UNIVERSIDAD CARLOS III DE MADRID working papers

UC3M Working papers
Economics
18-12
December, 2018
ISSN 2340-5031

Departamento de Economía
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 916249875

Counterfactual Analysis Using Censored Duration Data

Miguel A. Delgado[1] & Andrés F. García-Suaza[2]

Abstract

We propose standardization techniques for the duration distribution in a population with respect to another taken as standard using right censored data, which forms a basis for counterfactual comparisons between distributional features of interest. Alternative standardizations are based on either a proportional hazard semiparametric specification or a nonparametric specification of the underlying conditional distribution. Applications to the restricted mean survival time and the hazard rate are discussed in detail. The proposal is applied to the counterfactual analysis of spells of unemployment duration gender gaps in Spain between 2004-2007. The behavior in small samples is investigated using Monte Carlo experiments.

Keywords: Survival time; Right Censoring; Standardization; Proportional hazard; Kaplan-Meier; RMST; Spells of unemployment; Gender gaps.

JEL Codes: C14, C24, C41, J64.

[1] Universidad Carlos III de Madrid. Email: delgado@est.econ.uc3m.es. Corresponding Author. Research funded by Ministerio Economía y Competitividad (Spain), grant numbers ECO2017-86675-P and MDM 2014-0431, Comunidad de Madrid (Spain), and MadEco-CM grant number S2015/HUM-3444.

[2] Universidad EIA. Email: andres.garcia58@eia.edu.co

I. Introduction

When comparing some phenomena between two populations, the different populations composition, i.e. the distribution of characteristics causally related to the phenomenon, must be taken into account in order to avoid misleading conclusions. Standardization techniques are commonly used for adjusting this confounding effect and providing meaningful comparisons. These techniques, introduced by Neison (1844), known as direct standardization, assume a particular population as standard and recomputes the marginal distribution of the population of interest plugging-in the component distribution of the standard population into the joint distributions of the target variable and components, i.e. covariates, giving by granted that an estimator of the conditional distribution is available.

This methodology was formerly applied in demography to standardizing "crude" mortality rates using contingency tables in order to compare populations with different age distributions. Kitawaga (1955) is credited as the first formalization of standardization methods and proposed decomposing the total difference between crude rates into an effect related to population composition and a remaining structure effect (residual effect in Kitagawa's terms). Each effect, which is in fact the difference between an observed rate and its corresponding standardized rate, has a counterfactual interpretation. The composition effect is the difference that would have been observed if the populations had only differed in the components, while the structural effect is the difference that would have been observed if the components distribution in the two populations had been identical. For further discussion of standardization and decomposition in the context of mortality rates see Kigatawa (1964), DasGupta (1993) and Canudas-Romo (1993).

Estimators of the cumulative distribution function (CDF) of the variable of interest and components used in the standardizations are based on some specification of the underlying conditional distribution function (CCDF), which can be parametric, semiparametric or nonparametric. Standardizations of many distributional features does not need estimating the CCDF, e.g. a regression function estimate is enough for mean standardization. In fact, these techniques became popular in economics from Oaxaca (1973) and Blinder (1973), OB henceforth, who proposed a standardization of the mean based on ordinary least squares (OLS) fitted values, which in fact assumes an underlying linear regression model. When the regression function is nonlinear, OB method can provide misleading counterfactual decompositions, but it can also be applied using flexible regression specifications, e.g. polynomial with a fixed degree, semiparametric or pure nonparametric regression models. For decompositions based on nonparametric specifications see Kitawaga (1955, 1964), Stock (1989) or Rothe (2010). Kitawaga's method applied to mean decomposition using grouped data can be interpreted as based on regressogram estimates of the underlying nonparametric regression using fix bindwidths, while Stock's (1989) method uses Nadaraya-Watson kernel regression estimators with a suitable bandwidth choice. Rothe (2010) uses kernel estimates of the CCDF for CDF standardization. Machado and Mata (2005) and Chernozhukov, Fernández-Val and Melly (2013) considered semiparametric specifications of the conditional distribution, like quantile or distributional regression model CCDF specification.

Fortin et al. (2011) review applications of standardization and decomposition methods in economics. The most numerous applications study gender wage gaps, e.g Oaxaca, (1973), Blinder (1973), Cain (1986), Blau and Kahn (1992), Oaxaca and Ransom (1999), Machado and Mata (2005) or O'Neill and O'Neill (2006). There are also studies on other sources of wage differentials, like racial (Reimers, 1983; Melly, 2005), unions (Freeman 1980, 1984), returns to skills (Juhn et al., 1993), between countries (Donald et al., 2000), or immigrant/resident (Chiquiar and Hanson, 2005). See Lemieux (2002) for a comprehensive review on decomposing changes in wage distributions. There are also applications in other contexts, like gender differences in smoking behavior (Bauer 2007) or housing prices in terms of cleaning up a hazardious-waste site (Stock 1989).

In duration analysis, the hazard function (HF), which completely characterizes the

CDF, is most informative and often directly specified. The popular Cox's proportional hazard (PH) specification of the conditional hazard function (CHF) has been used to studying duration of unemployment spells by comparing the averaged CHF (ACHF), which is the integrated CHF with respect to the distribution of the relevant population components. See Ham et al. (1999), Gonzalo and Saarela (2000), Du and Dong (2008), Tansel and Tasci (2010), Baussola et al. (2015) or De la Rica and Rebollo-Sanz (2017). However, the ACHF and HF shapes are typically unrelated, and inferences based on the ACHF can be misleading. At the best of our knowledge, there are no counterfactual studies based on direct standardization of the HF, or other CDF's relevant features, using right censored data.

Since consistent estimation of the duration mean is not possible under censoring, because there is no information beyond the boundary of the censoring variable, we propose comparing the restricted mean survival time (RMST) that can always be consistently estimated for a suitable boundary restriction. We consider two specification settings, from which we propose alternative standardization and decomposition techniques. On one hand, under a semiparametric PH specification of the CHF, which results in a semiparametric specification of the underlying CCDF. On the other hand, under a pure nonparametric specification of the joint distribution of survival time and component variables that is consistently estimated buy the Kaplan and Meier (1958) (KM) method. This allows standardizing the RMST both applying the OB methodology and using a pure nonparametric estimator of the underlying regression function.

The proposal is applied to study spells of unemployment duration gender gaps in Spain during the period 2004-2007 using data from the European Survey on Income and Living Conditions (SILC). Our findings suggest that the distributional components play a minor role explaining gender gaps in spells of unemployment duration. We also show the substantial difference between the HF and the ACHF shapes for this data set.

The rest of this article is organized as follows. Next section introduces the main notation, standardization concepts and corresponding counterfactual decompositions. Section III and IV present estimators based on PH and nonparametric specifications of the underlying CCDF, respectively. The finite sample performance is studied by means of Monte Carlo experiments in section V. Section VI applies the proposed methodologies to analyze unemployment duration gender gaps in Spain. Conclusions and final remarks are in the last section.

II. Standardization and Counterfactual Decompositions in Duration Analysis

The challenge consists of standardizing duration distribution features of two populations, 0 and 1 say, when durations $T^{(j)}$ are observed under right censoring according to a variable $C^{(j)}$, using the information given by a $k \times 1$ vector of population components $X^{(j)}$, $j = 0, 1$. That is, inferences are based on an observed random vector $\left(Z^{(j)}, \delta^{(j)}, X^{(j)}\right)$, where $Z^{(j)} = \min\left(T^{(j)}, C^{(j)}\right)$ and $\delta^{(j)} = 1_{\left\{T^{(j)} \leq C^{(j)}\right\}}$ indicates whether or not the observed duration is censored with $1_{\{A\}}$ the indicator function of the event A.

Denote by $\mathbb{P}^{(j)}$ the induced probability of $\left(T^{(j)}, C^{(j)}, X^{(j)}\right)$, $F^{(j)}_{T,X}\left(T^{(j)}, X^{(j)}\right) = \mathbb{P}^{(j)}\left(T^{(j)} \leq t, X^{(j)} \leq x\right)$ is the joint CDF of $\left(T^{(j)}, X^{(j)}\right)$, inequalities are coordinatewise and $F^{(j)}_{T|X}\left(\cdot|\, X^{(j)}\right)$ is the CCDF of $T^{(j)}$ given $X^{(j)}$, *i.e.*

$$F^{(j)}(t, x) = \mathbb{E}\left[F^{(j)}_{T|X}\left(t|\, X^{(j)}\right) \cdot 1_{\left\{X^{(j)} \leq x\right\}}\right] = \int_{\{\bar{x} \leq x\}} F^{(j)}_{T|X}\left(t|\, \bar{x}\right) F^{(j)}_X(d\bar{x})$$

for all $(x, y) \in \mathbb{R}^+ \times \mathbb{R}^k$ and $j = 0, 1$, where, henceforth, $\mathbb{E}$ is the expectation operator applied both to $\mathbb{P}^{(0)}$ and $\mathbb{P}^{(1)}$, and for any random variable or random vector $\xi^{(j)}$, which can be $T^{(j)}$, $C^{(j)}$, $Z^{(j)}$ or $X^{(j)}$, and $F^{(j)}_\xi(w) = \mathbb{P}^{(j)}\left(\xi^{(j)} \leq w\right)$ is its marginal distribution. The standardized distribution of $F^{(j,s)}_T$, using population s as standard, is

$$F^{(j,s)}_T(t) = \mathbb{E}\left[F^{(j)}_{T|X}\left(t|\, X^{(s)}\right)\right] = \int_{\mathbb{R}^k} F^{(j)}_{T|X}\left(t|\, x\right) F^{(s)}_X(dx). \tag{1}$$

Therefore, $F^{(j,j)}_T(\cdot) \equiv F^{(j)}_T(\cdot) \equiv F^{(j)}(\cdot, \infty)$, using the notation $\infty = (\infty, ..., \infty)^{\mathrm{T}}$, where "T" means transpose. That is, $F^{(j,s)}_T$ represents the distribution of the duration outcome in population j that one would have observed if their population components had been distributed as in population s. Henceforth, we give by granted that integrals in (1) are well defined, which requires that the support of the components $X^{(s)}$ in the standard population is contained in the support of the standardized population. It is also possible to use the pooled distribution as standard, i.e. $\mathbb{P}^{(p)} = \lambda\mathbb{P}^{(0)} + (1-\lambda)\,\mathbb{P}^{(1)}$, where $\lambda \in (0,1)$ is the probability of picking up at random an individual of population 0 from the pooled population. In what follows we only consider the case where one of the two populations is the standard.

Any $F_T^{(j)\prime}s$ feature, like the mean, can be standardized. For instance, the standardized mean of population j, using population s as standard, is

$$\mu^{(j,s)} = \int_{\mathbb{R}^+} tF_T^{(j,s)}(dt) = \int_{\mathbb{R}^k} m^{(j)}\,(x)\,F_X^{(s)}(dx) = \int_{\mathbb{R}^+} \left(1 - F_T^{(j,s)}\,(t)\right) dt,$$

where $m^{(j)}\,(x) = \int_{\mathbb{R}^+} tF_{T|X}^{(j)}\,(\,dt|\,x)$ is the conditional expectation of the survival time in population j evaluated at x, and $\mu^{(j,j)} = \mu^{(j)} = \mathbb{E}\left(T^{(j)}\right)$.

Let $\vartheta^{(j)}(x) = a^{(j)} + x^{\mathrm{T}}b^{(j)}$ be the best linear predictor (BLP) of $T^{(j)}$ given $X^{(j)} = x$, i.e. $\left(a^{(j)}, b^{(j)\mathrm{T}}\right)^{\mathrm{T}}$ is a vector of parameters such that $T^{(j)} - \vartheta^{(j)}(X^{(j)})$ has zero mean and is uncorrelated with each component of $X^{(j)}$. OB exploited the fact that , since $\mu^{(j)} = \mathbb{E}\left(\vartheta^{(j)}(X^{(j)})\right)$ and $\mu^{(j)} = \mu^{(j,j)} = \ell^{(j,j)}$, where

$$\ell^{(j,s)} = \int_{\mathbb{R}^k} \vartheta^{(j)}(x)F_X^{(s)}(dx) = a^{(j)} + \mathbb{E}\left(X^{(s)}\right)^{\mathrm{T}} b^{(j)},$$

is the suggested standardized mean. In practical terms, this is equivalent to assume that $m^{(j)}$ is linear, which explains why the standardization can be misleading under non-linearity of the underlying conditional expectation (see Barsky et. al, 2002). In fact, whenever $\beta^{(j)} = 0$, $\ell^{(j,s)} = \mu^{(j)}$ for all $j, s = 0, 1$, i.e. there is no effective standardization. Of course, we can use a more flexible specification of $m^{(j)}$, e.g a polynomial of some degree. For instance, adding squares of the variables components in the least squares fit, i.e.

$$\vartheta^{(j)}(x) = a^{(j)} + \sum_{l=1}^{k} b_l^{(j)} X_l^{(j)} + \sum_{l=1}^{k} b_{l+k}^{(j)} X_l^{(j)^2},$$

where $T^{(j)} - \vartheta^{(j)}(X^{(j)})$ is a zero mean r.v. uncorrelated with $X_l^{(j)}$ and $\left(X_l^{(j)}\right)^2$, $l = 1, ..., k$, but other powers or interactive effects can be introduced. A nonparametric specification of $m^{(j)}$ is a reasonable alternative when there are doubts on the underlying regression functional form.

Henceforth, for any generic random variable $\xi^{(j)}$ in population $j = 0, 1$, $\tau_\xi^{(j)} = \inf\left(t : F_\xi^{(j)}(t) = 1\right)$, is the upper bound of its support. Therefore, consistent estimation of $F_T^{(j)}$ beyond $\tau_Z^{(j)} = \min\left(\tau_C^{(j)}, \tau_T^{(j)}\right)$ is no possible because there is no relevant information on $T^{(j)}$. In fact, $F^{(j)}(t, x)$ cannot be identified from the CDF of the observable random vector $\left(Z^{(j)}, X^{(j)}, \delta^{(j)}\right)$. Hence, consistent estimation of the sub-CDF,

$$\bar{F}^{(j)}(t, x) = \begin{cases} F^{(j)}(t, x) \ \text{ for } t < \tau_Z^{(j)} \\ F^{(j)}\left(\tau_Z^{(j)}-, x\right) + 1_{\left\{\tau_Z^{(j)} \in A^{(j)}\right\}} F^{(j)}\left(\left\{\tau_Z^{(j)}\right\}, x\right) \ \ t \geq \tau_Z^{(j)}, \end{cases}$$

is the best we can hope for, where $A^{(j)} = \left\{z : \mathbb{P}^{(j)}\left(Z^{(j)} = z\right) > 0\right\}$ is the number of jumps in $F_Z^{(j)}$. Notice that $\bar{F}^{(j)}(t, \cdot) = F^{(j)}(t, \cdot)$ for all $t < \tau_Z^{(j)}$, and also for all $t \leq \tau_Z^{(j)}$, when

$$\text{either, } F^{(j)}(\{t\}, \cdot) = 0 \quad \text{or} \quad F^{(j)}(\{t\}, \cdot) > 0 \ \text{ and } \mathbb{P}^{(j)}\left(C^{(j)} \leq \tau_Z^{(j)}\right) < 1.$$

Of course, $\bar{F}^{(j)} = F^{(j)}$ when no censoring is present. Furthermore, assuming $T^{(j)}$ independent of $C^{(j)}$, $\bar{F}^{(j)} = F^{(j)}$ when $\tau_T^{(j)} < \tau_C^{(j)}$, e.g when $\tau_C^{(j)} = \infty$, irrespective of whether or not $\tau_T^{(j)}$ is finite. However, these restrictions cannot be tested using the sample information. If $\tau_T^{(j)} > \tau_C^{(j)}$, $\bar{F}^{(j)} \neq F^{(j)}$ and it is not possible consistently estimating the mean, among other distributional features.

We can always use the RMST, introduced by Irvin (1949), as a relevant summary parameter for comparing durations in two distributions. For any $\tau^* < \tau_Z^{(j)}$, we can consistently estimating the RMST

$$\mu_{\tau^*}^{(j)} = \int_0^{\tau^*} \left(1 - F_T^{(j)}(t)\right) dt = \mathbb{E}\left(\min\left(T^{(j)}, \tau^*\right)\right),$$

which is the average duration in the first τ^* periods. Notice that $\mu_{\tau_Z^{(j)}}^{(j)}$ is the closest possible approximation to $\mu^{(j)} = \mu_\infty^{(j)}$, but other values of τ^* might be of interest. For instance, the average unemployment duration during the first 24 months ($\tau^* = 24$) which is related to short-term unemployment. Karrison (1997), Zhang and Schanbel (2011), Zucker(1998) or Chen and Tsiatis (2001) provide applications of RMST to different contexts.

We can apply the OB method for $\mu_{\tau^*}^{(j)}$ using the standardization $\ell_{\tau^*}^{(j,s)} = \alpha_{\tau^*}^{(j)} + \beta_{\tau^*}^{(j)T}\mathbb{E}\left(X^{(s)}\right)$, for $\tau^* \leq \tau_Z^{(j)}$,

$$\begin{pmatrix} \alpha_{\tau^*}^{(j)} \\ \beta_{\tau^*}^{(j)} \end{pmatrix} = \arg\min_{a,b} \mathbb{E}\left(\min\left(T^{(j)}, \tau^*\right) - a - b^{\mathrm{T}} X^{(j)}\right)^2, \tag{2}$$

i.e $\mu_{\tau^*}^{(j)} = \ell_{\tau^*}^{(j,j)}$.

A very informative distributional feature in survival analysis is the hazard rate. If $T^{(j)}$ is discrete,

$$\begin{aligned} \lambda^{(j)}(t) &= \mathbb{P}\left(T^{(j)} > t+1 \left| T^{(j)} \geq t\right.\right) \\ &= \frac{F_T^{(j)}\{t\}}{1 - F_T^{(j)}(t-)}, \ t = 0, 1, 2, ..., \end{aligned} \tag{3}$$

where for any generic CDF G, $G\{t\} = G(t) - G(t-)$. It $T^{(j)}$ is continuous and $F_T^{(j)}$ is absolutely continuous with pdf $f_T^{(j)}$,

$$\begin{aligned} \lambda^{(j)}(t) &= \lim_{h \to 0} \frac{1}{h} \mathbb{P}\left(T^{(j)} > t+h \left| T^{(j)} \geq t\right.\right) \\ &= \frac{f_T^{(j)}(t)}{1 - F_T^{(j)}(t)}, \ t \in \mathbb{R}^+ \end{aligned} \tag{4}$$

The hazard rate $\lambda^{(j)}(t)$ is interpreted as the instantaneous probability that an individual does not survive longer than a period of length t given that she/he is alive after t periods. However, counterfactual analysis in the existing empirical research is based on comparing the ACHF

$$\gamma^{(j)}(t) = \mathbb{E}\left[\lambda^{(j)}_{T|X}(t|\,X^{(j)})\right],\ t < \tau^{(j)}_Z,$$

where $\lambda^{(j)}_{T|X}(\cdot|\,X^{(j)})$ is the conditional hazard rate of $T^{(j)}$ given $X^{(j)}$. The corresponding standardized ACHF is

$$\gamma^{(j,s)}(t) = \mathbb{E}\left[\lambda^{(j)}_{T|X}(t|\,X^{(s)})\right],\ t < \tau^{(j)}_Z,$$

The HF and ACHF shapes are, in general, unrelated, and so $\lambda^{(j,s)}$ and $\gamma^{(j,s)}$ are. For instance, if $T^{(j)}$ given $X^{(j)}$ is exponentially distributed (memoryless), the CHF $\lambda^{(j)}_{T|X}(t|\,X^{(j)})$ is independent of t a.s., and so do $\gamma^{(j)}$ is, but $\lambda^{(j)}$ may variate with t (see the example in next section.)

A. Standardization Based on the Semiparametric Proportional Hazard Specification

The most popular CHF specification is Cox's (1972) PH model,

$$\lambda(t|\,X^{(j)}) = \lambda^{(j)}_b(t)\cdot \exp\left(X^{(j)\mathrm{T}}\eta^{(j)}\right), \tag{5}$$

where $\lambda^{(j)}_b$ is the nonparametric baseline hazard function, and $\eta^{(j)}$ is a vector of unknown parameters. Tsiatis (1981) shows that $\eta^{(j)}$ and the cumulative baseline hazard $\Lambda^{(j)}_b(t) = \int_0^t \lambda^{(j)}_b(\bar{t})d\bar{t}$ can be identified from the joint CDF of the observed $\left(Z^{(j)}, X^{(j)}, \delta^{(j)}\right)$, for $t < \tau^{(j)}_Z$, assuming that,

A.0. $T^{(j)}$ is independent of $C^{(j)}$ conditionally on $X^{(j)}$, $j = 0, 1$.

Cox (1972) proposed the partial likelihood method to estimate $\eta^{(j)}$, which only uses the non censored observed durations, while Breslow (1974) proposed a consistent estimator of $\Lambda^{(j)}_b(t)$ for $t < \tau^{(j)}_Z$ by a function with jumps at each non-censored observed duration. See Kalbfleisch and Prentice (1973) and Andersen and Gill (1982) for further discussion and Tsiatis (1981) for formal justification of the asymptotic properties of these estimators under assumption A.0. These two estimators provide consistent estimators of $F^{(j)}_{T|X}$ and

$F_T^{(j)}$ when the PH specification is correct, which forms a basis for consistently estimating the standardized distribution

$$F_T^{(j,s)}(t) = 1 - \mathbb{E}\left[\exp\left\{-\Lambda_b^{(j)}(t)\exp\left(X^{(s)\mathrm{T}}\eta^{(j)}\right)\right\}\right],$$

and the corresponding HF

$$\lambda^{(j,s)}(t) = \lambda_b(t)\cdot\frac{\mathbb{E}\left[\exp\left(X^{(s)\mathrm{T}}\eta^{(j)}\right)\left(1 - F_{T|X}^{(j)}(t|\, X^{(s)})\right)\right]}{1 - F_T^{(j,s)}(t)}.$$

Note that, when $\lambda_b(t) = 1$, which corresponds to an exponential CCDF,

$$F_{T|X}^{(j)}\left(t|\, X^{(j)}\right) = 1 - \exp\left(-t\exp\left(X^{(j)\mathrm{T}}\eta^{(j)}\right)\right) \quad a.s.,$$

and $\gamma^{(j,s)}(t) = \mathbb{E}\left(\exp\left(X^{(s)\mathrm{T}}\eta^{(j)}\right)\right)$ is constant for all t. However, in this case,

$$F_T^{(j,s)}(t) = 1 - \mathbb{E}\left[\exp\left(-t\exp\left(X^{(s)\mathrm{T}}\eta^{(j)}\right)\right)\right],$$

and

$$\lambda^{(j,s)}(t) = \frac{\mathbb{E}\left[\exp\left(X^{(s)\mathrm{T}}\eta^{(j)} - t\exp\left(X^{(s)\mathrm{T}}\eta^{(j)}\right)\right)\right]}{\mathbb{E}\left[\exp\left(-t\exp\left(X^{(s)\mathrm{T}}\eta^{(j)}\right)\right)\right]},$$

which varies with t, but the ACHF $\gamma^{(j,s)} = \mathbb{E}\left[\exp\left(X^{(j)\mathrm{T}}\eta^{(j)}\right)\right]$ is constant.

B. Standardization under a Nonparametric Specification.

Identification of the nonparametric $\bar{F}^{(j)}$ requires the following conditions,

A.1 $T^{(j)}$ is independent of $C^{(j)}$, $j = 0, 1$.

A.2 $\delta^{(j)}$ and $X^{(j)}$ are independent conditional on $T^{(j)}$ $a.s.$, $j = 0, 1$.

Condition A.1 is the standard identification condition for the nonparametric $\bar{F}_T^{(j)}$, which justifies consistency of the KM product limit estimator using censored data. Assumption A.2 is the extra condition, provided by Stute (1993), for $\bar{F}^{(j)}$ identification. In Stute (1993) words, "...This is a convenient way to remind you of the uneasy fact that once $T^{(j)}$ is known, things that had been considered of some importance in your life then become irrelevant."

The KM estimator of $\bar{F}^{(j)}$ forms a basis to estimate $\ell_{\tau^*}^{(j,s)}$ in (2) using the OB approach. Moreover, a nonparametric identification of $\bar{F}_T^{(j)}$ can be used from Kigatawa (1955) type of standardization, which is designed for grouped data. Consider a fixed partition of $\mathbb{R}^k$, $\mathbb{A}_m = \{\mathcal{A}_i\}_{i=1}^m$ say, i.e. $\mathcal{A}_i$ are disjoint with $\bigcup_{i=1}^m \mathcal{A}_i = \mathbb{R}^k$. The CDF of $T^{(j)}$ given $\{X^{(j)} \in \mathcal{A}\}$ is

$$F_{T|\mathcal{A}}^{(j)}(t|\mathcal{A}) = \mathbb{P}^{(j)}\left(T^{(j)} \leq t \middle| X^{(j)} \in \mathcal{A}\right) = \frac{p_{\mathcal{A}}^{(j)}(t)}{q_{\mathcal{A}}^{(j)}},$$

with $p_{\mathcal{A}}^{(j)}(t) = \int_{x\in\mathcal{A}} F^{(j)}(t, dx)$ and $q_{\mathcal{A}}^{(j)} = p_{\mathcal{A}}^{(j)}(\infty)$. Thus,

$$F_T^{(j)}(t) = \sum_{i=1}^m F_{T|\mathcal{A}_i}^{(j)}(t|\mathcal{A}_i)\, q_{\mathcal{A}_i}^{(j)} = \sum_{i=1}^m p_{\mathcal{A}_i}^{(j)}(t),$$

which suggests the standardization,

$$F_{T\mathbb{A}_m}^{(j,s)}(t) = \sum_{i=1}^m F_{T|\mathcal{A}}^{(j)}(t|\mathcal{A}_i)\, q_{\mathcal{A}_i}^{(s)} = \sum_{i=1}^m \frac{p_{\mathcal{A}_i}^{(j)}(t)}{q_{\mathcal{A}_i}^{(j)}} q_{\mathcal{A}_i}^{(s)},\ t < \tau_Z^{(j)}. \tag{6}$$

Notice that $F_{T\mathbb{A}_m}^{(j,j)} = F_T^{(j)}$, but $F_{T\mathbb{A}_m}^{(j,s)}$ and $F_T^{(j,s)}$ are different for $j \neq s$. When $X^{(j)}$ is a vector of discrete variables taking a finite number of values, or when data is grouped, the partition is natural.

The standardized RMST using partitions is

$$\mu_{\tau^*\mathbb{A}_m}^{(j,s)} = \int_0^\infty \min(t, \tau^*)\, F_{T\mathbb{A}_m}^{(j,s)}(dt) = \sum_{i=1}^m \frac{q_{\mathcal{A}_i}^{(s)}}{q_{\mathcal{A}_i}^{(j)}} \int_0^\infty \min(t, \tau^*) \cdot p_{\mathcal{A}_i}^{(j)}(dt),$$

and the corresponding standardized HF is

$$\lambda_{\mathbb{A}_m}^{(j,s)}(t) = \frac{F_{T\mathbb{A}_m}^{(j,s)}\{t\}}{1 - F_{T\mathbb{A}_m}^{(j,s)}(t-)},\ t < \tau_Z^{(j)},$$

in the discrete case, and

$$\lambda_{\mathbb{A}_m}^{(j,s)}(t) = \lim_{h\to 0} \frac{1}{h} \frac{F_{T\mathbb{A}_m}^{(j,s)}(t+h) - F_{T\mathbb{A}_m}^{(j,s)}(t)}{1 - F_{T\mathbb{A}_m}^{(j,s)}(t)},\ t < \tau_Z^{(j)},$$

in the continuous case.

C. Counterfactual Decompositions

The total difference of the CDFs between population j and s at t, $\Delta^{(j,s)}(t) = F_T^{(j)}(t) - F_T^{(s)}(t)$, can be written as

$$\begin{aligned} \Delta^{(j,s)}(t) &= \left[F_T^{(j,j)}(t) - F_T^{(j,s)}(t)\right] + \left[F_T^{(j,s)}(t) - F_T^{(s,s)}(t)\right] \\ &= \Delta^{(j,s)C}(t) + \Delta^{(j,s)S}(t),\ t < \tau_Z, \end{aligned} \tag{7}$$

where $\Delta^{(j,s)S}$ is the counterfactual structural effect, $\Delta^{(j,s)C}(t)$ is the counterfactual composition effect and $\tau_Z = \min\left(\tau_Z^{(j)}, \tau_Z^{(s)}\right)$. Likewise, we perform the decomposition using standardizations based on a partition $\mathbb{A}_m$ by substituting $F_T^{(j,s)}$ by $F_{T\mathbb{A}_m}^{(j,s)}$ in (7). We can also decompose any distributional characteristic summarized by a functional $\theta : G \mapsto \mathbb{R}$, where G can be either $F_T^{(j,s)}$ or $F_{T\mathbb{A}_m}^{(j,s)}(t)$. For instance, $\theta\left(F_T^{(j,s)}\right) = \int_0^{\tau^*}\left(1 - F_T^{(j,s)}(t)\right)dt = \mu_{\tau^*}^{(j,s)}$ or $\theta\left(F_T^{(j,s)}\right) = \ell_{\tau^*}^{(j,s)}$ with $\tau^* \leq \tau_Z$, or $\theta\left(F_T^{(j,s)}\right) = \lambda_T^{(j,s)}(t)$ for $t < \tau_Z$.

The corresponding decomposition is

$$\begin{aligned} \Delta_\theta^{(j,s)} &= \theta(F_T^{(j)}) - \theta(F_T^{(s)}) \\ &= \left[\theta(F_T^{(j,j)}) - \theta(F_T^{(j,s)})\right] + \left[\theta(F_T^{(j,s)}) - \theta(F_T^{(s,s)})\right] \\ &= \Delta_\theta^{(j,s)C} + \Delta_\theta^{(j,s)S}, \end{aligned} \tag{8}$$

e.g. $\Delta_{\mu_{\tau^*}}^{(j,s)}$ corresponds to $\theta\left(F_T^{(j,s)}\right) = \mu_{\tau^*}^{(j,s)}$. A similar decomposition can be performed using $F_{T\mathbb{A}_m}^{(j,s)}$, i.e. replace $F_T^{(j,s)}$ by $F_{T\mathbb{A}_m}^{(j,s)}$ in (8).

III. Estimation Under a Proportional Hazard Specification

The sample observed consists of *iid* $\left\{Z_i^{(j)}, X_i^{(j)}, \delta_i^{(j)}\right\}_{i=1}^{n_j}$ as $\left(Z^{(j)}, X^{(j)}, \delta^{(j)}\right)$, $j = 0, 1$, where $Z^{(j)} = \min\left(T^{(j)}, C^{(j)}\right)$, and $\delta^{(j)} = 1_{\left\{T^{(j)} \leq C^{(j)}\right\}}$. First of all, we must realize that it is not possible to estimate $F_T^{(j)}(t)$ for $t \geq Z_{n_j:n_j}^{(j)}$, where $Z_{1:n_j}^{(j)} \leq Z_{2:n_j}^{(j)} \leq ... \leq Z_{n_j:n_j}^{(j)}$ are the order statistics of $\left\{Z_i^{(j)}\right\}_{i=1}^{n_j}$. Henceforth, ties within censored or uncensored duration times are ordered arbitrarily and ties among uncensored and censored durations are treated as if the former precedes the latter.

Assuming the PH specification (5), the parameter $\eta^{(j)}$ can be consistently estimated using partial likelihood by

$$\tilde{\eta}_{n_j}^{(j)} = \arg\max_d \prod_{i=1}^{n} \left[\frac{\exp\left(d^{\mathrm{T}} X_i^{(j)}\right)}{\sum_{\ell \in \left\{j : Z_i^{(j)} \geq Z_\ell^{(j)}\right\}} \exp\left(d^{\mathrm{T}} X_\ell^{(j)}\right)} \right]^{\delta_j},$$

and the corresponding Breslow (1974) estimator of $\Lambda_b^{(j)}(t)$ is the step function, with jumps at each observed uncensored duration time,

$$\tilde{\Lambda}_{bn_j}^{(j)}(t) = \sum_{i \leq t} \frac{\delta_i^{(j)}}{\sum_{\ell \in \left\{j : Z_i^{(j)} \geq Z_\ell^{(j)}\right\}} \exp\left(\hat{\eta}_{n_j}^{(j)\mathrm{T}} X_\ell^{(j)}\right)} \text{ for } t < Z_{n_j:n_j}^{(j)}.$$

Consistency and weak convergence of these estimators have been justified in Tsiatis (1981) under A.0. The associated estimator of $\bar{F}_T^{(i,s)}(t)$ using continuous time is,

$$\tilde{F}_{Tn_j}^{(j)}(t) = 1 - \frac{1}{n_s} \sum_{i=1}^{n_s} \exp\left(-\tilde{\Lambda}_{bn_j}^{(j)}(t) \exp\left(\tilde{\eta}_{n_j}^{(j)\mathrm{T}} X_i^{(s)}\right)\right), t < Z_{n_j:n_j}^{(j)}, with n = (n_0, n_1).$$

Weak convergence of $\tilde{F}_{Tn}^{(j,s)}$ is an immediate consequence of the well developed asymptotic theory for $\left(\tilde{\eta}_{n_j}^{(j)}, \tilde{\Lambda}_{bn_j}^{(j)}\right)$ and, hence, for $\tilde{F}_{Tn}^{(j,j)}(t)$. Tsiatis (1981) showed that, under $A.0$, $\tilde{\eta}_{n_j}^{(j)} = \eta^{(j)} + O_p\left(n_j^{-1/2}\right)$ and $\sup_{t \in [0,\tau_Z]} \left|\left(\tilde{\Lambda}_{bn_j}^{(j)} - \Lambda_b^{(j)}\right)(t)\right| = O_p\left(n_j^{-1/2}\right)$. Thus, applying Theorem 4.1 in Chernozukov, Fernández-Val and Melly (2013), for $j, s = 0, 1$,

$$\sup_{t \in [0,\tau_Z]} \left|\left(\tilde{F}_{Tn}^{(j,s)} - F_T^{(j,s)}\right)(t)\right| = O_p\left(\sqrt{\frac{n_0 + n_1}{n_0 n_1}}\right) \text{ as } n_0, n_1 \to \infty \text{ and } \frac{n_j}{n_0 + n_1} \to s_j \in (0,1). \tag{9}$$

Asymptotic confidence intervals for $\tilde{F}_{Tn}^{(j,s)}(t)$ can be obtained as in Andersen and Gill (1982). Bootstrap confidence intervals can also be obtained using techniques designed for the PH model, as in Burr (1994), which can be justified in the lines of Chernozukov, Fernández-Val and Melly (2013) Theorem 4.2.

Consequently, the natural estimator of $\mu_{\tau^*}^{(j,s)}$ for $\tau^* \leq \tau_Z^{(j)}$is

$$\begin{aligned}
\tilde{\mu}_{\tau^* n}^{(j,s)} &= \int_0^\infty \min\left(Z_{i:n_j}^{(j)}, \tau^*\right) \tilde{F}_n^{(j,s)}(dt) \\
&= \sum_{i=1}^{n_j} \min\left(Z_{i:n_j}^{(j)}, \tau^*\right)\left[\tilde{F}_{Tn}^{(j:s)}\left(Z_{i:n_j}^{(j)}\right) - \tilde{F}_{Tn}^{(j:s)}\left(Z_{i-1:n_j}^{(j)}\right)\right].
\end{aligned}$$

Consistency of $\tilde{\mu}_{\tau^* n}^{(j,s)}$ follows straightforwardly from (9).

$\lambda^{(j,s)}(\lambda)$ can be estimated by the jump function,

$$\tilde{\lambda}_n^{(j,s)}(t) = \frac{\tilde{F}_{Tn}^{(j,s)}\{t\}}{1 - \tilde{F}_{Tn}^{(j,s)}(t-)},\ t < Z_{n_j:n_j}^{(j)},$$

and $\gamma^{(j,s)}$ by

$$\tilde{\gamma}_n^{(j,s)}(t) = \left[\tilde{\Lambda}_{bn_j}^{(j)}(t) - \tilde{\Lambda}_{bn_j}^{(j)}(t-)\right]\frac{1}{n_s}\sum_{i=1}^{n_s}\exp\left(X_i^{(s)\mathrm{T}}\tilde{\eta}_{n_j}\right),\ t < Z_{n_j:n_j}^{(j)}.$$

A smooth version of $\tilde{\lambda}_n^{(j,s)}(t)$ is,

$$\check{\lambda}_n^{(j,s)}(t) = \frac{\check{f}_{nh}^{(j,s)}(t)}{1 - \tilde{F}_{Tn}^{(j,s)}(t-)},\ t < Z_{n_j:n_j}^{(j)},$$

with $\check{f}_n^{(j,s)}(t) = \int_0^\infty K_h(\bar{t} - t)\tilde{F}_{Tn}^{(j,s)}(d\bar{t})$, where $K_h(\cdot) = h^{-1}k(\cdot/h)$, k is an even kernel and h a suitable chosen bandwidth. The corresponding estimated decompositions are obtained by plugging-in these estimates in (8), i.e. for any functional θ, $\tilde{\Delta}_{\theta n}^{(j,s)} = \theta(\tilde{F}_{Tn}^{(j)}) - \theta(\tilde{F}_{Tn}^{(s)})$ estimates $\Delta_\theta^{(j,s)}$ and is decomposed as

$$\begin{aligned}
\tilde{\Delta}_{\theta n}^{(j,s)} &= \left[\theta(\tilde{F}_T^{(j,j)}) - \theta(\tilde{F}_T^{(j,s)})\right] + \left[\theta(\tilde{F}_T^{(j,s)}) - \theta(\tilde{F}_T^{(s,s)})\right] \\
&= \tilde{\Delta}_{\theta n}^{(j,s)C} + \tilde{\Delta}_{\theta n}^{(j,s)S}.
\end{aligned}$$

IV. Estimation under a Non-Parametric Specification

Aunder A.1, $\bar{F}_T^{(j)}$ is consistently estimated by the KM estimator,

$$\hat{F}_{Tn_j}^{(j)}(t) = 1 - \prod_{i=1}^{n_j}\left[1 - \frac{\delta_{[i:n_j]}^{(j)}}{n_j - i + 1}\right]^{1_{\{Z_{i:n_j}^{(j)} \leq t\}}} = \sum_{i=1}^{n}\omega_{n_j i}^{(j)} 1_{\left\{Z_{i:n_j}^{(j)} \leq t\right\}},\ \tau^* < \tau_Z^{(j)},$$

where

$$\omega_{n_j i}^{(j)} = \frac{\delta_{[i:n_j]}^{(j)}}{n_j - i + 1} \prod_{\ell=1}^{i-1} \left[\frac{n_j - \ell}{n_j - \ell + 1} \right]^{\delta_{[\ell:n_j]}^{(j)}}, \; i = 1, ..., n, \; j = 0, 1,$$

are KM weights, and for any generic sequence $\{\xi_i\}_{i=1}^n$, $\xi_{[i:n]}$ is the $i-th$ $\xi-$concomitant of the ordered $Z-values$, i.e. $\xi_{[i:n]} = \xi_j$ if $Z_{i:n} = Z_j$. Then, $\omega_{n_j i}^{(j)}$ is the mass attached to $\left(Z_{i:n_j}^{(j)}, X_{[i:n_j]}^{(j)}\right)$. Likewise, under A.1 and A.2, $F^{(j)}$ can be consistently estimated by

$$\hat{F}_{n_j}^{(j)}(t, x) = \sum_{i=1}^{n_j} \omega_{n_j i}^{(j)} 1_{\left\{Z_{i:n_j}^{(j)} \leq t, \; X_{[i:n_j]} \leq x\right\}} \text{ for } t < \tau_Z^{(j)},$$

which suggests estimating $\mu_{\tau^*}^{(j)}$ by

$$\hat{\mu}_{\tau^* n_j}^{(j)} = \sum_{i=1}^{n_j} \omega_{n_j i}^{(j)} \min\left(Z_{i:n_j}^{(j)}, \tau^*\right), \; \tau^* \leq \tau_Z^{(j)}, \; j = 0, 1.$$

Notice that, for $j = 0, 1$, $\hat{\mu}_{\tau^* n_j}^{(j)} = \hat{\alpha}_{\tau^* n_j}^{(j)} + \hat{\beta}_{\tau^* n_j}^{(j)\mathrm{T}} \hat{\mu}_{X n_j}^{(j)}$, with $\hat{\mu}_{X n_j}^{(j)} = n_j^{-1} \sum_{i=1}^{n_j} X_i^{(j)} \omega_{n_j i}^{(j)}$ and

$$\begin{aligned}
\begin{pmatrix} \hat{\alpha}_{\tau^* n_j}^{(j)} \\ \hat{\beta}_{\tau^* n_j}^{(j)} \end{pmatrix} &= \arg\min_{a,b} \int_0^\infty \int_{\mathbb{R}^k} \left(\min(t, \tau^*) - a - b^{\mathrm{T}} x\right)^2 \hat{F}_{n_j}^{(j)}(dt, dx) \\
&= \arg\min_{a,b} \sum_{i=1}^{n_j} \left(\min\left(Z_{i:n_j}^{(j)}, \tau^*\right) - a - b^{\mathrm{T}} X_{[i:n_j]}^{(j)}\right)^2 \omega_{n_j i}^{(j)} \\
&= \begin{pmatrix} \hat{\mu}_{\tau^* n_j}^{(j)} - \hat{\beta}_{\tau^* n_j}^{(j)\mathrm{T}} \hat{\mu}_{X n_j}^{(j)} \\ \left[\sum_{i=1}^{n_j} X_{[i:n]}^{(j)} X_{[i:n]}^{(j)\mathrm{T}} \omega_{n_j i}^{(j)}\right]^{-1} \sum_{i=1}^{n_j} X_{[i:n]}^{(j)} \min\left(Z_{i:n_j}^{(j)}, \tau^*\right) \omega_{n_j i}^{(j)} \end{pmatrix}.
\end{aligned}$$

Therefore, the standardized BLP under $\hat{F}_{n_j}^{(j,s)}(t, x)$ is $\hat{\ell}_{\tau^* n_j}^{(j,s)} = \hat{\alpha}_{\tau^* n_j}^{(j)} + \hat{\beta}_{\tau^* n_j}^{(j)\mathrm{T}} \hat{\mu}_{X n_s}^{(s)}$ and $\hat{\ell}_{\tau^* n_j}^{(j,j)} = \hat{\mu}_{\tau^* n_j}^{(j)}$.

In turn, regarding the standardization based on Kitawaga's (1955) method, the KM analog of (6) based on a partition $\mathbb{A}_m$ is,

$$\hat{F}_{T\mathbb{A}_m, n}^{(j,s)}(t) = \sum_{k=1}^{m} \frac{\hat{q}_{\mathcal{A}_k n_s}^{(s)}}{\hat{q}_{\mathcal{A}_k n_j}^{(j)}} \hat{p}_{\mathcal{A}_k n_j}^{(j)}(t), \; j, s = 0, 1,$$

where $\hat{p}_{\mathcal{A} n_j}^{(j)}(t) = \sum_{i=1}^{n_j} \omega_{n_j i}^{(j)} 1_{\left\{Z_{i:n_j}^{(j)} \leq t, X_{[i:n_j]}^{(j)} \in \mathcal{A}\right\}}$ and $\hat{q}_{\mathcal{A} n_j}^{(j)} = \hat{p}_{\mathcal{A} n_j}^{(j)}(\infty)$. Notice that $\hat{F}_{Tn}^{(j)}(t) =$

$\hat{F}^{(j,j)}_{T\mathbb{A}_m,n}(t) = \sum_{i=1}^{n_j} \omega^{(j)}_{n_j i} 1_{\left\{Z^{(j)}_{(i:n)} \leq t\right\}}$ does not depend on the partition $\mathbb{A}_m$.

Stute (1993) provided, under $A.1$ and $A.2$, a Glivenko-Cantelli theorem for $\hat{F}^{(j)}_{n_j}$ as an estimator of $\bar{F}$, and showed that for any function $\varphi : \mathbb{R}^+ \times \mathbb{R}^k \to \mathbb{R}$ and $S^{(j)}_{\varphi} = \varphi\left(T^{(j)}, X^{(j)}\right)$, such that $\mathbb{E}S^{(j)}_{\varphi} < \infty$, $\hat{S}^{(j)}_{\varphi n_j} = \sum_{i=1}^{n_j} \varphi\left(Z^{(j)}_{i:n_j}, X^{(j)}_{[i:n_j]}\right) \omega^{(j)}_{n_j}$ is an $a.s.$-consistent estimator of $\bar{S}^{(j)}_{\varphi} = \int_{\mathbb{R}^+ \times \mathbb{R}^k} \varphi(t,x) \bar{F}(dt,dx)$, which shows that $\hat{\ell}^{(j,s)}_{\tau^* n_j}$ and $\hat{p}^{(j)}_{\mathcal{A}_k n_j}(t)$ are $a.s.$-consistent estimators of $\ell^{(j,s)}_{\tau^* n_j}$and $p^{(j)}_{\mathcal{A}n_j}(t)$, respectively. This justifies that $\hat{F}^{(j,s)}_{T\mathbb{A}_m,n}$ is a consistent estimator of $F^{(j,s)}_{T\mathbb{A}_m}$. Stute (1996) found the asymptotic distribution of $\hat{S}^{\varphi}_{n_j}$. Bootstrap confidence intervals can be obtained as in Stute, González-Manteiga and Sellero (2000). Based on $\hat{F}^{(j,j)}_{T\mathbb{A}_m,n}$, we can estimate $\mu^{(j,s)}_{\tau^*}$, when $\mathbb{E}\left(\min\left(\tau^*, T^{(j)}\right) \middle| X^{(j)}\right)$ is nonparametric, by

$$\hat{\mu}^{(j,s)}_{\tau^* \mathbb{A}_m n} = \int_{\mathbb{R}^+} t \cdot \hat{F}^{(j,s)}_{T\mathbb{A}_m,n}(dt) = \sum_{k=1}^{m} \frac{\hat{q}^{(s)}_{\mathcal{A}n_s}}{\hat{q}^{(j)}_{\mathcal{A}n_j}} \sum_{i=1}^{n_j} \min\left(Z^{(j)}_{i:n_j}, \tau^*\right) 1_{\left\{X^{(j)}_{[i:n]} \leq \mathcal{A}_k\right\}} \omega^{(j)}_{n,i},$$

with $\hat{\mu}^{(j,j)}_{\tau^* \mathbb{A}_m n} = \hat{\mu}^{(j)}_{\tau^* n}$. The corresponding $\lambda^{(j,s)}_{\mathbb{A}_m}(t)$ estimator is

$$\hat{\lambda}^{(j,s)}_{\mathbb{A}_m n}(t) = \frac{\hat{F}^{(j,s)}_{T\mathbb{A}_m,n}\{t\}}{1 - \hat{F}^{(j,s)}_{T\mathbb{A}_m,n}(t-)},$$

and its smooth version is

$$\breve{\lambda}^{(j,s)}_{\mathbb{A}_m n}(t) = \frac{\breve{f}^{(j,s)}_{T\mathbb{A}_m,n}(t)}{1 - \hat{F}^{(j,s)}_{T\mathbb{A}_m,n}(t-)},$$

$\breve{f}^{(j,s)}_{T\mathbb{A}_m,n}(t) = \int_0^\infty K_h(\bar{t} - t) \hat{F}^{(j,s)}_{T\mathbb{A}_m,n}(d\bar{t})$.

Regarding the decompositions, we estimate $\Delta^{(j,s)}(t) = \left(F^{(j)}_T - F^{(s)}_T\right)(t)$ by $\hat{\Delta}^{(j,s)}_n(t) = \hat{F}^{(j)}_{Tn}(t) - \hat{F}^{(s)}_{Tn}(t)$ and the corresponding decomposition is

$$\begin{aligned}\hat{\Delta}^{(j,s)}_n(t) &= \left[\hat{F}^{(j,j)}_{T\mathbb{A}_m,n}(t) - \hat{F}^{(j,s)}_{T\mathbb{A}_m,n}(t)\right] + \left[\hat{F}^{(j,s)}_{T\mathbb{A}_m,n}(t) - \hat{F}^{(s,s)}_{T\mathbb{A}_m,n}(t)\right] \\ &= \hat{\Delta}^{(i,s)C}_{\mathbb{A}_m n}(t) + \hat{\Delta}^{(i,s)S}_{\mathbb{A}_m n}(t).\end{aligned}$$

Also (8) can also be estimated nonparametrically substituting $F^{(j,s)}_T$ by $\hat{F}^{(j,s)}_{T\mathbb{A}_m,n}$. For instance both $\theta\left(\hat{F}^{(j,s)}_{T\mathbb{A}_m,n}(t)\right) = \hat{\ell}^{(j,s)}_{T\mathbb{A}_m n}$ and $\theta\left(\hat{F}^{(j,s)}_{T\mathbb{A}_m,n}(t)\right) = \hat{\mu}^{(j,s)}_{\mathbb{A}_m \tau^*,n}$ result in decompositions of $\hat{\Delta}^{(j,s)}_{\mu_{\tau^*}n}(t) = \hat{\mu}^{(j)}_{\tau^*n} - \hat{\mu}^{(s)}_{\tau^*n}$, and $\theta\left(\hat{F}^{(j,s)}_{T\mathbb{A}_m,n}(t)\right) = \hat{\lambda}^{(j,s)}_{T\mathbb{A}_m n}(t)$ results in a decomposition of $\Delta^{(j,s)}_{\lambda_{\mathbb{A}_m}n}(t) = \hat{\lambda}^{(j)}_n(t) - \hat{\lambda}^{(s)}_n(t)$.

V. Monte Carlo Simulations

This section provides evidence on the finite sample performance of the alternative standardization methods, and the resulting counterfactual decompositions, for the RMST. Henceforth, PML stands for the partial maximum likelihood estimator based on the PH specification, OB-KM refers to the estimation method based on the classical OB decomposition using KM weights, and NP(m) refers to the nonparametric method with m partitions; we use $m = 3$ and 10. We also consider the OB-KM method using the OLS fit of a polynomial of order 3, rather than a linear model, named OB-KM-Pol3. By simplicity, we only consider a single covariate as component. In the case of NP(m), partitions are defined according to classes of equal size.

We consider three designs described in Table 1. In the data generating process (DGP) DGP1, $T^{(j)}$ are generated using a linear regression model with independent Gaussian errors. In DGP2, $T^{(j)}$ is distributed as a Weibull random variable, i.e. the specification corresponds to the Cox's PH model. Finally, in DGP3, $T^{(j)}$ are generated according to a nonlinear regression model with independent Gaussian errors. Parameter values of distributions in the different designs were chosen to produce a censoring level of 30%.

TABLE 1 ABOUT HERE

In all simulations we consider sample sizes of 200, 800 and 3,200 and Monte Carlo experiments based on 1,000 replications. We make comparisons of the alternative standardization methods are based on the mean absolute error (MAB) and the root mean squared error (RMSE). Since $C^{(j)}$ and $T^{(j)}$ are unbounded, we can consistently estimate the means $\mu^{(j)}$ and the corresponding counterfactual decomposition of their difference.

However, simulations are performed using $\mu_{\tau_Z}^{(j)}$ as target parameter, without exploiting the fact that $T^{(j)}$ and $C^{(j)}$ are unbounded and, hence, $\tau_Z = \infty$ and $\mu_{\tau_Z}^{(j)} = \mu^{(j)}$. The parameter τ_Z is consistently estimated by $\hat{\tau}_{Zn} = \min\left(Z_{n_0:n_0}^{(0)}, Z_{n_1:n_1}^{(1)}\right)$ and the RMST is calculated with $\tau^* = \hat{\tau}_{Zn}$.

Table 2 presents results under DGP1 design. Since the PH specification is incorrect, the PML is inconsistent for $\mu^{(j,s)}$, $j, s = 0, 1$, which explains the large biases observed. The regression function is linear and, hence, both OB-KM and OB-KM-Pol3 are consistent. As expected, OB-KM, which is asymptotically the most efficient, performs best in finite samples, but there are no significant loses using the over-parameterized OB-KM-Pol3. Of course, the nonparametric NP(3) and NP(10), which does not use any information on the DGP, perform worse.

TABLE 2 ABOUT HERE

Table 3 reports results under DGP2 specification. The OB-KM and OB-KM-Pol 3 are still consistent estimators of $\mu^{(j)} = \mu^{(j,j)}$, but are inconsistent estimators of $\mu^{(j,s)}$ for $j \neq s$. In turn, since the specification corresponds to the Cox's model, the PML is consistent and more efficient than NP(m). This is confirmed by the simulations. Interestingly, OB-KM-Pol3 seems to be a fairly robust alternative to OB-KM and performs similarly to NP(m).

TABLE 3 ABOUT HERE

Table 4 reports results under DGP3 specification. In this case, all $\mu^{(j,s)}$ estimators for $j \neq s$, but NP(m), are inconsistent, though OB-KM and OB-KM-Pol3 are consistent for $\mu^{(j)} = \mu^{(j,j)}$. NP(m) performs much better than both PML and OB-KM in this case, particularly for the larger sample sizes. However, there is sensitivity with respect to the number of classes choices, i.e. NP(3) versus NP(10). OB-KM-Pol3 shows to be a robust alternative to OB-KM that captures fairly well the underlying regression nonlinearity.

TABLE 4 ABOUT HERE

We also study the effect of ignoring the presence of censoring. First, assessing the effect on the estimates when a correction of censoring is not taken into account and, second, analyzing the effect of trying to estimate the mean when $C^{(j)}$ is bounded. Table 5 provides the mean absolute bias and RMSE, using DGP1 design, for the decomposition effects based on the OLS fits using all observations but ignoring the presence of censoring, i.e. assuming that $Z^{(j)}$ is the actual duration. And also, an OLS fits dropping out the censored observations. These biased estimates are compared with the corresponding OB-KM estimator. Simulation results show serious biases when ignoring censoring.

TABLE 5 ABOUT HERE

Table 6 illustrates the effect of neglecting the fact that $\tau_T^{(j)} > \tau_C^{(j)}$ when estimating the duration mean. In this case, F_T cannot be consistently estimated beyond $\tau_Z^{(j)}$ and, unlike previous experiments where $\tau_T = \tau_C = \infty$, it is not possible estimating $\mu^{(j,s)}$ or $\mu_{\tau^*}^{(j,s)}$ for $\tau^* > \tau_Z^{(j)}$. We compare PML estimators of $\mu_{\tau_Z}^{(j,s)}$ under design DGP2, i.e. PML is efficient, when $C^{(j)}$ is censored, using $\widetilde{C}^{(j)} = \min\left(C^{(j)}, 6.5\right)$ as censored variable, i.e. $\tau_C^{(0)} = \tau_C^{(1)} = 6.5 < \tau_T^{(0)} = \tau_T^{(1)} = \infty$. Table 6 confirms high biases for estimating the mean when $\tau_T^{(j)} > \tau_C^{(j)}$, but the PML estimator still performs well as an $\mu_{\tau_Z}^{(j,s)}$ estimator. This shows the importance of focusing the statistical inference in truncated or restricted parameters as the RMST.

TABLE 6 ABOUT HERE

VI. Unemployment Duration Gender Gaps in Spain

This section investigates the causes of unemployment durations gender gaps using counterfactual decompositions of HF and RMST differences in Spain. We also provide a comparison between ACHF and HF estimates using the alternative specifications. The Spanish case is particularly interesting because it has experienced one of the highest unemployment rates among OECD countries in recent decades. According to official statistics (see OECD 2013), for the period 1995-2005, the average unemployment rate was around 6.8% in OECD countries and 5% in the US, while it was 14% in Spain. Moreover, the difference in unemployment rates by gender has also been important. For the same period, women exhibited on average an unemployment rate 9 percentage points (p.p.) higher, while in the US this gap was around 0.04 p.p.

The existing literature has mainly paid attention to gender gaps in the aggregated unemployment rate, e.g. Niemi, (1974), Johnson (1983), Azmat et al. (2006), or Queneau (2007), but gender gaps in other unemployment features, like spells of unemployment duration have received less attention. Research devoted to study unemployment duration gender gaps has almost exclusively focused in explaining the gender differences in the ACHF, e.g. Ham (1999) for the Czech and Slovak Republic, Gonzalo (2000) for Finland, Du (2009) for China, Tansel (2010) for Turkey, Baussola (2014) for Italy and UK, and De la Rica and Robledo-Saenz (2017) for Spain. As we have discussed, the ACHF and HF shapes are generally unrelated and may lead to misleading analyses of spells of unemployment durations.

We implement the proposed methodologies to perform counterfactual decompositions of the HF and RMST differences using data from the Survey of Income and Living Conditions (SILC) for the period 2004-2007. This survey, carried out by the European Commission, is a rotative household panel that collects information on socioeconomic characteristics, including the occupational status (monthly) for a period of 4 years. Our population consists of unemployed workers older than 25 starting an spell of unemployment during the period 2004-2007. We measure unemployment duration as the number of months that a worker is not employed, which is usually referred as non-employment duration.

We consider as composition variables those commonly used in unemployment duration analysis such as age, educational level, tenure, marital status, whether the individual is head of the household, the number of unemployed in the household, city size (according to three levels of urbanization in the SILC as big city, medium size city and small city) and region, see e.g., Addison (2003), Kuhn (2004), or Tansel (2010). The first three variables control by human capital characteristics, while the others are related to the opportunity cost of being unemployed and the reservation wage. Table 7 provides the means and standard deviations for these variables and Figure 1 the corresponding QQ plots for the continuous variables, i.e. age and tenure. Composition variables are similar in the two populations, except in the case of tenure, which implies that the composition effect should not be particularly important explaining gender gaps. We observe censoring levels of 21.4% for women and 16.2% for men.

TABLE 7 and FIGURE 1 ABOUT HERE

Henceforth, population 0 corresponds to women and population 1 to men. First we analyze the HF differences between the two populations. Figure 2 provides KM estimates of the marginal HF, $\lambda^{(j)}$, and the corresponding nonparametric estimates of $\lambda_{\mathbb{A}_m}^{(0,1)}$ based on a partition $\mathbb{A}_m$, as well as the HF difference, $\Delta_\lambda^{(0,1)}$, into counterfactural effects. The partition is based on $m = 8$ classes and $\lambda_{\mathbb{A}_8}$ is estimated using kernel smoothing, $\check{\lambda}_{\mathbb{A}_m n}^{(0,1)}$, with an Epanechnikov's kernel and the bandwidth chosen by the classical plug-in method. The number of classes in the partition was established according to the structure of the composition variables using some natural thresholds for the continuous ones. For instance, we grouped workers between 25 and 40 years old (prime-age workers) and workers with less than 10 years tenure. Some might have very few observations due to the high number of possible partitions and the correlation between covariates, e.g. there are few observations in a partition class with young workers of more than 10 years tenure, married and with high education level. Therefore, by using age, labor market size and marital status, we construct 8 classes. This method has the advantage that the classes in a partition can be chosen in a natural way accounting by the observed relation between duration and covariates.

The nonparametric KM estimates of the HF show a clear acceleration of women's HF after 24 months of unemployment, which is consistent with the unemployment compensation normative in Spain during the analyzed period. Unemployment benefits expired after 24 months and workers received 70% of their salary during the first 6 months and 60% up to 24 months. Figure 2 suggests that, on average, women exhaust all unemployment benefits producing a spike in the HF starting around the expiration rate. That is, women and men exhibit different optimal delays in job acceptance. This phenomenon has also been documented by Roed and Zhang (2003) and Boone and Ours (2012) using Norwegian and Slovenian data, respectively. We observe that the estimated composition effect is very small at any period. That is, HF differences between women and men do not seem to be explained by socioeconomic characteristics, which are almost identical, but by anything else, like circumstances of the labor market tightness and discrimination related to institutional factors, labor circumstances or behavioral aspects (Bachmann and Sinning 2016).

FIGURE 2 ABOUT HERE

We have checked Cox's model specification for the two populations using the popular Schoenfeld's (1982) residual based goodness-of-fit test, which result in p-values of 0.031 and 0.654 for women and men, respectively. Therefore, Cox's specification is rejected for women, but no for men, at 5% of significance. Figure 3 shows the smooth version of PML HF estimates, $\check{\lambda}_n^{(0,1)}$, with the same kernel and bandwidths used in the nonparametric case. The estimates are quantitatively very different to the nonparametric ones in Figure 2, possibly because of misspecification, though it also shows a composition effect close to zero at any period and an acceleration of the HF for women after 24 months of unemployment.

FIGURE 3 ABOUT HERE

Figure 4 and 5 provide ACHF estimates using NP(8) and PML, respectively. HF and ACHF shapes are very different for each estimator. However, the counterfactual composition effect is also close to zero. The PLM's ACHF estimates of the counterfactual decomposition is particularly hard to interpret.

FIGURE 4 and 5 ABOUT HERE

Next, we analyze $\mu_{\tau^*}^{(j)}$ and $\mu_{\tau^*}^{(0,1)}$ estimates for $\tau^* = \hat{\tau}_{Zn} = 42$ using the OB-KM, PML and NP(8), which produces estimates of the duration means when $\tau_C^{(j)} \geq \tau_T^{(j)}$, and also for $\tau^* = 24, 36$, which are of interest when studying short and medium term unemployment differences. Estimates of RMST and the corresponding counterfactual decompositions can be found in tables 8 and 9, respectively. The estimate of the non standardized RMST during the first 42 months is around 10.8 months for women and 7.9 months for men. The estimate of the standardized RMST, which is interpreted as the average unemployment duration during the first 42 months if women would had the same component distribution than men, is around 10.3 months, close to the corresponding non standardized value. Results across methods are qualitatively similar and reveal a reduction in the corresponding standardized RMST.

TABLE 8 and 9 ABOUT HERE

The counterfactual decompositions of RMST using the OB-KM and PML estimates for $\tau^* = \hat{\tau}_{Zn} = 42$ (see Table 9) are fairly different to the NP(8), which may indicate a misspecification of the underlying structures assumed. For instance, the OB-KM estimator might be biased because $\mathbb{E}\left(\min(\tau^*, Z^{(j)}) \middle| X^{(j)}\right)$ is non-linear and the PML is also biassed because the underlying CCDF does not follow a PH specification, which was confirmed by Schoenfeld's test.

The counterfactural composition effect for any τ^* is close to zero for any of the three methods. This indicates that the difference in workers' characteristics slightly increases the severity of women unemployment duration with respect to men, which might be driven mainly through tenure, as suggested in previous descriptive analysis. Notice that the counterfactual effects using OB-KM and PML methods are similar, but the composition effect using NP(8) is much smaller, as we have already seen for the HF in Figures 2 and 3. Similar results are obtained for $\tau^* = 24$ and $\tau^* = 36$. However, based on NP(8) and OB-KM, it seems that the composition effect is more important when explaining unemployment duration gender gaps in the first stages of unemployment, but is always much smaller than the structure effect.

VII. Conclusions

We have provided a methodological approach for standardizing lifetime distributions when durations are observed under right censoring. Our research is motivated from the analysis of unemployment gender gaps, which has been carried out in previous research using ACHF comparisons. We have shown that ACHF is not necessarily related to HF and can produce misleading conclusions. Since estimation of duration means is not possible because of censoring, we have focussed on the RMST. This has allowed examining counterfactual decompositions of mean unemployment duration time differences during different time periods. We have also discussed the standardization of the HF, under a semiparametric PH and nonparametric specifications, which has been compared with the corresponding ACHF.

We have proposed a natural extension of the OB type decomposition of the RMST to the censored data case using KM weights. Monte Carlo studies show that under nonlinearity, the OB method may produce serious biases. Nonparametric estimates of the CCDF based on partitions of the components support forms a basis for standardizing the CDFs, or any feature of interest, without imposing more restrictions than those needed for the identification of the underlying nonparametric joint CDF of duration and components. The number and size of the classes in the partition are fixed and does not need to shrink as the sample size increases for consistency of the corresponding estimates.

The proposed methodology has been applied to investigate the causes of unemployment duration gender gaps in Spain. Findings indicate that the composition of men and women populations is not relevant explaining HF and RMST gender gaps, which are mainly explained by the structure effect related to market and workers, or employers, behaviors.

REFERENCES

Addison, John T., and Pedro Portugal, "Unemployment Duration Competing and Defective Risks," *Journal of Human Resources* 38 (2003), 156-191.

Andersen, Per Kragh, and Richard David Gill, "Cox's Regression Model for Counting Processes: A Large Sample Study." *The Annals of Statistics* 10 (1982), 1100-1120.

Azmat, Ghazala, Maia Güell, and Alan Manning, "Gender Gaps in Unemployment Rates in OECD Countries," *Journal of Labor Economics* 24 (2006), 1-37.

Bachmann, Ronald, and Mathias Sinning, "Decomposing the Ins and Outs of Cyclical Unemployment," *Oxford Bulletin of Economics and Statistics* 78 (2016), 853-876.

Barsky, Robert, John Bound, Kerwin Charles, and Joseph Lupton, "Accounting for the black–white wealth gap: a nonparametric approach. ", *Journal of the American Statistical Association*, 97 (2002), 663-673.

Bauer, Thomas K., and Mathias Sinning, "An Extension of the Blinder-Oaxaca Decomposition to Nonlinear Models." *AStA Advances in Statistical Analysis* 92 (2008), 197-206.

Blau, Francine D., and Lawrence M. Kahn, "The Gender Earnings Gap: Learning from International Comparisons." *The American Economic Review* 82 (1992), 533-538.

Blinder, Alan S, "Wage Discrimination: Reduced Form and Structural Estimates," *Journal of Human Resources* 8 (1973), 436-455.

Boone, Jan, and Jan C. van Ours, "Why is there a spike in the job finding rate at benefit exhaustion?." *De Economist* 160 (2012), 413-438.

Breslow, Norman, "Covariance Analysis of Censored Survival Data," *Biometrics* 30 (1974), 89-99.

Burr, Deborah, "A comparison of certain bootstrap confidence intervals in the Cox model." *Journal of the American Statistical Association* 89 (1994), 1290-1302.

Cain, Glen, "The Economic Analysis of Labor Market Discrimination: A Survey," in Orley Ashenfelter and David Card, eds. *Handbook of Labor Economics* 1, (Elsevier Science Publishers, 1986, 693-785).

Canudas-Romo, Vladimir, *Decomposition Methods in Demography* (Amsterdam: Rozenberg Publishers, 2003).

Chen, Pei-Yun, and Anastasios A. Tsiatis. "Causal Inference on the Difference of the Restricted Mean Lifetime between Two Groups." *Biometrics* 57 (2001), 1030-1038.

Chernozhukov, Victor, Iván Fernández-Val, and Blaise Melly, "Inference on Counterfactual Distributions." *Econometrica* 81 (2013), 2205-2268.

Cox, David R., "Regression Models and Life-Tables," *Journal of the Royal Statistical Society. Series B* (Methodological) 34 (1972), 87-22.

Chiquiar, Daniel, and Gordon H. Hanson, "International migration, self-selection, and the distribution of wages: Evidence from Mexico and the United States." *Journal of Political Economy* 113 (2005), 239-281.

Das Gupta, Prithwis, *Standardization and Decomposition of Rates: A User's Manual.* (US Department of Commerce, Economics and Statistics Administration, Bureau of the Census, No. 183, 1993).

De la Rica, Sara, and Yolanda F. Rebollo-Sanz, "Gender Differentials in Unemployment Ins and Outs During the Great Recession in Spain," *De Economist* 165 (2017), 67-99.

DiNardo, John, Nicole M. Fortin, and Thomas Lemieux. "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach," *Econometrica* 64 (1996), 1001–1044.

Donald, Stephen G., and Yu-Chin Hsu, "Estimation and Inference for Distribution Functions and Quantile Functions in Treatment Effect Models," *Journal of Econometrics* 178 (2014), 383-397.

Du, Fenglian, and Xiao-yuan Dong, "Why do Women Have Longer Durations of Unemployment than Men in Post-restructuring Urban China?," *Cambridge Journal of Economics* 33.2 (2008), 233-252.

Efron, Bradley, "Censored data and the bootstrap." *Journal of the American Statistical Association* 76 (1981), 312-319.

Fortin, Nicole, Thomas Lemieux, and Sergio Firpo, "Decomposition Methods in Economics," in Orley Ashenfelter and David Card, eds. *Handbook of Labor Economics* 4 (Elsevier Science Publishers, 2011, 1-102).

Freeman, Richard B, "Unionism and the Dispersion of Wages," *ILR Review* 34.1 (1980), 3-23.

Gonzalo, María Teresa, and Jan Saarela, "Gender differences in Exit Rates from Unemployment: Evidence from a Local Finnish Labour Market," *Finnish Economic Papers* 13 (2000), 129-139.

Ham, John C., Jan Svejnar, and Katherine Terrell. "Women's Unemployment During Transition: Evidence from Czech and Slovak Micro-data," *Economics of Transition* 7 (1999), 47-78.

Irwin, J. O., "The Standard Error of An Estimate of Expectation of Life, with Special Reference to Expectation of Tumourless Life in Experiments with Mice," *Epidemiology and Infection* 47 (1949), 188-189.

Johnson, Janet L, "Sex Differentials in Unemployment Rates: A Case for No Concern," *Journal of Political Economy* 91 (1983), 293-303.

Juhn, Chinhui, Kevin M. Murphy, and Brooks Pierce, "Accounting for the Slowdown in Black-White Wage Convergence", in M.H. Kosters (Ed.), *Workers and Their Wages* (Washington DC: American Enterprise Institute, 1991, 107–143).

Kaplan, Edward L., and Paul Meier, "Nonparametric Estimation from Incomplete Observations," *Journal of the American Statistical Association* 53 (1958), 457-481.

Kalbfleisch, John D., and Ross L. Prentice, "Marginal Likelihoods based on Cox's Regression and Life Model," *Biometrika* 60 (1973), 267-278.

Karrison, Theodore, "Restricted Mean Life with Adjustment for Covariates," *Journal of the American Statistical Association* 82 (1987), 1169-1176.

Kitagawa, Evelyn M., "Components of a Difference between Two Rates," *Journal of the American Statistical Association* 50.272 (1955), 1168-1194.

Kitagawa, Evelyn M., "Standardized Comparisons in Population Research," *Demography* 1 (1964), 296-315.

Kuhn, Peter, and Mikal Skuterud, "Internet Job Search and Unemployment Durations," *American Economic Review* 94 (2004), 218-232.

Lemieux, Thomas, "Decomposing Changes in Wage Distributions: A Unified Approach," *Canadian Journal of Economics* 35 (2002), 646-688.

Machado, José AF, and José Mata, "Counterfactual Decomposition of Changes in Wage Distributions using Quantile Regression," *Journal of Applied Econometrics* 20 (2005), 445-465.

Melly, Blaise, "Decomposition of Differences in Distribution using Quantile Regression," *Labour Economics* 12 (2005), 577-590.

Neison, F. G. P., "On a Method Recently Proposed for Conducting Inquiries into the Comparative Sanatory Condition of Various Districts, with Illustrations, Derived from Numerous Places in Great Britain at the Period of the Last Census," J*ournal of the Statistical Society of London* 7 (1844), 40-68.

Niemi, Beth, "The Female-Male Differential in Unemployment Rates," *ILR Review* 27 (1974), 331-350.

Oaxaca, Ronald, "Male-Female Wage Differentials in Urban Labor Markets," *International Economic Review* 14 (1973), 693-709.

Oaxaca, Ronald L., and Michael R. Ransom, "Identification in Detailed Wage Decompositions," *The Review of Economics and Statistics* 81.1 (1999), 154-157.

O'Neill, June E., and Dave M. O'Neill, "What do Wage Differentials Tell us About Labor Market Discrimination?", in *The Economics of Immigration and Social Policy* (Emerald Group Publishing Limited, 2006, 293-357).

Queneau, Herve, and Amit Sen, "Evidence Regarding Persistence in the Gender Unemployment Gap based on the Ratio of Female to Male Unemployment Rate," *Economics Bulletin* 5 (2007), 1-10.

Reimers, Cordelia W. "Labor market discrimination against Hispanic and black men," *The Review of Economics and Statistics*, 1983, p. 570-579.

Roed, Knut, and Tao Zhang, "Does unemployment compensation affect unemployment duration?." *The Economic Journal* 113 (2003), 190-206.

Rothe, Christoph, "Nonparametric Estimation of Distributional Policy Effects," *Journal of Econometrics* 155 (2010), 56-70.

Schoenfeld, David, "Partial residuals for the proportional hazards regression model," *Biometrika* 69 (1982), 239-241.

Stock, James H., "Nonparametric Policy Analysis," *Journal of the American Statistical Association* 84 (1989), 567-575.

Stute, Winfried, "Consistent Estimation under Random Censorship when Covariables are Present," *Journal of Multivariate Analysis* 45 (1993), 89-103.

Stute, W., W. González Manteiga, and C. Sánchez Sellero, "Nonparametric model checks in censored regression," *Communications in Statistics-theory and Methods* 29 (2000), 1611-1629.

Tansel, Aysit, and H. Mehmet Taşçi, "Hazard Analysis of Unemployment Duration by Gender in a Developing Country: The case of Turkey," *Labour* 24 (2010), 501-530.

Tsiatis, Anastasios A., "A Large Sample Study of Cox's Regression Model," *The Annals of Statistics* 9 (1981), 93-108.

Zhang, Min, and Douglas E. Schaubel, "Estimating Differences in Restricted Mean Lifetime using Observational Data Subject to Dependent Censoring," *Biometrics* 67 (2011), 740-749.

Zucker, David M., "Restricted Mean Life with Covariates: Modification and Extension of a useful Survival Analysis Method," *Journal of the American Statistical Association* 93 (1998), 702-709.

Tables

Table 1 DGPs

	Population 0	Population 1
	$T^{(0)} = 5 + X^{(0)} + \varepsilon_Y^{(0)} \;\; \varepsilon_T^{(0)} \sim \mathcal{N}(0,1)$	$T^{(1)} = 5 + X^{(1)} + \varepsilon_T^{(1)} \;\; \varepsilon_Y^{(1)} \sim \mathcal{N}(0,1)$
DGP1	$X^{(0)} \sim \mathcal{N}(1.5, 0.5)$	$X^{(1)} \sim \mathcal{N}(1, 0.5)$
	$C^{(0)} = 7.5 + \varepsilon_C^{(0)} \;\; \varepsilon_C^{(0)} \sim \mathcal{N}(v_0, 1.5)$	$C^{(1)} = 7 + \varepsilon_C^{(1)} \;\; \varepsilon_C^{(1)} \sim \mathcal{N}(v_1, 1.5)$
	$T^{(0)} \vert X^{(0)} \sim \mathcal{WB}\left(e^{2-x_0}, 5\right)$	$T^{(1)} \vert X^{(1)} \sim \mathcal{WB}\left(e^{2-x_1}, 5\right)$
DGP2	$C^{(0)} \sim \mathcal{WB}\left(e^{2+v_0}, 4\right), v_0 = -0.26$	$C^{(1)} \sim \mathcal{WB}\left(e^{2+v_1}, 4\right), v_1 = -0.22$
	$X^{(0)} \sim \mathcal{U}(0,1)$	$X^{(1)} \sim \sum_{i=1}^{3} \mathcal{U}(0, 1/3)$
	$T^{(0)} = 5 + 2\vert X^{(0)}\vert + \varepsilon_T^{(0)}, \varepsilon_T^{(0)} \sim \mathcal{N}(0,1)$	$T^{(1)} = 5 + 2\vert X^{(1)}\vert + \varepsilon_T^{(1)}, \varepsilon_T^{(1)} \sim \mathcal{N}(0,1)$
DGP3	$C^{(0)} = 9.5 + \varepsilon_C^{(0)} \;\; \varepsilon_C^{(0)} \sim \mathcal{N}(0,2)$	$C^{(1)} = 7.7 + \varepsilon_C^{(1)} \;\; \varepsilon_C^{(1)} \sim \mathcal{N}(0, 1.5)$
	$X^{(0)} \sim \mathcal{U}(-3,3)$	$X^{(1)} \sim \mathcal{N}(0,1)$

Data generating process as for two populations. Parameters in the distributions are chosen to produce 30% of censoring.

Table 2 Comparison of alternative procedures under DGP1

		MAE				
Method	**Sample size**	$\Delta_\mu^{(0,1)C}$	$\Delta_\mu^{(0,1)S}$	$\mu^{(0)}$	$\mu^{(1)}$	$\mu^{(0,1)}$
PML	**200**	0.481	0.444	0.064	0.069	0.469
	800	0.492	0.488	0.034	0.034	0.490
	3200	0.497	0.492	0.017	0.018	0.494
NP(3)	**200**	0.244	0.240	0.065	0.067	0.240
	800	0.146	0.149	0.035	0.034	0.146
	3200	0.109	0.108	0.017	0.018	0.109
NP(10)	**200**	0.249	0.254	0.065	0.067	0.250
	800	0.133	0.139	0.035	0.034	0.135
	3200	0.066	0.067	0.017	0.018	0.066
OB-KM	**200**	0.079	0.113	0.070	0.069	0.100
	800	0.045	0.066	0.037	0.035	0.057
	3200	0.022	0.032	0.018	0.018	0.027
OB-KM-Pol3	**200**	0.090	0.117	0.070	0.067	0.107
	800	0.050	0.069	0.037	0.035	0.060
	3200	0.024	0.033	0.018	0.018	0.028

		RMSE				
Method	**Sample size**	$\Delta_\mu^{(0,1)C}$	$\Delta_\mu^{(0,1)S}$	$\mu^{(0)}$	$\mu^{(1)}$	$\mu^{(0,1)}$
PML	**200**	0.481	0.457	0.080	0.087	0.476
	800	0.492	0.492	0.042	0.044	0.492
	3200	0.497	0.493	0.022	0.022	0.495
NP(3)	**200**	0.313	0.315	0.081	0.084	0.312
	800	0.184	0.189	0.043	0.042	0.186
	3200	0.128	0.128	0.022	0.022	0.128
NP(10)	**200**	0.321	0.327	0.081	0.084	0.320
	800	0.169	0.174	0.043	0.042	0.170
	3200	0.084	0.086	0.022	0.022	0.084
OB-KM	**200**	0.099	0.142	0.088	0.087	0.127
	800	0.057	0.082	0.046	0.044	0.072
	3200	0.029	0.040	0.023	0.023	0.035
OB-KM-Pol3	**200**	0.114	0.148	0.087	0.085	0.136
	800	0.064	0.087	0.046	0.044	0.077
	3200	0.030	0.041	0.023	0.023	0.036

MAB and RMSE for the standardization and decomposition components of the RMST under DGP1 in Table 1.

Table 3 Comparison of alternative procedures under DGP2

		MAE				
Method	**Sample size**	$\Delta_\mu^{(0,1)C}$	$\Delta_\mu^{(0,1)S}$	$\mu^{(0)}$	$\mu^{(1)}$	$\mu^{(0,1)}$
PML	**200**	0.014	0.105	0.080	0.070	0.084
	800	0.009	0.057	0.046	0.035	0.046
	3200	0.006	0.029	0.024	0.017	0.023
NP(3)	**200**	0.163	0.154	0.083	0.070	0.154
	800	0.091	0.089	0.048	0.035	0.087
	3200	0.063	0.059	0.025	0.018	0.059
NP(10)	**200**	0.266	0.259	0.083	0.070	0.257
	800	0.141	0.136	0.048	0.035	0.132
	3200	0.070	0.067	0.025	0.018	0.065
OB-KM	**200**	0.082	0.122	0.097	0.070	0.116
	800	0.085	0.106	0.056	0.035	0.107
	3200	0.099	0.109	0.029	0.018	0.110
OB-KM-Pol3	**200**	0.098	0.121	0.091	0.070	0.110
	800	0.060	0.065	0.055	0.035	0.061
	3200	0.032	0.035	0.029	0.018	0.032

		RMSE				
Method	**Sample size**	$\Delta_\mu^{(0,1)C}$	$\Delta_\mu^{(0,1)S}$	$\mu^{(0)}$	$\mu^{(1)}$	$\mu^{(0,1)}$
PML	**200**	0.017	0.132	0.101	0.088	0.106
	800	0.011	0.072	0.058	0.043	0.058
	3200	0.007	0.036	0.030	0.022	0.029
NP(3)	**200**	0.207	0.196	0.106	0.088	0.198
	800	0.116	0.114	0.060	0.044	0.110
	3200	0.075	0.073	0.031	0.023	0.072
NP(10)	**200**	0.353	0.339	0.106	0.088	0.340
	800	0.177	0.172	0.060	0.044	0.168
	3200	0.089	0.086	0.031	0.023	0.084
OB-KM	**200**	0.102	0.150	0.125	0.088	0.143
	800	0.097	0.120	0.071	0.044	0.120
	3200	0.102	0.114	0.036	0.022	0.114
OB-KM-Pol3	**200**	0.123	0.150	0.118	0.088	0.138
	800	0.076	0.083	0.070	0.044	0.077
	3200	0.040	0.045	0.036	0.022	0.041

MAB and RMSE for the standardization and decomposition components of the RMST under DGP2 in Table 1.

Table 4 Comparison of alternative procedures under DGP3

MAE						
Method	**Sample size**	$\Delta_\mu^{(0,1)C}$	$\Delta_\mu^{(0,1)S}$	$\mu^{(0)}$	$\mu^{(1)}$	$\mu^{(0,1)}$
PML	**200**	1.221	1.112	0.104	0.102	1.174
	800	1.297	1.285	0.056	0.050	1.297
	3200	1.347	1.342	0.029	0.027	1.349
NP(3)	**200**	0.565	0.465	0.106	0.096	0.511
	800	0.569	0.562	0.056	0.049	0.568
	3200	0.599	0.596	0.029	0.026	0.600
NP(10)	**200**	0.212	0.176	0.106	0.096	0.180
	800	0.111	0.099	0.056	0.049	0.097
	3200	0.062	0.058	0.029	0.026	0.059
OB-KM	**200**	1.214	1.323	0.166	0.092	1.322
	800	1.295	1.422	0.125	0.052	1.406
	3200	1.347	1.416	0.068	0.027	1.410
OB-KM-Pol3	**200**	0.225	0.208	0.108	0.092	0.188
	800	0.145	0.173	0.059	0.055	0.156
	3200	0.140	0.164	0.033	0.030	0.155

RMSE						
Method	**Sample size**	$\Delta_\mu^{(0,1)C}$	$\Delta_\mu^{(0,1)S}$	$\mu^{(0)}$	$\mu^{(1)}$	$\mu^{(0,1)}$
PML	**200**	1.224	1.125	0.131	0.128	1.183
	800	1.298	1.289	0.070	0.063	1.300
	3200	1.348	1.343	0.036	0.033	1.350
NP(3)	**200**	0.590	0.496	0.134	0.120	0.536
	800	0.576	0.568	0.070	0.062	0.574
	3200	0.601	0.598	0.036	0.033	0.601
NP(10)	**200**	0.262	0.221	0.134	0.120	0.228
	800	0.139	0.124	0.070	0.062	0.121
	3200	0.077	0.071	0.036	0.033	0.072
OB-KM	**200**	1.217	1.336	0.206	0.115	1.332
	800	1.296	1.426	0.151	0.065	1.409
	3200	1.347	1.417	0.081	0.034	1.411
OB-KM-Pol3	**200**	0.247	0.244	0.138	0.114	0.219
	800	0.158	0.190	0.072	0.069	0.168
	3200	0.144	0.169	0.041	0.037	0.158

MAB and RMSE for the standardization and decomposition components of the RMST under DGP3 in Table 1.

Table 5 Effect of ignoring censoring

		MAE					
		No censoring			Censoring		
	Sample size	**200**	**800**	**3200**	**200**	**800**	**3200**
$\Delta^{(j,s)C}_{\ell_{\hat{\tau}Zn}}$	**Ignoring censoring**	0.071	0.035	0.017	0.152	0.147	0.147
	Dropping censored obs.	0.071	0.035	0.017	0.101	0.077	0.074
	OB-KM	0.071	0.035	0.017	0.086	0.047	0.023
$\Delta^{(j,s)S}_{\ell_{\hat{\tau}Zn}}$	**Ignoring censoring**	0.097	0.049	0.023	0.162	0.149	0.149
	Dropping censored obs.	0.097	0.049	0.023	0.130	0.085	0.075
	OB-KM	0.097	0.049	0.023	0.122	0.063	0.032
		RMSE					
		No censoring			**Censoring**		
	Sample size	**200**	**800**	**3200**	**200**	**800**	**3200**
$\Delta^{(j,s)C}_{\ell_{\hat{\tau}Zn}}$	**Ignoring censoring**	0.090	0.044	0.022	0.169	0.153	0.149
	Dropping censored obs.	0.090	0.044	0.022	0.123	0.088	0.077
	OB-KM	0.090	0.044	0.022	0.108	0.058	0.028
$\Delta^{(j,s)S}_{\ell_{\hat{\tau}Zn}}$	**Ignoring censoring**	0.121	0.062	0.029	0.191	0.160	0.152
	Dropping censored obs.	0.121	0.062	0.029	0.158	0.103	0.082
	OB-KM	0.121	0.062	0.029	0.153	0.078	0.040

MAB and RMSE for the decomposition components of the RMST under DGP1 in Table 1.

Table 6 Estimation of the RMST and the mean with $\tau_C^{(j)}$ fixed, $\tau_C^{(j)} < \tau_T^{(j)}$

	MAE					
		μ			μ_{τ^*}	
Sample size	**200**	**800**	**3200**	**200**	**800**	**3200**
$\Delta^{(0,1)C}_{\mu_{\tau^*}}$	0.081	0.076	0.076	0.013	0.007	0.003
$\Delta^{(0,1)S}_{\mu_{\tau^*}}$	0.101	0.052	0.026	0.095	0.052	0.025
$\mu^{(0,0)}_{\tau^*}$	0.072	0.038	0.024	0.067	0.035	0.017
$\mu^{(1,1)}_{\tau^*}$	0.112	0.094	0.094	0.077	0.041	0.020
$\mu^{(0,1)}_{\tau^*}$	0.104	0.094	0.095	0.063	0.033	0.016
	RMSE					
		μ			μ_{τ^*}	
Sample size	**200**	**800**	**3200**	**200**	**800**	**3200**
$\Delta^{(0,1)C}_{\mu_{\tau^*}}$	0.084	0.076	0.076	0.016	0.008	0.004
$\Delta^{(0,1)S}_{\mu_{\tau^*}}$	0.127	0.065	0.032	0.118	0.065	0.031
$\mu^{(0,0)}_{\tau^*}$	0.089	0.048	0.029	0.084	0.044	0.022
$\mu^{(1,1)}_{\tau^*}$	0.137	0.106	0.098	0.097	0.051	0.025
$\mu^{(0,1)}_{\tau^*}$	0.126	0.102	0.097	0.078	0.041	0.020

Mean absolute bias and RMSE for the decomposition components of the mean and the RMST, with $\tau^* = \tau_C^{(0)} = \tau_C^{(1)} = 6.5$ and $\tau_T^{(0)} = \tau_T^{(1)} = \infty$, using PML method. Simulations under DGP2 in Table 1.

Table 7 Descriptive statistics composition variables

Variable	Total	Women	Men	p-value
21.5in**Age**	39.727	40.014	39.218	2.4in0.236
	(0.322)	(0.381)	(0.585)	
21.5in**Houshold head**	0.323	0.237	0.476	2.4in0.000
	(0.012)	(0.014)	(0.022)	
21.5in**Tenure**	13.466	11.003	17.834	2.4in0.000
	(0.32)	(0.334)	(0.615)	
21.5in**Marital Status**	0.640	0.680	0.569	2.4in0.000
	(0.013)	(0.015)	(0.022)	
21.5in**High education (Dummy)**	0.182	0.193	0.162	2.4in0.000
	(0.01)	(0.013)	(0.016)	
21.5in**No. Unemployed**	1.129	1.120	1.145	2.4in0.212
	(0.01)	(0.012)	(0.016)	
21.5in**Big city**	0.330	0.316	0.354	2.4in0.136
	(0.012)	(0.015)	(0.021)	
21.5in**Medium size city**	0.244	0.236	0.260	2.4in0.140
	(0.011)	(0.014)	(0.019)	
No. Obs	1,473	942	531	

Standard erros in parenthesis.

Table 8 Standardized RMST

	$\hat{\tau}_{Zn} = 42$			$\tau^* = 36$			$\tau^* = 24$		
Method	**Women**	**Men**	$\mu^{(0,1)}_{\hat{\tau}_{Zn}}$	**Women**	**Men**	$\mu^{(0,1)}_{\tau^*}$	**Women**	**Men**	$\mu^{(0,1)}_{\tau^*}$
PML	10.721	8.090	10.275	10.398	7.886	10.016	9.463	7.258	9.220
NP(8)	10.864	7.804	10.831	10.594	7.596	10.476	9.726	7.043	9.545
OB-KM	10.864	7.804	10.427	10.594	7.596	10.094	9.726	7.043	9.193

RMST estimates for women and men for different τ^*. Women is Population 0 and men refers to Population 1.

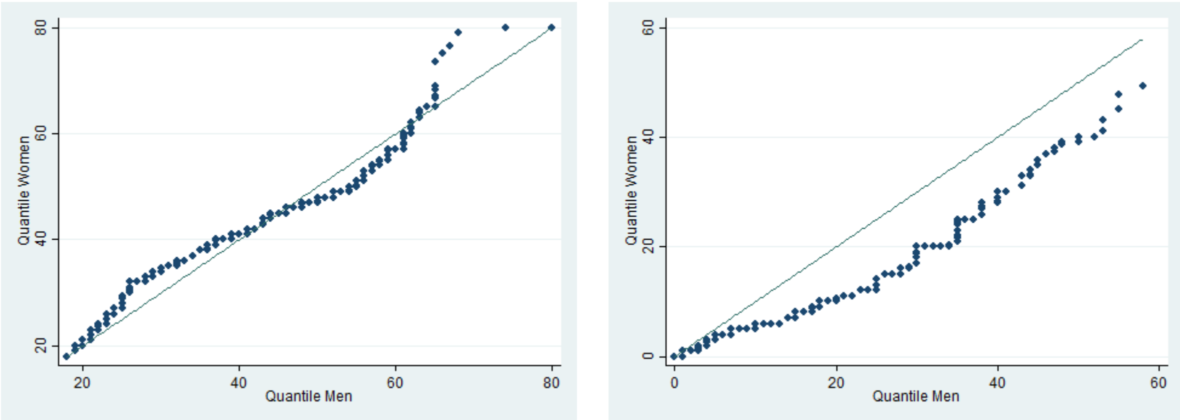
Table 9 Decomposition Components RMST

	$\hat{\tau}_{Z_n} = 42$			$\tau^* = 36$			$\tau^* = 24$		
Method	$\Delta^{(0,1)}_{\mu_{\hat{\tau}_{Zn}}}$	$\Delta^{(0,1)C}_{\mu_{\hat{\tau}_{Zn}}}$	$\Delta^{(0,1)S}_{\mu_{\hat{\tau}_{Zn}}}$	$\Delta^{(0,1)}_{\mu_{\tau^*}}$	$\Delta^{(0,1)C}_{\mu_{\tau^*}}$	$\Delta^{(0,1)S}_{\mu_{\tau^*}}$	$\Delta^{(0,1)}_{\mu_{\tau^*}}$	$\Delta^{(0,1)C}_{\mu_{\tau^*}}$	$\Delta^{(0,1)S}_{\mu_{\tau^*}}$
PML	2.631	0.446	2.185	2.512	0.382	2.130	2.204	0.242	1.962
NP(8)	3.060	0.033	3.027	2.997	0.118	2.880	2.683	0.180	2.502
OB-KM	3.060	0.438	2.622	2.997	0.499	2.498	2.683	0.533	2.150

Decomposition components of the RMST for different τ^*. Women is Population 0 and men refers to Population 1.

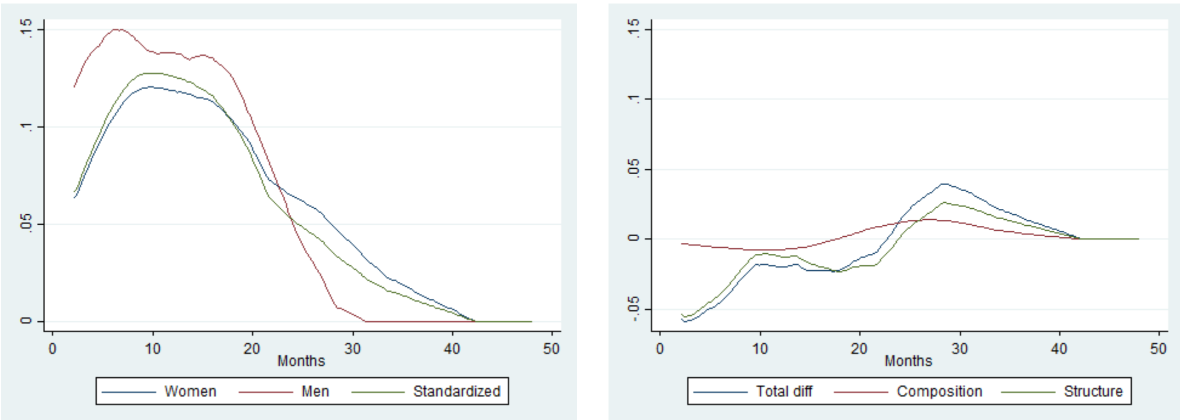
Figures

Figure 1. QQ plots for Age and Tenure by Gender



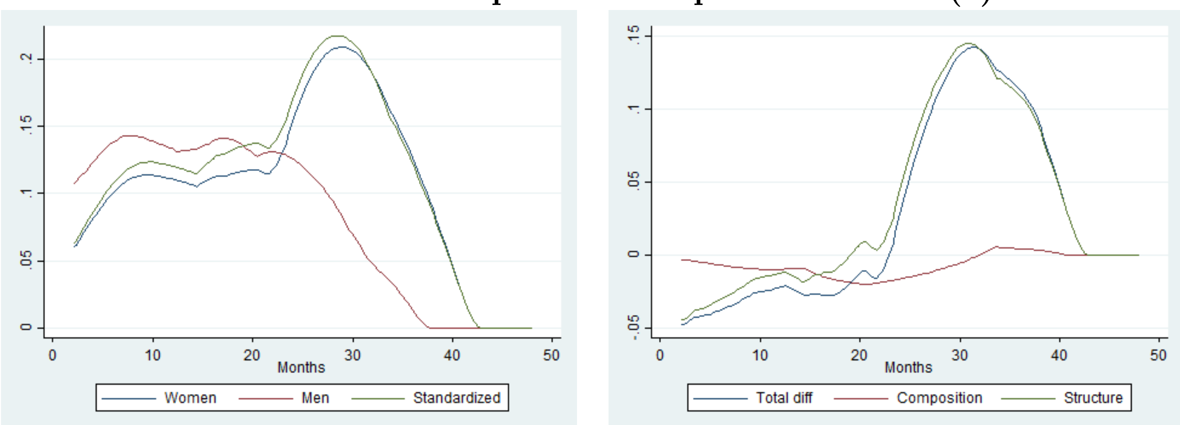
Age   Tenure

Figure 2. Standardization and Decomposition of the Averaged Hazard Function based on a Nonparametric Specification -NP(8)-



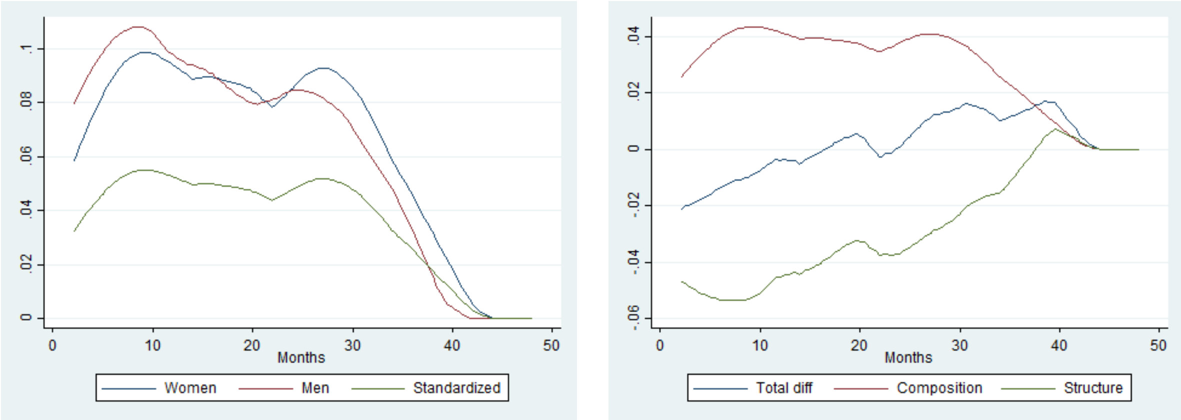
Averaged Conditional Hazard   Counterfactual Components

Figure 3. Standardization and Decomposition of the Marginal Hazard Function based on a Nonparametric Specification -NP(8)-



Marginal Hazard   Counterfactual Components

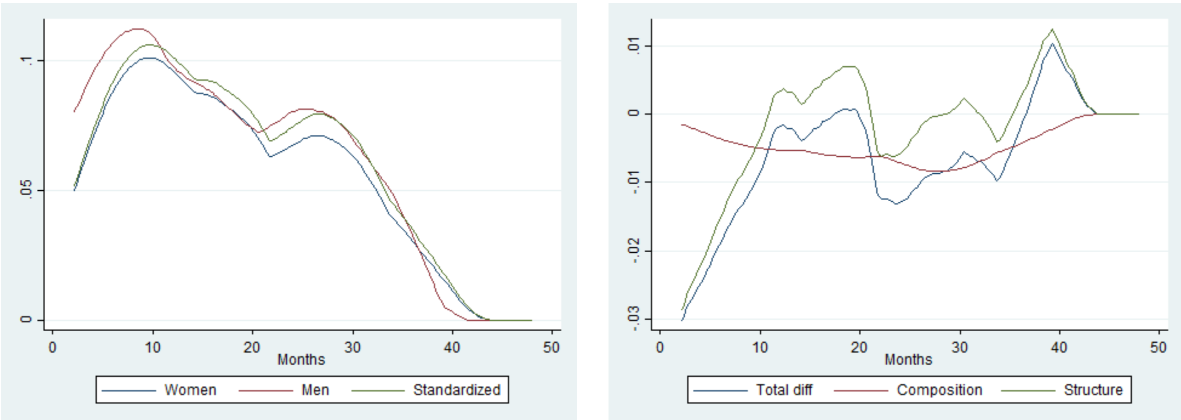
Figure 4. Standardization and Decomposition of the Averaged Hazard Function based on the PH Specification



Averaged Conditional Hazard

Counterfactual Components

Figure 5. Standardization and Decomposition of the Marginal Hazard Function based on the PH Specification



Marginal Hazard

Counterfactual Components