$\mathbf{F_X}$: the probability distribution of the full data X.
G: the conditional probability distribution of C, given X, also called the censoring mechanism, treatment mechanism, or action mechanism, depending on what C stands for. When we define CAR, we do this in terms of the conditional distribution of Y, given X, which is determined by G, and, by CAR, it is the identifiable part of G. In this book we frequently denote the conditional distribution of Y, given X, with G as well.
$\mathbf{P_{F_X,G}}$: the distribution of the observed data Y, which only depends on G through the conditional distribution of Y, given X.
$\mathcal{G}$: a model for the censoring mechanism G (i.e., it is known that $G \in \mathcal{G}$).
$\mathcal{G}(\mathbf{CAR})$: all conditional distributions G of C, given X, satisfying coarsening at random (CAR).
$\mathcal{M}^{\mathbf{F}}$: a model for F_X (i.e., the full data model).
$\mathcal{M} = \{P_{F_X,G} : F_X \in \mathcal{M}^{F,w}, G \in \mathcal{G}(CAR)\} \cup \{P_{F_X,G} : F_X \in \mathcal{M}^F, G \in \mathcal{G}\}$, the observed data model allowing that either the working model $\mathcal{M}^{F,w}$ for F_X or the censoring model $\mathcal{G}$ for G is misspecified, but not both.
$\mathcal{M}(\mathcal{G}) = \{P_{F_X,G} : F_X \in \mathcal{M}^F, G \in \mathcal{G}\}$: the observed data model when assuming a correctly specified model $\mathcal{G}$ for G.
$\mathcal{M}(\mathbf{G}) = \{P_{F_X,G} : F_X \in \mathcal{M}^F\}$: the observed data model if the censoring mechanism G is known.
$\mathcal{M}(\mathbf{CAR}) = \{P_{F_X,G} : F_X \in \mathcal{M}^F, G \in \mathcal{G}(CAR)\}$: the observed data model if the censoring mechanism is only known to satisfy CAR.
$\mu = \mu(\mathbf{F_X}) \in \mathbb{R}^k$: the Euclidean parameter of F_X of interest.
$\mathbf{Z} = \mathbf{g}(\mathbf{X}^* \mid \alpha) + \epsilon$: a multivariate generalized regression model (a particular choice of full data model), where Z is a p-variate outcome, X^* is a vector of covariates, $g(X^* \mid \alpha)$ is a p-dimensional vector whose components are regression curves parametrized with a regression parameter $\alpha \in \mathbb{R}^k$, ϵ is a p-variate residual satisfying $E(K(\epsilon_j) \mid X^*) = 0$, $j = 1, \ldots, p$, for a given monotone nondecreasing function K.
$\mathbf{K}(\cdot)$: monotone function specifying the location parameter (e.g., mean, median, truncated mean, smooth median) of the conditional distribution of Z, given X^*, modeled by $g(X^* \mid \alpha)$. For example, 1) $K(\epsilon) = \epsilon$, 2) $K(\epsilon) = I(\epsilon > 0) - (1-p)$, 3) $K(\epsilon) = \epsilon$ on $[-\tau, \tau]$ and $K(\epsilon) = \tau$ for $\epsilon > \tau$, $K(\epsilon) = -\tau$ for $\epsilon < -\tau$ correspond with mean regression, pth quantile regression (e.g., $p = 0.5$ gives median regression), and truncated mean regression, respectively.
$\mathbf{N(t)}$: a counting process being a part of the full data X.
$\lambda(t)dt = E(dN(t) \mid \bar{Z}(t-)) = Y(t)\lambda_0(t)\exp(\beta W(t))dt$: a multiplicative intensity model (a particular full data model), where $\bar{Z}(t)$ is a given function of $\bar{X}(t)$ including the past $\bar{N}(t)$ of the counting process N, $Y(t)$ is an indicator function of $\bar{Z}(t-)$ (indicator that $N(\cdot)$ is at risk of jumping at time t), and $W(t)$ is a vector of covariates extracted from $\bar{Z}(t-)$. We also consider the case where $\bar{Z}(t-)$ does not include the past of N. In this case, we refer to these models as proportional rate models. We also consider discrete multiplicative intensity models, where $\lambda(t) = Y(t)\Lambda_0(dt)\exp(betaW(t))$ is now a conditional probability.

$\bar{\mathbf{F}} = 1 - F$.
$< \mathbf{S_1}, \ldots, \mathbf{S_k} >$: linear span of k elements (typically scores in $L_0^2(F_X)$ or $L_0^2(P_{F_X,G})$) in a Hilbert space.
$< \vec{\mathbf{S}} > \equiv < S_1, \ldots, S_k >$: linear span of the k components of $\vec{S}$.
$\langle \mathbf{f}, \mathbf{g} \rangle$: inner product defined in a Hilbert Space.
$\langle \mathbf{f}, (\mathbf{g_1}, \ldots, \mathbf{g_k}) \rangle \equiv (\langle f, g_1 \rangle, \ldots, \langle f, g_k \rangle)$.
$\mathbf{H_1} \oplus \mathbf{H_2} = \{h_1 + h_2 : h_j \in H_j, j = 1, 2\}$: the sum space spanned by two orthogonal sub-Hilbert spaces H_1, H_2 of a certain Hilbert space.
$\mathbf{H_1} + \mathbf{H_2} = \{h_1 + h_2 : h_j \in H_j, j = 1, 2\}$: the sum space spanned by two sub-Hilbert spaces H_1, H_2 of a certain Hilbert space.
$\mathbf{\Pi}(\cdot \mid \mathbf{H})$: the projection operator onto a subspace H of a certain Hilbert space.
$\mathbf{L_0^2(F_X)}$: Hilbert space of functions $h(X)$ with $E_{F_X} h(X) = 0$ with inner product $\langle h, g \rangle_{F_X} = E_{F_X} h(X)g(X)$ and corresponding norm $\| h \|_{F_X} = \sqrt{E_{F_X} h^2(X)}$.
$\mathbf{T^F(F_X)} \subset L_0^2(F_X)$: the tangent space at F_X in the full data model $\mathcal{M}^F$. This is the closure of the linear space spanned by scores of a given class of one-dimensional submodels $\epsilon \to F_\epsilon$ that cross F_X at $\epsilon = 0$.
$\mathbf{T^F_{nuis}(F_X)} \subset L_0^2(F_X)$: the nuisance tangent space at F_X in the full data model $\mathcal{M}^F$. This is the closure of the linear space spanned by scores of a given class of one-dimensional submodels $\epsilon \to F_\epsilon$ that cross F_X at $\epsilon = 0$ and satisfy $d/d\epsilon \mu(F_\epsilon)|_{\epsilon=0} = 0$.
$\mathbf{T^{F,\perp}_{nuis}(F_X)} \subset L_0^2(F_X)$: the orthogonal complement of the nuisance tangent space $T^F_{nuis}(F_X)$ in model $\mathcal{M}^F$, where μ is the parameter of interest. The class of full data estimating functions $D_h(\cdot \mid \mu, \rho)$, $h \in \mathcal{H}^F$, is chosen so that $T^{F,\perp}_{nuis}(F_X) \supset \{D_h(X \mid \mu(F_X), \rho(F_X)) : h \in \mathcal{H}^F\}$, where the right hand side is chosen as rich as possible so that we might even have equality.
$\mathbf{D_h}$, defined in full data model $\mathcal{M}^F$: full data estimating function $D_h : \mathcal{X} \times \{(\mu(F_X), \rho(F_X)) : F_X \in \mathcal{M}^F\} \to \mathbb{R}$ for parameter μ with nuisance parameter ρ. Here $h \in \mathcal{H}^F$ indexes different possible choices of full data estimating functions.
$\mathcal{H}^{\mathbf{F}}$: index set providing a rich class of full data estimating functions satisfying:

$$D_h(X \mid \mu(F_X), \rho(F_X)) \in T^{F,\perp}_{nuis}(F_X) \text{ for all } h \in \mathcal{H}^F.$$

$D_h, h \in \mathcal{H}^{F,k}$: for $h = (h_1, \ldots, h_k) \in \mathcal{H}^{Fk}$,

$$D_{h_1,\ldots,h_k}(X \mid \mu, \rho) = (D_{h_1}(X \mid \mu, \rho), \ldots, D_{h_k}(X \mid \mu, \rho)).$$

A full data structure estimating function D_h, $h \in \mathcal{H}^{Fk}$, defines an estimating equation for μ: given an estimate of ρ, one can estimate μ with the solution of the k-dimensional equation $0 = \sum_{i=1}^n D_h(X_i \mid \mu, \rho_n)$.
$\mathbf{D_h(X \mid \mu, \rho)}$: full data estimating function D_h evaluated at X, μ, ρ. Sometimes, $D_h(X \mid \mu, \rho)$ is used to denote the actual estimating function,