Explaining and Forecasting Homelessness in the United States

Konstantinos Christopher Tsiolis

Friday, April 19, 2024

Abstract

We implement a Bayesian state-space model with the dual objectives of projecting future homelessness rates in each of the 50 US states and investigating the association between homelessness and a collection of demographic and economic factors. Homelessness data is obtained from the US Department of Housing and Urban Development, while covariates are extracted from the US Census Bureau, the US Environmental Protection Agency, and Forbes Media data. We model fluctuations over time via second-order random walks and temporal smoothing with P-splines, finding that the former yields superior prediction performance on held-out data. Our final model identifies median monthly housing cost as having a significant association with state homelessness rate.

1 Introduction

Homelessness is a major issue across the United States. According to the most recent estimates from the US Department of Housing and Urban Development (HUD), 653,100 Americans experienced homelessness in January 2023, the largest total since reporting began in 2007 [de Sousa et al., 2023]. States of emergency have been declared in various jurisdictions — including Los Angeles [City of Los Angeles, 2023], Denver [Zelinger, 2023], Seattle [Daniels, 2023], and Alameda County [Brinkley, 2023] — with the aim of streamlining the access to financial resources and the enactment of policies to respond to the crisis.

People experiencing homelessness are at higher risk for a variety of health problems and in particular are at a higher risk of mortality [Wyse et al., 2023]. They are often subject to social stigma that blames them for their situation, limits their ability to participate in society, and impedes their access to health care Belcher and DeForge [2012]. Their experience is a direct violation Article 25 of the Universal Declaration of Human Rights, which recognizes "the right to a standard of living adequate for the health and well-being of [themselves] and of [their] family, including food, clothing, housing and medical care and necessary social services" [United Nations General Assembly, 1948].

The objective of this work is to develop a statistical model in a Bayesian framework to identify predictors associated with homelessness in the United States and to produce state-level forecasts of the rate of homelessness in the population. We take inspiration from [Alexander et al., 2022], who develop a Bayesian state-space model that simultaneously achieves these two objectives of identification and forecasting in the context of the American foster care system. The above paper highlights the two primary benefits of such a model, which translate naturally to our application. First, the identification of predictors most associated with homelessness can give federal, state, and local governments a sense of what policies are more likely to effectively prevent homelessness (e.g., subsidized housing, infrastructure projects, social security reform). Second, forecasting can lead to a more informed allocation of resources (e.g., shelters, food banks) to respond to demand.

Previous work has used survey data to identify factors affecting homelessness in youth. For example, Shelton et al. [2009] identify those who have experienced homelessness out of the 14,888 people surveyed in the National Longitudinal Study of Adolescent Health and employ a logistic regression model with covariates such as childhood experiences of abuse and criminal history of the parents. Closer to our work, Byrne et al. [2013] leverage population-level homelessness data by state and CoC (sub-state-level administrative unit for coordinating homelessness response) from HUD to develop a log-linear hierarchical model that predicts homelessness at a fixed point in time based on rental cost, vacancy rate, and poverty rate, among others. In a more recent study, Glynn and Fox [2019] use the same data — along with population data from the US Census Bureau and rental data from Zillow — to fit a "dynamic Bayesian hierchical model for time-varying homeless data". They model the relationship between rent increases and increases in the homeless population in the 25 largest US metropolitan areas.

2 Data

We combine the aforementioned annual data on homelessness by state from HUD with data from other US government agencies (primarily the Census Bureau) on population, housing, and poverty, as well as data compiled by Forbes on affordability. The homelessness data consists of annual Point-in-Time (PiT) estimates of the number of people experiencing homelessness in each state from 2007 to 2023 [de Sousa et al., 2023]. Here, PiT means that estimates were made on a specific night of the given year between January 22nd and 31st. The count groups people experiencing homelessness based on whether they are sheltered, their age, their gender, and their race.[1] Counts are conducted separately by each of the 381 continuums of care (CoCs). The CoCs are granted a degree of flexibility with regards to the methodology, but a national standard is set by HUD. This standard permits counting unsheltered individuals by sending enumerators to cover the CoC's entire area, by covering only "known locations where people who are unsheltered are located at night", or by taking a random sample of areas within the CoC and extrapolating [United States Department of Housing and Urban Development, 2014]. Sheltered counts can similarly be conducted via census or random sample and extrapolation, but these have the added benefit of exploiting the Homeless Management Information System (HMIS), which many shelters participate in and whose data conforms to HUD's standards.

Both due to the methodological discrepancies outlined above and the difficulty of conducting a count in a single night, these counts — especially the unsheltered counts — are likely to be underestimates. This is most pronounced in 2021 and 2022, when the COVID-19 pandemic forced shelters to implement social distancing measures or close altogether. Additionally, the data contains artifacts such as when in 2023, 22 CoCs did not conduct an unsheltered count in 2023 and used the count from 2022 [de Sousa et al., 2023].

We obtain population estimates for July 1st of each year from the US Census Bureau (USCB), which we use to calculate *homelessness rate* — our dependent variable of interest — the proportion of the population experiencing homelessness. The population counts from the census years of 2010 and 2020 are the most reliable, while intercensal estimates leverage vital registration systems for birth and death count estimation, income tax, medicare, and social security data for intra-US migration, the American Community Survey (ACS) for immigration, and registries of other countries for emigration estimation [United States Census Bureau, 2023]. Additionally, we incorporate estimates of the proportion of the population in living in urban areas by state during each of the census years. The USCB currently dictates that an urban area "must encompass at least 2000 housing units or at least 5000 people", in addition to a set of density requirements [United States Census Bureau, 2020]. However, it is worth noting that the criteria in the 2010 Census were stricter. Furthermore, to incorporate population density as a covariate, we obtain the land area of each state from the USCB. The choice of these two covariates is motivated by the often-made assumption that homelessness is primarily an issue in large cities, a notion that has increasingly been refuted in recent years [Meehan, 2019].

Our remaining covariates are economic indicators which naturally lead to an analysis of homelessness through the lens of availability of affordable housing. We incorporate annual data on the poverty rate (only available until 2022) and gross vacancy rate (across all housing units) by state from the USCB. The poverty threshold increases from year to year and is dependent on age, family size, and number of children (but not on state). For example, the poverty threshold in 2022 for an individual under the age of 65 living on their own was 15,225 USD. Lastly, we consider data compiled by Forbes on median monthly housing cost in 2021 (adjusted to 2023 USD) and cost of living in 2023 [Rothstein and Jennings, 2024]. The former is sourced to the USCB

[1] For our purposes we only consider the total count for each state. We were unable to easily extract data on state population by age, race, and gender from the US Census Bureau's interface. Incorporating these categories into our hierarchical models is an interesting avenue for future work.

(though we were unable to recover the original data) and the latter is an aggregation of data from the Council for Community and Economic Research (C2ER), KFF, the MIT Living Wage Calculator, and the USCB.

Unfortunately, the Forbes datasets do not account for the District of Columbia (which is not a state but is treated equivalently by the USCB). It is worth noting that DC is anomalous in several ways. Its population density (which peaked in 2019 with over 10,000 people per square mile) is almost an order of magnitude larger than any of the states (New Jersey had a density of 1065 people per square mile in 2023), and its entire area is urban. In 2016, its homelessness rate was more than twice as high as any of the states. We elect to filter out DC from our data and focus on the 50 states.

In Figure 1, we observe interesting patterns when plotting the evolution of the homelessness rate for the 50 states. While trends for individual states vary, it is noteworthy that states within the same region generally exhibit similar trends. All of the southern states saw a steady decline in the homelessness rate during the 2010s. On the other hand, the states on the Pacific coast of the mainland have seen steady increases since the mid-2010s interrupted only by a large drop in 2021. This drop can be explained both by the aforementioned changes in the policies of homeless shelter and " the strengthening of safety net programs, income protections, and eviction moratoria" during the COVID-19 pandemic [de Sousa et al., 2023]. Because of this, we choose to treat 2021 as an anomaly and do not input the homelessness rate from that year into our models (though we do keep the covariates for that year). Indeed, the time series models introduced in the next section display a much improved fit to the data when we treat the 2021 rates as missing and interpolate. In 2022 and 2023, the Pacific states returned to their pre-pandemic trend. Meanwhile, two northeastern states — Vermont and New York — exhibit substantial post-pandemic spikes. Vermont's public news agency highlights "[an] extremely tight housing market[,] soaring rents, the end of the pandemic-era eviction ban[, and] the tightening of eligibility for Vermont's emergency hotel housing program" as the key contributing factors [Elder-Connors, 2023]. Meanwhile, New York City has attributed a marked increase in homelessness from 2022 to 2023 to a large influx of asylum seekers sent to the city by the State of Texas [Office of the New York City Comptroller Brad Lander, 2023]. For easier visualization of the above-mentioned points, we also include a plot for a selection of six states in Figure 2.

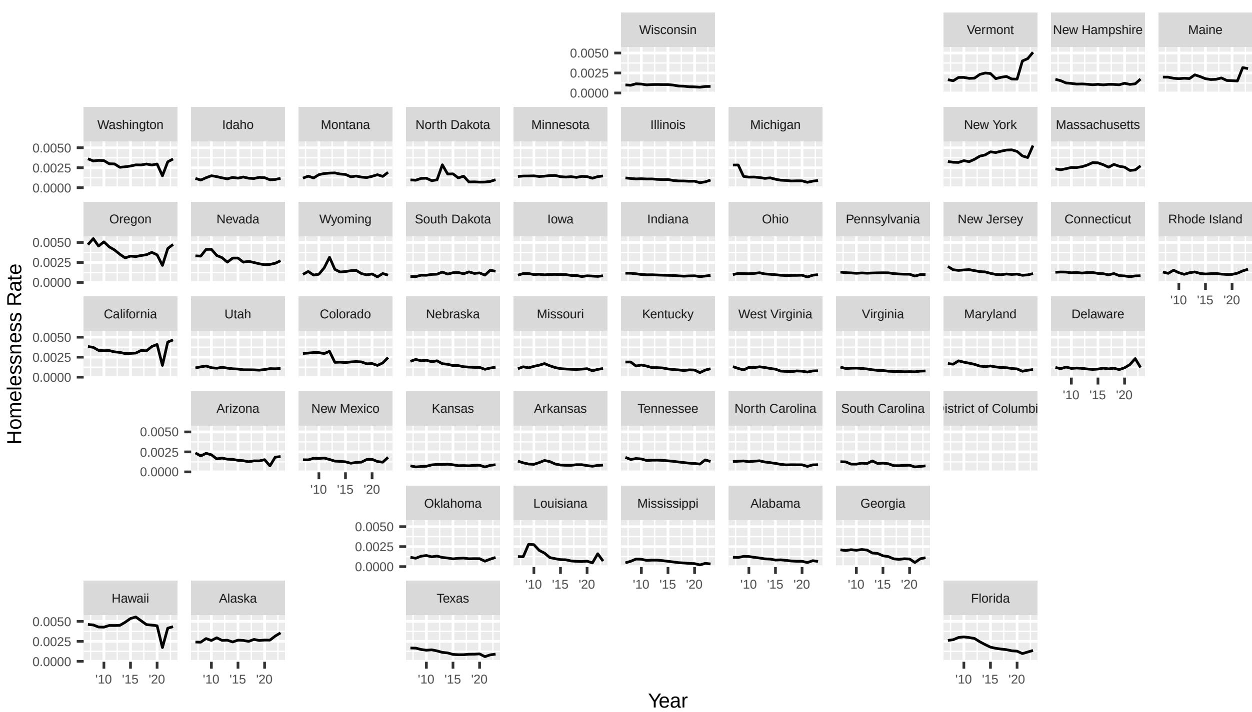


Figure 1: The evolution of homelessness rate by state for the years HUD has collected data (2007-2023).

In investigating the correlation between our covariates (Figure 3), we find a near-perfect positive correlation

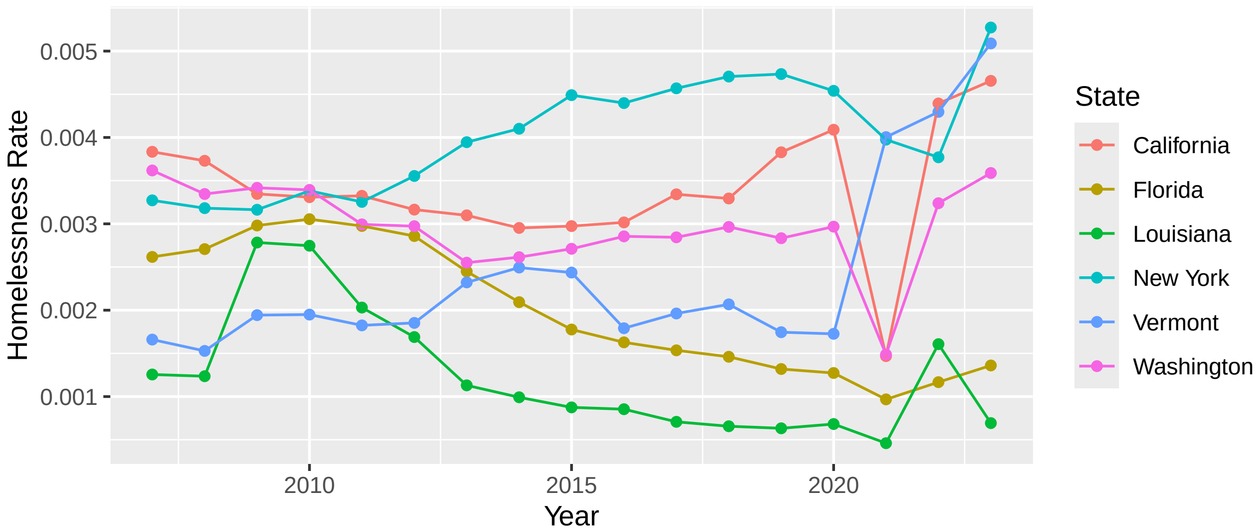


Figure 2: The evolution of homelessness rate in select states, 2007-2023.

(0.956) between median monthly housing cost and cost of living. We also observe notable correlations between urban proportion and population density, housing cost, and vacancy rate (0.512, 0.682, and -0.448 respectively). This motivates us to drop cost of living and urban proportion in our models. Of the four covariates that remain, only median monthly housing cost exhibits a noticeable association with homelessness rate in our exploratory data analysis. We plot this association in Figure 4.

3 Method

We implement a Bayesian state space model similar to the one in Alexander et al. [2022]. This class of model is designed to simultaneously handle responses and covariates that vary with time and is naturally suited to hierarchical modelling. The first point has a caveat in our case. Though all of our covariates are time-dependent, three of them are only available for specific years (urban proportion in 2010 and 2020, housing cost in 2021, and cost of living in 2023). Having eliminated two of these during our exploratory data analysis, we treat the remaining one — median housing cost — as time-independent. As for the second point, hierarchical modelling is particularly useful since we have a relatively small dataset (50 states and 17 years per state). We find that pooling states by region (the UCSB defines four regions: West, South, Northeast, Midwest) leads to better convergence of the posterior sampler (as measured by effective sample size and the Gelman-Rubin diagnostic $\hat{R}$). Moreover, this pooling is very natural given the observation in our EDA that states within the same region tend to exhibit similar trends.

We model the natural logarithm of the homelessness rate in a given state and year $(y_{s,t})$ as a normal random variable with mean $(\mu_{s,t})$ parametrized as a combination of a state-specific intercept (α_s), a linear combination of the time-dependent covariates with time-varying slopes $(x_{s,t}^T \beta_t)$, a linear combination of the time-independent covariates $(z_s^T \gamma)$, and a state and time dependent fluctuation term $(\varepsilon_{s,t})$. We place second-order random walk priors on the time-dependent coefficients β and the state-time fluctuation terms ε. We briefly attempted having separate β coefficients for each region, but the resulting sampler did not converge. We model variances hierarchically, grouping states by region as in Alexander et al. [2022]. We denote the region of a state s by $r(s)$. We write the full model specification below:

$$\begin{aligned}
\log y_{s,t} &\sim \mathcal{N}(\mu_{s,t}, \sigma^2_{y,r(s)}) \\
\mu_{s,t} &\sim \alpha_s + x_{s,t}^T \beta_t + z_s^T \gamma + \varepsilon_{s,t} \\
\alpha_s &\sim \mathcal{N}(\mu_{\alpha,r(s)}, \sigma^2_{\alpha,r(s)}) \\
\mu_\alpha &\sim \mathcal{N}(\lambda_{\mu_\alpha}, \tau^2_{\mu_\alpha})
\end{aligned}$$

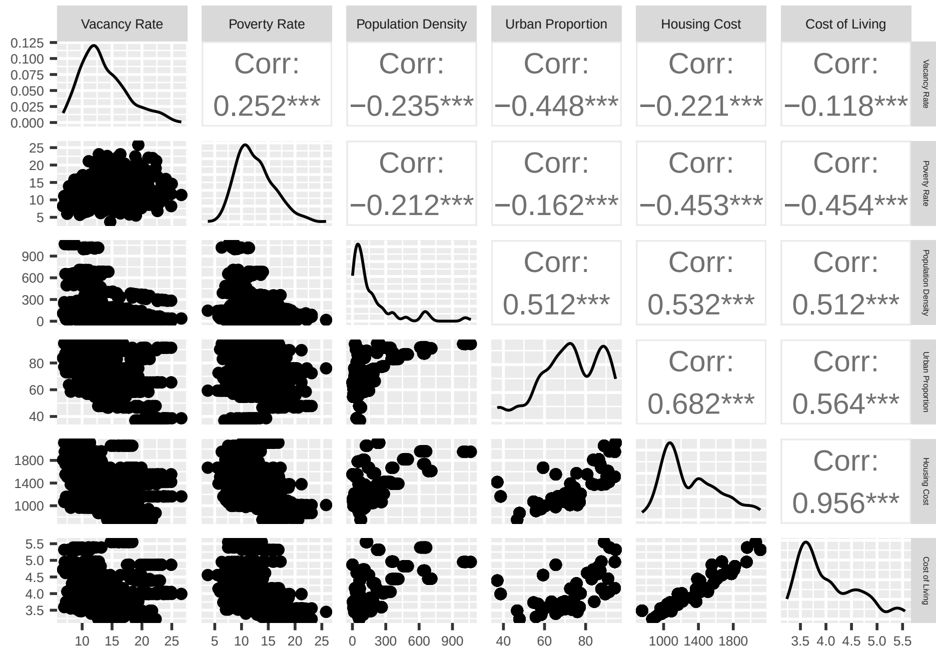


Figure 3: Pairs plot of the covariates under consideration. Cost of living is in tens of thousands of 2023 US dollars.

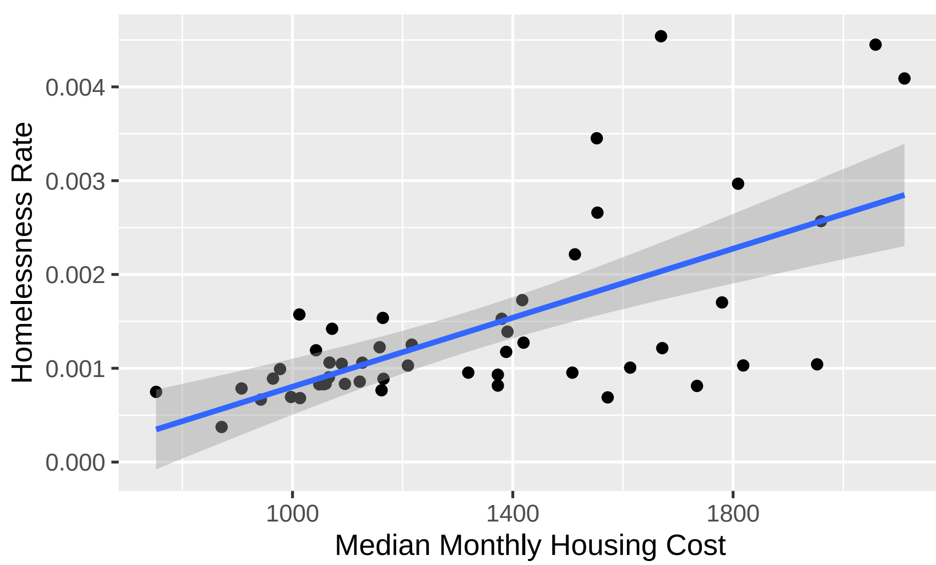


Figure 4: Plot of homelessness rate (in 2020) against median monthly housing cost (in 2021, reported in 2023 USD). We plot the response for 2020 instead of 2021 due to the aforementioned underreporting. $R^2 = 0.41$.

$$\begin{aligned}
\log \sigma_\alpha &\sim \mathcal{N}(\nu_{\sigma_\alpha}, \tau^2_{\sigma_\alpha}) \\
\beta_{1,k} &\sim \mathcal{N}(0, \sigma^2_{\beta,k}) \\
\beta_{2,k} &\sim \mathcal{N}(\beta_{1,k}, \sigma^2_{\beta,k}) \\
\beta_{t,k} &\sim \mathcal{N}(2\beta_{t-1,k} - \beta_{t-2,k}, \sigma^2_{\beta,k}) \text{ for } t > 3 \\
\log \sigma_\beta &\sim \mathcal{N}(\nu_{\sigma_\beta}, \tau^2_{\sigma_\beta}) \\
\gamma &\sim \mathcal{N}(\mu_\gamma, \sigma^2_\gamma) \\
\varepsilon_{1,s} &\sim \mathcal{N}(0, \sigma^2_{\varepsilon,r(s)}) \\
\varepsilon_{2,s} &\sim \mathcal{N}(\varepsilon_{1,s}, \sigma^2_{\varepsilon,r(s)}) \\
\varepsilon_{t,s} &\sim \mathcal{N}(2\varepsilon_{t-1,s} - \varepsilon_{t-2,s}, \sigma^2_{\varepsilon,r(s)}) \text{ for } t > 3 \\
\log \sigma_\varepsilon &\sim \mathcal{N}(\nu_{\sigma_\varepsilon}, \tau^2_{\sigma_\varepsilon}) \\
\log \sigma_y &\sim \mathcal{N}(\nu_{\sigma_y}, \tau^2_{\sigma_y}) \\
\lambda_{\mu_\alpha} &\sim \mathcal{N}(-6, 1) \\
\nu, \mu_\gamma &\sim \mathcal{N}(0, 1) \\
\tau, \sigma_\gamma &\sim \mathcal{N}^+(0, 1).
\end{aligned}$$

The choice of the $\mathcal{N}(-6, 1)$ prior for λ_{μ_α} is motivated by the observation that state homelessness rates tend to be between $e^{-7} \approx 0.0009$ and $e^{-5} \approx 0.0067$. We observe better convergence compared to the less informative $\mathcal{N}(0, 1)$ prior.

We additionally consider temporal smoothing with P-splines as an alternative to the second-order random walk prior on the state-time fluctuations. In this setting, we model $\varepsilon_{s,t}$ as a linear combination of cubic B-spline basis functions evaluated at t. Then, the spline coefficients — instead of the fluctuations themselves — are modelled as a second-order random walk over knot points. That is, letting J denote the number of knots in the spline basis[2] and $\{B_j\}_{j=1}^J$ denote the cubic spline basis functions, we model

$$\begin{aligned}
\varepsilon_{s,t} &= \sum_{j=1}^{J} \eta_{s,j} B_j(t) \\
\eta_{s,j} &\sim \mathcal{N}(2\eta_{s,j-1} - \eta_{s,j-2}, \sigma^2_{\eta,r(s)}) \\
\log \sigma_{\eta,s} &\sim \mathcal{N}(\nu_{\sigma_\eta}, \tau^2_{\sigma_\eta}) \\
\nu_{\sigma_\eta} &\sim \mathcal{N}(0, 1) \\
\tau_{\sigma_\eta} &\sim \mathcal{N}^+(0, 1).
\end{aligned}$$

We fit all models using Stan, which runs a variant of Hamiltonian Monte Carlo (HMC) to sample from the joint posterior distribution of the model parameters. For each model, we run 4 chains with 1000 warmup iterations and 1000 sampling iterations. We take particular interest in the marginal posterior of the means $\mu_{s,t}$ (since this is the logarithm of our prediction for state s at time t) and of the coefficients β_t and γ (for the two models with covariates). As such, we report the smallest effective sample size across these three sets of parameters. We also report the largest Gelman-Rubin diagnostic $\hat{R}$ observed across all parameters and generated quantities to measure the mixing of the four chains. We also investigated trace plots of parameters with large $\hat{R}$, but we omit these from this report for the sake of brevity (and since we were able to achieve $\hat{R} < 1.05$ for all four models).

Recalling our goal of interpretation, we consider the 95% credible intervals of the parameters β and γ for our models with covariates to conclude whether a significant (linear) interaction exists between covariates and the (log) homelessness rate. As for our goal of projection, to generate predictions $e^{\mu_{s,t}}$ for future years,

[2] We take evenly spaced knots with a spacing of 2.5 years.

we leverage the random walk priors on time-dependent parameters. As in Alexander et al. [2022], we take a three-year rolling average of time-dependent covariates. We fit our models on the data from 2007 to 2022 (excluding 2021) and project forward five years. We compare models quantitatively via the mean squared prediction error on the 2023 data. We perform a qualitative comparison by inspecting the posterior distribution of $e^{\mu_{s,t}}$, with the aim of identifying the model that smoothly fits the data from 2007 to 2022 (i.e., does not overfit) and exhibits the lowest variance in its projections.

4 Results

We summarize our main quantitative results in Table 1. As mentioned in the previous section, we observe $\hat{R} < 1.05$ for all models, indicating satisfactory mixing of the Markov chains in the posterior sampler. Effective sample size is lower for the two models with covariates (perhaps since these models have a larger number of parameters than the no-covariate baselines), indicating greater uncertainty in our parameter estimates. This is most pronounced for the spline model (lowest ESS of 101 with all covariates, compared to 455 without covariates). Of the four models, the random walk model with all covariates achieves the lowest prediction error on the 2023 data (MSE 1.995×10^{-5}), followed by its counterpart without covariates (MSE 2.322×10^{-5}). The two spline models lag far behind.

Method	ESS	$\hat{R}$	Validation MSE $\times 10^{-5}$
Random Walk, No Covariates	413	1.034	2.322
Spline, No Covariates	455	1.017	5.468
Random Walk, All Covariates	323	1.024	1.995
Spline, All Covariates	101	1.030	3.975

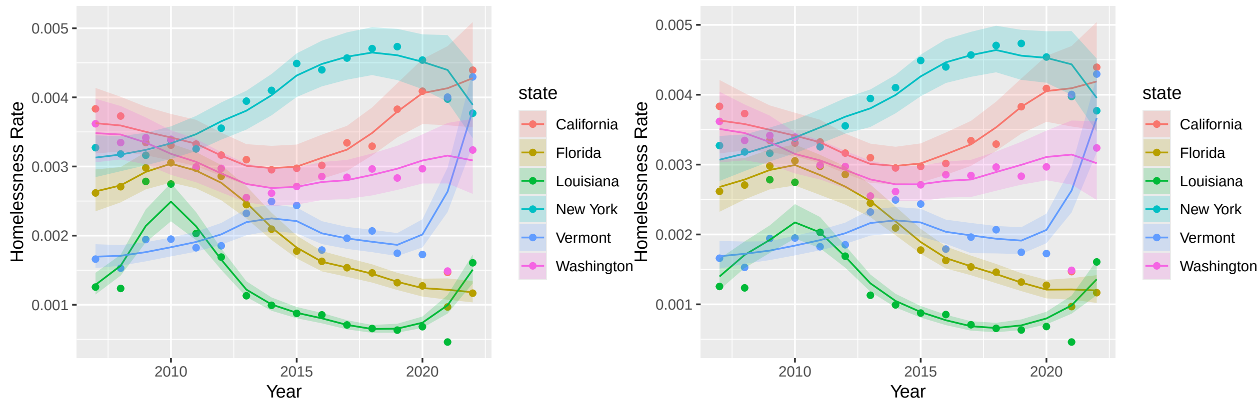
Table 1: Summary of quantitative diagnostics and prediction performance for all four models. ESS denotes the smallest effective sample size across the posterior means $\mu_{s,t}$ and the coefficients β_t and γ (where applicable), $\hat{R}$ denotes the largest Gelman-Rubin diagnostic across all parameters, and validation MSE is measured on the 2023 data.

Focusing on the random walk model with all covariates due to its lower prediction error and lower estimation uncertainty, we list the 95% credible intervals for significant coefficients in Table 2. The coefficient for the median housing cost covariate, γ, stands out as indicating a significant positive association with our response. This is expected from our EDA (recall Figure 4). Moreover, the coefficient for the poverty rate covariate is significant only for select years (2010, 2011, 2013, 2021, 2022) and the coefficient for the vacancy rate is significant in 2022 only. The coefficient for population density is never significant.

Coefficient	95% Credible Interval
$\beta_{2010,\text{poverty}}$	(0.0029, 0.0667)
$\beta_{2011,\text{poverty}}$	(0.0009, 0.0644)
$\beta_{2013,\text{vacancy}}$	(0.0001, 0.0938)
$\beta_{2021,\text{vacancy}}$	(0.0010, 0.1984)
$\beta_{2022,\text{vacancy}}$	(0.0664, 0.3049)
$\beta_{2022,\text{poverty}}$	(-0.1950,-0.0074)
γ	(0.1382, 0.4011)

Table 2: List of the 95% credible intervals that do not overlap with zero for parameters β and γ in the random walk model with all covariates.

Next, our visual inspection of the posterior distribution of $\exp(\mu_{s,t})$ on the same selection of states as in Figure 2 leads us to conclude that all models behave similarly on the data that they are fitted to (i.e, from 2007 to 2022). See Figure 5 for an illustration. Interestingly, this distribution is much more concentrated around its mean for the two southern states (Florida and Louisiana), though we suspect that this is due to their relatively low homelessness rate compared to the others. The spline model appears slightly less sensitive to outliers compared to the random walk model. In the case of Louisiana in 2009 and 2010, the random walk model overcompensates for a rate spike in these years while the spline model displays a smoother trend.

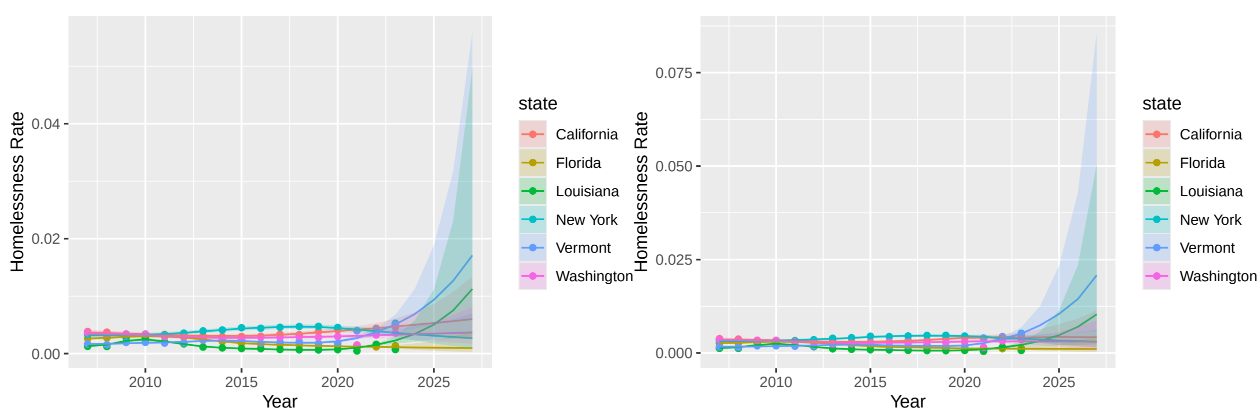


(a) Random walk model with covariates

(b) Spline model with covariates

Figure 5: Illustration of model fits for a selection of six states. The fits without covariates are similar. Points indicate the true data, lines indicate posterior mean predicted homelessness rate, and ribbons indicate 95% credible intervals for this posterior.

When it comes to forecasting five years in the future, due to the nature of the second-order random walk (be it on the state-time fluctuations directly or on the spline coefficients), the models simply project forward the most recent observed trend (i.e., between 2020 and 2022). This proves to be problematic for states such as Vermont and Louisiana that saw a spike in homelessness rate during this period. For these states, the models project sharp increases and large ranges of uncertainty (see Figure 6). These increases are even sharper for the models with covariates compared to those without. This is because of the random walk on the slopes β. Changes in these slopes are amplified by the second-order random walk when projecting forward. Thus, these coefficients — all of which are insignificant in most years — can factor substantially into the predictions (see Figure 7). On the other hand, in Figure 8, we remove Vermont and Louisiana, observing a more reasonable projection for the remaining states where the change from 2020 to 2022 was less dramatic. For the sake of brevity, we only show the plots for the random walk models, but the observations for the spline models are the same.



(a) Random walk model without covariates

(b) Random walk model with covariates

Figure 6: Illustration of model fits (2007-2022) and projections (2023-2027) on a selection of states for the two random walk models.

Since the time-dependent coefficients β are insignificant in the majority of years, this motivates us to fit a fifth model: a random walk model that keeps only the time-independent covariate (median monthly housing

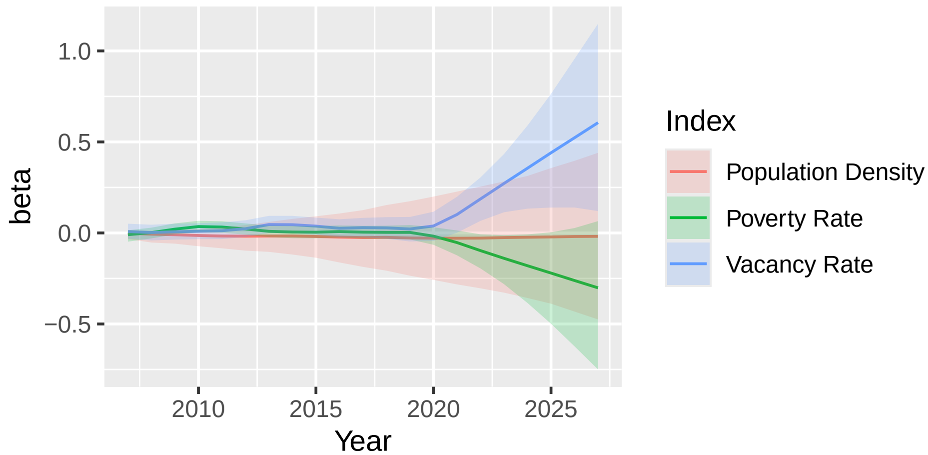
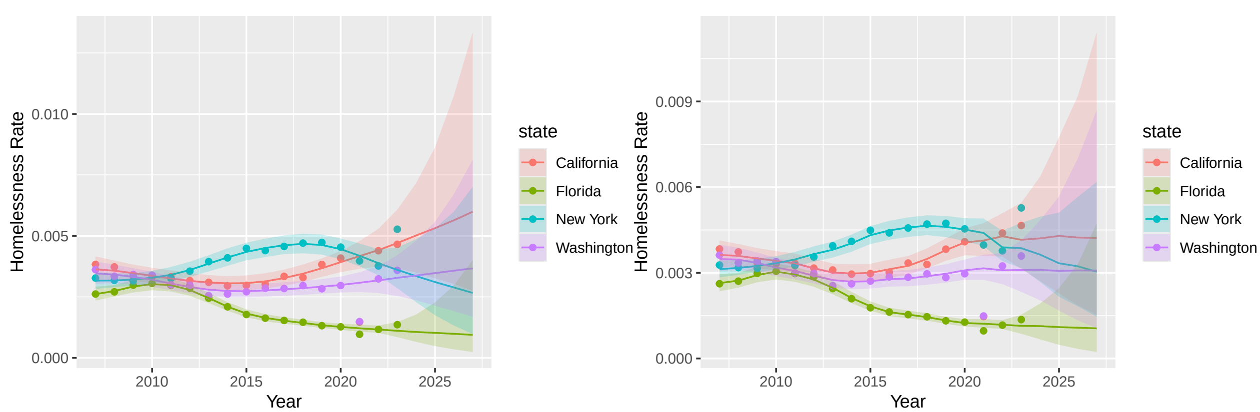


Figure 7: Evolution of the slopes for the time-dependent covariates in the random walk model with covariates.



(a) Random walk model without covariates

(b) Random walk model with covariates

Figure 8: Illustration of model fits (2007-2022) and projections (2023-2027) on a selection of states (excluding Vermont and Louisiana) for the two random walk models.

cost), the coefficient of which was found to be significant. This model has a minimum effective sample size of 545, a maximum $\widehat{R}$ of 1.028, and a validation MSE of 2.243×10^{-5}, which lies in between the MSE for the random walk model with covariates and the one without. The plot of its fit to the data from 2007 to 2022 and its projections from 2023 to 2027 appears indistinguishable from the plot for the random walk model without covariates, so we omit it.

5 Discussion

Of the five models that we fit in this project, the random walk model with all covariates appears to be best based on its validation MSE. However, its projections of the homelessness rate — particularly in the case of Vermont and Louisiana — grow unreasonably fast in part due to undesirable behaviour from its time-dependent coefficients. Because of this, we maintain that the random walk model that retains only the median monthly housing cost covariate is best. Indeed, this is the only covariate that we can definitively conclude has a significant association with homelessness rate. While this association is natural and expected, it is interesting to note that according to our model, the availability of housing (reflected by the vacancy rate) is not (consistently) significant once we account for cost. In other words, the bigger roadblock in addressing homelessness is not the lack of housing, but rather the unaffordability of the housing that is available. The lack of significance for the poverty rate covariate is unsurprising since the poverty line is constant across states and does not account for differences in cost of living. The credible intervals of the slope for population density overlap with zero in all years, reinforcing the notion that homelessness is not a primarily urban issue.

The projections generated by our chosen best model should be taken with a substantial grain of salt. The lack of significant time-dependent covariates leaves the projections as nothing more than the natural extrapolation of the most recent trend. The situation could perhaps be improved with a higher-order random walk, but ultimately, in order to make a more informed forecast, we require more covariates. IPUMS[3] hosts an extensive archive of detailed data from the decennial US census and the annual American Community Survey. Unfortunately, storing and manipulating the very large datasets from this resource may require more computational resources than are at our disposal. With that said, covariates that could be considered in future work include health insurance coverage rate (since for uninsured individuals, healthcare expenses must be accounted for in addition to other necessities like housing), high school graduation rate, unemployment rate, inflation rate, median disposable income, median income tax rate, the governing party in each of the executive and legislative branches (as this influences homelessness policy), and weather (including temperature, precipitation, and frequency of natural disasters such as wildfires, floods, and hurricanes). Moreover, recall that the HUD PiT counts are grouped by sheltered/unsheltered status, age, gender, and race. Hence, it is natural to extend our hierarchical model to account for these groupings. We could for instance have a common prior across the coefficients for each gender and a first-order random walk prior on age groups.

References

Monica Alexander, Christopher Wildeman, and Alexander Roehrkasse. Future projections of child welfare outcomes in the united states using a bayesian state-space model. In *PAA 2022 Annual Meeting*. PAA, 2022.

John R Belcher and Bruce R DeForge. Social stigma and homelessness: The limits of social change. *Journal of Human Behavior in the Social Environment*, 22(8):929–946, 2012.

Leslie Brinkley. Unhoused crisis in alameda co. leads superviors to declare state of emergency: Here's what next. *ABC 7 News*, 2023.

Thomas Byrne, Ellen A Munley, Jamison D Fargo, Ann E Montgomery, and Dennis P Culhane. New perspectives on community-level determinants of homelessness. *Journal of Urban Affairs*, 35(5):607–625, 2013.

[3]https://www.ipums.org

City of Los Angeles. Mayor bass signs updated state of emergency on housing and homelessness to confront ongoing crisis. 2023.

Chris Daniels. Seattle spent nearly $1 billion on homelessness, but number of unsheltered grew. *KOMO News*, 2023.

Tanya de Sousa, Alyssa Andrichik, Ed Prestera, Katherine Rush, Colette Tano, Micaiah Wheeler, and Abt Associates. The 2023 annual homelessness assessment report (ahar) to congress. *The U.S. Department of Housing and Urban Development Office of Community Planning and Development*, 2023.

Liam Elder-Connors. Homelessness in vermont has increased — and towns are responding differently to the problem. *Vermont Public*, 2023.

Chris Glynn and Emily B Fox. Dynamics of homelessness in urban america. *The Annals of Applied Statistics*, 13(1):573–605, 2019.

Mary Meehan. Unsheltered and uncounted: Rural america's hidden homeless. *National Public Radio*, 2019.

Office of the New York City Comptroller Brad Lander. Review of the new york city department of homeless services' programs and services. *Office of the New York City Comptroller Brad Lander*, 2023.

Robin Rothstein and Chris Jennings. Examining the cost of living by state in 2024. 2024.

Katherine H Shelton, Pamela J Taylor, Adrian Bonner, and Marianne Van Den Bree. Risk factors for homelessness: Evidence from a population-based study. *Psychiatric services*, 60(4):465–472, 2009.

United Nations General Assembly. Universal declaration of human rights. *UN General Assembly*, 302(2): 14–25, 1948.

United States Census Bureau. 2020 census urban areas faqs. 2020.

United States Census Bureau. Methodology for the united states population estimates: Vintage 2023. 2023.

United States Department of Housing and Urban Development. Point-in-time count methodology guide. *United States Department of Housing and Urban Development*, 2014.

Angela Wyse, Bruce Meyer, and Ilina Logani. Life and death at the margins of society: The mortality of the us homeless population. In *2023 APPAM Fall Research Conference*, 2023.

Marshall Zelinger. Denver mayor mike johnston declares state of emergency over homelessness. *KUSA-TV*, 2023.