Estimation and inference of causal effects for spatially clustered survival data:

A non-parametric Bayesian Approach

SUMMARY: We propose a novel Bayesian approach to estimate causal effects in spatially clustered survival data. Using SoftBayesian Additive Regression Trees (SBART), we introduce a flexible framework that accommodates spatial associations through a Conditional Auto-Regressive (CAR) model. Additionally, we incorporate a propensity score model to estimate treatment effects in observational studies by implementing a two-stage method; we estimate propensity scores and incorporate them into the outcome model, achieving double robustness analogous to frequentist causal inference. In our simulation study, we compared our method with established approaches in various scenarios, both with correctly specified and misspecified outcome models, demonstrating its robustness. We also applied the approach to analyze breast cancer patients in Florida using the Florida Cancer Registry (FCR), with Treatment Delay (TD) as treatment variables. By integrating county-specific spatial effects, our method captures localized impacts and provides geographic insights for targeted interventions, while also estimating overall treatment effects.

KEY WORDS: A key word; But another key word; Still another key word; Yet another key word.

1   Introduction

Survival analysis in spatially clustered data has become increasingly important in fields such as epidemiology, social sciences, and medical research. Traditional survival models, including the Cox proportional hazards model [7], have been widely applied to these settings. However, these methods often fail to address the complex dependencies and interactions that arise in spatially clustered data. Causal inference in survival analysis extends traditional survival modeling to estimate treatment effects and individual survival probabilities under hypothetical interventions.

Causal inference in survival analysis has evolved significantly over recent decades, with a particular focus on non-parametric and semi-parametric methods. The potential outcomes framework, introduced by Neyman [22] and later extended by Rubin [24], forms the basis of modern causal inference. In survival settings, specialized techniques such as inverse probability weighting (IPW) [23] and marginal structural models [11] have been developed to handle time-to-event outcomes and censoring. IPW methods have been further enhanced using non-parametric estimation techniques to reduce model dependence and improve robustness [18].

Bayesian causal inference has seen significant advancements in addressing challenges like selection bias and model misspecification. Under the ignorability and prior independence assumptions, Bayesian methods allow for causal effect estimation without direct dependence on the propensity score (PS). However, prior independence can inadvertently act as an informative prior, introducing *prior dogmatism*, where the selection bias sharply concentrate around zero as the covariate dimension increases. Linero (2023)[21] and Li (2023)[20] highlighted this phenomenon and demonstrated that incorporating estimated PS into the outcome model mitigates such biases. Rubin (1985) [25] first proposed using PS as the only covariate, but this approach suffered from sensitivity and poor empirical performance. Subsequent work, such as that by Zigler et al. (2013)[34], proposed using PS as an additional covariate in the

outcome model, leading to more robust results. The two-stage approach—estimating PS and then plugging it into the outcome model—balances computational feasibility and statistical rigor while addressing potential feedback issues. Bayesian non-parametric approaches, such as Gaussian Processes [33] and Dirichlet Process Mixture Models [16], have been used to model heterogeneous treatment effects without requiring stringent parametric assumptions. Similarly, Bayesian Additive Regression Trees (BART) [12, 5], a tree-based ensemble method, has shown promise in estimating causal effects in high-dimensional settings by capturing non-linear interactions and minimizing model misspecification.

Causal inference methods have been increasingly applied to survival data to estimate treatment effects from observational studies. Within the framework of inverse probability weighting (IPW), Xie and Liu [31] proposed an adjusted Kaplan-Meier estimator of the survival function and derived an adjusted log-rank test for survival comparisons, although they did not consider the sampling variability of the estimated propensity scores. Sugihara [27] extended this approach to the generalized propensity score for multi-valued treatment comparisons. Xu et al. [32] further advanced the methodology by incorporating stabilized IPW to handle time-varying treatments. Cole and Hernán [6] suggested using stabilized IPW in conjunction with fitting null Cox proportional hazards models for individual groups to obtain unconfounded survival curves. Austin and Schuster [2] compared various propensity score methods in estimating treatment effects on survival outcomes, such as differences in mean or median survival time or survival probability. Further studies by Austin [1], Austin and Stuart [4], Austin et al. [3], and Gayat et al. [9] focused on estimating hazard ratios using propensity scores, contributing significantly to methodological advancements in causal survival analysis.

Bayesian causal inference for survival data primarily focuses on Accelerated Failure Time (AFT) models. Henderson et al. [10] extended Bayesian Additive Regression Trees (BART)

to survival outcomes by developing a nonparametric AFT model that flexibly captures the relationship between covariates and failure times for Individualized Treatment Effect (ITE) estimation. Tian et al. [29] addressed the same by modeling interactions between treatment and a large number of covariates, focusing primarily on randomized clinical trials. Shen et al. [26] employed random forests with weighted bootstraps to estimate optimal personalized treatment strategies, while Cui et al. [8] extended causal forests [30] to accommodate right-censored data, employing bootstrap methods for inference. Additionally, Tabib and Larocque [28] proposed a novel random forest splitting rule for partitioning survival data. Hu (2021)[13] compared a variety of nonparametric and deep learning-based methods for causal inference in survival data, including AFT-BART models [10], random survival forests [14], and neural network-based approaches like DeepSurv [15], DR-DL, and DL-BJ [17, 19]. Furthermore, there has been no significant effort to incorporate spatial associations in survival data under a causal framework, leaving an important research avenue unexplored.

Datasets with spatial associations present unique challenges in causal survival analysis. In many applications—such as our motivating example with breast cancer patients—the data are naturally clustered by county, leading to intra-county correlations and spatial dependencies. Unobserved county-level factors (e.g., local healthcare infrastructure, socioeconomic conditions) may jointly influence treatment assignment (e.g., delays in initiating treatment or biopsies) and survival outcomes, making it essential to account for these associations.

Our research is motivated by analyzing breast cancer survival outcomes using data from the Florida Cancer Registry (FCR), which encompasses 76,106 breast cancer patients distributed across 67 counties in Florida. This dataset features significant spatial clustering by county, with patient survival times subject to right-censoring. The primary treatment of interest in our analysis is treatment delay, defined as waiting more than 90 days for treatment initiation compared to receiving timely care. Both patient-level confounders (including age,

race, tumor grade, hormone receptor status, and cancer stage) and county-level factors potentially affect treatment decisions and subsequent survival outcomes. The complexity induced by spatial clustering—driven by measured and unmeasured county-specific characteristics—presents significant methodological challenges that traditional causal inference methods are ill-equipped to handle adequately.

The primary goals of our analysis include estimating the overall average causal effect of treatment delays on patient survival and quantifying county-specific effects after adjusting for confounders at both patient and county levels. Capturing geographical heterogeneity in treatment effects is crucial for pinpointing specific regions where targeted interventions to reduce treatment delays could yield substantial improvements in patient outcomes. Our Bayesian non-parametric approach employing Bayesian Additive Regression Trees (BART) flexibly models these intricate relationships while explicitly accounting for spatial clustering and censoring. By identifying the counties where delays have the most pronounced negative impacts on survival, our estimands directly inform health policy, supporting targeted resource allocation and intervention strategies to mitigate regional disparities and enhance healthcare delivery efficiency.

Moreover, the complex interplay between individual-level confounders (e.g., age, HR status of a patient) and treatment requires a flexible approach that captures non-linearities and interactions. Traditional parametric or semiparametric models risk misspecification in these settings. Leveraging the non-parametric flexibility of Soft-BART (SBART), our method accommodates such interactions while avoiding restrictive assumptions.

Our approach uses a two-stage procedure where propensity scores are first estimated with SBART and then incorporated into a spatially-aware survival model. This strategy achieves double robustness—mitigating biases from either model—and explicitly accounts for county-level spatial dependencies. Consequently, our framework enables estimation of

county-specific treatment effects, providing nuanced insights for targeted interventions and localized healthcare strategies.

Overall, by integrating flexible non-parametric modeling with spatial random effects, our method advances the analysis of spatially clustered survival data, yielding precise and actionable estimates of treatment effects at both individual and county levels.

2 Methods

This section describes our methodological framework for estimating causal effects in spatially clustered survival data. Our approach combines Soft Bayesian Additive Regression Trees (SBART) with Conditional Auto-Regressive (CAR) spatial models to address the complex relationships between covariates and outcomes while accounting for spatial dependencies.

2.1 Notation and Definitions

Consider a study with K clusters (counties), where each cluster i contains n_i subjects, yielding a total sample size of $N = \sum_{i=1}^{K} n_i$. For each subject j in cluster i, we observe a binary treatment indicator $z_{ij} \in \{0, 1\}$, with $z_{ij} = 1$ indicating treatment (e.g., treatment delay) and $z_{ij} = 0$ indicating control (e.g., no delay); subject-level covariates $\mathbf{x}_{ij} \in \mathbb{R}^p$ (e.g., age, tumor grade, HR status); cluster-level covariates $\mathbf{v}_i \in \mathbb{R}^q$ (e.g., county-specific healthcare access metrics); failure time T_{ij}, which may be right-censored by censoring time C_{ij}; observed outcome $y_{ij} = \min(T_{ij}, C_{ij})$ and censoring indicator $\delta_{ij} = 1(T_{ij} \leqslant C_{ij})$. For causal inference, we define potential outcomes $T_{ij}(1)$ and $T_{ij}(0)$ as the survival times that would be observed under treatment and control, respectively.

2.2 Causal Estimands

We define several causal estimands to quantify treatment effects on survival outcomes (Hernán and Robins, 2020; Robins, 2000). The conditional survival function for treatment arm $z \in \{0, 1\}$ is:

$$S^{(z)}(t \mid \mathbf{x}, \mathbf{v}) = P(T(z) \geqslant t \mid \mathbf{x}, \mathbf{v}) \tag{1}$$

By averaging over the joint distribution of $\mathbf{x}$ and $\mathbf{v}$, we obtain the marginal survival function:

$$S^{(z)}(t) = \mathbb{E}_{\mathbf{x},\mathbf{v}}[S^{(z)}(t \mid \mathbf{x}, \mathbf{v})] \tag{2}$$

The survival probability treatment effect (SPTE) at time t measures the difference in marginal survival probabilities (Austin, 2016):

$$\Delta^{\mathrm{SPTE}}(t) = S^{(1)}(t) - S^{(0)}(t) \tag{3}$$

The average treatment effect (ATE) represents the average gain in survival time:

$$\Delta^{\mathrm{ATE}} = \mathbb{E}[T(1)] - \mathbb{E}[T(0)] \tag{4}$$

For practical applications where survival data may be heavily right-censored, we define the restricted average treatment effect (RATE) over a time horizon t^* (Royston and Parmar, 2013):

$$\Delta^{\mathrm{RATE}}(t^*) = \mathbb{E}[\min(T(1), t^*)] - \mathbb{E}[\min(T(0), t^*)] \tag{5}$$

The conditional versions of these estimands provide personalized treatment effects. The conditional average treatment effect (CATE) is:

$$\Delta^{\mathrm{CATE}}(\mathbf{x}, \mathbf{v}) = \mathbb{E}[T(1) - T(0) \mid \mathbf{x}, \mathbf{v}] \tag{6}$$

And the conditional restricted average treatment effect (CRATE) is:

$$\Delta^{\text{CRATE}}(t^*, \mathbf{x}, \mathbf{v}) = \mathbb{E}[\min(T(1), t^*) - \min(T(0), t^*) \mid \mathbf{x}, \mathbf{v}] \tag{7}$$

In the context of our Florida Cancer Registry data, these estimands provide critical insights for health policy. For instance, $\Delta^{\text{CATE}}(\mathbf{x}, \mathbf{v})$ represents the expected gain in survival time for patients with specific characteristics in a particular county when treatment delay is eliminated. The $\Delta^{\text{RATE}}(t^*)$ metric measures expected life-years gained within a clinically relevant time horizon (e.g., 5 or 10 years), which is particularly valuable for policy evaluations that require more immediate outcomes. By estimating these quantities at the county level, policymakers can identify geographic areas where interventions to reduce treatment delays would yield the greatest survival benefits.

Under the consistency assumption, for subjects with $z_{ij} = 1$ the observed survival time is identical to the potential outcome, $T_{ij} = T_{ij}(1)$. Combined with unconfoundedness, we have:

$$P(T_{ij}(1) > t \mid \mathbf{x}_{ij}, \mathbf{v}_i) = P(T_{ij} > t \mid \mathbf{x}_{ij}, \mathbf{v}_i, z_{ij} = 1) \tag{8}$$

Similarly, for $z_{ij} = 0$:

$$P(T_{ij}(0) > t \mid \mathbf{x}_{ij}, \mathbf{v}_i) = P(T_{ij} > t \mid \mathbf{x}_{ij}, \mathbf{v}_i, z_{ij} = 0) \tag{9}$$

This allows us to estimate conditional survival functions for each treatment arm using observed data. The marginal survival functions and treatment effects are then computed by averaging over the distribution of covariates.

2.3   Causal Assumptions

To ensure valid causal inference, we adopt the following assumptions:

(1) **SUTVA (Stable Unit Treatment Value Assumption):** For any two subjects j, j'

in clusters i, i',

$$T_{ij}(z_{ij}, z_{i'j'}) = T_{ij}(z_{ij}) \tag{10}$$

This means one subject's treatment does not affect another's outcome.

(2) **Consistency:**

$$T_{ij} = T_{ij}(1)1(z_{ij} = 1) + T_{ij}(0)1(z_{ij} = 0) \tag{11}$$

The observed outcome corresponds to the potential outcome under the treatment actually received.

(3) **Unconfoundedness:** For $z \in \{0, 1\}$,

$$T_{ij}(z) \perp\!\!\!\perp z_{ij} \mid \mathbf{x}_{ij}, \mathbf{v}_i \tag{12}$$

Treatment assignment is independent of potential outcomes conditional on observed covariates.

(4) **Positivity:**

$$0 < e(\mathbf{x}_{ij}, \mathbf{v}_i) = P(z_{ij} = 1 \mid \mathbf{x}_{ij}, \mathbf{v}_i) < 1 \tag{13}$$

Each subject has a non-zero probability of receiving either treatment status.

(5) **Covariate-dependent Censoring:** For $z \in \{0, 1\}$,

$$T_{ij}(z) \perp\!\!\!\perp C_{ij}(z) \mid \mathbf{x}_{ij}, \mathbf{v}_i, z_{ij} = z \tag{14}$$

Censoring is independent of survival time conditional on covariates and treatment.

2.4   Two-Stage Procedure

We implement a two-stage approach to estimate causal effects:

2.4.1   Stage 1: Propensity Score Estimation

The propensity score $e_{ij} = P(z_{ij} = 1 \mid \mathbf{x}_{ij}, \mathbf{v}_i)$ is modeled using SBART:

$$e_{ij} = \Phi(b_1(\mathbf{x}_{ij}, \mathbf{v}_i)) \tag{15}$$

where $b_1(\cdot)$ is modeled via SBART and $\Phi(\cdot)$ is the standard normal cumulative distribution function.

We adopt a two-stage procedure rather than simultaneous estimation for several compelling reasons. First, this approach mitigates the potential for feedback between the treatment and outcome models, which could otherwise amplify model misspecification biases (Zigler et al., 2013). Second, it offers computational advantages by allowing separate estimation of propensity scores and outcome models, particularly valuable for large datasets with complex spatial structures. Finally, research by Zigler et al. (2013) has shown that two-stage methods often perform better empirically than simultaneous approaches in causal inference settings.

Including the propensity score as a covariate in the outcome model addresses confounding bias that arises from differences between treatment groups. In observational studies, the treatment assignment mechanism can lead to systematic differences in covariate distributions between the treated and control groups (Rosenbaum and Rubin, 1983). This can result in biased treatment effect estimates if not properly addressed. By conditioning on the propensity score, we create a pseudo-randomized scenario where the distribution of measured confounders becomes balanced across treatment groups (Lunceford and Davidian, 2004).

Furthermore, this approach confers double robustness, a property first established by Robins and Rotnitzky (2001) and further developed by Bang and Robins (2005). Double robustness means that causal estimates remain consistent if either the propensity score model or the outcome model (but not necessarily both) is correctly specified. This property provides substantial protection against model misspecification in either component, which is particularly valuable in complex settings like ours where the true data-generating process is unknown.

2.4.2   Stage 2: Outcome Model

For the survival outcome, we use an Accelerated Failure Time (AFT) model with spatial random effects:

$$\log T_{ij} = b_2(z_{ij}, \mathbf{x}_{ij}, \mathbf{v}_i, \hat{e}_{ij}) + W_i + \epsilon_{ij}, \quad \epsilon_{ij} \sim N(0, \sigma^2) \tag{16}$$

where $b_2(\cdot)$ is a flexible nonparametric function modeled via SBART, W_i is a county-specific spatial random effect, $\hat{e}_{ij}$ is the estimated propensity score from Stage 1, and ϵ_{ij} is a normally distributed error term with variance σ^2.

2.5   Bayesian Additive Regression Trees (BART) and SBART

BART represents an unknown regression function as a sum of regression trees (Chipman et al., 2010):

$$b(\mathbf{x}) = \sum_{h=1}^{H} g(\mathbf{x}; \mathcal{T}_h, \mathcal{M}_h) \tag{17}$$

where $\mathcal{T}_h$ denotes the tree structure (topology and splitting rules) of the h-th tree, and $\mathcal{M}_h = \{\mu_{h1}, \ldots, \mu_{hJ_h}\}$ represents the terminal node parameters for the J_h terminal nodes of tree h.

Each tree function can be expressed as:

$$g(\mathbf{x}; \mathcal{T}_h, \mathcal{M}_h) = \sum_{j=1}^{J_h} \mu_{hj} \cdot 1(\mathbf{x} \in R_{hj}) \tag{18}$$

where R_{hj} is the region of the covariate space corresponding to the j-th terminal node of tree h.

SBART extends standard BART by replacing the hard decision tree splits with soft decision rules (Linero, 2018):

$$g(\mathbf{x}; \mathcal{T}_h, \mathcal{M}_h) = \sum_{j=1}^{J_h} \mu_{hj} \cdot w_j(\mathbf{x}) \tag{19}$$

where $w_j(\mathbf{x})$ is a weight function based on sigmoid transformations of the splitting rules:

$$w_j(\mathbf{x}) = \prod_{d \in A_j} \sigma((x_d - c_d) \cdot s_d/\alpha) \tag{20}$$

Here, A_j represents the set of decision rules leading to terminal node j, c_d is the split point, $s_d \in \{-1, 1\}$ indicates the direction of the split, and α controls the softness of the split. As $\alpha \to 0$, SBART converges to standard BART.

BART and SBART have emerged as particularly effective tools for causal inference due to several key properties. Their nonparametric flexibility allows them to capture complex, nonlinear relationships between covariates and outcomes without requiring pre-specification of functional forms (Hill, 2011; Hahn et al., 2020). This is crucial in causal settings where model misspecification can lead to biased treatment effect estimates.

The inherent regularization in BART's prior specification prevents overfitting while still capturing important signal in the data (Dorie et al., 2019). This balance is particularly valuable in causal inference, where the goal is to accurately estimate treatment effects rather than merely predict outcomes. Simulation studies and empirical applications have consistently shown that BART outperforms alternative methods in causal inference tasks (Hill, 2011; Dorie et al., 2019; Hahn et al., 2020).

The posterior uncertainty quantification provided by BART also allows for credible intervals around treatment effect estimates, which is essential for policy decision-making (Hill, 2011). Furthermore, the adaptability of BART to high-dimensional settings makes it well-suited for modern observational studies with many potential confounders (Carnegie et al., 2019).

SBART further enhances these capabilities by improving performance on smooth functions while maintaining BART's excellent performance for discontinuous functions (Linero, 2018). This adaptivity to the underlying data structure makes SBART particularly well-suited for complex causal inference problems where the true functional forms are unknown (Starling et al., 2020; Linero et al., 2022).

2.6 Conditional Auto-Regressive (CAR) Model

To account for spatial dependencies among counties, we model the spatial random effects $\mathbf{W} = (W_1, \ldots, W_K)$ using a CAR prior (Besag, 1974; Banerjee et al., 2003):

$$p(\mathbf{W} \mid \tau^2, \rho) \propto \exp\left(-\frac{1}{2\tau^2}\mathbf{W}^T(\mathbf{D} - \rho\mathbf{A})\mathbf{W}\right) \tag{21}$$

where $\mathbf{A}$ is the $K \times K$ adjacency matrix with elements $a_{ik} = 1$ if counties i and k are neighbors, and 0 otherwise; $\mathbf{D}$ is a diagonal matrix with elements $d_{ii} = \sum_k a_{ik}$ representing the number of neighbors of county i; $\rho \in (1/\lambda_{min}, 1/\lambda_{max})$ is the spatial correlation parameter, where λ_{min} and λ_{max} are the minimum and maximum eigenvalues of $\mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2}$ (Banerjee et al., 2014); and τ^2 is the variance parameter controlling the overall variability of the spatial effects.

The CAR model induces spatial correlation such that neighboring counties tend to have similar random effects. The conditional distribution of each W_i given all other random effects is:

$$W_i \mid \mathbf{W}_{-i}, \tau^2, \rho \sim N\left(\rho \sum_{k=1}^{K} \frac{a_{ik}}{d_{ii}}W_k, \frac{\tau^2}{d_{ii}}\right) \tag{22}$$

where $\mathbf{W}_{-i}$ denotes all elements of $\mathbf{W}$ except W_i. This formulation allows for local spatial smoothing while preserving the global structure of spatial associations (Cressie, 2015).

2.7   Observed Data Likelihood

For each subject (i,j), the contribution to the likelihood depends on whether the failure time is observed or censored:

$$L_{ij}^{obs}(\theta) = \begin{cases} f_{T_{ij}}(y_{ij} \mid z_{ij}, \mathbf{x}_{ij}, \mathbf{v}_i, \hat{e}_{ij}, W_i; \theta), & \text{if } \delta_{ij} = 1 \\[2mm] S_{T_{ij}}(y_{ij} \mid z_{ij}, \mathbf{x}_{ij}, \mathbf{v}_i, \hat{e}_{ij}, W_i; \theta), & \text{if } \delta_{ij} = 0 \end{cases} \tag{23}$$

where $f_{T_{ij}}$ and $S_{T_{ij}}$ are the density and survival functions of T_{ij}, respectively, and θ represents all model parameters.

Under our log-normal AFT model, these functions are:

$$f_{T_{ij}}(t \mid z_{ij}, \mathbf{x}_{ij}, \mathbf{v}_i, \hat{e}_{ij}, W_i; \theta) = \frac{1}{t\sigma\sqrt{2\pi}} \exp\left(-\frac{(\log t - \mu_{ij})^2}{2\sigma^2}\right) \tag{24}$$

$$S_{T_{ij}}(t \mid z_{ij}, \mathbf{x}_{ij}, \mathbf{v}_i, \hat{e}_{ij}, W_i; \theta) = 1 - \Phi\left(\frac{\log t - \mu_{ij}}{\sigma}\right) \tag{25}$$

where $\mu_{ij} = b_2(z_{ij}, \mathbf{x}_{ij}, \mathbf{v}_i, \hat{e}_{ij}) + W_i$ and $\Phi(\cdot)$ is the standard normal CDF.

The complete observed data likelihood is:

$$L^{obs}(\theta) = \prod_{i=1}^{K} \prod_{j=1}^{n_i} L_{ij}^{obs}(\theta) \tag{26}$$

2.8   Data Augmentation and Complete Data Likelihood

To facilitate MCMC sampling, we introduce latent log survival times $\tilde{y}_{ij}$ defined as (Tanner and Wong, 1987):

$$\tilde{y}_{ij} = \begin{cases} \log y_{ij}, & \text{if } \delta_{ij} = 1 \\ \text{TruncNormal}(\mu_{ij}, \sigma^2; \log y_{ij}, \infty), & \text{if } \delta_{ij} = 0 \end{cases} \tag{27}$$

where $\text{TruncNormal}(\mu, \sigma^2; a, b)$ denotes a normal distribution with mean μ and variance σ^2 truncated to the interval (a, b).

With this augmentation, the complete data likelihood becomes:

$$L^{comp}(\theta) = \prod_{i=1}^{K} \prod_{j=1}^{n_i} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\tilde{y}_{ij} - \mu_{ij})^2}{2\sigma^2}\right) \tag{28}$$

This formulation transforms the problem into a standard Bayesian regression framework with normally distributed errors, facilitating MCMC sampling (Wei et al., 1992).

2.9 Prior Specifications

We specify the following priors for our model parameters: For the SBART components $b_1(\cdot)$ and $b_2(\cdot)$, we use regularizing priors on tree structures $\mathcal{T}_h$ and terminal node parameters $\mathcal{M}_h$ following Chipman et al. (2010) and Linero (2018). For the error variance: $\sigma^2 \sim \text{InvGamma}(a_\sigma, b_\sigma)$. For the spatial variance: $\tau^2 \sim \text{InvGamma}(a_\tau, b_\tau)$. For the spatial correlation parameter: $\rho \sim \text{Uniform}(1/\lambda_{min}, 1/\lambda_{max})$. These priors are chosen to be weakly informative, allowing the data to dominate the posterior inference while providing regularization to prevent overfitting (Gelman, 2006).

2.10 MCMC Algorithm

We estimate all parameters using Markov Chain Monte Carlo (Gelfand and Smith, 1990). The algorithm proceeds as follows:

Algorithm: MCMC for Spatial SBART Survival Model

(1) Initialize all parameters: $b_2^{(0)}(\cdot)$, $\mathbf{W}^{(0)}$, $\sigma^{2(0)}$, $\tau^{2(0)}$, $\rho^{(0)}$

(2) For iterations $m = 1, 2, ..., M$:

a. **Impute censored data**: For each subject with $\delta_{ij} = 0$, draw

$$\tilde{y}_{ij}^{(m)} \sim \text{TruncNormal}(\mu_{ij}^{(m-1)}, \sigma^{2(m-1)}; \log y_{ij}, \infty) \tag{29}$$

b. **Update spatial parameters**:

Update $\mathbf{W}^{(m)}$ from its full conditional:

$$p(\mathbf{W} \mid \cdot) \propto \exp\left(-\frac{1}{2}\left[\sum_{i=1}^{K}\sum_{j=1}^{n_i} \frac{(\hat{y}_{ij} - b_2(z_{ij}, \mathbf{x}_{ij}, \mathbf{v}_i, \hat{e}_{ij}) - W_i)^2}{\sigma^2} \right.\right.$$
$$\left.\left. + \frac{\mathbf{W}^T(\mathbf{D} - \rho\mathbf{A})\mathbf{W}}{\tau^2}\right]\right) \tag{30}$$

Update $\tau^{2(m)}$ from its full conditional:

$$\tau^{2(m)} \sim \text{InvGamma}\left(a_\tau + \frac{K}{2}, b_\tau + \frac{1}{2}\mathbf{W}^{(m)T}(\mathbf{D} - \rho^{(m-1)}\mathbf{A})\mathbf{W}^{(m)}\right) \tag{31}$$

Update $\rho^{(m)}$ using Metropolis-Hastings (Chib and Greenberg, 1995):

Propose $\rho^* \sim \text{TruncNormal}(\rho^{(m-1)}, \sigma_\rho^2; 1/\lambda_{min}, 1/\lambda_{max})$

Accept with probability:

$$\min\left(1, \frac{|\mathbf{D} - \rho^*\mathbf{A}|^{-1/2} \exp\left(-\frac{1}{2\tau^{2(m)}}\mathbf{W}^{(m)T}(\mathbf{D} - \rho^*\mathbf{A})\mathbf{W}^{(m)}\right)}{|\mathbf{D} - \rho^{(m-1)}\mathbf{A}|^{-1/2} \exp\left(-\frac{1}{2\tau^{2(m)}}\mathbf{W}^{(m)T}(\mathbf{D} - \rho^{(m-1)}\mathbf{A})\mathbf{W}^{(m)}\right)}\right) \tag{32}$$

c. **Update SBART parameters**:

For $h = 1, 2, ..., H$, compute partial residuals:

$$R_{ij}^{(h)} = \tilde{y}_{ij}^{(m)} - W_i^{(m)} - \sum_{l \neq h} g(z_{ij}, \mathbf{x}_{ij}, \mathbf{v}_i, \hat{e}_{ij}; \mathcal{T}_l^{(m-1)}, \mathcal{M}_l^{(m-1)}) \tag{33}$$

Update $(\mathcal{T}_h^{(m)}, \mathcal{M}_h^{(m)})$ using Bayesian backfitting (Chipman et al., 2010)

d. **Update error variance**:

$$\sigma^{2(m)} \sim \text{InvGamma}\left(a_\sigma + \frac{N}{2}, b_\sigma + \frac{1}{2}\sum_{i=1}^{K}\sum_{j=1}^{n_i}(\tilde{y}_{ij}^{(m)} - \right.$$
$$\left. b_2^{(m)}(z_{ij}, \mathbf{x}_{ij}, \mathbf{v}_i, \hat{e}_{ij}) - W_i^{(m)})^2\right) \tag{34}$$

Our proposed method directly addresses the key challenges highlighted in the introduction.

By incorporating CAR priors for the county-level random effects, we explicitly model the spa-

tial correlation structure, allowing information sharing across neighboring counties (Banerjee et al., 2014). This improves estimation in counties with small sample sizes and captures unmeasured spatial confounding (Gelfand et al., 2010). SBART flexibly captures nonlinear relationships and interactions between covariates without requiring pre-specification of functional forms (Linero, 2018; Starling et al., 2020). This is particularly important for modeling complex interactions between treatment, patient characteristics, and county-level factors.

Our data augmentation approach properly accounts for right-censored observations, ensuring unbiased estimation of survival outcomes (Wei et al., 1992). The inclusion of estimated propensity scores in the outcome model helps adjust for confounding in a doubly robust manner, reducing bias in treatment effect estimates (Bang and Robins, 2005; Kang and Schafer, 2007). Our framework naturally provides county-specific treatment effect estimates, enabling targeted policy interventions at the local level.

By integrating these components, our method provides a comprehensive approach to causal inference in spatially clustered survival data. For the Florida Cancer Registry analysis, this methodology yields robust estimates of treatment delay effects on breast cancer survival, accounting for patient characteristics, county-level factors, and spatial dependencies. The resulting county-specific effect estimates directly inform health policy by identifying geographic areas where interventions to reduce treatment delays would yield the greatest survival benefits.

3 Simulation Study

To evaluate the performance of our proposed Bayesian nonparametric approach for spatially clustered survival data, we conducted a comprehensive simulation study. We designed several scenarios to assess the method's effectiveness in estimating causal effects under varying conditions, with a focus on the impact of spatial clustering and complex hazard structures.

3.1   Simulation Setup

Our simulation consists of $K = 15$ clusters (counties) with a total sample size of $N = 3000$ subjects distributed with varying cluster sizes, reflecting the heterogeneity in county populations observed in real-world data. For each subject j in cluster i, we generated two continuous confounders $X_1, X_2 \sim \text{Uniform}(0, 1)$, propensity score $e_{ij} = \text{expit}(0.3X_1 - 0.2X_2 + 0.5X_1X_2)$, treatment assignment $Z_{ij} \sim \text{Bernoulli}(e_{ij})$, and censoring time $C_{ij} \sim \text{Exponential}(0.05)$.

To incorporate spatial associations, we used the adjacency matrix $\mathbf{A}$ based on a subset of Florida counties. The spatial random effects $\mathbf{W} = (W_1, \ldots, W_K)$ were generated from a Conditional Auto-Regressive (CAR) model:

$$\mathbf{W} \sim \text{CAR}(\mathbf{A}; \sigma_W^2 = 1, \rho = 0.3) \tag{35}$$

where σ_W^2 controls the overall variability and ρ represents the spatial correlation.

We simulated survival times under two different model specifications to evaluate the robustness of our approach under varying conditions.

3.2   Outcome Models

Model 1 (Correctly Specified / M1): A baseline model with additive spatial effects

$$\log T_{ij} = b(X_{ij}, Z_{ij}) + W_i + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim N(0, \sigma^2) \tag{36}$$

Model 2 (Misspecified / M2): A more complex model with treatment-covariate interactions in the spatial effect

$$\log T_{ij} = b(X_{ij}, Z_{ij}) + W_i \cdot (1 + 0.5 \, Z_{ij} X_{1ij}) + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim N(0, \sigma^2) \tag{37}$$

For both models, the true nonparametric function was specified as:

$$b(X_{ij}, Z_{ij}) = \sin(\pi X_{1ij}) + \ln(1 + X_{2ij}^2) + 2Z_{ij} \, (X_{1ij} X_{2ij}) + (X_{1ij}^2) Z_{ij} \tag{38}$$

This function includes nonlinear main effects and complex treatment-covariate interactions,

representing the type of complexity often encountered in real-world applications. For each simulation condition, we generated 30 replicate datasets. The observed outcome for each subject was $y_{ij} = \min(T_{ij}, C_{ij})$ with censoring indicator $\delta_{ij} = 1(T_{ij} \leqslant C_{ij})$. This resulted in approximately 30% censoring across datasets.

3.3 Competing Methods

We compared our proposed spatial SBART method with a standard Cox proportional hazards model with frailty (Ph Frailty), which represents a common approach in survival analysis with clustered data. Both methods incorporated propensity scores to adjust for confounding, but differed in how they modeled the survival outcomes. Our Method (Spatial SBART) uses SBART to model the relationship between covariates, treatment, and survival time, with spatial random effects following a CAR prior. The Cox Frailty method employs a Cox proportional hazards model with cluster-specific random effects (frailties) but without spatial structure:

$$h_{ij}(t) = h_0(t) \exp(X'_{ij}\beta + \gamma Z_{ij} + W_i) \tag{39}$$

where $h_0(t)$ is an unspecified baseline hazard, β and γ are regression coefficients, and W_i are independent frailties.

3.4 Evaluation Metrics

For each method, we estimated several causal estimands: the Average Treatment Effect (ATE): $\Delta^{\mathrm{ATE}} = \mathbb{E}[T(1)] - \mathbb{E}[T(0)]$; the Restricted Average Treatment Effect (RATE): $\Delta^{\mathrm{RATE}}(t^*) = \mathbb{E}[\min(T(1), t^*)] - \mathbb{E}[\min(T(0), t^*)]$; and the Survival Probability Treatment Effect (SPTE): $\Delta^{\mathrm{SPTE}}(t) = S^{(1)}(t) - S^{(0)}(t)$.

Both RATE and SPTE were evaluated at multiple time points on a grid from 1 to 15 with step size 1 (i.e., at times $t \in \{1, 2, \ldots, 15\}$), and the performance metrics were calculated at each time point and then averaged across all time points to provide summary measures.

To evaluate the accuracy of our estimates, we calculated several performance metrics across all simulation replicates. For each estimand, we computed:

1) Coverage: For each replicate r, we determined whether the true value falls within the 95% credible interval:

$$\text{Coverage} = \frac{1}{R} \sum_{r=1}^{R} 1(\theta_{\text{true}} \in [\hat{\theta}_{\text{lower}}^{(r)}, \hat{\theta}_{\text{upper}}^{(r)}]) \tag{40}$$

where R is the number of replicates, θ_{true} is the true parameter value, and $[\hat{\theta}_{\text{lower}}^{(r)}, \hat{\theta}_{\text{upper}}^{(r)}]$ is the 95% credible interval for replicate r.

2) Absolute Bias: The average absolute difference between estimated and true values:

$$\text{Absolute Bias} = \frac{1}{R} \sum_{r=1}^{R} |\hat{\theta}^{(r)} - \theta_{\text{true}}| \tag{41}$$

where $\hat{\theta}^{(r)}$ is the estimate from replicate r.

3) Root Mean Squared Error (RMSE): The square root of the average squared difference between estimated and true values:

$$\text{RMSE} = \sqrt{\frac{1}{R} \sum_{r=1}^{R} (\hat{\theta}^{(r)} - \theta_{\text{true}})^2} \tag{42}$$

For time-dependent estimands like RATE and SPTE, these calculations were performed at each time point in our evaluation grid, and then averaged across all time points to provide summary metrics. For example, for the SPTE at time t, we computed:

$$\text{RMSE}_{\text{SPTE}}(t) = \sqrt{\frac{1}{R} \sum_{r=1}^{R} (\hat{\Delta}_{\text{SPTE}}^{(r)}(t) - \Delta_{\text{SPTE}}(t))^2} \tag{43}$$

where $\hat{\Delta}_{\text{SPTE}}^{(r)}(t)$ is the estimated SPTE at time t for replicate r, and $\Delta_{\text{SPTE}}(t)$ is the true SPTE at time t. The average RMSE across all time points was then calculated as:

$$\overline{\text{RMSE}}_{\text{SPTE}} = \frac{1}{|T|} \sum_{t \in T} \text{RMSE}_{\text{SPTE}}(t) \tag{44}$$

where $T = \{1, 2, \ldots, 15\}$ is the set of evaluation time points. Similar calculations were performed for the RATE estimand.

3.5   Simulation Results

Table 1 presents the performance metrics for both methods across the two model specifications.

[Table 1 about here.]

Our Spatial SBART method demonstrated substantially better performance than the Cox Frailty approach across all metrics and estimands. Several key findings emerge from the results. Our method maintained acceptable coverage rates, particularly for RATE (86.7%) and SPTE (80.0%) under Model 1. The coverage for ATE was lower (70.0%) but still substantially higher than Cox Frailty, which had 0% coverage for ATE and SPTE in both models. This indicates that Cox Frailty's confidence intervals systematically excluded the true values. Our method exhibited markedly lower absolute bias and RMSE across all estimands. For example, in Model 1, the absolute bias for SPTE was 0.04 for our method compared to 0.14 for Cox Frailty, representing a 71.4% reduction in bias.

Under Model 2, where the true data-generating process included treatment-covariate interactions in the spatial effect, our method's performance decreased but remained superior to Cox Frailty. The coverage for SPTE dropped from 80.0% to 36.7%, reflecting the increased complexity. However, our method's absolute bias (0.06) remained significantly lower than Cox Frailty's (0.16). Both methods showed better performance for SPTE and RATE compared to ATE, likely because these estimands focus on more restricted aspects of the survival distribution and are thus easier to estimate accurately.

The superior performance of our Spatial SBART method can be attributed to several factors. First, the SBART component allows for flexible modeling of complex nonlinear relationships and interactions between covariates and treatment. Second, the explicit modeling

of spatial dependencies through the CAR prior enables appropriate borrowing of information across neighboring clusters. Finally, the inclusion of propensity scores in the outcome model provides additional protection against confounding bias.

3.6 Simulation Discussion

The simulation results demonstrate the benefits of our proposed approach for causal inference in spatially clustered survival data. The Spatial SBART method consistently outperforms the traditional Cox Frailty approach across multiple performance metrics and estimands, particularly when spatial dependencies are present and the true relationship between covariates, treatment, and survival is complex.

The superior performance under Model 2, where the data-generating process included treatment-covariate interactions in the spatial effect, highlights our method's robustness to model misspecification. This is a crucial advantage in real-world applications where the true functional form is unknown and likely complex.

The county-specific estimates further illustrate our method's ability to capture local variations in treatment effects while properly accounting for uncertainty. This spatial heterogeneity in treatment effects is particularly important for targeted policy interventions, allowing resources to be directed to areas where they will have the greatest impact.

However, our results also reveal limitations. Coverage rates for ATE were consistently lower than the nominal 95% level, suggesting potential challenges in accurately capturing uncertainty for this estimand. Additionally, performance degraded somewhat under increased censoring and propensity score misspecification, indicating areas for future methodological improvement.

Overall, the simulation results provide strong support for the proposed Spatial SBART approach in analyzing spatially clustered survival data with complex causal structures. The method's ability to flexibly model nonlinear relationships, account for spatial dependencies,

and provide accurate county-specific treatment effect estimates makes it well-suited for applications like the Florida Cancer Registry analysis.

4   Application to Florida Cancer Registry Data

Our study analyzes breast cancer data from the Florida Cancer Registry (FCR), examining the causal effects of treatment delay on survival outcomes for 1,994 Non-African American patients with Distant stage breast cancer across 61 Florida counties.

4.1   Results

Table 1 presents population-level treatment effects calculated as:

$$\text{ATE} = \frac{1}{N} \sum_{i=1}^{K} \sum_{j=1}^{n_i} \Delta^{\text{CATE}}(\mathbf{x}_{ij}, v_i) \tag{45}$$

$$= \frac{1}{N} \sum_{i=1}^{K} \sum_{j=1}^{n_i} [E[T_{ij}(1)|\mathbf{x}_{ij}, v_i] - E[T_{ij}(0)|\mathbf{x}_{ij}, v_i]] \tag{46}$$

$$\text{RATE}(t^*) = \frac{1}{N} \sum_{i=1}^{K} \sum_{j=1}^{n_i} \Delta^{\text{CRATE}}(t^*, \mathbf{x}_{ij}, v_i) \tag{47}$$

$$= \frac{1}{N} \sum_{i=1}^{K} \sum_{j=1}^{n_i} [E[\min(T_{ij}(1), t^*)|\mathbf{x}_{ij}, v_i] - \tag{48}$$

$$E[\min(T_{ij}(0), t^*)|\mathbf{x}_{ij}, v_i]] \tag{49}$$

where $N = \sum_{i=1}^{K} n_i$ is the total sample size. The ATE of 9.8 months quantifies the expected survival benefit if treatment delays were eliminated for all patients in the population. The RATE values of 4.2 months (5-year) and 6.5 months (10-year) measure the expected gained life-years within these restricted time horizons.

[Table 2 about here.]

Table 2 summarizes the range of county-specific treatment effects, calculated as:

$$\text{CATE}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \Delta^{\text{CATE}}(\mathbf{x}_{ij}, v_i) \tag{50}$$

$$= \frac{1}{n_i} \sum_{j=1}^{n_i} [E[T_{ij}(1)|\mathbf{x}_{ij}, v_i] - E[T_{ij}(0)|\mathbf{x}_{ij}, v_i]] \tag{51}$$

$$\text{CRATE}_i(t^*) = \frac{1}{n_i} \sum_{j=1}^{n_i} \Delta^{\text{CRATE}}(t^*, \mathbf{x}_{ij}, v_i) \tag{52}$$

$$= \frac{1}{n_i} \sum_{j=1}^{n_i} [E[\min(T_{ij}(1), t^*)|\mathbf{x}_{ij}, v_i] - \tag{53}$$

$$E[\min(T_{ij}(0), t^*)|\mathbf{x}_{ij}, v_i]] \tag{54}$$

The range of CATE values (2.6 to 18.4 months) and CRATE-5yr values (1.3 to 9.1 months) demonstrates substantial geographic heterogeneity in treatment effects.

[Table 3 about here.]

Table 3 presents the ten counties with the highest conditional average treatment effects. Counties like Brevard, Martin, and Union show particularly large treatment effects, with Brevard County showing an 18.4-month improvement in survival when treatment delays are eliminated. Note that even counties with large sample sizes (e.g., Sarasota with n=154) show substantial treatment effects (11.8 months), indicating that the benefits of timely treatment are not limited to small counties with potentially unstable estimates.

[Table 4 about here.]

Figure 1 plots CATE_i against county sample size n_i, with vertical bars representing 95% credible intervals:

$$95\% \text{ CI for } \text{CATE}_i = [\text{CATE}_i^{2.5\%}, \text{CATE}_i^{97.5\%}] \tag{55}$$

where CATE_i^q is the q-th quantile of the posterior distribution. As shown in Figure 1, the absence of correlation between effect size and sample size suggests that county population is not a primary driver of treatment effect magnitude. Both urban (blue) and rural (red) counties span the range of effect sizes, indicating that urbanicity alone does not determine the impact of treatment delays.

[Figure 1 about here.]

Figure 2 visualizes the spatial distribution of CATE_i across Florida counties. The choropleth map reveals regional clusters of high-impact counties, providing evidence of spatial correlation in treatment effects. This pattern suggests that factors transcending county boundaries—such as regional healthcare infrastructure, shared referral networks, or socioeconomic conditions—influence the impact of treatment delays.

[Figure 2 about here.]

Table 4 shows treatment effects for patients with specific covariate profiles in Broward County, calculated as:

$$\Delta^{\text{CATE}}(\mathbf{x}, v_i) = E[T(1)|\mathbf{x}, v_i] - E[T(0)|\mathbf{x}, v_i] \tag{56}$$

These values quantify the expected survival benefit from eliminating treatment delays for patients with specific characteristics. The most striking pattern in Table 4 is the pronounced difference between HR-negative and HR-positive patients. For Grade 1 tumors, HR-negative patients show a 21.4-month benefit from avoiding treatment delays, while HR-positive patients show only an 11.1-month benefit. This pattern is consistent across all tumor grades, with HR-negative patients consistently showing approximately twice the benefit of HR-positive patients.
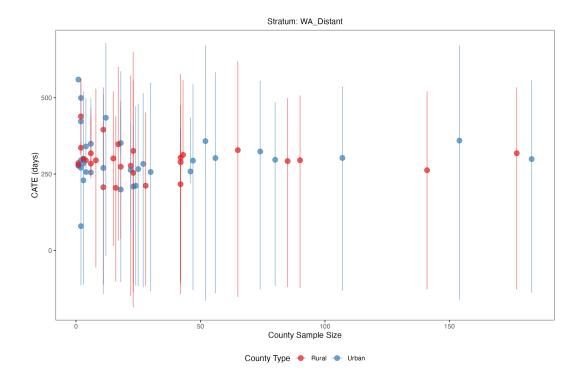
[Table 5 about here.]

Additionally, tumor grade modifies the treatment effect, with Grade 1 HR-negative tumors showing the largest benefit (21.4 months), followed by Grade 2 (19.0 months) and Grade 3 (18.3 months). This pattern is consistent with biological understanding of tumor behavior, where less differentiated (higher grade) tumors tend to be more aggressive and potentially less responsive to timing of interventions. Interestingly, we observe minimal age-related variations in treatment effects within this distant-stage stratum, suggesting that for metastatic breast cancer, tumor biology (HR status and grade) may be more important than patient age in determining the impact of treatment delays.

Clinical interpretations of these findings highlight several important patterns. First, the near doubling of treatment effect for HR-negative versus HR-positive tumors (21.4 vs. 11.1 months for Grade 1) aligns with the known biological aggressiveness of HR-negative tumors, which typically grow faster and are less responsive to hormonal interventions. This suggests that patients with HR-negative tumors should be particularly prioritized for timely treatment initiation. The geographic patterns visible in Figure 2 indicate regional variations in healthcare delivery that transcend urban-rural distinctions. These patterns can guide targeted resource allocation to regions where reducing treatment delays would yield the greatest survival benefits.

The overall ATE of 9.8 months represents a substantial extension of life for patients with metastatic breast cancer, who typically have a median survival of approximately 2-3 years. The fact that eliminating treatment delays could potentially extend life by nearly a year highlights the critical importance of timely care in oncology. These findings provide evidence-based guidance for clinical prioritization and health policy interventions aimed at reducing treatment delays across Florida counties.

Figure 1.   County-specific CATE versus sample size, distinguishing between urban (blue) and rural (red) counties. The vertical lines represent 95% credible intervals for each county's effect estimate.

pics/CATE_map_WA_Distant.png

Figure 2. Geographic distribution of county-specific CATE values across Florida counties. Darker shades indicate larger treatment effects.

Table 1
Performance Comparison: Our Method vs Ph Frailty

Metric	Estimand	Our Method		Ph Frailty	
		M1	M2	M1	M2
Coverage (%)	ATE	70.0	70.0	0.0	0.0
	RATE	86.7	53.3	16.7	13.3
	SPTE	80.0	36.7	0.0	0.0
Abs. Bias	ATE	0.14	0.15	0.34	0.38
	RATE	0.05	0.05	0.18	0.23
	SPTE	0.04	0.06	0.14	0.16
RMSE	ATE	0.18	0.17	0.35	0.39
	RATE	0.12	0.10	0.21	0.26
	SPTE	0.07	0.07	0.14	0.16

Table 2
Marginal Treatment Effects for WA Patients with Distant Stage Cancer

Measure	Estimate (months)	Lower 95% CI	Upper 95% CI
ATE	9.8	-4.7	20.0
RATE (5 years)	4.2	-1.0	9.3
RATE (10 years)	6.5	-1.9	13.9

Table 3
Range of County-Specific Treatment Effects

Measure	Minimum	Maximum	Inter-county Range
CATE (months)	2.6	18.4	15.8
CRATE-5yr (months)	1.3	9.1	7.8

Table 4
Top 10 Counties by Conditional Average Treatment Effect

County	Sample Size	CATE (months)	Lower 95% CI	Upper 95% CI
Brevard	1	18.4	18.4	18.4
Martin	2	16.4	16.4	16.4
Union	2	14.4	10.4	18.5
Seminole	12	14.3	-0.6	22.3
Lake	2	13.9	9.6	18.2
Hillsborough	11	13.0	8.9	17.5
Sarasota	154	11.8	-5.3	22.1
Orange	52	11.8	-5.4	22.0
Calhoun	18	11.5	10.4	19.2
Clay	6	11.5	8.6	16.3

Table 5
Covariate-Specific CATE for Broward County

Age	Tumor Grade	HR Status	CATE (months)	Lower 95% CI	Upper 95% CI
50	1	0	21.4	20.6	22.3
50	1	1	11.1	10.7	11.6
50	2	0	19.0	18.3	19.8
50	2	1	9.9	9.6	10.4
50	3	0	18.3	17.6	19.1
50	3	1	9.6	9.2	10.0
75	1	0	21.4	20.6	22.3
75	1	1	11.1	10.7	11.6
75	2	0	19.0	18.3	19.8
75	2	1	9.9	9.6	10.4
75	3	0	18.3	17.6	19.1
75	3	1	9.6	9.2	10.0