Supplemental Material for "Modeling the effect of drug courts in North Carolina counties"

1 Data

We discretized the targeting variable into three bins with cutoffs based on the terciles such that *low targeting* indicates counties whose targeting ratio was in the bottom third (below 0.107), *high targeting* indicates counties in the upper tercile (ratio above 0.217), and *average targeting* indicating counties in the middle tercile. Figure 1 contains maps of the targeting index for all counties for each year.

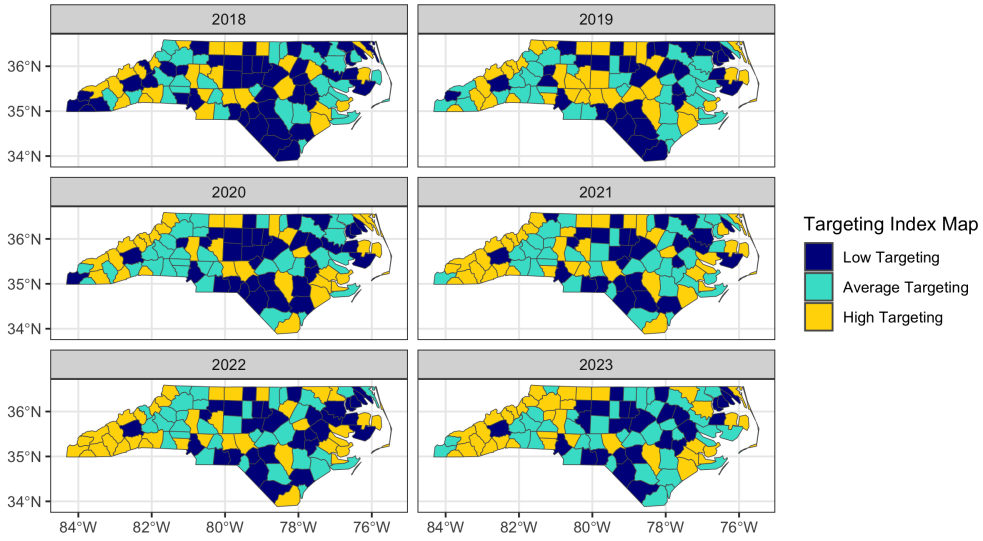

Figure 1: Map of targeting index by county by year. Values are binned into three categories to accommodate extreme outliers.

Figure 2 shows time series plots of the response variable, observed illicit opioid overdose death rates, for each county in NC. We observe a sharp increase from 2019 to 2020, coinciding with the onset of the COVID-19 pandemic. The positive trend continues until 2023, when rates begin to fall. Maps of the county-level rates for each year can be found in the appendix.

Table 3 contains the standardized mean differences between counties which do and do

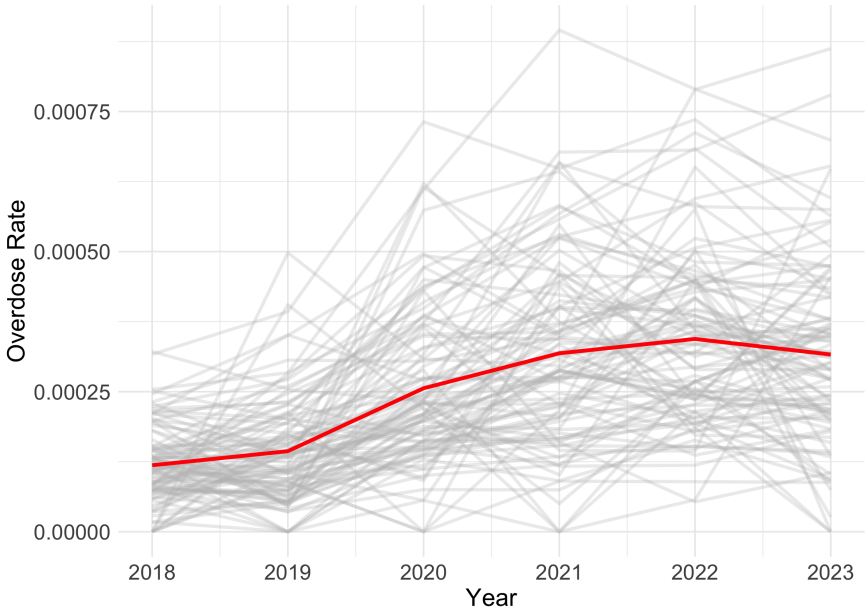

Figure 2: Observed illicit opioid deaths per capita for each county (gray lines) in NC. The red line shows the state average rate.

not contain drug courts. Note that each county is counted six times here, once for each observation year, and thus our clusters share an underlying temporal correlation. However, many counties began a drug court program within this window. It is unclear from prior literature and this study the time frame which a drug court impacts the rate change of overdose rates in the county which it is present in.

Below are the weights for our PCA vectors, representing socioeconomic status and healthcare accessibility.

	Values
% Disabled	0.39
% Without High School Diploma	0.46
% on Food Stamps	0.50
% Under Poverty Line	0.49
% Unemployed	0.38
Proportion of Variance	0.66

Table 1: The first PCA vector measuring poverty. These are the weights for each element and the variance captured by the PCA.

Table 4 contains the framework by which counties were designated as urban or rural.

Characteristic	Drug Court N = 130[1]	No Drug Court N = 470[1]	Difference[2]	95% CI[2]
Death Count	33 (11, 89)	11 (4, 21)	0.93	0.73, 1.1
Population	152,908 (44,481, 329,849)	45,136 (21,958, 80,547)	0.86	0.66, 1.1
Crime Rate	0.25 (-0.79, 1.37)	-0.16 (-0.61, 0.42)	0.42	0.22, 0.61
HIDTA	36 (28%)	54 (11%)	0.42	0.22, 0.61
DA Party	59 (45%)	173 (37%)	0.17	-0.02, 0.37
Urban/Rural	90 (69%)	180 (38%)	0.65	0.45, 0.85
% White and Asian	0.50 (-0.75, 1.09)	-0.05 (-0.77, 0.89)	0.23	0.04, 0.43
Poverty Index	-1.09 (-2.23, -0.08)	0.16 (-0.90, 1.48)	-0.81	-1.0, -0.62
Healthcare Index	-1.32 (-1.75, 0.42)	0.33 (-0.28, 0.83)	-0.89	-1.1, -0.69
Targeting Index Bin			0.38	0.19, 0.58
Average	28 (22%)	173 (37%)		
High	43 (33%)	156 (33%)		
Low	59 (45%)	141 (30%)		
Death Rate	0.0002 (0.0001, 0.0003)	0.0002 (0.0001, 0.0004)	-0.07	-0.26, 0.13

[1] Median (Q1, Q3); n (%)

[2] Standardized Mean Difference

Abbreviation: CI = Confidence Interval

Figure 3: Standardized mean differences between counties with and without drug courts. We had 130 county-year combinations with drug courts present out of 600 observations. Our first two columns contain the median and inter-quartile range of our counties with and without drug courts across each covariate. The third and fourth columns contain the standardized mean difference between each DC and NDC counties, along with a 95% confidence interval for this estimate.

	Values
Medically Underserved Area Designation	0.69
Number of OTPs	-0.41
% Uninsured	0.60
Proportion of Variance	0.43

Table 2: The second PCA vector measuring medical accessibility. These are the weights for each element and the variance captured by the PCA.

2 Supplemental Statistical Modeling Details

Let $Y_{i,t}$ denote the observed illicit opioid overdose death count in county $i = 1, ..., n = 100$ during year $t = 1, ..., T = 7$ where $t = 1$ corresponds to 2017 and $t = 7$ corresponds to 2023. We assume a Poisson autoregressive model such that $Y_{i,t} \mid \mu_{it} \sim \text{Poisson}(\mu_{i,t})$, where $\mu_{i,t}$ is the expected death count for that county in that year. To first quantify an overall effect of drug courts, we let

$$\log(\mu_{i,t}) = \begin{cases} \log(P_{i,1}) + \alpha_i & \text{for } t = 1 \\ \log\left(\frac{P_{i,t}}{P_{i,t-1}}\right) + \log(\mu_{i,t-1}) + \beta_{0i} + \mathbf{X}_{i,t}\boldsymbol{\beta} + \mathrm{I}_{i,t}\delta_0 & \text{for } t = 2, ..., T, \end{cases} \tag{1}$$

where $P_{i,t}$ denotes the population of county i in year t. For the first year, we establish a baseline rate for each county and assume $\alpha_i \mid \alpha \sim N(\alpha, \sigma^2)$. The mean α relates to the statewide average death rate while α_i captures county-level variability about the mean for county i in 2017. Note that the baseline year is not scientifically of interest, but it must be defined so that we can analyze how the death rate changes over time. $\boldsymbol{\beta}$ is a p-dimensional vector of regression coefficients. $\mathrm{I}_{i,t}$ is an indicator of whether county i had a drug court in year t or not, and δ_0 the corresponding coefficient. We let each county have it's own intercept, β_{0i}, which will be specified to also account for spatial dependence between counties.

$\mathbf{X}_{i,t}$ is a p-dimensional vector of covariates for county i in year t. It includes an indicator variable for the presence of a drug court, an indicator for low and high targeting, the scaled crime rate, an indicator for HIDTA designation, an indicator for the district attorney's

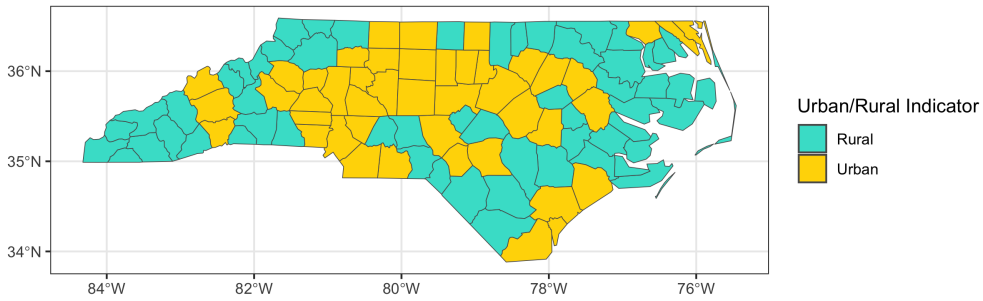

Figure 4: 2023 Rural-Urban Continuum Codes. Criteria for each level of classification.

political affiliation where Democrats are ones and Republicans are zeros, an indicator for the county's urban/rural designation where urban areas are ones and rural areas are zeros, and the poverty and healthcare indices. We assume the spatially-varying intercept is of the form $\beta_{0i} = \beta_0 + \mathbf{K}_i\mathbf{\Gamma}$, where β_0 represents an overall state-level intercept, $\mathbf{K}_i$ is a q-dimensional vector of Moran's I basis functions, and $\mathbf{\Gamma}$ a q-dimensional vector of coefficients. Moran's I basis functions are used to reduce the dimension of the spatial random effect while accounting for spatial autocorrelation and also eliminating confounding with covariates (Hughes and Haran, 2013). $\mathbf{K}_{i,t}$ is a vector containing the first $q = 20$ eigenvectors with

positive eigenvalues of $\mathbf{QAQ}$, where A is the adjacency matrix and

$$Q = \mathbf{I} - \mathbf{X}_{2018}(\mathbf{X}_{2018}^T\mathbf{X}_{2018})^{-1}\mathbf{X}_{2018}^T.$$

Thus Q forms the projection of our errors onto the orthogonal basis for the complement of the design matrix from 2018, $\mathbf{X}_{2018}$, ensuring spatial random effects which are linearly independent of $\mathbf{X}_{2018}$ in line with the methodology described in Hughes and Haran (2013).

Observe that for $t = 2, ..., T$, Equation (1) is equivalent to

$$\log\left(\frac{\mu_{i,t}/P_{i,t}}{\mu_{i,t-1}/P_{i,t-1}}\right) = \beta_{0i} + \mathbf{X}_{i,t}\boldsymbol{\beta} + \mathrm{I}_{i,t}\boldsymbol{\delta}_0.$$

Thus, this model allows us to quantify the relationship between the covariates and the *change in* the overdose death rate. That is, this model is already accounting for baseline heterogeneity across counties, allowing us to quantify how drug courts relate to improving or worsening illicit opioid overdose death rates, relative to each county's historical rate.

We fit our model in the Bayesian framework and must assume prior distributions for all model parameters. We assumed weakly informative conjugate priors throughout. More specifically, we assume all regression coefficients, $\boldsymbol{\beta}$, β_0, α, δ_0, and $\boldsymbol{\Gamma}$ are independent, mean-zero normal distributions with standard deviation 10. We assume the baseline heterogeneity variance parameter σ^2 has an inverse gamma prior distribution with shape and scale parameters of 0.1. We fit our model by implementing a Metropolis-within-Gibbs Markov chain Monte Carlo (MCMC) algorithm using the R package NIMBLE (Valpine et al., 2024). We performed 1,000,000 iterations with a burn-in of 500,000. The output was thinned such that only every 5th iteration was stored. This was run on the DEAC High Performance Computing cluster (Information Systems and Wake Forest University, 2021).

The second model we fit modifies equation (1) to instead be

$$\log(\mu_{i,t}) = \begin{cases} \log(P_{i,1}) + \alpha_i & \text{for } t = 1 \\ \log\left(\frac{P_{i,t}}{P_{i,t-1}}\right) + \log(\mu_{i,t-1}) + \beta_{0i} + \mathbf{X}_{i,t}\boldsymbol{\beta} + \mathrm{I}_{i,t}\delta_0 + \mathbf{X}_{i,t}\mathrm{I}_{i,t}\boldsymbol{\delta} & \text{for } t = 2, ..., T, \end{cases} \tag{2}$$

where $\boldsymbol{\delta}$ is a p-dimensional vector of regression coefficients for the interaction terms. That is, $\boldsymbol{\delta}$ quantifies how the effect of drug courts vary according to covariates in $\mathbf{X}_{i,t}$. We assume weakly informative normal prior distributions with standard deviation 10 for each regression coefficient in $\boldsymbol{\delta}$. The second model is also fit in the Bayesian paradigm under the same specifications previously described.

To simulate our first counterfactual outcomes, we set $I_{i,t} = 0$ for all i and t. We simulate from the posterior predictive distribution of the expected count under this assumption, given by

$$\log\left(\mu_{i,t}^{(-D)}\right) = \log\left(\frac{P_{i,t}}{P_{i,t-1}}\right) + \log\left(\mu_{i,t-1}^{(-D)}\right) + \beta_{0i} + \mathbf{X}_{i,t}\boldsymbol{\beta}, \tag{3}$$

for $t = 2, ..., T$. We estimate the expected difference in death counts, $\mu_{i,t}^{(-D)} - Y_{i,t}$, and the difference in death rates, $(\mu_{i,t}^{(-D)} - Y_{i,t})/P_{i,t}$, for each county by simulating from the posterior predictive distribution of these quantities.

For our second counterfactual, we set $\mathrm{I}_{i,t} = 1$ for all i, t, and simulate from the posterior predictive distribution of the expected count

$$\log\left(\mu_{i,t}^{(+D)}\right) = \log\left(\frac{P_{i,t}}{P_{i,t-1}}\right) + \log\left(\mu_{i,t-1}^{(+D)}\right) + \beta_{0i} + \mathbf{X}_{i,t}\boldsymbol{\beta} + \delta_0 + \mathbf{X}_{i,t}\boldsymbol{\delta}, \tag{4}$$

for $t = 2, ..., T$. We estimate the effect of universal adoption by estimating $\mu_{i,t}^{(+D)} - Y_{i,t}$, and the difference in death rates, $(\mu_{i,t}^{(+D)} - Y_{i,t})/P_{i,t}$, for each county.

For our third counterfactual, we modified the high targeting indicator to all equal 0's in addition to using $\mathrm{I}_{i,t} = 1$.

3 Supplemental Results

Table 3 shows the estimated regression coefficients for model 1, along with the 95% credible intervals and percentage of posterior probability that is positive. Table 4 shows estimates of the regression coefficients for the main effects terms in model 2, corresponding to estimates for counties with no drug courts. Table 5 shows the estimates of $\boldsymbol{\beta} + \boldsymbol{\delta}$, which corresponds to the estimated regression coefficients for counties with drug courts.

Variable	Estimate	95% CI	% Positive
Drug Court	-0.03	-0.06, 0.01	0.05
Intercept	0.16	0.14, 0.18	1.00
Low Targeting	0.02	-0.01, 0.05	0.91
High Targeting	0.05	0.02, 0.08	1.00
Crime Rate	0.04	0.03, 0.05	1.00
HIDTA	-0.01	-0.03, 0.01	0.22
DA Party	-0.01	-0.03, 0.02	0.32
Urban	0.04	0.01, 0.06	1.00
Poverty Index	0.02	0.01, 0.02	1.00
Healthcare Index	0.02	0.01, 0.03	1.00

Table 3: Estimates of the first model's $\boldsymbol{\beta}$ regression coefficients along with the lower and upper bounds of the 95% credible intervals. The last column is the estimated percentage of the posterior distribution that was positive. The variables that are bold correspond to regression coefficients where the 95% credible interval does not contain zero.

Variable	Estimate	95% CI	% Positive
Intercept	0.17	0.14, 0.20	1.00
Low Targeting	0.04	-0.01, 0.08	0.95
High Targeting	0.04	0.01, 0.08	0.99
Crime Rate	0.05	0.03, 0.07	1.00
HIDTA	-0.02	-0.05, 0.02	0.17
DA Party	-0.01	-0.04, 0.02	0.25
Urban/Rural Code	0.03	0.00, 0.07	0.99
Poverty Index	0.01	0.00, 0.02	0.99
Healthcare Index	0.03	0.02, 0.05	1.00

Table 4: Estimates of the second model's $\boldsymbol{\beta}$ regression coefficients along with the lower and upper bounds of the 95% credible intervals. The last column is the estimated percentage of the posterior distribution that was positive. The variables that are bold correspond to regression coefficients where the 95% credible interval does not contain zero. These parameters represent associations with overdose deaths in counties without drug courts.

Variable	Estimate	95% CI	Percent Positive
Intercept	-0.00	-0.08, 0.08	0.48
Low Targeting	-0.00	-0.06, 0.06	0.49
High Targeting	0.23	0.14, 0.33	1.00
Crime Rate	0.07	0.04, 0.09	1.00
HIDTA	-0.05	-0.09, -0.00	0.02
DA Party	0.04	-0.01, 0.10	0.94
Urban/Rural Code	0.08	0.00, 0.15	0.98
Poverty Index	0.01	-0.00, 0.03	0.95
Healthcare Index	-0.02	-0.05, 0.01	0.14

Table 5: Summary of distribution for the random variable $\boldsymbol{\beta} + \boldsymbol{\delta}$. This represents the coefficients for counties with drug courts. The HDI and positive posterior probability are calculated from the distribution obtained by taking the sum of each coefficient for each sample in our MCMC.

4 Model Fit

Figure 5 shows the predicted time series for each county under our model, and Table 6 shows the average quantile residuals for each year. Note that our model largely overestimates death counts in 2018, 2019, and 2023, and it underestimated in 2020 - 2022. We hypothesize that the COVID-19 pandemic is related to the jump in overdose deaths in 2020, and our model did not explicitly account for that. We also note that the observed data showed a decrease in death counts for most counties in 2023, but our model did not allow for declining trends. With such a short time period, we wanted to avoid overfitting the data. If more years of data become available, we might want to consider a model that allows non-monotonic temporal trends.

Year	Mean rQ
2018	-0.47
2019	-0.85
2020	0.50
2021	0.83
2022	0.35
2023	-0.81

Table 6: Average Residual by Year.

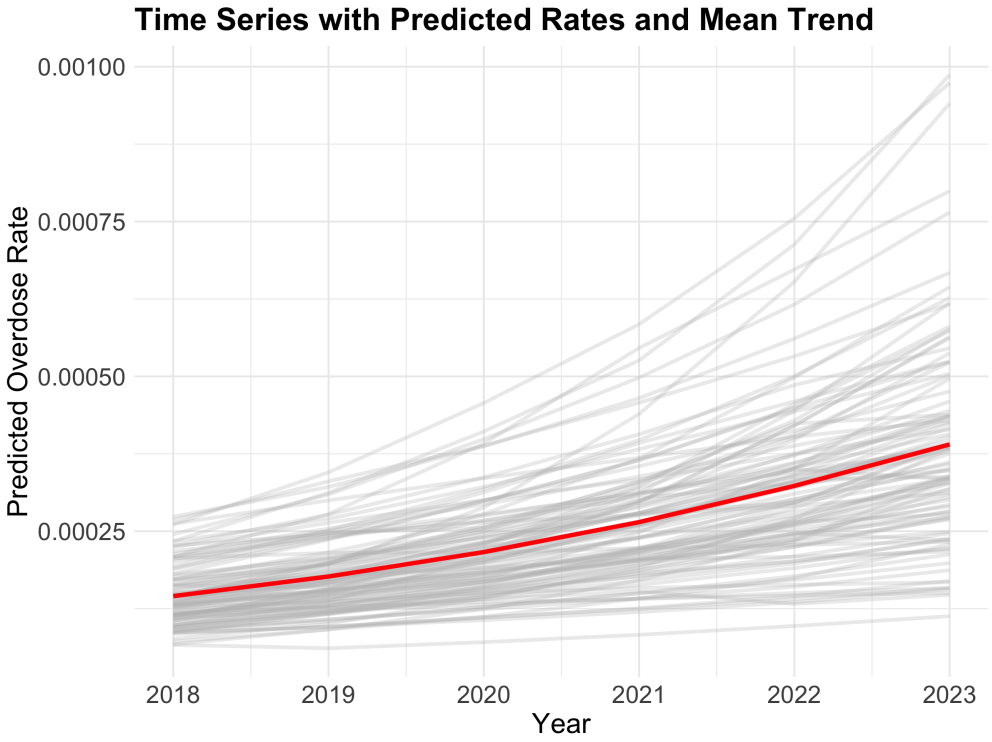

Figure 5: Time series of predicted illicit opioid deaths per capita for each county in NC. The red line shows a predicted state average rate.