparison approach relative to approaches based on likelihood-ratio or Lagrange multiplier methods is that the result does not depend on specific parameter values, but is valid for all possible parameter values.

3.2 Data and data sources

Data for the dependent variable vector y come from the US Centers for Disease Control and Prevention (CDC) Wonder Online Databases of Multiple Cause of Death, filtered for drug-related causes of deaths, estimated per 10,000 population for each state and year. Drug overdose deaths were identified using the *International Classification of Diseases, Tenth Revision* (ICD-10) underlying cause-of-death codes X40-X44 (unintentional), X60-X64 (suicide), X85 (homocide), or Y10-Y14 (undetermined intent).[15] Among deaths with drug overdose as the underlying cause, the opioid subcategory was determined by the following ICD-10 multiple cause-of-death codes: opium (T40.0); heroin (T40.1); natural and semisynthetic opioids (T40.2); methadone (T40.3); synthethic opioids other than methadone (T40.4); or other and unspecified narcotics (T40.6) as a contributing cause.[16]
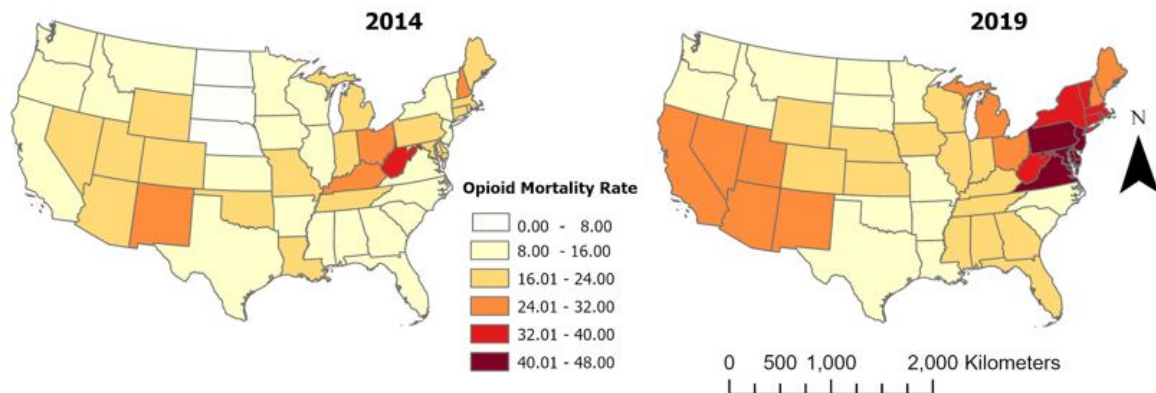

Fig. 1 Opioid mortality rates in US, 2014 and 2019

Figure 1 shows a marked temporal growth of the opioid mortality rates over the time frame, and increasing geographic heterogeneity in the spatial distribution of the rates in 2019, with hot spots in the Rust Belt and the Midwest. While the opioid crisis initially hit the Midwest states and New Mexico hardest, Western states such as California, Nevada and Arizona have also experienced

---

[15]See https://www.cdc.gov/nchs/icd/icd10.htm. The data are based on death certificates for US residents. Each death certificate contains a single underlying cause of death and up to twenty additional multiple causes. Drug-specific overdose numbers and rates might have changed substantially from 2017 to 2018 as a result of changes in reporting in some states (see Wilson et al. 2020).

[16]Note that the opioid subcategories are not mutually exclusive. Some deaths involved more than one opioid subcategory and were included in the opioid mortality rates.

increases in opioid mortality rates. Based on prior opioid research we selected the following set of sociodemographic and socioeconomic X-variables:[17]

(a) Five variables are considered to capture the racial/ethnic structure of a state: percentages of the population who identify as *White, African American, Native American, Asian* or *Hispanic/Latino* races. The data were derived from the US Census Bureau's Population Estimates Program called American Community Survey (ACS).

(b) To uncover the impact of age on opioid mortality rates, we include age in years based on standard definitions: percentages of people with *18-24 years, 25-34 years, 35-44 years, 45-64 years*, and with *65 years and more* (Data source: ACS).

(c) The level of educational attainment of the population is widely viewed as a relevant factor affecting opioid mortalities (see, for example, Blake-Gonzalez et al. 2021). Educational attainment is measured in terms of two variables, the percentage of people with 25 years or older who hold a *high school's degree and above* and the percentage of people having a *bachelor's degree or higher* (Data source: ACS).

(d) Economic conditions of a state could account for up to one-tenth of the rise of opioid mortality rates (Blake-Gonzalez et al. 2021). Following Nosrati et al. (2019) we consider three variables that address economic distress: the *poverty rate*, the variable of *lower income* and the *unemployment rate*. The data for these variables come from the American Community Survey. The *poverty rate* is calculated as the percentage of the population whose income falls below the state-specific level per year, while *lower income* is measured in terms of the number of people with income of less than $ 34,999 per 10,000 persons. The *unemployment rate* denotes the percentage of people aged 16 years and older who are unemployed in the resident labor force.[18]

(e) In addition to the above sociodemographic and socioeconomic factors, we include the *opioid*

---

[17]For the rationale choosing these variables, see Sun (2022), Blake-Gonzalez et al (2021), Wilson et al. (2020), Bohnert et al. (2011), Marotta et al. (2019), Nechuta et al. (2018) and King et al. (2014) among others.

[18]It is worth noting that the US Department of Health and Human Services utilizes unemployment in its opioid study (Ghertner and Groves 2018)

prescription rate to account for one of the opioid supply factors,[19] and *health insurance coverage*. The prescription rate is calculated by dividing the total number of opioid prescriptions dispensed annually by the annual resident population and multiplying the number by 100. The data come from the US Opioid Dispensing Rate Maps provided by IQVIA Xponent from the CDC in 2020. Health insurance coverage is measured in terms of the percentage of people with health insurance. To the extent that individuals are covered by health insurance, they may better afford to purchase prescriptions. Hence, health insurance may have the potential to increase the demand for drugs (Blake-Gonzalez et al. 2021).

3.3 Estimation results

Estimation is based on Markov Chain Monte Carlo estimation as described in Section 2.2. The resulting estimates are presented in Table 1. The table contains posterior means for the parameter estimates of the 17 explanatory variables along with the associated (asymptotic) t-statistics and z-probabilities. t-statistics represent the posterior mean divided by the posterior standard deviation both of which are calculated from the retained MCMC draws. The t-statistics are then used to determine the associated z-probabilities. R-square for the estimates is very high (0.94) and sigma-square low (0.0502). We see evidence of significant spatial effects working through the dependent variable, since $\hat{\rho} = 0.2939$ which is significant at the 99% level. The parameter estimates for the β-coefficients on the X-variables and the θ-coefficients on the WX-variables are not meaningful if one is interested in how changes in explanatory variables impact the dependent variable outcomes which is the focus of interest in this study. As emphasized in Section 2.3, it is is necessary to properly calculate the direct, indirect and total effects associated with changes in the explanatory variables on opioid mortality rates, given our spatial panel data model.

Table 2 reports the posterior mean effects estimates along with the 95 percent credible intervals and t-statistics that provide a valid inference with respect to the significance of the impact estimates. The scalar summary measures reflect: (i) a direct impact, (ii) an indirect (spatial spillover) impact and (iii) a total impact on opioid mortality rates for ceteribus paribus changes in the explanatory

---

[19]The recent spike in the number of opioid deaths has been reported widely and linked to the expanded use of prescription opioids and potentially inappropriate opioid prescribing practices (see Stewart et al. 2017 and Stopka et al. 2019).

Table 1 Posterior parameter estimates of the spatial Durbin panel data model

Variable (Parameter)	Posterior Mean	Asymp. t-stat.	z-prob.
Percent White (β_1)	−0.0163	−0.7389	0.4600
Percent Black/African American (β_2)	0.0078	0.1706	0.8645
Percent Native (β_3)	−0.2221	−1.8458	0.0649
Percent Asian (β_4)	−0.2059*	−2.3068	0.0211
Percent Hispanic/Latino (β_5)	0.0360	0.7658	0.4438
Percent Age 18-24 (β_6)	−0.1057	−1.2978	0.1944
Percent Age 25-34 (β_7)	0.1372	1.4731	0.1407
Percent Age 35-44 (β_8)	0.0610	0.5755	0.5650
Percent Age 45-64 (β_9)	−0.0594	−0.8160	0.4145
Percent Age $\geq$ 65 (β_{10})	−0.1694*	−2.2128	0.0269
Percent High School and above (β_{11})	−0.0807*	−2.4654	0.0137
Percent Bachelor and higher (β_{12})	0.0466	1.6510	0.0987
Poverty Rate (β_{13})	−0.0282*	−1.9677	0.0491
Lower Income(β_{14})	0.0238	1.3788	0.1680
Unemployment Rate (β_{15})	0.0210	0.6234	0.5330
Opioid Prescription Rate (β_{16})	−0.0064*	−2.3763	0.0175
Percent Health Insurance (β_{17})	−0.0145	−1.6671	0.0955
$W*$ Percent White (θ_1)	0.2098*	2.5127	0.0120
$W*$ Percent Black/African American (θ_2)	−0.3091	−1.8291	0.0674
$W*$ Percent Native (θ_3)	−0.7714	−1.7840	0.0744
$W*$ Percent Asian (θ_4)	0.7766*	2.4292	0.0151
$W*$ Percent Hispanic/Latino (θ_5)	0.6277*	3.1032	0.0019
$W*$ Percent Age 18-24 (θ_6)	0.0487	0.2115	0.8325
$W*$ Percent Age 25-34 (θ_7)	−0.3329	−1.1208	0.2624
$W*$ Percent Age 35-44 (θ_8)	−0.7836*	−2.9728	0.0030
$W*$ Percent Age 45-64 (θ_9)	−0.2515	−1.0777	0.2812
$W*$ Percent Age $\geq$ 65 (θ_{10})	−0.8490*	−4.7356	0.0000
$W*$ Percent High School and above (θ_{11})	−0.0220	−0.2096	0.8340
$W*$ Percent Bachelor and higher (θ_{12})	0.0962	1.0993	0.2716
$W*$ Poverty Rate (θ_{13})	−0.1078*	−2.1844	0.0289
$W*$ Lower Income (θ_{14})	0.1852*	3.7613	0.0002
$W*$ Unemployment Rate (θ_{15})	−0.4948*	−5.7702	0.0000
$W*$ Opioid Prescription Rate (θ_{16})	−0.0222*	−2.5986	0.0094
$W*$ Percent Health Insurance (θ_{17})	−0.0158	−0.9359	0.3493
Rho (ρ)	0.2939*	4.0535	0.0001

Notes: Results are based on a k-nearest neighbor matrix W with $k = 8$ nearest neighboring states; $N = 49$, $T = 6$ and $Q = 17$; parameter estimates calculated using LeSage's (2021) Panel Data Toolbox for MATLAB; performance: log-marginal likelihood= −174.8609, R-square= 0.9353, corr-square= 0.6507 and sigma-square= 0.0502; * indicates significantly different from zero.

variable concerned. A positive (negative) mean with positive (negative) lower and upper credible intervals should be interpreted as a positive (negative) effect. Effects whose credible interval span zero are not significant. These statistics for the scalar summary effects estimates are calculated using the retained MCMC draws for the (direct, indirect and total) effects (for each explanatory variable). The standard deviation of the draws is used to construct t-statistics that are taken to determine the associated t-probabilities. The lower 0.05 and upper 0.95 intervals are also based on the MCMC draws.

Focusing first on the direct effects estimates shown in Table 2, we see five out of 17 explanatory variables that are significantly different from zero. All these impact estimates are negative. The direct impact of changes in the senior (65 years and older) population share on opioid mortality is negative with a magnitude of 0.2042. This result matches our intuition, as we should expect that a larger percentage of senior people in the typical state i would reduce the opioid mortality rate in this state. A bit larger in magnitude is the direct effect of the *Native American* population share (-0.2545) which suggests that a larger percentage of this population in state i would reduce opioid mortality in i, since members of this population tend to die of drug overdose less likely than those from the White, African American and Hispanic subpopulations (for differences in opioid mortality by race see Nechuta et al. 2018).

Opioid Prescribing[20] representing the supply side has a negative direct effect of -0.0073, suggesting that this rate in state i would reduce the opioid mortality rate in this state i, in contrast to our expectation. This is likely to arise from the fact that our prescribing data includes only prescriptions for Medicare and Medicaid patients and, thus, underscores the use of prescription opioids, leading to biasing downward the effects estimates to a magnitude close to zero.

The direct impact of changes in the *Poverty Rate* is significant and negative on the opioid mortality rate in i, reducing this too, but with a much lower magnitude of 0.0327 as compared to the sociodemographic variables. The estimated direct effect of the *Unemployment Rate* is positive, but not significantly different from zero, suggesting that this rate has no impact on own-state opioid mortality. Its indirect effects estimate, however, indicates a significant and negative spillover impact, implying that increasing unemployment in neighboring state j decreases opioid mortality in state i.

---

[20]Prescription opioids include natural, semisynthetic opioids and methadone (Scholl et al. 2019).

Table 2 Direct, indirect and total impact estimates for changes in X- and WX-variables

Impact Estimates	Posterior Mean	t-stat.	t-prob.	Lower 0.05	Upper 0.95
A. Direct Impact Estimates					
Percent White	−0.0085	−0.3738	0.7087	−0.0524	0.0353
Percent Black/African American	−0.0041	−0.0842	0.9330	−0.1000	0.0873
Percent Native	−0.2545*	−2.0072	0.0453	−0.5028	−0.0083
Percent Asian	−0.1785	−1.9235	0.0550	−0.3623	0.0016
Percent Hispanic/Latino	0.0606	1.2011	0.2303	−0.0356	0.1582
Percent Age 18-24	−0.1050	−1.2992	0.1945	−0.2639	0.0500
Percent Age 25-34	0.1262	1.3160	0.1888	−0.0583	0.3103
Percent Age 35-44	0.0316	0.2972	0.7664	−0.1774	0.2446
Percent Age 45-64	−0.0698	−0.9563	0.3394	−0.2075	0.0750
Percent Age ≥ 65	−0.2042*	−2.6293	0.0088	−0.3535	−0.0522
Percent High School and above	−0.0825*	−2.4486	0.0147	−0.1488	−0.0167
Percent Bachelor and higher	0.0508	1.7462	0.0814	−0.0075	0.1085
Poverty Rate	−0.0327*	−2.2147	0.0272	−0.0610	−0.0037
Lower Income	0.0312	1.7802	0.0757	−0.0025	0.0672
Unemployment Rate	0.0021	0.0630	0.9498	−0.0654	0.0688
Opioid Prescription Rate	−0.0073*	−2.6367	0.0086	−0.0127	−0.0018
Percent Health Insurance	−0.0153	−1.7790	0.0759	−0.0321	0.0012
B. Indirect Impact Estimates					
Percent White	0.2851*	2.3991	0.0168	0.0569	0.5318
Percent Black/African American	−0.4279	−1.7446	0.0817	−0.9173	0.0489
Percent Native	−1.1673	−1.8483	0.0652	−2.4187	0.0171
Percent Asian	0.9952*	2.1655	0.0308	0.1122	1.9368
Percent Hispanic/Latino	0.8889*	2.9519	0.0033	0.3175	1.5079
Percent Age 18-24	0.0244	0.0776	0.9382	−0.5871	0.6369
Percent Age 25-34	−0.4046	−0.9602	0.3374	−1.2489	0.4292
Percent Age 35-44	−1.0646*	−2.8531	0.0045	−1.8139	−0.3602
Percent Age 45-64	−0.3749	−1.1445	0.2530	−1.0468	0.2564
Percent Age ≥ 65	−1.2530*	−4.4361	0.0000	−1.8542	−0.7393
Percent High School and above	−0.0640	−0.4300	0.6674	−0.3648	0.2361
Percent Bachelor and higher	0.1535	1.2188	0.2235	−0.0835	0.4039
Poverty Rate	−0.1619*	−2.2505	0.0249	−0.3084	−0.0266
Lower Income	0.2680*	3.5939	0.0004	0.1306	0.4227
Unemployment Rate	−0.6805*	−5.0752	0.0000	−0.9641	−0.4371
Opioid Prescription Rate	−0.0336*	−2.6377	0.0086	−0.0594	−0.0097
Percent Health Insurance	−0.0280	−1.2301	0.2193	−0.0754	0.0158

Impact Estimate	Posterior Mean	*t*-stat.	*t*-prob.	Lower 0.05	Upper 0.95
C. Total Impact Estimates					
Percent White	0.2767*	2.2097	0.0276	0.0392	0.5319
Percent Black/African American	−0.4319	−1.6138	0.1072	−0.9697	0.0891
Percent Native	−1.4218*	−2.0513	0.0408	−2.7932	−0.1314
Percent Asian	0.8167	1.6403	0.1016	−0.1490	1.8213
Percent Hispanic/Latino	0.9495*	2.9070	0.0038	0.3358	1.6187
Percent Age 18-24	−0.0806	−0.2494	0.8031	−0.7277	0.5337
Percent Age 25-34	−0.2784	−0.6080	0.5435	−1.1690	0.6221
Percent Age 35-44	−1.0330*	−2.6249	0.0089	−1.8313	−0.2848
Percent Age 45-64	−0.4447	−1.3038	0.1929	−1.1340	0.2234
Percent Age ≥ 65	−1.4571*	−4.8200	0.0000	−2.0865	−0.8960
Percent High School and above	−0.1465	−0.9060	0.3654	−0.4669	0.1834
Percent Bachelor and higher	0.2044	1.4858	0.1380	−0.0535	0.4814
Poverty Rate	−0.1946*	−2.5234	0.0119	−0.3508	−0.0488
Lower Income	0.2992*	3.7691	0.0002	0.1508	0.4650
Unemployment Rate	−0.6784*	−4.9108	0.0000	−0.9747	−0.4277
Opioid Prescription Rate	−0.0409*	−2.9591	0.0032	−0.0697	−0.0149
Percent Health Insurance	−0.0432	−1.8770	0.0611	−0.0908	0.0020

Notes: Impact estimates are calculated using LeSage's (2021) Panel Data Toolbox for MATLAB;
* indicates not significantly different from zero.

The final direct effects estimate that is significantly different from zero, is the direct impact of changes in the percentage of people with 25 years or older who hold at least a *high school's degree.* This negative direct effect has a magnitude of 0.0825, implying to reduce the opioid mortality in state i. The direct effects of all other explanatory variables are not significant, suggesting that they have no impact on own-state opioid mortality.

The direct impact estimates differ from the corresponding coefficient estimates outlined in Table 1. The difference is due to some feedback effect that comes into play in the direct effects estimates. The discrepancy is positive since the impact estimates exceed the coefficients, reflecting some positive feedback. Note that it would be a mistake to interpret the β-parameters as representing direct effects estimates. If we would incorrectly view, for example, the model coefficient β_3 on the *Percent Native American* variable as representing the direct impact, this would lead to an effect

16

not significantly different from zero. But the true direct impact estimate points to effects estimates that are negative and significant.

We now turn to considering the indirect (spatial spillover) effects estimates. These effects are measures for the magnitudes of spatial spillovers from the model. We emphasize that it would be a mistake to interpret the θ-coefficients as representing spatial spillover magnitudes in our model. To see how incorrect this is, consider the difference between the θ-coefficients and the true indirect effects correctly calculated from the cross-partial derivatives of the model, $\partial y_i / \partial X_j^q$, accumulated across all states $j \neq i$ and then averaged. These estimates are referred to as a measure of cumulative spatial spillovers, representing the impact of changes in other-state explanatory variables on own-state outcomes.

Table 2 shows five variables with statistically significant negative indirect effects estimates. The indirect impact of changes in the percentage of the middle-aged (*35-44 years*) and senior (*60 years and older*) populations in neighboring states j on the opioid mortality rate in state i is negative with a magnitude of 1.06476 and 1.2530, respectively, implying to reduce opioid mortality in i. This could arise from either product market impacts of the populations staying in neighboring states or from neighboring states economies, or perhaps networks between suppliers in neighboring states and own states. Similar arguments apply for the poverty, the unemployment and opioid prescription rates, resulting, however, in much lower magnitudes (-0.1619, -0.6805 and -0.0336) in comparison with the age-specific indirect effects.

Moreover, we see four positive spillover effects that are significantly different from zero. The effects of *White*, *Asian* and *Hispanic* subpopulations in neighboring states j is positive (0.2851, 0.9952 and 0.8889), suggesting to increase the opioid mortality rate in state i. The same happens with the *Lower Income* variable that shows a spillover effect 0.2680. These results are not entirely expected, but the reader should note that the relationship being estimated is an unidentified reduced form relationship.

Indirect effects have larger magnitudes than the direct effects. But it must be noted that the indirect effects are cumulated spillovers where cumulation takes place over all neighboring states $j \neq i$, neighbors to the neighboring states, and so on. While this may seem counter-intuitive, the indirect effect falling on any single state would be generally much smaller, consistent with spillovers

being a 'second-order' effect. Also the largest indirect effects would fall on nearby states. It is the cumulation of the spatial spillovers over all the states that leads to relatively larger indirect than direct effects.

Finally it should be noted that for matter of completeness Table 2 also lists the total effects that are defined as sum of the direct and indirect effects estimates. Inference statistics is based on the sum of two coefficients, $\hat{\beta} + \hat{\theta}$, and a proper measure for dispersion of this sum of coefficients. This measure can be constructed from the MCMC draws. A final point worth mentioning is that a focus on direct rather than total impacts that include spatial spillover effects would lead to an underestimate of the impacts.

4   Closing Remarks

This paper emphasizes the importance of employing a spatial Durbin panel data model for quantifying the impact of a set explanatory variables on opioid mortality rates in the US. The empirical model uses a pool of 49 US states over six years from 2014 to 2019, and a $k=8$ nearest-neighbor matrix to represent the topological structure between the states. Calculation of the direct (own-state) and indirect (spatial spillover) effects of the explanatory variables is based on Bayesian estimation and inference. The effects estimates reflect a proper interpretation of the marginal effects of the nonlinear model that involves spatial lags of the dependent and independent variables. This study's methodology and resulting findings provide a useful template for future empirical work using other geographic locations or shifting interest to other pandemics.

Our study has produced several interesting empirical results that carry some policy implications of interest for policy-makers and government agencies alike. *First*, evidence suggests that opioid mortality depends on characteristics of the state itself (direct impact), and on those of nearby states (indirect impact). *Second*, direct impacts are important, but externalities or (cross-state) spatial spillovers are more critical. In fact, there is empirical evidence that indirect effects of the poverty rate in neighboring states, for example, has an impact five times larger than the direct impact. When assessing the relative costs versus benefits of regional policies aiming to reduce opioid mortality, focus on direct (beneficial) rather than total (beneficial) effects, including spatial spillover benefits,

would lead to underestimating the benefits to society or the returns to such policies. *Third*, a state's sociodemographic structure (age and race) is essential. Economic distress of a state covered by the poverty rate, variable of lower income and unemployment rate are significant too, but comparatively less so as indicated by the total impact magnitudes.

A few limitations to our study should be noted. *First*, all the inferences are made conditional on the data used. The analysis is based on data reported on death certificates which we assume to be correct, but acknowledge they are imperfect. Death certificates may classify causes of death, and results may be biased by geographic heterogeneity in cause-of-death reporting if misclassification existed and/or changed over time and across states (see Alexander et al. 2018). *Second*, our prescribing opioid data covers only prescriptions for Medicare and Medicaid patients, and underscores the use of prescription opioids. This underscoring may be downward biasing the effects estimates of the prescription opioid supply variable. *Finally*, scalar summary effect measures provide information on the typical state rather than individual states. This may obscure variations in the direct, indirect, and total effects over space that would be of special interest to policy-makers and government agencies. Further explorations with observation-level direct, indirect, and total effects would produce additional insights to better understand regional differences in impact estimates that would be important for developing prevention and intervention strategies.

References

Alexander MJ, Kiang MV, Barbieri M (2018) Trends in Black and White opioid mortality in the United States, 1979-2015, *Epidemiology*, vol. 29(5), 707-715

Baltagi BH, Song SH, Jung BC, Koh W (2007) Testing for serial correlation, spatial autocorelation and random effects using panel data, *Journal of Econometrics*, vol. 140(1), 5-51

Blake-Gonzalez B, Cebula RJ, Koch JV (2021) Drug-overdose death rates: The economic misery explanation and its alternatives, *Applied Economics*, vol. 53(6), 730-741

Buchanich JM, Balmert LC, Pringle JL, Williams KE, Burke DS, Marsh GM (2016) Patterns and

trends in accidental poisoning death rates in the US, 1979-2014, *Preventive Medicine*, vol. 89, 317-323

Cossman JS, Cossman RE, James WL, Campbell CR, Blanchard TC, Cosby AG (2007) Persistent clusters of mortality in the United States, *American Journal of Public Health*, vol. 97(12), 2148-2150

Cushing BJ, Erfanian E, Peter D (2020) The national drug crisis – What have we learned from the regional science disciplines, *The Review of Regional Studies*, vol. 50(3), 353-382

Elhorst JP (2010) Applied spatial econometrics: Raising the bar, *Spatial Economic Analysis*, vol. 5(1), 9-28

Elhorst JP (2021) Spatial panel models and common factors. In Fischer MM and Nijkamp (eds.) *Handbook of Regional Science*, 2nd and extended edition, pp. 2141-2159. Springer, Berlin and Heidelberg

Fischer MM, LeSage JP (2020) Network dependence in multi-indexed data on international trade flows, *Journal of Spatial Econometrics*, vol 1(1): article 4, available online, doi: 10.1007/s43071-020-00005-w

Fischer MM, Wang J (2011) *Spatial Data Analysis. Methods and Techniques* [Springer Briefs in Regional Science Book Series]. Springer, Berlin, Heidelberg and New York

Ghertner R, Groves L (2018) The opioid crisis and economic opportunity: Geographic and economic trends. U.S. Department of Health and Human Services, *ASPE Research Brief*, 1-22

Hsiao C (2007) Panel data analysis – advantages and challenges, *Test*, vol. 16, 1-22

King NB, Fraser V, Boikos C, Richardson R, Harper S (2014) Determinants of increased opioid-related mortality in the United States and Canada, 1990-2013: A systematic review, *American Journal of Public Health*, vol. 104, e32-e42

Kopel J (2019). Opioid mortality in rural communities, *The Southwest Respiratory and Critical Care Chronicle*, vol. 7(31), 59-62

LeSage JP (2014) Spatial econometric panel data model specification: A Bayesian approach, *Spatial Statistics*, vol. 9, 122-145

LeSage JP (2020) Fast MCMC estimation of multiple W-matrix spatial regression models and Metropolis-Hastings Monte Carlo log-marginal likelihoods, *Journal of Geographical Systems*, vol. 22(1), 47-75

LeSage JP (2021) A panel data tool box for MATLAB, Department of Economics, University of Toledo

LeSage JP, Fischer MM (2008) Spatial growth regressions. Model specification, estimation and interpretation, *Spatial Economic Analysis*, vol. 3(3), 275-304

LeSage JP, Fischer MM (2020) Cross-sectional dependence specifications in a static trade panel data setting, *Journal of Geographical Systems*, vol. 22(1), 5-46

LeSage JP, Pace RK (2009) *Introduction to Spatial Econometrics*. CRC Press, Taylor & Francis Group, Boca Raton

LeSage JP, Pace RK (2021) Interpreting spatial econometric models. In Fischer MM and Njikamp P (eds.) *Handbook of Regional Science*, 2nd and extended edition, pp. 2201-2218. Springer, Berlin and Heidelberg

Marotta PL, Hunt T, Gilbert L, Wu E, Goddard E, El-Bassel N (2019) Assessing spatial relationships between prescription drugs, race, and overdose in New York State from 2013 to 2015, *Journal of Psychoactive Drugs*, vol. 51(4), 360-370

Mills JA, Parent O (2021) Bayesian Markov Chain Monte Carlo estimation. In Fischer MM and Njikamp P (eds.) *Handbook of Regional Science*, 2nd and extended edition, pp. 2073-2096. Springer, Berlin and Heidelberg

Nechuta SJ, Tyndall BD, Mukhopadhyay S, McPheeters ML (2018) Sociodemographic factors, prescription history and opioid overdose deaths: A statewide analysis using linked PDMP and mortality data, *Drugs and Alcohol Dependence*, vol. 190, 62-71

Nosrati E, Kang-Brown J, Ash M, McKee M, Marmot M, King L (2019) Economic decline, incarceration, and mortality from drug disorders in the U.S.A. between 1983 and 2014: An observational analysis, *Lancet Public Health*, vol. 4(7), e326-e333

Rossen LM, Khan D, Warner M (2014) Hot spots in mortality from drug poisoning in the United States, 2007-2009, *Health & Place*, vol. 26, 14-20

Scholl L, Seth P, Kariisa M, Wilson N, Baldwin G (2019) Drug and opioid-involved overdose deaths — United States, 2013–2017, *Morbidity and Mortality Weekly Report*, vol. 67, 1419–1427

Sparks PJ, Sparks CS (2010) An application of spatially autoregressive models to the study of US county mortality rates, *Population, Space and Place*, vol. 16(6), 465-481

Stewart K, Cao Y, Hsu MH, Artigiani E, Wish E (2017) Geospatial analysis of drug poisoning deaths involving heroin in the USA, 2000-2014, *Journal of Urban Health*, vol. 94(4), 572-586

Stopka TJ, Amaravadi H, Kaplan AR, Hoh R, Bernson D, Chui KKH, Land T, Walley AY, LaRochelle MR, Rose AJ (2019) Opioid overdose deaths and potentially inappropriate opioid prescribing practices (PIP): A spatial epidemiological study, *International Journal of Drug Policy*, vol. 68, 37-45

Sun F (2022) Rurality and opioid prescribing rates in U.S. counties from 2006 to 2018: A spatiotemporal investigation, *Social Science & Medicine*, vol. 296, 114788, available online, doi: 10.1016/j.socscimed.2022.114788

Wilson N, Kariisa M, Seth P, Smith H, Davis NL (2020) Drug and opioid-involved overdose deaths — United States, 2017-2018, *Morbidity and Mortality Weekly Report*, vol. 69, 290-297

Yang T-C, Noah A, Shoff C (2015) Exploring geographic variation in US mortality rates using a spatial Durbin approach, *Population, Space and Place*, vol. 21(1), 18-37