Socio–Spatial Network Dynamics of Public Health Crises in the US

by

Kushagra Tiwari

B.Tech, Vellore Institute of Technology (VIT), 2020

Submitted to the Graduate Faculty of

the Swanson School of Engineering in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2025

UNIVERSITY OF PITTSBURGH

SWANSON SCHOOL OF ENGINEERING

This dissertation was presented

by

Kushagra Tiwari

It was defended on

September 3, 2025

and approved by

Lisa Maillart, PhD, Professor, Department of Industrial Engineering

Jayant Rajgopal, PhD, Professor, Department of Industrial Engineering

Arun Kumar Kuchibhotla, PhD, Associate Professor, Department of Statistics and Data Science, Carnegie Mellon University

Dissertation Director: M. Amin Rahimian, PhD, Assistant Professor, Department of Industrial Engineering

Socio–Spatial Network Dynamics of Public Health Crises in the US

Kushagra Tiwari, PhD

University of Pittsburgh, 2025

The opioid epidemic and suicide are persistent U.S. public health crises with pronounced socio-spatial patterning. This dissertation develops socio–spatial exposure measures derived from Facebook's Social Connectedness Index (SCI) to quantify how county-to-county social ties and geographic proximity structure opioid overdose and suicide mortality. We define SCI–based social exposure and distance-decay spatial exposure and analyze their associations with county mortality using population-level statistical models. The objective is to characterize socio-spatial dynamics in these crises; all estimates are associational.

Chapter 2 investigates whether opioid overdose mortality is statistically associated with mortality levels in socially connected counties, consistent with a framework of social influence operating through large-scale social networks. Using SCI, we construct a measure of social influence defined as the SCI–weighted average opioid overdose mortality in socially connected counties. We assess its statistical association with focal-county overdose mortality while controlling for mortality in spatially proximate counties and comprehensive demographic and clinical covariates. Associations are estimated using cluster-robust linear regressions, with robustness evaluated through network and spatial autocorrelation models, two-way fixed effects, and two-stage least squares (2SLS).

Chapter 3 examines how social networks modulate geographic and sociodemographic heterogeneity in U.S. suicide mortality. We evaluate how risk and preventive policy information diffuse through inter-county social ties. Specifically, we test whether SCI–weighted average suicide mortality in socially connected counties is positively associated with focal county suicide rates, and whether social exposure to Extreme Risk Protection Orders (ERPOs)—the SCI-weighted share of population in ERPO–implementing counties—is negatively associated with suicide rates. Models include time-varying covariates and county and state-year fixed effects.

Chapter 4 analyzes monthly county-level data (2018–2021) to test whether socio–spatial

exposure improves short-term forecasting of opioid overdose mortality. We benchmark autoregressive baselines against models augmented with lagged mortality in socially connected and geographically proximate counties, using fixed effects and XGBoost. Time-ordered cross-validation shows consistent predictive gains, with rapid decay through social networks (half-life $\approx$ 9 days) and slower geographic decay (half-life $\approx$ 15 days).

Collectively, these chapters present an integrated framework for constructing network-reach exposure metrics, estimating mortality associations, and improving prediction for targeted public health intervention.

Table of Contents

List of Tables

List of Figures

To my mother, for always giving me strength; to my father, for his constant guidance; and to my sister—my second mother—for her unwavering belief and support. ,

Preface

First and foremost, I thank my advisor, Professor M. Amin Rahimian, for constant support and guidance throughout my PhD. He instilled the habit of asking better questions and fostered an open forum for discussion with rigorous, ongoing feedback. He taught me how to initiate collaborations, present work effectively, and articulate methodological arguments—both concisely and, when appropriate, in depth. Over the last four years, his insights were pivotal to my academic development and personal growth, and his consistent availability for discussion and steady mentorship decisively shaped this dissertation. Beyond research, I served as his teaching assistant for *Design of Experiments* and *Data for Good*, developing instructional materials, leading hands-on coding labs on model implementation in Python and R, and mentoring student project teams; this experience provided firsthand teaching practice, expanded my engagement with students, and broadened my doctoral training beyond research.

I am grateful to Professor Lisa M. Maillart, whose clear guidance consistently oriented me toward productive, actionable directions. During the development of my second paper, she urged me to think more broadly about figures—their role in narrative continuity and their capacity to answer central research questions—which materially strengthened the work. Her probing questions and feedback were invaluable.

I also thank Professor Jayant Rajgopal, with whom I interacted extensively throughout my PhD. He provided clear blueprints for organizing my workflow and decomposing problems to meet requirements efficiently. He also shaped my way of thinking by emphasizing first-principles reasoning, explicit assumptions, and disciplined prioritization. His guidance on systematically tracking the literature kept me current, and his openness to questions and discussion was invaluable.

I am particularly grateful to Professor Arun Kumar Kuchibhotla, whom I met in my third year. He helped surface conceptual and methodological gaps, encouraged adopting a Bayesian analytical lens in future work, and pressed for sharper framing and greater rigor in my second study; these contributions were pivotal to completing this dissertation.

I also had the pleasure of meeting Swapnil Keshari at the start of my PhD. We explored parallels between his work in computational biology and my work in operations research, which we both approached through machine learning and statistical methods. These cross-disciplinary exchanges were highly instructive and broadened my perspective.

I gratefully acknowledge the leadership and administrative staff of the University of Pittsburgh—across the Swanson School of Engineering and the Department of Industrial Engineering—for cultivating a collegial environment that enabled productive research, teaching, and collaboration. I also thank my colleagues, classmates, and friends for their steady support and companionship throughout these years.

1.0 Introduction & Overview

Public health crises in the United States continue to be a significant concern. Over the last four decades, drug overdose mortality has risen exponentially [59], with opioid-involved deaths now dominating drug-related mortality [52]. Polysubstance use [62] and disparities across age and geographic groups [110] compound the upward trend, which persists despite periods of modest decline. Suicide remains a leading cause of death and, like opioid use disorder (OUD), disproportionately affects young and middle-aged adults [92]. Both crises are no longer viewed solely as individual clinical conditions; instead, they are emergent properties of socio-spatial systems where interpersonal ties and spatial context shape risk.

Recent work demonstrates that social networks shape opioid misuse [71]. Peer networks containing non-users exert protective effects, whereas networks rich in substance users promote initiation and continuation of use [119]. Data from the CDC's State Unintentional Drug Overdose Reporting System (SUDORS) indicate that almost half of overdose deaths occur in the presence of a potential bystander [1], challenging the perception that overdoses occur in isolation.

Likewise, exposure to suicide can increase suicide ideation and capacity for suicide by providing concrete uses of lethal means that participate in suicide planning [20] – a phenomenon termed *suicide contagion* [45, 91, 109]. Suicide and OUD share risk factors such as mental illness, substance use, and social isolation, and exhibit overlapping socio-spatial patterns. Durkheim's sociological theory and the social-ecological model predict that reduced social integration elevates suicide risk; modern empirical studies confirm that deaths in one's social network increase one's suicide risk [46].

This thesis investigates how socio-spatial networks drive opioid overdose mortality and suicide at the county level in the United States and how such knowledge can inform prediction and prevention. In Chapter 2, we estimate how deaths in socially connected counties propagate opioid mortality. Using SCI to approximate county-level friendship ties, we construct

[1] https://www.cdc.gov/overdose-prevention/data-research/facts-stats/sudors-dashboard-fatal-overdose-data.html

a "socially lagged" measure of exposure and show that a one-standard-deviation increase in deaths in social proximity significantly increases opioid overdose deaths in ego counties, even after controlling for spatial proximity and socioeconomic covariates. Chapter 2 also discusses limitations of using mortality data and suggests leveraging agent-based models and structural causal models to simulate social contagion more mechanistically.

Chapter 3 extends this framework to suicide mortality. Counties are embedded in a socio-spatial system, and the models show that a one-standard-deviation increase in deaths in social proximity is associated with an increase in suicide deaths by roughly 3.3 per $100,000$ population; the effect remains significant when adjusting for deaths in spatial proximity. We further examine Extreme Risk Protection Orders (ERPOs) and find that social exposure to ERPO-adopting states reduces suicide mortality, whereas spatial exposure alone carries little independent signal. These findings imply that preventive policies diffuse along social ties and complement evidence that network structure can both amplify risk or confer resilience [119, 90]. Chapter 3 also recommends that prevention strategies incorporate SCI-derived network metrics to target counties with high potential to transmit or receive risk and protective influences.

Chapter 4 addresses prediction. A county–month panel of OOD (2018–2021) is assembled and evaluated with gradient-boosting and two-way fixed-effects models incorporating socio-spatial exposures. Augmenting autoregressive baselines with socio-spatial features yields modest but consistent improvements across one- and two-month horizons. An exponentially weighted moving average (EWMA) framework demonstrates that the predictive content of social and spatial exposures is concentrated within submonthly windows: social influence decays with a half-life of about nine days, whereas geographic spillover decays over approximately fifteen days. These results suggest that overdose diffusion is driven by short-lived excitation and that surveillance systems should prioritize submonthly data to detect emerging clusters. They also align with calls for process models such as multivariate Hawkes processes that explicitly capture self- and mutually-exciting dynamics [51]. Chapter 4 provides a roadmap for such modelling and emphasizes the practical implications for early-warning systems and resource allocation.

Collectively, these chapters demonstrate that socio-spatial networks are critical determi-

nants of both opioid overdose and suicide mortality. Social influence not only magnifies risk but also transmits the protective effects of policies, offering actionable leverage for public health. Through a combination of statistical inference, machine learning, and network science, this thesis provides new evidence that social connectedness shapes the spatial dynamics of contemporary public health crises and offers tools to harness these insights for prediction and intervention.

2.0 Measuring Network Dynamics of Opioid Overdose Deaths in the United States

The US opioid overdose epidemic has been a major public health concern in recent decades. There has been increasing recognition that its etiology is rooted in part in the social contexts that mediate substance use and access; however, reliable statistical measures of social influence are lacking in the literature. We use Facebook's social connectedness index as a proxy for real-life social networks across diverse spatial regions that help quantify social connectivity across different spatial units. This is a measure of the relative probability of connections between localities that offers a unique lens to understand the effects of social networks on health outcomes. We use SCI to develop a variable, called "deaths in social proximity", to measure the influence of social networks on opioid overdose deaths (OODs) in US counties. Our results show a statistically significant effect size for deaths in social proximity on OODs in counties in the United States, controlling for spatial proximity, as well as demographic and clinical covariates. The effect size of standardized deaths in social proximity in our cluster-robust linear regression model indicates that a one-standard-deviation increase, equal to 11.70 more deaths per $100,000$ population in the social proximity of ego counties in the contiguous United States, is associated with thirteen more deaths per $100,000$ population in ego counties. To further validate our findings, we performed a series of robustness checks using a network autocorrelation model to account for social network effects, a spatial autocorrelation model to capture spatial dependencies, and a two-way fixed-effect model to control for unobserved spatial and time-invariant characteristics. These checks consistently provide statistically robust evidence of positive social influence on OODs in US counties. Our analysis provides a pathway for public health interventions informed by social network structures. The statistical robustness of our primary variable of interest, deaths in social proximity, supports the hypothesis of a social network effect on OODs. Using agent-

based modeling (ABM) to simulate social networks can offer an effective method to design interventions that incorporate the dynamics of social networks for maximum impact.

2.1 Introduction

The opioid overdose epidemic is a major public health crisis in the US, with an exponentially increasing number of drug overdose deaths in the last four decades [57, 59]. Alpert et al. (2022) report that opioid overdose deaths (OODs) account for 75% of the increase in drug overdose deaths [4]. Addressing this crisis requires planned interventions that focus on supply-side regulations and the dynamics of social connections. The rate of initiation of opioid misuse is known to increase due to social influence [75]. Costello et al. (2021) report that of the 370 participants who entered an opioid withdrawal program, 97% identified knowing the individual with whom they initiated opioid use, with friendship being the most reported relationship between participants and their initiation partners [30]. Similarly, Rigg et al. (2018) note that two-thirds of misused drugs are obtained from friends and family [106]. Guarino et al. (2018) study of 539 young adults who misuse opioids and heroin indicates that their first experiences with the misuse of prescription opioids typically occur in a social setting with peers [48]. The misuse of prescription opioids has been growing among young people [58]. Syvertsen et al. (2017) make similar observations about young people who experiment with drugs and the initiation of drug use [116]. Social networks can have positive and negative impacts on substance use. Empirical results have indicated that peer networks with subjects who do not use substances have a positive influence on curbing drug abuse; however, networks consisting of more substance users are likely to increase substance use [119]. In intervention designs, recovery support strategies, including peer recovery, have shown encouraging results. Peer workers who have completed their recovery help others in recovery from substance addiction. This form of peer-supported recovery is found to be more effective in reducing the prevalence of opioid use disorder (OUD) [111].

Providing reliable measures of social influence on the opioid epidemic is complicated by the confounding factors that influence opioid misuse and social interactions, on the one hand,

and the ethical barriers to randomized experiments, on the other. Studies measuring social influence use complex systems-based generative models to understand these phenomena in areas such as voting contagion [18]. Specifically, Braha and Aguiar (2017) [18] use a generalized social voter model combined with spatial-statistical analysis to examine how social influence has shaped voting behavior in the US presidential election over the past century. They distinguish between social contagion and external influences (e.g., media and opinion leaders) to assess their impact on county-level vote share distributions over time and geography. By analyzing spatial patterns, they demonstrate that social influence is geographically clustered and spreads like a contagion across county borders. In contrast, applying similar modeling techniques to the opioid epidemic to design targeted interventions has encountered certain limitations. Homer et al. (2021) [56] discuss the complexity of modeling OUD and highlight the limitations of their models in accurately capturing real-life scenarios due to their simplicity. Recognizing the complexities of modeling the opioid epidemic through generative processes, our research aims to address these gaps using controlled regressions.

From the traditional setting to the new digital era of social networks, we have witnessed a significant shift in informal social interactions. Facebook is the largest informal online social network globally, with 2.1 billion active users and 239 million active users in the US and Canada. Given its broad reach and prevalence, Facebook connections can provide insight into real-life social networks in many geographical regions. Facebook has released a social connectedness index dataset that measures how intensely geographical locations are connected according to the relative probability of connections [9]. In Supplementary Information section A.1 we provide a review of SCI use cases as a proxy for real-life social connections in public health, epidemiology, economics and development applications. In the following paragraph, we justify our use of SCI for measuring the network dynamics of OODs in the US.

We use SCI to construct a measure of OODs in the social proximity of counties in the United States and investigate the statistical significance of its effect on county-level OODs after controlling for clinical, spatiotemporal, and socioeconomic confounders. Our analysis seeks to provide information on the role of social influence on OODs. Using SCI to represent the intensity of interpersonal networks between counties reflects the possibility

of physical interactions between county residents. Bailey et al. (2020) [10, 12] show that SCI is predictive of travel patterns within urban areas and throughout Europe. Coven et al. (2023) [31] show that counties with higher social connections to New York City are the most preferred destinations for those moving away from the city during the pandemic, highlighting the association of physical interaction with SCI. Kuchler et al. (2022) [69] use SCI to show that COVID-19 is more likely to spread between regions with stronger social networks and highlighted the potential of SCI to improve the prediction of epidemics. Although SCI provides a robust measure of social connectedness, some studies have explored alternative data sources, such as social media platforms, to examine trends in the context of the opioid epidemic. However, these approaches have faced significant challenges due to demographics and other data limitations [43, 94, 13]. For example, Pandrekar et al. (2018) [94] use Reddit data to analyze psychological categories and patterns of opioid abuse on a national scale. A major limitation of their study is that the data collected through the Reddit API do not provide access to user location information. In addition, the Reddit data do not indicate whether users are friends, which restricts the ability to analyze the structure of social networks and their association with OODs.

In contrast, SCI offers distinct advantages. SCI is location-based and provides detailed information about the structure of social networks in US counties. SCI measures friendship networks, serving as a proxy for real-world social connections. Unlike Reddit data, SCI allows for the analysis of location-specific friendship networks, which makes it particularly useful for studying how social networks influence the opioid epidemic on a population scale. Our choice of SCI as the network measure is informed by previous use cases that reflect the real-life dynamics of social connections in different domains such as education [36] and public health [69]. Our objective is to measure how these social connections contribute to heterogeneous patterns of opioid overdose deaths in US counties.

In our study, we present a novel perspective on analyzing the association between friendship networks and opioid overdose deaths at the population level in counties within three distinct geographical regions: the eastern United States, the central and western United States, and the entire contiguous United States. We provide statistical evidence linking the geographical spread of OODs with the structure of social connections.

To achieve this, we use a range of statistical methodologies. We use cluster-robust linear regression to account for intra-cluster correlation between counties within the same states, network and spatial autocorrelation methods to address the autocorrelation in error terms arising from unobserved factors shared among spatially or socially connected units, and two-way fixed effects models to control for unobserved spatial and temporal heterogeneities. Additionally, we perform robustness checks to account for the distance decay of proximity weights and apply a two-stage least squares method to jointly address spatial and network effects, discussed in the Supplementary Information sections A.5 and A.4.5. Our multifaceted statistical analysis demonstrates that our variable of interest, deaths in social proximity, is statistically significant across the three distinct geographical regions.

Understanding the role of social networks is important in designing interventions to combat opioid misuse behavior. Research measuring the network dynamics of opioid overdose death on the US population scale and at the resolution of counties remains limited, with the exception of [79, 32] that measure social network drivers of the opioid epidemic and use natural experiments to support their claims. Mäckle and Reunzi (2023) examine changes in county-level overdose deaths due to the reformulation of OxyContin and the must-access Prescription Drug Monitoring Program (PDMP) [79]. They analyze policy-induced shocks to estimate the indirect effect of friendship networks (measured by SCI) on OODs. Their analysis includes the correlation of "social proximity" with OODs by using a two-way fixed-effects model. Compared to Mäckle and Reunzi (2023) [79], our study differs in the methodological approach. Along with cluster-robust linear regression and two-way fixed effects, our analysis uses a network and spatial autocorrelation model to account for autocorrelation in error terms, which Mäckle and Reunzi (2023) [79] do not consider when investigating the association of "social proximity" with OODs. We also include a robustness check to account for spatial and network autocorrelation together and provide evidence that the statistical significance of the "deaths in social proximity" variable is not sensitive to the choice of distance decay function. Furthermore, our regression includes domain-specific covariates such as availability of naloxone and buprenorphine, opioid dispensing rate, mental health distress rate and state-level fentanyl and analog seizure data to account for illicit opioids which Mäckle and Reunzi (2023) [79] do not consider when investigating the association

between "social proximity" and OOD. Furthermore, Mäckle and Reunzi (2023) use the National Vital Statistics System (NVSS) public database and employ a backout procedure to recover mortality data points that the NVSS otherwise suppresses if the county has fewer than ten deaths. In contrast, our study uses mortality data from the National Center for Health Statistics (NCHS) which provides access to these suppressed data points. In this regard, our results complement their findings, as their coefficient for "social proximity" is also statistically significant. Concurrently with our work, Cuttler and Donahoe (2024) [32] have explored the dynamics of opioid death rates, focusing on SCI and the distance between counties to analyze spillovers. They posit that the increase in opioid demand is endogenous, resulting from spillovers between affected populations. Their study underscores the importance of social and spatial spillovers in estimating opioid demand, which is correlated with increased mortality rates. However, our work differs in its methodological approach, concentrating on constructing a socially lagged variable to assess the impact of social influence on overdose death rates.

In the following section, we formally introduce our variable of interest, "deaths in social proximity". The descriptive statistics serve as a starting point for later estimating the effect size of deaths in social proximity in the Results section.

2.2 Results

2.2.1 OOD Rates in Social and Spatial Proximity

The root of the opioid epidemic is partially associated with social contexts that mediate substance use and accessibility. Existing studies integrating social network analysis with the geographical spread of overdose deaths have demonstrated how social characteristics influence the trajectory of substance use, for example, geographic discordance, which means that the community in which the overdose death occurs is different from the community of residence [40]. Using data on overdose deaths from 2017 to 2020 in Milwaukee County, Wisconsin, Forati et al. (2023) build a social-spatial network framework to detect network

interaction hotspots for overdose deaths and analyze their geographical discordance [40]. However, their study is limited to Milwaukee County and does not extend to US population scales.

As discussed in the Introduction, the significant contribution of social influence in initiating substance use is well documented in the literature. Researchers have highlighted the impact of social networks on the regulation of substance use patterns based on the ego network [68]. However, applying these findings to a geographical context presents challenges, as geographical proximity substantially influences social connections and communication patterns across varying distances [123]. In our model, we control for the inherent spatial patterns of geographical proximity to refine our estimate of the effect of social networks on OODs.

Assessing the strength of social ties within every individual's network in a wide geographical area is very resource-intensive. To address this challenge, one potential approach is to aggregate and estimate the social networks of metapopulations residing in different localities, such as ZIP codes or counties in the United States. This offers a broader perspective and alleviates some of the constraints associated with individual-level analysis. In 2018, Meta Platforms, which operates the Facebook social network, released a data set that measures the distribution of social ties of location-specific networks globally. SCI is a surrogate for real-life friendship networks between registered Facebook users at each location [9]. It quantifies the strength of friendship ties in various locations using relative probability and is available at the ZIP code and county level for the United States. Formally, the SCI between two locations i and j is defined as follows [9]:

$$\text{SCI}_{ij} = \frac{\text{Facebook Connections}_{ij}}{\text{Facebook Users}_i \times \text{Facebook Users}_j}. \tag{1}$$

Here, Facebook Users_i represents the number of Facebook users in the county i. $\text{Facebook Connections}_{ij}$ is the total number of Facebook friendship connections between individuals in counties i and j.

Building on the established link between SCI and real-world social connections, we introduce two variables, "deaths in social proximity" and "deaths in spatial proximity", to capture the influence of social and spatial factors on the distribution of OOD rates in counties in the United States. The term "deaths in social proximity" indicates the average number of death

rates in alter locations weighted by their SCI to the focal node, also known as "ego". This variable operates as a socially lagged variable, accounting for the influence of death rates in "alter" locations, referring to the direct connections of the ego. On the other hand, "deaths in spatial proximity" measures the average number of death rates in the alters weighted by their inverse geographical distance to the ego. Unlike the socially lagged variable, this is a spatially lagged variable to account for the effects of deaths in the spatial vicinity of the ego. These factors work together to provide a comprehensive picture of how deaths are distributed and influenced by social and geographical factors in a given location. Quantitatively, deaths in social and spatial proximity, denoted by s_{-i} and d_{-i} for an ego location i, are defined as follows:

$$s_{-i} = \sum_{j \neq i} w_{ij} y_j, \quad \text{and} \quad d_{-i} = \sum_{j \neq i} a_{ij} y_j, \tag{2}$$

where y_i is OOD rate in US county i. The social and spatial proximity weights (w_{ij} and a_{ij}) are defined as follows:

$$w_{ij} = \frac{n_j \mathrm{SCI}_{ij}}{\sum_{k \neq i} n_k \mathrm{SCI}_{ik}}, \quad \text{and} \quad a_{ij} = \frac{1 + \frac{1}{d_{ij}}}{\sum_{k \neq i} (1 + \frac{1}{d_{ik}})}, \tag{3}$$

where n_j is the population of county j and d_{ij} is the distance between county locations i and j. Using $1 + 1/d_{ij}$ instead of $1/d_{ij}$ in the definition of a_{ij} improves numerical stability when dealing with long distances (large d_{ij} values) but does not change the decreasing nature of the spatial proximity weights with the increasing distances. To capture the effect of far away counties that would be otherwise down-weighted heavily by $1/d_{ij}$, in Supplementary Information section S5 we repeat our analysis with a slowly decaying distance function ($1/d_{ij}^{1/10}$) and observe that our main conclusions remain unchanged and are not sensitive to the choice of the distance decay function. To visualize the social and spatial adjacency weights across the United States, we aggregate the county-level data to the state level. This aggregation allows us to effectively depict the dense networks. We give the details of the state level aggregation in Supplementary Information section S2. It is important to note that our analysis of estimating the effect of "deaths in social proximity" on OODs is still conducted at the county level. The visualization of the proximity weights and proximity values, however, is presented at the state level to help understand the geographical dispersion of the social

and spatial proximity weights in relation to OOD rates. For comparison, in Supplementary Information section S2 we give the county-level visualization of social and spatial adjacency weights in Pennsylvania (PA). Figure 1A illustrates the state-level social network, measured by the social proximity weights (w_{xz}) in the contiguous United States. Figures 1B and 1C show the spatial dispersion of the proximity weights for the socially and spatially lagged variables for two ego states, California (1B) and Pennsylvania (1C), in the contiguous US.

Having formally introduced the socially and spatially lagged variables, we use mortality data from the National Center for Health Statistics (NCHS) for 2018-2019 to measure the state-level OOD rates. Figure 2 shows the state-level spatial distributions of our main variables of interest in the contiguous US: OOD rates (Figure 2A), deaths in social proximity (Figure 2B), deaths in spatial proximity (Figure 2C), and their differences (Figure 2D). Figure 3 shows the scatter plots of death rates (y_i), deaths in social proximity (s_{-i}), and deaths in spatial proximity (d_{-i}) for all counties in the contiguous US. The scatter plot matrix reveals a moderate linear dependence between death rates and spatially and socially lagged variables. In addition, there is a strong correlation between spatial and social proximity. However, the histograms that represent the distribution of these two variables exhibit differences. This contrast helps us identify the spatial effects of social influence on OOD and estimate an effect size for social proximity using controlled regressions described in the Results. Our choice of controls comes from domain knowledge consisting of clinical covariates and factors of social determinants of health (SDOH). The SDOH covariates are selected using the least absolute shrinkage and selection operator (LASSO) from an array of 17 variables. The details of SDOH and clinical covariates are in Methods.

2.2.2 Estimating the Effect Size of "Deaths in Social Proximity" Variable

Our outcome of interest is the county-level OOD rates that we measure using NCHS data from 2018 to 2019. We use cluster-robust linear regression to estimate the coefficient of the socially lagged variable. Robust estimators and clustering by state help us correct for correlation among counties, which might be higher for counties in the same state than between different states. Figure 4 shows the significant coefficient of the socially lagged

variable and provides statistical evidence for the effect of social networks on the spatial spread of OOD rates. The estimate of the effect size of the socially lagged variables is statistically significant across the eastern, western-central, and the entire contiguous US. The positive coefficient aligns with the theoretical proposition of the literature on the importance of social influence in the opioid epidemic [30]. The positive and significant magnitude of this effect size can originate from the dissemination of information through social networks about the initiation of use and availability of substances, leading to more OODs. The effect sizes for s_{-i} and d_{-i}, derived from the cluster-robust standard error model, along with other covariates, are in Supplementary Tables 7, 8, and 9 for the eastern, western-central, and contiguous US.

We standardize s_{-i} and d_{-i} prior to regression analysis. Consequently, the effect size for s_{-i} indicates that an increase of one standard deviation, equal to 11.69523 and 12.2417 more "deaths in social proximity" per $100,000$ population in the contiguous and eastern United States, respectively, is associated with an increase of nine deaths per $100,000$ population in ego counties within the eastern United States and thirteen deaths per $100,000$ population in the contiguous United States. For counties in the western and central United States, a similar increase of one standard deviation, equal to 5.7145 more deaths per $100,000$ population in the social proximity of ego counties, corresponds to six more deaths per $100,000$ population in ego counties. Despite observing a significant effect for the socially lagged variable on OOD rates, it is crucial to address potential issues arising from the inherent nature of the primary variable of interest. We recognize the methodological challenges due to correlated residuals when using statistical models to analyze social and spatial effects and perform robustness checks using network and spatial autocorrelation, as well as fixed-effects models to substantiate our findings.

Given that lag variables have a social and spatial component, we expect the error terms in our regression model to be correlated. To address this, we implemented a spatial error model (SEM) to test and correct for network and spatial autocorrelations in error terms. The autocorrelation in error terms arises from unobserved factors shared among spatially or socially connected units. Cluster-robust linear regression may fail to capture these autocorrelation stemming from unobserved factors, which may lead to bias and inefficient estimates. Hence, to add robustness to our analysis, we utilize spatial and network aurocorrelation

models. The methodological frameworks for the network and spatial autocorrelation models are explained in Supplementary Information sections A.4.2 and A.4.3. Supplementary Tables 10, 11 and 12 provide the SEM regression results for the network and spatial autocorrelation models for the eastern, western-central and contiguous United States, in which we find statistical evidence for correlated error. It is important to note that we performed two distinct models to test and correct for correlated error terms originating from both spatial and network dependence separately. Network autocorrelation might come from the structure of the socially lagged variable s_{-i}, while spatially correlated error terms could be attributed to the spatially lagged variable d_{-i}. Our autocorrelation models provide statistical evidence for the significance of social proximity in OODs.

To address spatial and temporal heterogeneity and enhance robustness, we employed a two-way fixed-effects model. Specifically, we included state-fixed effects to control for unobserved state-specific characteristics, such as regulations and policy environments, that are constant over time, but vary between states. We also incorporate year-fixed effects to absorb nationwide shocks or trends that could influence our outcome of interest. This modeling strategy enables us to robustly assess the statistical significance of "deaths in social proximity" while accounting for unobserved spatial and temporal confounders. The results of this model are in Supplementary Tables 13, 14 and 15 for the eastern, western-central and contiguous United States. Our findings consistently demonstrate a significant positive effect size for "deaths in social proximity". For a detailed explanation of the two-way fixed effect model, refer to Supplementary Information section A.4.4.

Figure 4 shows the confidence intervals (CI) for the cluster-robust standard error linear model, network autocorrelation, spatial autocorrelation, and the two-way, fixed-effects models. We consistently observe a positive and significant coefficient for s_{-i}, indicating the effect of social influence on the spread of OOD, while the effect size for d_{-i} has a varying CI, changes sign, and is not always significant. We also observe cluster-robust linear regression has a broader confidence interval compared to the other models, it is primarily because cluster-robust standard errors adjust for the intra-cluster correlation by accounting for the fact that there is less independent information than the total number of observations suggests. This adjustment often results in larger standard errors compared to conventional

ones, reflecting the reduced amount of independent information. There is also a loss of statistical power when doing this analysis, as the effective sample size becomes closer to the number of clusters rather than the total number of observations. However, in spite of accounting for intra-cluster correlation, we observe a statistically significant coefficient for "deaths in social proximity". A key takeaway from this result is that social connections are predominantly more significant than the effect of spatial proximity on OOD rates. This robustness check adds to the consistency of our statistical evidence for the size of the effect of "deaths in social proximity". In our analysis, given the network and spatial configurations inherent in s_{-i} and d_{-i}, we suspected and addressed the correlated errors that stem from the endogeneity of both variables simultaneously, using a two-stage least squares approach. The results of our implementation are discussed in the Supplementary Information section A.4.5 and confirm the positive and significant effect size of our social proximity variable in the eastern, western-central and entire contiguous United States (Supplementary Figure 13 and Supplementary Table 16). To end our analysis, we show the robustness of the effect of "deaths in social proximity" by accounting for the rate at which spatial adjacency weights decay with increasing distance. By accounting for the "deaths in spatial proximity" decay rate, we allow the effect of distant counties far from the focal county to be more pronounced. We use cluster-robust linear regression to test the significance of the coefficient s_{-i}. We observe a statistically significant effect for the coefficient of "deaths in social proximity" between counties in the eastern, western-central, and contiguous United States. The results of this implementation are discussed in the Supplementary Information Section A.5. The confidence interval plot for the model is illustrated in Supplementary Figure 14 and the effect sizes are shown in Supplementary Tables 17, 18, and 19.

2.3 Materials and Methods

2.3.1 Data Pipeline and Preprocessing

Our data sources are shown in Figure 5 and are explained in detail in the following paragraphs.

2.3.2 Mortality Data and Census Demographics

We measure the OOD rate from mortality data obtained from the National Center for Health Statistics (NCHS) for the years 2018-2019. This data set includes demographic details of individuals who have lost their lives to opioid-related overdoses. To first identify overdose related deaths we utilise the following International Classification of Disease (ICD) codes "X40","X41", "X42", "X43", "X44","X60", "X61", "X62","X63", "X64", "X85", "Y10", "Y11", "Y12", "Y13", "Y14". The X and Y codes provide information about deaths that have occurred due to substance overdose. Furthermore, to specifically target opioid overdose deaths we use the ICD T codes "T400", "T401", "T402", "T403", "T404", and "T406". The T codes determine the cause of death in the specification of opioids from other substances. Despite the comprehensive nature of the NCHS mortality data and the use of specific ICD codes to identify opioid overdose deaths, there are important limitations associated with this dataset. Death certificates may not always specify the drugs involved in an overdose, and some overdose deaths involve multiple drugs, making it difficult to determine which substance was primarily responsible. In addition, the analysis also incorporates demographic data on the broader population of the United States of America. The data is stratified at the county levels. To extract this information, we utilize the R package "tidycensus" to systematically retrieve data from data.census.gov.

2.3.3 Clinical and Mental Health Covariates

Clinical factors such as the opioid dispensing rate (ODR), availability of naloxone, and access to buprenorphine for the treatment of opioid use disorder are used as controls in our

regression. Morgan et al. (2018) [86] underscore the role of naloxone in reducing opioid-related harm, while Pendergrass et al. (2019) [95] emphasize the importance of buprenorphine availability in mitigating overdose fatalities. The data source for clinical factors such as naloxone, buprenorphine, and ODR is the IQVIA Xponent database, which provides the number of prescriptions for naloxone and buprenorphine distributed throughout counties in the US through retail pharmacy channels. The ODR, which captures the total number of opioid doses dispensed, is measured as morphine milligram equivalents (MME). The current wave of the opioid epidemic from 2013 to the present has witnessed an increase in the illicit use of synthetic opioids such as fentanyl. Furthermore, Kuehn (2023) [70] discuses the impact of fentanyl on overdose deaths, particularly among adolescents. Pergolizzi et al. (2018) [96] also discuss the role of fentanyl in exacerbating the opioid epidemic. To control for the supply of illegal substances that contribute to OODs, we incorporate state-level fentanyl and analog seizure data from the National Forensic Laboratory Information System as a control variable. Our selection of clinical and illicit-supply covariates is comprehensive based on the data sources available to us. In addition, several studies have shown associations between mental health and opioid overdose deaths [37, 107, 120]. Thus, we used data on frequent mental health distress from County Health Rankings and Roadmaps (CHRR). Frequent mental distress is the percentage of adults who reported poor mental health for more than 14 days in response to the question "How many days during the past 30 days was your mental health not good?" [124]. We use this measure to control for the effect of mental health-related issues in counties.

2.3.4 Population Density and Political Affiliation

We also control for population density and political affiliation that are identified in the literature on the opioid epidemic and the structure of social networks. To account for the heterogeneity associated with SCI and opioid use among populations residing in urban and rural counties, we include population density in our regression model [10, 98]. The risk status of opioid misuse is also associated with the political affiliations and liberal status of states [49]. Therefore, we control for the effect of political affiliation in our regression by accounting

for counties' political leanings using 2016 general election data at the county level.

2.3.5 SCI Data

SCI is available at the ZIP code and county levels in the United States. We chose counties as our analysis unit instead of ZIP codes because the latter had significant limitations. A considerable number of ZIP codes have missing SCI data. Typically, ZIP codes without SCI data are those with low populations or those designated exclusively for institutions. Institutional ZIP codes are assigned to areas predominantly occupied by specific institutions, such as hospitals, universities, or military bases. These institutional ZIP codes can also introduce spatial bias; for example, ZIP codes with hospitals are more likely to report higher overdose death rates. To ensure a more continuous and representative spatial framework, we use counties as our unit of analysis. This approach mitigates the issues of missing data and spatial bias, providing a more robust basis for analysis.

2.3.6 Social Determinants of Health

Socioeconomic factors shape social structures and ties; therefore, incorporating these factors is essential to interpret the influence of social networks on OOD. Shared socioeconomic conditions can also influence behaviors that mirror peer influence, and therefore, we include socioeconomic covariates as controls in our analysis to mitigate the risk of bias due to missing confounders when estimating the effect size of our deaths in the social proximity variable (s_{-i}). Social determinants of health (SDOH) encompass aspects of physical infrastructure, economic context, healthcare context, and social environment of counties. Our selection of SDOH covariates is based on socioeconomic predictors of the opioid epidemic. Liu et al. (2023) [76] demonstrate the effects of SDOH measures on drug overdose death locations. Therefore, we use SDOH variables to control for socioeconomic factors that may confound the impact of s_{-i} on county-level OOD rates. These covariates are selected from a pool of SDOH variables using the Least Absolute Shrinkage and Selection Operator (LASSO) technique to avoid multicollinearity issues. We select a subset of covariates from an array of 17 SDOH variables that are listed in Supplementary Table 6. We provide details of the

LASSO selection process in the Supplementary Information section A.3.

2.3.7 Statistical Models

The model coefficients are evaluated for statistical significance at $p < 0.05$ level. Using county-level opioid overdose death rates as the outcome variable, we set up regression models to test the significance of deaths in social proximity to explain the county-level death rate after controlling for deaths in spatial proximity, clinical (mental health, availability of naloxone and buprenorphine and opioid dispensing rates in pharmacies) and fentanyl covariates, and socioeconomic covariates selected using LASSO from SDOH variables, as well as population density and political affiliation covariates. Linear regression and cluster-robust linear regression are analyzed for statistical significance (details in Supplementary Information section A.4.1). The residuals in the regression are weighted by population size of the counties, so that the inference is reflecting the population residing in the counties across the US. To account for correlated error terms from the spatial and network structure of the variables d_{-i} and s_{-i}, respectively, we use spatial and network autocorrelation models (details in Supplementary Information sections A.4.3 and A.4.2). To account for unobserved space- and time-invariant characteristics that might be associated with the covariates in our regression, we use a two-way fixed-effects model. This model includes fixed temporal effects for 2018 and 2019 and fixed spatial effects for states in the eastern United States, the western and central United States, and the contiguous United States (details in Supplementary Information section A.4.4). As a robustness check, we also use a two-stage least squares regression to account for correlated error terms that arise from the simultaneous estimation of effects from s_{-i} and d_{-i} (details in the Supplementary Information section A.4.5). As a final robustness check, we use distance decay weight to account for the effect of faraway counties from the focal county (details in the Supplementary Information section A.5).

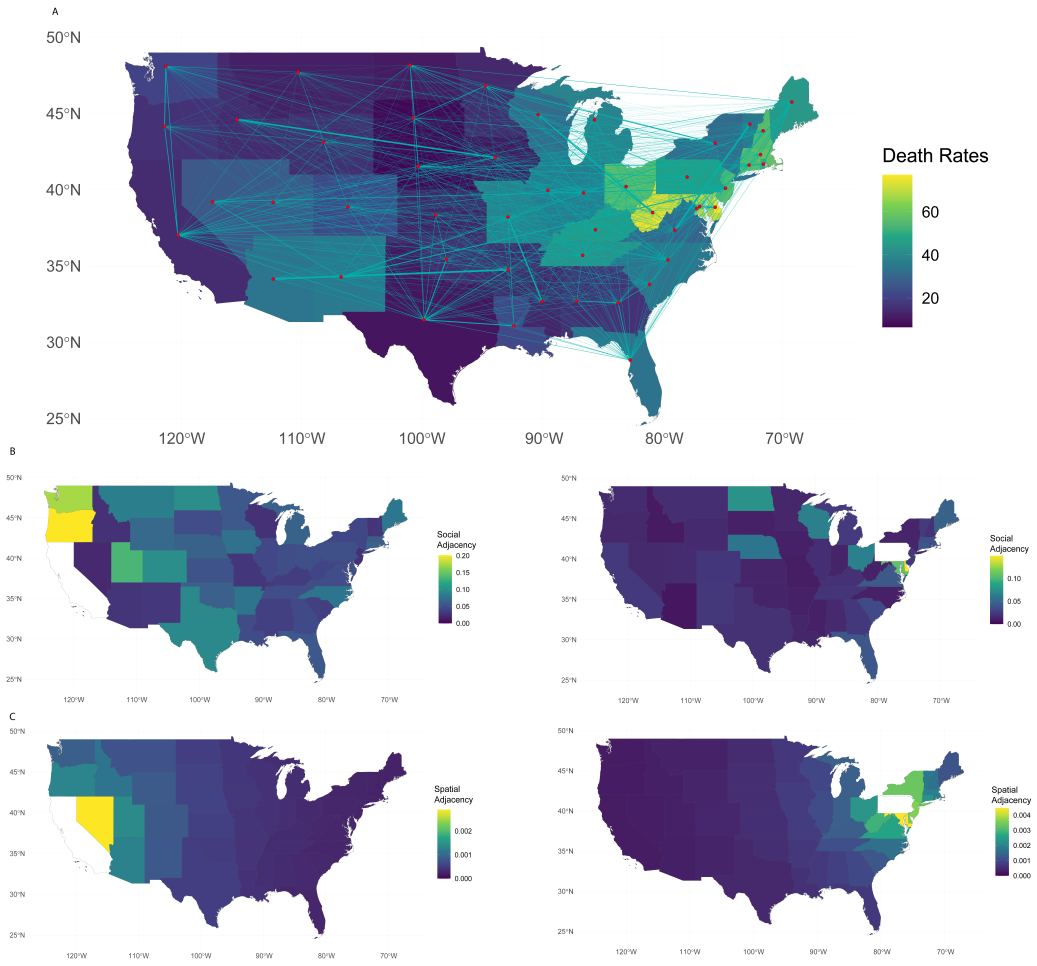

Figure 1: A) The spatial distribution of overdose death rates per 100,000 population in contiguous US from 2018 to 2019. Superimposed on this map is a social network diagram with edge widths representing the state-level social proximity weights . B) The two middle maps show the social proximity weights of alter states to California (on the left) and Pennsylvania (on the right). C) The bottom two maps show the spatial proximity weights of alter states to California (on the left) and Pennsylvania (on the right). *Maps generated using R version 4.2.2 (2022-10-31, available at* `https://cran.r-project.org`. *Network diagram created using Gephi 0.10.1 (2023-01-17), available at* `https://gephi.org`. *Image editing performed with Adobe Illustrator 27.9.5.*

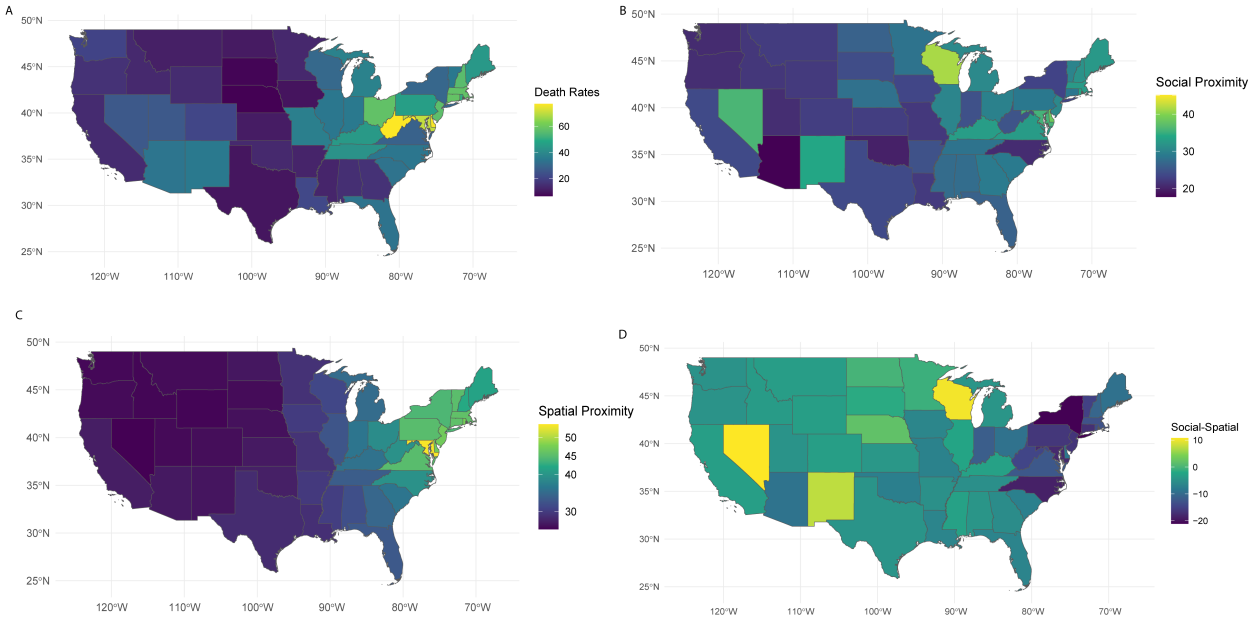

Figure 2: A) The top left map shows the spatial spread of state-level opioid overdose death rates in the contiguous US. B) The top right map shows the spatial dispersion of "deaths in social proximity" for states in the contiguous US. C) The bottom left map shows the geographical spread of "deaths in spatial proximity". D) The bottom right map shows the difference between deaths in social and spatial proximity from top right and bottom left maps. *Maps generated using R version 4.2.2 (2022-10-31, available at https://cran.r-project.org). Image editing performed with Adobe Illustrator 27.9.5.*

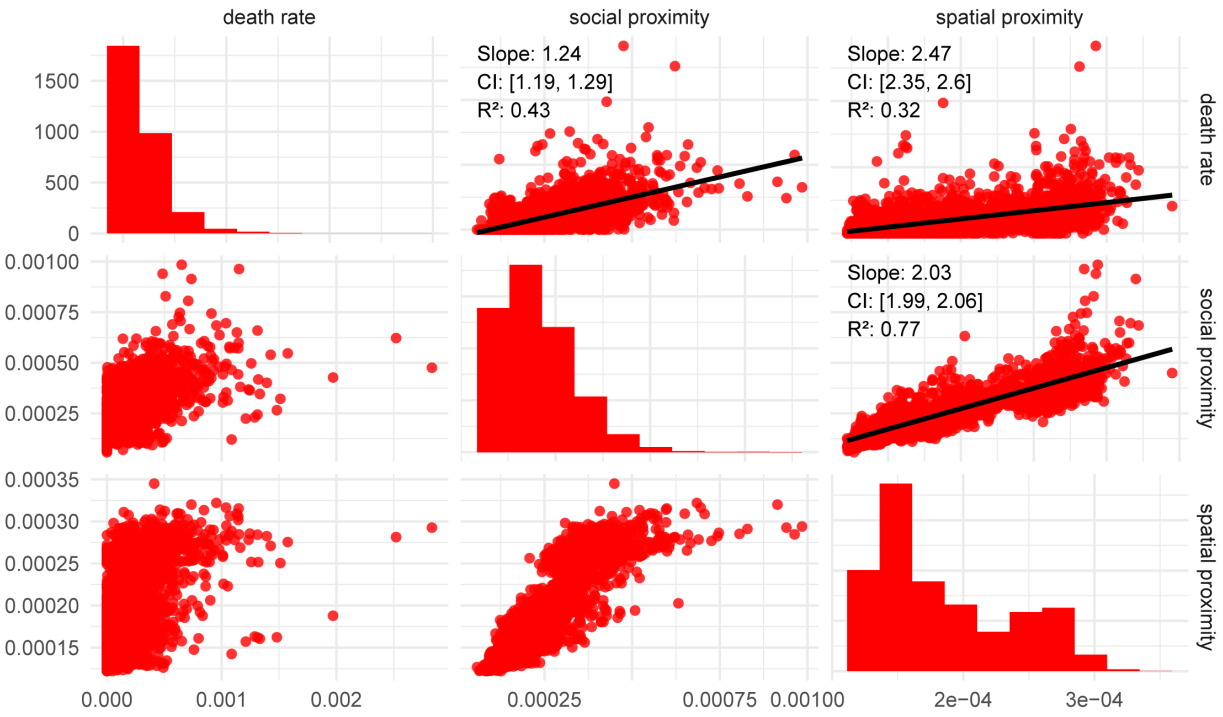

Figure 3: The figure shows the distribution and relationships between the primary variables of interest (death rates y_i, deaths in social proximity s_{-i}, and deaths in spatial proximity d_{-i}). The histograms on the main diagonal depict the distributions of y_i, s_{-i}, and d_{-i}. Moving to the upper triangle, we observe the degree of linear dependence between these variables, while the lower triangle displays scatter plots. *Scatter plot matrix generated using the R version 4.2.2 (2022-10-31, available at* `https://cran.r-project.org`*). Image editing performed with Adobe Illustrator 27.9.5.*

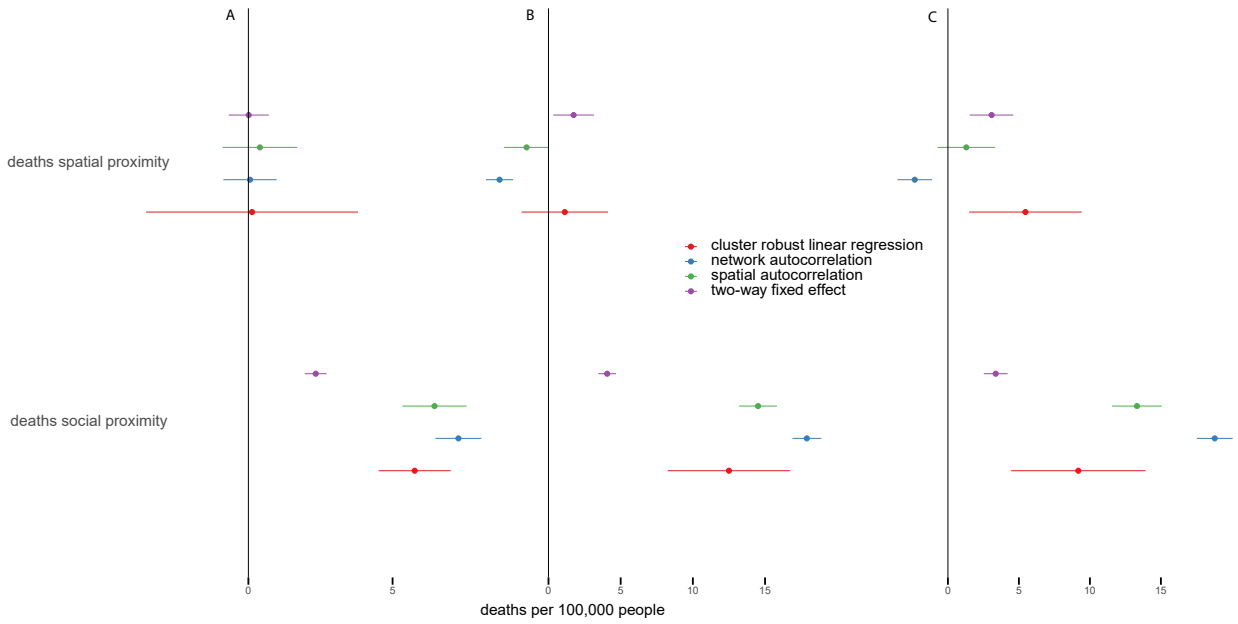

Figure 4: A) The plot shows the coefficient confidence interval plots for western and central US counties. The coefficient for s_{-i} for cluster-robust linear regression (Supplementary Table 8), network and spatial autocorrelation (Supplementary Table 11), and two-way fixed effects model (Supplementary Table 14), all indicate a positive, significant ($p<0.001$) coefficient for s_{-i}. B). Shows the coefficient plot for social and spatial proximity for counties in the contiguous US. The coefficient for s_{-i} for cluster-robust linear regression (Supplementary Table 9), network and spatial autocorrelation (Supplementary Table 12), and two-way fixed effects models (Supplementary Table 15) are all positive and significant ($p< 0.001$). C) Shows the coefficient plot for s_{-i} and d_{-i} for counties in the eastern US. The coefficient for s_{-i} for cluster-robust linear regression (Supplementary Table 7), network and spatial autocorrelation models (Supplementary Table 10), and two-way fixed effects models (Supplementary Table 13) are all positive and significant ($p < 0.001$). The effect sizes for standardized s_{-i} in the cluster-robust linear regression models indicate that a one-standard-deviation increase, equal to 11.69523, 12.2417, and 5.7145 more deaths per $100,000$ population in the social proximity of the ego counties in contiguous, eastern and western-central United States, respectively, is associated with thirteen more deaths per $100,000$ population in contiguous and nine more deaths eastern and six more deaths in western-central US counties.

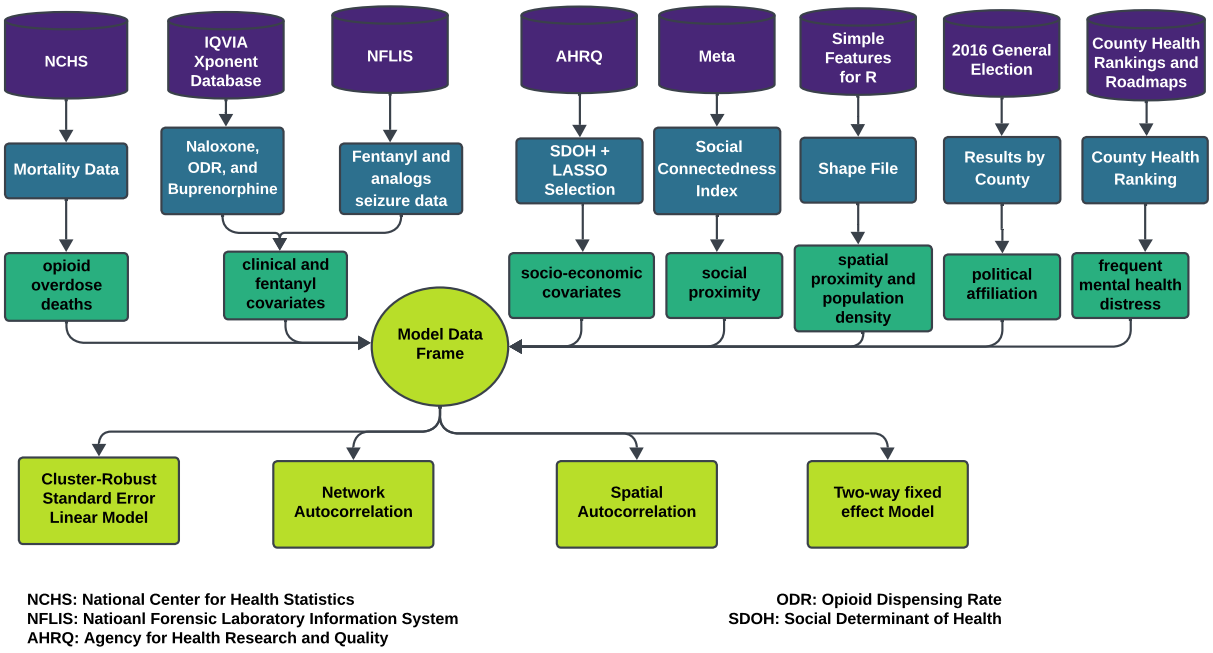

Figure 5: The diagram depicts the data pipeline for our analysis, including the data streams for the primary variable of interests, s_{-i} and d_{-i}, as well as the relevant socioeconomic and clinical covariates. It also outlines our regression models for estimating the effect of peer influence as measured by SCI on county-level OOD outcomes.

3.0 Socio-spatial Patterns of Suicide Mortality in the United States

Suicide mortality in the United States exhibits substantial geographical and sociodemographic heterogeneity. Yet the role of large-scale social networks in shaping this variation remains underexplored. We integrate data on county-level suicide mortality (2010–2022) and Facebook's Social Connectedness Index (SCI) to assess how both the risk of suicide mortality and the effect of firearm restriction policies propagate through inter-county social ties. First, using two-way fixed effects regression models with sociodemographic, economic, and spatial controls, we find that a one-standard-deviation increase in (SCI-weighted) suicide mortality in socially connected counties is associated with an increase of 2.78 suicide deaths per 100,000 people in the focal county (95% CI: 1.06- 4.50). Second, we examine Extreme Risk Protection Orders (ERPOs) — state-level firearm policies that allow temporary restriction of firearm access for individuals at risk of self-harm — and show that counties with stronger (Facebook) social ties to ERPO-adopting states experience reductions in suicide mortality, even without local policy implementation. Our findings suggest that a one-standard-deviation increase in ERPO social exposure is associated with a decrease of 0.214 suicide deaths per 100,000 people in the focal county (95% CI: 0.342-0.086). This protective association persists after adjusting for geographical proximity and including state-by-year fixed effects that capture time-varying state-level factors. In sum, our findings suggest that social networks can facilitate the diffusion of both harmful exposures and protective interventions. This socio-spatial structuring of suicide mortality underscores the need for network-driven prevention strategies that incorporate social network topology (e.g., SCI-derived influence metrics), alongside more traditional approaches based on geographical targeting.

3.1 Introduction

Suicide mortality is a persistent public health crisis in the United States (US). In 2023, suicide was the eleventh leading cause of death in the US, accounting for an estimated 49,000

fatalities [23]. However, this national statistic does not capture substantial age-specific heterogeneity. Adolescents and working-age adults are the most vulnerable. Among individuals aged 10–34 and 35–44, suicide was the second and fourth leading cause of death, respectively [23]. To better understand and address these disparities across different demographics, the social ecological framework provides a valuable lens, categorizing risk factors and preventive measures for suicide at the individual, community, or societal level [102]. Examples of risk factors include social isolation [19], mental health and substance use disorders [**?**, 21], bullying and cyberbullying [54], and access to lethal means (particularly firearms) [83]. Further, knowing someone who died by suicide [115], or being exposed to suicide-related information through social media may amplify one's suicide risk [101].

In this paper, we focus on the effect of social networks, which span the individual level (close interpersonal connections) and the community level (indirect and broader social ties). Social networks can either amplify or mitigate the risk of suicide, depending on their structure and nature. For instance, social isolation or interaction with negative peers may influence an individual toward unhealthy coping mechanisms, such as substance use and misuse. In addition, exposure to suicide death can increase suicidal ideation and increase capability via exposure to method-specific examples of lethal means that guide planning [20] – a phenomenon termed *suicide contagion* [45, 91, 109]. Durkheim [61] posited that suicide mortality rates in a given population are affected by the degree of social integration, with elevated levels under conditions of social disconnection. Building on this foundational theory, empirical studies have demonstrated that the risk of suicide mortality increases following exposure to suicidal behavior within one's social network [46]. In sum, theoretical and empirical studies suggests that suicidal behaviors can propagate through both interpersonal and community networks [100, 88, 109]. The role of such social exposures is especially prominent among adolescents and young adults who frequently encounter suicide-related content via digital media [8, 90]. Despite these negative influences, social connection is protective: family, peer, and school connectedness and a sense of belonging are associated with lower suicidal ideation and attempts, and simple connection-building interventions (e.g., caring contacts) reduce subsequent suicide risk [60, 121, 82, 6, 66, 87]. Given the established role of social ties in suicide risk, adopting a network science perspective to examine suicide mortality offers a

valuable interdisciplinary approach for understanding complex public health issues such as suicide [97].

Until recently, efforts to assess population-level associations between social networks and suicide mortality were limited by data constraints. Researchers relied on small samples or qualitative studies due to the lack of high-resolution social network data. The release of the Social Connectedness Index (SCI) by Meta in 2018 enabled the quantification of social ties between geographic regions using aggregated Facebook friendship data [9]. The SCI measures the relative probability that a Facebook user in region i is a friend of a user in region j, normalized by the product of the number of active Facebook users in each region. This formulation allows for comparisons of social connectedness between pairs of regions, where higher SCI values indicate stronger interpersonal ties. Constructed from billions of anonymized friendship links and periodically updated, the SCI provides a scalable proxy for real-world social connectedness, facilitating the study of how network structure relates to spatial patterns in behavioral and health outcomes. Bailey et al. [9] demonstrated the empirical relevance of SCI by showing that it correlates with inter-county migration patterns, trade volumes, job search behavior, patent citations, and specific public health outcomes. The availability of SCI data has made possible studies examining how behavioral and health outcomes propagate across socially connected regions. For instance, Charoenwong et al. [26] found that counties socially connected to early COVID-19 outbreak areas exhibited earlier reductions in mobility. Holtz et al. [55] showed that compliance with public health mandates varied systematically with SCI-based social proximity to counties with more stringent policies. Similarly, Tiwari et al. [117] demonstrated that opioid overdose death rates in socially proximate counties were positively associated with rates in the focal county. Together, these findings underscore the utility of SCI to capture the diffusion of behavioral and outcomes across social networks.

Further, research shows that social network structure is a modifiable target for suicide prevention—specifically, that increasing supportive ties to low-risk peers, disrupting high-risk clustering, and leveraging prosocial central connectors (e.g.,gatekeeper programs) can shift norms and facilitate help-seeking, thereby reducing suicidal ideation, attempts, and deaths [25, 35]. However, most studies have been restricted to small samples or specific

settings (e.g., school-based cohorts, military units, clinically high-risk youth) and have not evaluated whether large-scale inter-county friendship networks across the United States are associated with geographic variation in suicide mortality or with the diffusion of preventive policies. As a result, the broader socio-spatial dynamics of suicide risk–particularly the indirect transmission of protective interventions through social ties–remain insufficiently understood. One prominent example of a socially mediated preventive intervention is the implementation of Extreme Risk Protection Orders (ERPOs), also known as "red flag laws". ERPOs are civil court orders that allow temporary restriction of firearm access from individuals deemed to present a risk to themselves or others [84]. This socially mediated approach leverages interpersonal networks to identify and minimize suicide risks proactively [113]. Specifically, it enabling law enforcement, family members, friends, and close associates to file petitions based on observed behaviors that indicate a heightened risk of suicide, which may result in temporarily restricting firearm access. Given that concerned individuals often initiate ERPO interventions within a person's social network, they offer one potential avenue to examine how social ties influence suicide prevention efforts empirically. Empirical evidence from multiple states shows ERPOs' efficacy in reducing suicide mortality. Connecticut's issuance of 762 ERPOs between 1999 and 2013 was estimated to avert per 10 to 20 orders, with individuals targeted by these orders initially exhibiting suicide rates approximately 40 times higher than the general population [113]. These features make ERPOs a useful case to examine in our study, where we assess whether exposure to such policies through social connectedness influences suicide mortality across counties. Building on these insights, our study systematically evaluates the role of social connectedness in shaping county-level suicide mortality.

The contributions of our study, which integrates data on county-level suicide mortality and social connectedness, are three-fold:

1. We define and quantify county-level social exposure and spatial exposure to suicide mortality.
2. We estimate associations between county-level suicide mortality and social exposure, with and without adjustment for spatial exposure.
3. We evaluate whether network-mediated exposure to a firearm-restriction policy is associ-

ated with reduced suicide mortality in non-adopting counties that are socially connected to counties that adopted the policy.

We test two hypotheses to evaluate whether network-mediated exposures are associated with county-level suicide mortality in the US:

H1 a one-standard-deviation increase in the SCI-weighted average suicide death rate in socially connected counties is positively associated with the focal county's suicide death rate, controlling for spatial exposure.

H2 a one-standard-deviation increase in social exposure to firearm-restriction policy is negatively associated with the focal county's suicide death rate.

We estimate county–month panel models with county and state–year fixed effects, controlling for time-varying sociodemographic covariates. Although the associations identified in our study do not imply causality, they highlight how social networks can function as pathways for both detrimental socio-spatial influence and beneficial dissemination of preventive policies. In conclusion, our findings highlight the role of social networks in suicide mortality and support the integration of network-based approaches into the design of public health interventions.

3.2 Results

3.2.1 Socio-Spatial Patterns of Suicide Mortality in the US

We analyzed county-level suicide death data from the National Vital Statistics System (NVSS) Multiple Cause of Death files for the period 2010–2022, encompassing 40,794 county-year observations across the US. Our primary outcome of interest was the overall suicide mortality rate per county (per 100,000 people), across all age groups. Our primary objective was to assess whether exposure to suicide deaths in socially connected counties is positively associated with suicide mortality in the focal county (**H1**). To test whether the estimated association between social connectedness to other counties and suicide mortality in the focal county was confounded by geographical distance, we compared two approaches, with and

without controlling for spatial proximity. To that end, we considered a spatial exposure variable defined as the weighted average of suicide mortality rates in all US counties other than the focal county, with county-level weights inversely proportional to the geographical distance to the focal county's centroid.

To estimate these associations, we utilized two-way fixed effects regression models with county and year indicators, effectively controlling for time-invariant local characteristics and for annual patterns common across all counties. Exposure to suicide mortality outcomes in other counties was captured through two metrics. The first metric, denoted by s_{-it}, corresponds to "deaths in social proximity". It is defined as the weighted average of suicide mortality rates in year t in counties other than focal county i, where the weight of a given county is equal to its SCI with the focal county. The second metric, denoted by d_{-it}, corresponds to "deaths in spatial proximity". It is defined as the weighted average of suicide mortality rates in year t in counties other than focal county i, where the weight of a given county is equal to the inverse of its geographical distance to the focal county's centroid. Detailed mathematical formulae for these metrics are provided in equations 4 and 5 in the Methods Section.

First, we estimated the relationship between suicide mortality in the focal county and "deaths in social proximity" without controlling for spatial proximity. In this model, a one-standard-deviation (1-SD) increase in "deaths in social proximity" was associated with an increase of 3.34 suicides per 100,000 people in the focal county (cluster-robust 95% CI: [1.76, 4.93], $p < 0.01$; Fig. 6, red estimate). Given that the strength of social connections is correlated with geographical proximity, we next evaluated whether this observed population-level association was confounded by spatial proximity.

Second, we estimated a model including both variables, i.e., "deaths in social proximity" and "deaths in spatial proximity", to disentangle their independent associations with suicide mortality. In this combined model accounting for both social and spatial proximity, the association between the suicide mortality rate of the focal county and "deaths in social proximity" remained statistically significant and substantial (2.78 suicides per 100,000, cluster-robust 95% CI: [1.06, 4.50], $p < 0.01$; Fig. 6). Simultaneously, the suicide mortality rate of the focal county was also positively associated with "deaths in spatial proximity",

albeit at a smaller magnitude (0.88 suicides per 100,000, SE = 0.316, $p < 0.05$). This finding suggests that social ties independently influence the risk of suicide mortality, beyond spatial proximity.

Complete regression results, including the covariate effects of population density, age structure, racial/ethnic composition, median household income, unemployment, educational attainment, and English proficiency, are detailed in Table 1. The robustness of the s_{-it} coefficient after controlling for deaths in spatial proximity suggests that social influence is positively associated with elevated suicide risk independent of spatial proximity. This provides the first national-scale evidence that suicide mortality in the U.S. exhibits socio-spatial structuring: risk is influenced not only by physical location but also by networked social ties (**H1**).

To test the robustness of these associations to outcome specification, we re-estimated the models using age-adjusted suicide mortality as the dependent variable, thereby accounting for cross-county differences in population age structure. Results from these sensitivity analyses (see Supplementary Results; Table 20) confirmed the positive and statistically significant relationship between deaths in social proximity and suicide mortality, though the estimated magnitudes were attenuated by approximately 60% relative to the unadjusted outcome. This attenuation is expected, as age adjustment removes demographic variability in baseline suicide risk. Yet, the consistency in direction and significance of the estimates across specifications underscores the robustness of our findings.

The subsequent subsection explores whether protective influences, specifically exposure to firearm restriction policies (**H2**), similarly propagate through these socio-spatial channels.

3.2.2 ERPO Policy Exposure through Social Networks

In this section, we investigate whether suicide mortality is associated with indirect exposure to firearm access restrictions through social networks connecting a county of interest to counties in other states that have enacted ERPOs. We hypothesize that counties with greater social connectedness to ERPO-adopting counties experience lower suicide mortality, even in the absence of local ERPO adoption. To test hypothesis H2, we construct a

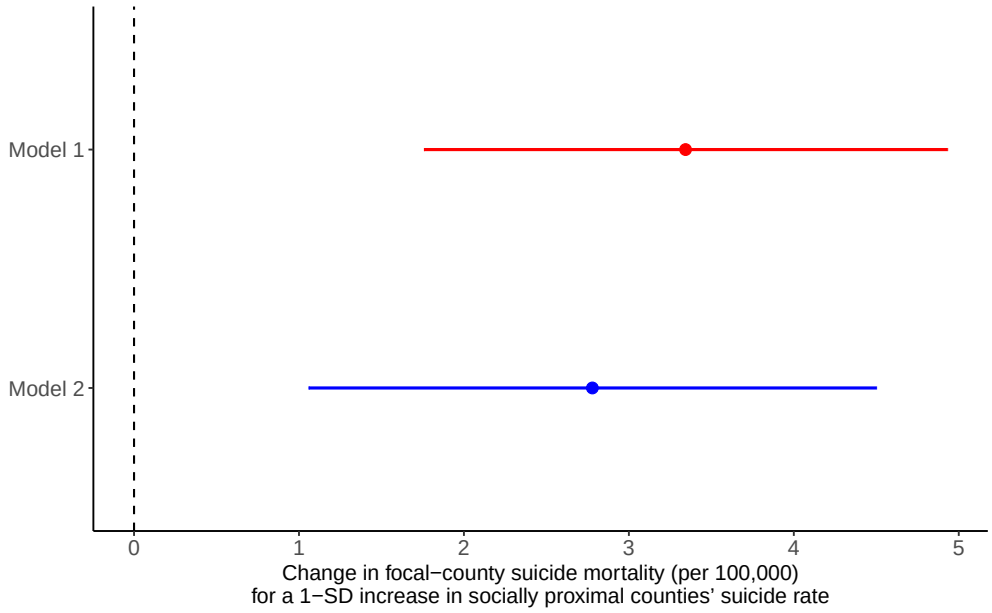

Figure 6: Role of social ties in county-level suicide mortality. Estimated regression coefficients ($\hat{\zeta}_1$) for suicide death rates in socially proximal counties (s_{-it}) in two models: without controlling for spatial proximity (red) and with deaths in spatial proximity (d_{-it}) as control (blue). Horizontal lines denote 95% confidence intervals; vertical dashed line indicates the null hypothesis ($\zeta_1 = 0$). Point estimate without controlling for spatial proximity: 3.34 (95% CI: [1.76, 4.93]); point estimate controlling for spatial proximity: 2.78 (95% CI: [1.06, 4.50]). Both models include county and year fixed effects and sociodemographic control variables (see Table 1).

Table 1: Estimates of socio-spatial influences on county-level suicide mortality obtained via two-way fixed effect regression. Column (1) presents estimates from a model regressing county-level suicide mortality on standardized deaths in social proximity (s_{-it}). Column (2) additionally controls for standardized deaths in spatial proximity (d_{-it}) to disentangle social influence from geographical proximity. Both models include county and year fixed effects and adjust for population density, age distribution (percent aged 0–17, 18–44 and 45–64), racial composition (percent Asian, Black, and Other racial subgroups), ethnicity (percent Hispanic), median household income, percent with limited English proficiency, percent unemployed, and percent with less than high school education. Standard errors are clustered at the state level.

	Dependent variable: County-level suicide deaths per 100k	
	Model 1	Model 2
	(1)	(2)
s_{-it}	3.344***	2.779***
	(0.790)	(0.856)
d_{-it}		0.783**
		(0.316)
Population Density	−1.181***	−0.987**
	(0.372)	(0.375)
Percent aged below 18	−12.818	−13.944
	(10.075)	(10.309)
Percent aged 18-44	−18.720	−20.628
	(15.324)	(15.658)
Percent aged 45-64	−12.087	−13.136
	(9.031)	(9.242)
Precent aged above 65	−16.951	−18.475
	(13.716)	(14.031)
Percent Black	−1.385*	−1.421*
	(0.746)	(0.766)
Percent Asian	−0.839***	−0.822***
	(0.293)	(0.289)
Percent Other	0.371**	0.308*
	(0.180)	(0.163)
Percent Hispanic	−3.709***	−3.590***
	(0.800)	(0.808)
Median Household Income	−0.699***	−0.692***
	(0.166)	(0.168)
Percent of population who do not speak English that well	−0.066	−0.036
	(0.075)	(0.072)
Percent unemployed	0.018	0.029
	(0.158)	(0.150)
Percent with less than high school education	−0.012	−0.022
	(0.125)	(0.124)
Observations	40,794	40,794
R^2	0.946	0.946
Adjusted R^2	0.941	0.941

standardized metric of ERPO social exposure using the SCI, which quantifies cross-county social ties. We then estimate the association between ERPO social exposure and suicide mortality using a two-way fixed effects regression, accounting for county- and time-specific heterogeneity. Finally, to disentangle the role of social exposure from that of geographical proximity (and assess the robustness of our results), we incorporate two regressions, i.e., with and without control for the same spatial exposure variable to assess robustness. The analysis proceeds by first describing ERPOs and their empirical effectiveness, then estimating their direct effects, followed by results for indirect social exposure and robustness to spatial confounding. The staggered implementation of ERPOs across states provides a chance for a statistical framework to explore the effect of social networks on suicide rates.

Since ERPO implementation is specific to each state, its *direct* impact is limited to counties within the states that enact it. This geographic separation of direct effects allows for a thorough evaluation of *indirect* social network influences. Specifically, the absence of direct ERPO exposure in non-implementing states enables us to isolate and measure the indirect effects of social networks on suicide mortality, which arise solely from social connectivity rather than direct legislative actions.

Since ERPO implementation is specific to each state, its *direct* impact is limited to counties within the states that enact it. This geographic separation of direct effects allows for a thorough evaluation of *indirect* social network influences. Specifically, the absence of direct ERPO exposure in non-implementing states enables us to isolate and measure the indirect effects of social networks on suicide mortality, which arise solely from social connectivity rather than direct legislative actions.

We begin by estimating the association of the ERPO policy in the counties of the implementation states with suicide rates. Table 2 presents the estimated coefficients from the direct-effect specification as shown in equation 7 in the Methods section. The ERPO indicator exhibits a statistically significant association of $\widehat{\psi} = -0.528$ (cluster-robust 95% CI: [-0.93, -0.126], $p < 0.01$). Standard errors are clustered by state (51 clusters), providing conservative inference that is robust to within-state correlation. Since the ERPO variable changes only at the state–year level, we include county fixed effects to control for county-specific characteristics that do not change over time, and year indicators to account for

nationwide shocks such as macroeconomic cycles.

After demonstrating that the ERPO has a statistically significant negative association with suicide mortality in implementing states' counties, we aim to investigate whether this effect indirectly impacts counties that have not implemented the ERPO through friendship networks. To measure ERPO exposure through social networks, we defined the *ERPO Social Exposure*$_{it}$ metric. The metric quantifies the share of the focal county's friendship ties directed toward counties in ERPO-adopting states in year t. A formal definition of this metric is provided by equation 8 in the Methods section. We hypothesize that *ERPO Social Exposure*$_{it}$ will be negatively associated with focal counties' suicide mortality. This may occur because suicide-related behaviors and outcomes are interlinked with social ties. To test this hypothesis, we implement a two-way fixed-effects regression as shown in equation 9. Figure 7 provides the empirical context for the regression analysis in equation 9. The map illustrates spatial heterogeneity in the change (Δ) in ERPO Social Exposure from 2010 to 2022. Counties within states that enacted ERPO statutes during this interval (e.g., CA, WA, OR, CT, NY, NJ, MA, RI, IL, FL) exhibit elevated exposure (highlighted in yellow) due to a substantial increase in the proportion of their cross-state SCI ties with counties in other ERPO-adopting states relative to 2010. Additionally, several counties within states that did not adopt ERPO laws but share borders or strong social connections with adopting states (e.g., selected counties in NV, PA, VT) similarly demonstrate increased exposure. In these instances, we observe a rise in the exposure metric, as socially proximal states implement ERPO statutes, despite the absence of local legislative change. Consequently, the yellow shading identifies areas of indirect ERPO exposure propagated through interstate social ties, providing the necessary variation to evaluate the association between increased ERPO exposure and subsequent reductions in suicide mortality.

Figure 8 presents the 95% CI of the estimated coefficient $\widehat{\delta}_1$ associated with ERPO social exposure. Table 2, Column 2 presents estimates from the indirect-exposure specification. The *ERPO Social Exposure*$_{it}$ coefficient is $\widehat{\delta}_1 = -0.214$ (cluster-robust 95% CI: [-0.342, -0.0866], $p < 0.01$), indicating a statistically significant negative association between suicide mortality and social connectedness to ERPO-adopting states after accounting for state-by-year fixed effects (γ_{st}). Because *ERPO Social Exposure*$_{it}$ is z-standardised (mean = 0,

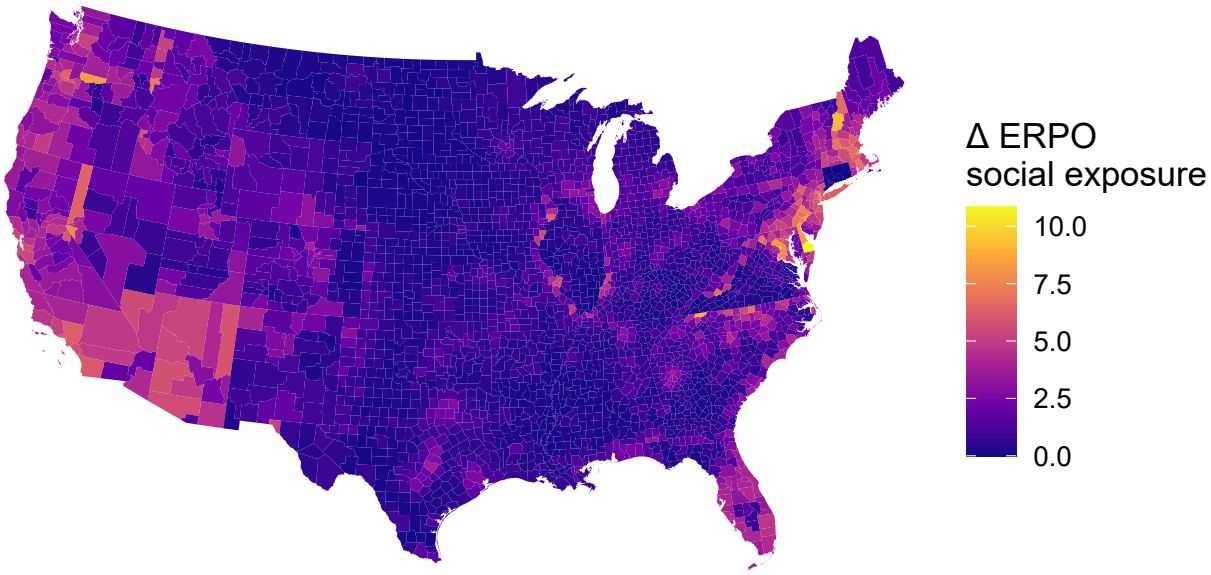

Figure 7: **County–level change in ERPO social exposure, 2010–2022.** Colours show the change (Δ) in standardised ERPO Social Exposure between 2010 and 2022, measured in within–sample standard–deviation units. Positive (yellow) values indicate that a county's social ties have become more concentrated in states that adopted ERPOs, whereas positive (purple) values indicate declining exposure. Although the underlying analysis covers *all* U.S. counties, the map shows only the 48 contiguous states and the District of Columbia; Alaska, Hawaii, and U.S. territories are not shown.

SD = 1), $\widehat{\delta}_1$ measures the change in suicide deaths per 100,000 population produced by a one-standard-deviation increase in out-of-state exposure. Raising a county's exposure by one SD is associated with a reduction of 0.214 suicide deaths per 100,000. The inclusion of county-fixed effects (ϕ_i) and state-by-year fixed effects (γ_{st}) ensures that the estimate is identified from within-state, cross-county variation in ties to ERPO-adopting states, net of time-invariant local characteristics and contemporaneous state-level shocks. Regressions are population-weighted, and standard errors are clustered at the state level.

To assess the robustness of this association and to differentiate the role of social networks from geographic proximity, we introduce ERPO spatial exposure. Similar to *ERPO Social Exposure*$_{it}$ the *ERPO Spatial Exposure*$_{it}$ metric quantifies the share of focal county's spatial ties measured by the inverse distance between focal and alter county directed toward counties in ERPO-adopting states in year t. A formal description of this metric is shown in equation 10 in the Methods Section.

Figure 8 presents the 95% CI of the estimated coefficient $\widehat{\theta}_1$ as shown in equation 11. Column 3 of Table 2 presents the results estimated from two-way fixed effect regression as shown in equation 11. The coefficient on *ERPO Social Exposure*$_{it}$ remains statistically significant at $\widehat{\theta}_1 = -0.298$ (cluster-robust 95% CI: [-0.475, -0.12], $p < 0.01$), while the coefficient on ERPO spatial exposure is $\widehat{\theta}_2 = 0.507$ (cluster-robust 95% CI: [-0.122, 1.14], $p < 0.10$).

To evaluate the robustness of our findings, we re-estimate the indirect social network exposure model using age-adjusted suicide mortality as the dependent variable. As shown in supplementary Figure 16 and Table 21, the coefficient on *ERPO Social Exposure*$_{it}$ remains negative and statistically significant in both the indirect social network exposure model and the robust indirect social network exposure model that additionally adjusts for *ERPO Spatial Exposure*$_{it}$. We further assess robustness by controlling for deaths in social proximity to address potential confounding from correlated suicide mortality patterns across the social network. As shown in supplementary Figure 17 and Tables 22 and 23, the coefficient on *ERPO Social Exposure*$_{it}$ remains negative and statistically significant in both the crude mortality and age-adjusted mortality specifications after including this control. Through these analyses, we show the existence of a negative association between

social ties and suicide-related events. This validates our hypothesis that ERPO social exposure is negatively associated with suicide mortality. We posit that the effect stems from the spread of information diffusion regarding suicidal norms, precaution, and events within social networks.

3.3 Materials and Methods

This study integrates multiple county level datasets to construct time-varying measures of suicide mortality, socio-spatial influence, spatial influence, social exposure, spatial exposures and county-level covariates.

3.3.1 Suicide Mortality

Data on suicide deaths were obtained from the National Vital Statistics System (NVSS), managed by the National Center for Health Statistics (NCHS). We extracted yearly counts of county-level mortality from 2010 to 2022 using the International Classification of Diseases, Tenth Revision (ICD-10) codes: `X60`–`X84`, `Y87.0`, which are conventionally used to identify deaths due to intentional self-harm. For confidentiality and consistency, all death counts were aggregated to the annual level and standardized per 100,000 population using county-level denominators from the U.S. Census Bureau.

3.3.2 Social Connectedness Index

The SCI was obtained from the Meta Data for Good program. SCI quantifies the relative strength of Facebook friendship ties between counties, normalized by population size. Specifically, SCI_{ij} reflects the likelihood of a Facebook user in county i being friends with a user in county j. This measure provides a high-resolution, empirical proxy for social ties throughout US counties. Formally, the SCI is defined in equation 1. Where, Facebook Users_i represents the number of Facebook users in the county i. $\text{Facebook Connections}_{ij}$ is the total

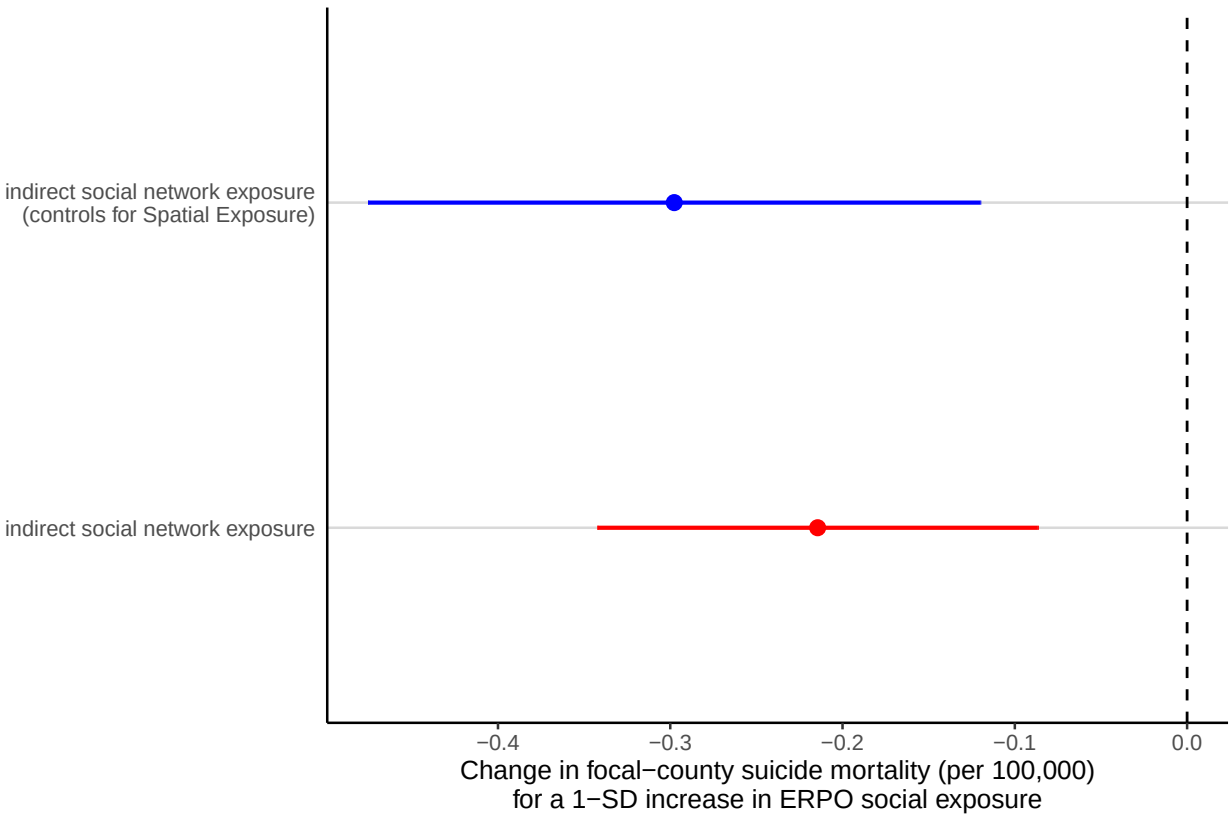

Figure 8: Estimated coefficients $(\hat{\delta}_1, \hat{\theta}_1)$ for ERPO social exposure in two specifications. Red point indicates estimate from the baseline model without spatial exposure ($\hat{\delta}_1 = -0.214$, cluster-robust 95% CI: $[-0.342, -0.0866]$); blue point indicates estimate from the specification controlling for $\textit{ERPO Spatial Exposure}_{it}$ ($\hat{\theta}_1 = -0.298$, cluster-robust 95% CI: $[-0.475, -0.120]$). Horizontal lines denote 95% confidence intervals; vertical dashed line denotes the null hypothesis ($\delta_1 = 0$). Both models include county and state-year fixed effects (ϕ_i, γ_{st}) and sociodemographic controls ($\overline{X}_{it}$). Consistent negative and statistically significant estimates indicate the association between suicide mortality and indirect social exposure to ERPO policies is robust to spatial confounding.

Table 2: Estimated effects of ERPO policy exposure on county-level suicide mortality (deaths per 100,000 people). Column (1) reports direct effects of local ERPO adoption with county and year fixed effects. Column (2) reports indirect effects of ERPO social exposure, measured through cross-county social ties, estimated with county and state–year fixed effects. Column (3) reports the indirect social exposure model with an additional control for ERPO spatial exposure, also estimated with county and state–year fixed effects, to assess robustness to spatial confounding. All models adjust for population density, age distribution (percent aged 0–17, 18–44 and 45–64), racial composition (percent Asian, Black, and Other racial subgroups), ethnicity (percent Hispanic), median household income, percent with limited English proficiency, percent unemployed, percent with less than high school education and political affiliation. Standard errors are clustered at the state level.

	Dependent variable: Deaths per 100K		
	Direct Effect	Indirect Social Network Exposure	Indirect Social Network Exposure (controls for Spatial Exposure)
	(1)	(2)	(3)
ERPO	−0.528**		
	(0.200)		
ERPO Social Exposure		−0.214***	−0.298***
		(0.064)	(0.088)
ERPO Spatial Exposure			0.507
			(0.313)
Population density	−1.453***	−0.598	−0.639
	(0.373)	(0.600)	(0.577)
Percent aged below 18	−0.103	−0.218	−0.201
	(0.361)	(0.375)	(0.377)
Percent aged 18-44	0.441	−0.214	−0.134
	(0.559)	(0.545)	(0.541)
Percent aged 45-64	−0.956***	−0.573	−0.513
	(0.331)	(0.387)	(0.399)
Percent Asian	−1.010***	−0.649**	−0.629**
	(0.338)	(0.285)	(0.270)
Percent Black	−1.323*	−2.878***	−2.792***
	(0.722)	(0.687)	(0.679)
Percent Other	0.474**	−0.109	−0.116
	(0.198)	(0.281)	(0.284)
Percent Hispanic	−3.913***	−2.697***	−2.816***
	(0.983)	(0.844)	(0.880)
Median household income	−0.658***	−0.629***	−0.597***
	(0.166)	(0.172)	(0.180)
Percent with limited English proficiency	−0.087	−0.034	−0.052
	(0.077)	(0.132)	(0.128)
Percent unemployed	0.010	−0.116	−0.101
	(0.163)	(0.124)	(0.126)
Percent with less than high school education	0.004	−0.046	−0.054
	(0.144)	(0.124)	(0.123)
Political affiliation	−0.095	0.156	0.134
	(0.167)	(0.159)	(0.159)
Observations	40,794	40,794	40,794
R^2	0.946	0.947	0.947
Adjusted R^2	0.941	0.941	0.941

Robust standard errors in parentheses. * p<0.1; ** p<0.05; *** p<0.01

number of Facebook friendship connections between individuals in counties i and j. Data were accessed for the year 2020 and assumed temporally stable over the study period.

3.3.3 Socioeconomic Covariates

Time-varying county-level sociodemographic characteristics were drawn from two sources. First, the Agency for Healthcare Research and Quality (AHRQ) provides cross-sectional community-level indicators, including education, unemployment, income, and racial/ethnic composition, for 2010–2020. Second, for years not covered by AHRQ, we accessed the American Community Survey (ACS) via the `tidycensus` package in `R`, using 5-year estimates for all U.S. counties. The harmonized covariates include population density, age structure, median income, educational attainment, unemployment rate, and proportion of residents with limited English proficiency.

3.3.4 Political Affiliation

County-level political affiliation was obtained from the MIT Election Data and Science Lab [85] for the 2008, 2012, 2016, and 2020 presidential election cycles. Political affiliation was operationalized as a binary indicator, coded as Republican-leaning (1) if the Republican candidate received a plurality of votes, and Democratic-leaning (0) otherwise. For non-election years, we assigned political affiliation based on the most recent preceding election. Alaska counties with incomplete county-level election reporting were assigned political affiliation based on district-level returns where available (FIPS codes 02013, 02016, 02020) or imputed based on statewide voting patterns for the remaining counties.

3.3.5 Socio-Spatial Influence Metrics

To quantify the influence of suicide mortality in socio-spatially connected counties, we construct two metrics: *social proximity to suicide deaths* (s_{-it}) and *spatial proximity to suicide deaths* (d_{-it}). These are formally defined as:

$$s_{-it} = \sum_{j \neq i} w_{ij} y_{jt}, \; d_{-it} = \sum_{j \neq i} a_{ij} y_{jt} \tag{4}$$

where y_{jt} denotes the suicide death rate in county j at time t. The weights w_{ij} and a_{ij} represent social and spatial proximity, respectively, defined as:

$$w_{ij} = \frac{n_j \cdot \mathrm{SCI}_{ij}}{\sum_{k \neq i} n_k \cdot \mathrm{SCI}_{ik}}, \quad a_{ij} = \frac{1/d_{ij}}{\sum_{k \neq i} (1/d_{ik})} \tag{5}$$

where n_j is the population of county j, SCI_{ij} is the Social Connectedness Index between counties i and j, and d_{ij} is the great-circle distance between them. Consistent with prior literature [69], we account for correlation between these metrics due to the spatial clustering of social networks by including both terms jointly in the model specification.

To examine the association between socio-spatial influence and suicide mortality, we estimate the following two-way fixed effects regression model:

$$y_{it} = \zeta_1 s_{-it} + \zeta_2 d_{-it} + \overline{\zeta}_3^T \overline{X}_{it} + \mu_i + \phi_t + \varepsilon_{it}, \tag{6}$$

where y_{it} is the suicide death rate in county i and year t; $\overline{X}_{it}$ is a vector of time-varying sociodemographic covariates (e.g., age distribution, racial/ethnic composition, income, education, unemployment, limited English proficiency); μ_i and ϕ_t are county and year fixed effects, respectively. Estimation is conducted using population-weighted least squares, and standard errors are clustered at the state level to account for within-state correlation in the error structure.

3.3.6 ERPO Policy Social and Spatial Exposure Metric

We estimate a two-way fixed effects model to assess the direct influence of Extreme Risk Protection Order (ERPO) policies on suicide mortality at the U.S. county level. The baseline specification is:

$$y_{it} = \psi_1 ERPO_{it} + \overline{\psi}_2^T \overline{X}_{it} + \phi_i + \gamma_t + \varepsilon_{it}, \tag{7}$$

where y_{it} denotes the suicide death rate per 100,000 population in county i and year t, and $ERPO_{it}$ is a binary indicator equal to 1 if the ERPO policy was enacted in state $s(i)$ at time t. The vector $\overline{X}_{it}$ includes socioeconomic and demographic covariates (e.g., age distribution, racial/ethnic composition, income, education, unemployment, limited English proficiency). County (ϕ_i) and year (γ_t) fixed effects control for unobserved, time-invariant heterogeneity and national temporal shocks, respectively. We exclude state-by-year fixed effects from this model, given that $ERPO_{it}$ varies at the state-year level. Estimation is performed using population-weighted ordinary least squares, and standard errors are clustered at the state level.

To quantify indirect influence transmitted through social networks, we define the following metric:

$$ERPO\ Social\ Exposure_{it} = \sum_{s(i)\neq s(j)}^{S} \mathbb{1}(\text{ERPO in state } s(j))_t \times \frac{SCI_{ij}}{\sum_h SCI_{ih}} \tag{8}$$

where SCI_{ij} denotes the Social Connectedness Index between counties i and j, and the indicator function reflects ERPO policy adoption in other states $s(j) \neq s(i)$. This standardized metric captures the share of social ties that county i maintains with counties located in ERPO-adopting states.

To estimate the association between social exposure and suicide mortality, we specify:

$$y_{it} = \delta_1 ERPO\ Social\ Exposure_{it} + \overline{\delta}_3^T \overline{X}_{it} + \phi_i + \gamma_{st} + \varepsilon_{it}, \tag{9}$$

where ϕ_i and γ_{st} represent county and state-by-year fixed effects, respectively. This specification accounts for unobserved heterogeneity across counties and time-varying factors

at the state level, including a county's own state-level ERPO implementation. The model is estimated using population-weighted OLS, with standard errors clustered at the state level to allow for arbitrary correlation within states. The inclusion of state-by-year fixed effects ensures that estimated effects are identified from variation in out-of-state ERPO policies among socially connected counties, while controlling for time-varying enforcement capacity, socioeconomic context, and other latent shocks at the state level.

As a robustness check, we define a *spatial exposure* metric that captures geographic proximity to ERPO implementation in neighboring states:

$$ERPO\ Spatial\ Exposure_{it} = \sum_{s(i)\neq s(j)}^{S} \mathbb{1}(\text{ERPO in state } s(j))_t \times \left(\frac{1/d_{ij}}{\sum_{k\neq i}(1/d_{ik})} \right) \quad (10)$$

where d_{ij} denotes the great-circle distance between counties i and j, excluding counties in the same state $s(i)$. This formulation captures whether ERPO implementation in geographically proximate counties affects suicide mortality via spatial diffusion.

To evaluate the joint influence of both social and spatial spillovers, we estimate the following model:

$$y_{it} = \theta_1 ERPO\ Social\ Exposure_{it} + \theta_2 ERPO\ Spatial\ Exposure_{it} + \overline{\theta}_3^T \overline{X}_{it} + \phi_i + \gamma_{st} + \varepsilon_{it}, \quad (11)$$

where ϕ_i and γ_{st} denote county and state-by-year fixed effects, and $\overline{X}_{it}$ represents time-varying covariates. This model allows for simultaneous estimation of the independent associations between suicide mortality and social vs. spatial channels of ERPO influence. Population-weighted estimation and state-level clustered standard errors ensure inference is robust to differential population sizes and intra-state dependence in error terms.

4.0 Socio–Spatio–Temporal Predictions of Opioid Overdose Deaths in the United States Counties

In this chapter, we analyze the socio–spatial patterns contribution in the improvement of short–horizon forecasts of U.S. county–month opioid–involved mortality and quantify the temporal scale of such influence. Using NVSS mortality for 3,138 counties (Jan 2018–Dec 2021), we predict log1p rates at horizons $h \in \{1, 2\}$ with gradient–boosted trees under three feature sets: (i) Baseline (two own monthly lags + static covariates), (ii) Baseline + Social proximity (SCI–weighted mortality in connected counties), and (iii) Baseline + Spatial proximity (inverse–distance–weighted mortality). Social proximity provides the lowest out–of–sample error across horizons, with RMSE decreasing from 0.710 to 0.702 at $h = 1$ and from 0.707 to 0.706 at $h = 2$, while the spatial variant yields little or no improvement (0.709, 0.710). To explain the modest gains, we estimate exponentially weighted moving–average (EWMA) exposures in a two–way fixed–effects design and map decay to half–lives. A 2D grid search identifies $(\alpha_s, \alpha_d) = (0.90, 0.75)$, implying half–lives of ≈ 0.30 and ≈ 0.50 months (~ 9 and ~ 15 days) and kernels concentrating $\approx 90\%$ (social) and $\approx 76\%$ (spatial) mass on the contemporaneous month. Specifications that encode near–contemporaneous exposure (EWMA or direct) explain far more within–county variation than monthly–lag models ($R^2_{\text{within}} = 0.763$–$0.770$ vs 0.004). Results indicate that network–mediated excitation operates on submonthly time scales; monthly lags under–resolve diffusion timing. We therefore motivate a continuous–time socio–spatial Hawkes formulation as a process–based extension aligned with the observed short memory and cross–location triggering.

4.1 Introduction

The U.S. opioid overdose epidemic remains a persistent public health crisis. During 2021–2023, drug–involved fatalities exceeded 100,000 annually, and opioid–involved deaths nearly doubled from 2019 to 2022 before a modest decline in 2023 [39, 89]. Mortality re-

mains historically high and unevenly distributed across places, motivating early–warning systems that identify emerging hotspots and support targeted response. Overdose risk is embedded in socio–spatial structure. Communities connected through interpersonal ties transmit risk through shared access, norms, and information flows. County–level analyses show that mortality increases with *social proximity* to high–burden areas, beyond geographic adjacency and demographics, consistent with network–mediated influence [117]. Forecasting and mapping work across multiple spatial scales–explicitly in the context of overdose and the opioid epidemic—-has demonstrated concrete mechanisms for borrowing strength over space and time and for detecting localized clustering. At the national county scale, Marks et al. (2021) [81] model U.S. drug overdose mortality (including opioid–involved deaths) with county–level predictors and neighbor–based spillover structure to rank and forecast next–year mortality, showing gains over trend–only baselines and early identification of emerging high–burden counties. At the metropolitan ZIP Code Tabulation Area scale, Sauer et al. (2021) [108] analyze opioid–involved emergency department overdose visits in the Baltimore region using Bayesian spatio–temporal hierarchical models with spatial random effects and temporal smoothing, producing risk surfaces and revealing persistent local clusters. At the state-–county scale, Acharya et al. (2022) [1] examine Virginia emergency department opioid overdose (EDOOD) visits, combining global and local spatial autocorrelation with multilevel regression to document hotspot persistence and socio–ecological correlates after adjustment. At the small–area (ZIP code) scale in Massachusetts, Bauer et al. (2023) [15] develop Bayesian dynamic spatio–temporal models that couple neighborhood (CAR–type) spatial structure with temporal dependence to improve one–year–ahead forecasts of opioid overdose burden and to delineate fine–grained hotspots. Collectively, these overdose–focused studies show that explicitly modeling spatial adjacency, temporal dependence, and hierarchical pooling improves short–horizon prediction while clarifying the geography of risk.

Recent work emphasizes decision–centric evaluation for resource prioritization, explicitly in overdose settings. Heuton et al. (2024) [53] evaluate forecasting systems by their usefulness for allocating limited public health resources (e.g., outreach, naloxone distribution), emphasizing ranking and early–warning performance over purely statistical fit and showing that models which pool information across space–time yield more actionable priority

lists. A growing literature also makes inter–community linkages explicit to capture spillovers that propagate along social connections. Liao et al. (2022) [74] develop network–aware models–including mutually exciting point processes and network–weighted features—that embed cross–county or cross–community linkages and demonstrate improved short–horizon prediction and clearer identification of diffusion pathways in OOD contexts.

A central timing challenge is that diffusion effects often attenuate rapidly. Kermack and McKendrick (1927) [63] introduce compartmental epidemic models in which infectious influence declines exponentially; Anderson and May (1991) [5] systematize this framework and connect decay to epidemiologically interpretable time scales. Hawkes (1971) [51] formalizes self–exciting point processes whose exponential kernels generate bursty, clustered events—an analogue of rapid contagion. Rasmussen et al. (2013) [103] develop Bayesian inference and kernel learning for Hawkes–type models, while Zhou et al. (2013) [126] recover sparse influence networks and excitation strengths from event data. Reinhart (2018) [105] reviews spatio–temporal Hawkes applications in crime, seismology, and epidemiology, emphasizing short–memory kernels; Arena et al. (2023) [7] highlight memory effects and rapid decay in contagion–like processes. In discrete time, an exponentially weighted moving average (EWMA) provides a practical proxy for this fast attenuation, yielding an interpretable half–life for exposure influence without committing to a continuous–time specification. This perspective connects epidemiological theory, social–contagion modeling, and pragmatic feature construction for forecasting; we adopt it empirically in the Results and formalize it in the Methods. Exponential weighting with interpretable half–life parameters has also been studied in general forecasting and time–series modelling [78].

This chapter operationalizes social connectivity at national scale using the SCI and contrasts it with geographic proximity [9]. Using a U.S. county level mortality data observed at monthly level, we evaluate short–horizon forecasting with gradient–boosted trees (XGBoost) under three feature sets: a baseline with lagged outcomes and static covariates, a baseline augmented with deaths in social proximity called as social proximity, and a baseline augmented with deaths in spatial proximity called as spatial proximity. Consistent with prior evidence on network–mediated diffusion, adding social proximity yields the lowest out–of–sample error within each architecture and horizon, though absolute gains over the baseline

and spatial–only variants are modest. We then diagnose temporal attenuation directly by smoothing exposures with EWMA and mapping the implied memory to epidemiological half–lives, finding rapid decay—on the order of days—that explains why monthly, lag–based features under–resolve diffusion timing.

Taken together, the empirical findings in this chapter—consistent yet modest forecasting gains from SCI–based social proximity and EWMA evidence of day–scale attenuation—establish the need for process–based modeling of socio–spatial spread in overdose mortality. We therefore propose a socio–spatial–temporal Hawkes process as a promising next step: a continuous–time, network–propagation framework aligned with the short–memory dynamics documented here and with prior contagion modeling (Hawkes (1971) [51]; Rasmussen et al. (2013) [103]; Zhou et al. (2013) [126]; Reinhart (2018) [105]) and overdose–focused network applications (Liao et al. (2022) [74]). This recommendation follows directly from the results and analysis presented in this chapter and is offered as a forward–looking direction for future modeling.

In summary, we document the incremental predictive value of SCI–based social proximity in national overdose forecasting, show that improvements are modest at monthly resolution because exposure influence decays within weeks, and motivate a process–based extension aligned with observed short–memory dynamics. Subsequent sections describe data and feature construction, model design and evaluation, EWMA–based decay diagnostics, and the prospective Hawkes specification.

4.2 Results

4.2.1 Predicting U.S. County Opioid Overdose Mortality with Socio–Spatial Proximity

We analyzed monthly county–level opioid overdose (OOD) mortality rates (per 100,000) from the NVSS Multiple Cause of Death files for January 2018–December 2021, yielding a county-–month panel for 3,138 U.S. counties. The prediction target was the *log–transformed*

OOD mortality rate. Forecasts were generated using Extreme Gradient Boosting (XGBoost) under three specifications: a Baseline with two monthly lags of the dependent variable and static sociodemographic covariates; a Social–proximity model augmenting Baseline with two monthly lags of mortality in socially connected counties, s_{-it}; and a Spatial–proximity model augmenting Baseline with two monthly lags of inverse–distance–weighted mortality, d_{-it}. Formal definitions for s_{-it} and d_{-it} are given in Equations.5 12 in the Methods Section. Models were trained on January 2018–February 2021 and evaluated on the remaining months for horizons $h \in \{1, 2\}$. Out–of–sample accuracy was assessed using RMSE of the natural–log–transformed rate. Across both horizons, the Social–proximity model achieved the lowest error (Table 3; Fig. 9), reducing RMSE from 0.710 to 0.702 at $h = 1$ and from 0.707 to 0.706 at $h = 2$ relative to Baseline. Absolute differences were modest (order 10^{--2} in log–RMSE), indicating that monthly autoregressive terms capture a large share of the predictable signal, while socio–spatial features provide additional but limited gains at monthly resolution. Full details of the modeling specification, feature engineering, and XGBoost hyperparameters and tuning are provided in the Methods Section.

The small but consistent gains from socially weighted features align with independent evidence that overdose risk clusters on submonthly timescales and at fine spatial radii, producing near–contemporaneous aggregation that month–to–month lags only partially capture. Spatiotemporal analyses report excess overdoses within 7–21 days and 100–500 meters following supply–side disruptions such as drug seizures [104]. Operational surveillance systems are designed accordingly: ODMAP issues spike alerts based on rolling 24–hour windows [122], and CDC's DOSE–SYS is used to "detect changes quickly and identify local drug overdose clusters" from near–real–time emergency department data [24]. Early–warning research in infectious disease surveillance has shown that exponentially weighted moving averages are well suited for short–term historical data and are particularly sensitive to small changes, an attribute that underpins systems like ESSENCE [3]. At the individual level, repeat overdoses are disproportionately concentrated within 30 days after an index event [112]. Together, these findings support the interpretation that a sizable component of socio–spatial excitation unfolds on submonthly horizons, attenuating when represented solely by discrete monthly lags. Consistent with our use of socially weighted features, our prior work shows

that SCI ties are associated with county OOD rates after adjusting for geographic distance and sociodemographic factors [117].

Motivated by these observations, the next subsection quantifies the effective memory of social and spatial exposures using exponentially weighted moving averages with tunable decay. This analysis tests whether influence dissipates too rapidly for discrete monthly lags to capture. We do not estimate Hawkes or other point–process models in this study; rather, our decay estimates provide empirical guidance for future process–based formulations that encode short–lived excitation and cross–location triggering.

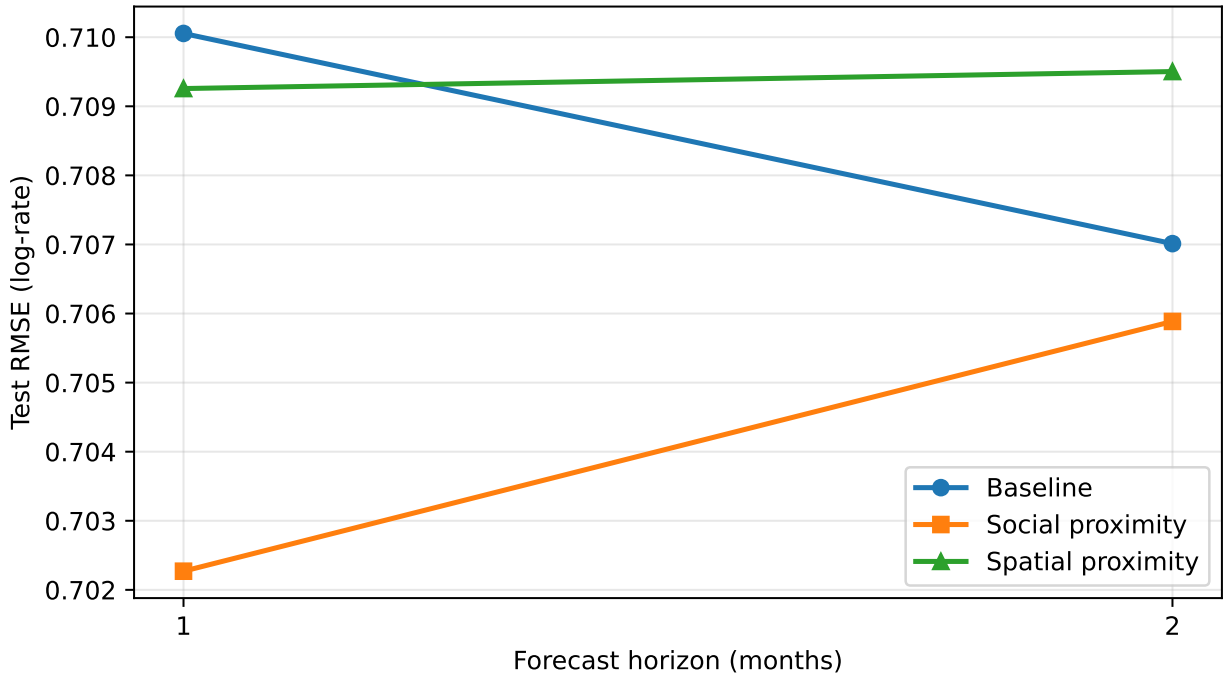

Figure 9: Out–of–sample forecast performance by model specification and horizon. Test RMSE for XGBoost models predicting log overdose death rates. The Social proximity model, which incorporates both social and spatial network features, achieves the lowest prediction error across both forecast horizons.

Table 3: Out–of–sample RMSE (log–rate) by horizon and feature set for XGBOOST. Lower is better. Bold indicates the best per horizon within each model.

Model	h	Baseline	Social proximity	Spatial proximity
XGBOOST	1	0.710	**0.702**	0.709
XGBOOST	2	0.707	**0.706**	0.710

4.2.2 Temporal Decay of Socio–Spatial Exposure

We investigated whether fixed monthly lags understate short–lived socio–spatial excitation in county–level *log* overdose rate mortality by quantifying the temporal decay of social and spatial exposures. The analysis uses the same NVSS Multiple Cause of Death panel as the forecasting exercise: monthly county observations for January 2018–December 2021 covering 3,138 counties; the outcome is the *log* mortality rate per 100,000. Social proximity s_{-it} is the SCI–weighted mortality in socially connected counties; spatial proximity d_{-it} is the inverse–distance–weighted mortality in all other counties; static covariates $\mathbf{X}_{it}$ are as defined in Methods. Prior work shows that contagion–like influence typically attenuates rapidly and approximately exponentially in continuous time: the SIR removal process yields exponential decay in infectiousness [63, 5], and self– and mutually–exciting point–process models encode the same idea by allowing recent events to temporarily elevate the likelihood of new events, with that elevation fading quickly as time passes [51, 103, 126, 7, 42, 16]. If overdose diffusion exhibits similar short memory, month–to–month lags will obscure near–contemporaneous clustering. We therefore encode temporal decay in discrete time using exponentially weighted moving averages (EWMA) of s_{-it} and d_{-it} (Eqs. 22-–23), and relate the discrete decay parameter α to a continuous–time half–life via $\gamma = -\ln(1-\alpha)$ and $h = \ln 2/\gamma$ (Eq. 24). Similar half–life interpretations appear in general exponentially weighted moving models and provide an intuitive measure of memory [78]. All regressions use a two–way fixed–effects design (Eq. 25) with county and state–year fixed effects, population weights, and state–clustered SEs.

We first constrained the two channels to share a common decay $\alpha \in \{0.1, 0.2, \ldots, 0.9\}$ while fixing the support to $K = 2$ months to align with the discrete–lag benchmark. Table 4 reports, for each α: the social and spatial coefficients with clustered SEs and t–statistics; the model's in–sample (full) R^2; and the kernel weights (w_0, w_1, w_2) on contemporaneous, one–month, and two–month components (deeper lags omitted for brevity). Fit increases monotonically with α, peaking at $\alpha = 0.9$ with $R^2 = 0.8977$. At $\alpha = 0.9$, the kernel assigns $\approx$90% of its mass to the contemporaneous month and $< 1\%$ to two months prior, indicating extremely short memory. Across all α, $\hat{\beta}_s > 0$ and highly significant, whereas $\hat{\beta}_d < 0$ and significant conditional on $\log s_{-it}$. This common–α sweep demonstrates that the data favor kernels that concentrate weight at t, consistent with rapid attenuation posited by epidemic and event–process theories.

To allow different temporal profiles for social versus spatial exposure, we performed a two–dimensional grid search over $(\alpha_s, \alpha_d) \in \{0.10, 0.15, \ldots, 0.90\}^2$. To avoid information leakage during tuning, we restricted estimation to the training window that excludes the last six months (Methods), computed EWMA features with support $K = 2$, and evaluated the in–sample R^2 of Eq. 25. Figure 10 summarizes this surface. The optimum occurs at $(\alpha_s, \alpha_d) = (0.90, 0.75)$. Using Eq. 24, these map to half–lives $h_s \approx 0.30$ months ($\sim$9 days) and $h_d \approx 0.50$ months ($\sim$15 days). The implied three–point kernels place mass (0.901, 0.090, 0.009) for social and (0.762, 0.190, 0.048) for spatial, i.e., both channels decay within weeks, with social influence attenuating faster. The R^2 surface increases monotonically in α_s and peaks at mid–high α_d, indicating asymmetric temporal memory consistent with faster social than geographic propagation.

Finally, we compared three specifications using the *within* R^2 from `fixest::feols`, which measures variation explained after absorbing fixed effects. Table 5 reports coefficients (SEs) and R^2_{within} for: (i) a contemporaneous model with $\log s_{-it}$ and $\log d_{-it}$; (ii) a lag–1 model with $\log s_{-i,t-1}$ and $\log d_{-i,t-1}$; and (iii) the EWMA–optimal model at $(\alpha_s, \alpha_d) = (0.90, 0.75)$. The contemporaneous model explains substantial residual variation ($R^2_{\text{within}} = 0.770$; $\hat{\beta}_s = 2.762\,(0.032)$, $\hat{\beta}_d = -1.816\,(0.046)$). The lag–1 model explains essentially none ($R^2_{\text{within}} = 0.004$; $\hat{\beta}_s = -0.049\,(0.014)$, $\hat{\beta}_d = 0.380\,(0.029)$), demonstrating that monthly lags capture little of the predictive signal. The EWMA–optimal specification recovers per-

formance comparable to contemporaneous exposures while explicitly encoding rapid decay ($R^2_{\text{within}} = 0.763$; $\hat{\beta}_s = 3.049\,(0.036)$, $\hat{\beta}_d = -2.244\,(0.053)$), with kernel weights and half–lives reported for reproducibility. Coefficient signs are interpreted as conditional associations given joint inclusion of social and spatial exposures and high collinearity, with predictive fit as the primary criterion.

Taken together, Table 4 shows that sharper exponential down–weighting improves in–sample fit; Figure 10 pinpoints distinct decay rates that concentrate influence within days–weeks; and Table 5 demonstrates that models which encode near–contemporaneous exposure—either directly (contemporaneous terms) or via EWMA—explain markedly more residual variation than discrete monthly lags. These findings provide reduced–form evidence that socio–spatial excitation in overdose mortality operates on submonthly timescales. While we do not estimate Hawkes processes here, the inferred half–lives and kernel mass functions supply empirical calibration targets for future process–based models that explicitly encode rapidly decaying cross–location triggering.

Table 4: Common–decay EWMA sweep ($K = 2$ months) applied simultaneously to social (s_{-it}) and spatial (d_{-it}) proximity in two–way fixed–effects models. All models include control for population density, age distribution(percentages aged 18–44, 45–64), racial composition (percentage Asian, Black and other races), ethinicity (Pecentage Hispanic), median household inclide, percentage with limited English proficiency, unemployement rate and percentage with less than a high school education county and state–year fixed effects. Residuals are population–weighted, and standard errors are clustered by state.; R^2 is the in–sample (full) value. w_0, w_1, and w_2 are the normalized kernel weights on t, $t-1$, and $t-2$, respectively. Increasing α concentrates mass at t; fit rises monotonically and peaks at $\alpha = 0.9$ ($R^2 = 0.8977$), assigning ≈ 0.90, 0.09, and 0.009 to $(t, t-1, t-2)$. Across α, the social coefficient is positive and highly significant, whereas the spatial coefficient is negative and significant conditional on joint inclusion.

α	$\hat{\beta}_s$	SE($\hat{\beta}_s$)	t_s	$\hat{\beta}_d$	SE($\hat{\beta}_d$)	R^2	Sig$_s$	w_0	w_1	w_2
0.1	3.0522	0.0361	84.63	–2.0926	0.0583	0.7001	Yes	0.10	0.09	0.081
0.2	3.3268	0.0399	83.43	–2.3252	0.0630	0.7283	Yes	0.20	0.16	0.128
0.3	3.5502	0.0430	82.57	–2.5110	0.0661	0.7591	Yes	0.30	0.21	0.147
0.4	3.6946	0.0449	82.29	–2.6253	0.0668	0.7909	Yes	0.40	0.24	0.144
0.5	3.7402	0.0453	82.57	–2.6519	0.0648	0.8216	Yes	0.50	0.25	0.125
0.6	3.6819	0.0442	83.21	–2.5892	0.0606	0.8490	Yes	0.60	0.24	0.096
0.7	3.5308	0.0421	83.93	–2.4504	0.0554	0.8714	Yes	0.70	0.21	0.063
0.8	3.3100	0.0392	84.54	–2.2594	0.0506	0.8878	Yes	0.80	0.16	0.032
0.9	3.0472	0.0359	84.96	–2.0419	0.0471	0.8977	Yes	0.90	0.09	0.009

Table 5: Model comparison for *log* overdose mortality with two–way fixed effects. All models include control for population density, age distribution(percentages aged 18–44, 45–64), racial composition (percentage Asian, Black and other races), ethinicity (Pecentage Hispanic), median household inclide, percentage with limited English proficiency, unemployement rate and percentage with less than a high school education county and state–year fixed effects. Residuals are population–weighted, and standard errors are clustered by state. R^2_{within} from `fixest::feols`.

Specification	Exposure terms	$\hat{\beta}_s$ (SE)	$\hat{\beta}_d$ (SE)	R^2_{within}	EWMA (α_s, α_d)
Contemporaneous	$\log s_{-it}$, $\log d_{-it}$	2.762 (0.032)	−1.816 (0.046)	0.770	——
Lag—-1 only	$\log s_{-i,t-1}$, $\log d_{-i,t-1}$	−0.049 (0.014)	0.380 (0.029)	0.004	——
EWMA—-optimal	$\log \widetilde{s}_{-it}(\alpha_s)$, $\log \widetilde{d}_{-it}(\alpha_d)$	3.049 (0.036)	−2.244 (0.053)	0.763	(0.90, 0.75)

Notes: $\hat{\beta}_s$ and $\hat{\beta}_d$ are coefficients on social and spatial proximity; SEs clustered by state. EWMA parameters apply only to the EWMA row. All reported coefficients are statistically significant at conventional levels: Contemporaneous $\hat{\beta}_s, \hat{\beta}_d$ ($p < 10^{-15}$); Lag–1 $\hat{\beta}_s$ ($p \approx 0.001$), $\hat{\beta}_d$ ($p < 10^{-15}$); EWMA–optimal $\hat{\beta}_s, \hat{\beta}_d$ ($p < 10^{-15}$). Half–lives implied by (α_s, α_d): $h_s = 0.30$ mo (∼9 d), $h_d = 0.50$ mo (∼15 d).

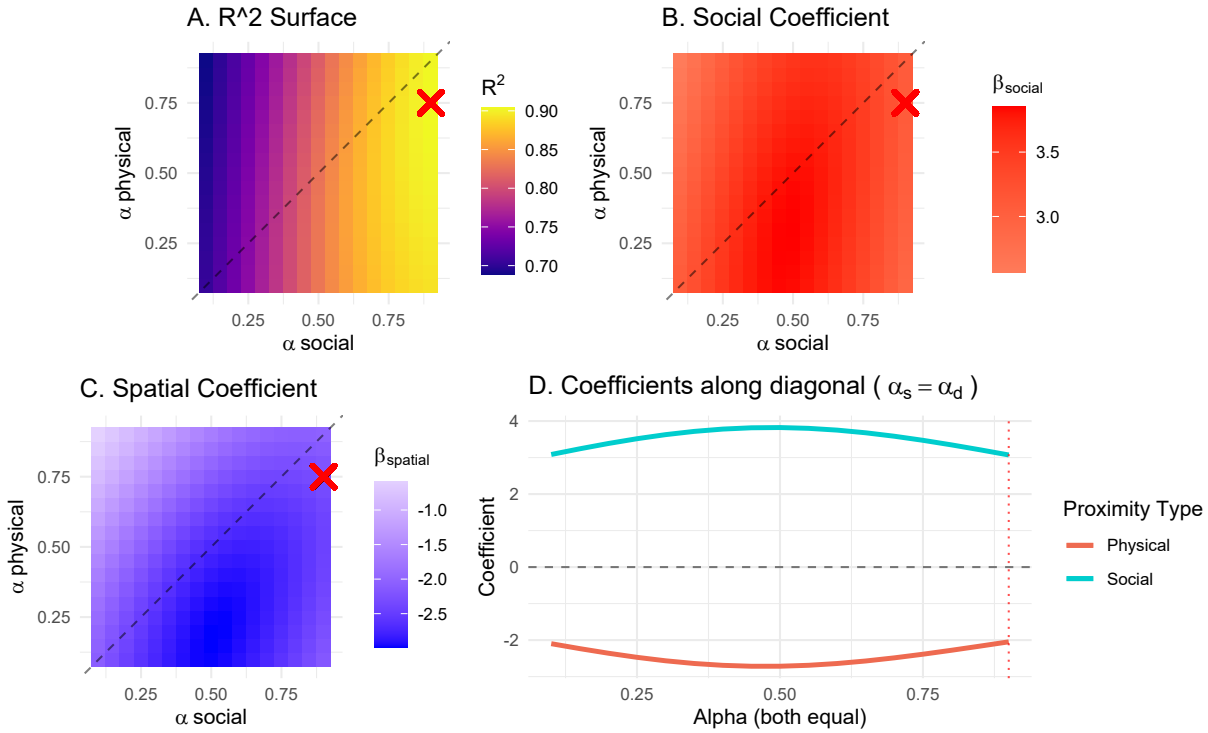

Figure 10: **Grid search over EWMA decay parameters.** Two–dimensional sweep over social and spatial decay rates (α_s, α_d) for EWMA–smoothed exposures in the two–way fixed–effects model (county and state–year FEs; population weights; state–clustered SEs). EWMA support is $K = 2$ months; tuning is performed on the training window that excludes the final six months (Methods). **(A)** In–sample R^2 across (α_s, α_d); the red "X" marks the optimum at $(0.90, 0.75)$, and the dashed line indicates the constraint $\alpha_s = \alpha_d$. **(B–C)** Estimated coefficients for the smoothed social (β_s) and spatial (β_d) exposures at each (α_s, α_d). **(D)** Coefficients along the diagonal $\alpha_s = \alpha_d$. At the optimum, the implied half–lives are $h_s \approx 0.30$ months ($\sim$9 days) and $h_d \approx 0.50$ months ($\sim$15 days), with three–point kernel weights $(w_0, w_1, w_2) \approx (0.901, 0.090, 0.009)$ for social and $(0.762, 0.190, 0.048)$ for spatial, indicating that influence is concentrated in the contemporaneous month and decays within weeks.

4.3 Materials and Methods

4.3.1 Study Design and Setting

We conduct short–horizon forecasting of U.S. county–month opioid–involved mortality for January 2018–December 2021. Indices i, j denote counties; t denotes month; $h \in \{1, 2\}$ is the forecast horizon (months ahead). We apply a log1p transformation, $y_{it}^{\log} = \log(1 + y_{it})$, i.e., a unit offset, to accommodate zero rates and avoid the $-\infty$ singularity at $y_{it} = 0$. This monotone transform reduces right–skewness; back–transformation is $y_{it} = \exp(y_{it}^{\log}) - 1$.

4.3.2 Data Sources

Monthly opioid–involved deaths and population denominators are from NVSS; county sociodemographic covariates (population density, age structure, race/ethnicity, income, education, unemployment, limited English proficiency) are compiled from ACS. Social connectivity is measured using the SCI [9]. Great–circle distances d_{ij} are computed between county centroids.

4.3.3 Proximity Measures

The SCI between counties i and j is defined in equation 1. Let n_j denote county j's population and d_{ij} the distance (miles) between i and j. The formal definition of the weight metrices are defined in equation 5.

Contemporaneous exposures average alters' outcomes, are formally defined as,

$$s_{-it} = \sum_{j \neq i} w_{ij}\, y_{jt}, \qquad d_{-it} = \sum_{j \neq i} a_{ij}\, y_{jt}, \tag{12}$$

and where y_{jt} represents the overdose deaths in in county j in month t. Further, we write $z_{-it} \in \{s_{-it}, d_{-it}\}$ for brevity.

4.3.4 Model Architecture: Predicting U.S. County Opioid Overdose Mortality with Socio–Spatial Proximity

We forecast short–horizon county–month overdose mortality at $h \in \{1, 2\}$ months ahead. The prediction target is the log1p–transformed rate per 100,000,

$$Y_{i,t+h} \equiv y^{\log}_{i,t+h} = \log\big(1 + y_{i,t+h}\big),$$

which accommodates zero rates and reduces right–skewness; the inverse map for interpretation is $y_{i,t+h} = \exp\big(Y_{i,t+h}\big) - 1$. For each horizon h, features $\mathbf{x}_{it}$ are constructed using only information available at time t (no look–ahead). Social and spatial exposures s_{-it} and d_{-it} are defined in Eqs. 12.

For horizon h, each training example is a pair $(\mathbf{x}_{it}, Y_{i,t+h})$ using only information available at time t; $\mathbf{z}_i$ denotes static county covariates. We consider three nested specifications:

$$\mathbf{x}^B_{it} = \big[y^{\log}_{i,t-1},\ y^{\log}_{i,t-2},\ \mathbf{z}_i\big] \qquad \text{(Baseline)}, \tag{13}$$

$$\mathbf{x}^S_{it} = \big[\mathbf{x}^B_{it},\ \log s_{-i,t-1},\ \log s_{-i,t-2}\big] \qquad \text{(Baseline + Social)}, \tag{14}$$

$$\mathbf{x}^D_{it} = \big[\mathbf{x}^B_{it},\ \log d_{-i,t-1},\ \log d_{-i,t-2}\big] \qquad \text{(Baseline + Spatial)}, \tag{15}$$

where s_{-it} and d_{-it} are the social– and inverse–distance–weighted contemporaneous exposures defined in Eqs. 5–12. (We report $h = 1, 2$; all features are computed without look–ahead.)

Let $\mathcal{X}$ be the feature space and $\mathbb{R}$ the target space. We learn a function $f_{\boldsymbol{\theta}} : \mathcal{X} \to \mathbb{R}$ from a parametric hypothesis class $\{f_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \Theta\}$, where $\boldsymbol{\theta}$ collects all model parameters. For Extreme Gradient Boosting (XGBoost), $f_{\boldsymbol{\theta}}$ is an additive ensemble of M regression trees,

$$f_{\boldsymbol{\theta}}(\mathbf{x}) = \sum_{m=1}^{M} \eta\, b_m\big(\mathbf{x};\, T_m, \mathbf{a}_m\big), \tag{16}$$

with learning rate η, tree structure T_m (splits/leaves), and leaf weights $\mathbf{a}_m$. The parameter vector $\boldsymbol{\theta} = \big(\eta, \{T_m, \mathbf{a}_m\}_{m=1}^{M}\big)$ fully indexes the model.

Given features $\mathbf{x}_{it}$ from one of (13)–(15), the point prediction h months ahead is

$$\widehat{Y}_{i,t+h} = f_{\boldsymbol{\theta}}(\mathbf{x}_{it}), \qquad \mathbf{x}_{it} \in \{\mathbf{x}^B_{it}, \mathbf{x}^S_{it}, \mathbf{x}^D_{it}\}. \tag{17}$$

For each horizon h, define the training set $\mathcal{D}_h^{\text{train}} = \{(\mathbf{x}_{it}, Y_{i,t+h})\}_{(i,t)\in\mathcal{I}_h^{\text{train}}}$ of size $N_h^{\text{train}} = |\mathcal{I}_h^{\text{train}}|$. We estimate $\widehat{\boldsymbol{\theta}}_h$ by minimizing regularized mean squared error on the log scale:

$$\widehat{\boldsymbol{\theta}}_h \in \arg\min_{\boldsymbol{\theta}\in\Theta} \left\{ \frac{1}{N_h^{\text{train}}} \sum_{(i,t)\in\mathcal{I}_h^{\text{train}}} \big(Y_{i,t+h} - -f_{\boldsymbol{\theta}}(\mathbf{x}_{it})\big)^2 \; + \; \Omega(\boldsymbol{\theta}) \right\}, \tag{18}$$

where $\Omega(\boldsymbol{\theta})$ is the standard XGBoost structural penalty on tree complexity (combining leaf–count and ℓ_2 penalties on $\mathbf{a}_m$). We use squared–error loss (`reg:squarederror`) as in the Results.

On the held–out test set $\mathcal{D}_h^{\text{test}}$ of size N_h^{test}, we report root mean squared error (RMSE) on the log scale:

$$\text{RMSE}_h = \sqrt{\frac{1}{N_h^{\text{test}}} \sum_{(i,t)\in\mathcal{I}_h^{\text{test}}} \big(Y_{i,t+h} - -\widehat{Y}_{i,t+h}\big)^2}, \qquad \widehat{Y}_{i,t+h} = f_{\widehat{\boldsymbol{\theta}}_h}(\mathbf{x}_{it}). \tag{19}$$

This metric aligns with the training loss (18) and with the Results tables/figures (log–scale accuracy).

The training window is January 2018—February 2021; the test window comprises subsequent months in 2021. For each (h, feature set), we perform forward–chaining time–series cross–validation with $K = 3$ folds entirely within the training window to select hyperparameters and early–stopping rounds, avoiding look–ahead (rows used to form leads/lags are excluded per horizon). Continuous features are standardized to zero mean and unit variance using statistics fit on each training fold and applied forward to its validation and the final test set.

We use histogram–based splits with row subsampling (`subsample`= 0.8) and column subsampling (`colsample_bytree`= 0.7). Tuned hyperparameters include boosting rounds $M \in \{300, 500\}$, learning rate $\eta \in \{0.01, 0.05\}$, and maximum depth $d_{\max} \in \{4, 6\}$. Early stopping is selected on validation folds; the final model is refit on the full training window with the chosen configuration and evaluated once on the held–out test months.

$\boldsymbol{\theta}$ denotes the model parameter vector (trees and leaf weights for XGBoost); N_h^{train} and N_h^{test} are the numbers of county–month observations used to train and test at horizon h; $\mathcal{I}_h^{\text{train}}$ and $\mathcal{I}_h^{\text{test}}$ are their index sets. All predictions and losses are computed on the log1p scale $Y_{i,t+h} = y_{i,t+h}^{\log}$.

4.3.5 Temporal Decay Modeling and EWMA–TWFE Estimation

The central question is whether near–real–time socio–spatial excitation is short–lived enough that discrete month–to–month lags understate its effect on county overdose mortality. Theory and evidence from epidemiology and social contagion point to rapid, approximately exponential attenuation of influence: infectiousness decays swiftly after infection in classical epidemic models [63, 5], and in self– and mutually exciting event models, recent events temporarily elevate the likelihood of new events with an effect that fades quickly as time passes [51, 103, 126, 7]. If overdoses diffuse through social and geographic channels with similar short memory, kernels that place most weight on very recent exposures should improve fit relative to coarse monthly lags.

Using the outcome and exposure definitions already given above, we estimate two two–way fixed–effects (TWFE) baselines with county fixed effects (α_i) and state–year fixed effects (γ_{st}), weighting by county population and clustering standard errors at the state level. The contemporaneous specification is

$$\log y_{it} = \alpha_i + \gamma_{st} + \beta_s \log s_{-it} + \beta_d \log d_{-it} + \mathbf{X}'_{it}\boldsymbol{\delta} + \varepsilon_{it}. \tag{20}$$

The one–month lag analogue is

$$\log y_{it} = \alpha_i + \gamma_{st} + \beta_s^{(1)} \log s_{-i,t-1} + \beta_d^{(1)} \log d_{-i,t-1} + \mathbf{X}'_{it}\boldsymbol{\delta} + \varepsilon_{it}. \tag{21}$$

To encode short–lived excitation in discrete time, we smooth each exposure with an exponentially weighted moving average over a K–month window (main analysis $K = 2$, to align with the discrete–lag benchmark while allowing intra–window down–weighting). For $\alpha \in (0, 1)$, define normalized geometric weights

$$w_k(\alpha) = \frac{\alpha(1-\alpha)^k}{\sum_{j=0}^{K} \alpha(1-\alpha)^j} = \frac{\alpha(1-\alpha)^k}{1-(1-\alpha)^{K+1}}, \qquad k = 0, \ldots, K, \tag{22}$$

and the smoothed exposure

$$\widetilde{z}_{-it}(\alpha) = \sum_{k=0}^{K} w_k(\alpha)\, z_{-i,t-k}, \qquad z_{-it} \in \{s_{-it}, d_{-it}\}. \tag{23}$$

Larger α concentrates mass at $k = 0$ (current month); smaller α spreads mass over recent months [77]. The discrete factor $(1-\alpha)^k$ maps to a continuous–time decay rate $\gamma = -\ln(1-\alpha)$ and half–life

$$h = \frac{\ln 2}{\gamma} = \frac{\ln 0.5}{\ln(1-\alpha)} \quad \text{(months)}, \tag{24}$$

providing an interpretable epidemiological time scale for memory length.

Replacing raw exposures by their EWMAs yields

$$\log y_{it} = \alpha_i + \gamma_{st} + \beta_s \log \widetilde{s}_{-it}(\alpha_s) + \beta_d \log \widetilde{d}_{-it}(\alpha_d) + \mathbf{X}'_{it}\boldsymbol{\delta} + \varepsilon_{it}. \tag{25}$$

We consider both a common–decay model with $\alpha_s = \alpha_d = \alpha$ and a channel–specific model allowing (α_s, α_d) to differ, thereby testing for asymmetric temporal memory in social versus spatial diffusion.

All regressions are estimated by OLS with fixed effects absorbed; observations are weighted by county population and standard errors are clustered by state. We report (i) the in–sample *full* R^2 (used for hyperparameter selection on training data) and (ii) the *within* R^2 (R^2_{within}) after partialling out fixed effects (used for model comparison). Implementations use `lfe::felm` and `fixest::feols` in R.

We perform two searches with $K = 2$:

1. *Common–decay sweep:* impose $\alpha_s = \alpha_d = \alpha$ and evaluate $\alpha \in \{0.1, 0.2, \ldots, 0.9\}$. For each α, we construct $\widetilde{s}_{-it}(\alpha)$ and $\widetilde{d}_{-it}(\alpha)$ and fit Eq. (25) on the full sample, tabulating coefficients, clustered SEs, t–statistics, full R^2, and the weight triplet (w_0, w_1, w_2) (Table 4).
2. *Channel–specific grid:* allow (α_s, α_d) to vary on $A = \{0.10, 0.15, \ldots, 0.90\}^2$. To avoid information leakage, we define a training window that excludes the last six months of data, compute EWMA features within training only, and evaluate Eq. (25) at each pair using the full R^2 as the tuning criterion. The R^2 surface and the associated coefficient surfaces are summarized in Figure 10. The maximizing pair $(\alpha_s^\star, \alpha_d^\star)$ is then used to refit Eq. (25) on the full sample; coefficients, R^2_{within}, implied half–lives h_s, h_d from Eq. (24), and (w_0, w_1, w_2) are reported in Table 5.

Because s_{-it} and d_{-it} are correlated exposures, $\hat{\beta}_s$ and $\hat{\beta}_d$ are interpreted as conditional associations in the joint specification with $\mathbf{X}_{it}$. Model comparison emphasizes R^2_{within} given the rich fixed–effects structure. The estimated half–lives and kernel weight profiles summarize empirically supported memory lengths and provide calibration targets for future process–based models that explicitly encode rapidly decaying cross–location triggering [51, 103, 126, 7].

5.0 Conclusions and Future Work

5.1 Summary & Discussion of Chapter 2

In **Chapter 2** our research underscores that social influence, measured by the coefficient of s_{-i}, exhibits a statistically significant impact on the spatial spread of OOD rates in US counties. This finding paves the way for a more in-depth exploration of the mechanisms driving the opioid epidemic, specifically within the social contagion framework. Harmon et al. (2015) [50] and Braha and Aguiar (2017) [18] have utilized complex generative modeling to understand social contagion in the context of economic crises and voting behavior. Building on their work in modeling social contagion, future studies can adopt a similar framework to model peer influence in the opioid epidemic using SCI data.

Our analysis demonstrates that deaths in social proximity, as measured by the SCI, are associated with opioid overdose deaths in counties in the United States. Although previous studies recognize the tendency for OODs to occur in isolated spaces [44, 73], data from the CDC's State Unintentional Drug Overdose Reporting System (SUDORS)[1] indicate that almost half of overdose deaths had a potential bystander present, suggesting that the occurrence of OODs in isolation may be less pronounced than previously thought.

Furthermore, our measure of social influence is based on the social connections derived from the social network structure of Facebook's SCI across US counties and serves as a proxy for aggregated friendship network structures and their association with OODs; however, it does not capture the strength of individual friendship ties among Facebook users in the counties. Focusing on these social connections and their relation to the spatial patterns of OODs, such as geographic discordance (the community in which overdose death occurs

[1] https://www.cdc.gov/overdose-prevention/data-research/facts-stats/sudors-dashboard-fatal-overdose-data.html

being different from the community of residence), can provide novel insights into the complex interplay of social and spatial factors in perpetuating this public health crisis.

In addressing the opioid epidemic, accounting for social networks is crucial in understanding and predicting OOD patterns at the community level to aid policy intervention. Studies using complex generative models with spatial clustering have yielded valuable insight into the dynamics of social influence and can be crucial to designing targeted interventions. For example, Braha and Aguiar (2017) [18] use such models to identify spatial clusters of voting contagion. In the context of the opioid epidemic, Liao et al. (2024) [74] introduce the Spatio-TEMporal Mutually Exciting Point Process (STEMMED) model to quantify the interconnections between historical and future events, reflecting space-time clustering in OODs across various communities. This methodological approach can improve the prediction of OOD trends within localized settings by modeling unique community-specific OOD event streams, considering spatial and social influences. Our research complements this perspective by bridging the gap in understanding the role of social influence within these dynamics. Although STEMMED captures the spatiotemporal clustering of OOD events, our approach highlights the importance of social structures in shaping these patterns. By estimating the coefficient of the socially lagged variable, we measure the effect of social interactions in different counties in the US, which in conjunction with the spatial focus of Liao et al. [74] can offer a more comprehensive view of the factors driving OOD clusters. For policymakers, this comprehensive approach can provide a solid foundation for designing targeted interventions that address both the spatial and social dimensions of the opioid epidemic.

One of our motivations for this study was to facilitate the creation of proxy networks for social interactions in agent-based models of the US population based on SCI data. Agent-based models with SCI-calibrated social networks can provide valuable information about peer effects in the initiation of opioid use and the development of use disorder. When used in conjunction with powerful agent-based modeling tools such as FRED (Framework for Reconstructing Epidemic Dynamics) [47], such models can enable epidemiologists and policymakers to simulate the spread of the opioid epidemic in great mechanistic detail and evaluate interventions. Chu et al. (2020) use FRED to study social contagion in the use of e-cigarettes [29], showing the potential of FRED to simulate social contagion of addictive be-

haviors. We expect that SCI-calibrated social networks, combined with agent-based models of OUD progression in FRED (cf., e.g., [2]), can reveal important mechanistic details about the opioid epidemic in the US. Our results may be particularly useful for calibrating social contagion parameters. The contagion parameters can be tuned so that they reflect the same association between "deaths in social proximity" and OODs as we measure in the mortality data.

SCI can also help design and evaluate intervention strategies based on social networks. Macmadu et al. (2022) highlight the need for interventions tailored to social networks [80]. Their research shows that after an overdose, network members exhibit observable behavioral changes, including increased risk taking to manage feelings of bereavement and trauma, protective actions, and some cases showing no change in drug use behavior. Based on these findings, opioid reduction efforts can be optimized in areas with greater social connectivity. Measurement of the effects of social networks across locations can also help optimize the allocation of limited medications, such as naloxone, buprenorphine, and methadone. In all these cases, social network-based strategies offer important opportunities for policy makers to mitigate the spread of the opioid epidemic.

5.2 Summary & Discussion of Chapter 3

In **Chapter 3** we model suicide mortality as a function of influence on socially and spatially proximate death rates, embedding counties within a socio-spatial system of interdependent risk. This framework enables the quantification of influence-induced increases in suicide rates through structured network-weighted exposure, one based on socio-spatial ties (s_{-it}) and the other on geographic distance (d_{-it}). The two-way fixed effects estimation model indicates that a one-standard-deviation increase in s_{-it} is associated with a statistically significant increase of 3.34 suicide deaths per 100,000 population, without controlling for spatial proximity. When controlling for spatial proximity (d_{-it}), the estimated effect of "deaths in social proximity" (s_{-it}) remains significant ($\widehat{\zeta_1} = 2.78$, SE $= 0.856$). The reduction in the estimate of s_{-it} reflects the underlying correlation between social and spatial

structures, but does not eliminate the observed influence of "deaths in social proximity". The results indicate that information about suicide related events, in this case, suicide mortality, diffuses not only along geographical proximity but also through nonlocal social ties. We speculate that this association is suggestive of a network-mediated contagion process.

We further extend our approach further to estimate the indirect effect of a policy intervention–Extreme Risk Protection Orders (ERPOs)–via social network exposure. In a two-way fixed specification with county and state-by-year fixed effects, we find that a one-standard-deviation increase in social exposure to ERPO-adopting states is associated with a reduction of 0.227 suicides per 100,000 population ($\widehat{\delta}_1 = -0.214$, SE = 0.063). This effect remains statistically significant after accounting for inverse-distance spatial exposure ($\widehat{\theta}_1 = -0.3$, SE = 0.089), further supporting the role of socio-spatial exposures. By contrast, the coefficient on spatial ERPO exposure was positive but not statistically distinguishable from zero ($\widehat{\theta}_2 = 0.549$, SE = 0.311; cluster-robust 95% CI: $[-0.08, 1.17]$). Conditional on social exposure and state–year fixed effects, the spatial ERPO exposure term carries little independent signal. We therefore place emphasis on the robust negative association for social ERPO exposure.

These associations indicate that suicide mortality and exposure to preventive policies are patterned along measured social connectedness, beyond geographic proximity. This pattern is consistent with diffusion operating through social ties, but we do not claim that networks actively transform the content or the mechanisms of transmission. The magnitude and directionality of effects depend jointly on the structure of connectivity, the nature of the exposure, and county-level sociodemographic and economic conditions (see Table 1), which we adjust for in all models. The observed associations in our study suggest that suicide risk and its mitigation are governed by system-level processes shaped by the topology of inter-county social ties. These results indicate that suicide prevention strategies should complement geography-based targeting with decentralized, locally adaptive coordination across socially connected counties, a direction already reflected in several state policy frameworks.

Existing models such as the Suicide Prevention Resource Center (SPRC) and the CDC's Community-Based Suicide Prevention (CSP) programs already recognize the importance of cross-sectoral coordination, localized coalitions, and mental health infrastructure at the

county level [22, 38]. Notably, the San Diego County Suicide Prevention Council exemplifies a multi-stakeholder, network-informed approach by convening municipalities, behavioral health providers, and faith-based institutions to form an integrated response system [93].

Current prioritization schemes typically target counties based on demographic or deprivation indicators. Our results suggest complementing these approaches with *SCI-derived network reach measures*—-for example, the SCI-weighted strength of a county's ties (the sum of its SCI weights to all other counties) or eigenvector centrality computed on the SCI matrix–to identify counties with high potential to receive or transmit risk and protective influences. Strategically positioning interventions in these counties has the potential to amplify indirect effects via social spillovers.

Network-informed infrastructures are critical for three operational reasons: (1) targeted resource deployment—such as gatekeeper training, telepsychiatry expansion, and school-based prevention—in structurally central counties can maximize indirect impact; (2) socially connected populations enable faster diffusion of behavioral norms, including help-seeking and firearm safety practices; and (3) early-warning systems built on socially connected surveillance networks can detect emerging suicide clusters before geographic clustering becomes apparent [67, 17, 125].

5.3 Summary & Discussion of Chapter 4

In **Chapter 4** we assembled a county–month panel of OOD mortality for the United States (January 2018–December 2021; 3,138 counties) and evaluated whether socio-spatial information improves short-horizon prediction of the *log*-transformed OOD rate. In forecast models (XGBoost), augmenting a strong autoregressive baseline with socio-spatial features yielded modest but consistent out-of-sample gains, with the Social-proximity specification performing best across two–month horizon ($h \in \{1, 2\}$). To diagnose the temporal scale of that predictive signal, we then quantified the decay of social and spatial exposures in a two-way fixed-effects (TWFE) framework. A common-decay sweep over exponentially weighted moving averages (EWMA) showed monotonic gains in in-sample fit as weight was concen-

trated in the contemporaneous month, peaking at $\alpha = 0.9$ (Table 4). A two-dimensional grid search that allowed distinct decays identified $(\alpha_s, \alpha_d) = (0.90, 0.75)$ (Figure 10), mapping to half-lives of ~ 9 and ~ 15 days, respectively. At those optima, the EWMA-optimal TWFE explained nearly as much within-county variation as the contemporaneous model, whereas a pure lag-1 model explained essentially none (Table 5). Together, these results indicate that the predictive content in socio-spatial exposures is realized on submonthly time scales, with social influence attenuating faster than geographic exposure to OOD.

The EWMA analysis provides a minimal yet effective representation of short-lived excitation in discrete time. First, it replaces an arbitrary choice of a single monthly lag with a *parametric* and *interpretable* temporal kernel whose mass is concentrated on recent history. Second, by relating the discrete decay parameter α to a continuous-time half-life (Equation 24), it bridges reduced-form estimation and mechanistic timing. Empirically, the common-α sweep shows that sharpening the kernel (larger α) improves fit, and the unconstrained search reveals asymmetric memory across channels (faster social decay than spatial). Third, because EWMA exposures and TWFE are computationally light and estimable with standard software, the approach is reproducible at national scale and extensible to other outcomes (e.g., stimulant-involved mortality) without bespoke point-process tooling. Importantly, our interpretation is associational rather than causal: coefficients reflect conditional correlations given joint inclusion of social and spatial exposures and high collinearity, while the decay parameters summarize the temporal window in which those associations carry predictive content. In sum, we provide national-scale evidence that socio-spatial excitation in overdose mortality is short-lived. Simple exponentially weighted kernels recover this timing and improve model performance relative to discrete monthly lags, and they furnish empirically grounded starting values for future Hawkes-process models of overdose diffusion.

5.4 Directions for Future Work of Chapter 2

For future work regarding our work as discussed in **Chapter 2** can be of the usage of SCI to investigate the social contagion of the opioid epidemic in mechanistic agent-based

models and structural causal models for causal discovery and inference. One can also plan to use SCI to improve our predictive accuracy in detecting OOD hotspots in the US. We have used SCI to measure the dynamics of social networks and their effect on the opioid epidemic. Our choice of SCI is guided by existing literature and our objective to study network effects at the population scale. We recognize the limitations of using SCI because Facebook users are not the same as the general population. SCI only acts as a proxy for real-life friendship networks, and it may not be a true representative of real-life friendship connections and their association with opioid use behavior. Other datasets that capture real-world interactions within a social network framework can also be valuable to understand the effects of social networks in the opioid epidemic. For example, the Add Health dataset [2], which includes information on best friends during high school and opioid misuse in adulthood (14 years later), offers an alternative means of measuring these effects. Although this dataset allows analysis within the age-specific bracket of 25 to 35 years, potential problems of population selection bias may limit the generalizability of the findings to the broader population.

5.5 Directions for Future Work of Chapter 3

In **Chapter 3** a central component of our study is the use of the SCI to quantify inter-county social exposure to both suicide mortality and firearm-restrictive policy interventions, specifically ERPOs. SCI captures the relative probability of Facebook friendship ties between U.S. counties, offering a scalable and validated measure of population-level social connectivity [9, 69]. Prior research has shown that SCI predicts a range of real-world outcomes including migration flows, economic mobility, and infectious disease spread, underscoring its utility as a proxy for structural social networks [9, 69]. Moreover, the SCI matrix is nearly complete: all U.S. counties have non-zero SCI values, enabling comprehensive analysis of spatially distributed social influence. In our setting, this allowed estimation of suicide mortality exposure and ERPO exposure effects through non-geographic social ties, leveraging the broad geographic coverage and empirical granularity of SCI.

[2] https://data.cpc.unc.edu/projects/2/view#public_li

However, the SCI's reliance on Facebook user data introduces critical limitations. Because SCI reflects only the activity of Facebook users, it excludes individuals who are not on the platform, particularly older adults (65+), rural residents, and those with lower socioeconomic status or limited internet access—demographic groups known to be underrepresented on Facebook [99]. For instance, only 50% of U.S. adults aged 65 and older report using Facebook, compared to 73% among those aged 18–29. Similarly, platform participation is lower among individuals with less than a high school education and those residing in rural areas. Consequently, SCI-derived exposure metrics may disproportionately reflect the social interactions of younger, more urban, and digitally connected populations, potentially underestimating the strength or pattern of influence in underserved subgroups. While this does not compromise the structural completeness of the SCI network, it raises concerns regarding representational bias and the generalizability of estimated social contagion effects. Future extensions should consider integrating auxiliary data sources—such as the National Longitudinal Study of Adolescent to Adult Health (Add Health)—to validate findings in age- or cohort-specific networks. Furthermore, methodological innovations, including dynamic or data-driven influence weighting (e.g., via latent space models or dynamic Bayesian kernels), may allow more flexible modeling of inter-county exposure beyond the fixed SCI structure.

While our current specification defines the social proximity weights w_{ij} through a parametric kernel based on SCI and population scaling, this structure imposes a fixed functional form on how social influence decays across counties. Although the current model incorporates a dynamic specification for social exposure by adjusting the strength of social influence based on suicide rates, the overall structure of the social proximity weights remains pre-specified and fixed. This parametric formulation is transparent and computationally efficient but may restrict the model's ability to reflect the true heterogeneity of influence processes, particularly in settings where socially central counties exert disproportionate or nonlinear effects on others. In future work, these limitations could be addressed by adopting a fully Bayesian hierarchical framework in which the social influence weights are not assumed *a priori*, but instead are learned from the data. For example, one could introduce latent interaction structures or flexible prior distributions that allow the strength and topology of social exposure to vary across space and time. This would permit more granular and context-sensitive

modeling of how influence propagates through the social network. Such an approach would enable the model to capture emergent patterns of contagion or resistance that may be overlooked under fixed-weight assumptions. Advancing toward data-driven, nonparametric, or semi-parametric Bayesian inference represents a rigorous path for improving the fidelity and generalizability of network-based models of suicide mortality in population health research.

One way to address this gap is to *investigate* a hierarchical Bayesian dynamic-weight specification that learns how influence concentrates across social ties rather than fixing it *a priori*. Let counties $i \in \{1, \ldots, N\}$ and years $t \in \{1, \ldots, T\}$; let Y_{it} be the suicide rate (per 100,000) in county i at year t; $\mathbf{x}_{it} \in \mathbb{R}^K$ the K-vector of covariates; SCI_{ij} the Social Connectedness Index between i and j; d_{ij} the great-circle distance (km); n_j the mean population of county j; and $\bar{Y}_{\cdot t} = N^{-1} \sum_{i=1}^{N} Y_{it}$ the national mean at t. A *candidate* dynamic concentration parameter $\kappa_t = \kappa_0 + \kappa_1 \log \bar{Y}_{\cdot t}$ (with κ_0 baseline concentration, κ_1 responsiveness to national burden) defines the pre-normalized kernel $g_{ijt} = n_j \, \mathrm{SCI}_{ij}^{\kappa_t}$ and normalized weights $w_{ijt} = g_{ijt} / \sum_{k \neq i} g_{ikt}$ (so $\sum_{j \neq i} w_{ijt} = 1$). Social and geographic lag exposures are $s_{it} = \sum_{j \neq i} w_{ijt} Y_{j,t-1}$ and $d_{it} = \sum_{j \neq i} a_{ij} Y_{j,t-1}$ with $a_{ij} = d_{ij}^{-1}$. Outcomes could be modeled as $Y_{it} \mid \cdot \sim \mathcal{N}(\theta_{it}, \sigma^2)$, $\theta_{it} = \beta_0 + \beta_s s_{it} + \beta_d d_{it} + \mathbf{x}_{it}^{\top} \boldsymbol{\beta} + \alpha_i + \gamma_t$, where β_0 is an intercept; β_s, β_d index associations with social and geographic exposure; $\boldsymbol{\beta}$ are covariate slopes; $\alpha_i \sim \mathcal{N}(0, \tau_\alpha^{-1})$ and $\gamma_t \sim \mathcal{N}(0, \tau_\gamma^{-1})$ capture county and year effects (extendable to state–year). Weakly informative priors (e.g., $\beta_0, \beta_s, \beta_d, \boldsymbol{\beta} \sim \mathcal{N}(0, 10^2)$; $\kappa_0, \kappa_1 \sim \mathcal{N}(0, 10^2)$; $\sigma^{-2}, \tau_\alpha, \tau_\gamma \sim \mathrm{Gamma}(1, 0.1)$) would regularize estimation and propagate uncertainty in κ_t; inference via HMC (NUTS) would then *explore* whether associations concentrate on stronger ties as national burden rises ($\kappa_1 > 0$) and whether social exposure remains predictive conditional on geography ($\beta_s > 0$). This route preserves interpretability through normalized w_{ijt} while allowing the strength and topology of social influence to be *learned*, offering a concrete yet non-committal roadmap that directly responds to the limitations articulated above.

5.6 Directions for Future Work of Chapter 4

In **Chapter 4** our evidence is derived from monthly aggregates and should be interpreted as associational. Monthly aggregation obscures day–level excitation and can bias identification of kernel time scales; reporting lags and measurement error in SCI, together with strong collinearity between social and spatial exposures, further affect coefficient magnitudes. These limitations motivate a process model at finer temporal resolution that separates baseline risk from triggering and explicitly represents cross–county excitation.

A natural next step is to *investigate* a multivariate Hawkes framework with three components: (i) a network kernel aligned to the Social Connectedness Index (capturing social triggering across counties), (ii) a geographic kernel decaying with distance (capturing spatial spillover), and (iii) a temporal kernel initialized at scales implied by our inferred half–lives. Baseline intensities can incorporate static covariates via log–linear or semiparametric terms; cross–kernels can encode directional excitation (social→spatial, spatial→social). Estimation may proceed via expectation—maximization or stochastic variational inference with regularization to manage dimensionality and identifiability in a 3,138—node system. Model adequacy should be assessed by time—rescaling residual tests, information criteria (AIC/BIC), and true out—of—sample prediction at weekly or daily resolution. In this roadmap, EWMA—derived half–lives provide informative priors/initializations for kernel scales, the (α_s, α_d) surface informs cross–channel asymmetry, and TWFE results allocate covariates to baseline versus excitation. This progression—from reduced–form decay to explicit self–mutually—exciting dynamics—-enables counterfactual simulations (e.g., temporary supply shocks, targeted harm–reduction deployment) and quantification of secondary event multipliers.

Appendix A Supplementary Information, Chapter 2

A.1 Additional related work

SCI data has been used in different social settings. For example, in counties where people have more concentrated social networks within the counties than outside, they also tend to have worse socioeconomic outcomes. Specifically, these counties exhibit lower average income, lower levels of education, higher rates of teenage births, lower life expectancy, less social capital, and reduced social mobility [9]. Diemer and Regan (2021) use SCI data to establish informal social interaction and highlight where social learning and knowledge flow might occur independently of geographical constraints [36]. In public health applications, Binod Khanal (2021) uses SCI to explain the variability of bottled water as avoidance behavior in the United States during the Flint water crisis [64]. Charonweng et al. (2020) use SCI to show how social networks are conduits of information about the pandemic and an economically important factor that affects compliance with and impact of mobility restrictions [26]. The findings of their work also suggest that SCI is likely to be a good proxy for real-life social connections. SCI data has also been used to understand better public health policies, for example, by studying the social network spillovers of regional lockdown policies during the pandemic [55]. Others have used SCI data to show the effect of social connections on beliefs and behaviors toward social distancing [11] and vaccination [14]. The results of Basu et al. [14] highlight that increased vaccination rates among Facebook friends correlate with higher vaccination rates in the individual's county. Vahedi et al. (2021) show the utility of SCI in spatio-temporal models to predict hot spots and incidence rates of COVID-19 at the county level in the United States [118]. In all these recent works, SCI is a primary tool for

understanding real-life social connections in various domains and acts as an essential tool to represent both economic and social interactions. As indicated in the analysis by Chetty et al.(2022) [27, 28], SCI serves as a proxy for friendship links that mediate economic mobility [27] and affect other outcomes ranging from education to health [28]. Following this growing literature, we use SCI as a proxy for real-life social connections and not as merely a metric of online connections. Through our analysis, we argue that the usage of SCI is not limited to economic outcomes and can be used to model intricate human networks that underlie public health crises such as the opioid epidemic.

A.2 Social and Spatial Proximity at the State Level and Across Counties in PA

Figures 1 and 2 display the state-level dispersion of the proximity weights and variables of interest. We formally define the aggregated state-level social and spatial proximity weights as follows:

$$w_{xz} = \frac{\sum_{i\in x}\sum_{j\in z} w_{ij}}{\sum_{z'\neq x}\sum_{i\in x}\sum_{j\in z'} w_{ij}}, \quad \text{and} \quad a_{xz} = \frac{1+\frac{1}{d_{xz}}}{\sum_{m\neq x}(1+\frac{1}{d_{xm}})},$$

where w_{ij} is the social proximity weights (defined in the main text) between county i in state x and county j in state z, z' represents the set of states excluding state x and d_{xz} represent the distance between states x and z. Similarly, we calculate "deaths in social proximity" and "deaths in spatial proximity" at the state level as follows:

$$s_{-x} = \sum_{z\neq x} w_{xz}y_z, \quad \text{and} \quad d_{-i} = \sum_{z\neq x} a_{xz}y_z,$$

where y_z is the death rate per $100,000$ in state z. It is important to note that our analysis of the effect of "deaths in social proximity" on overdose deaths is conducted at the county level. Due to the unavailability of state-level Social Connectedness Index (SCI) measures, we aggregated the social proximity weights for visualization of state-level variables in Figures 1 and 2. Using state-level outcomes in these figures allows us to observe how the aggregated county units contribute to differences across states. This approach also aids in visualizing the

network layout, which is otherwise infeasible given the density of the connectivity between counties.

For comparison, we also visualize the layout of social networks and the dispersion of proximity weights between counties in Pennsylvania (PA). Figure 11A illustrates the social network at the county level measured by their social proximity weights (w_{ij}) in PA counties. Figures 11B and 11C show the spatial dispersion of the proximity weights for the socially and spatially lagged variables for two ego counties, Allegheny (11B) and Philadelphia (11C). In Figure 12, we show the spatial distributions of our primary variables of interest: the OOD rates, s_{-i}, d_{-i}, and their differences $(s_{-i} - d_{-i})$, in PA counties.

A.3 Covariate Selection

The opioid epidemic is a complex interplay between social networks and socioeconomic status. Studies have shown that prescription opioids became a means of establishing and maintaining social capital within social networks in low-socioeconomic-status communities [33]. Similarly, Gupta et al. (2018) [34] underscore the relationship between the opioid epidemic and its association with socioeconomic factors. To control the influence of socioeconomic factors on "deaths in social proximity", we used data on social determinants of health (SDOH) available from the Agency for Healthcare Research and Quality (AHRQ). This data set includes information on the physical and health infrastructure and socioeconomic indicators for each county in the United States. We further select the most relevant SDOH covariates using the Least Absolute Shrinkage and Selection Operator (LASSO) to avoid overfitting and multicollinearity in our model. To incorporate domain knowledge variables informed by literature, we specifically use the partially penalized LASSO regression [41].

The LASSO objective function for the model we utilize for the analysis:

$$\underset{\beta}{\text{minimize}} \left[\frac{1}{2n} \sum_{i=1}^{n} (y_i - \sum_{j=1}^{p} x_{ij}\beta_j)^2 + \lambda \sum_{j=1}^{p} \pi_j |\beta_j| \right]$$

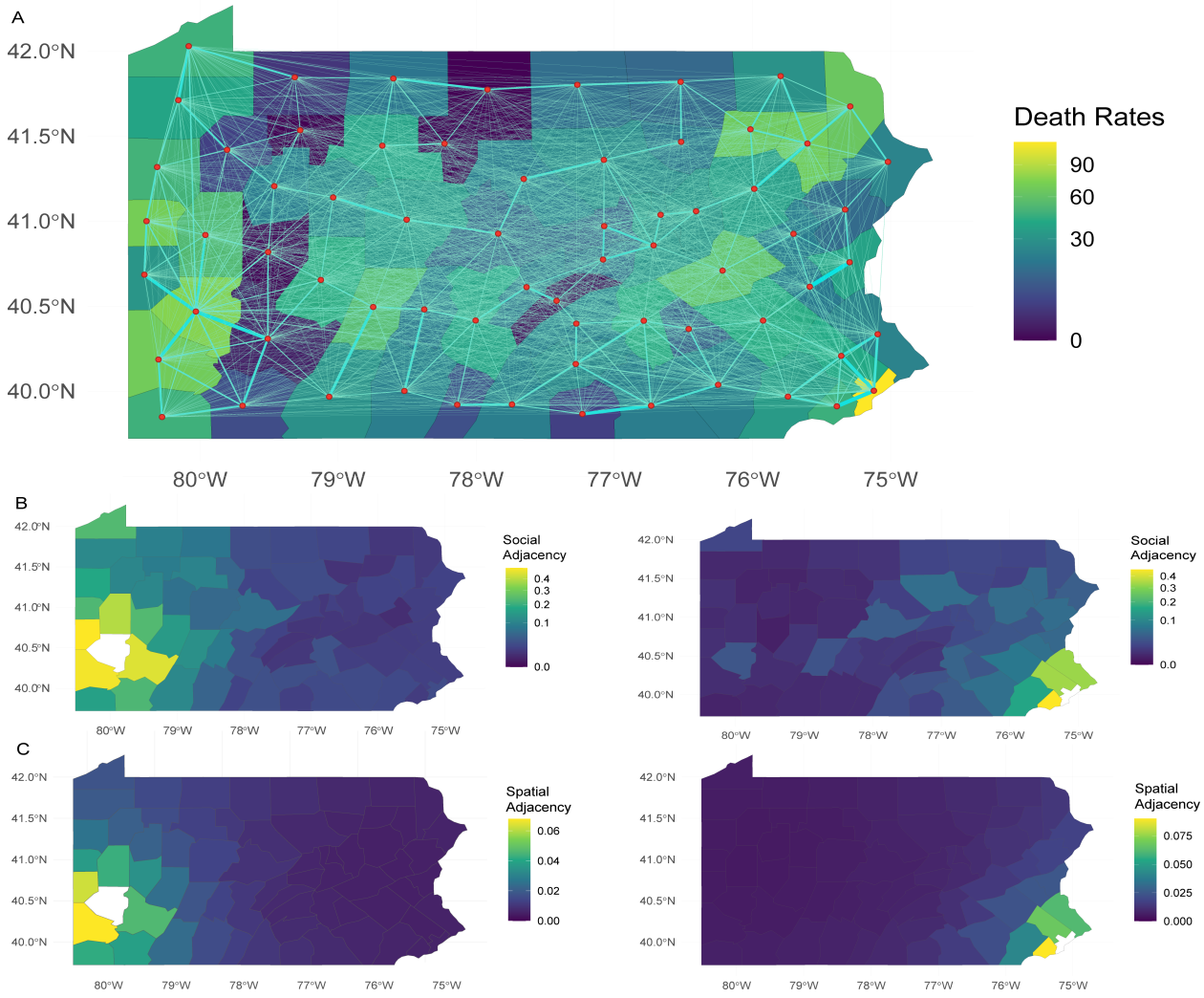

Figure 11: A) The spatial distribution of overdose death rates in 67 counties in Pennsylvania (PA) in 2018 and 2019. The map highlights the higher proportion of opioid overdose death rate in Philadelphia County. Superimposed on this map is a social network diagram with edge widths representing the social proximity weights (w_{ij}). B) The two middle maps show the social proximity weights of alter counties to Allegheny County (on the left) and Philadelphia County (on the right). C) The bottom two maps show the spatial proximity weights of alter counties to Allegheny County (on the left) and Philadelphia County (on the right).

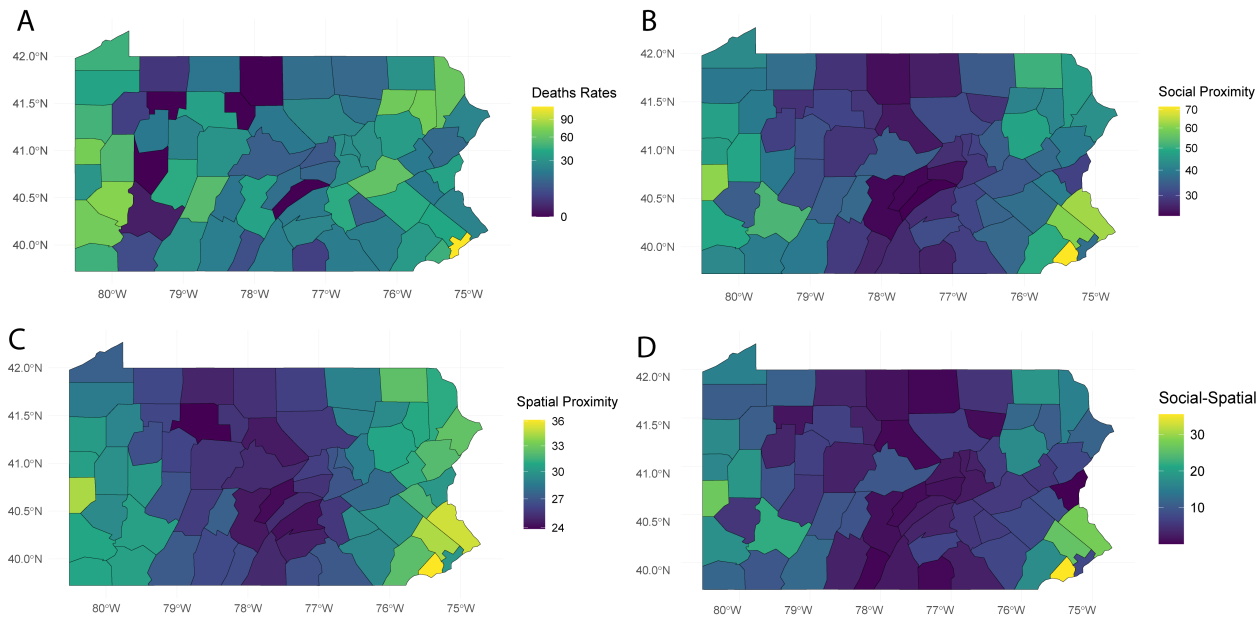

Figure 12: A) The top left map shows the spatial spread of opiod overdose death rates across PA counties. B) The top right map shows the spatial dispersion of death in social proximity (s_{-i}) for PA counties. C) The bottom left map shows the geographical spread of "deaths in spatial proximity" (d_{-i}) for PA counties. D) The bottom right map shows the difference between deaths in social and spatial proximity ($s_{-i} - d_{-i}$) in PA counties.

Here β is the vector of regression coefficients, x is the design matrix containing SDOH variables, π is an indicator variable for the penalized factor, if $\pi_j = 0$ then the corresponding covariates are not reduced to zero ensuring that they stay in the model and if $\pi_j = 1$ their corresponding covariates are subject to shrinkage, enabling variable selection. We use this LASSO framework to ensure that covariates informed by domain knowledge remain in the regression model, while the LASSO selection operates on other SDOH covariates. Specifically, variables such as the Opioid Dispensing Rate (ODR), naloxone availability, buprenorphine availability, population density, mental health distress rate, political affiliation, and median household income are considered crucial based on domain expertise and are therefore not penalized. The remaining covariates listed in Supplementary Table 6 are subject to selec-

tion through the LASSO method. In the model y is the vector of the response variable "death rates" aggregated over four years (2013-2017) for counties in Pennsylvania. The data source is from the Commonwealth of Pennsylvania Department of Health, Bureau of Health Statistics and Registries. The data source lacks information on 'T codes'; consequently, the LASSO selection for SDOH covariates utilizes overall overdose death rates instead of specifically targeting opioid overdose death rates. Here, n is the number of observations and p is the number of predictors. Performing selection using a different dataset and outcome variable helps avoid post-selection inference issues, ensuring the confidence interval derived from our models has a valid interpretation. The bias introduced by the L_1 penalty term forces some regression coefficients to shrink to zero. The zero coefficients are removed from the model, which serves as a way to perform variable selection. To estimate the optimal value of the tuning parameter λ, we performed a k-fold cross-validation ($k = 6$), using the mean squared error (MSE) performance metric.

Supplementary Table 6 shows the set of coefficient estimates of SDOH variables after selection of LASSO. The selected variables indicate their significance for the death rates. Thus, this Supplementary Table collectively provides a comprehensive understanding of the influential variables that play a significant role in OODs. To isolate the effect of "deaths in social proximity" on opioid overdose deaths, we used the variables selected by LASSO regression (see Table 6) as controls in our model. We excluded state-level data on the count of illicit opioids reported from the variable selection pool because this measure lacks variation within Pennsylvania counties; consequently, LASSO would automatically drop this variable even with penalized regression. In Supplementary Table 6, a "–" indicates variables that were not selected by the LASSO regression. When analyzing the National Center for Health Statistics (NCHS) data to estimate the effect of "deaths in social proximity" on OODs, we included only the covariates selected by the LASSO, along with state-level control. It is important to note that in our analysis, counties lacking reported values for clinical covariates, specifically the availability of naloxone and buprenorphine, are assumed to have zero values. This assumption is based on the observation that these counties recorded negligible overdose deaths during the years 2018-2019 and have very small populations. In the next section, we discuss the specifics of the robustness checks conducted in our study.

Table 6: Columns one to four list SDOH variable names, their descriptions, LASSO model coefficients (- values are dropped post-LASSO), and the corresponding covariates in our models.

	Variable	Description	Coefficient	Covariates
1	ACS_TOT_POP_POV	Total population for whom poverty status is determined	-	poverty status
2	AMFAR_MHFAC_RATE	Total number of facilities that provide mental health services per 1,000 population	-	mental health service rate
3	ACS_PCT_HU_NO_VEH	Percentage of housing units with no vehicle available	-	percentage no vehicle
4	ACS_PCT_LT_HS	Percentage of population with less than high school education (ages 25 and over)	-	education level (below high school)
5	POS_MEAN_DIST_ALC	Mean distance to nearest hospital with alcohol/drug abuse care, using pop-weighted centroids in county	-7.518093×10^{-5}	mean distance to hospital
6	ACS_PCT_OTHER_INS	Percentage with other health insurance coverage	-	health insurance coverage
7	CCBP_TOT_BWLSTORES_Rate	Number of beer, wine, liquor stores per 1,000 population	-	alcohol stores per 1,000 Pop
8	AHRF_TOT_COM_HEALTH_GRANT	Total community health centers, grantees only	-	total community health centers
9	ACS_PCT_UNEMPLOY	percentage unemployed (civilian labor force, ages 16+)	6.249815×10^{-5}	percentage unemployed
10	ACS_MEDIAN_HH_INC	median household income	8.190151×10^{-5}	median household income
11	ACS_MEDIAN_AGE	Median age of population	-	median age
12	ACS_PCT_WHITE	Percentage of population white	-	percentage White
13	ACS_PCT_BLACK	Percentage of population black	-	percentage Black
14	ACS_PCT_AIAN	Percentage American Indian and Alaska Native	-2.546380×10^{-5}	percentage AIAN
15	ACS_PCT_NHPI	Percentage Native Hawaiian and Other Pacific Islander	1.307321×10^{-5}	percentage NHPI
16	ACS_PCT_MULT_Race	Percentage two or more races	-	percentage multi-race
17	ACS_PCT_ASIAN	Percentage of population Asian	-	percentage Asian
18	ODR	Opioid Dispensing Rate	1.993769×10^{-4}	ODR
19	Naloxone_Available	Availability of naloxone	7.221718×10^{-4}	naloxone available
20	buprenorphine_Available	Availability of buprenorphine	6.462634×10^{-3}	buprenorphine available
21	population_density	Population per square kilometers	-1.385327×10^{-8}	population density
22	frequent_mental_health_distress	Rate of frequent mental health distress	3.000371×10^{-3}	mental health distress rate
23	political_affiliation	political affiliation measure	2.921279×10^{-4}	political affiliation

A.4 Robustness checks

As robustness checks, we set up linear regression, spatial and network autocorrelation, two-way fixed-effect models, two-stage least squares estimates. These steps help us ensure consistent reporting and further bolster our claims regarding the significance of our socially lagged variable.

A.4.1 Linear Regression with Cluster–Robust Standard Error Correction

As a first robustness check model, we utilize a simple linear regression as shown in the equation:

$$y_i = \beta_0 + \beta_1 s_{-i} + \beta_2 d_{-i} + \overline{\beta_3}^T \overline{C}_i + \overline{\beta_4}^T \overline{X}_i + \epsilon_i$$

Here, β_1 is the coefficient for "deaths in social proximity", β_2 is the coefficient for "deaths in spatial proximity", $\overline{\beta_3}^T$ denotes the vector of the coefficients associated with clinical covariates and $\overline{\beta_4}^T$ is the vector of the coefficient associated with SDOH variables. The model allows us to estimate the effect of our socially lagged variable on death rates.

To ensure that the inference is reflecting the population residing in the counties across the US, we weight the residuals by population size. Weighing the residuals also enhances the ability of the model to capture the heterogeneity between different spatial units. In our simple linear model, we ensure that we account for the clustering of standard errors using cluster-robust standard error correction.

Supplementary Tables 7, 8, and 9 illustrate the results of the linear regression model for the eastern, western-central, and contiguous United States. The socially lagged variable remains significant. However, we lose the statistical significance for "deaths in spatial proximity" when the geographical focus shifts from counties in the eastern US to counties in the contiguous US. Our previous discussion has highlighted the underlying spatial and network structure of our variable of interest. Spatial and social variables are often susceptible to correlated error terms because of their autocorrelated construction. Hence, as a mitigating measure, we use the spatial error model to account for network and spatial autocorrelation in the model. In the next section, we discuss the network autocorrelation model.

Table 7: Linear regression for measuring the significance of the effect size of "deaths in social proximity", for counties in the eastern United States

	Dependent variable: death rate per 100,000 people	
	OLS	*Cluster-Robust OLS*
	(1)	(2)
deaths social proximity	9.187^{***}	9.187^{***}
	(0.978)	(2.414)
deaths spatial proximity	5.462^{***}	5.462^{***}
	(0.921)	(2.021)
ODR	54.106^{***}	54.106^{**}
	(12.571)	(23.513)
naloxone available	162.003^{***}	162.003^{*}
	(17.037)	(85.602)
buprenorphine available	42.685^{***}	42.685
	(13.212)	(40.431)
state count illicit opioid reported	$6{,}303.967^{**}$	6,303.967
	(3,193.753)	(8,982.550)
population density	-0.001^{***}	-0.001^{*}
	(0.0002)	(0.001)
mental health distress rate	182.710^{**}	182.710
	(77.001)	(185.534)
political affiliation	-3.649^{***}	-3.649
	(1.340)	(2.706)
percentage unemployed	68.805^{***}	68.805^{**}
	(9.677)	(29.286)
mean distance to hospital	-72.700^{***}	-72.700^{***}
	(7.093)	(11.279)
median household income	-24.389^{***}	-24.389
	(6.898)	(15.699)
percentage AIAN	-41.111	-41.111^{*}
	(35.741)	(23.164)
percentage NHPI	-15.308	-15.308
	(12.245)	(13.041)
constant	2.770	2.770
	(10.378)	(26.515)
observations	1,606	
R^2	0.464	
adjusted R^2	0.460	
residual std. error	3.179 (df = 1590)	
F statistic	94.409^{***} (df = 14; 1590)	
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 8: Linear regression for measuring the significance of the effect size of "deaths in social proximity", for counties in the western and central United States

	Dependent variable: death rate per 100,000 people	
	OLS	*Cluster-Robust OLS*
	(1)	(2)
deaths social proximity	5.768***	5.768***
	(0.340)	(0.637)
deaths spatial proximity	0.128	0.128
	(0.528)	(1.873)
ODR	17.101***	17.101**
	(3.847)	(8.698)
naloxone available	69.584***	69.584**
	(10.675)	(30.149)
buprenorphine available	43.633***	43.633**
	(8.586)	(19.757)
state count illicit opioid reported	6,468.861	6,468.861
	(4,288.239)	(9,546.581)
population density	0.006***	0.006***
	(0.0005)	(0.001)
mental health distress rate	58.461	58.461
	(38.647)	(147.098)
political affiliation	0.578	0.578
	(0.723)	(2.070)
percentage unemployed	−6.495	−6.495
	(6.179)	(9.489)
mean distance to hospital	−15.183***	−15.183
	(3.604)	(10.792)
median household income	−4.777	−4.777
	(3.467)	(10.669)
percentage AIAN	−7.516	−7.516
	(6.795)	(12.271)
percentage NHPI	−2.114	−2.114
	(3.862)	(6.128)
constant	0.926	0.926
	(5.349)	(22.011)
observations	1,502	
R^2	0.459	
adjusted R^2	0.454	
residual std. error	1.306 (df = 1486)	
F statistic	90.049*** (df = 14; 1486)	
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 9: Linear regression for measuring the significance of the effect size of "deaths in social proximity", for counties in the contiguous United States.

	Dependent variable: death rate per 100,000 people	
	OLS	*Cluster-Robust OLS*
	(1)	(2)
deaths social proximity	12.503^{***}	12.503^{***}
	(0.706)	(2.163)
deaths spatial proximity	1.132^{*}	1.132
	(0.648)	(1.526)
ODR	49.014^{***}	49.014^{***}
	(8.187)	(18.479)
naloxone available	153.300^{***}	153.300^{***}
	(11.252)	(57.614)
buprenorphine available	56.048^{***}	56.048
	(10.038)	(41.411)
state count illicit opioid reported	$11{,}396.620^{***}$	11,396.620
	(2,391.373)	(8,899.337)
population density	0.0003^{*}	0.0003
	(0.0002)	(0.001)
mental health distress rate	21.961	21.961
	(45.829)	(119.935)
political affiliation	-3.589^{***}	-3.589^{*}
	(0.829)	(1.893)
percentage unemployed	40.118^{***}	40.118^{*}
	(6.755)	(23.199)
mean distance to hospital	-50.508^{***}	-50.508^{***}
	(5.560)	(12.930)
median household income	-17.588^{***}	-17.588
	(4.330)	(10.696)
percentage AIAN	-17.954	-17.954
	(10.987)	(13.417)
percentage NHPI	-9.727^{*}	-9.727
	(5.689)	(10.393)
constant	9.849	9.849
	(6.194)	(16.294)
observations	3,108	
R^2	0.502	
adjusted R^2	0.500	
residual std. error	1.899 (df = 3092)	
F Statistic	222.608^{***} (df = 14; 3092)	
Note:		*p<0.1; **p<0.05; ***p<0.01

A.4.2 Network Autocorrelation

Autocorrelation, in essence, refers to the tendency of a variable to be correlated with itself at different points in space or time. In the context of our study, we anticipate encountering correlated error terms when analyzing death rates related to social proximity. This likelihood comes from the inherent structural characteristics and the defining attributes of our variable of interest. Thus, we utilize the Spatial Error Model (SEM) to account for spatial and network autocorrelation. SEM is a specific linear regression model that accounts for spatial dependence in error terms. The so-called network autocorrelation model for our study is defined by, $y_i = \gamma_0 + \gamma_1 s_{-i} + \gamma_2 d_{-i} + \overline{\gamma_3}^T \overline{C}_i + \overline{\gamma_4}^T \overline{X}_i + u_i$, where γ_1 represents the coefficient associated with "deaths in social proximity", γ_2 illustrates the coefficient size related to "deaths in spatial proximity", $\overline{\gamma_3}^T$ denotes the vector of the coefficient associated with clinical covariates, and $\overline{\gamma_4}^T$ be the vector of the coefficient associated with SDOH variables. The error term u is assumed to be spatially autocorrelated, which means that geographically close observations are more likely to have similar error terms than distant observations .

To account for autocorrelation in error terms, the SEM assumes that error terms can be decomposed into two components: an autocorrelated error term u and a noise error term $\bar{\epsilon}$, where $\bar{\epsilon} \sim N(0, \sigma^2 I)$. The autocorrelated error term is modeled as a linear combination of the error terms of neighboring observations , weighted by a spatial weight matrix W. The error model is given by:

$$u_i = \lambda \sum_j w_{ij} u_j + \epsilon_i.$$

In the weight matrix W, each element

$$w_{ij} = \frac{n_j \mathrm{SCI}_{ij}}{\sum_{k \neq i} n_k \mathrm{SCI}_{ik}},$$

where λ is a parameter that captures the strength of network autocorrelation, and W calculates the weight for the connection between two areas i and j in a network. The noise in error term ϵ is assumed to be independently and identically distributed (IID) with mean zero and variance σ^2. To estimate the model, we utilize the `spatialreg` package in R, specifically employing the `errorsarlm` function. The function implements maximum-likelihood estimation

for spatial simultaneous autoregressive error models. The parameter λ is initially estimated using the `optimize` function. Subsequently, the β and other parameters are estimated using generalized least squares. Same estimation method is used when accounting for spatial autocorrelation model.

The findings of the network autocorrelation model are shown in the Supplementary Tables 10, 11 and 12 for the eastern, western-central and contiguous United States. To further solidify our analysis, we account for spatial dependence employing spatial autocorrelation.

A.4.3 Spatial Autocorrelation

From a methodological perspective, the network and spatial autocorrelation share many similarities. However, a key difference is induced through the definition of weights. The weights are designed to encapsulate the underlying network structure in a network autocorrelation model. In contrast, in the case of spatial autocorrelation, the weights are formulated to capture the geographical dependencies.

The error model is defined as follows, where a_{ij} is the spatial weights,

$$u_i = \lambda \sum_j a_{ij} u_j + \epsilon_i \,,$$
$$a_{ij} = \frac{1 + \frac{1}{d_{ij}}}{\sum_{k \neq i} (1 + \frac{1}{d_{ik}})} \,.$$

It is important to note that the equations are used similarly for both the network and the spatial autocorrelation models. However, within the scope of this study, the only distinction between the two models is the definition of these weights. Supplementary Tables 10, 11 and 12 show the results of the autocorrelation model for the eastern, western-central, and contiguous United States. This result reinforces the validity of our findings derived from simple linear regression and supports the significance obtained for s_{-i}. The models we have considered thus far account for heterogeneity across spatial boundaries but do not fully capture variations across space and time. Therefore, as a final robustness check, we will incorporate a two-way fixed effect model into our analysis to better capture these nuances.

Table 10: Autocorrelation Models for the counties in the eastern United States

	Dependent variable: death rate per 100,000 people	
	Network	*Spatial*
	(1)	(2)
deaths social proximity	18.791^{***}	13.314^{***}
	(0.644)	(0.894)
deaths spatial proximity	-2.335^{***}	1.296
	(0.624)	(1.035)
ODR	47.971^{***}	53.261^{***}
	(6.504)	(7.209)
naloxone available	31.139^{***}	31.797^{***}
	(10.486)	(11.219)
buprenorphine available	7.028	17.375^{***}
	(5.426)	(6.631)
state count illicit opioid reported	−1,945.999	−130.128
	(1,756.263)	(2,477.706)
population density	0.001^{**}	0.002^{**}
	(0.001)	(0.001)
mental health distress rate	-160.796^{***}	−89.597
	(42.584)	(56.496)
political affiliation	-6.124^{***}	-4.957^{***}
	(1.136)	(1.330)
percentage unemployed	7.073	6.073
	(4.791)	(5.503)
mean distance to hospital	-19.737^{***}	-26.527^{***}
	(3.362)	(3.983)
median household income	-22.393^{***}	-27.317^{***}
	(5.596)	(7.087)
percentage AIAN	47.620^{***}	45.101^{***}
	(13.005)	(14.542)
percentage NHPI	1.963	4.928
	(6.104)	(6.411)
constant	53.460^{***}	43.923^{***}
	(6.393)	(8.606)
observations	1,606	1,606
log likelihood	−6,931.226	−6,944.193
σ^2	322.701	333.249
λ	−0.81195	0.68115
Akaike Inf. Crit.	13,896.450	13,922.390
Wald test (df = 1)	65.324^{***}	11.167^{***}
LR Test (df = 1)	28.489^{***}	2.554
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 11: Autocorrelation models for the counties in the western and central United States

	Dependent variable:	
	death rates	
	Network	*Spatial*
	(1)	(2)
deaths social proximity	7.284^{***}	6.454^{***}
	(0.405)	(0.566)
deaths spatial proximity	0.058	0.400
	(0.473)	(0.662)
ODR	6.515^{***}	6.862^{***}
	(2.436)	(2.559)
naloxone available	-1.754	-3.460
	(5.690)	(5.905)
buprenorphine available	24.510^{***}	30.127^{***}
	(5.681)	(6.021)
state count illicit opioid reported	568.756	712.539
	(2,520.982)	(3,707.215)
population density	0.009^{***}	0.010^{***}
	(0.001)	(0.002)
mental health distress rate	-81.327^{***}	-54.738^{*}
	(23.426)	(30.530)
political affiliation	-1.077	-0.716
	(0.980)	(1.093)
percentage unemployed	6.172^{*}	5.312
	(3.741)	(4.075)
mean distance to hospital	-1.382	-1.576
	(2.355)	(2.587)
median household income	-7.234^{**}	-5.256
	(3.648)	(4.007)
percentage AIAN	7.932^{**}	6.295
	(3.299)	(4.160)
percentage NHPI	-0.985	-0.638
	(4.439)	(4.760)
constant	18.843^{***}	14.717^{***}
	(3.716)	(4.621)
observations	1,502	1,502
log likelihood $-5{,}860.207$	$-5{,}868.434$	
σ^2	142.595	144.851
λ	-0.54591	0.52449
Akaike Inf. Crit.	11,754.410	11,770.870
Wald test (df = 1)	22.991^{***}	3.504^{*}
LR test (df = 1)	18.083^{***}	1.629
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 12: Autocorrelation models for counties in the contiguous United States

	Dependent variable: death rate per 100,000 people	
	Network	*Spatial*
	(1)	(2)
deaths social proximity	17.894^{***}	14.517^{***}
	(0.509)	(0.671)
deaths spatial proximity	-3.386^{***}	-1.516^{*}
	(0.481)	(0.795)
ODR	36.510^{***}	39.822^{***}
	(4.320)	(4.596)
naloxone available	21.914^{***}	19.275^{***}
	(6.056)	(6.314)
buprenorphine available	13.723^{***}	26.807^{***}
	(4.535)	(5.338)
state count illicit opioid reported	363.577	1,052.167
	(1,446.788)	(1,916.814)
population density	0.003^{***}	0.003^{***}
	(0.0005)	(0.001)
mental health distress rate	-124.751^{***}	-92.180^{***}
	(23.424)	(30.657)
political affiliation	-4.311^{***}	-4.086^{***}
	(0.790)	(0.892)
percentage unemployed	1.457	1.790
	(3.439)	(3.787)
mean distance to hospital	-6.814^{***}	-9.530^{***}
	(2.585)	(2.948)
median household income	-12.265^{***}	-14.344^{***}
	(3.653)	(4.286)
percentage AIAN	20.064^{***}	17.316^{***}
	(3.815)	(4.736)
percentage NHPI	−0.048	1.835
	(5.213)	(5.787)
constant	35.723^{***}	32.721^{***}
	(3.625)	(5.584)
observations	3,108	3,108
log likelihood	−13,021.780	−13,025.660
σ^2	253.641	255.359
λ	−0.58906	0.90999
Akaike Inf. Crit.	26,077.560	26,085.310
Wald test (df = 1)	47.679^{***}	214.894^{***}
LR test (df = 1)	24.566^{***}	16.817^{***}
Note:		*p<0.1; **p<0.05; ***p<0.01

A.4.4 Two–Way Fixed–Effect Model

The two-way fixed-effects model accounts for cross-sectional heterogeneity between counties in the US and periods (year). It controls for any unobserved space- and time-invariant characteristics that may be correlated with the covariates, furthering more accurate and unbiased estimates. Mathematically, we can express the two-way fixed effect model by

$$y_{it} = \zeta_0 + \zeta_1 s_{-it} + \zeta_2 d_{-it} + \overline{\zeta_3}^T \overline{C}_{it} + \overline{\zeta_4}^T \overline{X}_{it} + \ \mu_i + \phi_t + \epsilon_{it}.$$

Here, μ_i and ϕ_t account for spatial and temporal heterogeneity. Supplementary Tables 13, 14 and 15 show the result obtained from the two-way fixed effect model in the eastern, western-central and the entire contiguous US. After controlling for covariates from SDOH, "deaths in spatial proximity", and clinical covariates, we still witness statistically significant effect sizes for "deaths in social proximity."

Table 13: Two-way fixed-effect model for counties in the eastern United States

	Dependent variable: death rate per 100,000 people
deaths social proximity	3.366^{***}
	(0.428)
deaths spatial proximity	3.078^{***}
	(0.784)
ODR	16.025
	(32.745)
naloxone available	$7{,}593.117^{***}$
	(748.879)
buprenorphine available	0.002^{***}
	(0.0001)
state count illicit opioid reported	$-1{,}220.069$
	(3,935.194)
population density	-0.001^{***}
	(0.0001)
mental health distress rate	351.109^{***}
	(35.178)
political affiliation	1.988^{***}
	(0.475)
percentage unemployed	16.336^{***}
	(3.852)
mean distance to hospital	-23.773^{***}
	(2.462)
median household income	-7.408^{***}
	(2.748)
percentage AIAN	-33.687^{***}
	(12.145)
percentage NHPI	0.112
	(5.062)
observations	3,212
R^2	0.552
adjusted R^2	0.546
residual std. error	3,405.648 (df = 3170)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 14: Two-way fixed-effect model for counties in the western and central United States

	Dependent variable: death rate per 100,000 people
deaths social proximity	2.335^{***}
	(0.193)
deaths spatial proximity	0.011
	(0.355)
ODR	471.881^{***}
	(54.213)
naloxone available	$3{,}058.806^{***}$
	(610.337)
buprenorphine available	0.0002^{***}
	(0.0001)
state count illicit opioid reported	-365.014
	(11,282.540)
population density	0.003^{***}
	(0.0002)
mental health distress rate	174.873^{***}
	(25.425)
political affiliation	-0.162
	(0.275)
percentage unemployed	3.459
	(2.526)
mean distance to hospital	-6.863^{***}
	(1.357)
median household income	8.489^{***}
	(1.786)
percentage AIAN	-23.412^{***}
	(3.021)
percentage NHPI	-1.874
	(1.157)
observations	3,004
R^2	0.510
adjusted R^2	0.504
residual std. error	1625.446 (df = 2967)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 15: Two-way fixed-effect model for counties in the contiguous United States

	Dependent variable: death rate per 100,000 people
deaths social proximity	4.065^{***}
	(0.317)
deaths spatial proximity	1.741^{**}
	(0.720)
ODR	141.232^{***}
	(26.094)
naloxone available	$8{,}212.523^{***}$
	(538.807)
buprenorphine available	0.001^{***}
	(0.0001)
state count illicit opioid reported	-101.839
	(3,171.807)
population density	-0.00004
	(0.0001)
mental health distress rate	301.861^{***}
	(24.263)
political affiliation	0.258
	(0.304)
percentage unemployed	16.511^{***}
	(2.750)
mean distance to hospital	-16.874^{***}
	(2.026)
median household income	6.457^{***}
	(2.019)
percentage AIAN	-30.523^{***}
	(4.315)
percentage NHPI	0.343
	(1.918)
observations	6,216
R^2	0.565
adjusted R^2	0.560
residual std. error	2,889.289 (df = 6152)
Note:	*p<0.1; **p<0.05; ***p<0.01

A.4.5 Two–stage Least Squares Estimation

Our regression analysis identified two endogenous variables, s_{-i} and d_{-i}. Initially, we attempted to correct the correlation in error terms separately, employing network and spatial autocorrelation methods. However, these approaches do not account for correlated error terms that emerge from the simultaneous estimation of effects from socially and spatially lagged variables. To enhance the robustness of our findings, we adopt a two-stage least squares (2SLS) methodology. This approach effectively handles the correlated error terms associated with including endogenous variables in the regression model. The implementation of this model proceeds in two stages. In the first stage, we regress s_{-i} and d_{-i} on a matrix of instrumental variables, denoted by Q. This matrix is defined as $Q = (Z_n, WZ_n, AZ_n, W^2Z_n, A^2Z_n, WAZ_n, AWZ_n)$, where Z_n is an $n \times k$ matrix. Here, n represents the total number of spatial units, and k signifies the number of covariates. Furthermore, W and A are the social and spatial weight matrices, each with dimensions $n \times n$. Our choice of instruments is based on the results of Lee et al. (2010) [72] for higher-order lags.

The equations for the first stage are given by:

$$s_{-i} = \omega_0 + \overline{\omega_1}^T \overline{Q}_i + u_{s_{-i}},$$
$$d_{-i} = \delta_0 + \overline{\delta_1}^T \overline{Q}_i + u_{d_{-i}}.$$

were $u_{s_{-i}}$ and $u_{d_{-i}}$ represent the error term for endogenous variable s_{-i} and d_{-i} and Q_i represents the set of instrumental variables for i-th county observation.

In the second stage, the predicted values of s_{-i} and d_{-i} derived from the first stage are used to fit the final model for y_i. The model is defined by:

$$y_i = \alpha_0 + \alpha_1 \hat{s}_{-i} + \alpha_2 \hat{d}_{-i} + \overline{\alpha_3}^T \overline{C}_i + \overline{\alpha_4}^T \overline{X}_i + \epsilon_i.$$

Supplementary Figure 13 shows the CIs for our 2SLS estimates of s_{-i} and d_{-i} effect sizes in the eastern, western-central and the entire contiguous US. The estimated effect sizes for s_{-i} in all three regions are significant and positive. Supplementary Table 16 provides the 2SLS estimation results and effect sizes.

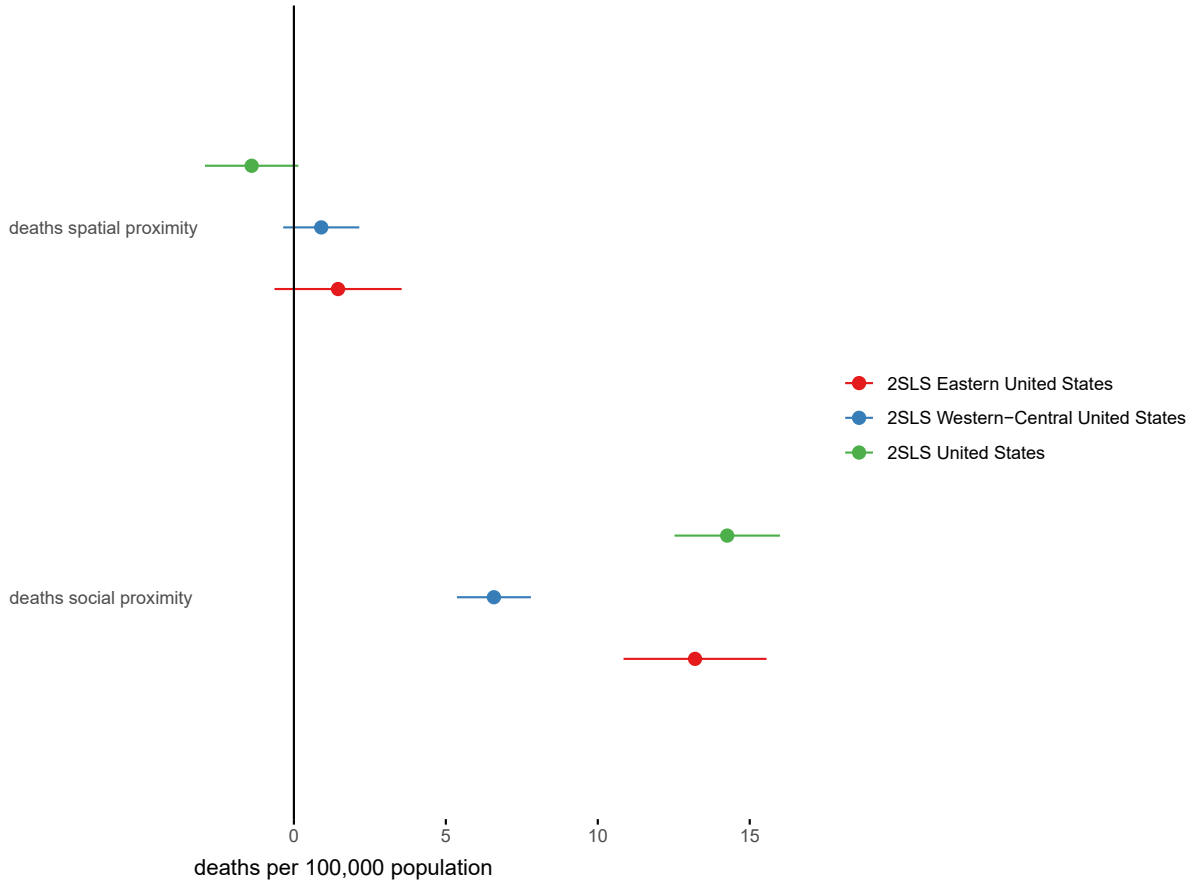

Figure 13: This figure shows the confidence intervals for the two-stage least squares estimates, comparing the counties in the eastern, western, and entire contiguous United States. The effect sizes for s_{-i} are statistically significant ($p < 0.001$) for the three regions. The estimation results are in Supplementary Table 16.

Table 16: This table shows the results of the two-stage least squares estimation, comparing the impact of social and spatial variables on overdose death rates in distinct geographical regions. Column (1) presents the estimates for the counties in the eastern United States , column (2) for the counties in the western and central United States, and column (3) encompasses counties in the entire contiguous United States.

	Dependent variable: death rate per 100,000 people		
	(1)	(2)	(3)
deaths social proximity	13.20017***	6.58148***	14.25920***
	(1.19966)	(0.62021)	(0.88471)
deaths spatial proximity	1.45265	0.89995	−1.38995*
	(1.06666)	(0.63798)	(0.78416)
ODR	52.22437***	6.17012**	38.43051***
	(7.53112)	(2.57692)	(4.80499)
naloxone available	32.04615***	−3.78452	21.51028***
	(11.78907)	(5.97941)	(6.64167)
buprenorphine available	15.08753**	27.46304***	23.85361***
	(6.84219)	(6.06131)	(5.48752)
state count illicit opioid reported	−483.76320	860.05570	1,156.91900
	(2,496.26400)	(3,439.58200)	(1,898.00800)
population density	0.00161**	0.00955***	0.00294***
	(0.00070)	(0.00158)	(0.00058)
mental health distress rate	−104.88350*	−75.99799***	−111.36910***
	(57.89841)	(28.63051)	(29.82314)
political affiliation	−5.34279***	−0.47798	−4.48322***
	(1.38088)	(1.09760)	(0.92130)
percentage unemployed	7.35985	4.65552	2.53671
	(5.72459)	(4.06029)	(3.92320)
mean distance to hospital	−26.03972***	−2.17998	−7.79723***
	(4.18689)	(2.58041)	(2.99407)
median household income	−25.27024***	−7.75322*	−12.02835***
	(7.36773)	(4.00888)	(4.44148)
percentage AIAN	45.87866***	8.02707**	19.49589***
	(15.17279)	(3.97502)	(4.72510)
percentage NHPI	5.78337	−2.46150	1.34235
	(6.75925)	(4.80422)	(6.01157)
constant	45.35172***	18.23200***	33.05425***
	(8.79892)	(4.39869)	(4.59876)
observations	1,606	1,502	3,108
R^2	0.40527	0.30141	0.42091
adjusted R^2	0.40003	0.29483	0.41829
residual std. error	19.28330 (df = 1591)	12.20299 (df = 1487)	16.92010 (df = 3093)
Note:			*p<0.1; **p<0.05; ***p<0.01

A.5 Robustness Check on the Spatial Weight Matrix

To enhance the robustness of our findings, we analyze the spatial adjacency matrix with the distance decay function. The major reason to do so is to ensure that the modified distance decay function reduces the rate at which adjacency weights decrease with distance. This allows us to capture the effect of farther away counties that would otherwise have been under-weighted. We define the the updated weight matrix as follows:

$$a'_{ij} = \frac{1 + \frac{1}{d_{ij}^{1/10}}}{\sum_{k \neq i} \left(1 + \frac{1}{d_{ik}^{1/10}}\right)},$$

where d_{ij} is the distance matrix between county i and j. We run the cluster robust linear regression for counties in eastern US, western and central US pooled together, and counties in the contiguous US. Table 17, 18 and 19 consistently show the coefficient for "deaths in social proximity" is statistically significant and adds additional robustness to our analysis. Supplementary figure 14 shows the confidence interval plot highlighting the statistically significant effect of coefficient for "deaths in social proximity".

Table 17: Linear regression for measuring the significance of the effect size of "deaths in social proximity", for counties in the eastern United States with slowly decaying distance weights

	Dependent variable: death rate per 100,000 people	
	OLS	*Cluster-Robust OLS*
	(1)	(2)
deaths social proximity	13.643***	13.643***
	(0.863)	(2.243)
deaths spatial proximity	−1.604**	−1.604
	(0.780)	(1.546)
ODR	39.663***	39.663
	(12.750)	(25.848)
naloxone available	172.358***	172.358**
	(17.162)	(83.825)
buprenorphine available	57.944***	57.944
	(13.297)	(44.091)
state count illicit opioid reported	10,648.380***	10,648.380
	(3,208.401)	(10,006.390)
population density	−0.0001	−0.0001
	(0.0002)	(0.001)
mental health distress rate	154.389**	154.389
	(77.943)	(152.686)
political affiliation	−3.258**	−3.258
	(1.354)	(2.850)
percentage unemployed	77.127***	77.127**
	(9.729)	(30.586)
mean distance to hospital	−69.209***	−69.209***
	(7.184)	(11.786)
median household income	−12.072*	−12.072
	(6.896)	(14.549)
percentage AIAN	−28.784	−28.784
	(36.110)	(21.705)
percentage NHPI	−19.562	−19.562
	(12.359)	(14.146)
Constant	−0.219	−0.219
	(10.465)	(22.483)
observations	1,606	
R^2	0.454	
adjusted R^2	0.449	
residual std. error	3.210 (df = 1590)	
F statistic	94.369*** (df = 14; 1590)	
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 18: Linear regression for measuring the significance of the effect size of "deaths in social proximity", for counties in the western and central United States with slowly decaying distance weights

	Dependent variable: death rate per 100,000 people	
	OLS	*Cluster-Robust OLS*
	(1)	(2)
deaths social proximity	6.013***	6.013***
	(0.234)	(0.781)
deaths spatial proximity	−3.694***	−3.694*
	(0.375)	(2.108)
ODR	14.954***	14.954
	(3.734)	(9.451)
naloxone available	93.746***	93.746**
	(10.214)	(39.755)
buprenorphine available	46.277***	46.277**
	(8.315)	(19.070)
state count illicit opioid reported	−8,620.923**	−8,620.923
	(4,324.188)	(12,994.870)
population density	0.005***	0.005***
	(0.0005)	(0.001)
mental health distress rate	92.493**	92.493
	(37.035)	(126.872)
political affiliation	−0.472	−0.472
	(0.704)	(2.299)
percentage unemployed	15.921***	15.921
	(5.979)	(17.045)
mean distance to hospital	−9.119***	−9.119
	(3.513)	(10.910)
median household income	6.268*	6.268
	(3.249)	(10.881)
percentage AIAN	−12.689*	−12.689
	(6.560)	(11.559)
percentage NHPI	12.089***	12.089
	(3.693)	(10.008)
Constant	−10.105**	−10.105
	(5.011)	(20.221)
observations	1,502	
R^2	0.492	
adjusted R^2	0.487	
residual std. error	1.004 (df = 1486)	
F statistic	102.823*** (df = 14; 1486)	
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 19: Linear regression for measuring the significance of the effect size of "deaths in social proximity", for counties in the contiguous United States with slowly decaying distance weights

	Dependent variable: death rate per 100,000 people	
	death rates	
	OLS	*Cluster-Robust OLS*
	(1)	(2)
deaths social proximity	15.013***	15.013***
	(0.622)	(1.787)
deaths spatial proximity	−2.357***	−2.357**
	(0.582)	(1.035)
ODR	48.719***	48.719**
	(8.169)	(19.212)
naloxone available	150.050***	150.050***
	(11.250)	(56.853)
buprenorphine available	60.425***	60.425
	(10.014)	(42.605)
state count illicit opioid reported	13,248.850***	13,248.850
	(2,365.555)	(9,362.013)
population density	0.001***	0.001
	(0.0002)	(0.001)
mental health distress rate	69.993	69.993
	(45.879)	(112.128)
political affiliation	−3.194***	−3.194*
	(0.828)	(1.884)
percentage unemployed	42.330***	42.330*
	(6.718)	(23.953)
mean distance to hospital	−49.510***	−49.510***
	(5.545)	(13.203)
median household income	−12.136***	−12.136
	(4.271)	(9.640)
percentage AIAN	−26.900**	−26.900*
	(11.020)	(14.641)
percentage NHPI	−18.952***	−18.952*
	(5.748)	(10.007)
constant	1.248	1.248
	(6.158)	(15.344)
observations	3,108	
R^2	0.504	
adjusted R^2	0.502	
residual std. error	1.895 (df = 3090)	
F statistic	224.521*** (df = 14; 3090)	
Note:		*p<0.1; **p<0.05; ***p<0.01

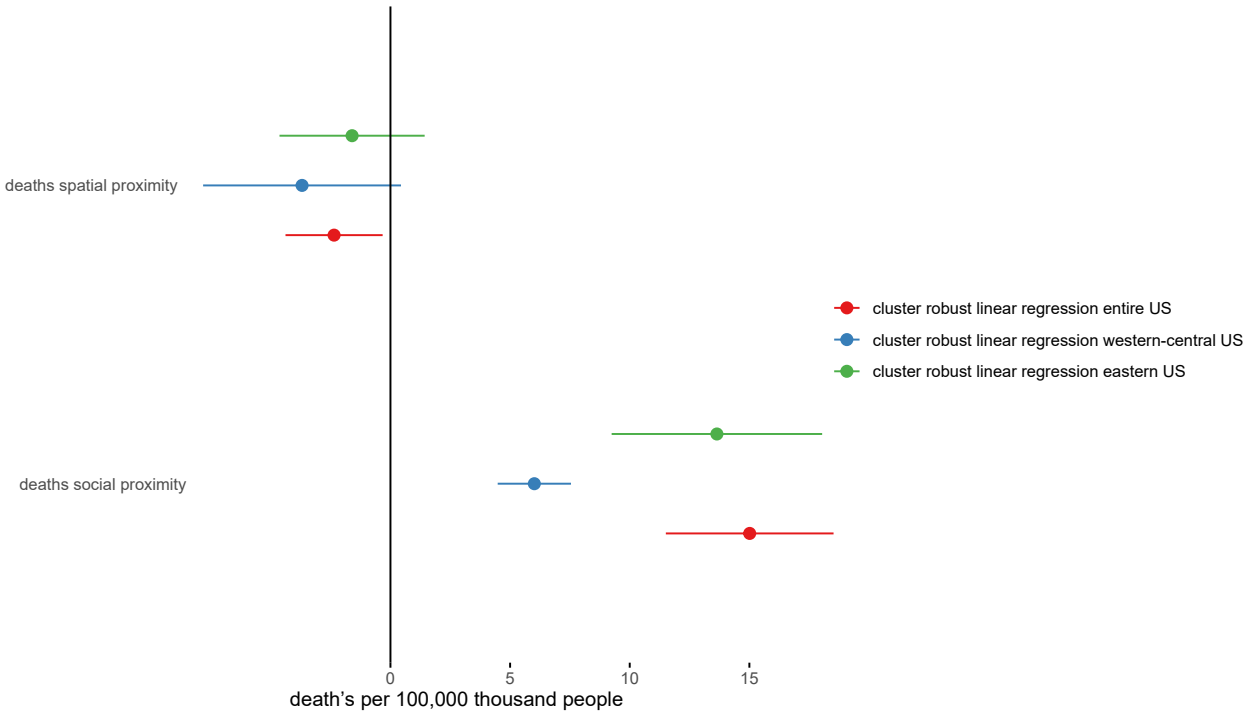

Figure 14: This figure shows the confidence intervals for the cluster-robust linear regression with spatial proximity based on the slow-decaying distance weights, comparing the counties in the eastern, western, and entire contiguous United States. The effect sizes for s_{-i} are statistically significant ($p< 0.001$) for the three regions. The estimation results are in Supplementary Table 17, 18 and 19.

Appendix B Supplementary Information, Chapter 3

B.1 Data Description

All analyses use data at the U.S. county-year level from 2010 to 2022. Suicide mortality data were obtained from the National Vital Statistics System (NVSS) Multiple Cause-of-Death files, restricted to records with ICD-10 underlying cause-of-death codes X60–X84 and Y87.0. Annual population denominators stratified by 18 five-year age groups were drawn from the CDC Bridged-Race Postcensal Population Estimates. These estimates are harmonized for intercensal consistency and aligned with official age groupings used by the U.S. standard population. Demographic and socioeconomic covariates, including median household income, racial/ethnic composition, educational attainment, unemployment rate, and limited English proficiency, were derived from the American Community Survey (ACS) 5-Year Estimates. Social proximity metrics were calculated using the 2018 version of Meta's Social Connectedness Index (SCI). Geospatial coordinates and land area were sourced from the U.S. Census Bureau's TIGER/Line shapefiles. All datasets were merged using 5-digit FIPS county codes.

B.2 Sensitivity Analysis Using Age-Adjusted Suicide Mortality Rates

This supplementary analysis evaluates whether the associations identified in the main manuscript are robust to the use of directly age-adjusted suicide death rates. The primary analysis used crude mortality rates and included county-level age structure as controls. However, considerable heterogeneity exists in the age composition across counties. To ensure that our findings are not an artifact of demographic confounding, we re-estimate all models using age-adjusted suicide mortality rates, which allow comparison of counties standardized to the same reference population. This approach avoids collinearity between age covariates and the outcome, and addresses structural differences in age composition that might otherwise bias

estimates.

Accordingly, age distribution variables, specifically the percentage of the population aged below 18, aged 18–44, aged 45–64, and aged 65 and older are excluded from the covariate set in these models. Age effects are absorbed into the outcome itself.

B.3 Direct Age Standardization of Mortality Rates

Age-adjusted suicide death rates were computed using direct standardization with the *2000 U.S. Standard Population*. Mortality and population counts were stratified using NVSS age recodes (27-category), harmonized to the following **18 age groups**: 0–4, 5–9, 10–14, 15–19, 20–24, 25–29, 30–34, 35–39, 40–44, 45–49, 50–54, 55–59, 60–64, 65–69, 70–74, 75–79, 80–84, 85+.

Let $d_{it}^{(a)}$ and $p_{it}^{(a)}$ denote the number of suicide deaths and the population in age group a, for county i, year t. The standard population share for age group a is denoted $w^{(a)}$. The age-adjusted suicide death rate $\tilde{y}_{it}$ is computed as:

$$\tilde{y}_{it} = 100{,}000 \times \sum_{a} w^{(a)} \cdot \frac{d_{it}^{(a)}}{p_{it}^{(a)}} \tag{26}$$

Estimates are suppressed for any county-year cell with fewer than 10 suicide deaths to ensure reliability and confidentiality in accordance with CDC guidelines.

B.4 Socio-spatial Patterns of Suicide Mortality in the US

We conducted a sensitivity analysis by re-estimating the main models with extitage-adjusted suicide mortality ($\tilde{y}_{it}$) as the outcome. This adjustment accounts for heterogeneity in population age structure across counties and enables us to evaluate the robustness of our findings to different outcome specifications. Age-standardization was performed using direct standardization over 18 age groups: 0–4, 5–9, 10–14, 15–19, 20–24, 25–29, 30–34, 35–39,

40–44, 45–49, 50–54, 55–59, 60–64, 65–69, 70–74, 75–79, 80–84, 85+, as defined in NVSS age recodes (27-category). Age-adjusted rates are expressed per 100,000 population.

Consistent with the main paper, we estimated two-way fixed effects regression models, incorporating county and year indicators to account for unobserved time-invariant county characteristics and nationwide time shocks. The specification included "deaths in social proximity" ($\tilde{s}_{-it}$) as the main exposure, and "deaths in spatial proximity" ($\tilde{d}_{-it}$) as a control in the second model to isolate the role of social connectedness. Sociodemographic covariates included population density, racial composition (Asian, Black, Other), ethnicity (Hispanic), median household income, unemployment rate, limited English proficiency, and education levels. Age composition covariates were excluded, as age adjustment directly accounts for differences in age distribution across counties.

Using the age-adjusted suicide rates ($\tilde{y}_{jt}$), we recompute the two proximity-based exposure metrics:

$$\tilde{s}_{-it} = \sum_{j \neq i} w_{ij} \tilde{y}_{jt} \tag{27}$$

$$\tilde{d}_{-it} = \sum_{j \neq i} a_{ij} \tilde{y}_{jt} \tag{28}$$

where w_{ij} and a_{ij} represent SCI-based and inverse-distance-based weights, respectively, defined identically to those in the main manuscript.

The model estimating the association between suicide mortality and deaths in social and spatial proximity is specified as:

$$\tilde{y}_{it} = \eta_1 \tilde{s}_{-it} + \eta_2 \tilde{d}_{-it} + \boldsymbol{\eta}_3^\top \overline{X}_{it} + \mu_i + \phi_t + \varepsilon_{it} \tag{29}$$

Here, $\tilde{y}_{it}$ denotes the age-adjusted suicide mortality rate in county i and year t; $\tilde{s}_{-it}$ represents the standardized suicide mortality in counties socially connected to i via Facebook SCI; $\tilde{d}_{-it}$ represents the inverse-distance-weighted suicide mortality in spatially proximal counties; $\overline{X}_{it}$ is a vector of county-year sociodemographic covariates; μ_i and ϕ_t denote county and year fixed effects respectively; and ε_{it} is the error term.

We first estimated the association between $\tilde{y}_{it}$ and $\tilde{s}_{-it}$ without controlling for spatial exposure. A one-standard-deviation increase in deaths in socially connected counties was associated with 1.29 additional age-adjusted suicides per 100,000 population in the focal county (cluster-robust 95% CI: [0.87, 1.70], $p < 0.01$; Fig. 15, red point). The magnitude of this effect is sizable relative to the average suicide mortality rate across U.S. counties.

We next included deaths in spatial proximity ($\tilde{d}_{-it}$) in the model. The coefficient for social exposure remained statistically significant, though the point estimate decreased to 1.09 (95% CI: [0.66, 1.52], $p < 0.01$; Fig. 15, blue point). Compared to the main model results in Table 1, where the estimated effects were 3.34 and 2.78 respectively, the effect sizes in the age-adjusted outcome specification are attenuated by approximately 60%. This reduction likely reflects differences in outcome construction: the raw suicide mortality rate incorporates variation from both exposure and demographic composition, while the age-adjusted rate isolates risk conditional on a standardized age distribution.

Despite the reduction, the direction and statistical significance of the coefficients are consistent across specifications, providing evidence for the robustness of the association between social exposure and suicide mortality. Full regression estimates are provided in Table 20. These results strengthen the conclusion that suicide risk is shaped by patterns of social connectedness between counties, independent of both spatial proximity and demographic composition.

B.5 ERPO Exposure: Related Evidence and Sensitivity Analyse

Asessment of the impact of intensified ERPO enforcement following high-profile violent incidents reported significant reductions in firearm suicide mortality—7.5% in Indiana and 13.7% in Connecticut [65]. A recent study examined responses from 4,583 ERPO respondents across California, Connecticut, Maryland, and Washington [114]. The findings estimated that roughly 17 to 23 ERPOs could prevent one suicide, with even greater effectiveness, about 13 to 18 ERPOs per suicide prevented when the petitions explicitly documented suicidal behavior. Additionally, county-level evidence from Oregon highlights the social

Table 20: Two-way fixed effect estimates of socio-spatial influence on *age-adjusted* suicide mortality. Column (1) presents estimates from a model regressing county-level age-adjusted suicide mortality ($\tilde{y}_{it}$) on standardized deaths in socially connected counties ($\tilde{s}_{-it}$). Column (2) additionally controls for standardized deaths in spatial proximity ($\tilde{d}_{-it}$) to isolate social influence from geographic proximity. Both models include county and year fixed effects and control for population density, racial composition (percentages Asian, Black, and Other races), ethnicity (percentage Hispanic), median household income, percentage with limited English proficiency, unemployment rate, and percentage with less than a high school education. Age distribution variables are excluded, as the dependent variable is already adjusted for population structure. Standard errors are clustered at the state level

	Dependent variable: Age-adjusted deaths per 100K	
	Model 1 (1)	Model 2 (2)
$\tilde{s}_{-it}$	1.287^{***}	1.089^{***}
	(0.207)	(0.213)
$\tilde{d}_{-it}$		0.537^{**}
		(0.238)
Population Density	-1.190^{***}	-1.082^{**}
	(0.400)	(0.413)
Percent Black	-1.222^{**}	-1.291^{**}
	(0.577)	(0.604)
Percent Asian	-0.766^{***}	-0.764^{***}
	(0.196)	(0.190)
Percent Other	0.134	0.093
	(0.176)	(0.170)
Percent Hispanic	-3.284^{***}	-3.313^{***}
	(0.817)	(0.773)
Median Household Income	-0.748^{***}	-0.727^{***}
	(0.148)	(0.149)
Percent of population who do not speak English that well	-0.065	-0.051
	(0.077)	(0.076)
Percent unemployed	0.164	0.154
	(0.162)	(0.153)
Percent with less than high school education	0.029	0.021
	(0.102)	(0.102)
Observations	40,794	40,794
R^2	0.566	0.566
Adjusted R^2	0.529	0.530

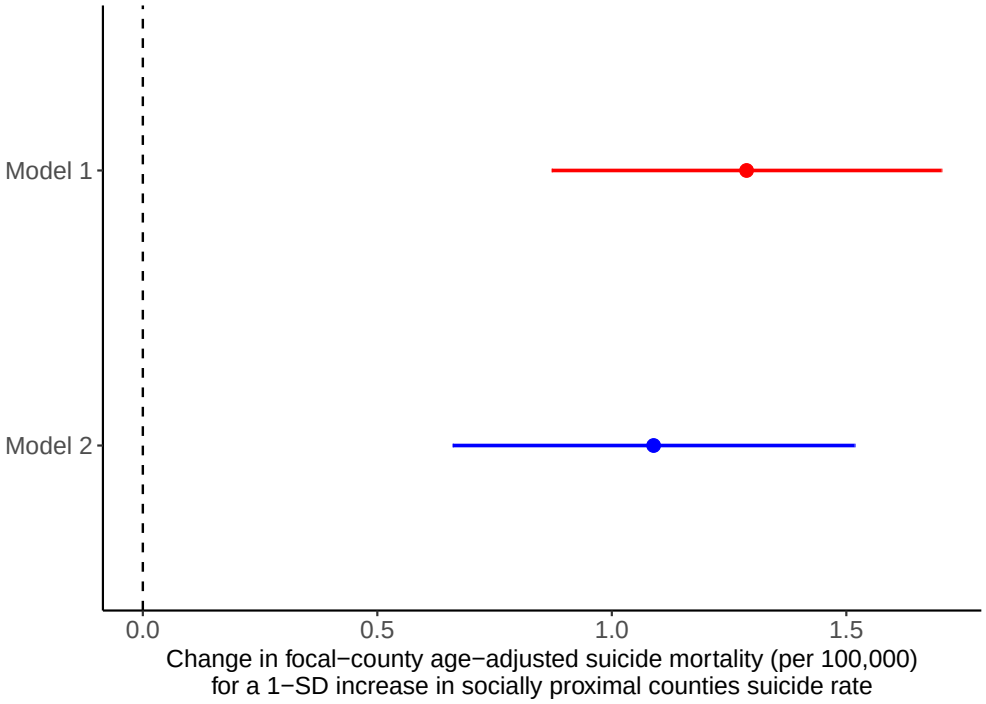

Figure 15: Estimated coefficients ($\hat{\eta}_1$) for deaths in socially connected counties ($\tilde{s}_{-it}$) in two models of age-adjusted suicide mortality: one without adjusting for spatial proximity (red) and one with deaths in spatial proximity ($\tilde{d}_{-it}$) included as a control (blue). Horizontal lines represent 95% confidence intervals; the vertical dashed line indicates the null hypothesis ($\eta_1 = 0$). Estimates correspond to the two-way fixed effects model: Point estimate without controlling for spatial proximity: 1.29 (95% CI: [0.87, 1.70]); with spatial proximity: 1.09 (95% CI: [0.66, 1.52]). All models include county and year fixed effects and time-varying sociodemographic covariates (see Table 20).

aspects of ERPO effectiveness, indicating that approximately 73% of petitions explicitly mentioned concerns about suicidal behavior [125].

To assess the robustness of our findings to outcome specification, we re-estimate the models evaluating ERPO exposure using *age-adjusted suicide mortality* ($\tilde{y}_{it}$) as the dependent variable. This outcome specification accounts for demographic heterogeneity across counties and isolates variation in suicide risk from age structure. Age standardization is performed using direct standardization across 18 age groups defined by the NVSS age recode system. Rates are expressed per 100,000 population.

We estimate three fixed effects models: (1) the direct effect of ERPO implementation; (2) the indirect social exposure model leveraging social ties; and (3) a robustness model adjusting for both social and spatial ERPO exposure. The direct exposure model includes county (ϕ_i) and year (ϕ_t) fixed effects, while the indirect and robust specifications include county fixed effects and state-by-year fixed effects (γ_{st}).

The regression equation for the direct effect model is:

$$\tilde{y}_{it} = \kappa_1 ERPO_{it} + \boldsymbol{\kappa}_2^\top \overline{X}_{it} + \phi_i + \gamma_t + \varepsilon_{it} \tag{30}$$

The regression equation for indirect ERPO social exposure is:

$$\tilde{y}_{it} = \tau_1 ERPO\ Social\ Exposure_{it} + \boldsymbol{\tau}_2^\top \overline{X}_{it} + \phi_i + \gamma_{st} + \varepsilon_{it} \tag{31}$$

The extended model including spatial exposure is:

$$\tilde{y}_{it} = \omega_1 ERPO\ Social\ Exposure_{it} + \omega_2 ERPO\ Spatial\ Exposure_{it} + \boldsymbol{\omega}_3^\top \overline{X}_{it} + \phi_i + \gamma_{st} + \varepsilon_{it} \tag{32}$$

Where $\tilde{y}_{it}$ is the age-adjusted suicide mortality rate in county i and year t; $ERPO\ Social\ Exposure_{it}$ captures the standardized share of social ties to ERPO-adopting counties; $ERPO\ Spatial\ Exposure_{it}$ reflects the inverse-distance-weighted spatial exposure to ERPO-adopting counties; $\overline{X}_{it}$ represents sociodemographic controls.

Direct ERPO Effects. The estimate for the direct policy indicator is $\widehat{\kappa} = -0.6112$ (cluster-robust 95% CI: [-1.00, -0.22]), indicating that ERPO enactment within a county's state is associated with a statistically significant reduction in age-adjusted suicide mortality.

This association is estimated using a two-way fixed effects model controlling for county (ϕ_i) and year (ϕ_t) effects. Because the ERPO variable varies at the state-year level, clustering of standard errors at the state level ensures conservative inference. A reduction of 0.462 deaths per 100,000 is substantial given national suicide baselines and highlights the potential population-level benefit of state-level ERPO legislation.

Social Exposure Effects. For the social ERPO exposure model, we observe $\widehat{\tau}_1 = -0.218$ (cluster-robust 95% CI: [-0.313, -0.123]). Because the exposure metric is standardized (mean = 0, SD = 1), this coefficient quantifies the change in age-adjusted suicide deaths per 100,000 for a one-standard-deviation increase in social connectedness to ERPO-adopting counties. This implies that greater social proximity to counties with active ERPO policies is associated with measurable reductions in suicide mortality, independent of local legislation. The model includes county fixed effects (ϕ_i) and state-by-year fixed effects (γ_{st}), allowing estimation to be identified from within-state, cross-county variation.

Robustness with Spatial Exposure. When additionally controlling for inverse-distance-weighted spatial exposure to ERPO-adopting states, we find $\widehat{\omega}_1 = -0.253$ (cluster-robust 95% CI: [-0.394, -0.111]) for social exposure and $\widehat{\omega}_2 = 0.276$ (cluster-robust 95% CI: [-0.260, -0.812]) for spatial exposure. The social exposure effect remains robust and statistically significant, while the spatial term becomes marginally significant and positive. This suggests that while social networks play a protective role, spatial diffusion may involve heterogeneous or countervailing mechanisms.

Comparison with Main Specification. The magnitude of social exposure effects in this age-adjusted analysis is slightly attenuated compared to the main analysis (which used crude suicide mortality), but direction and significance remain consistent. This highlights the robustness of our findings to alternate outcome specifications and reinforces the conclusion that both direct ERPO policy and social exposure to these policies contribute to reductions in suicide mortality.

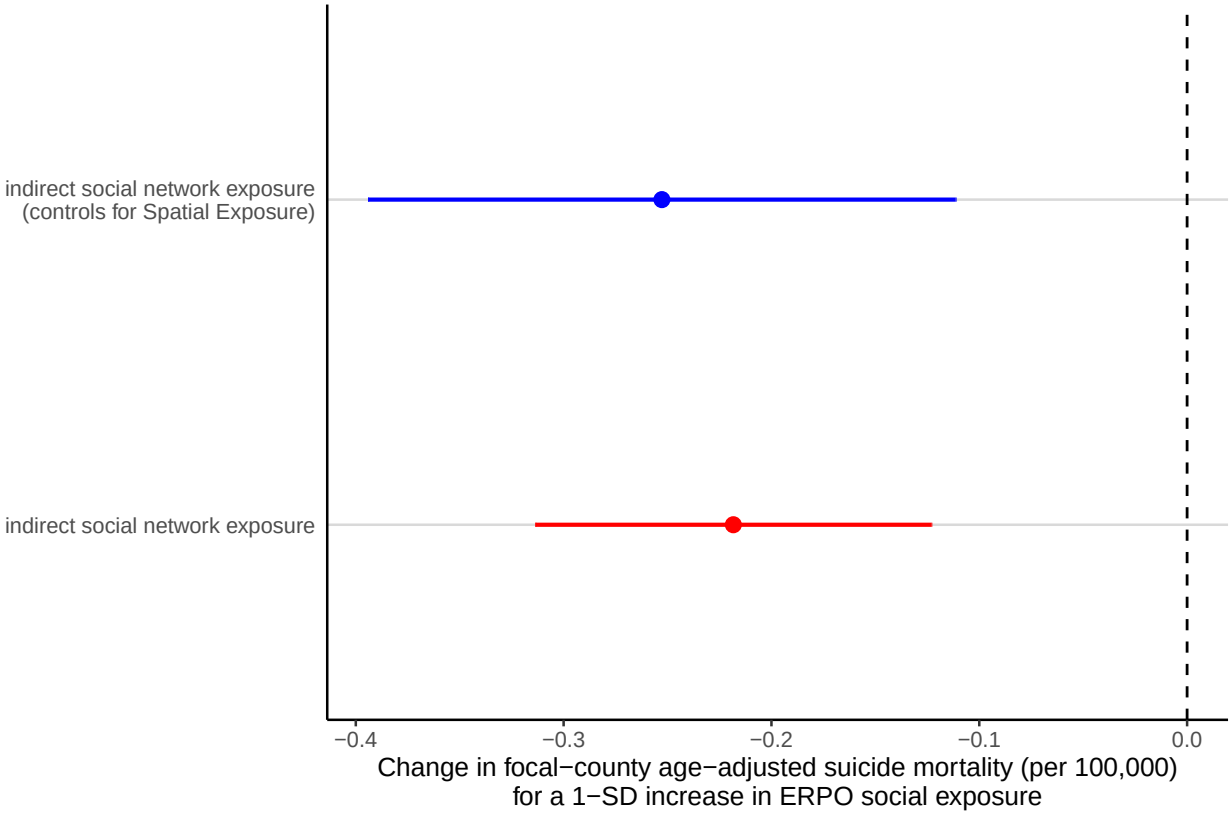

Figure 16: Estimated coefficients ($\hat{\tau}_1$, $\hat{\omega}_1$) for *ERPO Social Exposure* in two specifications. The red point indicates the estimate from the baseline model without spatial exposure ($\hat{\tau}_1 = -0.218$, cluster-robust 95% CI: $[-0.314, -0.123]$); the blue point indicates the estimate from the specification controlling for *ERPO Spatial Exposure* ($\hat{\omega}_1 = -0.253$, cluster-robust 95% CI: $[-0.394, -0.111]$). Horizontal lines denote 95% confidence intervals, and the vertical dashed line denotes the null hypothesis ($\tau_1 = 0$). Both models include county and state-year fixed effects (ϕ_i, γ_{st}) and sociodemographic controls ($\overline{X}_{it}$). The consistent negative and statistically significant estimates indicate that the association between *age-adjusted* suicide mortality and indirect social exposure to ERPO policies is robust to spatial confounding.

Table 21: Estimated Effects of Policy Exposure on *age-adjusted* suicide mortality Direct and Indirect Effects Models. The table reports coefficient estimates from three fixed-effects models assessing the impact of policy exposure on death rates per 100,000 population. The models include direct effects, indirect social network effects, and robust indirect social network effects. Robust standard errors are clustered at the state level.

	Dependent variable: Age-adjusted deaths per 100K		
	Direct Effect	Indirect Social Network Exposure	Robust Indirect Social Network Exposure
	(1)	(2)	(3)
ERPO	-0.611^{***}		
	(0.195)		
ERPO Social Exposure		-0.218^{***}	-0.253^{***}
		(0.047)	(0.070)
ERPO Spatial Exposure			0.276
			(0.267)
Population density	-1.553^{***}	-0.639	-0.671
	(0.422)	(0.532)	(0.516)
Percent Asian	-1.355^{**}	-2.411^{***}	-2.344^{***}
	(0.528)	(0.614)	(0.600)
Percent Black	-0.929^{***}	-0.507^{***}	-0.491^{***}
	(0.226)	(0.172)	(0.170)
Percent Other	0.213	-0.048	-0.062
	(0.191)	(0.188)	(0.191)
Percent Hispanic	-3.460^{***}	-1.546^{***}	-1.585^{***}
	(0.989)	(0.555)	(0.562)
Median household income	-0.729^{***}	-0.635^{***}	-0.611^{***}
	(0.166)	(0.224)	(0.221)
Percent with limited English proficiency	-0.092	-0.151^{*}	-0.157^{**}
	(0.082)	(0.076)	(0.073)
Percent unemployed	0.181	-0.102	-0.099
	(0.171)	(0.135)	(0.136)
Percent with less than high school education	0.057	-0.077	-0.094
	(0.124)	(0.109)	(0.109)
Political Affiliation	0.036	0.140	0.124
	(0.164)	(0.152)	(0.154)
Observations	40,794	40,794	40,794
R^2	0.565	0.575	0.577
Adjusted R^2	0.528	0.532	0.534

Robust standard errors in parentheses. * p<0.1; ** p<0.05; *** p<0.01

B.5.1 Robustness Check: Controlling for Deaths in Social Proximity

To further validate our findings, we conduct a robustness check that controls for suicide deaths in socially connected counties. This addresses potential confounding from correlated suicide mortality patterns across the social network. Specifically, social connectedness may proxy for shared suicide risk environments, contagion effects, or common cultural and structural determinants of suicide that operate independently of ERPO policies. Without accounting for these baseline correlations in suicide mortality, our estimated ERPO social exposure effects could spuriously reflect pre-existing network-level patterns rather than genuine policy impacts.

We estimate two fixed-effects models: (1) a crude mortality specification and (2) an age-adjusted mortality specification that leverages standardized social and spatial ERPO exposures. Both models include county fixed effects (ϕ_i) and state-by-year fixed effects (γ_{st}), with standard errors clustered at the state level. The deaths-in-social-proximity controls (s_{-it} or $\tilde{s}_{-it}$) capture the SCI-weighted average of contemporaneous suicide mortality in connected counties, excluding the focal county itself. By conditioning on these network-level suicide patterns, we isolate the effect of ERPO social exposure beyond what can be explained by correlated mortality across the social network.

The regression equation for indirect ERPO social exposure with crude mortality as an outcome is:

$$\begin{aligned} y_{it} = \alpha_1 \, ERPO\ Social\ Exposure_{it} + \alpha_2 \, ERPO\ Spatial\ Exposure_{it} \\ + \alpha_3 \, s_{-it} + \boldsymbol{\alpha}_4^\top \, \overline{X}_{it} + \phi_i + \gamma_{st} + \varepsilon_{it}. \end{aligned} \tag{33}$$

The regression equation for indirect ERPO social exposure with age adjusted mortality as an outcome:

$$\begin{aligned} \tilde{y}_{it} = \vartheta_1 \, \widetilde{ERPO\ Social\ Exposure}_{it} + \vartheta_2 \, \widetilde{ERPO\ Spatial\ Exposure}_{it} \\ + \vartheta_3 \, \tilde{s}_{-it} + \boldsymbol{\vartheta}_4^\top \, \overline{X}_{it} + \phi_i + \gamma_{st} + \varepsilon_{it}. \end{aligned} \tag{34}$$

Where y_{it} is the crude suicide mortality rate (per 100,000) in county i and year t; $\tilde{y}_{it}$ is the age-adjusted suicide mortality rate; $ERPO\ Social\ Exposure_{it}$ and $ERPO\ Spatial\ Exposure_{it}$

are social and inverse-distance spatial ERPO exposure measures; s_{-it} is the social connectedness weighted average of *crude* suicide deaths in socially connected counties ($j \neq i$); $\tilde{s}_{-it}$ is the corresponding *age-adjusted* measure; and $\overline{X}_{it}$ denotes sociodemographic controls. All specifications include county fixed effects (ϕ_i) and state-by-year fixed effects (γ_{st}); inference uses state-clustered standard errors.

The results demonstrate that the negative association between ERPO social exposure and suicide mortality remains robust after controlling for contemporaneous suicide deaths in socially connected counties. As shown in Figure 17 and Tables 22 and 23, the estimated coefficients for ERPO social exposure are consistently negative and statistically significant in both the crude mortality ($\hat{\alpha}_1 = -0.304$, 95% CI: $[-0.472, -0.136]$) and age-adjusted mortality ($\hat{\vartheta}_1 = -0.270$, 95% CI: $[-0.420, -0.120]$) specifications. Notably, the deaths-in-social-proximity control is statistically significant in the age-adjusted specification ($\hat{\vartheta}_3 = -0.428$, $p < 0.05$) but not in the crude specification. This divergence likely reflects the standardization procedure in age-adjusted rates, which removes age-specific heterogeneity and reveals underlying correlations in suicide mortality across socially connected counties that are otherwise masked by differential age structures. Critically, the inclusion of this control does not materially alter the magnitude or significance of the ERPO social exposure coefficient in either specification, indicating that the observed associations are not driven by confounding from correlated suicide risk across the social network. The robustness of our estimates to this control strengthens the interpretation that ERPO social exposure independently predicts reductions in suicide mortality beyond what can be explained by pre-existing network-level suicide patterns or shared risk environments.

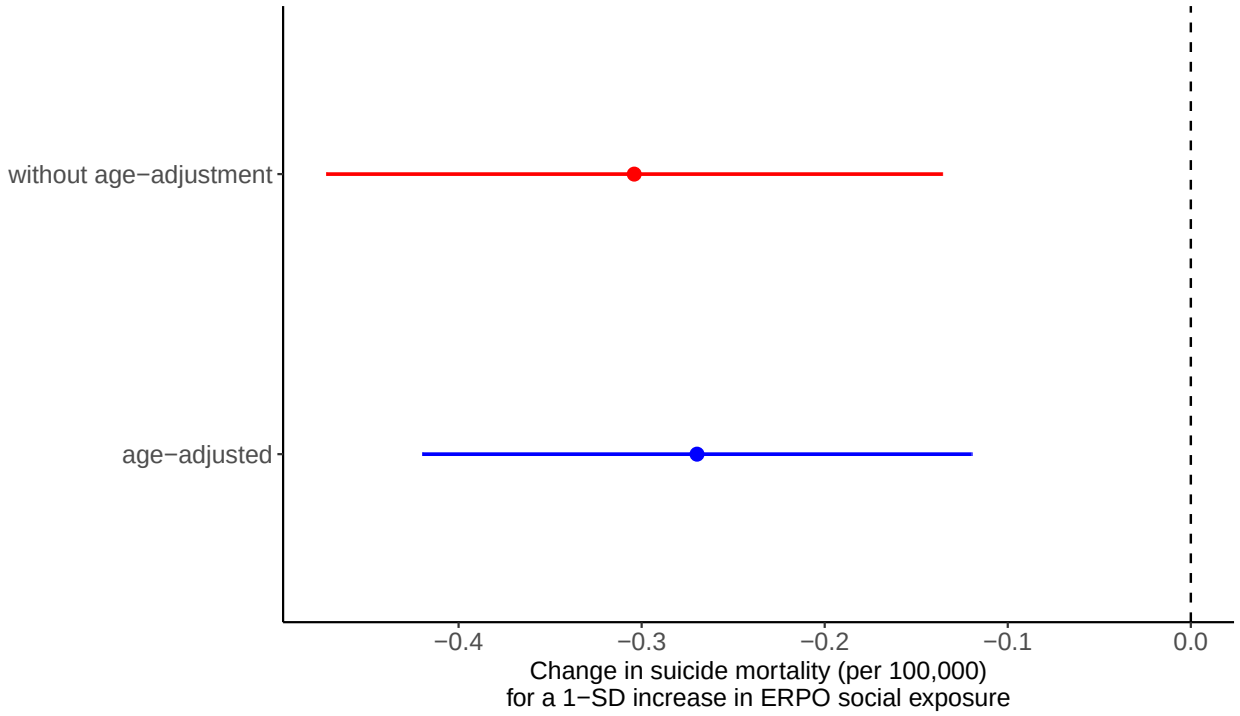

Figure 17: Estimated coefficients ($\hat{\alpha}_1$, $\hat{\vartheta}_1$) for *ERPO Social Exposure* in two specifications that control for deaths in social proximity. The red point indicates the estimate from the *age-adjusted* model ($\hat{\vartheta}_1 = -0.270$, cluster-robust 95% CI: $[-0.420, -0.120]$); the blue point indicates the estimate from the model *without age adjustment* ($\hat{\alpha}_1 = -0.304$, cluster-robust 95% CI: $[-0.472, -0.136]$). Horizontal lines denote 95% confidence intervals, and the vertical dashed line denotes the null hypothesis ($= 0$). Both models include county and state-by-year fixed effects (ϕ_i, γ_{st}), sociodemographic controls ($\overline{X}_{it}$), and a control for deaths in social proximity (s_{-it} or $\tilde{s}_{-it}$). The consistent negative and statistically significant estimates indicate that the association between *suicide mortality* and indirect social exposure to ERPO policies is robust to social-spillover confounding.

Table 22: Estimated effects of ERPO policy exposure on county-level suicide mortality (deaths per 100,000 people). The estimated effects includes county and state–year fixed effects. The model adjusts for s_{-it}, population density, age distribution (percent aged 0–17, 18–44 and 45–64), racial composition (percent Asian, Black, and Other racial subgroups), ethnicity (percent Hispanic), median household income, percent with limited English proficiency, percent unemployed, percent with less than high school education and political affiliation. Standard errors are clustered at the state level

	Dependent variable: death rates per 100K
ERPO Social Exposure	-0.304^{***}
	(0.084)
ERPO Spatial Exposure	0.514
	(0.312)
Deaths in social proximity s_{-it}	-0.491
	(1.149)
Population density	-0.652
	(0.587)
Percent aged below 18	-0.186
	(0.370)
Percent aged 18-44	-0.121
	(0.556)
Percent aged 45-64	-0.514
	(0.402)
Percent Asian	-0.646^{**}
	(0.254)
Percent Black	-2.790^{***}
	(0.681)
Percent Other	-0.122
	(0.280)
Percent Hispanic	-2.876^{***}
	(0.826)
Median household income	-0.603^{***}
	(0.186)
Percent with limited English proficiency	-0.050
	(0.129)
Percent unemployed	-0.098
	(0.125)
Percent with less than high school education	-0.052
	(0.124)
Political Affiliation	0.138
	(0.156)
Observations	40,794
R^2	0.947
Adjusted R^2	0.941

Table 23: Estimated effects of ERPO policy exposure on county-level age-adjusted suicide mortality (age-adjusted deaths per 100,000 people). The estimated effects includes county and state–year fixed effects. The model adjusts for $\widetilde{s}_{-it}$ population density, racial composition (percent Asian, Black, and Other racial subgroups), ethnicity (percent Hispanic), median household income, percent with limited English proficiency, percent unemployed, percent with less than high school education and political affiliation. Standard errors are clustered at the state level

	Dependent variable: Age-adjusted deaths per 100K
ERPO Social Exposure	-0.270^{***}
	(0.075)
ERPO Spatial Exposure	0.275
	(0.280)
Deaths in social proximity $\tilde{s}_{-it}$	-0.428^{**}
	(0.205)
Population density	-0.702
	(0.534)
Percent Asian	-2.361^{***}
	(0.609)
Percent Black	-0.529^{***}
	(0.173)
Percent Other	-0.073
	(0.195)
Percent Hispanic	-1.666^{***}
	(0.597)
Median household income	-0.621^{***}
	(0.221)
Percent with limited English proficiency	-0.159^{**}
	(0.074)
Percent unemployed	-0.094
	(0.137)
Percent with less than high school education	-0.094
	(0.110)
Political Affiliation	0.131
	(0.154)
Observations	40,794
R^2	0.577
Adjusted R^2	0.534

Bibliography

[1] Angeela Acharya, Alyssa M. Izquierdo, Stefanie F. Gonçalves, Rebecca A. Bates, Faye S. Taxman, Martin P. Slawski, Huzefa S. Rangwala, and Siddhartha Sikdar. Exploring county-level spatio-temporal patterns in opioid overdose related emergency department visits. *PLOS ONE*, 17(12):e0269509, 2022.

[2] Abdulrahman A Ahmed, M Amin Rahimian, and Mark S Roberts. Estimating treatment effects using costly simulation samples from a population-scale model of opioid use disorder. In *2023 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI)*, pages 1–4. IEEE, 2023.

[3] Yousef Alimohamadi, S. M. Zahraei, M. Karami, M. Yaseri, M. Lotfizad, and K. Holakouie. Global infectious disease early warning models: An updated review and lessons from the covid-19 pandemic. *Infectious Disease Modelling*, 9:1–14, 2024. Discusses early warning models and notes that exponentially weighted moving averages are suitable for short-term historical data and sensitive to small changes.

[4] Abby Alpert, William N Evans, Ethan MJ Lieber, and David Powell. Origins of the opioid crisis and its enduring impacts. *The Quarterly Journal of Economics*, 137(2):1139–1179, 2022.

[5] Roy M. Anderson and Robert M. May. *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press, 1991.

[6] Alejandra Arango, David Brent, Jacqueline Grupp-Phelan, Bradley J Barney, Anthony Spirito, Megan M Mroczkowski, Rohit Shenoi, Melinda Mahabee-Gittens, T Charles Casper, Cheryl King, et al. Social connectedness and adolescent suicide risk. *Journal of child psychology and psychiatry*, 65(6):785–797, 2024.

[7] Giulia Arena, Joris Mulder, and Roger Th. A. J. Leenders. How fast do we forget our past social interactions? understanding memory retention with parametric decays in relational event models. *Network Science*, 11(2):267–294, 2023.

[8] Florian Arendt, Sebastian Scherr, and Daniel Romer. Effects of exposure to self-harm on social media: Evidence from a two-wave panel study among young adults. *New Media & Society*, 21(11-12):2422–2442, 2019.

[9] Michael Bailey, Rachel Cao, Theresa Kuchler, Johannes Stroebel, and Arlene Wong. Social connectedness: Measurement, determinants, and effects. *Journal of Economic Perspectives*, 32(3):259–80, 2018.

[10] Michael Bailey, Patrick Farrell, Theresa Kuchler, and Johannes Stroebel. Social connectedness in urban areas. *Journal of Urban Economics*, 118:103264, 2020.

[11] Michael Bailey, Drew Johnston, Martin Koenen, Theresa Kuchler, Dominic Russel, and Johannes Stroebel. Social networks shape beliefs and behavior: Evidence from social distancing during the COVID-19 pandemic. *Journal of Political Economy Microeconomics*, 2(3):000–000, 2024.

[12] Michael Bailey, Drew Johnston, Theresa Kuchler, Dominic Russel, Bogdan State, and Johannes Stroebel. The determinants of social connectedness in europe. In *Social Informatics: 12th International Conference, SocInfo 2020, Pisa, Italy, October 6–9, 2020, Proceedings 12*, pages 1–14. Springer, 2020.

[13] Duilio Balsamo, Paolo Bajardi, Alberto Salomone, and Rossano Schifanella. Patterns of routes of administration and drug tampering for nonmedical opioid consumption: data mining and content analysis of reddit discussions. *Journal of Medical Internet Research*, 23(1):e21212, 2021.

[14] Arnab K Basu, Nancy H Chau, and Oleg Firsin. Social connections and COVID-19 vaccination. In *Discussion Paper Series*. Institute of Labor Economics (IZA), 2023.

[15] Cici Bauer, Kehe Zhang, Wenjun Li, Dana Bernson, Olaf Dammann, Marc R. LaRochelle, and Thomas J. Stopka. Small area forecasting of opioid-related mortality: Bayesian spatiotemporal dynamic modeling approach. *JMIR Public Health and Surveillance*, 9:e41450, 2023.

[16] Alba Bernabeu, Jiancang Zhuang, and Jorge Mateu. Spatio-temporal hawkes point processes: A review. *Journal of Agricultural, Biological and Environmental Statistics*, 30:89–119, 2025. Provides a review of spatio-temporal Hawkes processes, highlighting exponential kernels as common choices for temporal and spatial triggering functions.

[17] O. Bernal, A. Mueller, and R. Kessler. Public health surveillance of suicide: recommendations for a comprehensive approach. *American Journal of Public Health*, 112(4):579–586, 2022.

[18] Dan Braha and Marcus AM De Aguiar. Voting contagion: Modeling and analysis of a century of US presidential elections. *PLOS ONE*, 12(5):e0177970, 2017.

[19] John T. Cacioppo and Stephanie Cacioppo. The growing problem of loneliness. *The Lancet*, 391(10119):426, 2018.

[20] Silvia Calvo, Silvia García-Sanjuán, Bárbara Pérez, Luis A. Rico-Uribe, Philippe Courtet, Àngels Cebrià, and Jordi Alonso. Does suicide contagion (werther effect) take place in response to social media? a systematic review. *Spanish Journal of Psychiatry and Mental Health*, 2024. In press.

[21] Canadian Centre on Substance Use and Addiction. Intersections of substance use and suicide: Evidence and key take-aways, 2024. Accessed July 22, 2025.

[22] Suicide Prevention Resource Center. Strategic planning framework for suicide prevention, 2020. https://sprc.org/sites/default/files/resource-program/SPRC-Strategic-Planning-Framework.pdf.

[23] Centers for Disease Control and Prevention. Provisional mortality data — united states, 2023, 2024. Accessed July 22, 2025.

[24] Centers for Disease Control and Prevention. Dose-sys dashboard: Nonfatal overdose syndromic surveillance data. https://www.cdc.gov/overdose-prevention/data-research/facts-stats/dose-dashboard-nonfatal-surveillance-data.html, July 2025. Syndromic surveillance used to rapidly detect overdose clusters.

[25] Ian Cero, Munmun De Choudhury, and Peter A. Wyman. Social network structure as a suicide prevention target. *Social Psychiatry and Psychiatric Epidemiology*, 2023.

[26] Ben Charoenwong, Alan Kwan, and Vesa Pursiainen. Social connections with COVID-19–affected areas increase compliance with mobility restrictions. *Science Advances*, 6(47):eabc3054, 2020.

[27] Raj Chetty, Matthew O Jackson, Theresa Kuchler, Johannes Stroebel, Nathaniel Hendren, Robert B Fluegge, Sara Gong, Federico Gonzalez, Armelle Grondin, Matthew Jacob, et al. Social capital i: measurement and associations with economic mobility. *Nature*, 608(7921):108–121, 2022.

[28] Raj Chetty, Matthew O Jackson, Theresa Kuchler, Johannes Stroebel, Nathaniel Hendren, Robert B Fluegge, Sara Gong, Federico Gonzalez, Armelle Grondin, Matthew

Jacob, et al. Social capital ii: determinants of economic connectedness. *Nature*, 608(7921):122–134, 2022.

[29] Kar-Hai Chu, Ariel Shensa, Jason B Colditz, Jaime E Sidani, Beth L Hoffman, David Sinclair, Mary G Krauland, and Brian A Primack. Integrating social dynamics into modeling cigarette and e-cigarette use. *Health Education & Behavior*, 47(2):191–201, 2020.

[30] Barbara J Costello, Bradley J Anderson, and Michael Stein. Peer influence in initiation to heroin use. *Journal of Drug Issues*, 51(2):323–339, 2021.

[31] Joshua Coven, Arpit Gupta, and Iris Yao. JUE Insight: Urban flight seeded the COVID-19 pandemic across the United States. *Journal of Urban Economics*, 133:103489, 2023.

[32] David M Cutler and J Travis Donahoe. Thick market externalities and the persistence of the opioid epidemic. Technical report, National Bureau of Economic Research, 2024.

[33] Christina M Cutter, Richard C Larson, and Mahshid Abir. Social network theory—an underutilized opportunity to align innovative methods with the demands of the opioid epidemic. *The American Journal of Drug and Alcohol Abuse*, 47(3):305–310, 2021.

[34] Nabarun Dasgupta, Leo Beletsky, and Daniel Ciccarone. Opioid crisis: no easy fix to its social and economic determinants. *American Journal of Public Health*, 108(2):182–186, 2018.

[35] Annamarie B. Defayette, Steven M. Silverstein, and Anthony R. Pisani. Social network structure as a biopsychosocial suicide prevention target for young people at clinical high-risk for psychosis. *Schizophrenia Research*, 270:63–67, 2024.

[36] Andreas Diemer and Tanner Regan. No inventor is an island: social connectedness and the geography of knowledge flows in the us. *Research Policy*, 51(2):104416, 2022.

[37] Amanda T Dinwiddie. Reported non-substance-related mental health disorders among persons who died of drug overdose—United States, 2022. *MMWR. Morbidity and Mortality Weekly Report*, 73, 2024.

[38] Centers for Disease Control and Prevention. Comprehensive suicide prevention program, 2022. https://www.cdc.gov/suicide/programs/csp/index.html.

[39] Centers for Disease Control and Prevention. Provisional drug overdose death counts. *Online Resource*, 2022. Accessed: 2024-07-27.

[40] Amir Forati, Rina Ghose, Fahimeh Mohebbi, and John R Mantsch. The journey to overdose: Using spatial social network analysis as a novel framework to study geographic discordance in overdose deaths. *Drug and alcohol dependence*, 245:109827, 2023.

[41] Jerome Friedman, Trevor Hastie, and Rob Tibshirani. Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33(1):1, 2010.

[42] Michele Garetto, Emilio Leonardi, and Giovanni Luca Torrisi. A time-modulated hawkes process to model the spread of covid-19 and the impact of countermeasures. *Annual Review of Control*, 51:551–563, 2021. Shows the equivalence between SIR models with exponential infectious periods and Hawkes processes with exponential kernels.

[43] Yao Ge, Sudeshna Das, Karen O'Connor, Mohammed Ali Al-Garadi, Graciela Gonzalez-Hernandez, and Abeed Sarker. Reddit-impacts: A named entity recognition dataset for analyzing clinical and social effects of substance use derived from social media. *arXiv preprint arXiv:2405.06145*, 2024.

[44] Rachel E Gicquelais, Becky L Genberg, Jessica L Maksut, Amy SB Bohnert, and Anne C Fernandez. Prevalence and correlates of using opioids alone among individuals in a residential treatment program in michigan: implications for overdose mortality prevention. *Harm reduction journal*, 19(1):135, 2022.

[45] Madelyn S. Gould. Suicide and the media. *Annals of the New York Academy of Sciences*, 1008(1):200–224, 2003.

[46] Madelyn S Gould and Alison M Lake. The contagion of suicidal behavior. In *Contagion of violence: Workshop summary*, pages 68–73. Institute of Medicine and National Research Council Washington, DC, 2013.

[47] John J Grefenstette, Shawn T Brown, Roni Rosenfeld, Jay DePasse, Nathan TB Stone, Phillip C Cooley, William D Wheaton, Alona Fyshe, David D Galloway, Anuroop Sriram, et al. Fred (a framework for reconstructing epidemic dynamics): an open-source software system for modeling infectious diseases and control strategies using census-based populations. *BMC Public Health*, 13:1–14, 2013.

[48] Honoria Guarino, Pedro Mateu-Gelabert, Jennifer Teubl, and Elizabeth Goodbody. Young adults' opioid use trajectories: From nonmedical prescription opioid use to heroin, drug injection, drug treatment and overdose. *Addictive behaviors*, 86:118–123, 2018.

[49] Rebecca L Haffajee, Lewei Allison Lin, Amy SB Bohnert, and Jason E Goldstick. Characteristics of US counties with high opioid overdose mortality and low capacity to deliver medications for opioid use disorder. *JAMA network open*, 2(6):e196373–e196373, 2019.

[50] Dion Harmon, Marco Lagi, Marcus AM De Aguiar, David D Chinellato, Dan Braha, Irving R Epstein, and Yaneer Bar-Yam. Anticipating economic market crises using measures of collective panic. *PloS one*, 10(7):e0131871, 2015.

[51] Alan G. Hawkes. Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1):83–90, 1971.

[52] Holly Hedegaard, Arialdi M Miniño, and Melissa R Spencer. Drug overdose deaths in the united states, 1999–2020. *NCHS Data Brief*, 428, 2021. https://www.cdc.gov/nchs/products/databriefs/db428.htm.

[53] Kyle Heuton, Jyontika Kapoor, Shikhar Shrestha, Thomas J. Stopka, and Michael C. Hughes. Spatiotemporal forecasting of opioid-related fatal overdoses: towards best practices for modeling and evaluation. *American Journal of Epidemiology*, 2024.

[54] Sameer Hinduja and Justin W. Patchin. Connecting adolescent suicide to the severity of bullying and cyberbullying. *Journal of School Violence*, 18(3):333–346, 2018.

[55] David Holtz, Michael Zhao, Seth G Benzell, Cathy Y Cao, Mohammad Amin Rahimian, Jeremy Yang, Jennifer Allen, Avinash Collis, Alex Moehring, Tara Sowrirajan, et al. Interdependence and the cost of uncoordinated responses to COVID-19. *Proceedings of the National Academy of Sciences*, 117(33):19837–19843, 2020.

[56] Jack Homer and Wayne Wakeland. A dynamic model of the opioid drug epidemic with implications for policy. *The American Journal of Drug and Alcohol Abuse*, 47(1):5–15, 2021.

[57] Hawre Jalal, Jeanine M Buchanich, Mark S Roberts, Lauren C Balmert, Kun Zhang, and Donald S Burke. Changing dynamics of the drug overdose epidemic in the United States from 1979 through 2016. *Science*, 361(6408):eaau1184, 2018.

[58] Hawre Jalal, Jeanine M Buchanich, David R Sinclair, Mark S Roberts, and Donald S Burke. Age and generational patterns of overdose death risk from opioids and other drugs. *Nature medicine*, 26(5):699–704, 2020.

[59] Hawre Jalal and Donald S Burke. Exponential growth of drug overdose poisoning and opportunities for intervention. *Addiction*, 117(5):1200–1202, 2022.

[60] Delfina Janiri, Gaelle E Doucet, Maurizio Pompili, Gabriele Sani, Beatriz Luna, David A Brent, and Sophia Frangou. Risk and protective factors for childhood suicidality: a us population-based study. *The Lancet Psychiatry*, 7(4):317–326, 2020.

[61] Robert Alun Jones. *Emile Durkheim: An introduction to four major works*. Sage Publ., 1987.

[62] Mbabazi Kariisa, Nancy L Davis, Harlan Kim, and et al. Vital signs: Drug overdose deaths, by selected sociodemographic and social determinants of health characteristics — 25 states and the district of columbia, 2019–2020. *Morbidity and Mortality Weekly Report*, 71(29):940–947, 2022.

[63] W. O. Kermack and A. G. McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society A*, 115(772):700–721, 1927.

[64] Binod Khanal. Role of social connectedness in response to a public health crisis: The case study of the flint water crisis. *Available at SSRN 3969415*, 2021.

[65] Aaron J Kivisto and Peter Lee Phalen. Effects of risk-based firearm seizure laws in connecticut and indiana on suicide rates, 1981–2015. *Psychiatric Services*, 69(8):855–862, 2018.

[66] Evan M Kleiman and Richard T Liu. Social support as a protective factor in suicide: Findings from two nationally representative samples. *Journal of affective disorders*, 150(2):540–545, 2013.

[67] B. Klimes-Dougan, D. Klingbeil, and D. Mroczek. Gatekeeper training as a suicide prevention strategy: A systematic review. *Suicide and Life-Threatening Behavior*, 51(2):322–335, 2021.

[68] Nana Koram, Hongjie Liu, Jianhua Li, Jian Li, Jian Luo, and Jennifer Nield. Role of social network dimensions in the transition to injection drug use: actions speak louder than words. *AIDS and Behavior*, 15:1579–1588, 2011.

[69] Theresa Kuchler, Dominic Russel, and Johannes Stroebel. JUE insight: The geographic spread of COVID-19 correlates with the structure of social networks as measured by facebook. *Journal of Urban Economics*, 127:103314, 2022.

[70] Bridget M Kuehn. Fentanyl drives startling increases in adolescent overdose deaths. *JAMA*, 329(4):280–281, 2023.

[71] Navin Kumar, William Oles, Benjamin A Howell, Kamila Janmohamed, Selena T Lee, Melissa C Funaro, Patrick G O'Connor, and Marcus Alexander. The role of social network support in treatment outcomes for medication for opioid use disorder: A systematic review. *Journal of Substance Abuse Treatment*, 127:108367, 2021.

[72] Lung-fei Lee and Xiaodong Liu. Efficient GMM estimation of high order spatial autoregressive models with autoregressive disturbances. *Econometric Theory*, 26(1):187–230, 2010.

[73] Yuchen Li, Harvey J Miller, Elisabeth D Root, Ayaz Hyder, and Desheng Liu. Understanding the role of urban social and physical environment in opioid overdose events using found geospatial data. *Health & Place*, 75:102792, 2022.

[74] Che-Yi Liao, Gian-Gabriel Garcia, Kamran Paynabar, Zheng Dong, Yao Xie, and Mohammad S Jalali. Tides need stemmed: A locally operating spatio-temporal mutually exciting point process with dynamic network for improving opioid overdose death prediction. *arXiv preprint arXiv:2211.07570*, 2022.

[75] Tse Yang Lim, Erin J Stringfellow, Celia A Stafford, Catherine DiGennaro, Jack B Homer, Wayne Wakeland, Sara L Eggers, Reza Kazemi, Lukas Glos, Emily G Ewing, et al. Modeling the evolution of the US opioid crisis for national policy development. *Proceedings of the National Academy of Sciences*, 119(23):e2115714119, 2022.

[76] Mengyu Liu, Joel M Caplan, Leslie W Kennedy, Imelda K Moise, Daniel J Feaster, Viviana E Horigian, John M Roll, Sterling M McPherson, and J Sunil Rao. Geospatial risk factor analysis for drug overdose death in south florida from 2014 to 2019, and the independent contribution of social determinants of health. *Drug and alcohol dependence*, 248:109931, 2023.

[77] James M. Lucas. Exponentially weighted moving average control schemes: Properties and enhancements. *Technometrics*, 32(1):1–12, 1990.

[78] Eric Luxenberg and Stephen Boyd. Exponentially weighted moving models. arXiv:2404.08136, 2024. Describes exponential weighting and half-life interpretations in time-varying models.

[79] Kai Mäckle and Stefan Ruenzi. Friends with drugs: The role of social networks in the opioid epidemic. *Available at SSRN*, 2022.

[80] Alexandria Macmadu, Lisa Frueh, Alexandra B Collins, Roxxanne Newman, Nancy P Barnett, Josiah D Rich, Melissa A Clark, and Brandon DL Marshall. Drug use behaviors, trauma, and emotional affect following the overdose of a social network member: a qualitative investigation. *International Journal of Drug Policy*, 107:103792, 2022.

[81] Charles Marks, Daniela Abramovitz, Christl A. Donnelly, Gabriel Carrasco-Escobar, Rocío Carrasco-Hernandez, Daniel Ciccarone, Arturo González-Izquierdo, Natasha K. Martin, Steffanie A. Strathdee, Davey M. Smith, and Annick Bórquez. Identifying counties at risk of high overdose mortality burden throughout the emerging fentanyl epidemic in the united states: a predictive statistical modeling study. *Lancet Public Health*, 6(10):e720–e728, 2021.

[82] Marisa E Marraccini and Zoe MF Brier. School connectedness and suicidal thoughts and behaviors: A systematic meta-analysis. *School psychology quarterly*, 32(1):5, 2017.

[83] Matthew Miller, Deborah Azrael, and Catherine Barber. Suicide mortality in the united states: The importance of access to lethal means. *The New England Journal of Medicine*, 387:1454–1457, 2022.

[84] Matthew Miller, Yifan Zhang, David M Studdert, and Sonja Swanson. Updated estimate of the number of extreme risk protection orders needed to prevent 1 suicide. *JAMA network open*, 7(6):e2414864–e2414864, 2024.

[85] MIT Election Data and Science Lab. County presidential election returns 2000–2024. Dataset, 2018. County-level presidential election results; accessed 2025.

[86] Jody Morgan and Alison L Jones. The role of naloxone in the opioid crisis. *Toxicology Communications*, 2(1):15–18, 2018.

[87] Jerome A Motto and Alan G Bostrom. A randomized controlled trial of postcrisis suicide prevention. *Psychiatric services*, 52(6):828–833, 2001.

[88] Anna S Mueller and Seth Abrutyn. Suicidal disclosures among friends: using social network data to understand suicide contagion. *Journal of health and social behavior*, 56(1):131–148, 2015.

[89] National Institute on Drug Abuse. Drug overdose deaths: Facts and figures. https://nida.nih.gov/research-topics/trends-statistics/overdose-death-rates, 2024. Accessed August 2025.

[90] Thomas Niederkrotenthaler, Michael Braun, Jane Pirkis, Benedikt Till, Steven Stack, Mark Sinyor, Ulrich S Tran, Martin Voracek, Qijin Cheng, Florian Arendt, et al. The role of suicide-related media content in suicide contagion: a systematic review. *The Lancet Psychiatry*, 7(7):629–641, 2020.

[91] Thomas et al. Niederkrotenthaler. Association of increased youth suicides with a u.s. television series portraying suicide. *JAMA Psychiatry*, 77(2):236–243, 2020.

[92] National Institute of Mental Health. Suicide statistics. https://www.nimh.nih.gov/health/statistics/suicide, 2023. Accessed: 2025-08-16.

[93] County of San Diego Health and Human Services. San diego county suicide prevention council annual report, 2015. https://www.sandiegocounty.gov/content/dam/sdc/hhsa/programs/bhs/documents/SPC-Annual-Report.pdf.

[94] Sheetal Pandrekar, Xin Chen, Gaurav Gopalkrishna, Avi Srivastava, Mary Saltz, Joel Saltz, and Fusheng Wang. Social media based analysis of opioid epidemic using reddit. In *AMIA Annual Symposium Proceedings*, volume 2018, page 867. American Medical Informatics Association, 2018.

[95] Sarah A Pendergrass, Richard C Crist, Laney K Jones, Jason R Hoch, and Wade H Berrettini. The importance of buprenorphine research in the opioid crisis. *Molecular Psychiatry*, 24(5):626–632, 2019.

[96] Joseph V Pergolizzi Jr, Jo Ann LeQuang, Robert Taylor Jr, Robert B Raffa, and NEMA Research Group. Going beyond prescription pain relievers to understand the opioid epidemic: the role of illicit fentanyl, new psychoactive substances, and street heroin. *Postgraduate Medicine*, 130(1):1–8, 2018.

[97] Bernice A. Pescosolido. A network frame offers a promising transdisciplinary tool for understanding complex health and health care system problems like suicide. *Proceedings of the National Academy of Sciences*, 121(34), 2024.

[98] David J Peters, Shannon M Monnat, Andrew L Hochstetler, and Mark T Berg. The opioid hydra: Understanding overdose mortality epidemics and syndemics across the rural-urban continuum. *Rural Sociology*, 85(3):589–622, 2020.

[99] Pew Research Center. Social media use in 2021, 2021. Accessed: 2025-08-02.

[100] David P Phillips. The influence of suggestion on suicide: Substantive and theoretical implications of the werther effect. *American Sociological Review*, 39(3):340–354, 1974.

[101] Jane Pirkis, Jason Bantjes, Madelyn Gould, Thomas Niederkrotenthaler, Jo Robinson, Mark Sinyor, Michiko Ueda, and Keith Hawton. Public health measures related to the transmissibility of suicide. *The Lancet Public Health*, 9(10):e807–e815, 2024.

[102] Vicente Prades-Caballero, Juan-José Navarro-Pérez, and Álvaro Carbonell. Factors associated with suicidal behavior in adolescents: An umbrella review using the socio-ecological model. *Community Mental Health Journal*, 61:612–628, 2025.

[103] Jakob Gulddahl Rasmussen. Bayesian inference for hawkes processes. *Methodology and Computing in Applied Probability*, 15(3):623–642, 2013.

[104] Brad Ray and colleagues. Spatiotemporal analysis exploring the effect of law enforcement drug market disruptions on overdose, indianapolis, indiana, 2020–2021. *American Journal of Public Health*, 113(7):750–758, 2023.

[105] Alex Reinhart. A review of self-exciting spatio-temporal point processes and their applications. *Statistical Science*, 33(3):299–318, 2018.

[106] Khary K Rigg, Katherine McLean, Shannon M Monnat, Glenn E Sterner III, and Ashton M Verdery. Opioid misuse initiation: implications for intervention. *Journal of addictive diseases*, 37(3-4):111–122, 2018.

[107] Thomas Santo Jr, Gabrielle Campbell, Natasa Gisev, Daniel Martino-Burke, Jack Wilson, Samantha Colledge-Frisby, Brodie Clark, Lucy Thi Tran, and Louisa Degenhardt. Prevalence of mental disorders among people with opioid use disorder: A systematic review and meta-analysis. *Drug and Alcohol Dependence*, 238:109551, 2022.

[108] Jeffery Sauer, Kathleen Stewart, and Zachary D. W. Dezman. A spatio-temporal bayesian model to estimate risk and evaluate factors related to drug-involved emer-

gency department visits in the greater baltimore metropolitan area. *Journal of Substance Abuse Treatment*, 131:108534, 2021.

[109] Jeffrey Shaman, Sasikiran Kandula, Sen Pei, Marta Galanti, Mark Olfson, Madelyn Gould, and Katherine Keyes. Quantifying suicide contagion at population scale. *Science advances*, 10(31):eadq4074, 2024.

[110] Gopal K Singh and Irene E Kim. Rural–urban disparities in drug overdose mortality in the united states, 1999–2017. *International Journal of MCH and AIDS*, 8(2):89–101, 2019.

[111] Erin J Stringfellow, Tse Yang Lim, Keith Humphreys, Catherine DiGennaro, Celia Stafford, Elizabeth Beaulieu, Jack Homer, Wayne Wakeland, Benjamin Bearnot, R Kathryn McHugh, et al. Reducing opioid use disorder and overdose deaths in the United States: A dynamic modeling analysis. *Science Advances*, 8(25):eabm8147, 2022.

[112] Brian Suffoletto, Aaron Zeigler, Patricia M. Arenth, Clifton Callaway, and Tammy Chung. Risk and protective factors for repeated overdose after opioid use disorder detoxification. *Substance Abuse: Research and Treatment*, 14:1–9, 2020.

[113] Jeffrey W Swanson, Michael A Norko, Hsiu-Ju Lin, Kelly Alanis-Hirsch, Linda K Frisman, Madelon V Baranoski, Michele M Easter, Allison G Robertson, Marvin S Swartz, and Richard J Bonnie. Implementation and effectiveness of connecticut's risk-based gun removal law: does it prevent suicides? *Law and Contemporary Problems*, 80(2):179–208, 2017.

[114] Jeffrey W Swanson, April M Zeoli, Shannon Frattaroli, Marian Betz, Michele Easter, Reena Kapoor, Christopher Knoepke, Michael Norko, Veronica A Pear, Ali Rowhani-Rahbar, et al. Suicide prevention effects of extreme risk protection order laws in four states. *Journal of the American Academy of Psychiatry and Law*, 52(3):327–337, 2024.

[115] Sonya A. Swanson and Ian Colman. Association between exposure to suicide and suicidality outcomes in youth. *CMAJ*, 185(10):870–877, 2013.

[116] Jennifer L Syvertsen, Catherine E Paquette, and Robin A Pollini. Down in the valley: Trajectories of injection initiation among young injectors in california's central valley. *International Journal of Drug Policy*, 44:41–49, 2017.

[117] Kushagra Tiwari, M Amin Rahimian, Mark S Roberts, Praveen Kumar, and Jeanine M Buchanich. Measuring network dynamics of opioid overdose deaths in the united states. *Scientific Reports*, 14(1):29563, 2024.

[118] Behzad Vahedi, Morteza Karimzadeh, and Hamidreza Zoraghein. Spatiotemporal prediction of COVID-19 cases using inter-and intra-county proxies of human interactions. *Nature Communications*, 12(1):6440, 2021.

[119] Thomas W Valente, Anamara Ritt-Olson, Alan Stacy, Jennifer B Unger, Janet Okamoto, and Steve Sussman. Peer acceleration: effects of a social network tailored substance abuse prevention program among high-risk adolescents. *Addiction*, 102(11):1804–1815, 2007.

[120] Jenna van Draanen, Christie Tsang, Sanjana Mitra, Vanessa Phuong, Arata Murakami, Mohammad Karamouzian, and Lindsey Richardson. Mental disorder and opioid overdose: a systematic review. *Social Psychiatry and Psychiatric Epidemiology*, pages 1–25, 2022.

[121] Jorge V Verlenden. Mental health and suicide risk among high school students and protective factors—youth risk behavior survey, united states, 2023. *MMWR supplements*, 73, 2024.

[122] Washington/Baltimore High Intensity Drug Trafficking Area (HIDTA). Overdose spike response framework for communities and local health departments. https://odmap.org/Content/docs/ODMap-Overdose-Response-Framework-2018-3.29.18.pdf, 2018. Describes ODMAP 24-hour spike alerts and operational response.

[123] Yang Xu, Paolo Santi, and Carlo Ratti. Beyond distance decay: Discover homophily in spatially embedded social networks. *Annals of the American Association of Geographers*, 112(2):505–521, 2022.

[124] Tse-Chuan Yang, Stephen A. Matthews, Fei Sun, and Miguel Armendariz. Modeling the importance of within- and between-county effects in an ecological study of the association between social capital and mental distress. *Preventing Chronic Disease*, 16:E75, June 2019.

[125] April M Zeoli, Jennifer Paruk, Charles C Branas, Patrick M Carter, Rebecca Cunningham, Justin Heinze, and Daniel W Webster. Use of extreme risk protection orders to reduce gun violence in oregon. *Criminology & public policy*, 20(2):243–261, 2021.

[126] Ke Zhou, Hongyuan Zha, and Le Song. Learning social infectivity in sparse low-rank networks using multi-dimensional hawkes processes. In *Proceedings of the 16th International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 31 of *Proceedings of Machine Learning Research*, pages 641–649, 2013.