Appendix

In survival analysis, three main functions are defined, the probability density function of time (i.e., $f(t)$) which is the density of failure at time t, the survival function (i.e., $S(t)$) which provides the probability of surviving (not failing) until time t, the probability of failing at time t conditional on surviving until time t can be shown by a hazard function (i.e., $h(t)$).

$$f(t) = lim\,\Delta t \rightarrow 0(Pr(t \leq T \leq t + \Delta t)\,\Delta t)$$

$$F(t) = Pr(T \leq t)$$

$$S(t) = Pr(T \geq t) = 1 - F(t)$$

$$h(t) = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{S(t)}$$

In survival analysis, the estimation of model parameters is generally accomplished through the utilization of maximum likelihood estimation (MLE) technique. The likelihood function in survival analysis incorporates both uncensored and censored observations. Uncensored observations are represented by $f(t)$, while censored observations are represented by $S(t)$ in the below likelihood function. In this equation, δi takes the value of 1 if individual i has experienced an event or relocation, and 0 otherwise. N denotes the total number of individuals. By transforming the likelihood function into its logarithmic form, we obtain the next equation.

$$L = \prod_{i=1}^{N} f(T_i)^{\delta_i}\, S(T_i)^{1-\delta_i}$$

$$LL = \sum_{i=1}^{N} L_i = \sum_{i=1}^{N} [\delta_i logh(T_i) + logS(T_i)]$$

Accelerated failure time model (AFT)

The Accelerated Failure Time (AFT) model is a commonly employed method in survival analysis used to examine the timing of events. The AFT model directly focuses on modeling the actual event time without considering the hazard rate. It assumes that the event time can be accelerated or decelerated based on the covariates or predictors.

$$\ln(T) = \beta X + z$$

$$T = exp(\beta X) expz$$

By estimating the parameters of the AFT model, we can gain insights into how these covariates influence the timing of the event of interest, such as residential relocation duration. This model provides an alternative perspective for understanding the underlying dynamics and factors affecting event occurrences. In the context of Accelerated Failure Time (AFT) modeling, the inclusion of covariates in the model is achieved by exponentiating the linear combination of covariates (X) and their respective coefficients (β) ($\psi = exp(-\beta X)$). The formulations for incorporating ψ into the survival function is as $S(t, X) = S_0(t\psi)$, where $S_0(t)$ is the baseline function which is formulated based on the chosen parametric model. In our study, we utilize the Weibull and Log-normal distributions to establish these baseline functions.

Cox proportional hazard model (Cox-PH)

The Cox proportional hazards model is a semi-parametric model that assumes the hazard rate is proportional to a baseline hazard function. It can be written as:

$$h(t|X) = h_0(t) \times exp(\beta X)$$

where $h(t|X)$ represents the hazard function at time t for a given set of covariates X, $h_0(t)$ is the baseline hazard function, β is the vector of regression coefficients, and $exp(\beta X)$ represents the hazard ratio associated with the covariates. The estimation of parameters involves maximizing the partial log-likelihood function presented below.

$$LL(\beta) = \sum_{i=1}^{N} \delta i \left(X\beta - \log\left(\sum_{k:t_k \geq t_i} exp\,(Xk\beta) \right) \right)$$