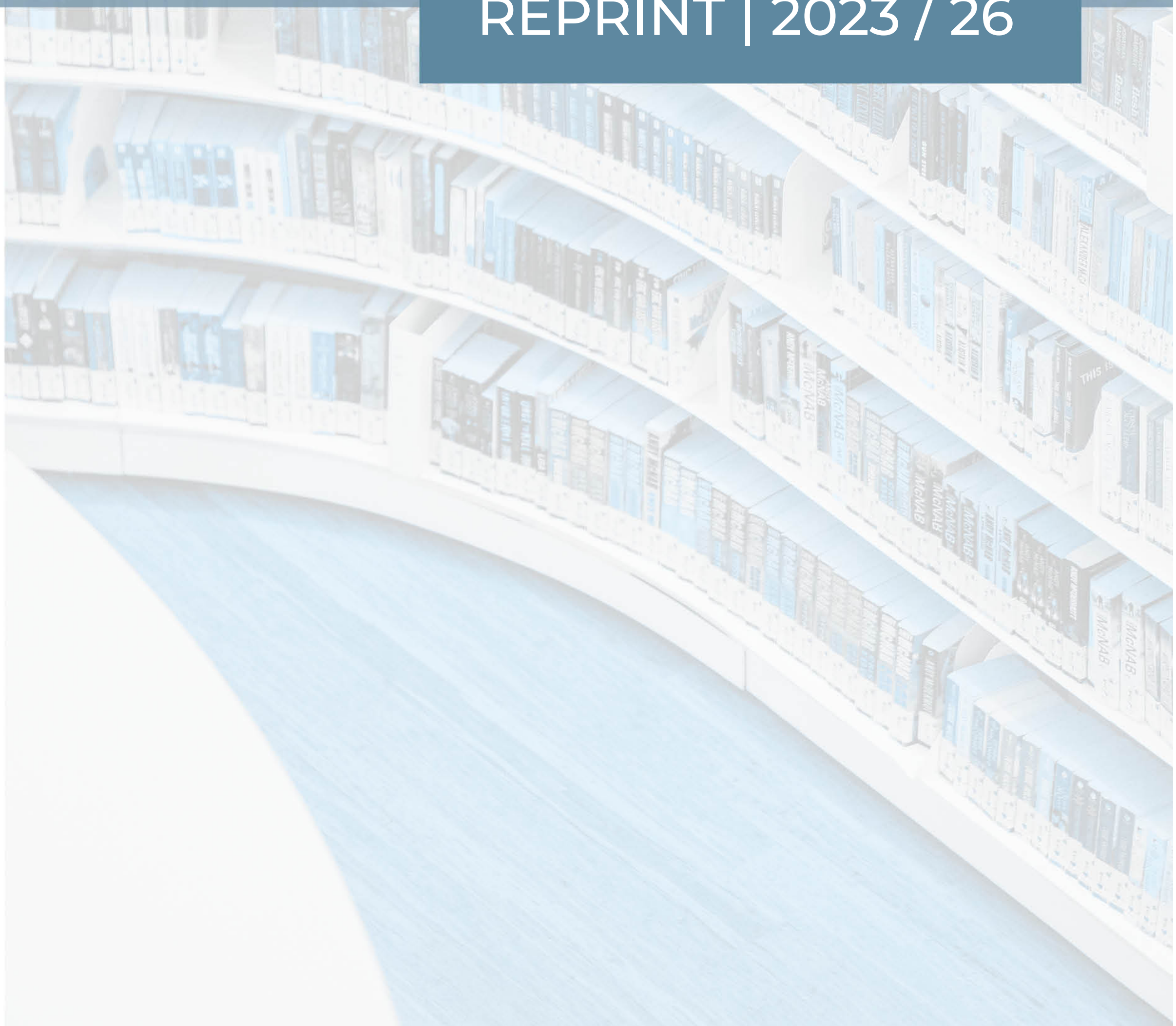
MORTALITY PROJECTIONS FOR HIGHER EDUCATIONAL ATTAINMENT WITH SEMI-PARAMETRIC ACCELERATED HAZARD RELATIONAL MODELS

Meitner Cadena, Michel Denuit

REPRINT | 2023 / 26



UCLouvain

LIDAM
ISBA | Louvain Institute of Data Analysis and Modeling in economics and statistics

ISBA
Voie du Roman Pays 20 - L1.04.01
B-1348 Louvain-la-Neuve
Email : lidam-library@uclouvain.be
https://uclouvain.be/en/research-institutes/lidam/isba/publication.html

Decisions in Economics and Finance (2023) 46:569–582
https://doi.org/10.1007/s10203-023-00391-4

Mortality projections for higher educational attainment with semi-parametric accelerated hazard relational models

Meitner Cadena[1] · **Michel Denuit**[2]

Received: 22 February 2021 / Accepted: 1 March 2023 / Published online: 25 March 2023

Abstract

The semiparametric accelerated hazard relational models proposed by Cadena and Denuit (Insurance Math. Econ. 68:1–16, 2016) are built from a reference life table by distorting the age scale. Individuals are made younger or older before performing the computations with the reference life table. In this paper, the mortality of individuals with higher educational attainment is modeled by means of an accelerated hazard relation specification. It is then projected by reference to the general population, using two different approaches. Numerical illustrations are performed on US mortality statistics.

Keywords Life tables · Accelerated hazard · Poisson regression · Mortality projection · Mortality differentials · Education

JEL Classification C14 · C01 · C02

1 Introduction and motivation

Relational methods are often used by demographers and actuaries to relate mortality in one population to that in others. A specific kind of relational model has been proposed by Cadena and Denuit (2016). Instead of transforming a set of reference death rates to fit a given mortality experience, as in the Cox proportional hazard model and its semi-parametric extensions, Cadena and Denuit (2016) adopted the accelerated hazard approach to induce a change in age scale between the forces of mortality. Specifically, the age scale is modified using a smooth unspecified age transform that is estimated by means of P(enalized)-splines techniques. This nonparametric approach avoids the specification of a rigid parametric form for the age transform, assuming only that this

✉ Meitner Cadena
mncadena2@espe.edu.ec

[1] Departamento de Ciencias Exactas, Universidad de las Fuerzas Armadas, Sangolquí, Ecuador

[2] Institut de Statistique, Biostatistique et Sciences Actuarielles - ISBA, Université Catholique de Louvain (UCL), Louvain-la-Neuve, Belgium

transform is a smooth function of age, i.e., it is continuous and possesses derivatives of any order.

Often, the analyst is interested in the mortality experience by a specific subgroup of the entire population. An insurance company, for instance, sells life insurance and life annuity products to individuals with higher socio-economic profile compared to the general population, resulting in improved longevity. Medical selection and product design tend to amplify this effect. Also, a forecast of mortality for the particular subgroup may be needed, such as a prediction of future longevity improvements for life annuitants. The same problem may be of interest for public policy or to assess mortality differentials inside the population, for instance. Theoretically speaking, the models used at the general population level could also be applied to the subgroup of interest. In practice, however, the smaller exposures and/or length of historical time series obscure the analysis. This is why different authors have proposed the so-called gravity approach (Cairns et al. 2011; Dowd et al. 2011) where the mortality dynamics experienced by the smaller group is linked to the general population one, helping to produce reliable, long-term forecasts. This is the approach followed in the present paper, in relation to the accelerated hazard specification. Other approaches for relating mortality of populations when measuring some risk have been proposed by, e.g., Villegas and Haberman (2014) and Villegas et al. (2017).

Precisely, the mortality of the general population and that of the sub-population corresponding to individuals with higher educational attainment are both modeled using the semi-parametric accelerated hazard specification proposed by Cadena and Denuit (2016). Then, two approaches to project the time components to the future are considered. A first analysis is based on a vector-autoregressive (VAR) model for each gender. Testing the different components for significance shows that the gravity construction applies, i.e., that the dynamics of the time component of the smaller population depends on the corresponding feature for the larger population. In the second approach, the modeling strategy takes advantage of the availability of a longer time series of observed mortality at the general population level. This produces more stable results since more data were available. Then, assuming a relationship similar to the gravity model, the dynamics of the time component of the smaller population is described by reference to the general population, the time component of which serving as explanatory variable, with ARMA error terms.

The remainder of the paper is organized as follows. In Sect. 2, we describe the semi-parametric accelerated hazard model proposed by Cadena and Denuit (2016). Section 3 then describes the problem of interest using data relating to the US population. The final Sect. 4 discusses the results and briefly concludes this paper.

2 Semiparametric accelerated hazard model

Let $\mu(x)$ be the force of mortality at age x in the population of interest and let $\mu^{\star}(x)$ be a reference force of mortality at the same age. The accelerated hazard model rescales the age axis inside the reference force of mortality, i.e.,

$$\mu(x) = \mu^{\star}(cx). \tag{2.1}$$

The parameter c alters the age scale of $\mu^\star$ and reflects the magnitude and direction of this alteration. The specification (2.1) is linear (on an appropriate scale). For applications, it appears to be useful to allow for nonlinear transformations in order to capture the particular features of mortality experience (that can be markedly nonlinear). This is why Cadena and Denuit (2016) proposed a semi-parametric accelerated hazard relational model to link the force or mortality μ for a population of interest, to a reference one $\mu^\star$ by modifying the age scale. A smooth unspecified function w to adjust the age at which $\mu^\star$ is computed, extending the linear specification of the standard accelerated hazard model. Specifically, the age scale is distorted by replacing (2.1) with

$$\mu(x) = \mu^\star\big(w(x)\big) \tag{2.2}$$

where the function $w : \mathbb{R}^+ \to \mathbb{R}^+$ acts on age x and specifies the age transformation to be applied to the reference mortality to reflect the actual mortality experience. In words, (2.2) means that the mortality specific to the group of interest at age x is the reference one at the corrected age $w(x)$.

Model (2.2) can be seen as a semiparametric extension of the accelerated hazard model where the age transformation w is no more linear but may be explicitly nonlinear. Such age corrections extend the age shifts that have been used so far in demography, which correspond to $w(x) = x - \delta$ where δ is the constant age correction accounting for the switch from the reference population to the one of interest. See, e.g., the shifting logistic model proposed by Bongaarts (2005). Model (2.2) extends this idea to general functions w, not necessarily linear.

Cadena and Denuit (2016) also proposed a dynamic version of the relational model (2.2) to produce mortality projections. The idea is to replace (2.2) with

$$\mu_t(x) = \mu^\star\big(w_t(x)\big) \tag{2.3}$$

where $\mu_t(x)$ is the force of mortality at age x during calendar year t and w_t is the age transformation for that year.

Empirical evidence in Cadena and Denuit (2016) suggests that $w_t(x)$ can be decomposed into the sum $w(x) + \theta_t$, i.e., into a static component $w(x)$ impacted by time-varying shifts θ_t, independent of age x. The model of interest then becomes

$$\mu_t(x) = \mu^\star\big(w(x) + \theta_t\big). \tag{2.4}$$

Specification (2.4) extends several approaches previously used by demographers. In the linear shift pattern of mortality, for instance, a baseline schedule of period mortality rates $\mu^\star$ is shifted every year along the age axis by a quantity r, i.e.,

$$\mu_t(x) = \mu^\star(x - rt)$$

which corresponds to $w(x) = x$ and a linearly decreasing θ_t in (2.4). See, e.g., Guillot and Kim (2011).

As in Brouhns et al. (2002), statistical inference is conducted by assuming that the observed number of deaths D_{xt} recorded at age x during calendar year t is Poisson distributed with mean $\text{ETR}_{xt}\mu_t(x)$, where ETR_{xt} is the corresponding (central) exposure-to-risk. The age transform w is estimated by Poisson regression for death counts, with P(enalized)-splines for w. P-splines combine the concept of B(asis)-splines and penalized likelihood. B-splines are made of polynomial pieces that are joined together at points called knots. In the P-spline approach, B-splines provide flexibility so that the fit to the available mortality statistics is accurate, while the penalty acts on the coefficients of B-splines to ensure that the resulting fit displays the assumed smooth behavior, see, e.g., Eilers and Marx (1996). The idea behind P-splines is to use a generous amount of equally spaced knots and to apply a penalty on adjacent coefficients to avoid over-fitting the data. We refer the interested reader to Cadena and Denuit (2016) for technical details.

In this paper, $\mu^\star$ corresponds to the limit life table proposed by Duchene and Wunsch (1988), i.e.,

$$\mu^\star(x) = \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1}$$

with $\alpha = 14.40198275$ and $\beta = 95$. The study performed by Cadena and Denuit (2016) has shown that the particular choice for $\mu^\star$ does not impact on the final output of the projection.

As small changes in θ_t may have an important impact on life expectancies, the θ_t are replaced with adjusted values so that the model exactly reproduces the observed period life expectancies at a given age for each calendar year. This adjustment procedure has been proposed by Lee and Miller (2001) for the model proposed by Lee and Carter (1992).

In this paper, we extend (2.4) to a sub-population of interest, in addition to the general population. Henceforth, we denote by "H" the sub-population of individuals having attained more than or equal to 13 years of study. The whole population is denoted by "P." Mortality in P and H is analyzed by age and time using the specification (2.4), separately for each gender. Specifically, the relations

$$\mu_{\text{P}t}(x) = \mu^\star\big(w_{\text{P}}(x) + \theta_{\text{P}t}\big) \tag{2.5}$$

$$\mu_{\text{H}t}(x) = \mu^\star\big(w_{\text{H}}(x) + \theta_{\text{H}t}\big), \tag{2.6}$$

are used for males and females, where P and H refer to the entire population and the subgroup with higher educational attainment, as explained previously. In (2.5)–(2.6), w_{P} and w_{H} capture the effects due to age, and the dynamic components $\theta_{\text{P}t}$ and $\theta_{\text{H}t}$ capture the effects due to calendar time. The latter are first estimated and then modeled as time series to forecast future mortality improvements. The interactions between the time series $\theta_{\text{P}t}$ and $\theta_{\text{H}t}$ drive the future evolution of the respective mortality experiences in P and H. This determines, for instance, the narrowing or widening of the mortality differentials between the two populations P and H. In the next section, we apply this approach to the US population.

3 Application to the US population

3.1 Available mortality statistics

We have at our disposal death counts together with exposures to risk by educational attainment for the US population. This country has been selected because it offers long series on mortality and exposed population according to educational attainment.

Information on decedents is registered in the US Standard Certificate of Death (SCD). Most of information of these certificates is publicly available in the Mortality Multiple Cause-of-Death Public Use Record (MMCDPUR) at http://www.cdc.gov/nchs/data_access/vitalstatsonline.htm. These records go from 1968 to 2018 (files accessed on January 27, 2021).

Unfortunately, inconsistencies on education declarations have been reported, e.g., by Sorlie and Johnson (1996). These authors conducted a study where education information from death certificates and that from the National Longitudinal Mortality Study (NLMS), a national longitudinal study on the effects of differentials in demographic and socio-economic characteristics on mortality, were related to each other, taking the latter information as baseline. This analysis was done for 10,433 individuals aged 25 or more who had died in 1989. These authors found that the informant would tend to assign a higher level on the death certificate than at baseline and this bias being more pronounced in older than younger decedents. Using data reported in Sorlie and Johnson (1996), noticing that when education categories are collapsed the bias in the true assignation of education grades reduces, scenarios with the lowest biases were identified. The scenario with only the two classes "$\leq$ 12 years" and "$\geq$ 13 years" was selected since this gives the lowest biases.

For estimating exposed populations, data from the Current Population Survey (CPS) were used, which is publicly available from the Integrated Public Use Microdata Series (IPUMS-CPS) that is a project dedicated to integrating and disseminating data from the CPS (see https://cps.ipums.org/cps-action/samples). See, e.g., Flood et al. (2015). This survey that collects monthly socio-economic information of households also includes respondents' educational attainment, as measured by the highest year of school or degree completed, available in the Annual Social and Economic Supplement (ASEC) of the CPS. An issue from the CPS is that data are aggregated for high ages, such aggregations varying from year to year. Hence, in order to have standardized data from 1989 to 2018, only disaggregated data for ages from 30 to 79 were considered.

Considering P, the same sources for deaths and exposed populations cited above are considered, to ensure consistency. With respect to deaths, the MMCDPUR provides information since 1968, but registers show inconsistencies from 1968 to 1978. Due to this fact, the data period for death information was restricted to 1979 to 2018. With respect to exposed populations, the CPS also gives details for each age between 30 and 79 for the period 1979 to 1988.

3.2 First analysis: joint VAR dynamics for θ_P and θ_H

The observation period common to P and H starts in 1989 to end in 2018. The dynamic version of the model is used here to produce mortality forecasts. The yearly age transformations w_t are decomposed into the sum of a time-independent w subject to annual shifts according to a time index θ_t. The time series are displayed in Fig. 1.

Let θ_{fKt} (resp. θ_{mKt}) be θ_K in year t associated to the female (resp. male) population $K \in \{P, H\}$. The estimated VAR equations are displayed next, with standard errors on coefficients appearing between brackets:

$$\theta_{fPt} = -0.0065_{(0.0391)} + 0.8892_{(0.0748)}\theta_{fP(t-1)} + 0.0106_{(0.1252)}\theta_{fH(t-1)} + \mathcal{E}_{fPt}$$
$$\theta_{fHt} = -0.2325_{(0.0654)} + 0.0670_{(0.0462)}\theta_{fP(t-1)} + 0.9060_{(0.0773)}\theta_{fH(t-1)} + \mathcal{E}_{fHt}$$

The error terms $\mathcal{E}_{fPt}$ and $\mathcal{E}_{fHs}$ are mutually independent random variables if $t \neq s$, with zero mean and variance σ^2_{fP} and σ^2_{fH}, respectively. For each t, $\mathcal{E}_{fPt}$ and $\mathcal{E}_{fHt}$ are possibly correlated. For males, we get

$$\theta_{mPt} = -0.0410_{(0.0371)} + 0.7250_{(0.0980)}\theta_{mP(t-1)} - 0.1050_{(0.1240)}\theta_{mH(t-1)} + \mathcal{E}_{mPt}$$
$$\theta_{mHt} = -0.2117_{(0.0471)} + 0.1790_{(0.0698)}\theta_{mP(t-1)} + 1.0090_{(0.0885)}\theta_{mH(t-1)} + \mathcal{E}_{mHt}$$

where $\mathcal{E}_{mPt}$ and $\mathcal{E}_{mHs}$ are mutually independent random variables if $t \neq s$, with zero mean and variance σ^2_{mP} and σ^2_{mH}, respectively.

Considering the large standard errors on the coefficient of $\theta_{fH(t-1)}$ (resp. $\theta_{mH(t-1)}$) in the dynamics of $\theta_{fP(t-1)}$ (resp. $\theta_{mP(t-1)}$), this effect does not seem to be significant. This is why we simplify the model by assuming that the secondary population H depends on the principal population P, but the principal population does not depend on the secondary population. The estimated equations for females then become

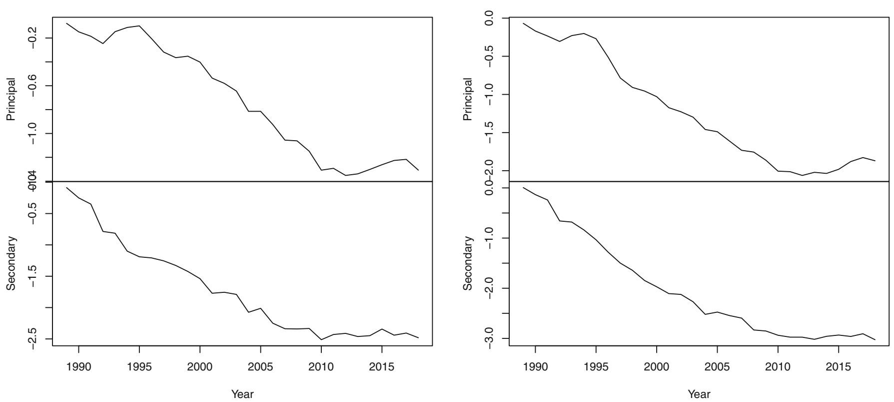


Fig. 1 Original time series for θ_P (principal population) and θ_H (secondary population). Females on the left, males on the right

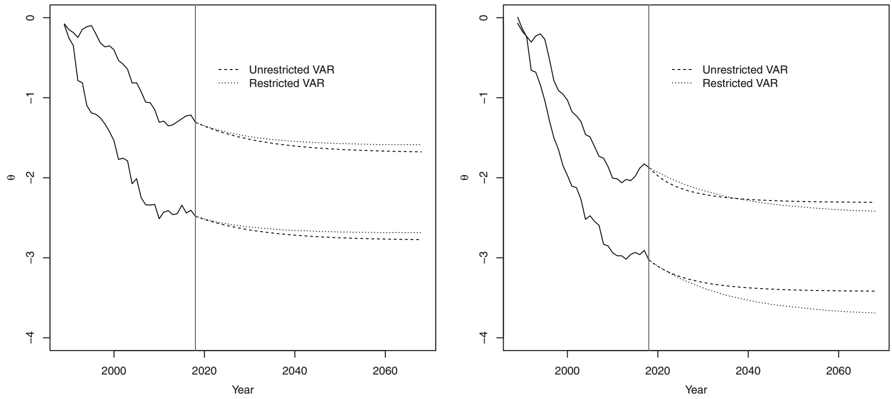


Fig. 2 Forecasts of θ_{fPt} and θ_{fHt} on the left, of θ_{mPt} and θ_{mHt} on the right

$$
\begin{aligned}
\theta_{fPt} &= 0.8769_{(0.0486)}\theta_{fP(t-1)} + \mathcal{E}_{fPt} \\
\theta_{fHt} &= -0.2251_{(0.0361)} + 0.0755_{(0.0228)}\theta_{fP(t-1)} + 0.9136_{(0.0213)}\theta_{fH(t-1)} + \mathcal{E}_{fHt}
\end{aligned}
$$

with contemporaneous covariance matrix of error terms $\mathcal{E}_{fPt}$ and $\mathcal{E}_{fHt}$ estimated to

$$
\widehat{\mathbf{\Sigma}}_f = \begin{pmatrix} 0.00416 & 0.00376 \\ 0.00376 & 0.01162 \end{pmatrix}.
$$

For males, we get

$$
\begin{aligned}
\theta_{mPt} &= 0.7059_{(0.0785)}\theta_{mP(t-1)} + \mathcal{E}_{mPt} \\
\theta_{mHt} &= -0.2089_{(0.0305)} + 0.2081_{(0.0485)}\theta_{mP(t-1)} + 0.9462_{(0.0148)}\theta_{mH(t-1)} + \mathcal{E}_{mHt}
\end{aligned}
$$

with contemporaneous covariance matrix of error terms $\mathcal{E}_{mPt}$ and $\mathcal{E}_{mHt}$ estimated to

$$
\widehat{\mathbf{\Sigma}}_m = \begin{pmatrix} 0.00545 & 0.00347 \\ 0.00347 & 0.00809 \end{pmatrix}.
$$

The deletion of the unsignificant effects greatly impacts on the forecast for θ, mainly for males, as it can be seen in Fig. 2. We can see there that omitting the terms in the VAR dynamics that are not significant generates faster mortality improvements in both populations. The careful analysis of the VAR specification thus appears to be an important part of the analysis.

3.3 Second analysis: ARIMA for θ_P and regression with ARMA errors for θ_H

In the second approach, we first fit the model (2.5) for the whole population P, in order to benefit from a longer time series. Here, we use mortality statistics available over

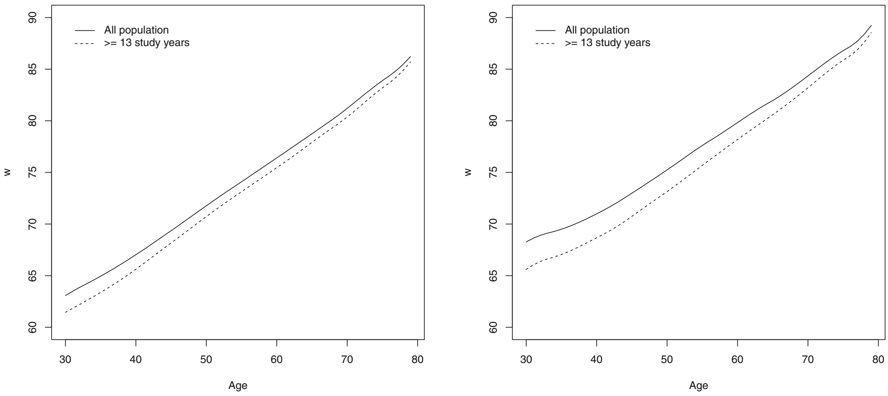


Fig. 3 Estimated w by gender, all population. Females on left, males on right

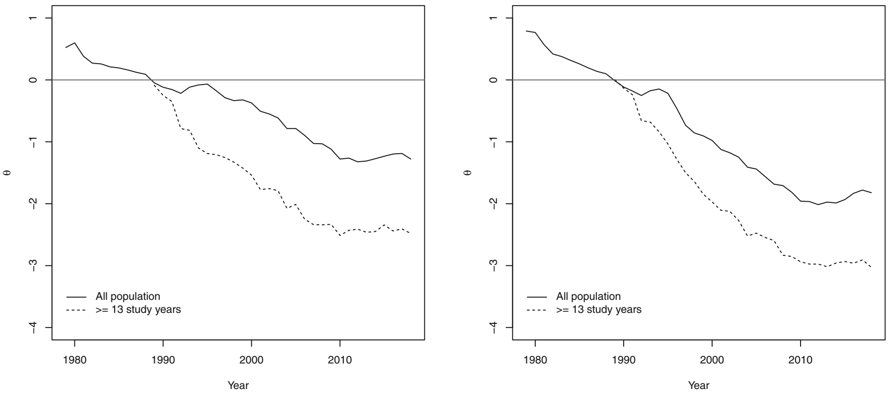


Fig. 4 Estimated θ by gender, all population. Females on left, males on right

the period 1979–2015. Then, we fit (2.6) over the restricted time period 1989–2015. Figures 3 and 4 display the estimated w and θ, respectively.

Different ARIMA models are then considered for describing the dynamics of θ_P over the period 1979–2015. Table 1 presents ARIMA fits to θ_P for $p, q \in \{0, 1\}$, where the goodness-of-fit Akaike information criterion (AIC), corrected AIC (AICc, see Hurvich and Tsai 1993), and Bayesian information criterion (BIC) are included. According to the results reported in this table, for females, the model ARIMA(0,1,0) appears to be optimal on all three penalized likelihood criteria considered, whereas for males such a selection changes to the model ARIMA(0,1,1). They are therefore retained for the remainder of the analysis. The corresponding forecasts are visible in Fig. 5.

Next, the time series of $\theta_{\mathrm{P}t}$ formed by the observed and the projected ones are used to describe $\theta_{\mathrm{H}t}$ using the following regression with ARMA errors:

$$\theta_{\mathrm{H}t} = \alpha\theta_{\mathrm{P}t} + \Delta_t \tag{3.1}$$

$$\Delta_t = c + \phi_1\Delta_{t-1} + \cdots + \phi_p\Delta_{t-p} - \xi_1\mathcal{E}_{t-1} - \cdots - \xi_q\mathcal{E}_{t-q} + \mathcal{E}_t, \tag{3.2}$$

Table 1 Summary of the ARIMA(p, 1, q) fits to θ_P for $p, q \in \{0, 1\}$

ARIMA order	Estimated dynamics	AIC	AICc	BIC
Females				
(0, 1, 0)	$\theta_{fPt} = -0.0488_{(0.0117)} + \theta_{fP(t-1)} + \epsilon_t$	– 85.04	– 84.68	– 81.88
(1, 1, 0)	$\theta_{fPt} = -0.0486_{(0.0122)} + 0.0342_{(0.1765)}\theta_{fP(t-1)} + \epsilon_t$	– 83.08	– 82.33	– 78.33
(0, 1, 1)	$\theta_{fPt} = -0.0486_{(0.0123)} + \epsilon_t + 0.0444_{(0.2053)}\epsilon_{t-1}$	– 83.09	– 82.34	– 78.34
(1, 1, 1)	$\theta_{fPt} = -0.0480_{(0.0127)} - 0.5879_{(0.3231)}\theta_{fP(t-1)} + \epsilon_t + 0.7521_{(0.2488)}\epsilon_{t-1}$	– 82.11	– 80.82	– 75.78
Males				
(0, 1, 0)	$\theta_{mPt} = -0.0758_{(0.0125)} + \theta_{mP(t-1)} + \epsilon_t$	– 80.53	– 80.16	– 77.36
(1, 1, 0)	$\theta_{mPt} = -0.0730_{(0.0180)} + 0.3557_{(0.1617)}\theta_{mP(t-1)} + \epsilon_t$	– 83.02	– 82.27	– 78.27
(0, 1, 1)	$\theta_{mPt} = -0.0730_{(0.0167)} + \epsilon_t + 0.4674_{(0.1790)}\epsilon_{t-1}$	– 84.34	– 83.59	– 79.58
(1, 1, 1)	$\theta_{fPt} = -0.0732_{(0.0161)} - 0.1051_{(0.3981)}\theta_{fP(t-1)} + \epsilon_t + 0.5585_{(0.3601)}\epsilon_{t-1}$	– 82.40	– 81.11	– 76.07

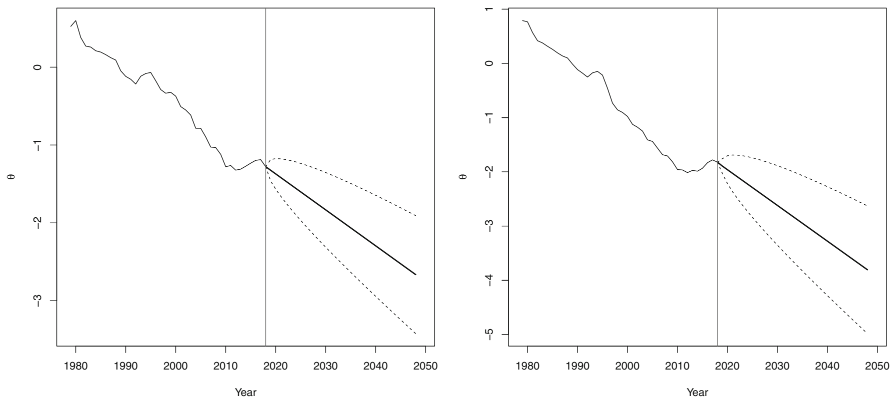


Fig. 5 Forecasts of θ_P. Females appear on the left, males on the right

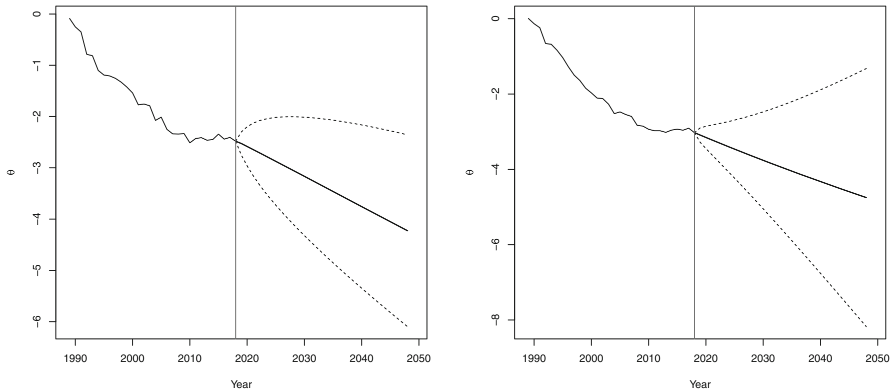


Fig. 6 Projection of $\theta_{\mathrm{H}t}$ from ARIMA(1, 1, 0) for females and ARIMA(1, 1, 1) for males, all population. Females on left, males on right

where Δ_t is an unknown process with ARMA(p,q) or ARMA(p,1,q) dynamics, and $\mathcal{E}_t$ is a white noise process, i.e., the $\mathcal{E}_t$ have zero mean and are independent and identically distributed. Constant c is introduced if it is detected that the auto-regressive part is non-stationary.

In this case, also several models (3.1)–(3.2) were assessed. Outputs of these models are presented in Table 2. As the rankings induced by the respective AIC, AICc and BIC values closely agree, we only reported the AIC. Considering the models with coefficients that are statistically significant, the model with the lowest AIC was selected, which corresponds, for females to the ARIMA(1, 1, 0) and for males to the ARIMA(1, 1, 1). Figure 6 presents forecasts of $\theta_{\mathrm{H}t}$ for an horizon of 30 years too, using the optimal model.

3.4 Resulting life expectancies

As an illustration, we compute the forecasts of period life expectancies, at age 25. Notice that we have to limit our calculations to age 80. Denoting as T_{25} the time to

Table 2 Summary of the regressions for $\theta_{\mathrm{H}t}$ with ARIMA$(p, 1, q)$ errors $\mathcal{E}_t$, for $p, q \in \{0, 1\}$, where $D\Delta_t$ denotes the first-order difference $\Delta_t - \Delta_{t-1}$

Estimated dynamics	AIC
Females	
$\theta_{fHt} = 1.2557_{(0.3006)}\theta_{fPt} + \Delta_t$	
$\Delta_t = -0.5930_{(0.3676)} + 0.9553_{(0.0523)}\Delta_{t-1} + \mathcal{E}_t$	−26.78
$\theta_{fHt} = 1.1238_{(0.2941)}\theta_{fPt} + \Delta_t$	
$\Delta_t = \Delta_{t-1} + \mathcal{E}_t$	−31.47
$\theta_{fHt} = 1.5260_{(0.1385)}\theta_{fPt} + \Delta_t$	
$\Delta_t = -0.6747_{(0.1075)} + \epsilon_t - 0.8297_{(0.1025)}\mathcal{E}_{t-1}$	−5.22
$\theta_{fHt} = 1.1611_{(0.2429)}\theta_{fPt} + \Delta_t$	
$D\Delta_t = 0.3939_{(0.1804)}D\Delta_{t-1} + \mathcal{E}_t$	−34.34
$\theta_{fHt} = 1.3595_{(0.3161)}\theta_{fPt} + \Delta_t$	
$D\Delta_t = \mathcal{E}_t + 0.4073_{(0.1856)}\mathcal{E}_{t-1}$	−33.67
$\theta_{fHt} = 1.2407_{(0.3944)}\theta_{fPt} + \Delta_t$	
$D\Delta_t = 0.6029_{(0.3633)}D\Delta_{t-1} + \mathcal{E}_t + 0.2510_{(0.4600)}\mathcal{E}_{t-1}$	−32.05
Males	
$\theta_{mHt} = 1.1885_{(0.2842)}\theta_{mPt} + \Delta_t$	
$\Delta_t = -0.4705_{(0.2964)} + 0.9341_{(0.1141)}\Delta_{t-1} + \mathcal{E}_t$	−32.32
$\theta_{mHt} = 1.0148_{(0.1963)}\theta_{mPt} + \Delta_t$	
$\Delta_t = \Delta_{t-1} + \mathcal{E}_t$	−36.89
$\theta_{mHt} = 1.6423_{(0.0506)}\theta_{mPt} + \Delta_t$	
$\Delta_t = \mathcal{E}_t - 1.0000_{(0.2209)}\mathcal{E}_{t-1}$	−8.71
$\theta_{mHt} = 0.9431_{(0.2366)}\theta_{mPt} + \Delta_t$	
$D\Delta_t = 0.2101_{(0.2025)}D\Delta_{t-1} + \mathcal{E}_t$	−35.97
$\theta_{mHt} = 0.9826_{(0.2132)}\theta_{mPt} + \Delta_t$	
$D\Delta_t = \mathcal{E}_t - 0.1509_{(0.1680)}\mathcal{E}_{t-1}$	−35.67
$\theta_{mHt} = 0.5169_{(0.2339)}\theta_{mPt} + \Delta_t$	
$D\Delta_t = 0.9770_{(0.0369)}D\Delta_{t-1} + \mathcal{E}_t - 0.7952_{(0.1105)}\mathcal{E}_{t-1}$	−39.28

death for an individual aged $x = 25$ subject to the mortality corresponding to calendar year t, this means that we use the formula

$$\begin{aligned}\mathrm{E}[\min\{T_{25}, 80-25\}] &= \mathrm{E}[\min\{T_{25}, 55\}|T_{25} < 55]\mathrm{P}[T_{25} < 55] + 55\mathrm{P}[T_{25} \geq 55] \\ &= \int_0^{55} \exp\left(-\int_0^{\xi} \mu_{25+\eta}(t)d\eta\right) d\xi \mathrm{P}[T_{25} < 55] \\ &\quad +55\mathrm{P}[T_{25} \geq 55]. \end{aligned} \tag{3.3}$$

The calculations of forecasts for these remaining life expectancies are performed as follows:

1. For the whole population P, using forecasts of θ_P, we compute the force of mortality μ_P for each year from 2019 to 2048 and each age from 30 to 79. Next, the

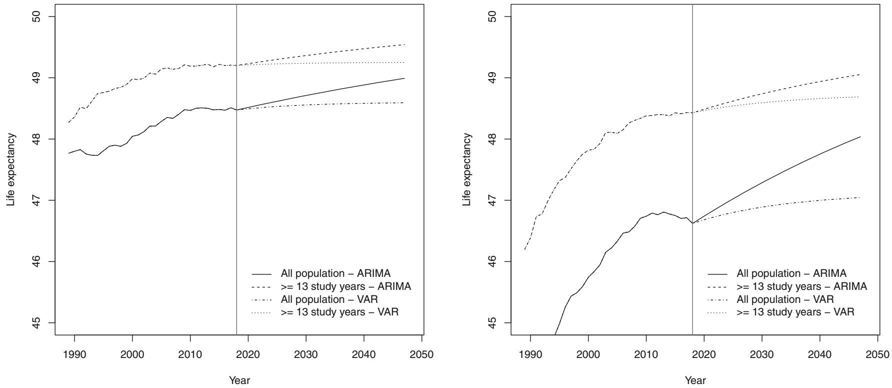


Fig. 7 Life expectancy at age 30 (truncated at age 80). Females on left, males on right

formula (3.3) is applied for each year, keeping t fixed as we work with period life expectancies.

2. For the sub-population with higher education H, using forecasts of θ_H, we compute the force of mortality μ_H for each year from 2019 to 2048 and each age from 30 to 79. Next, the formula (3.3) is applied for each year.

The results are displayed in Fig. 7. We can see there that the obtained forecasts greatly vary according to the model used for projecting θ_P and θ_H, especially for males.

Longevity improvements predicted by VAR appear to be smaller compared to the gains projected with the second approach. Another difference concerns the mortality differentials. Whereas VAR models produce a stable gap between the general population and the group with higher educational attainment, the second approach predicts a narrowing of these mortality differentials, more markedly for males.

The VAR models for the pair θ_P and θ_H have better AIC values, with AIC: -99.55503 and AICc: -95.46412 for females and AIC: -98.07383 and AICc: -93.98292 for males. Moreover, we can see in Fig. 7 that this specification predicts lower values for life expectancies in the future. This is in line with statistics released by the National Center for Health Statistics (NCHS) showing higher death rates and lower life expectancies compared to preceding years. Standard life expectancy at birth stopped increasing in 2015. Considering these recent trends, the VAR specification seems to be preferable to its competitor.

4 Discussion

In this paper, we assessed the impact of higher educational attainment on mortality by means of the semi-parametric accelerated hazard relational model proposed by Cadena and Denuit (2016). To this end, we use death statistics by education level from the US population.

A first analysis represents θ_P and θ_H appearing in (2.5)–(2.6) by means of a VAR model for each gender. The large estimated standard errors on some of the coefficients of this model suggest that the model can be further simplified. This leads to a restricted VAR model where the behavior of θ_H was described using the behavior of θ_P in line with the gravity construction recalled in the introduction: In the gravity approach, a feature of a small population is modeled using the corresponding feature of a larger population. In the second approach, the modeling strategy takes advantage of the availability of a longer time series of observed mortality for P. This produces more stable results since more data were available. Then, assuming the relationship similar to the gravity model between θ_P and θ_H, an AR model for θ_H incorporating θ_P as an explanatory variable was proposed, with ARMA error terms. In this last model, the dynamics of θ_P was represented by a random walk with drift for females, and by a moving average with drift for males.

Life expectancies are projected under both approaches. The results greatly vary in terms of longevity gains, which appear to be more pronounced with the second approach, as well as in terms of mortality differentials, which tend to remain at their current level with the first approach but narrow in the second approach. Based on the last statistics published in the USA, the VAR specification seems to deliver the most reasonable forecast.

The procedures developed in this paper have shown abilities to put in evidence differences on mortality behaviors among populations due to education levels of individuals. These kinds on impacts are important for populations since they involve transcendent human conditions as health and life quality. Thus, these results may be used for instance to configure strategies addressed to guarantee similar life conditions among individuals of regions or countries. Hence, we plan to apply these procedures to other factors affecting human mortality such as health access, incomes, work available and type of work among others. Furthermore, these results may be combined with recent measures of mortality differences as the ones proposed by Cadena and Denuit (2021).

References

Bongaarts, J.: Long-range trends in adult mortality: models and projection methods. Demography **42**, 23–49 (2005)

Brouhns, N., Denuit, M., Vermunt, J.: A Poisson log-bilinear approach to the construction of projected life tables. Insurance Math. Econ. **31**, 373–393 (2002)

Cadena, M., Denuit, M.: Semi-parametric accelerated hazard relational models with applications to mortality projections. Insurance Math. Econ. **68**, 1–16 (2016)

Cadena, M., Denuit, M.: A new measure of mortality differentials based on precedence probability. Eur. Actuarial J. **11**, 717–724 (2021)

Cairns, A.J.G., Blake, D., Dowd, K., Coughlan, G.D., Khalaf-Allah, M.: Bayesian stochastic mortality modelling for two populations. Astin Bull. **41**, 29–59 (2011)

Dowd, K., Cairns, A.J.G., Blake, D., Coughlan, G.D., Khalaf-Allah, M.: A gravity model of mortality rates for two related populations. North Am. Actuarial J. **15**, 334–356 (2011)

Duchene, J., Wunsch, G.: From the demographer's cauldron: single-decrement life tables and the span of life. Genus **44**, 1–17 (1988)

Eilers, P.H.C., Marx, B.D.: Flexible smoothing with B-splines and penalties. Stat. Sci. **11**, 89–121 (1996)

Flood, S., King, M., Ruggles, S., Warren, J.R.: Integrated public use microdata series, current population survey: version 5.0 [dataset]. Minneapolis, MN: University of Minnesota (2015)

Guillot, M., Kim, H.S.: On the correspondence between CAL and lagged cohort life expectancy. Demogr. Res. **24**, 611–632 (2011)

Hurvich, C.M., Tsai, C.-L.: A corrected Akaike information criterion for vector autoregressive model selection. J. Time Ser. Anal. **14**, 271–279 (1993)
Lee, R.D., Carter, L.: Modelling and forecasting the time series of US mortality. J. Am. Stat. Assoc. **87**, 659–671 (1992)
Lee, R., Miller, T.: Evaluating the performance of the Lee-Carter approach to modelling and forecasting. Demography **38**, 537–549 (2001)
Sorlie, P.D., Johnson, N.J.: Validity of education in formation on the death certificate. Epidemiology **7**, 437–439 (1996)
Villegas, A.M., Haberman, S.: On the modeling and forecasting of socioeconomic mortality differentials: an application to deprivation and mortality in England. North Am. Actuarial J. **18**, 168–193 (2014)
Villegas, A.M., Haberman, S., Kaishev, V.K., Millossovich, P.: A comparative study of two population models for the assessment of basis risk in longevity hedges. ASTIN Bull. **47**, 631–679 (2017)