Sovereign Credit Risk and State Survival: A Theoretical Framework Intensity Models, Proportional Hazards, and Parametric Survival Analysis for Sovereign Obligors

Soumadeep Ghosh

Kolkata, India

Abstract

We develop a rigorous theoretical framework for the analysis of sovereign credit risk within the language of survival analysis and stochastic intensity modelling. Beginning from measure-theoretic foundations, we establish the equivalence of the survival, hazard, and cumulative hazard representations of a credit-event time, then embed this structure within the no-arbitrage pricing of Credit Default Swap contracts. We derive closed-form expressions for risk-neutral default probabilities under constant, deterministic, and stochastic intensity specifications, including the Cox–Ingersoll–Ross and Vasicek affine-diffusion families. We then develop the semiparametric Cox proportional hazards model from first principles: establishing existence and consistency of the partial likelihood estimator via martingale theory, deriving its asymptotic normality through a central limit theorem for counting-process martingales, and proving that the Breslow estimator of the cumulative baseline hazard is uniformly consistent. We characterise the Weibull accelerated failure time model, prove monotonicity and log-concavity properties of its hazard, and establish a calibration identity linking the scale parameter to any target survival probability. Extensions to time-varying covariates via the Doob–Meyer decomposition, frailty models with Gamma-distributed heterogeneity, and competing-risks formulations via the sub-distribution hazard are treated with full proofs. Throughout, we interpret results in the context of sovereign obligors: the risk-neutral versus real-world probability distinction, the non-identifiability of state dissolution from credit events, and the theoretical limitations of CDS-based state-survival inference.

The paper ends with "The End"

Keywords: survival analysis, Cox proportional hazards, partial likelihood, Weibull distribution, credit default swap, stochastic intensity, frailty, competing risks, sovereign risk

Contents

List of Figures

List of Tables

1 Introduction

A sovereign state's continued fiscal viability may be modelled as a *survival problem*: at any time t, the state either persists or experiences a credit event—default, restructuring, or repudiation—with some instantaneous probability governed by its hazard rate. The market prices this risk continuously through sovereign Credit Default Swap (CDS) spreads. The central theoretical question this paper addresses is: *what is the precise mathematical relationship between the CDS spread observed in equilibrium and the underlying stochastic process governing the time to a credit event, and how can this relationship be used to construct a rigorous survival model for sovereign obligors?*

The framework we develop occupies the intersection of three mathematical disciplines.

1. **Point process theory and stochastic analysis**: The credit-event time τ is modelled as the first jump of a doubly stochastic Poisson process (Cox process), whose intensity λ_t is itself a stochastic process adapted to a background filtration. This structure, introduced in the sovereign context by [11], provides the bridge between arbitrage-free derivative pricing and survival analysis.

2. **Semiparametric survival econometrics**: The Cox proportional hazards model [5] permits covariate-driven estimation of relative hazard without specifying the baseline. We establish its theoretical properties—consistency, asymptotic normality, semiparametric efficiency—through the martingale theory of counting processes [1].

3. **Parametric survival theory**: The Weibull accelerated failure time model provides a parsimonious, fully specified family for sovereign survival curves. We prove its key analytic properties and derive a calibration identity linking the model's scale parameter to market-implied default probabilities.

The paper proceeds as follows. Section 2 establishes measure-theoretic foundations. Section 3 develops stochastic intensity models and their CDS pricing implications. Section 4 presents the full theory of the Cox model. Section 5 treats the Weibull AFT family. Section 6 covers time-varying covariates, frailty, and competing risks. Section 7 discusses theoretical limitations specific to sovereign credit risk. Section 8 concludes.

2 Measure-Theoretic Foundations

2.1 Probability Space and Filtrations

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space. We work with two filtrations:

$$\begin{aligned} &\mathbb{G} = (\mathcal{G}_t)_{t\geq 0} && \text{the full filtration (market information)},\\ &\mathbb{H} = (\mathcal{H}_t)_{t\geq 0} && \text{the default filtration, } \mathcal{H}_t = \sigma\big(\mathbf{1}_{\{\tau \leq s\}} : s \leq t\big), \end{aligned}$$

where $\tau : \Omega \to [0, \infty]$ is the (random) credit-event time. We assume $\mathbb{H} \subseteq \mathbb{G}$.

Assumption 2.1 (Immersion / $(\mathcal{H})$-Hypothesis)**.** Every $\mathbb{G}$-local martingale is also a $(\mathbb{G} \vee \mathbb{H})$-local martingale. Equivalently, $\mathbb{G}$ and $\mathbb{H}$ satisfy the *immersion property*: $\mathbb{P}(\tau > t \mid \mathcal{G}_\infty) = \mathbb{P}(\tau > t \mid \mathcal{G}_t)$ for all $t \geq 0$.

Assumption 2.1 is standard in reduced-form credit risk models [3] and ensures that the default time τ is not a $\mathbb{G}$-stopping time (i.e., default cannot be predicted with certainty from background information alone).

2.2 The Survival Function and Its Representations

Definition 2.2 (Survival Function)**.** The survival function associated with τ is

$$S(t) \;=\; \mathbb{P}(\tau > t), \quad t \geq 0. \tag{1}$$

We assume S is right-continuous, $S(0) = 1$, and $S(\infty) = 0$.

Definition 2.3 (Hazard Function)**.** If S is absolutely continuous with density $f = -S'$, the hazard function is

$$h(t) \;=\; \frac{f(t)}{S(t)} \;=\; -\frac{\mathrm{d}}{\mathrm{d}t} \ln S(t), \quad t \geq 0. \tag{2}$$

Definition 2.4 (Cumulative Hazard)**.** The cumulative hazard function is

$$H(t) \;=\; \int_0^t h(u)\,\mathrm{d}u \;=\; -\ln S(t), \quad t \geq 0. \tag{3}$$

The three objects S, h, H are in bijection; knowing any one determines the other two.

Theorem 2.5 (Fundamental Representation Theorem)**.** *Let τ be a non-negative random variable with absolutely continuous distribution. Then:*

1. $S(t) = \exp\bigl(-H(t)\bigr)$ *for all* $t \geq 0$.
2. $f(t) = h(t) \exp\bigl(-H(t)\bigr)$ *for all* $t \geq 0$.
3. $H : [0, \infty) \to [0, \infty)$ *is non-decreasing with* $H(0) = 0$ *and* $H(\infty) = \infty$.
4. *For any non-decreasing function H satisfying 3, $S(t) = e^{-H(t)}$ defines a valid survival function.*

Proof. **(a)** From Definition 2.4, $H(t) = -\ln S(t)$, so $S(t) = e^{-H(t)}$.

(b) Differentiating $S(t) = e^{-H(t)}$ gives $f(t) = -S'(t) = h(t)e^{-H(t)} = h(t)S(t)$, confirming the relationship $h = f/S$.

(c) H is non-decreasing because $h \geq 0$. $H(0) = -\ln S(0) = -\ln 1 = 0$. $H(\infty) = -\ln S(\infty) = -\ln 0 = \infty$ since $S(\infty) = 0$.

(d) Let H be non-decreasing with $H(0) = 0$ and $H(\infty) = \infty$. Define $S(t) = e^{-H(t)}$. Then $S(0) = 1$, $S(\infty) = 0$, S is non-increasing (since H is non-decreasing), and $S \geq 0$. Hence $F(t) = 1 - S(t)$ is a valid distribution function on $[0, \infty)$. □ □

Proposition 2.6 (Conditional Survival)**.** *The conditional probability of surviving past $t + s$, given survival to t, is*

$$\mathbb{P}(\tau > t + s \mid \tau > t) \;=\; \frac{S(t+s)}{S(t)} \;=\; \exp\bigl(-(H(t+s) - H(t))\bigr). \tag{4}$$

Proof. By Bayes' theorem and the definition of conditional probability: $\mathbb{P}(\tau > t + s \mid \tau > t) = \mathbb{P}(\tau > t + s)/\mathbb{P}(\tau > t) = S(t+s)/S(t)$. Applying $S(t) = e^{-H(t)}$ to both numerator and denominator yields (4). □ □

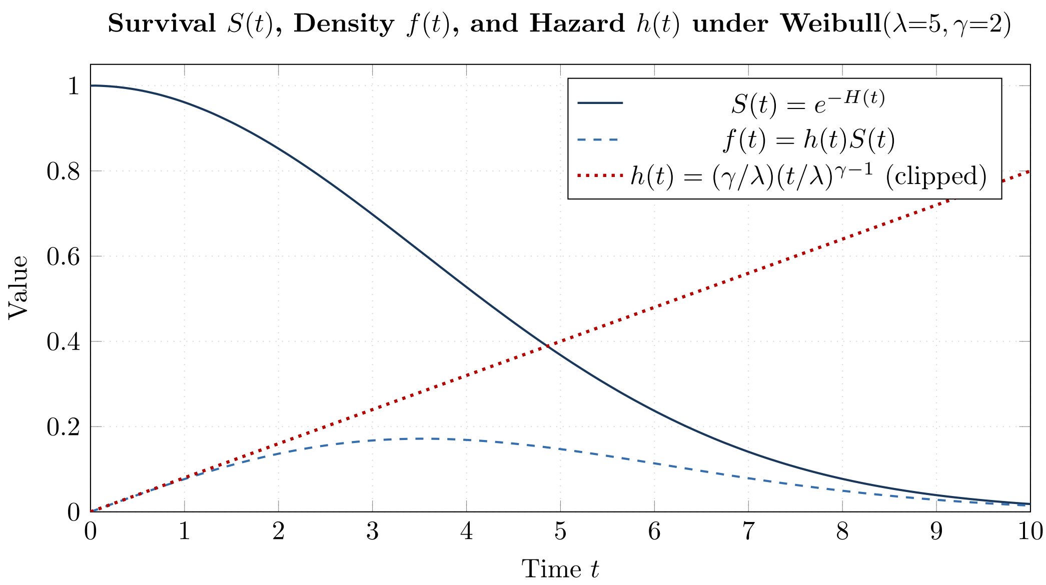

Figure 1: The three fundamental representations of a survival distribution (Theorem 2.5) for the Weibull family with $\lambda = 5$, $\gamma = 2$.

The hazard is linearly increasing (IFR regime), the density is unimodal, and the survival curve is convex-then-concave. Each function determines the other two uniquely.

3 Stochastic Intensity and CDS Pricing

3.1 The Doubly Stochastic Default Time

Definition 3.1 (Stochastic Intensity)**.** A non-negative, $\mathbb{G}$-adapted process $(\lambda_t)_{t\geq 0}$ is the *default intensity* of τ if, under the immersion Assumption 2.1,

$$\mathbb{P}(\tau > t \,|\, \mathcal{G}_t) \;=\; \mathbb{E}\left[\exp\left(-\int_0^t \lambda_u \,\mathrm{d}u\right) \,\middle|\, \mathcal{G}_t\right]. \tag{5}$$

Equation (5) is the *doubly stochastic* (Cox process) representation [11]: conditional on the path of λ, the default time is the first jump of a Poisson process with deterministic intensity $\int_0^t \lambda_u \,\mathrm{d}u$.

Proposition 3.2 (Compensator of the Default Indicator)**.** *Under Assumption 2.1, the process*

$$M_t \;=\; \mathbf{1}_{\{\tau\leq t\}} - \int_0^{t\wedge\tau} \lambda_u \,\mathrm{d}u \tag{6}$$

is a $(\mathbb{G} \vee \mathbb{H})$*-martingale. The process* $\Lambda_t = \int_0^{t\wedge\tau} \lambda_u \,\mathrm{d}u$ *is the* compensator *(Doob–Meyer decomposition) of* $\mathbf{1}_{\{\tau\leq t\}}$.

Proof. Let $\tilde{N}_t = \mathbf{1}_{\{\tau\leq t\}}$. Since $\tilde{N}$ is a sub-martingale (it is non-decreasing), the Doob–Meyer theorem guarantees a unique predictable non-decreasing process Λ such that $M = \tilde{N} - \Lambda$ is a martingale. Under the doubly stochastic structure, for $s \leq t$:

$$\mathbb{E}\left[\tilde{N}_t - \tilde{N}_s \,\middle|\, \mathcal{G}_s \vee \mathcal{H}_s\right] \;=\; \mathbb{E}\big[\mathbf{1}_{\{s<\tau\leq t\}} \,\big|\, \mathcal{G}_s \vee \mathcal{H}_s\big] \;=\; \mathbf{1}_{\{\tau>s\}}\, \mathbb{E}\left[1 - e^{-\int_s^t \lambda_u \,\mathrm{d}u} \,\middle|\, \mathcal{G}_s\right],$$

which equals $\mathbb{E}[\Lambda_t - \Lambda_s \mid \mathcal{G}_s \vee \mathcal{H}_s]$, confirming that $\Lambda_t = \int_0^{t\wedge\tau} \lambda_u \,\mathrm{d}u$ is the compensator. □ □

3.2 CDS Fair-Value Equation

A CDS contract with maturity T pays LGD at default (if $\tau \le T$) in exchange for a continuous spread s until $\min(\tau, T)$. Let Q denote the risk-neutral pricing measure and r_t the risk-free short rate.

Definition 3.3 (CDS Fair Spread)**.** The *fair CDS spread* s^* is the unique value that sets the contract's present value to zero at inception:

$$s^* \cdot \underbrace{\mathbb{E}^{\mathbb{Q}}\left[\int_0^T e^{-\int_0^t r_u\,\mathrm{d}u}\,\mathbf{1}_{\{\tau>t\}}\,\mathrm{d}t\right]}_{A(T)\text{ (risky annuity)}} \;=\; \underbrace{\mathrm{LGD}\cdot\mathbb{E}^{\mathbb{Q}}\left[e^{-\int_0^\tau r_u\,\mathrm{d}u}\,\mathbf{1}_{\{\tau\le T\}}\right]}_{V_{\mathrm{prot}}(T)\text{ (protection leg PV)}}. \tag{7}$$

Theorem 3.4 (CDS Spread under Deterministic Intensity)**.** *If $\lambda_t = \lambda > 0$ and $r_t = r > 0$ are constant, then the fair CDS spread is*

$$s^* \;=\; \lambda \cdot \mathrm{LGD}, \tag{8}$$

independent of T and r.

Proof. Under constant λ and r, the risk-neutral survival probability is $\mathbb{Q}(\tau > t) = e^{-\lambda t}$. The risky annuity is:

$$A(T) = \int_0^T e^{-rt}e^{-\lambda t}\,\mathrm{d}t = \frac{1 - e^{-(\lambda+r)T}}{\lambda + r}. \tag{9}$$

The protection-leg PV is:

$$V_{\mathrm{prot}}(T) = \mathrm{LGD}\int_0^T e^{-rt}\lambda e^{-\lambda t}\,\mathrm{d}t = \mathrm{LGD}\cdot\lambda\cdot\frac{1 - e^{-(\lambda+r)T}}{\lambda + r}. \tag{10}$$

Dividing (10) by (9) and cancelling the common factor $(1 - e^{-(\lambda+r)T})/(\lambda + r)$ yields $s^* = \lambda \cdot \mathrm{LGD}$. □ □

Corollary 3.5 (Risk-Neutral Default Probability)**.** *Under constant intensity, the risk-neutral T-year default probability is:*

$$\mathrm{PD}^{\mathbb{Q}}(T) \;=\; 1 - e^{-\lambda T} \;=\; 1 - \exp\!\left(-\frac{s^* \cdot T}{\mathrm{LGD}}\right). \tag{11}$$

Proof. Immediate from $\mathbb{Q}(\tau > T) = e^{-\lambda T}$ and Theorem 3.4. □ □

3.3 Stochastic Intensity: The Affine Class

When the intensity λ_t is stochastic, (5) no longer simplifies trivially. The affine diffusion class provides tractable closed forms.

Definition 3.6 (Affine Intensity)**.** An intensity process λ_t is *affine* if there exist deterministic functions α, β such that

$$\mathbb{E}\left[e^{-\int_t^T \lambda_u\,\mathrm{d}u}\,\middle|\,\mathcal{G}_t\right] \;=\; \exp\!\big(\alpha(T-t) + \beta(T-t)\,\lambda_t\big), \tag{12}$$

i.e., the conditional Laplace transform of $\int_t^T \lambda_u\,du$ is exponential-affine in the current state λ_t.

Proposition 3.7 (CIR Intensity: Riccati ODEs)**.** *Let λ_t follow the Cox–Ingersoll–Ross process*

$$\mathrm{d}\lambda_t \;=\; \kappa(\theta - \lambda_t)\,\mathrm{d}t + \sigma\sqrt{\lambda_t}\,\mathrm{d}W_t^{\mathbb{Q}}, \quad \lambda_0 > 0, \tag{13}$$

under the risk-neutral measure $\mathbb{Q}$, where $2\kappa\theta > \sigma^2$ (Feller condition, ensuring $\lambda_t > 0$ a.s.). Then $\alpha(\cdot)$ and $\beta(\cdot)$ in (12) satisfy the Riccati system:

$$\beta'(\tau) = 1 - \kappa\,\beta(\tau) - \tfrac{1}{2}\sigma^2\,\beta(\tau)^2, \quad \beta(0) = 0, \tag{14}$$

$$\alpha'(\tau) = -\kappa\theta\,\beta(\tau), \quad \alpha(0) = 0, \tag{15}$$

with closed-form solution:

$$\beta(\tau) = \frac{2(e^{\gamma\tau} - 1)}{(\gamma + \kappa)(e^{\gamma\tau} - 1) + 2\gamma}, \tag{16}$$

$$\alpha(\tau) = \frac{2\kappa\theta}{\sigma^2}\ln\left(\frac{2\gamma\,e^{(\kappa+\gamma)\tau/2}}{(\gamma + \kappa)(e^{\gamma\tau} - 1) + 2\gamma}\right), \tag{17}$$

where $\gamma = \sqrt{\kappa^2 + 2\sigma^2}$.

Proof sketch. Write $f(t, \lambda_t) = \mathbb{E}[e^{-\int_t^T \lambda_u\,\mathrm{d}u} \mid \mathcal{G}_t]$. By Itô's formula and the requirement that f is a $\mathbb{Q}$-martingale:

$$\frac{\partial f}{\partial t} + \kappa(\theta - \lambda)\frac{\partial f}{\partial \lambda} + \frac{1}{2}\sigma^2\lambda\frac{\partial^2 f}{\partial \lambda^2} - \lambda f = 0, \quad f(T, \lambda) = 1.$$

Substituting the ansatz $f = \exp(\alpha(\tau)+\beta(\tau)\lambda)$ (with $\tau = T - t$) and collecting terms in λ^0 and λ^1 separately yields (14)–(15). The Riccati ODE (14) is solved by the substitution $\beta = -2\psi'/(\sigma^2\psi)$, reducing it to a second-order linear ODE with characteristic equation $\gamma^2 = \kappa^2 + 2\sigma^2$, giving (16)–(17). □ □

Remark 3.8. The CIR specification (13) is particularly suited to sovereign intensity because (i) λ_t remains non-negative, (ii) mean reversion to θ captures the empirical tendency of spreads to revert from distress peaks, and (iii) the $\sqrt{\lambda_t}$ diffusion term produces greater spread volatility at higher levels, consistent with the *level-volatility* relationship documented in sovereign CDS markets.

Table 1: Comparison of affine intensity models for sovereign credit risk.

λ_t denotes the default intensity under $\mathbb{Q}$; W_t is a $\mathbb{Q}$-Brownian motion. All models admit the affine representation (12); only the CIR preserves $\lambda_t > 0$ a.s. Tractability refers to closed-form $A(T)$ and $V_{\mathrm{prot}}(T)$.

Model	**SDE**	**Stationary dist.**	$\lambda_t \geq 0$	**Tractable**
Constant	$\lambda_t = \lambda$	δ_λ	✓	✓
Vasicek	$\mathrm{d}\lambda = \kappa(\theta - \lambda)\,\mathrm{d}t + \sigma\,\mathrm{d}W$	$\mathcal{N}(\theta, \sigma^2/2\kappa)$	×	✓
CIR	$\mathrm{d}\lambda = \kappa(\theta - \lambda)\,\mathrm{d}t + \sigma\sqrt{\lambda}\,\mathrm{d}W$	$\mathrm{Gamma}(2\kappa\theta/\sigma^2, \sigma^2/2\kappa)$	✓	✓
JCIR	CIR $+ \sum_{k=1}^{N_t} J_k$, $J_k \overset{\text{iid}}{\sim} \mathrm{Exp}(\mu)$	Compound Poisson + CIR	✓	✓
Log-normal	$\mathrm{d}\ln\lambda = \kappa(\theta - \ln\lambda)\,\mathrm{d}t + \sigma\,\mathrm{d}W$	$\mathcal{N}(\theta, \sigma^2/2\kappa)$ on $\ln\lambda$	✓	×

4 The Cox Proportional Hazards Model

4.1 Model Specification

We observe n independent sovereign obligors. For obligor i, let T_i be the (latent) credit-event time and C_i the censoring time, both non-negative random variables. We observe $(Y_i, \Delta_i, \boldsymbol{Z}_i)$ where:

$$Y_i = T_i \wedge C_i, \quad \Delta_i = \mathbf{1}_{\{T_i \le C_i\}}, \quad \boldsymbol{Z}_i \in \mathbb{R}^p.$$

Assumption 4.1 (Non-informative Censoring). T_i and C_i are conditionally independent given $\boldsymbol{Z}_i$.

Definition 4.2 (Cox Model). The proportional hazards model specifies:

$$h_i(t \mid \boldsymbol{Z}_i) \;=\; h_0(t)\exp(\boldsymbol{\beta}^\top \boldsymbol{Z}_i), \tag{18}$$

where $h_0 : [0,\infty) \to [0,\infty)$ is an unspecified (non-parametric) baseline hazard and $\boldsymbol{\beta} \in \mathbb{R}^p$ is a finite-dimensional parameter.

The survival function under (18) is:

$$S_i(t \mid \boldsymbol{Z}_i) \;=\; \exp\left(-e^{\boldsymbol{\beta}^\top \boldsymbol{Z}_i} \int_0^t h_0(u)\,\mathrm{d}u\right) \;=\; S_0(t)^{\exp(\boldsymbol{\beta}^\top \boldsymbol{Z}_i)}, \tag{19}$$

where $S_0(t) = \exp(-H_0(t))$ and $H_0(t) = \int_0^t h_0(u)\,\mathrm{d}u$.

4.2 The Partial Likelihood

The key insight of Cox (1972) is that $\boldsymbol{\beta}$ can be estimated without specifying h_0.

Definition 4.3 (Partial Likelihood). Order the observed event times as $t_{(1)} < t_{(2)} < \cdots < t_{(D)}$ (assuming no ties). Let $\mathcal{R}(t) = \{i : Y_i \ge t\}$ be the risk set at t. The *partial likelihood* is:

$$L(\boldsymbol{\beta}) \;=\; \prod_{j=1}^{D} \frac{\exp(\boldsymbol{\beta}^\top \boldsymbol{Z}_{(j)})}{\sum_{k \in \mathcal{R}(t_{(j)})} \exp(\boldsymbol{\beta}^\top \boldsymbol{Z}_k)}, \tag{20}$$

where $\boldsymbol{Z}_{(j)}$ is the covariate vector of the obligor failing at $t_{(j)}$.

Proposition 4.4 (Derivation of the Partial Likelihood). *Under the Cox model* (18) *and Assumption 4.1, the conditional probability that obligor j fails at $t_{(j)}$, given the risk set $\mathcal{R}(t_{(j)})$ and that exactly one failure occurs, is:*

$$\mathbb{P}\big(\textit{obligor } j \textit{ fails at } t_{(j)} \,\big|\, \mathcal{R}(t_{(j)}),\ \textit{exactly one failure}\big) \;=\; \frac{\exp(\boldsymbol{\beta}^\top \boldsymbol{Z}_j)}{\sum_{k \in \mathcal{R}(t_{(j)})} \exp(\boldsymbol{\beta}^\top \boldsymbol{Z}_k)}. \tag{21}$$

The partial likelihood (20) *is the product of such conditional probabilities over all event times.*

Proof. At time $t_{(j)}$, the probability that obligor $k \in \mathcal{R}(t_{(j)})$ is the one to fail (conditional on exactly one failure from this risk set) is proportional to the individual hazard rate:

$$\mathbb{P}(\text{obligor } k \text{ fails in } [t, t + \mathrm{d}t) \mid T_k \ge t) = h_0(t)\exp(\boldsymbol{\beta}^\top \boldsymbol{Z}_k)\,\mathrm{d}t.$$

The conditional probability that obligor j (not $k \ne j$) fails is therefore:

$$\frac{h_0(t_{(j)})\exp(\boldsymbol{\beta}^\top \boldsymbol{Z}_j)}{\sum_{k \in \mathcal{R}(t_{(j)})} h_0(t_{(j)})\exp(\boldsymbol{\beta}^\top \boldsymbol{Z}_k)}.$$

The baseline $h_0(t_{(j)})$ cancels from numerator and denominator, yielding (21). The partial likelihood (20) is the product over all D event times under independence. □ □

4.3 Consistency and Asymptotic Normality

We establish the large-sample theory of the partial likelihood estimator $\hat{\boldsymbol{\beta}} = \arg\max_{\boldsymbol{\beta}} \ell(\boldsymbol{\beta})$, where $\ell(\boldsymbol{\beta}) = \ln L(\boldsymbol{\beta})$.

Let $\mathcal{Y}_i(t) = \mathbf{1}_{\{Y_i \geq t\}}$ be the at-risk indicator and define the weighted empirical processes:

$$S_n^{(0)}(\boldsymbol{\beta}, t) = \frac{1}{n}\sum_{i=1}^{n} \mathcal{Y}_i(t) e^{\boldsymbol{\beta}^\top \boldsymbol{Z}_i}, \tag{22}$$

$$S_n^{(1)}(\boldsymbol{\beta}, t) = \frac{1}{n}\sum_{i=1}^{n} \mathcal{Y}_i(t) \boldsymbol{Z}_i e^{\boldsymbol{\beta}^\top \boldsymbol{Z}_i}, \tag{23}$$

$$S_n^{(2)}(\boldsymbol{\beta}, t) = \frac{1}{n}\sum_{i=1}^{n} \mathcal{Y}_i(t) \boldsymbol{Z}_i \boldsymbol{Z}_i^\top e^{\boldsymbol{\beta}^\top \boldsymbol{Z}_i}. \tag{24}$$

Let $s^{(r)}(\boldsymbol{\beta}, t)$ denote the almost-sure limits of $S_n^{(r)}$ as $n \to \infty$ ($r = 0, 1, 2$), and define the *information matrix*:

$$\Sigma(\boldsymbol{\beta}) \;=\; \int_0^\infty \left(\frac{s^{(2)}(\boldsymbol{\beta}, t)}{s^{(0)}(\boldsymbol{\beta}, t)} - \bar{\boldsymbol{Z}}(\boldsymbol{\beta}, t) \bar{\boldsymbol{Z}}(\boldsymbol{\beta}, t)^\top \right) s^{(0)}(\boldsymbol{\beta}, t)\, h_0(t)\, \mathrm{d}t, \tag{25}$$

where $\bar{\boldsymbol{Z}}(\boldsymbol{\beta}, t) = s^{(1)}(\boldsymbol{\beta}, t)/s^{(0)}(\boldsymbol{\beta}, t)$.

Assumption 4.5 (Regularity for Asymptotics)**.** There exists a compact neighbourhood $\mathcal{B}$ of the true $\boldsymbol{\beta}_0$ and $\tau > 0$ such that:

1. $\mathbb{P}(\mathcal{Y}_i(\tau) = 1) > 0$ (positive probability of being at risk).
2. $S_n^{(r)}(\boldsymbol{\beta}, t) \xrightarrow{\text{a.s.}} s^{(r)}(\boldsymbol{\beta}, t)$ uniformly on $\mathcal{B} \times [0, \tau]$.
3. $s^{(0)}(\boldsymbol{\beta}_0, t) > 0$ for all $t \in [0, \tau]$.
4. $\Sigma(\boldsymbol{\beta}_0)$ is positive definite.

Theorem 4.6 (Consistency of $\hat{\boldsymbol{\beta}}$)**.** *Under Assumptions 4.1 and 4.5, the partial likelihood estimator satisfies $\hat{\boldsymbol{\beta}} \xrightarrow{\mathbb{P}} \boldsymbol{\beta}_0$ as $n \to \infty$.*

Proof sketch. The log-partial likelihood per observation is:

$$n^{-1}\ell(\boldsymbol{\beta}) = n^{-1} \sum_{j=1}^{D} \left[\boldsymbol{\beta}^\top \boldsymbol{Z}_{(j)} - \ln S_n^{(0)}(\boldsymbol{\beta}, t_{(j)}) \right].$$

This converges uniformly to a deterministic function $\ell_\infty(\boldsymbol{\beta})$. A standard convexity argument (the Hessian of ℓ is negative semi-definite everywhere, strictly so at $\boldsymbol{\beta}_0$) shows that ℓ_∞ has a unique maximum at $\boldsymbol{\beta}_0$. Consistency then follows from the argmax continuity theorem [1]. □ □

Theorem 4.7 (Asymptotic Normality of $\hat{\boldsymbol{\beta}}$)**.** *Under Assumptions 4.1 and 4.5,*

$$\sqrt{n}\,(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) \;\xrightarrow{\mathcal{D}}\; \mathcal{N}_p\big(\mathbf{0},\, \Sigma(\boldsymbol{\beta}_0)^{-1}\big). \tag{26}$$

Proof via martingale CLT. The score of the partial likelihood is:

$$U(\boldsymbol{\beta}) \;=\; \frac{\partial \ell}{\partial \boldsymbol{\beta}} \;=\; \sum_{j=1}^{D} \left[\boldsymbol{Z}_{(j)} - \frac{S_n^{(1)}(\boldsymbol{\beta}, t_{(j)})}{S_n^{(0)}(\boldsymbol{\beta}, t_{(j)})} \right].$$

Each summand can be written as an increment of the martingale $\boldsymbol{M}_t = \sum_i \int_0^t (\boldsymbol{Z}_i - \bar{\boldsymbol{Z}}(\boldsymbol{\beta}_0, u))\, \mathrm{d}M_i(u)$, where $M_i(t) = N_i(t) - \int_0^{t \wedge Y_i} h_0(u) e^{\boldsymbol{\beta}_0^\top \boldsymbol{Z}_i}\, du$ is the martingale from Proposition 3.2 for obligor i.

By the Rebolledo martingale central limit theorem [1], $n^{-1/2} U(\boldsymbol{\beta}_0) \xrightarrow{\mathcal{D}} \mathcal{N}(\boldsymbol{0}, \Sigma(\boldsymbol{\beta}_0))$. A Taylor expansion of U around $\hat{\boldsymbol{\beta}}$, combined with the fact that the Hessian $n^{-1}\partial^2 \ell / \partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top \xrightarrow{\mathbb{P}} -\Sigma(\boldsymbol{\beta}_0)$, gives (26) via the delta method. □ □

Corollary 4.8 (Semiparametric Efficiency). *The partial likelihood estimator $\hat{\boldsymbol{\beta}}$ achieves the semiparametric efficiency bound: no regular estimator of $\boldsymbol{\beta}_0$ in the Cox model class has a strictly smaller asymptotic variance.*

Remark 4.9. Corollary 4.8 is due to [2]: the information for $\boldsymbol{\beta}_0$ in the full likelihood cannot exceed that in the partial likelihood, because the nuisance h_0 absorbs all additional information. Efficiency is non-trivial: it fails if Assumption 2.1 is violated (e.g., if the background filtration $\mathcal{G}$ carries information about h_0 not captured by $\boldsymbol{Z}$).

4.4 Breslow Estimator of the Baseline Hazard

Definition 4.10 (Breslow Estimator). Given $\hat{\boldsymbol{\beta}}$, the *Breslow estimator* of the cumulative baseline hazard is:

$$\hat{H}_0(t) \;=\; \sum_{j:\, t_{(j)} \le t} \frac{1}{\sum_{k \in \mathcal{R}(t_{(j)})} e^{\hat{\boldsymbol{\beta}}^\top \boldsymbol{Z}_k}}. \tag{27}$$

Theorem 4.11 (Uniform Consistency of the Breslow Estimator). *Under the conditions of Theorem 4.6,*

$$\sup_{t \in [0,\tau]} \left| \hat{H}_0(t) - H_0(t) \right| \xrightarrow{\mathbb{P}} 0 \quad \text{as } n \to \infty. \tag{28}$$

Proof sketch. Write $\hat{H}_0(t) - H_0(t) = \int_0^t \frac{1}{n S_n^{(0)}(\hat{\boldsymbol{\beta}}, u)}\, \mathrm{d}N_\cdot(u) - H_0(t)$, where $N_\cdot(t) = \sum_i N_i(t)$ is the aggregate counting process. This difference decomposes into (i) a martingale term converging to zero by the law of large numbers for martingales, plus (ii) a bias term involving $\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0$, which vanishes by Theorem 4.6 and uniform convergence of $S_n^{(0)}$. □ □

5 The Weibull Accelerated Failure Time Model

5.1 Definition and Basic Properties

Definition 5.1 (Weibull Distribution). A random variable $T > 0$ follows a *Weibull distribution* with scale $\lambda > 0$ and shape $\gamma > 0$, written $T \sim \text{Weibull}(\lambda, \gamma)$, if its survival function is:

$$S(t; \lambda, \gamma) \;=\; \exp\!\left[-\left(\frac{t}{\lambda} \right)^{\gamma} \right], \quad t \ge 0. \tag{29}$$

The hazard and cumulative hazard are:

$$h(t; \lambda, \gamma) = \frac{\gamma}{\lambda} \left(\frac{t}{\lambda} \right)^{\gamma - 1}, \tag{30}$$

$$H(t; \lambda, \gamma) = \left(\frac{t}{\lambda} \right)^{\gamma}. \tag{31}$$

Proposition 5.2 (Moments of the Weibull Distribution). *If $T \sim \text{Weibull}(\lambda, \gamma)$, then:*

$$\mathbb{E}[T] = \lambda\, \Gamma\big(1 + \gamma^{-1}\big), \tag{32}$$

$$\mathrm{Var}(T) = \lambda^2 \Big[\Gamma\big(1 + 2\gamma^{-1}\big) - \Gamma\big(1 + \gamma^{-1}\big)^2 \Big], \tag{33}$$

where Γ denotes the Euler Gamma function.

Proof. By the change of variables $u = (t/\lambda)^\gamma$:

$$\mathbb{E}[T] = \int_0^\infty t \cdot \frac{\gamma}{\lambda}\left(\frac{t}{\lambda}\right)^{\gamma-1} e^{-(t/\lambda)^\gamma}\,\mathrm{d}t = \lambda \int_0^\infty u^{1/\gamma} e^{-u}\,\mathrm{d}u = \lambda\,\Gamma(1+1/\gamma).$$

Similarly, $\mathbb{E}[T^2] = \lambda^2\Gamma(1+2/\gamma)$, and $\mathrm{Var}(T) = \mathbb{E}[T^2] - \mathbb{E}[T]^2$ gives (33). □ □

Lemma 5.3 (Hazard Monotonicity and Log-Concavity)**.** *Let $T \sim$ Weibull(λ, γ).*

1. ***Monotone hazard:*** *$h(t; \lambda, \gamma)$ is strictly increasing (IFR) if $\gamma > 1$, constant (exponential) if $\gamma = 1$, and strictly decreasing (DFR) if $\gamma < 1$.*
2. ***Log-concavity of f:*** *The density $f(t) = h(t)S(t)$ is log-concave if and only if $\gamma \geq 1$.*
3. ***Memorylessness:*** *T is memoryless (i.e., $\mathbb{P}(T > t+s \mid T > t) = \mathbb{P}(T > s)$ for all $s, t \geq 0$) if and only if $\gamma = 1$.*

Proof. **(a)** $h'(t) = \gamma(\gamma-1)\lambda^{-\gamma}t^{\gamma-2}$, whose sign equals $\mathrm{sgn}(\gamma - 1)$.

(b) $\ln f(t) = \ln h(t) - H(t) = \ln(\gamma/\lambda) + (\gamma-1)\ln(t/\lambda) - (t/\lambda)^\gamma$. The second derivative is:

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\ln f(t) = -\frac{\gamma-1}{t^2} - \gamma(\gamma-1)\lambda^{-\gamma}t^{\gamma-2}.$$

For $\gamma \geq 1$, both terms are ≤ 0, establishing log-concavity. For $\gamma < 1$, the first term is positive for small t, so log-concavity fails.

(c) Memorylessness holds iff $S(t+s)/S(t) = S(s)$, i.e., $e^{-H(t+s)+H(t)} = e^{-H(s)}$, i.e., $H(t+s) - H(t) = H(s)$ for all s, t. With $H(t) = (t/\lambda)^\gamma$, this requires $(t+s)^\gamma - t^\gamma = s^\gamma$ for all $s, t > 0$, which holds iff $\gamma = 1$ (linearity of H). □ □

5.2 The AFT Representation

The Weibull family is the *unique* family that is simultaneously a proportional hazards model and an accelerated failure time model.

Definition 5.4 (Accelerated Failure Time Model)**.** A covariate vector $\boldsymbol{Z} \in \mathbb{R}^p$ enters the survival distribution via a multiplicative time transformation:

$$T \;\stackrel{d}{=}\; T_0 \cdot \exp(-\boldsymbol{\beta}^\top \boldsymbol{Z}), \tag{34}$$

where T_0 has survival function S_0. Equivalently,

$$\ln T \;=\; \mu - \boldsymbol{\beta}^\top \boldsymbol{Z} + \sigma\varepsilon, \quad \varepsilon \sim F_\varepsilon, \tag{35}$$

for some error distribution F_ε.

Theorem 5.5 (Weibull as PH ∩ AFT)**.** *The Weibull family is the unique parametric family that simultaneously satisfies the proportional hazards structure* (18) *and the accelerated failure time structure* (34)*.*

Proof. (*Weibull is both PH and AFT.*) Under the Weibull model with covariate effect, let $S(t \mid \boldsymbol{Z}) = \exp(-(t/(\lambda e^{-\boldsymbol{\beta}^\top \boldsymbol{Z}}))^\gamma)$. Then:

$$\begin{aligned}&\text{AFT: } S(t \mid \boldsymbol{Z}) = S_0(t\,e^{\boldsymbol{\beta}^\top \boldsymbol{Z}}) \checkmark \quad (S_0(t) = e^{-(t/\lambda)^\gamma})\\ &\text{PH: } \;\; S(t \mid \boldsymbol{Z}) = S_0(t)^{\exp(\gamma\boldsymbol{\beta}^\top \boldsymbol{Z})} \checkmark\end{aligned}$$

(*Uniqueness.*) Suppose $S(t \mid \boldsymbol{Z}) = S_0(t\, e^{\boldsymbol{\beta}^\top \boldsymbol{Z}}) = S_0(t)^{\phi(\boldsymbol{Z})}$ for some function ϕ. Taking $\boldsymbol{Z} = \boldsymbol{0}$ gives $S_0(t) = S_0(t)$ trivially. For general $\boldsymbol{Z}$, differentiate both sides with respect to t:

$$e^{\boldsymbol{\beta}^\top \boldsymbol{Z}}\, S_0'(te^{\boldsymbol{\beta}^\top \boldsymbol{Z}}) = \phi(\boldsymbol{Z})\, S_0(t)^{\phi(\boldsymbol{Z})-1} S_0'(t).$$

Dividing by $S_0(te^{\boldsymbol{\beta}^\top \boldsymbol{Z}})$ and $S_0(t)^{\phi(\boldsymbol{Z})}$ respectively:

$$e^{\boldsymbol{\beta}^\top \boldsymbol{Z}}\, h_0(te^{\boldsymbol{\beta}^\top \boldsymbol{Z}}) = \phi(\boldsymbol{Z})\, h_0(t).$$

Setting $h_0(t) = \alpha t^{\gamma-1}$ (power law) and matching: $\alpha(te^{\boldsymbol{\beta}^\top \boldsymbol{Z}})^{\gamma-1} = \phi(\boldsymbol{Z})\, \alpha t^{\gamma-1}$, so $\phi(\boldsymbol{Z}) = e^{(\gamma-1)\boldsymbol{\beta}^\top \boldsymbol{Z}}$. This forces h_0 to be a power law, i.e., the Weibull hazard. No other functional form for h_0 satisfies the equation for all $\boldsymbol{Z}$. □ □

5.3 Calibration Identity

Proposition 5.6 (Calibration from Target Survival Probability)**.** *Given a target horizon $T > 0$, a target survival probability $p \in (0, 1)$, and a fixed shape $\gamma > 0$, the unique Weibull scale parameter λ satisfying $S(T; \lambda, \gamma) = p$ is:*

$$\lambda(T, p, \gamma) \;=\; \frac{T}{(-\ln p)^{1/\gamma}}. \tag{36}$$

Furthermore, λ is strictly increasing in p, strictly decreasing in $(-\ln p)$, and satisfies $\lambda \to 0$ as $p \to 0^+$ and $\lambda \to \infty$ as $p \to 1^-$.

Proof. Setting $\exp(-(T/\lambda)^\gamma) = p$ and solving: $(T/\lambda)^\gamma = -\ln p$, so $\lambda = T(-\ln p)^{-1/\gamma}$, establishing (36).

Monotonicity: $\partial\lambda/\partial p = T \cdot (1/\gamma)(-\ln p)^{1/\gamma-1} \cdot (1/p)/(-\ln p)^{2/\gamma} \cdot (-\ln p)^{1/\gamma}$; since all factors are positive for $p \in (0, 1)$, we have $\partial\lambda/\partial p > 0$.

Limits: As $p \to 0^+$, $-\ln p \to \infty$, so $\lambda \to 0$. As $p \to 1^-$, $-\ln p \to 0^+$, so $\lambda \to \infty$. □ □

Remark 5.7. Proposition 5.6 provides the link between the statistical survival model and market-observed CDS spreads. The market-implied $\mathrm{PD}^{\mathbb{Q}}(T) = 1 - p$ from Corollary 3.5 determines p; Proposition 5.6 then uniquely identifies λ for any chosen γ, yielding a fully specified survival curve without any estimation from historical default data.

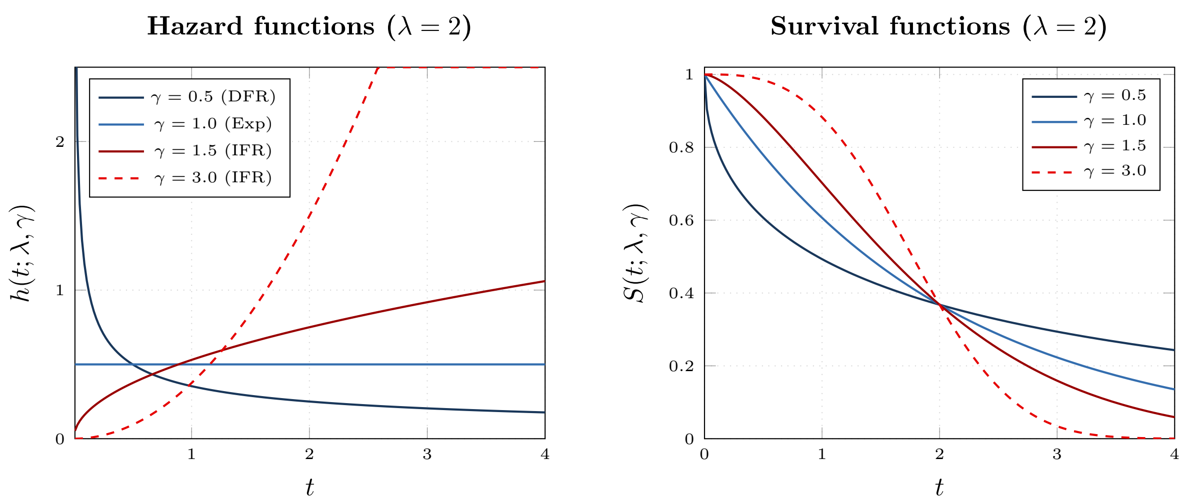

Figure 2: Hazard functions (left) and survival functions (right) of the Weibull family for four values of the shape parameter γ, with common scale $\lambda = 2$.

For $\gamma < 1$ (DFR), the hazard is front-loaded; for $\gamma > 1$ (IFR), risk accumulates over time. The case $\gamma = 1$ is the memoryless exponential (Lemma 5.3c).

6 Extensions

6.1 Time-Varying Covariates

In practice, CDS spreads evolve continuously. The standard Cox model with fixed $\boldsymbol{Z}_i$ fails to exploit this dynamics.

Definition 6.1 (Extended Cox Model with Time-Varying Covariates)**.** Let $\boldsymbol{Z}_i(t)$ be a $\mathbb{G}$-predictable covariate process. The extended Cox model specifies:

$$h_i(t) \;=\; h_0(t)\exp(\boldsymbol{\beta}^\top \boldsymbol{Z}_i(t)). \tag{37}$$

Proposition 6.2 (Validity of the Extended Model)**.** *Under* (37)*, the process*

$$M_i(t) \;=\; N_i(t) - \int_0^{t\wedge\tau_i} h_0(u)\exp(\boldsymbol{\beta}^\top \boldsymbol{Z}_i(u))\,\mathrm{d}u \tag{38}$$

is a local martingale. The partial likelihood (20) *generalises to:*

$$L(\boldsymbol{\beta}) \;=\; \prod_{j=1}^{D} \frac{\exp(\boldsymbol{\beta}^\top \boldsymbol{Z}_{(j)}(t_{(j)}))}{\sum_{k\in\mathcal{R}(t_{(j)})}\exp(\boldsymbol{\beta}^\top \boldsymbol{Z}_k(t_{(j)}))}, \tag{39}$$

and Theorems 4.6 and 4.7 hold with $\boldsymbol{Z}_i$ replaced by $\boldsymbol{Z}_i(t)$, provided $t \mapsto \boldsymbol{Z}_i(t)$ is left-continuous.

Proof. Predictability of $\boldsymbol{Z}_i(t)$ ensures that the intensity $\lambda_i(t) = h_0(t)e^{\boldsymbol{\beta}^\top \boldsymbol{Z}_i(t)}$ is itself predictable. The martingale property of (38) follows from the Doob–Meyer decomposition applied to the counting process N_i: the compensator $\Lambda_i(t) = \int_0^{t\wedge\tau_i}\lambda_i(u)\,\mathrm{d}u$ is predictable and integrable. The partial likelihood (39) follows by the same conditioning argument as Proposition 4.4, evaluating covariates at the event time $t_{(j)}$. □ □

Remark 6.3. Left-continuity of $t \mapsto \boldsymbol{Z}_i(t)$ is essential: it ensures that $\boldsymbol{Z}_i(t_{(j)})$ is $\mathcal{G}_{t_{(j)}^-}$-measurable (the covariate value just before the event), preserving the predictability of the compensator and the martingale structure. Using right-continuous covariates (e.g., a CDS spread observed *at* the event time) would introduce anticipation.

6.2 Gamma Frailty and Unobserved Heterogeneity

Sovereign obligors differ in unobservable ways—institutional quality, political cohesion, reserve management—not captured by any finite covariate vector. Frailty models incorporate this via a latent multiplicative factor.

Definition 6.4 (Frailty Model)**.** Let U_i be i.i.d. non-negative random variables (the *frailties*) independent of $\boldsymbol{Z}_i$, with $\mathbb{E}[U_i] = 1$. The frailty model specifies:

$$h_i(t \mid U_i, \boldsymbol{Z}_i) \;=\; U_i \cdot h_0(t)\exp(\boldsymbol{\beta}^\top \boldsymbol{Z}_i). \tag{40}$$

Theorem 6.5 (Gamma Frailty Marginalisation)**.** *Suppose $U_i \sim \mathrm{Gamma}(\theta^{-1}, \theta)$ (mean 1, variance θ). Marginalising over U_i, the population survival function is:*

$$S_{\mathrm{pop}}(t \mid \boldsymbol{Z}_i) \;=\; \big(1 + \theta\, e^{\boldsymbol{\beta}^\top \boldsymbol{Z}_i} H_0(t)\big)^{-1/\theta}. \tag{41}$$

As $\theta \to 0$, (41) *recovers the Cox model: $S_{\mathrm{pop}}(t \mid \boldsymbol{Z}_i) \to \exp(-e^{\boldsymbol{\beta}^\top \boldsymbol{Z}_i} H_0(t))$.*

Proof. The conditional survival function given $U_i = u$ is $S(t \mid u, \boldsymbol{Z}_i) = \exp(-u\, e^{\boldsymbol{\beta}^\top \boldsymbol{Z}_i} H_0(t))$. Marginalising over $U_i \sim \text{Gamma}(1/\theta, \theta)$ with density $p(u) = u^{1/\theta - 1} e^{-u/\theta} / (\theta^{1/\theta} \Gamma(1/\theta))$:

$$\begin{aligned} S_{\text{pop}}(t \mid \boldsymbol{Z}_i) &= \int_0^\infty e^{-u\, e^{\boldsymbol{\beta}^\top \boldsymbol{Z}_i} H_0(t)} p(u)\, \mathrm{d}u \\ &= \left(1 + \theta\, e^{\boldsymbol{\beta}^\top \boldsymbol{Z}_i} H_0(t)\right)^{-1/\theta}, \end{aligned}$$

where the last step uses the moment generating function of the Gamma: $\mathbb{E}[e^{-sU}] = (1+\theta s)^{-1/\theta}$ for $s > 0$.

For the limit, write $(1+\theta x)^{-1/\theta} = \exp(-\theta^{-1} \ln(1+\theta x))$. As $\theta \to 0$, $\theta^{-1} \ln(1+\theta x) \to x$, so the expression $\to e^{-x}$, recovering the Cox survival function. □ □

Corollary 6.6 (Frailty-Induced Positive Dependence)**.** *Within a group sharing a common frailty U, the marginal hazard* decreases *over time relative to the conditional hazard:*

$$h_{\text{pop}}(t \mid \boldsymbol{Z}_i) \;<\; h_0(t) e^{\boldsymbol{\beta}^\top \boldsymbol{Z}_i} \quad \textit{for all } t > 0,\ \theta > 0. \tag{42}$$

Proof. Differentiate $-\ln S_{\text{pop}}(t) = \theta^{-1} \ln(1 + \theta e^{\boldsymbol{\beta}^\top \boldsymbol{Z}_i} H_0(t))$:

$$h_{\text{pop}}(t) = \frac{e^{\boldsymbol{\beta}^\top \boldsymbol{Z}_i} h_0(t)}{1 + \theta\, e^{\boldsymbol{\beta}^\top \boldsymbol{Z}_i} H_0(t)} < e^{\boldsymbol{\beta}^\top \boldsymbol{Z}_i} h_0(t),$$

since $1 + \theta e^{\boldsymbol{\beta}^\top \boldsymbol{Z}_i} H_0(t) > 1$ for $t > 0$. □ □

Remark 6.7. Corollary 6.6 has an important sovereign interpretation: at the population level, the observed hazard rate *appears* to decline even if the individual conditional hazard does not. This is because high-frailty (fragile) sovereigns default early and are removed from the risk set, leaving a progressively more robust surviving cohort. Ignoring frailty causes the Cox model to underestimate the hazard for individual high-risk sovereigns while overestimating it for the population.

6.3 Competing Risks and the Sub-Distribution Hazard

A sovereign may exit the risk set via two distinct mechanisms: outright default (credit event type A) and coercive restructuring (type B), which carry different loss given default values. This constitutes a *competing risks* problem.

Definition 6.8 (Cause-Specific Hazard)**.** Let $\epsilon_i \in \{1, \dots, K\}$ denote the event type. The *cause-specific hazard* for type k is:

$$h^{(k)}(t) \;=\; \lim_{\delta \to 0^+} \frac{\mathbb{P}(t \le \tau < t + \delta,\ \epsilon = k \mid \tau \ge t)}{\delta}. \tag{43}$$

The overall hazard is $h(t) = \sum_k h^{(k)}(t)$, and the overall survival function is $S(t) = \exp(-\sum_k H^{(k)}(t))$.

Definition 6.9 (Sub-Distribution Hazard (Fine–Gray))**.** The *sub-distribution hazard* for cause k is:

$$\tilde{h}^{(k)}(t) \;=\; \lim_{\delta \to 0^+} \frac{\mathbb{P}(t \le \tau < t + \delta,\ \epsilon = k \mid \tau \ge t \text{ or } (\tau < t, \epsilon \ne k))}{\delta}. \tag{44}$$

The corresponding *cumulative incidence function* (CIF) is:

$$F^{(k)}(t) \;=\; \mathbb{P}(\tau \le t, \epsilon = k) \;=\; 1 - \exp\left(-\tilde{H}^{(k)}(t)\right), \tag{45}$$

where $\tilde{H}^{(k)}(t) = \int_0^t \tilde{h}^{(k)}(u)\, \mathrm{d}u$.

Proposition 6.10 (CIF–Cause-Specific Hazard Relationship)**.** *The cause-k cumulative incidence function satisfies:*

$$F^{(k)}(t) \;=\; \int_0^t S(u^-)\, h^{(k)}(u)\, \mathrm{d}u, \tag{46}$$

and hence $\sum_{k=1}^{K} F^{(k)}(t) = 1 - S(t)$*.*

Proof. By definition, $F^{(k)}(t) = \mathbb{P}(\tau \le t, \epsilon = k)$. Conditioning on the event time and type:

$$\mathbb{P}(\tau \in \mathrm{d}u, \epsilon = k) \;=\; h^{(k)}(u) S(u^-)\, \mathrm{d}u,$$

which integrates to (46). Summing over k: $\sum_k F^{(k)}(t) = \int_0^t S(u^-) h(u)\, \mathrm{d}u = 1 - S(t)$ by the fundamental identity (Theorem 2.5). □ □

Remark 6.11. The cause-specific hazard $h^{(k)}$ and the sub-distribution hazard $\tilde{h}^{(k)}$ give rise to different regression targets. The Cox model for $h^{(k)}$ addresses *aetiology* (what drives risk of type-k events?); the Fine–Gray model for $\tilde{h}^{(k)}$ addresses *prognosis* (what is the probability of eventually experiencing type-k before any other event?). For sovereign risk, the Fine–Gray model is more appropriate when pricing loss given event, since the expected loss depends on the probability of reaching each event type, not merely its cause-specific hazard.

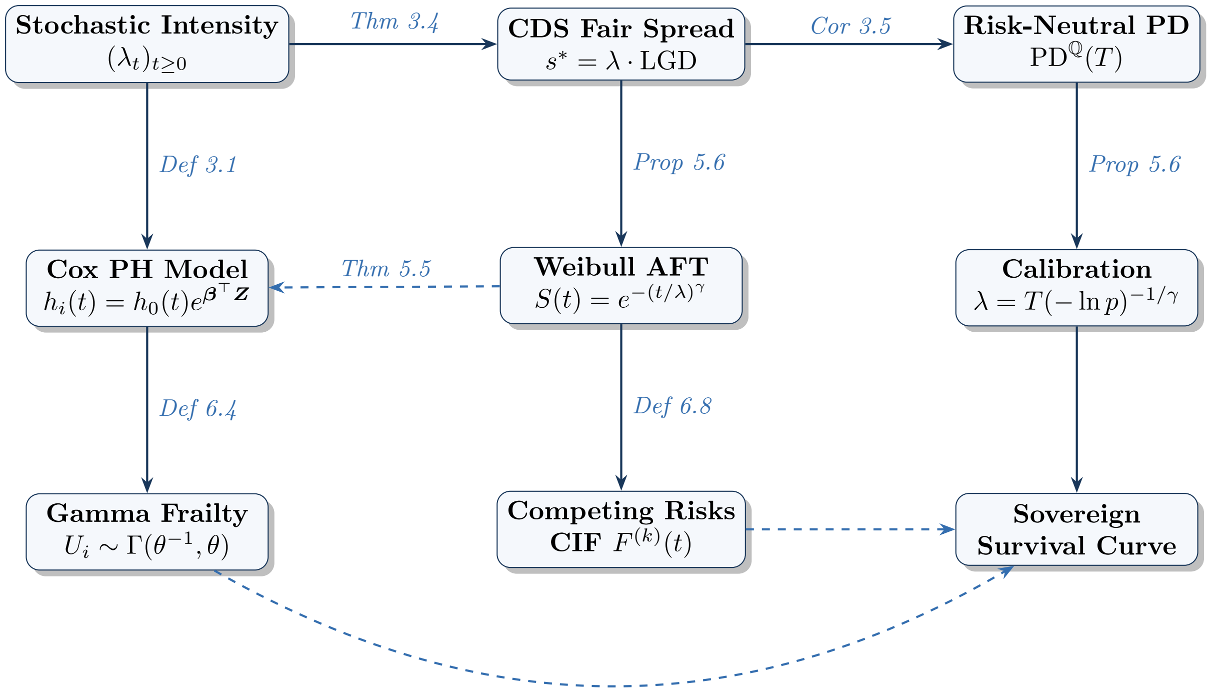

Figure 4: Theoretical dependency graph of the framework.

Solid arrows denote direct derivation; dashed arrows indicate model equivalence or joint contribution to the final output. The Weibull AFT and Cox PH models are linked by Theorem 5.5 (unique intersection of the two families). All paths converge to the sovereign survival curve, which aggregates intensity-model pricing, parametric estimation, and competing-risks adjustment.

7 Theoretical Limitations for Sovereign Obligors

7.1 The Risk-Neutral vs. Real-World Distinction

All quantities derived in Sections 3–4 that flow from CDS prices are risk-neutral. We now characterise the gap between $\mathbb{Q}$ and the real-world measure $\mathbb{P}$ precisely.

Proposition 7.1 (Measure Change and Credit Spread Decomposition)**.** *Suppose the Radon–Nikodým derivative* $\mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P}\big|_{\mathcal{G}_t}$ *is a strictly positive* $\mathbb{P}$*-martingale. Let* $\xi > 0$ *be the market price of default risk. Then the default intensity transforms as:*

$$\lambda_t^{\mathbb{Q}} = \xi_t \cdot \lambda_t^{\mathbb{P}}, \tag{47}$$

and the CDS spread decomposes as:

$$s^* = \underbrace{\lambda^{\mathbb{P}} \cdot \mathrm{LGD}}_{\textit{actuarial premium}} + \underbrace{(\xi - 1)\lambda^{\mathbb{P}} \cdot \mathrm{LGD}}_{\textit{risk premium}}. \tag{48}$$

Proof. Under the measure change, for any $\mathbb{P}$-martingale $M^{\mathbb{P}}$, the Girsanov–Jacod theorem for point processes gives the $\mathbb{Q}$-intensity as $\lambda_t^{\mathbb{Q}} = \xi_t \lambda_t^{\mathbb{P}}$, where ξ_t is the $\mathbb{P}$-predictable process that weights the default likelihood under $\mathbb{Q}$. Substituting into $s^* = \lambda^{\mathbb{Q}}\mathrm{LGD}$ yields (48). □ □

Corollary 7.2 (Systematic Overestimation of Real-World PD)**.** *Since* $\xi_t > 1$ *(investors are risk-averse and demand a premium), CDS-implied default probabilities systematically overstate actuarial default frequencies:*

$$\mathrm{PD}^{\mathbb{Q}}(T) > \mathrm{PD}^{\mathbb{P}}(T) \quad \textit{for all } T > 0. \tag{49}$$

Proof. From (47), $\lambda^{\mathbb{Q}} > \lambda^{\mathbb{P}}$ when $\xi > 1$. Since $\mathrm{PD}(T) = 1 - e^{-\lambda T}$ is strictly increasing in λ, (49) follows immediately. □ □

7.2 Non-Identifiability of State Dissolution

The survival model developed here targets *credit events* as defined by ISDA [8]. We now formalise the sense in which state dissolution is a distinct, non-identified phenomenon.

Proposition 7.3 (Non-Identification)**.** *Let* τ_C *denote the credit-event time and* τ_D *the state-dissolution time. From CDS data alone, the joint distribution* (τ_C, τ_D) *is* not identified*: for any bivariate distribution* $F(\tau_C, \tau_D)$ *consistent with the observed marginal of* τ_C*, there exists a continuum of distributions* $\tilde{F}(\tau_C, \tau_D)$ *with the same* τ_C*-marginal but different* τ_D*-marginal.*

Proof. CDS prices identify only the $\mathbb{Q}$-distribution of τ_C. The joint distribution of (τ_C, τ_D) requires knowledge of the copula $C(u, v)$ linking the marginals. Since the τ_D-marginal is unobservable from CDS data, the copula is unrestricted. For any candidate copula C and any marginal G for τ_D with support on $[0, \infty)$, the bivariate distribution $F(s, t) = C(F_{\tau_C}(s), G(t))$ is consistent with the observed CDS prices. Hence the joint distribution is not identified. □ □

Remark 7.4. Proposition 7.3 has a sharp practical corollary: it is *impossible in principle*, not merely difficult in practice, to infer the probability of state dissolution from CDS data alone, regardless of the sophistication of the survival model or the length of the CDS time-series. Additional data sources—political risk indices, conflict event databases, institutional quality measures—are *logically required* to identify the τ_D-marginal, and even then only the copula between fiscal and political fragility remains unidentified without further structural assumptions.

Table 2: Structural properties of the main model classes developed in this paper.

✓ denotes the property holds; ∘ denotes partial fulfilment; × denotes absence. "MLE equiv." = partial likelihood is asymptotically equivalent to full MLE in the semiparametric sense.

Property	**Const. Intensity**	**CIR Intensity**	**Cox PH**	**Weibull AFT**	**Gamma Frailty**
$\lambda_t \geq 0$ a.s.	✓	✓	∘	✓	✓
Closed-form $S(t)$	✓	✓	×	✓	✓
Semiparametric	×	×	✓	×	×
MLE equiv.	✓	✓	✓	✓	✓
Handles censoring	✓	✓	✓	✓	✓
Time-varying covariates	✓	✓	✓	×	∘
Non-parametric baseline	×	×	✓	×	×
Frailty / heterogeneity	×	×	×	×	✓
Competing risks	∘	∘	∘	×	×
Calibrated from market prices	✓	✓	×	✓	✓

8 Conclusion

We have developed a self-contained theoretical framework for analysing sovereign credit risk through the lens of survival analysis. The main contributions are as follows.

Measure-theoretic foundations. Theorem 2.5 establishes the bijective correspondence between the survival function, hazard function, and cumulative hazard; Proposition 3.2 embeds the credit-event time within the Doob–Meyer decomposition via a doubly stochastic intensity.

CDS pricing theory. Theorem 3.4 derives the exact spread-intensity identity under deterministic intensity; Proposition 3.7 extends this to the affine CIR class via a Riccati ODE system; Proposition 7.1 decomposes the observable spread into actuarial and risk-premium components.

Cox model asymptotics. Propositions 4.4 and 4.6–4.7 establish the derivation, consistency, asymptotic normality, and semiparametric efficiency (Corollary 4.8) of the partial likelihood estimator; Theorem 4.11 proves uniform consistency of the Breslow baseline hazard estimator.

Weibull AFT theory. Lemma 5.3 characterises hazard monotonicity, log-concavity, and memorylessness; Theorem 5.5 proves the Weibull family is the unique intersection of the PH and AFT classes; Proposition 5.6 derives the calibration identity linking the scale parameter to any target survival probability.

Extensions. Theorem 6.5 marginalises the Gamma frailty model to produce a closed-form population survival function; Corollary 6.6 characterises frailty-induced hazard reduction; Proposition 6.10 relates the cumulative incidence function to cause-specific hazards under competing risks.

Fundamental limitation. Proposition 7.3 establishes the logical non-identifiability of state dissolution from CDS data: no model of any sophistication can recover the joint distribution (τ_C, τ_D) from the CDS-implied marginal of τ_C alone. This is a statement about identification, not estimation, and applies regardless of sample size.

The framework presented here provides the rigorous theoretical substrate for any empirical deployment that seeks to rank or quantify sovereign survival probabilities from market-observable CDS prices.

References

[1] Andersen, P. K., Borgan, Ø., Gill, R. D., & Keiding, N. (1993). *Statistical Models Based on Counting Processes*. Springer-Verlag, New York.

[2] Bickel, P. J., Klaassen, C. A. J., Ritov, Y., & Wellner, J. A. (1993). *Efficient and Adaptive Estimation for Semiparametric Models*. Johns Hopkins University Press, Baltimore.

[3] Bielecki, T. R., & Rutkowski, M. (2002). *Credit Risk: Modelling, Valuation and Hedging*. Springer Finance, Berlin.

[4] Calvo, G. A. (1998). Capital flows and capital-market crises: The simple economics of sudden stops. *Journal of Applied Economics*, 1(1), 35–54.

[5] Cox, D. R. (1972). Regression models and life-tables (with discussion). *Journal of the Royal Statistical Society, Series B*, 34(2), 187–220.

[6] Cox, D. R. (1975). Partial likelihood. *Biometrika*, 62(2), 269–276.

[7] Fine, J. P., & Gray, R. J. (1999). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association*, 94(446), 496–509.

[8] International Swaps and Derivatives Association. (2014). *2014 ISDA Credit Derivatives Definitions*. ISDA Publications, New York.

[9] Kalbfleisch, J. D., & Prentice, R. L. (2002). *The Statistical Analysis of Failure Time Data* (2nd ed.). Wiley-Interscience, Hoboken.

[10] Kaminsky, G. L., & Reinhart, C. M. (1999). The twin crises: The causes of banking and balance-of-payments problems. *American Economic Review*, 89(3), 473–500.

[11] Lando, D. (1998). On Cox processes and credit-risky securities. *Review of Derivatives Research*, 2(2–3), 99–120.

[12] Longstaff, F. A., Pan, J., Pedersen, L. H., & Singleton, K. J. (2011). How sovereign is sovereign credit risk? *American Economic Journal: Macroeconomics*, 3(2), 75–103.

[13] Reinhart, C. M., & Rogoff, K. S. (2009). *This Time Is Different: Eight Centuries of Financial Folly*. Princeton University Press, Princeton.

[14] Vaupel, J. W., Manton, K. G., & Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, 16(3), 439–454.

[15] Wellner, J. A. (1983). Limit theorems for the ratio of the empirical distribution function to the true distribution function. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 45(1), 73–88.

Glossary

AFT model

Accelerated Failure Time model. A survival regression model in which covariates multiply the time scale: $T \overset{d}{=} T_0 \cdot e^{-\boldsymbol{\beta}^\top \boldsymbol{Z}}$. Covariates either accelerate (shorten) or decelerate (lengthen) expected event time. The Weibull is the unique parametric family satisfying both AFT and PH structure (Theorem 5.5).

Affine model
A class of stochastic intensity processes for which the conditional Laplace transform of the integrated intensity is exponential-affine in the current state (Definition 3.6). Encompasses the Vasicek, CIR, and jump-extended CIR models, all admitting closed-form CDS pricing.

Breslow estimator
A non-parametric estimator of the cumulative baseline hazard $H_0(t)$ in the Cox model, defined in (27) and proven uniformly consistent in Theorem 4.11. Enables recovery of individual survival curves after $\hat{\boldsymbol{\beta}}$ is obtained.

CDS
Credit Default Swap. A bilateral derivative paying LGD at default in exchange for a continuous spread. The fair spread $s^* = \lambda \cdot \mathrm{LGD}$ under constant intensity (Theorem 3.4). CDS prices identify the risk-neutral intensity $\lambda^{\mathbb{Q}}$ but not the real-world intensity $\lambda^{\mathbb{P}}$.

CIF *Cumulative Incidence Function.* $F^{(k)}(t) = \mathbb{P}(\tau \leq t, \epsilon = k)$: the probability of experiencing cause-k event by time t, accounting for the competing-risks structure. Related to cause-specific hazards via Proposition 6.10.

CIR process
Cox–Ingersoll–Ross process. A mean-reverting square-root diffusion $\mathrm{d}\lambda_t = \kappa(\theta - \lambda_t)\,\mathrm{d}t + \sigma\sqrt{\lambda_t}\,\mathrm{d}W_t$ preserving non-negativity when $2\kappa\theta > \sigma^2$. Its affine structure yields closed-form CDS pricing via Riccati ODEs (Proposition 3.7).

Compensator
The unique predictable non-decreasing process Λ_t in the Doob–Meyer decomposition $\tilde{N}_t = M_t + \Lambda_t$ of the default indicator $\tilde{N}_t = \mathbf{1}_{\{\tau \leq t\}}$. Under the doubly stochastic model, $\Lambda_t = \int_0^{t\wedge\tau} \lambda_u\,\mathrm{d}u$ (Proposition 3.2).

Cox process
A doubly stochastic Poisson process in which the intensity λ_t is itself random (a stochastic process). The default time τ is the first jump of a Cox process; conditional on λ, τ is exponentially distributed (Definition 3.1).

Cox PH model
Cox Proportional Hazards model. The semiparametric regression model $h_i(t \mid \boldsymbol{Z}_i) = h_0(t)e^{\boldsymbol{\beta}^\top \boldsymbol{Z}_i}$ with unspecified baseline h_0. Parameters estimated by partial likelihood (Definition 4.3), which is asymptotically efficient (Corollary 4.8).

Cumulative hazard
$H(t) = \int_0^t h(u)\,\mathrm{d}u = -\ln S(t)$. A non-decreasing function on $[0, \infty)$ in bijection with the survival function via Theorem 2.5.

Doob–Meyer decomposition
The unique decomposition of a submartingale X into a martingale M and a predictable non-decreasing process A: $X = M + A$. Applied to $\mathbf{1}_{\{\tau \leq t\}}$, A is the compensator of the default indicator.

Feller condition
The parameter constraint $2\kappa\theta > \sigma^2$ ensuring that a CIR process λ_t remains strictly positive almost surely, thereby guaranteeing a well-defined intensity.

Frailty
A latent non-negative random variable U_i multiplying the hazard rate, representing unobserved heterogeneity. Gamma-distributed frailty yields a closed-form marginal survival function (Theorem 6.5) and induces positive dependence within risk groups.

Hazard function

$h(t) = -(\mathrm{d}/\mathrm{d}t) \ln S(t) = f(t)/S(t)$: the instantaneous rate of the credit event at time t, conditional on survival to t. Central object of survival analysis; determines S via $S(t) = e^{-H(t)}$.

IFR / DFR

Increasing / Decreasing Failure Rate. A survival distribution has IFR (resp. DFR) if its hazard $h(t)$ is non-decreasing (resp. non-increasing). The Weibull has IFR iff $\gamma > 1$ (Lemma 5.3a).

Immersion

Also called the $(\mathcal{H})$-hypothesis. The condition that $\mathbb{P}(\tau > t \mid \mathcal{G}_\infty) = \mathbb{P}(\tau > t \mid \mathcal{G}_t)$: the background filtration contains no forward-looking information about the default time (Assumption 2.1).

LGD

Loss Given Default. The fraction $1 - \mathrm{RR}$ of notional principal lost at the credit event, where RR is the recovery rate. Enters the CDS fair-spread formula as the proportionality constant between intensity and spread.

Partial likelihood

A product of conditional event-time probabilities that eliminates the nuisance baseline hazard h_0 while retaining sufficient information for consistent, efficient estimation of $\boldsymbol{\beta}$ (Propositions 4.4 and 4.7, Corollary 4.8).

PH model

Proportional Hazards model. Any model in which the ratio of hazard rates between two subjects is constant over time: $h_i(t)/h_j(t) = \phi(\boldsymbol{Z}_i, \boldsymbol{Z}_j)$ for a time-independent function ϕ. The Cox model is the canonical semiparametric PH model.

Risk set

$\mathcal{R}(t) = \{i : Y_i \geq t\}$: the set of obligors still under observation (neither defaulted nor censored) at time t. Appears in the denominator of the partial likelihood.

Riccati ODE

A nonlinear first-order ODE of the form $y' = a + by + cy^2$. Arises in the derivation of affine-model bond and CDS prices (Proposition 3.7); often solvable in closed form via substitution or Bernoulli reduction.

Sub-distribution hazard

$\tilde{h}^{(k)}(t)$: the hazard of the cause-k cumulative incidence function, treating individuals who experienced a competing event as still at risk (Definition 6.9). Used in the Fine–Gray model for regression on cumulative incidence directly.

Survival function

$S(t) = \mathbb{P}(\tau > t)$: the probability of no credit event through time t. Right-continuous, non-increasing, with $S(0) = 1$ and $S(\infty) = 0$ (Definition 2.2).

Weibull distribution

A two-parameter family Weibull(λ, γ) with survival $S(t) = e^{-(t/\lambda)^\gamma}$, the unique parametric family satisfying both PH and AFT structures (Theorem 5.5). Admits IFR, DFR, and exponential special cases controlled by γ.

The End