Online Appendix for
"Inequality in Life Expectancies across Europe and the US"

Radim Boháček
CERGE-EI

Jesús Bueren
EUI

Laura Crespo
Banco de España

Pedro Mira
CEMFI

Josep Pijoan-Mas
CEMFI and *CEPR*

Contents

Appendix A: Funding acknowledgements

Radim Boháček acknowledges funding from the Ministry of Education, Youth, and Sports of the Czech Republic, grant LM2015052 and the SHARE-CZ+ project; Pedro Mira acknowledges support from grant ECO2014-57768-P from the Spanish *Ministerio de Economía y Competitividad*; Josep Pijoan-Mas acknowledges funding from from the Spanish *Ministerio de Educación* through the 2017 *Salvador de Madariaga* Mobility Grant and the hospitality of UCL Economics Department. The developers and funders of ELSA, HRS, and SHARE do not bear any responsibility for the analyses or interpretations presented here.

We use data from SHARE Waves 1, 2, 3 (SHARELIFE), 4, 5 and 6, see Börsch-Supan et al. (2013) for methodological details.[1] The SHARE data collection has been primarily funded by the European Commission through FP5 (QLK6-CT-2001-00360), FP6 (SHARE-I3: RII-CT-2006-062193, COMPARE: CIT5-CT-2005-028857, SHARELIFE: CIT4-CT-2006-028812) and FP7 (SHARE-PREP: N211909, SHARE-LEAP: N227822, SHARE M4: N261982). Additional funding from the German Ministry of Education and Research, the Max Planck Society for the Advancement of Science, the U.S. National Institute on Aging (U01_AG09740-13S2, P01_AG005842, P01_AG08291, P30_AG12815, R21_AG025169, Y1-AG-4553-01, IAG_BSR06-11, OGHA_04-064, HHSN271201300071C) and from various national funding sources is gratefully acknowledged (see www.share-project.org).

ELSA data were made available through the UK Data Archive. ELSA was developed by a team of researchers based at the NatCen Social Research, University College London and the Institute for Fiscal Studies. The data were collected by NatCen Social Research. The funding is provided by the National Institute of Aging in the United States, and a consortium of UK government departments coordinated by the Office for National Statistics.

Appendix B: Estimation

In order to reduce the uncertainty of estimated parameters from surveys with a small number of observed transitions, we rely on Bayesian techniques by constraining the space of possible $\boldsymbol{\beta}$ to satisfy a set of five regularity conditions $r_1(\boldsymbol{\beta}|a)$ to $r_5(\boldsymbol{\beta}|a)$ that we re-write

[1] (DOIs: 10.6103/SHARE.w1.610, 10.6103/SHARE.w2.610, 10.6103/SHARE.w3.610, 10.6103/SHARE.w4.610, 10.6103/SHARE.w5.610, 10.6103/SHARE.w6.610)

as a prior for $\boldsymbol{\beta}$ with pdf:

$$p(\beta) = \prod_{a=50}^{\bar{a}} r_1(\boldsymbol{\beta}|a) \cdot r_2(\boldsymbol{\beta}|a) \cdot r_3(\boldsymbol{\beta}|a) \cdot r_4(\boldsymbol{\beta}|a) \cdot r_5(\boldsymbol{\beta}|a) \tag{B.1}$$

These five regularity conditions are:

$$r_1(\boldsymbol{\beta}|a) = \begin{cases} 1 & \text{if } \dfrac{p_{22}(a)}{1-p_{20}(a)} \geq \dfrac{p_{22}(a+1)}{1-p_{20}(a+1)}, \\ 0 & \text{otherwise} \end{cases} \tag{B.2}$$

$$r_2(\boldsymbol{\beta}|a) = \begin{cases} 1 & \text{if } \dfrac{p_{12}(a)}{1-p_{10}(a)} \geq \dfrac{p_{12}(a+1)}{1-p_{10}(a+1)}, \\ 0 & \text{otherwise} \end{cases} \tag{B.3}$$

$$r_3(\boldsymbol{\beta}|a) = \begin{cases} 1 & \text{if } p_{20}(a+1) \geq p_{20}(a), \\ 0 & \text{otherwise} \end{cases} \tag{B.4}$$

$$r_4(\boldsymbol{\beta}|a) = \begin{cases} 1 & \text{if } p_{10}(a+1) \geq p_{10}(a), \\ 0 & \text{otherwise} \end{cases} \tag{B.5}$$

$$r_5(\boldsymbol{\beta}|a) = \begin{cases} 1 & \text{if } p_{10}(a) \geq p_{20}(a), \\ 0 & \text{otherwise} \end{cases} \tag{B.6}$$

and allow us to restrict the parameter space such that: conditional on surviving, the probability of remaining in good health decreases with age, equation (B.2); conditional on surviving, the probability of moving from bad to good health decreases with age, equation (B.3); the probability of surviving (conditional on both good and bad health) decreases with age, equations (B.4) and (B.5); and the probability of dying is larger when in bad health than in good health, equation (B.6).

The posterior distribution of $\boldsymbol{\beta}$ is

$$p(\boldsymbol{\beta}|H) \propto p(H|\boldsymbol{\beta}) \cdot p(\boldsymbol{\beta}), \tag{B.7}$$

In order to sample from the posterior distribution, we use Markov Chain Monte-Carlo (MCMC) methods with a random-walk Metropolis algorithm:

1. Initialize at $\boldsymbol{\beta}^{t=0}$
2. Propose candidate $\boldsymbol{\beta}^c = \boldsymbol{\beta}^t + \epsilon$, where $\epsilon \sim N(0, \sigma_\epsilon^2)$
3. Accept $\boldsymbol{\beta}^c$ with probability:

$$\alpha(\boldsymbol{\beta}^c|\boldsymbol{\beta}^t) = \min\left\{1, \frac{p(\boldsymbol{\beta}^c|H)}{p(\boldsymbol{\beta}^t|H)}\right\}$$

4. If candidate is accepted $\boldsymbol{\beta}^{t+1} = \boldsymbol{\beta}^c$, otherwise $\boldsymbol{\beta}^{t+1} = \boldsymbol{\beta}^t$.
5. Set t=t+1 and go back to 2 until convergence in the posterior distribution.

The empirical results are based on 500,000 draws for each sample. The first 40,000 draws are disregarded as burn-in and the remaining 460,000 provide a posterior distribution of the vector of parameters $\boldsymbol{\beta}$ for each country-gender-education sample. We adjust the variance σ_ε^2 of the proposal for every country-gender-education sample to ensure acceptance rates are around 30%.

Appendix C: Building life expectancies

In this Appendix we explain how we compute the life expectancy, the healthy life expectancy and the unhealthy life expectancy from our estimated multi-state life tables. Given the parameter estimates, we recover $p_{ij}(a)$, the probability that an individual with health $i \in \{1, 2\}$ transits into health $j \in \{0, 1, 2\}$ from age a to age $a + 1$. We define Γ_a as the three-state transition matrix containing these probabilities,

$$\Gamma_a = \begin{bmatrix} p_{11}(a) & p_{21}(a) & 0 \\ p_{12}(a) & p_{22}(a) & 0 \\ p_{10}(a) & p_{20}(a) & 1 \end{bmatrix}$$

where each matrix entry is the probability of transiting between any two states at age a (of course, dead is an absorbing state). Now, let's define the 3×1 vector z_a as the vector describing the fraction of individuals in each state $(z_0(a), z_1(a), z_2(a))$. Given an initial health distribution at age $a = 50$ (our initial age) we can compute, $z_{a+1} = \Gamma_a z_a$ for all ages.[2]

[2]To compute the health distribution at age 50 we use the average share of individuals in good health between ages 50 and 54

To derive the expected duration in each health status, we start by computing the expected years lived in each health status in the interval $(a, a+1)$. The expected years lived in status $i \in \{1, 2\}$ is

$$z_i(a)p_{ii}(a) + \frac{1}{2}z_i(a)\Big[p_{ij}(a) + p_{i0}(a)\Big] + \frac{1}{2}z_j(a)p_{ji}(a)$$

where $j \in \{1, 2\}$ and $j \neq i$. The first term counts a full year for those individuals who were in health i at age a and remain in health i at age $a+1$, the second term counts half-year for those individuals who were in health i at age a and change state (either to health j or to death) before age $a+1$, and the third term counts half-year for those individuals who were in health j at age a and transit to state i before age $a+1$. Thus the expected duration at age 50 in status i is given by

$$ED_i = \sum_{a=50}^{\bar{a}} z_i(a)p_{ii}(a) + \frac{1}{2}z_i(a)\Big[p_{ij}(a) + p_{i0}(a)\Big] + \frac{1}{2}z_j(a)p_{ji}(a)$$

where $\bar{a} = 90$. Keeping with our notation, HLE $\equiv ED_2$, ULE $\equiv ED_1$, and LE = HLE+ULE.

Appendix D: Additional Tables

TABLE D.1: Sample statistics

	Waves	First year	Last year	Individuals
Austria*	6	2004	2015	5139
Belgium	6	2004	2015	6557
Czechia*	5	2006	2015	6441
Denmark*	6	2004	2015	4453
Estonia*	3	2010	2015	6322
France*	6	2004	2015	5964
Germany	6	2004	2015	5723
Greece	4	2004	2015	3394
Israel	4	2005	2015	3041
Italy*	6	2004	2015	5248
Netherlands	5	2004	2013	3474
Poland*	4	2006	2015	2175
Slovenia*	3	2011	2015	3035
Spain*	6	2004	2015	6927
Sweden*	6	2004	2015	5242
Switzerland	6	2004	2015	3557
England*	6	2002	2013	14242
US*	6	2002	2013	27198

Notes: "First year" and "Last year" refer to year of interview or death in our sample. An * mark indicates that the country sample has been selected for the main analysis of the paper.

Table D.2: Educational Attainment

	EUROSTAT (Age 55-74)		Sample (Age 55-74)		Sample (Age 50-90)	
	Males	Females	Males	Females	Males	Females
Austria*	19.6	7.9	20.0	14.2	20.3	14.6
Belgium	22.7	17.2	29.1	24.4	29.7	25.6
Czechia*	13.9	7.8	20.6	11.7	20.3	12.1
Denmark*	25.1	23.1	36.5	36.5	37.4	37.7
Estonia*	26.8	31.4	21.4	20.5	21.7	21.2
France*	15.9	13.2	20.5	16.5	20.8	16.9
Germany	30.1	13.8	33.5	20.4	33.6	20.7
Greece	15.3	7.7	18.6	8.8	19.2	9.5
Italy*	8.9	6.3	8.6	5.8	8.9	6.1
Netherlands	29.1	17.0	26.5	16.5	27.2	17.7
Poland*	12.9	10.9	6.4	5.4	6.4	5.3
Slovenia*	17.3	13.1	15.2	10.7	14.9	10.9
Spain*	17.3	9.6	10.8	6.8	11.1	7.9
Sweden*	21.6	26.6	21.8	22.9	21.8	22.9
Switzerland	33.8	12.7	19.3	10.4	19.6	11.3
England*	26.4	23.8	16.4	12.5	16.4	12.6
US*	31.3	24.1	27.7	20.4	27.8	21.1

Notes: An * mark indicates that the country sample has been selected for the main analysis of the paper. Sources: EUROSTAT (Population by educational attainment level, sex and age [edat_lfs_9903]) share of population age 55-74 with ISCED11 tertiary education (levels 5-8), average for years 2004-2012. For the US: OECD (2016) share of population age 55-64 with at least a bachelor's degree. Sample: In continental Europe, share of population with at least 15 years of education. In England, share of population finished full-time education after age 19. In the US, share of population with completed college.

Table D.3: Life expectancies: College

	A. Males			B. Females			C. Difference		
	LE	HLE	ULE	LE	HLE	ULE	LE	HLE	ULE
Western Europe	34.1	30.9	3.2	35.6	31.9	3.7	-1.5	-1.0	-0.5
	(0.6)	(0.6)	(0.3)	(0.6)	(0.6)	(0.4)	(0.8)	(0.8)	(0.5)
Austria	33.5	30.3	3.2	34.3	31.1	3.2	-0.8	-0.8	-0.0
	(0.8)	(0.8)	(0.4)	(0.9)	(0.9)	(0.5)	(1.2)	(1.2)	(0.7)
France	34.7	31.4	3.3	36.7	32.5	4.2	-2.0	-1.1	-0.9
	(0.8)	(0.9)	(0.5)	(0.7)	(0.8)	(0.6)	(1.1)	(1.2)	(0.8)
Eastern Europe	30.2	26.9	3.3	35.7	31.8	3.9	-5.5	-4.9	-0.6
	(0.7)	(0.6)	(0.3)	(0.5)	(0.5)	(0.3)	(0.9)	(0.8)	(0.4)
Czechia	29.6	26.6	2.9	36.9	32.1	4.7	-7.3	-5.5	-1.8
	(1.2)	(1.1)	(0.4)	(0.9)	(1.1)	(0.8)	(1.5)	(1.5)	(0.9)
Estonia	30.1	26.5	3.5	35.2	31.3	3.9	-5.1	-4.8	-0.4
	(1.1)	(1.0)	(0.4)	(0.6)	(0.7)	(0.4)	(1.3)	(1.2)	(0.6)
Poland	30.8	26.1	4.5	31.6	26.7	4.7	-0.9	-0.6	-0.1
	(1.8)	(1.8)	(1.1)	(2.5)	(2.5)	(1.7)	(3.1)	(3.0)	(2.0)
Slovenia	32.1	29.3	2.7	37.0	34.5	2.3	-4.8	-5.2	0.4
	(1.5)	(1.4)	(0.7)	(1.4)	(1.6)	(0.9)	(2.0)	(2.1)	(1.1)
Mediterranean	33.0	30.6	2.4	34.4	32.7	1.6	-1.3	-2.1	0.8
	(0.9)	(0.9)	(0.5)	(1.1)	(1.1)	(0.4)	(1.4)	(1.5)	(0.6)
Italy	33.5	30.3	3.1	36.3	34.5	1.6	-2.7	-4.2	1.5
	(1.3)	(1.4)	(0.8)	(1.6)	(1.8)	(0.8)	(2.0)	(2.2)	(1.1)
Spain	32.5	30.5	1.9	33.4	31.5	1.9	-0.9	-1.0	0.1
	(1.3)	(1.3)	(0.5)	(1.3)	(1.4)	(0.6)	(1.9)	(1.9)	(0.7)
Scandinavia	32.7	30.2	2.5	34.6	32.2	2.4	-1.9	-2.0	0.1
	(0.5)	(0.5)	(0.3)	(0.5)	(0.5)	(0.3)	(0.7)	(0.8)	(0.4)
Denmark	31.6	29.3	2.3	33.1	31.3	1.7	-1.4	-2.0	0.5
	(0.8)	(0.8)	(0.3)	(0.8)	(0.8)	(0.3)	(1.1)	(1.1)	(0.4)
Sweden	33.6	30.8	2.7	36.0	32.9	3.1	-2.4	-2.0	-0.4
	(0.7)	(0.7)	(0.4)	(0.6)	(0.7)	(0.4)	(0.9)	(1.0)	(0.6)
England	32.6	28.2	4.4	33.9	28.5	5.4	-1.3	-0.3	-0.9
	(0.5)	(0.5)	(0.3)	(0.5)	(0.5)	(0.4)	(0.7)	(0.7)	(0.5)
US	31.0	28.2	2.8	33.9	29.8	4.1	-2.9	-1.5	-1.4
	(0.4)	(0.4)	(0.1)	(0.3)	(0.3)	(0.2)	(0.5)	(0.5)	(0.2)

Notes: LE stands for life expectancy, HLE for healthy life expectancy, and ULE for unhealthy life expectancy. For each country we report the median (and the standard deviation in parenthesis) of the distribution of the corresponding life expectancies that arises from the posterior distribution of the estimated $\boldsymbol{\beta}$ parameters.

Table D.4: Life expectancies: No College

	A. Males			B. Females			C. Difference		
	LE	HLE	ULE	LE	HLE	ULE	LE	HLE	ULE
Western Europe	30.2	26.2	4.0	33.9	28.9	5.0	-3.7	-2.6	-1.0
	(0.5)	(0.4)	(0.2)	(0.3)	(0.3)	(0.2)	(0.6)	(0.5)	(0.2)
Austria	29.6	26.2	3.4	32.6	28.4	4.2	-3.0	-2.2	-0.8
	(0.7)	(0.7)	(0.3)	(0.5)	(0.5)	(0.2)	(0.9)	(0.8)	(0.4)
France	30.7	26.3	4.4	34.9	29.2	5.7	-4.2	-2.9	-1.3
	(0.6)	(0.5)	(0.2)	(0.4)	(0.4)	(0.2)	(0.7)	(0.6)	(0.3)
Eastern Europe	26.2	22.6	3.6	31.8	26.6	5.2	-5.6	-4.1	-1.6
	(0.4)	(0.3)	(0.1)	(0.3)	(0.2)	(0.1)	(0.4)	(0.4)	(0.2)
Czechia	26.0	22.6	3.5	31.0	26.4	4.5	-4.9	-3.8	-1.1
	(0.6)	(0.5)	(0.2)	(0.4)	(0.4)	(0.2)	(0.7)	(0.7)	(0.3)
Estonia	25.5	21.9	3.6	32.7	26.8	5.9	-7.3	-4.9	-2.4
	(0.7)	(0.6)	(0.2)	(0.5)	(0.4)	(0.3)	(0.8)	(0.7)	(0.3)
Poland	25.4	21.6	3.8	30.6	25.4	5.3	-5.3	-3.8	-1.5
	(0.8)	(0.7)	(0.3)	(0.6)	(0.6)	(0.4)	(1.0)	(0.9)	(0.5)
Slovenia	29.6	25.6	3.9	34.0	29.2	4.7	-4.4	-3.5	-0.8
	(0.9)	(0.8)	(0.4)	(0.7)	(0.7)	(0.4)	(1.2)	(1.1)	(0.5)
Mediterranean	30.0	27.2	2.8	33.6	28.6	5.1	-3.7	-1.4	-2.3
	(0.3)	(0.3)	(0.1)	(0.3)	(0.2)	(0.1)	(0.4)	(0.4)	(0.2)
Italy	30.9	28.2	2.7	33.7	28.6	5.1	-2.8	-0.4	-2.4
	(0.4)	(0.4)	(0.2)	(0.3)	(0.3)	(0.2)	(0.6)	(0.5)	(0.3)
Spain	29.0	26.2	2.8	33.4	28.5	5.0	-4.4	-2.3	-2.1
	(0.5)	(0.5)	(0.2)	(0.4)	(0.4)	(0.2)	(0.6)	(0.6)	(0.3)
Scandinavia	30.5	27.3	3.2	32.3	28.5	3.8	-1.8	-1.2	-0.6
	(0.4)	(0.4)	(0.2)	(0.4)	(0.4)	(0.2)	(0.5)	(0.5)	(0.3)
Denmark	29.6	26.3	3.3	31.0	27.3	3.7	-1.3	-0.9	-0.4
	(0.6)	(0.6)	(0.3)	(0.6)	(0.6)	(0.3)	(0.9)	(0.9)	(0.4)
Sweden	31.1	27.9	3.2	33.2	29.3	3.9	-2.1	-1.4	-0.7
	(0.5)	(0.5)	(0.2)	(0.4)	(0.5)	(0.2)	(0.7)	(0.7)	(0.3)
England	29.2	23.5	5.7	32.7	25.5	7.2	-3.5	-2.0	-1.5
	(0.3)	(0.3)	(0.1)	(0.2)	(0.2)	(0.2)	(0.3)	(0.3)	(0.2)
US	27.5	23.0	4.5	30.7	24.6	6.1	-3.3	-1.6	-1.6
	(0.2)	(0.2)	(0.1)	(0.2)	(0.2)	(0.1)	(0.3)	(0.3)	(0.1)

Notes: LE stands for life expectancy, HLE for healthy life expectancy, and ULE for unhealthy life expectancy. For each country we report the median (and the standard deviation in parenthesis) of the distribution of the corresponding life expectancies that arises from the posterior distribution of the estimated $\boldsymbol{\beta}$ parameters.

TABLE D.5: Life expectancies: Pooled education

	A. Males			B. Females			C. Difference		
	LE	HLE	ULE	LE	HLE	ULE	LE	HLE	ULE
Western Europe	31.3	27.4	3.8	34.3	29.5	4.8	-3.0	-2.0	-1.0
	(0.4)	(0.3)	(0.2)	(0.3)	(0.3)	(0.2)	(0.5)	(0.4)	(0.2)
Austria	30.8	27.5	3.3	32.9	28.9	4.0	-2.1	-1.4	-0.7
	(0.6)	(0.5)	(0.2)	(0.5)	(0.4)	(0.2)	(0.7)	(0.7)	(0.3)
France	31.7	27.4	4.2	35.2	29.8	5.4	-3.5	-2.3	-1.2
	(0.5)	(0.4)	(0.2)	(0.3)	(0.3)	(0.2)	(0.6)	(0.6)	(0.3)
Eastern Europe	26.9	23.3	3.6	32.4	27.4	5.0	-5.5	-4.0	-1.4
	(0.3)	(0.3)	(0.1)	(0.2)	(0.2)	(0.1)	(0.4)	(0.4)	(0.2)
Czechia	26.7	23.3	3.3	31.5	27.0	4.5	-4.8	-3.7	-1.2
	(0.5)	(0.5)	(0.2)	(0.4)	(0.4)	(0.2)	(0.7)	(0.6)	(0.3)
Estonia	26.4	22.9	3.5	33.3	27.9	5.5	-6.9	-5.0	-1.9
	(0.6)	(0.5)	(0.2)	(0.4)	(0.4)	(0.2)	(0.7)	(0.6)	(0.3)
Poland	26.0	22.1	3.8	30.8	25.6	5.2	-4.8	-3.4	-1.4
	(0.7)	(0.7)	(0.3)	(0.6)	(0.6)	(0.4)	(0.9)	(0.9)	(0.5)
Slovenia	30.1	26.4	3.7	34.3	29.8	4.5	-4.2	-3.4	-0.8
	(0.8)	(0.8)	(0.3)	(0.6)	(0.6)	(0.3)	(1.1)	(1.0)	(0.5)
Mediterranean	30.2	27.5	2.8	33.7	28.8	4.9	-3.4	-1.3	-2.1
	(0.3)	(0.3)	(0.1)	(0.2)	(0.2)	(0.1)	(0.4)	(0.4)	(0.2)
Italy	31.2	28.4	2.8	33.8	28.8	5.0	-2.7	-0.4	-2.2
	(0.4)	(0.4)	(0.2)	(0.3)	(0.3)	(0.2)	(0.5)	(0.5)	(0.3)
Spain	29.4	26.6	2.8	33.4	28.6	4.8	-4.1	-2.0	-2.0
	(0.5)	(0.4)	(0.1)	(0.4)	(0.3)	(0.2)	(0.6)	(0.5)	(0.2)
Scandinavia	31.1	28.2	3.0	33.0	29.7	3.3	-1.9	-1.5	-0.3
	(0.3)	(0.3)	(0.1)	(0.3)	(0.3)	(0.1)	(0.4)	(0.4)	(0.2)
Denmark	30.3	27.4	2.9	31.8	28.8	3.0	-1.5	-1.4	-0.1
	(0.5)	(0.5)	(0.2)	(0.5)	(0.5)	(0.2)	(0.7)	(0.7)	(0.3)
Sweden	31.8	28.8	3.0	33.9	30.4	3.5	-2.1	-1.6	-0.5
	(0.4)	(0.4)	(0.2)	(0.3)	(0.4)	(0.2)	(0.5)	(0.6)	(0.3)
England	29.8	24.4	5.4	32.8	26.0	6.9	-3.0	-1.6	-1.5
	(0.2)	(0.2)	(0.1)	(0.2)	(0.2)	(0.1)	(0.3)	(0.3)	(0.2)
US	28.3	24.3	4.1	31.3	25.5	5.8	-3.0	-1.3	-1.7
	(0.2)	(0.2)	(0.1)	(0.2)	(0.2)	(0.1)	(0.3)	(0.2)	(0.1)

Notes: LE stands for life expectancy, HLE for healthy life expectancy, and ULE for unhealthy life expectancy. For each country we report the median (and the standard deviation in parenthesis) of the distribution of the corresponding life expectancies that arises from the posterior distribution of the estimated $\boldsymbol{\beta}$ parameters.

Table D.6: Education gradients

	A. Males			B. Females			C. Difference		
	LE	HLE	ULE	LE	HLE	ULE	LE	HLE	ULE
Western Europe	3.9	4.6	-0.8	1.7	3.0	-1.3	2.2	1.6	0.6
	(0.7)	(0.7)	(0.4)	(0.6)	(0.7)	(0.4)	(1.0)	(1.0)	(0.6)
Austria	3.8	4.1	-0.3	1.7	2.7	-1.0	2.2	1.4	0.8
	(1.1)	(1.0)	(0.5)	(1.1)	(1.0)	(0.5)	(1.5)	(1.4)	(0.7)
France	3.9	5.0	-1.1	1.8	3.2	-1.5	2.2	1.8	0.4
	(1.0)	(1.0)	(0.5)	(0.8)	(0.9)	(0.7)	(1.3)	(1.4)	(0.9)
Eastern Europe	4.0	4.3	-0.3	3.9	5.1	-1.2	0.1	-0.8	0.9
	(0.8)	(0.7)	(0.3)	(0.6)	(0.6)	(0.4)	(1.0)	(0.9)	(0.5)
Czechia	3.5	4.0	-0.5	5.9	5.7	0.1	-2.3	-1.6	-0.7
	(1.3)	(1.2)	(0.5)	(1.0)	(1.1)	(0.8)	(1.7)	(1.7)	(0.9)
Estonia	4.6	4.6	-0.0	2.5	4.5	-2.0	2.1	0.1	2.0
	(1.3)	(1.2)	(0.5)	(0.8)	(0.8)	(0.5)	(1.5)	(1.4)	(0.7)
Poland	5.4	4.5	0.8	1.0	1.3	-0.6	4.4	3.2	1.3
	(2.0)	(1.9)	(1.1)	(2.6)	(2.6)	(1.7)	(3.2)	(3.2)	(2.0)
Slovenia	2.5	3.7	-1.3	3.0	5.3	-2.5	-0.5	-1.6	1.2
	(1.7)	(1.6)	(0.8)	(1.5)	(1.7)	(1.0)	(2.3)	(2.4)	(1.2)
Mediterranean	3.0	3.4	-0.4	0.7	4.1	-3.5	2.3	-0.7	3.1
	(1.0)	(1.0)	(0.5)	(1.1)	(1.1)	(0.4)	(1.5)	(1.5)	(0.7)
Italy	2.5	2.1	0.3	2.5	5.9	-3.6	0.1	-3.8	3.9
	(1.3)	(1.4)	(0.8)	(1.6)	(1.8)	(0.8)	(2.1)	(2.3)	(1.2)
Spain	3.4	4.3	-0.9	-0.0	3.0	-3.1	3.5	1.3	2.2
	(1.4)	(1.4)	(0.5)	(1.4)	(1.5)	(0.6)	(2.0)	(2.0)	(0.8)
Scandinavia	2.1	2.9	-0.8	2.3	3.7	-1.4	-0.1	-0.8	0.7
	(0.7)	(0.7)	(0.3)	(0.6)	(0.6)	(0.3)	(0.9)	(0.9)	(0.4)
Denmark	2.0	3.0	-1.0	2.1	4.0	-1.9	-0.1	-1.0	0.9
	(1.0)	(1.0)	(0.4)	(1.0)	(1.0)	(0.4)	(1.4)	(1.4)	(0.6)
Sweden	2.5	2.9	-0.5	2.8	3.6	-0.8	-0.3	-0.7	0.3
	(0.8)	(0.9)	(0.5)	(0.8)	(0.8)	(0.5)	(1.1)	(1.2)	(0.7)
England	3.4	4.7	-1.3	1.2	3.0	-1.8	2.2	1.7	0.5
	(0.6)	(0.6)	(0.3)	(0.6)	(0.6)	(0.4)	(0.8)	(0.8)	(0.5)
US	3.6	5.3	-1.7	3.2	5.1	-2.0	0.4	0.1	0.3
	(0.4)	(0.4)	(0.2)	(0.4)	(0.4)	(0.2)	(0.6)	(0.6)	(0.3)
Average	3.4	4.0	-0.6	2.2	3.9	-1.7	1.2	0.1	1.1
	(0.4)	(0.3)	(0.2)	(0.4)	(0.4)	(0.2)	(0.5)	(0.5)	(0.3)

Notes: LE stands for life expectancy, HLE for healthy life expectancy, and ULE for unhealthy life expectancy, all at age 50. The education gradient is the difference in the corresponding life expectancy between college and non-college individuals. Panel A refers to males, Panel B to females, and Panel C is the difference between the male and female gradients. For each country we report the median (and the standard deviation in parenthesis) of the distribution of the corresponding life expectancy that arises from the posterior distribution of the estimated $\boldsymbol{\beta}$ parameters.

Table D.7: Gender Gaps

	A. Low educated			B. High educated			C. Difference		
	LE	HLE	ULE	LE	HLE	ULE	LE	HLE	ULE
Western Europe	3.7	2.6	1.0	1.5	1.0	0.5	2.2	1.6	0.6
	(0.6)	(0.5)	(0.2)	(0.8)	(0.8)	(0.5)	(1.0)	(1.0)	(0.6)
Eastern Europe	5.6	4.1	1.6	5.5	4.9	0.6	0.1	-0.8	0.9
	(0.4)	(0.4)	(0.2)	(0.9)	(0.8)	(0.4)	(1.0)	(0.9)	(0.5)
Mediterranean	3.7	1.4	2.3	1.3	2.1	-0.8	2.3	-0.7	3.1
	(0.4)	(0.4)	(0.2)	(1.4)	(1.5)	(0.6)	(1.5)	(1.5)	(0.7)
Scandinavia	1.8	1.2	0.6	1.9	2.0	-0.1	-0.1	-0.8	0.7
	(0.5)	(0.5)	(0.3)	(0.7)	(0.8)	(0.4)	(0.9)	(0.9)	(0.4)
England	3.5	2.0	1.5	1.3	0.3	0.9	2.2	1.7	0.5
	(0.3)	(0.3)	(0.2)	(0.7)	(0.7)	(0.5)	(0.8)	(0.8)	(0.5)
US	3.3	1.6	1.6	2.9	1.5	1.4	0.4	0.1	0.3
	(0.3)	(0.3)	(0.1)	(0.5)	(0.5)	(0.2)	(0.6)	(0.6)	(0.3)
Average	3.9	2.5	1.4	2.7	2.4	0.3	1.2	0.1	1.1
	(0.2)	(0.2)	(0.1)	(0.5)	(0.5)	(0.3)	(0.5)	(0.5)	(0.3)

Notes: LE stands for life expectancy, HLE for healthy life expectancy, and ULE for unhealthy life expectancy, all at age 50. The gender gap is the difference in the corresponding life expectancy between females and males. Panel A refers to individuals without college, Panel B to individuals with a college degree, and Panel C is the difference between the non-college and the college gender gaps. For each country we report the median (and the standard deviation in parenthesis) of the distribution of the corresponding life expectancy that arises from the posterior distribution of the estimated $\boldsymbol{\beta}$ parameters.

Appendix E: Validation Appendix

E.1 Introduction

The use of survey data such as SHARE, HRS, or ELSA for survival analysis is not without problems. Despite the careful design of their sampling frames, there is a variety of reasons that may cause the survival rates computed with survey data not to align well with the ones computed in the population. For instance, there is the possibility of biases in sample design, in the response rates at baseline, or in the sample retention, which are all problems inherent in survey data. The importance of these potential problems is likely to differ across countries. Therefore, this Appendix explores the quality for survival analysis of each SHARE country data alongside the ELSA and HRS surveys. Section E.2 compares non-parametric Kaplan-Meier survival functions from our survey data to those obtained from the population, showing that HRS, ELSA, and several SHARE country samples do quite well. Next, Section E.3 provides another set of results assessing the quality of a parametric logit model to approximate the non-parametric survival functions. Finally, in Section E.4 we discuss the observed attrition in the SHARE country samples and to which extent it may be correlated with survival.

E.2 Validation

In this Section we compare the Kaplan-Meier survival functions computed on the survey data for each country-gender to the ones obtained in the population. Our *long sample* uses all observations of any individual, even if they are non-consecutive, as we do in the main paper (see Section 3 of the main paper). SHARE samples suffered substantial attrition in the first two waves and a special effort was undertaken to recover lost households. This effort to recover individuals lost to attrition may have potentially added new sources of bias if its success is related to the survival status of the individual. For this reason, we also use a second sample (*short sample*) that only uses the first set of consecutive observations.

The data for the survival functions in the population are extracted from different sources. For each country we use the life tables of the range of years for which we have survey data and take the unweighted average.[3]

We display the survival functions for every country and gender in Figures E.1 to E.18 and the associated life expectancies at ages 50 and 65 in Table E.1. The results from the

[3]We use EUROSTAT data for all European countries. Data for England comes from the Office of National Statistics, data for Israel comes from the Central Bureau of Statistics, and data for the US comes from the Center for Disease Control and Prevention. See Table D.1 in the main paper for the range of years that applies to each country.

validation exercise are the following.

First, the survival curves of the *short sample* are systematically above the ones of the *long sample* for all SHARE countries except for Greece, Israel and Spain —as well as for the non-SHARE countries England and the US— where the opposite is true.[4] This difference is clearly reflected in the life expectancies computed with each sample. The differences between the survival functions computed with the different samples is large for some cases, like in Greece, Israel, Denmark, Netherlands, and Sweden.

Second, the survival curves of both the *short sample* and the *long sample* are above the ones in the population for all countries except for England, Israel, Spain, and the US. All in all, the *long sample* produces a better approximation to the population life tables in England, the US, and all SHARE countries except Greece and Spain. Given this and that by construction the *long sample* contains more transitions than the *short sample*, we conclude that the *long sample* is preferred for survival analysis.

Third, the difference between the survival functions computed with the *long sample* and the ones from the population tends to be substantially larger for males than females.

And fourth, we observe a large amount of heterogeneity in the difference between the survival functions computed in the population and the ones computed with the *long sample*. We use life expectancies at age 50 as a way to quantify these differences. In particular, we classify our country samples as producing an excellent, good, or poor approximation to the population survival functions if the difference in life expectancy at age 50 is 1 or less, above 1 and up to 2.5, or above 2.5. According to this criterion the excellent country samples are: Austria females, Czech Republic, Denmark females, England, Italy, Israel females, Poland, Spain, and the US; the good country samples are: Austria males, Denmark males, Estonia, France, Greece females, Israel males, Slovenia, Sweden, and Switzerland females; and the poor country samples are Belgium, Germany, Greece males, Netherlands, and Switzerland males. For our main analysis we drop the countries where at least one of the sexes has a poor approximation to life tables: Belgium, Germany, Greece, Netherlands, and Switzerland. In addition, we drop Israel because there is a large incidence of errors in the dates of death.[5] Therefore, the list of countries we work with is given by: Austria, Czech Republic, Denmark, England, Estonia, France,

[4]The distinction between *short sample* and *long sample* for England and the US is not so easy to interpret because these surveys experienced neither the strong attrition suffered by SHARE nor the posterior effort to recover individuals.

[5]These errors are uncovered because more than 10% of recorded death events have a date of death previous to an actual interview where the respondent was alive. We set these dates of death to missing, but other errors in the date of death are impossible to detect. This problem may be behind the odd fact that Israel is the only SHARE country where the survival rate in the sample is clearly below the one in the population.

Italy, Poland, Slovenia, Spain, Sweden, and the US.

E.3 Parametric versus non-parametric survival functions

The non-parametric Kaplan-Meier survival functions give plenty of degrees of freedom to adjust the age-profile of survival functions. However, when exploring the socio-economic gradient of survival functions we need a more parsimonious approach. For this reason, in this Section we compare the Kaplan-Meier survival functions to the estimates based on a logit model. We do so only for the *long sample* in each country-gender group. In Figures E.1 to E.18 we plot for each country and gender the Kaplan-Meier survival functions together with a logit survival function estimated with a linear term in age. In Table E.1 we summarize these survival functions with the life expectancies at age 50 and 65. Our results are very clear: the logit model produces survival functions that follow very well the non-parametric Kaplan-Meier estimates, and the differences in life expectancies computed with one or the other method are very similar.

E.4 Attrition

One potential problem regarding the use of SHARE data for survival analysis is that panel attrition rates for those who survive might be different from those who die. If for instance panel attrition rates are larger among deceased respondents, then our survival estimates would be upward biased. Since it is difficult to know the survival status of those who attrite between waves, implementing a formal test for the existence of that bias is not possible. Alternatively, and assuming that subjective survival probabilities correlate with actual mortality, we can exploit the data that SHARE collects on respondents' survival expectations to explore the relationship between attrition and survival.[6]

SHARE respondents' expectations are elicited as point expectations through the following question: "What are the chances that you will live to be age [75 / 80 / 85 / 90 / 95 / 100 / 105 / 110 / 120]? We analyse whether panel attrition, conditional on age and education, correlates with subjective survival probability through the estimation of logit models where the dependent variable is given by a binary indicator that equals to one if the respondent did not participate in the next wave ("attrited"). This binary variable is regressed on the respondent's own subjective survival probability ("explive_dm"), which has been demeaned using for each observation the sample mean for its country and age

[6]Using HRS data, Hurd, McFadden and Merrill (1999) show that subjective survival probabilities correlate with actual mortality. Other studies also based on the HRS have also shown that these indicators are highly correlated with other measures of expected longevity such as smoking and body mass index (Hurd and McGarry (1995)).

group in order to account for the fact that the target age in the SHARE question on survival expectation varies by the respondent's age. Additional controls are given by a quadratic function on age and a dummy variable for less than tertiary education.

Regarding the interpretation of the results, a negative and significant coefficient of the subjective survival probability would imply that individuals with higher survival expectations have a lower likelihood to drop out from the survey. The estimates shown in Panel A of Table 3 are based on the pool of observations for all waves at the country-gender level. Our results provide evidence that, conditional on age and education, subjective survival probability predicts panel attrition only in some cases. Among our country-gender samples, the coefficient of the variable "explive_dm" is negative and significant for Austria females, Germany females, Italy females, Greece males, Switzerland both males and females, Poland males, and Estonia females, which means that "survivors" are less likely to attrit in those cases. As mentioned above, this suggests that panel attrition might bias (upwards) the estimates of life expectancy based on these samples. Instead, country samples like Belgium or Netherlands, whose life expectancies are clearly above their population counterparts, do not show any evidence of survival-related attrition.

An additional important issue is whether attrition could induce biases in the education gradient for life expectancy. This would happen as long as panel attrition differentials between "survivors" and "deceased" individuals varied by education level. To check this out, we include in our logit regressions an interaction term between the subjective survival probability and the dummy for low educated ("exp_educ"). A significant coefficient would suggest that subjective survival probability predicts attrition differently for different education groups which would bias the gradient. Panel B shows the estimated coefficients of interest for this version of the logit model. In particular, we get significant, or close to significant coefficients only for Austria females, France females, and Belgium females. For Austria females the sign is positive, which would imply that we might be overestimating the education gradient in life expectancy. For France females and Belgium females the sign is negative, hence the bias would run in the opposite direction.

E.5 Tables and Figures

TABLE E.1: Life expectancies

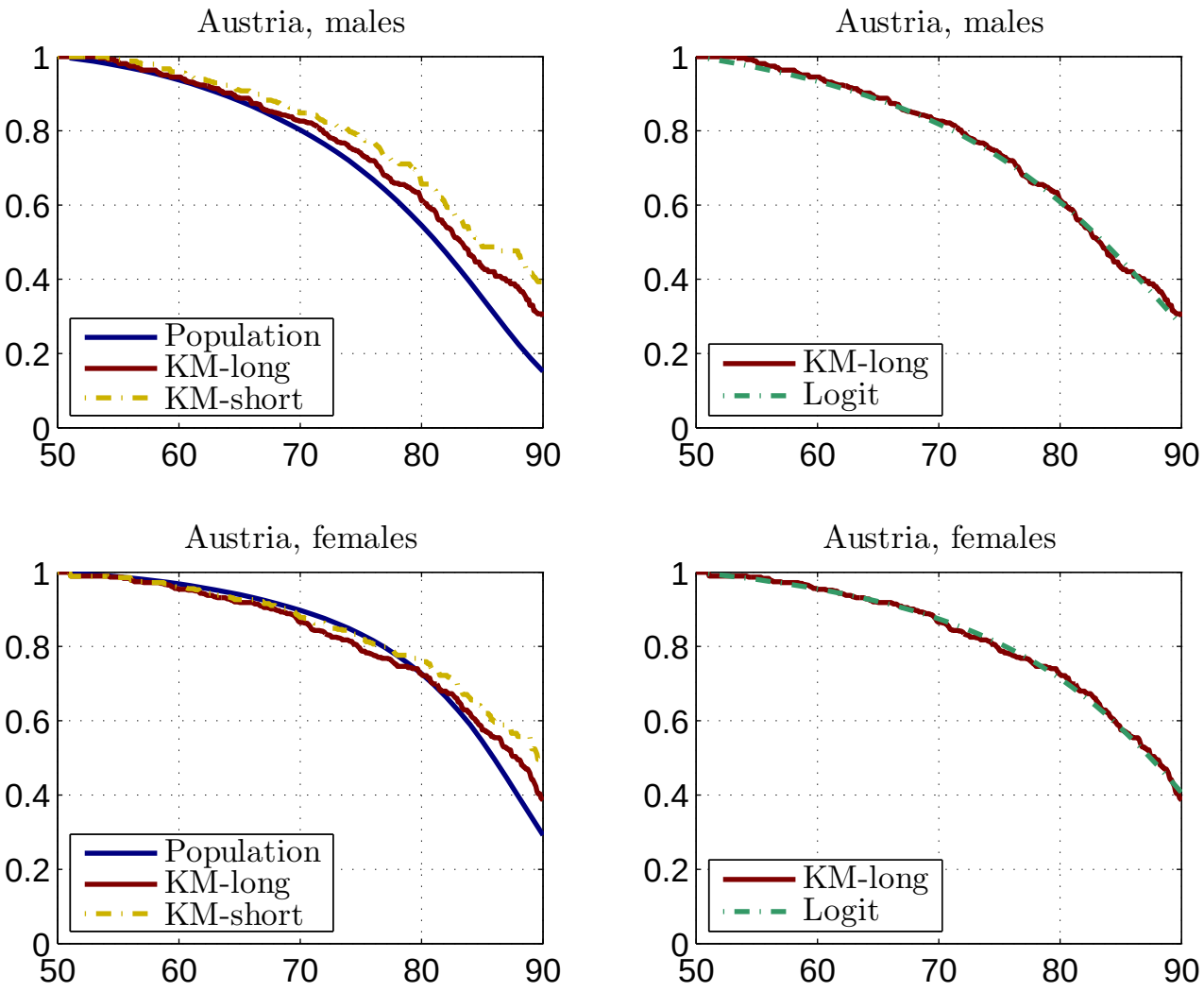
		A. Life expectancy at 50				B. Life expectancy at 65			
Country	G	Pop	km_S	km_l	Logit	Pop	km_S	km_l	Logit
Austria	m	28.9	3.2	1.7	1.4	16.6	2.6	1.6	1.6
Austria	f	33.0	0.9	0.0	0.0	19.5	1.4	0.6	0.5
Belgium	m	28.7	4.3	2.9	2.8	16.4	3.6	2.4	2.5
Belgium	f	32.6	3.1	2.3	2.2	19.4	2.7	2.1	2.0
Czech	m	26.2	1.1	0.6	0.3	14.7	0.8	0.8	0.6
Czech	f	31.0	1.3	0.8	0.6	17.8	1.0	0.6	0.5
Denmark	m	28.3	3.7	2.1	2.2	16.0	1.9	1.0	1.0
Denmark	f	31.4	2.3	0.7	0.7	18.3	1.8	0.5	0.4
Estonia	m	24.8	1.5	1.3	1.5	14.1	1.3	1.3	1.0
Estonia	f	31.7	2.1	1.5	1.3	18.6	1.9	1.6	1.5
France	m	29.4	3.1	2.5	2.6	17.4	2.8	2.2	2.1
France	f	34.1	2.1	1.6	1.6	20.6	1.5	1.0	1.1
Germany	m	28.8	4.6	4.0	4.1	16.5	3.8	3.3	3.4
Germany	f	32.6	2.8	2.7	2.6	19.3	1.8	1.8	1.9
Greece	m	29.1	2.8	4.2	4.3	16.8	2.4	3.0	3.1
Greece	f	32.9	2.0	2.3	2.4	19.2	1.7	1.8	1.8
Israel	m	30.3	-3.0	-1.5	-1.6	17.5	-1.7	-0.5	-0.6
Israel	f	33.0	-1.8	-0.3	-0.5	19.3	-1.0	0.0	0.0
Italy	m	30.1	1.5	1.0	1.0	17.2	1.2	0.6	0.8
Italy	f	33.7	0.9	0.1	0.1	20.0	0.5	-0.2	-0.3
Netherlands	m	29.3	3.9	2.4	2.2	16.4	3.5	2.3	2.2
Netherlands	f	32.4	4.1	3.2	3.0	19.1	3.3	2.4	2.4
Poland	m	24.8	0.4	0.3	0.5	14.3	0.0	0.0	0.3
Poland	f	30.9	0.1	-0.8	-0.6	18.1	-0.1	-0.8	-0.7
Slovenia	m	28.3	2.3	1.8	1.6	16.2	1.8	1.5	1.4
Slovenia	f	33.0	2.2	1.5	1.5	19.5	2.0	1.6	1.8
Spain	m	29.5	-0.7	-0.6	-0.6	17.1	0.2	0.1	0.1
Spain	f	34.1	0.1	-0.6	-0.6	20.3	-0.1	-0.4	-0.4
Sweden	m	30.2	3.7	1.8	1.7	17.2	2.5	1.1	1.2
Sweden	f	33.0	3.3	1.2	1.1	19.5	2.4	0.8	1.0
Switzerland	m	30.6	4.1	3.1	3.2	17.7	3.0	2.6	2.5
Switzerland	f	33.9	3.1	2.5	2.6	20.2	2.6	2.4	2.3
England	m	29.3	-1.3	-0.6	0.2	16.7	-0.7	-0.2	0.2
England	f	32.2	-0.3	0.2	0.6	18.8	-0.3	0.0	0.4
US	m	28.4	-0.7	-0.4	-0.3	16.6	-0.2	-0.2	-0.1
US	f	31.5	-0.3	-0.2	-0.3	18.6	-0.4	-0.3	-0.2

TABLE E.2: Attrition analysis

	A. Specification 1		B. Specification 2			
	explive_dm		explive_dm		exp_educ	
Country	Males	Females	Males	Females	Males	Females
AT	-0.0031	-0.0066**	-0.0037	-0.0172***	0.0008	0.0136**
	(0.0028)	(0.0026)	(0.0058)	(0.0044)	(0.0065)	(0.0054)
DE	-0.0006	-0.0032*	-0.0003	-0.0070	-0.0004	0.0047
	(0.0020)	(0.0019)	(0.0035)	(0.0043)	(0.0042)	(0.0048)
SE	-0.0027	-0.0009	0.0035	-0.0031	-0.0073	0.0028
	(0.0025)	(0.0025)	(0.0063)	(0.0050)	(0.0069)	(0.0058)
NL	-0.0037	0.0017	-0.0081	0.0069	0.0062	-0.0063
	(0.0037)	(0.0045)	(0.0080)	(0.0101)	(0.0089)	(0.0114)
ES	0.0020	-0.0006	0.0056	0.0037	-0.0039	-0.0046
	(0.0029)	(0.0026)	(0.0080)	(0.0090)	(0.0086)	(0.0094)
IT	-0.0040	-0.0044**	-0.0042	-0.0119*	0.0003	0.0083
	(0.0024)	(0.0021)	(0.0087)	(0.0070)	(0.0090)	(0.0073)
FR	-0.0022	-0.0029	-0.0016	0.0061	-0.0007	-0.0102
	(0.0022)	(0.0019)	(0.0053)	(0.0065)	(0.0058)	(0.0068)
DK	-0.0010	-0.0005	-0.0014	-0.0082	0.0006	0.0104
	(0.0035)	(0.0033)	(0.0057)	(0.0067)	(0.0072)	(0.0076)
GR	-0.0086**	-0.0024	-0.0027	0.0038	-0.0083	-0.0071
	(0.0041)	(0.0040)	(0.0076)	(0.0100)	(0.0091)	(0.0110)
CH	-0.0061**	-0.0044*	-0.0040	-0.0030	-0.0032	-0.0018
	(0.0025)	(0.0026)	(0.0043)	(0.0054)	(0.0053)	(0.0062)
BE	-0.0028	-0.0055	0.0011	0.0067	-0.0049	-0.0152*
	(0.0027)	(0.0041)	(0.0054)	(0.0079)	(0.0062)	(0.0091)
IL	0.0030	-0.0002	0.0048	0.0031	-0.0068	-0.0080
	(0.0068)	(0.0058)	(0.0087)	(0.0067)	(0.0133)	(0.0124)
CZ	0.0023	-0.0011	0.0086*	0.0020	-0.0077	-0.0036
	(0.0020)	(0.0019)	(0.0048)	(0.0047)	(0.0053)	(0.0051)
PL	-0.0069**	-0.0012	-0.0067	-0.0078	-0.0002	0.0073
	(0.0028)	(0.0025)	(0.0077)	(0.0093)	(0.0082)	(0.0097)
SI	0.0000	-0.0002	0.0001	-0.0021	-0.0001	0.0022
	(0.0025)	(0.0023)	(0.0062)	(0.0067)	(0.0067)	(0.0071)
EE	-0.0011	-0.0045***	0.0014	-0.0068***	-0.0040	0.0040
	(0.0018)	(0.0017)	(0.0029)	(0.0026)	(0.0037)	(0.0034)

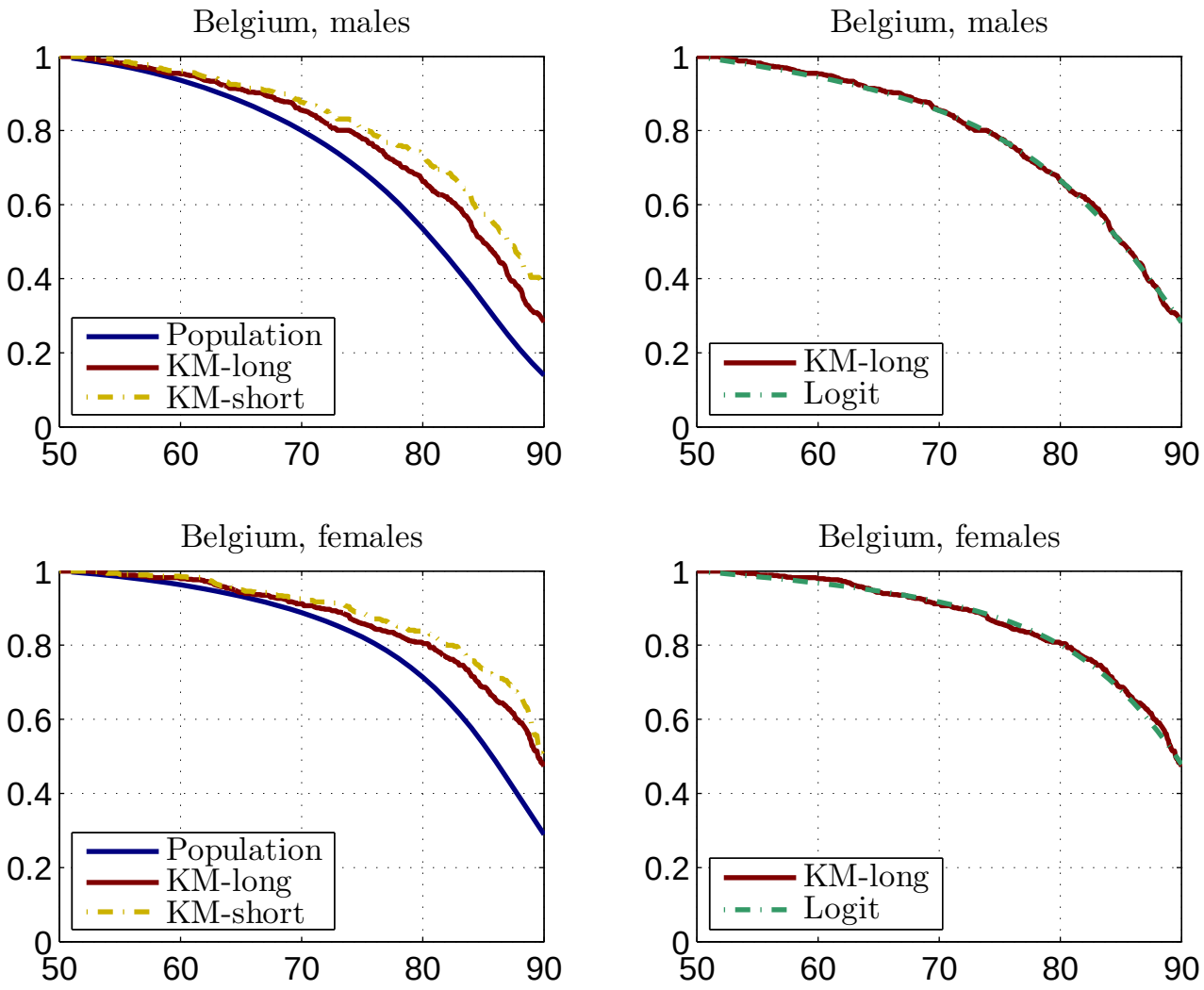
Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Logit regressions that control for age, age2, and loweduc3. Weighted results.

FIGURE E.1: Survival rates: Austria



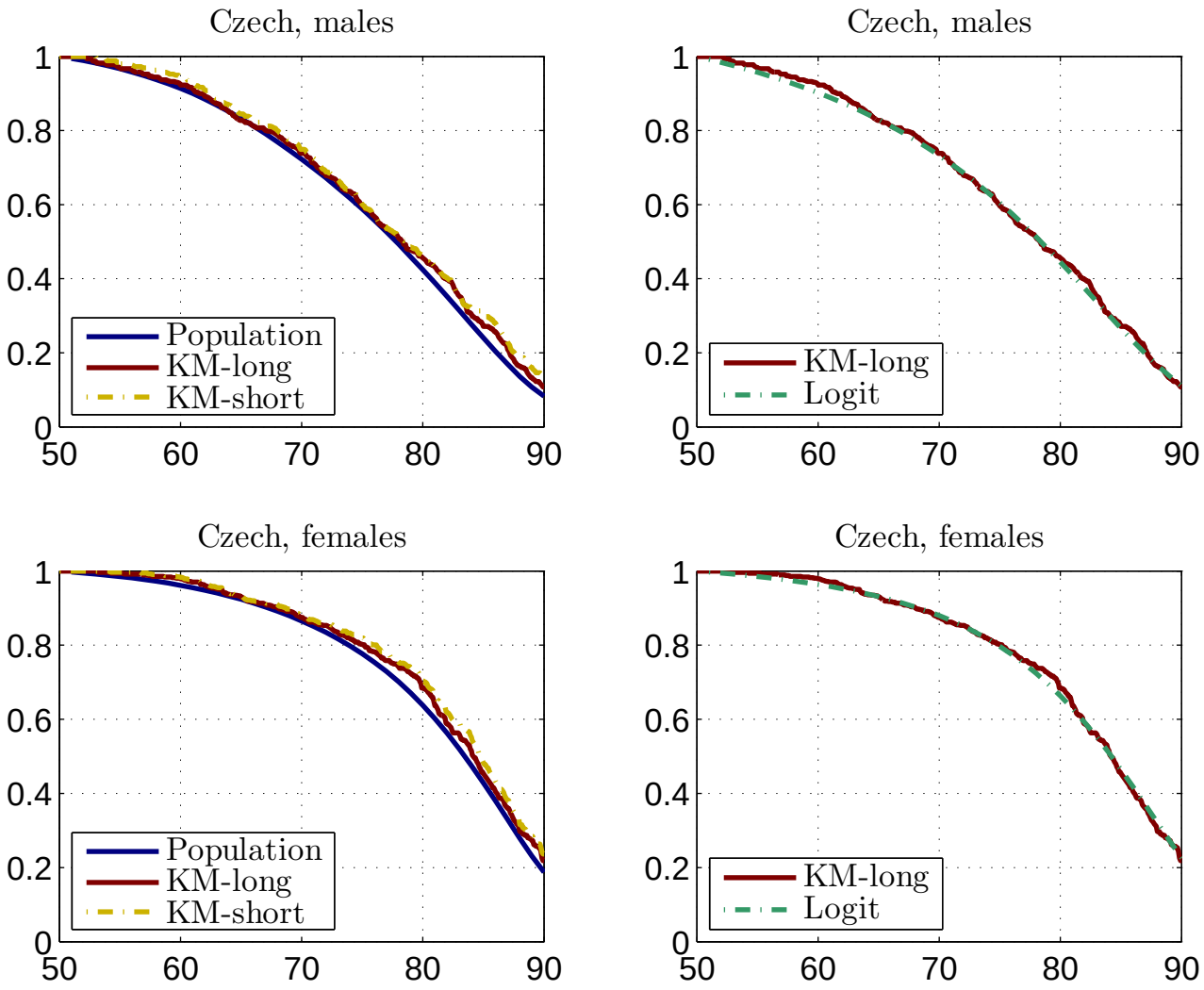
Note.

FIGURE E.2: Survival rates: Belgium



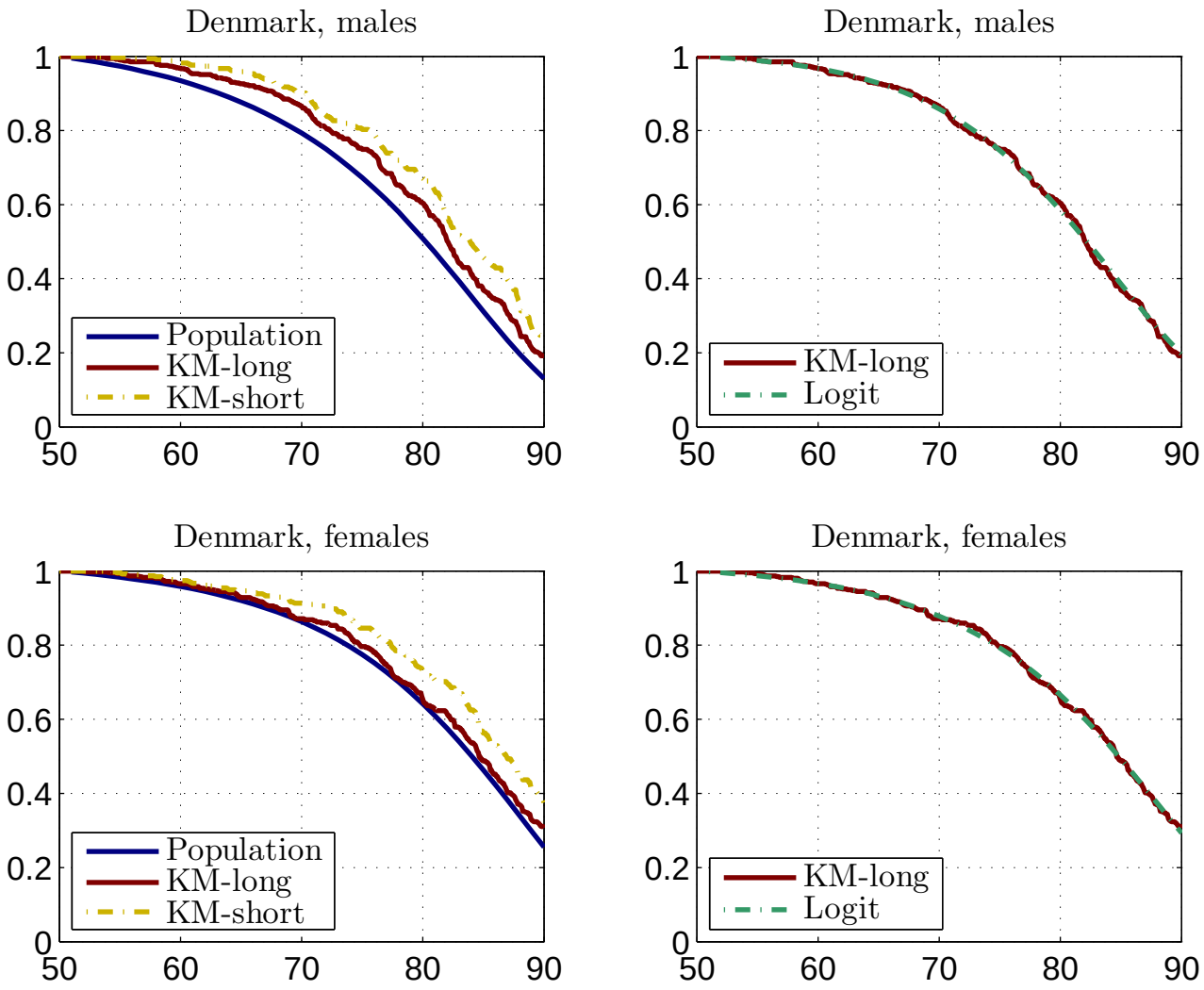
Note.

FIGURE E.3: Survival rates: Czech



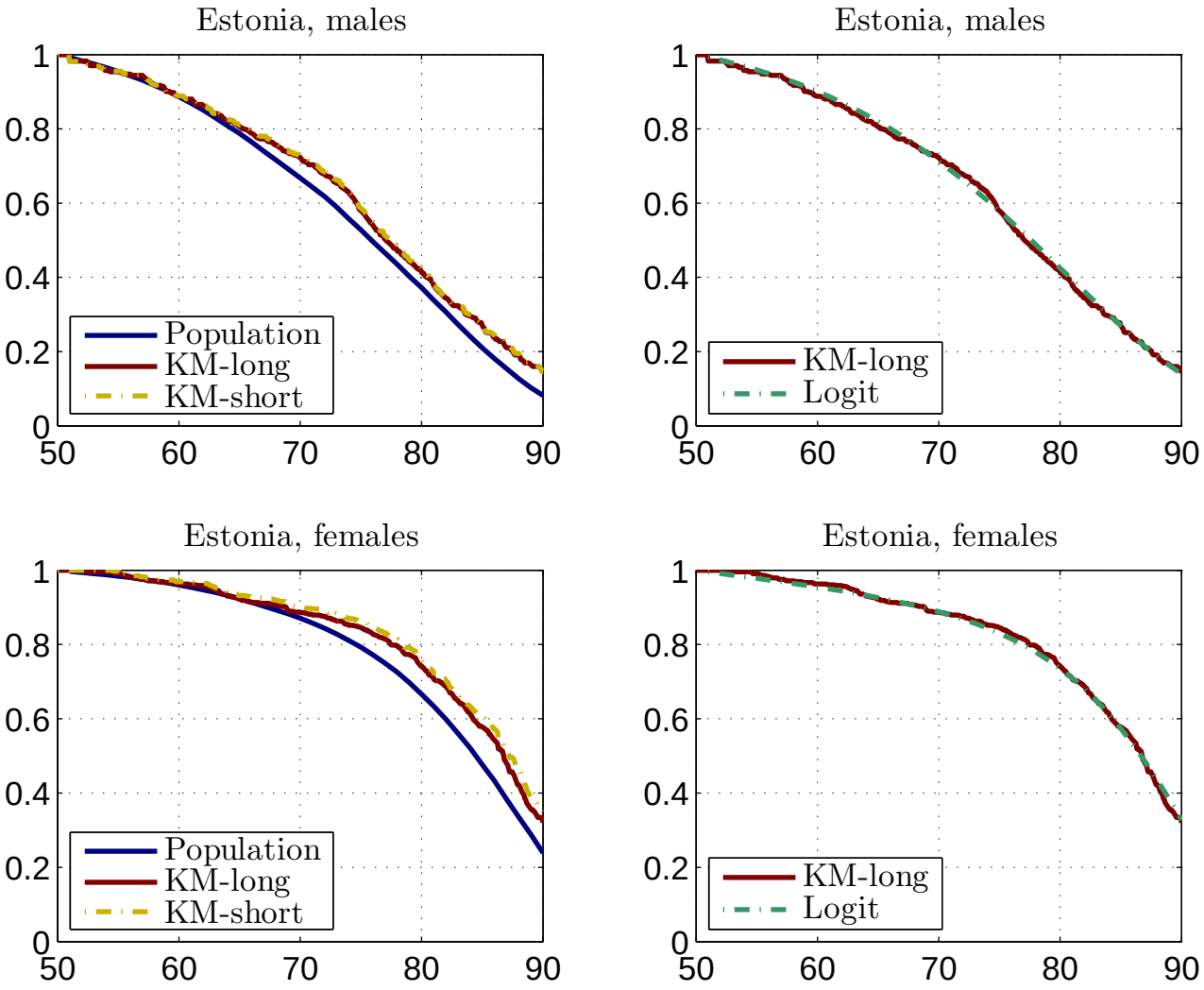
Note.

FIGURE E.4: Survival rates: Denmark



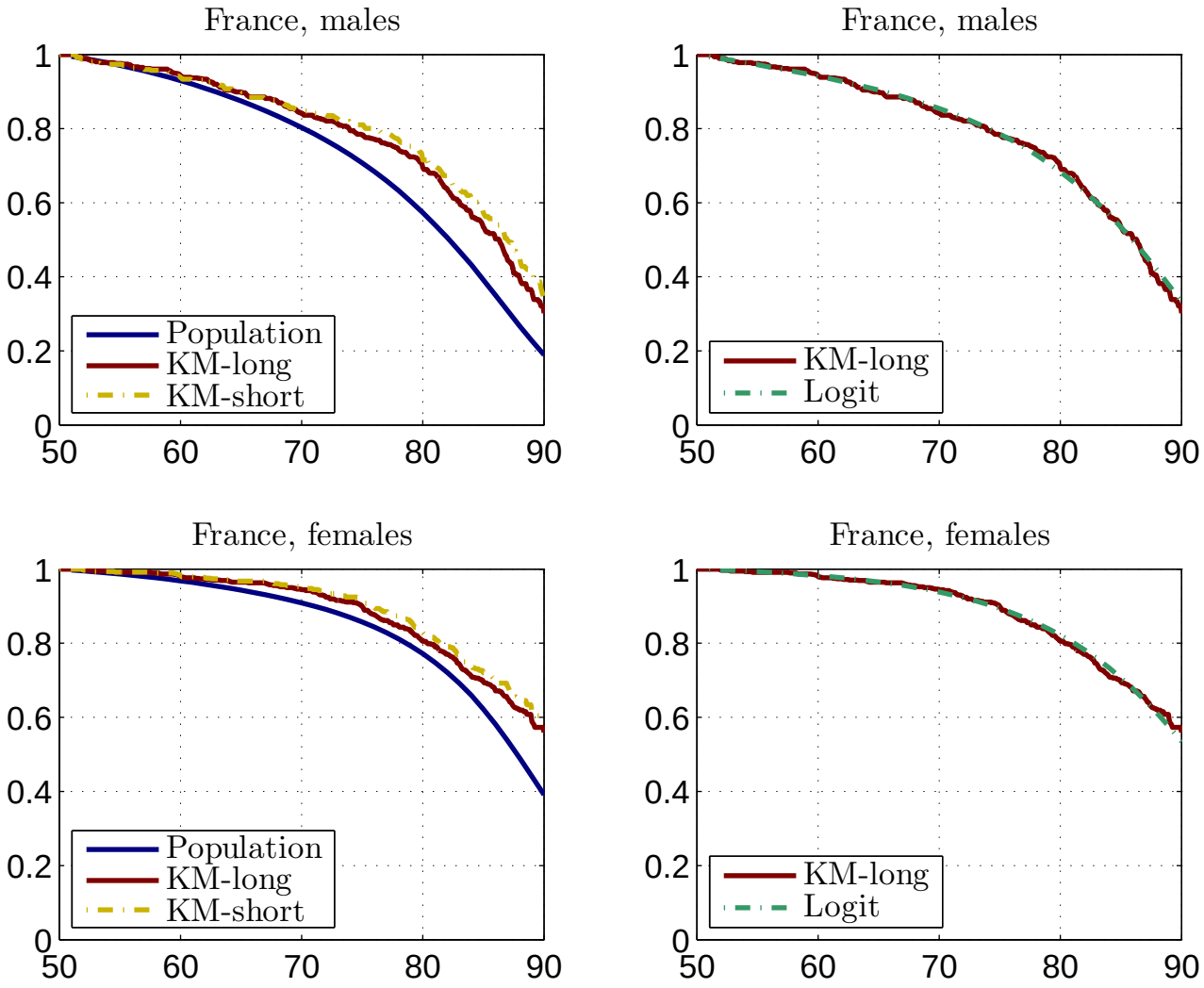
Note.

FIGURE E.5: Survival rates: Estonia



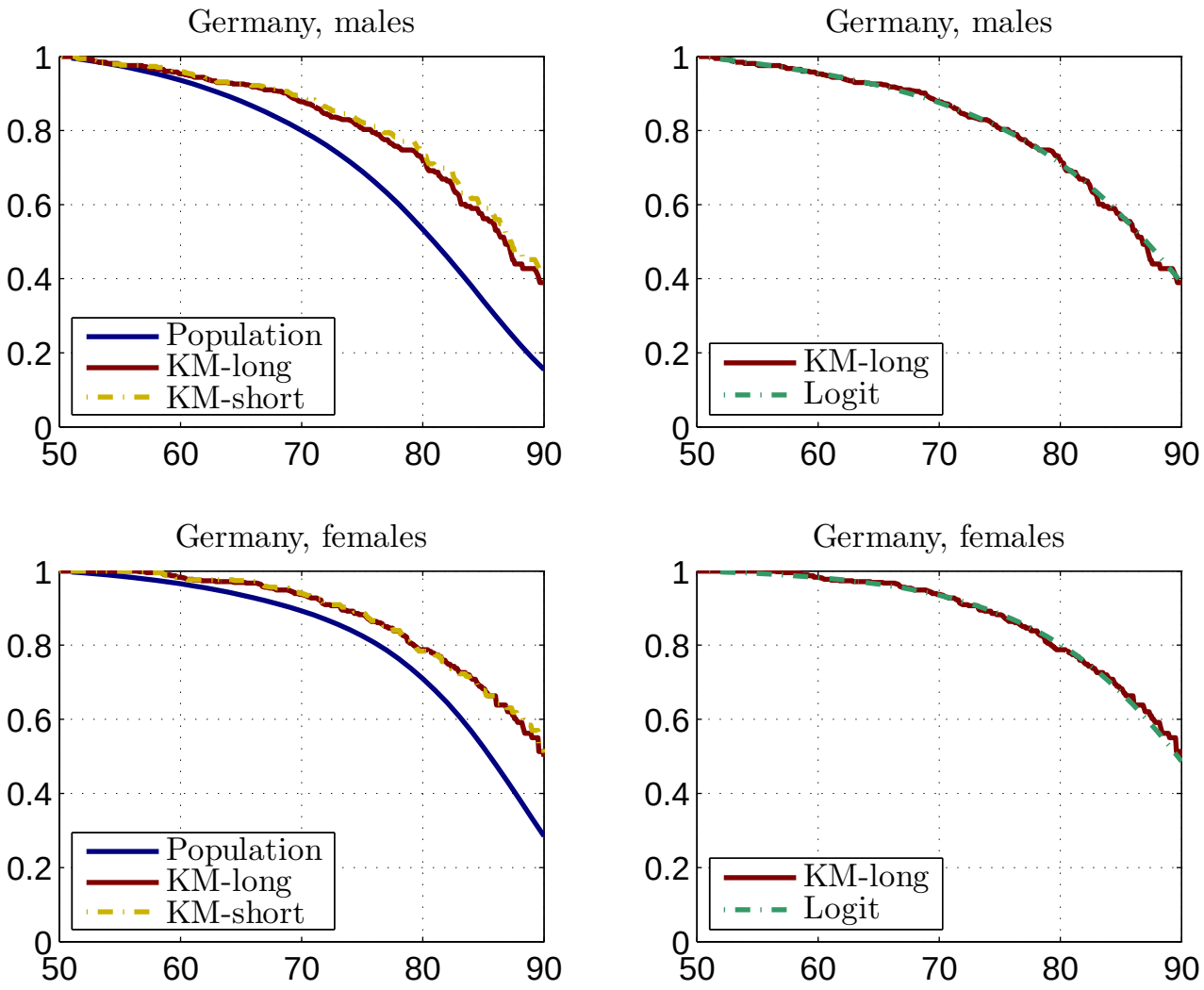
Note.

FIGURE E.6: Survival rates: France



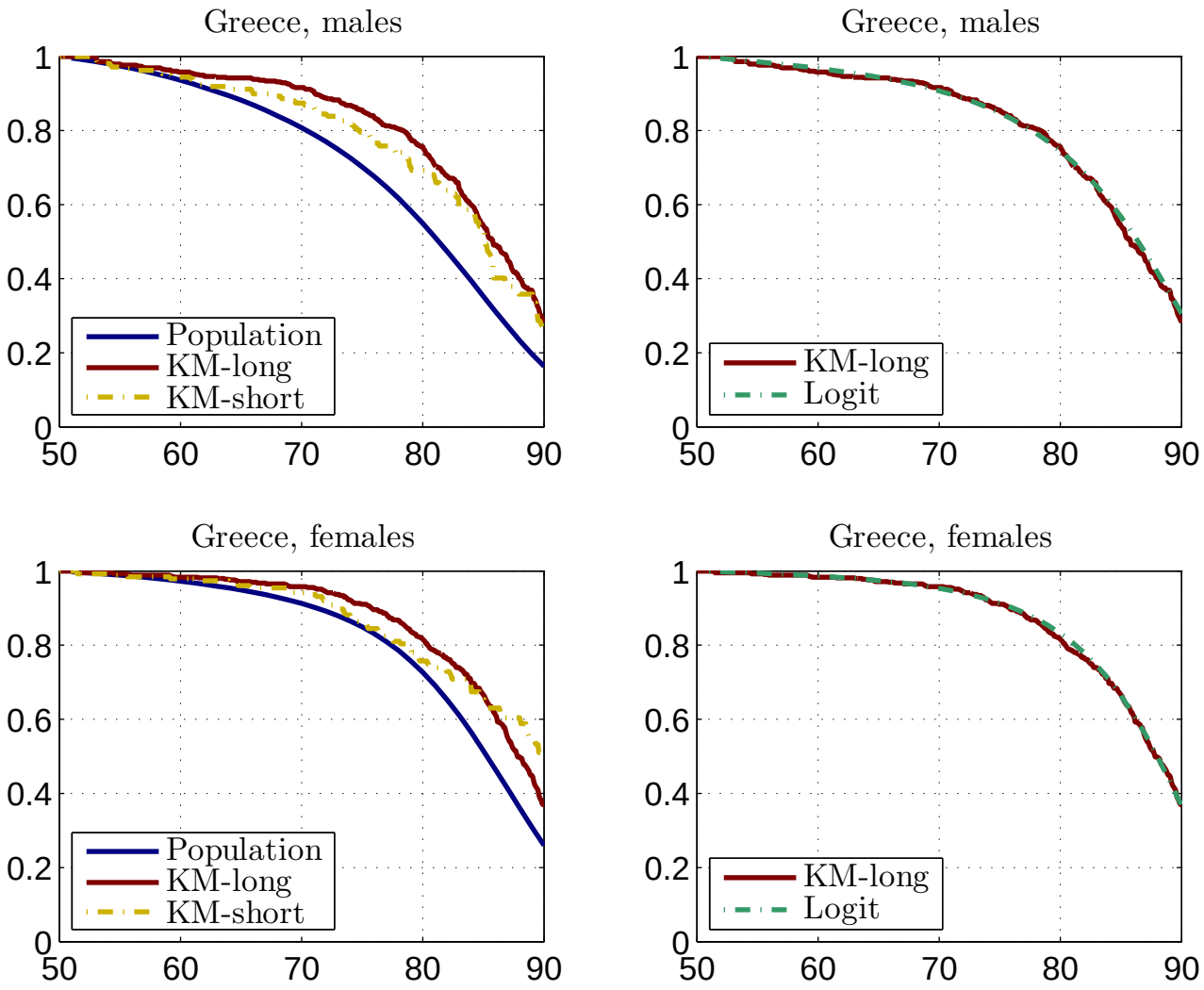
Note.

FIGURE E.7: Survival rates: Germany



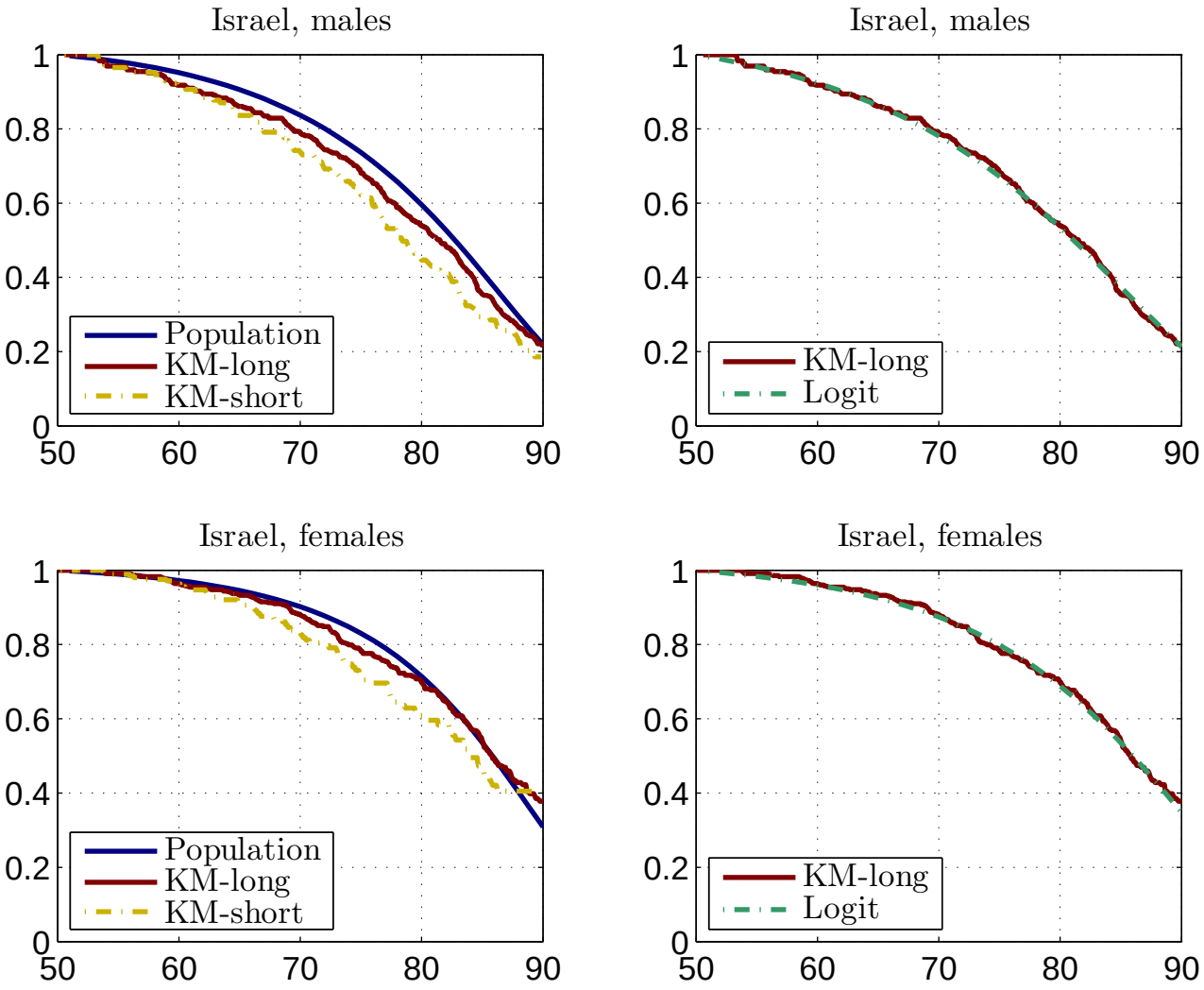
Note.

FIGURE E.8: Survival rates: Greece



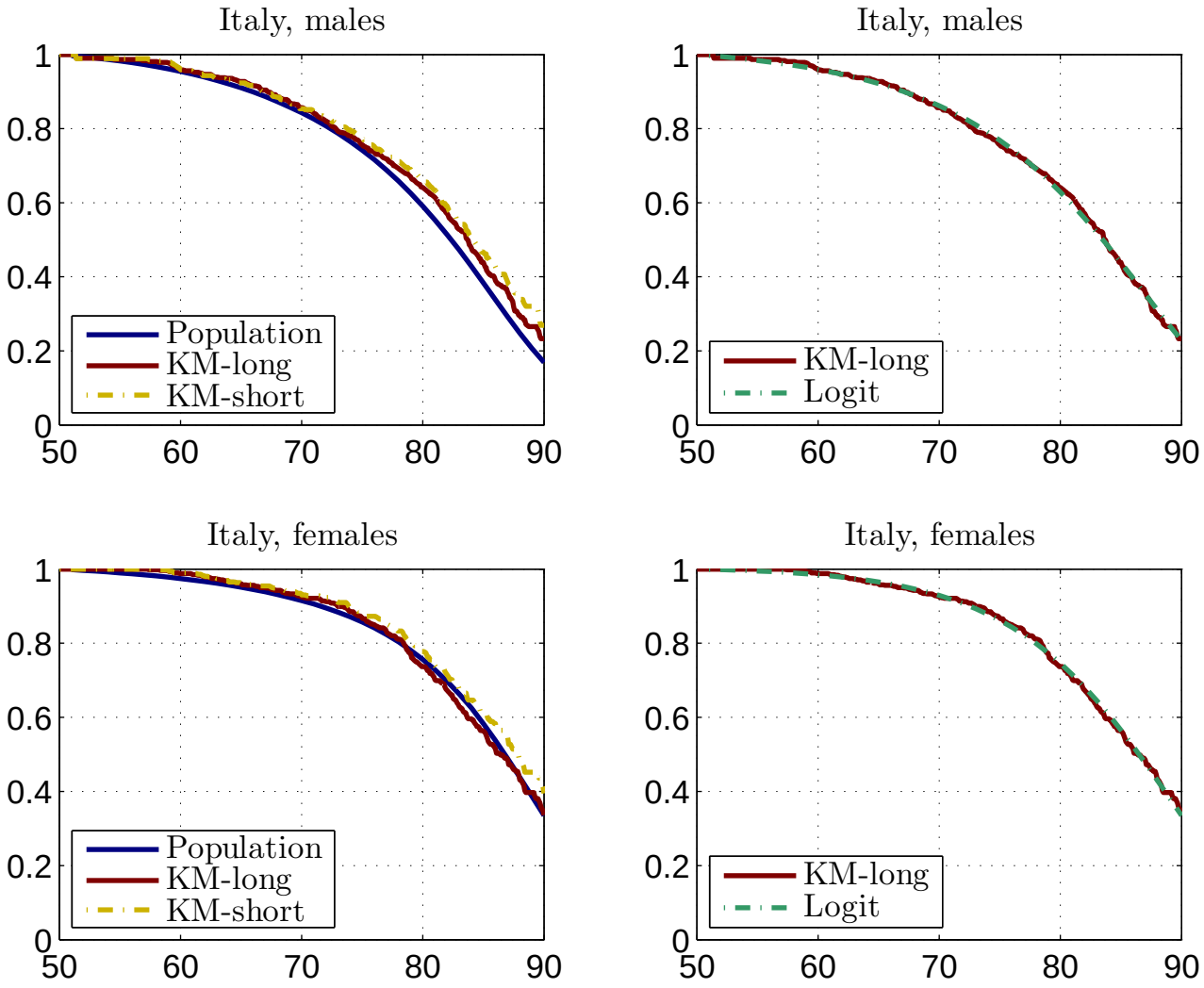
Note.

FIGURE E.9: Survival rates: Israel



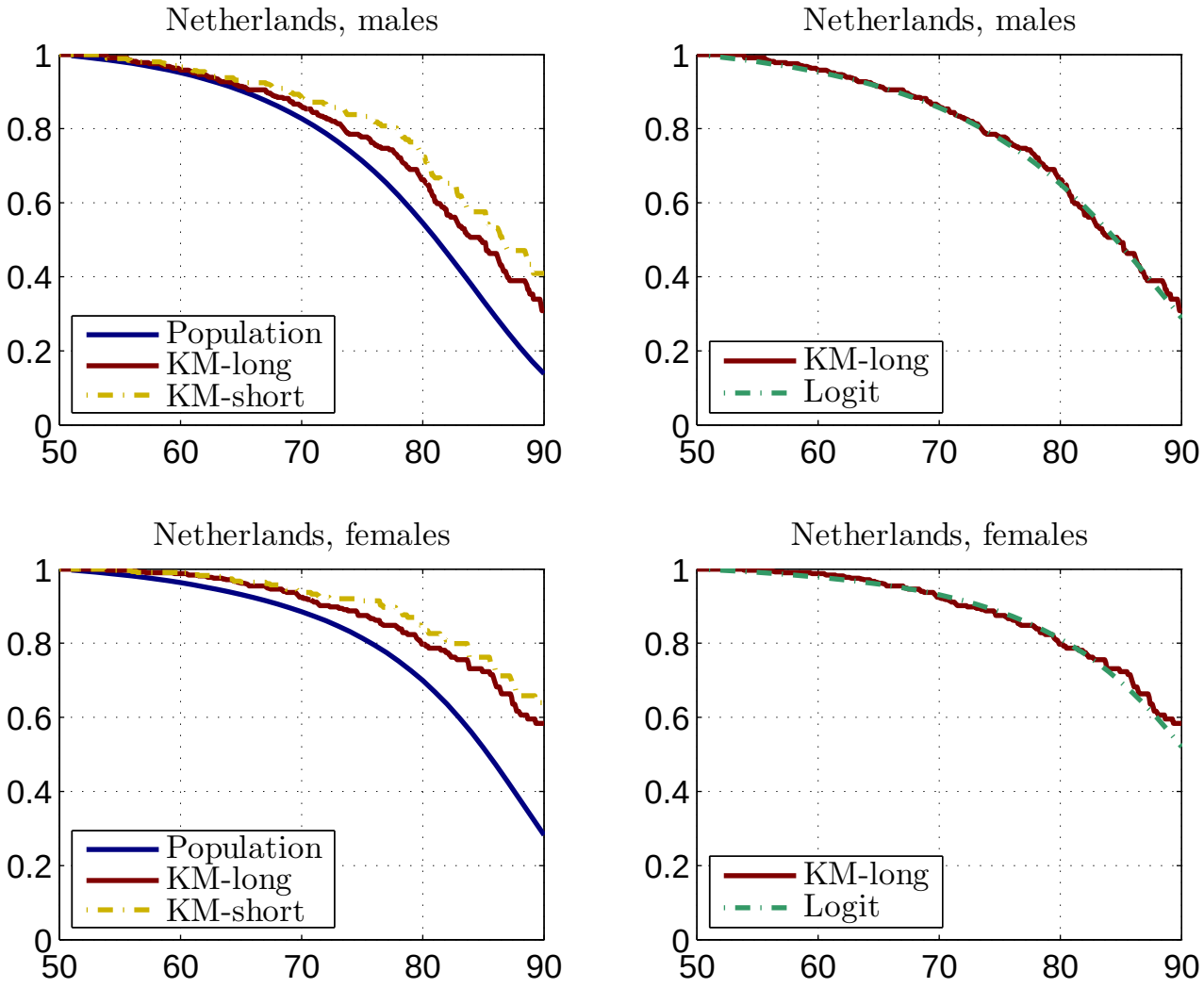
Note.

FIGURE E.10: Survival rates: Italy



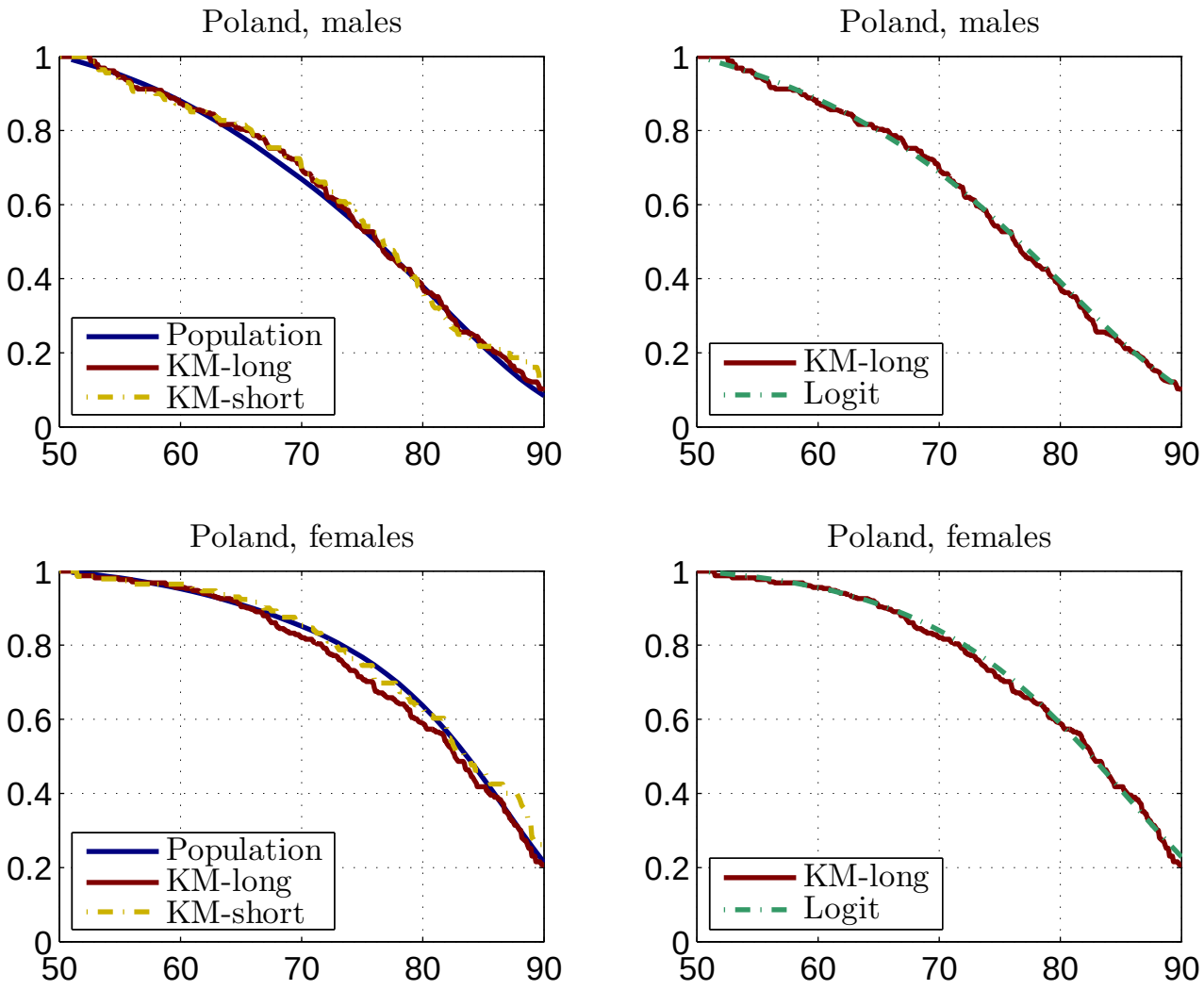
Note.

FIGURE E.11: Survival rates: Netherlands



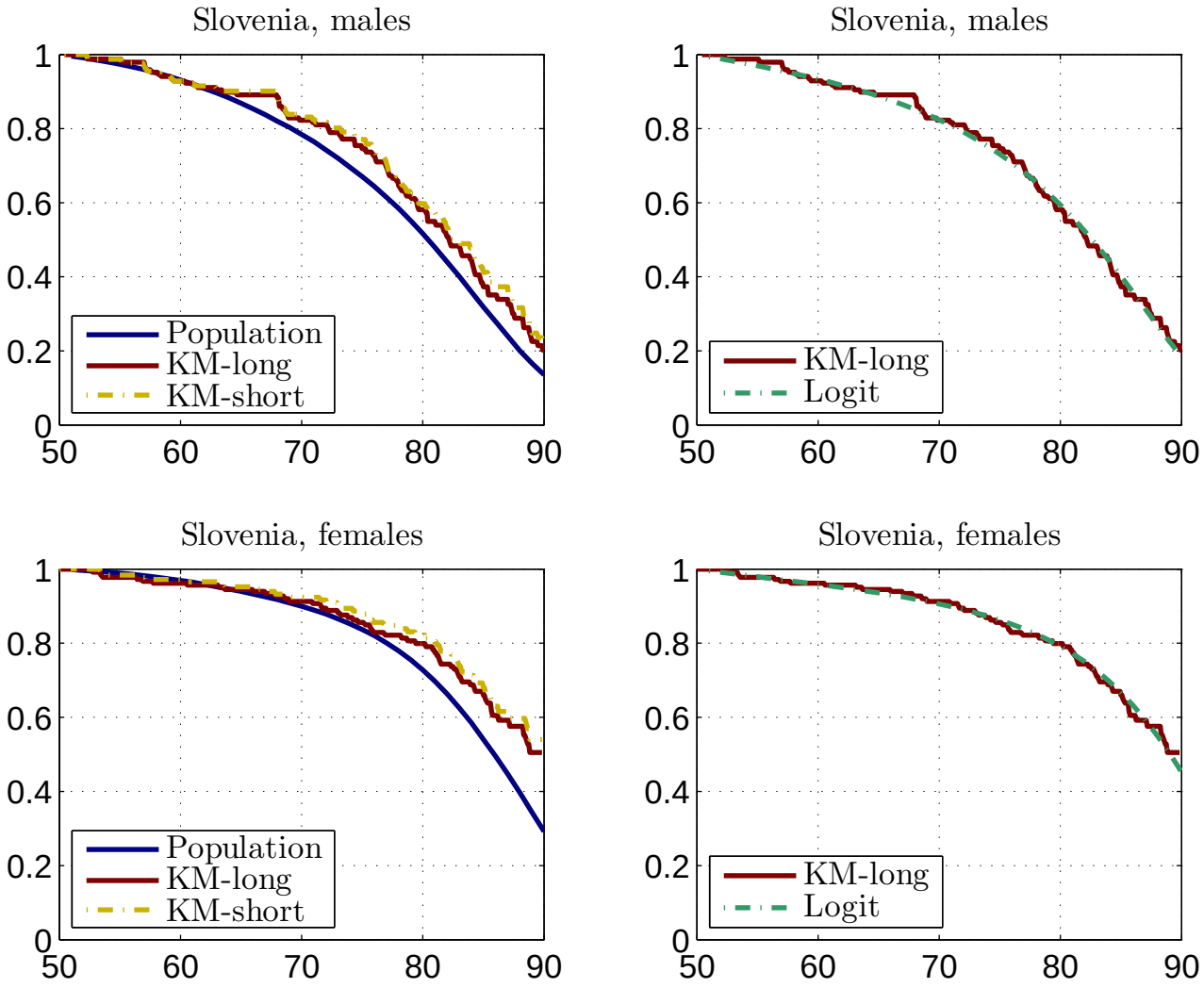
Note.

FIGURE E.12: Survival rates: Poland



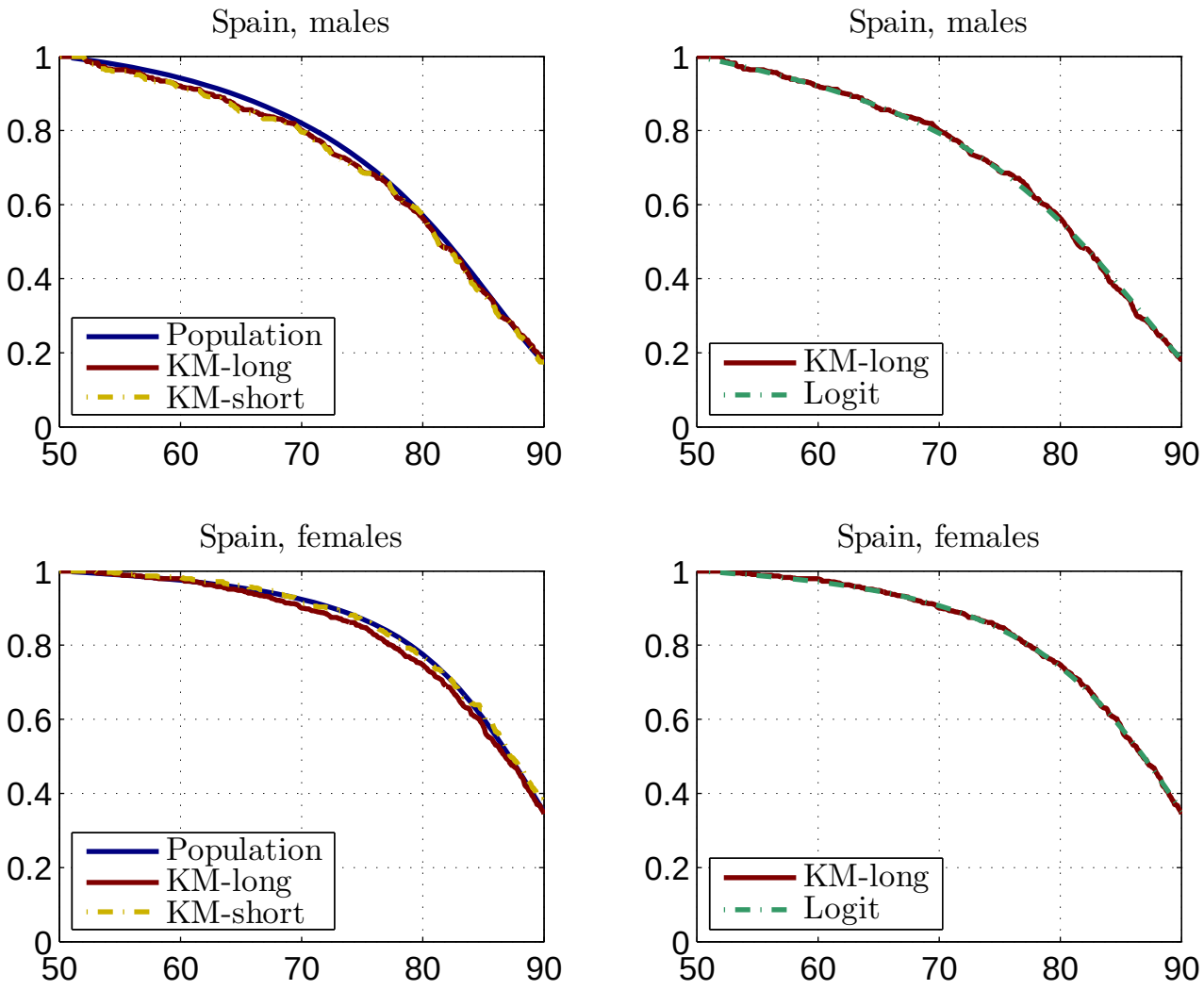
Note.

FIGURE E.13: Survival rates: Slovenia



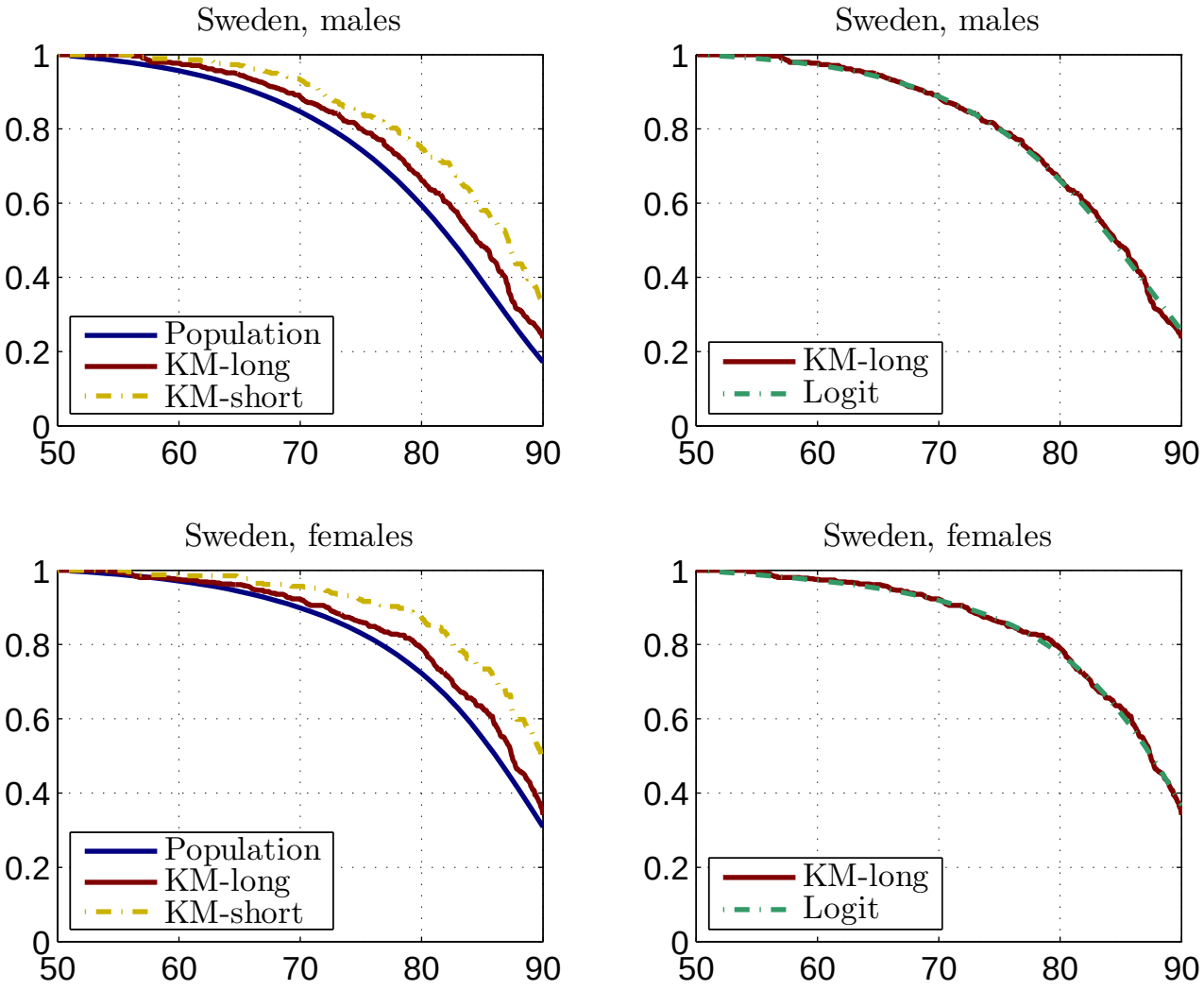
Note.

FIGURE E.14: Survival rates: Spain



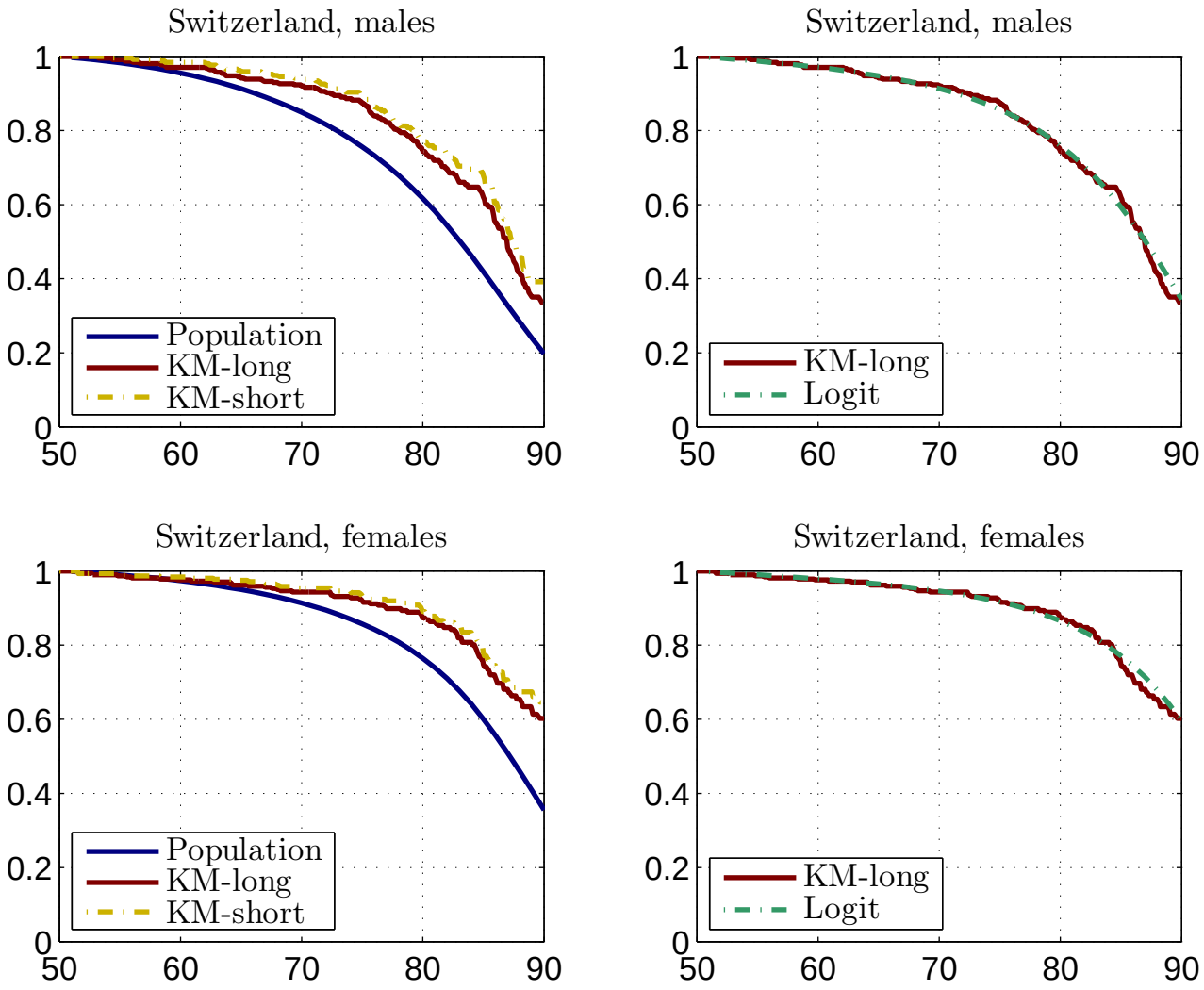
Note.

FIGURE E.15: Survival rates: Sweden



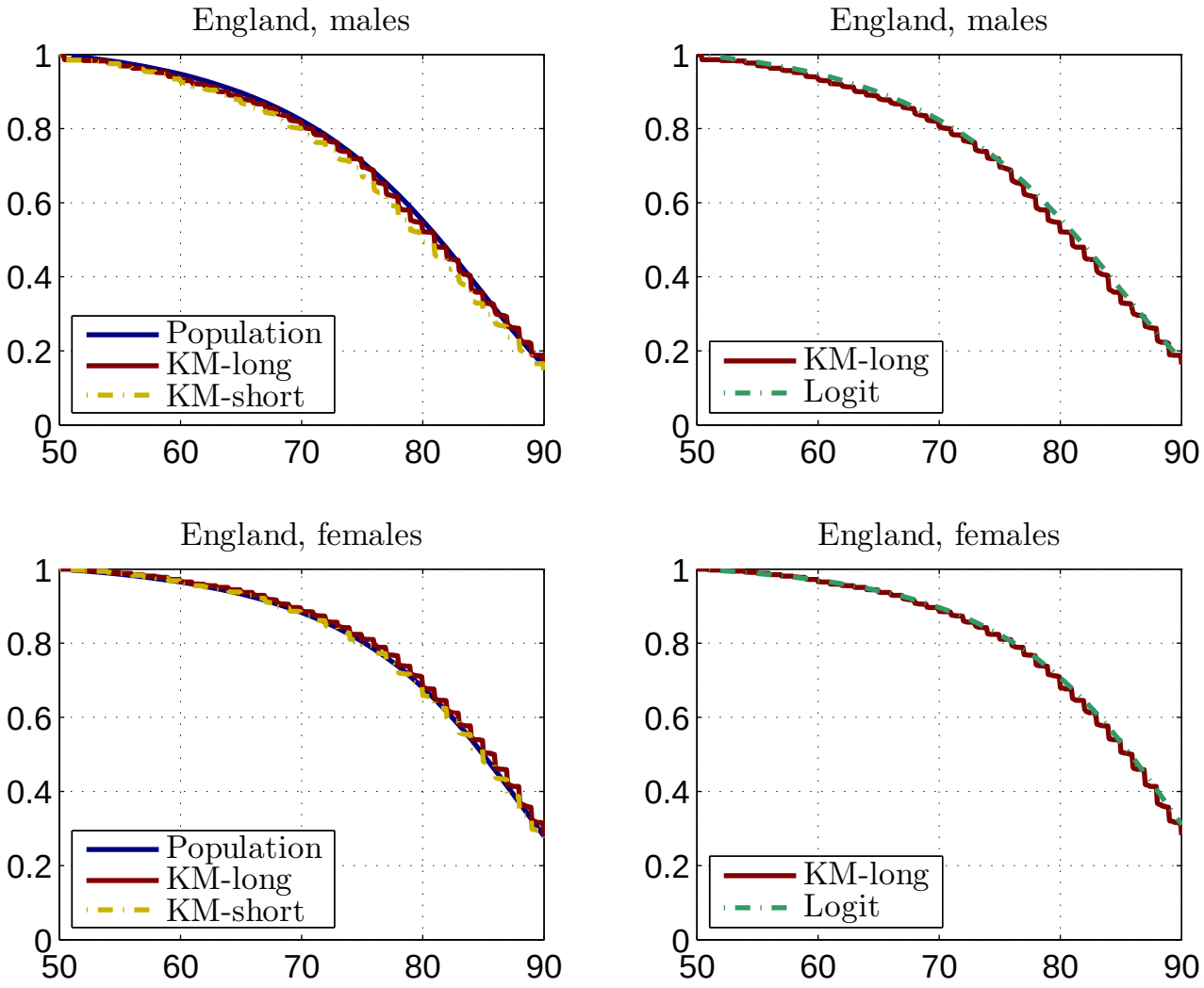
Note.

FIGURE E.16: Survival rates: Switzerland



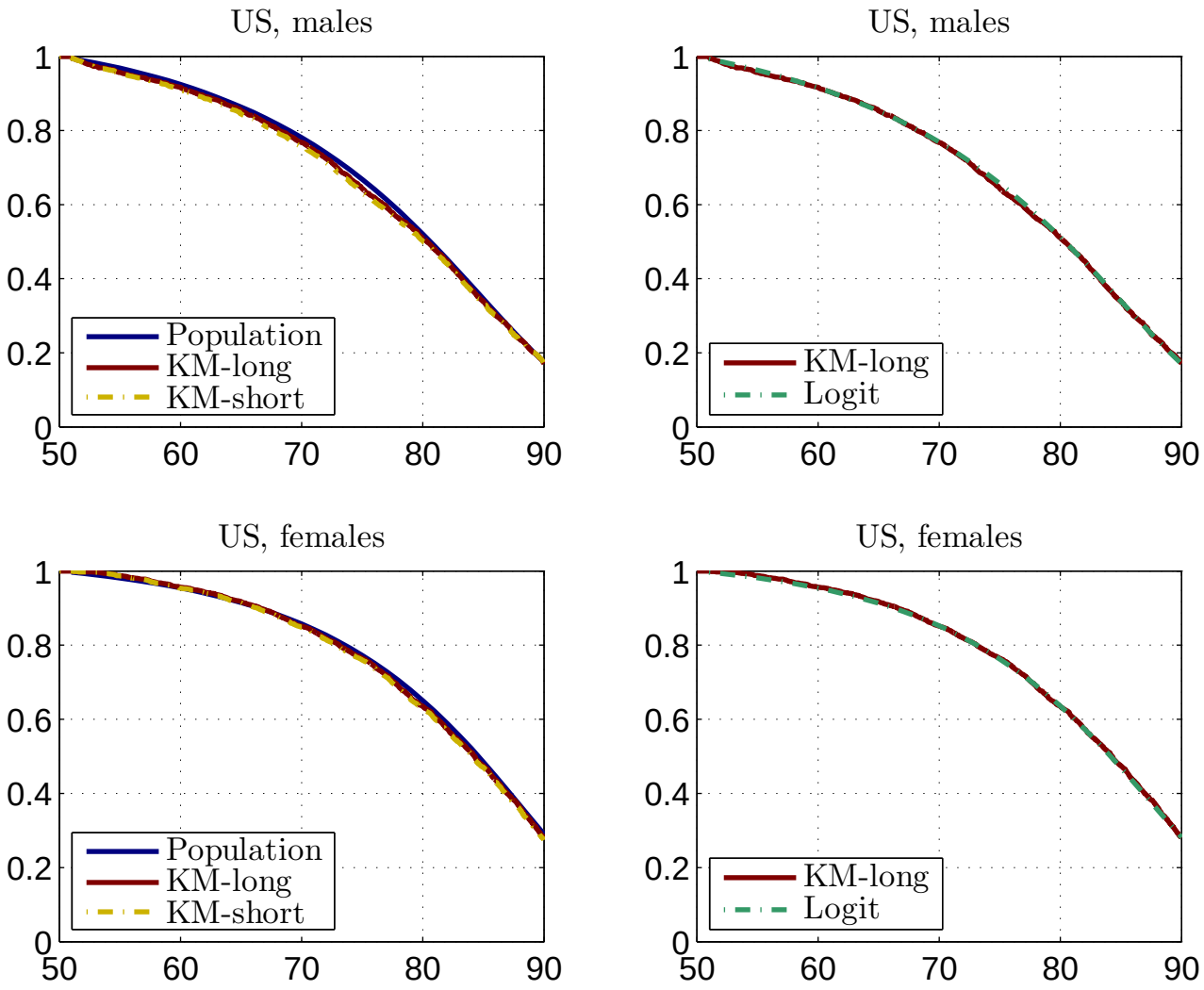
Note.

FIGURE E.17: Survival rates: England



Note.

FIGURE E.18: Survival rates: US



Note.

References

Börsch-Supan, A., M. Brandt, C. Hunkler, T. Kneip, J. Korbmacher, F. Malter, B. Schaan, S. Stuck, and S. Zuber (2013): "Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE)," *International Journal of Epidemiology*, 42, 992–1001, dOI: 10.1093/ije/dyt088.