Mortality forecasts by age and cause of death: How to forecast both components coherently?

Marie-Pier Bergeron-Boucher

Danish Center for Demographic Research (DCem)
Interdisciplinary Center on Population Dynamics (CPop)
Southern Denmark University

Odense, 28 October 2020
Project financed by SCOR Foundation for Science

Introduction

Forecasts are generally produced so that the mortality risk is known by age-group. But, mortality forecasts for other components are useful to...

- know the risk for diverse factors (age, sex, cause of death, etc.);
- plan spending more efficiently (on research, capital investment, preventive measures or palliative care);
- have better specification of the morbidity process;
- capture mortality dynamics specific to an insured population.

But ...

Introduction

There is a consensus that forecasts by cause of death are subject to many limitations:

1. Inherent pessimism (especially in linear models);
2. Can lead to unrealistic trends;
3. Modifications to the International Classification of Diseases (ICD) create discontinuities over time;
4. Inconsistent with an all-cause forecast;
5. Trajectories of causes of death are considered to be independent, yet in reality they are interconnected.

How can dependance between causes be considered?

1. Coherent models applied to mortality by cause;
2. Compositional Data Analysis (CoDA).

What is CoDA?

Instead of forecasting mortality rates, Oeppen (2008) suggested forecasting the death distributions (d_x).

By using d_x, we forecast a redistribution of deaths across ages and/or causes, such that the decrease in proportion in one age-cause will lead to an increase in at least one other.

The CoDA model

$$clr(d_{t,x*c} \ominus \alpha_{x*c}) = \beta_{x*c}\kappa_t + \epsilon_{t,x*c}$$

where

$d_{t,x*c}$ are the life table deaths at time t, age x and cause c

clr is the centered log-ratio transformation

α_{x*c} is the age-and-cause-specific average

β_{x*c} is the age-pattern of change

κ_t is the time-index

$\epsilon_{t,x*c}$ are the errors

How should the age and cause dimension be forecast?

Three options:

- Forecast the age-and-cause-of-death distribution (ACDD) simultaneously (Oeppen 2008);
- Forecast the cause-of-death distribution (CDD) at each age (top-down approach);
- Forecast the age-at-death distribution (ADD) within each cause.

Aim

To test the accuracy of the three approaches and discuss how they address the limitations listed previously.

Approach 1 (A1): Forecast the age-and-cause-of-death distribution (ACDD) simultaneously

We compared 2 models:

1. One common time-trend for age and cause

$$A1_{1T}: \qquad clr(d_{t,x*c} \ominus \alpha_{x*c}) = \beta_{x*c}\kappa_t + \epsilon_{t,x*c}, \tag{1}$$

2. Multiple time-trends, one for each cause

$$A1_{MT}: \qquad clr(d_{t,x} \ominus \alpha_x) = \sum_{c=1}^{C} \beta_x^c \kappa_t^c + \epsilon_{t,x}^c. \tag{2}$$

Approach 2 (A2): Forecast the CDD at each age

2 steps:

- Forecast mortality by age only, using a standard CoDA model

$$clr(d_{t,x} \ominus \alpha_x) = \beta_x \kappa_t + \epsilon_{t,x}. \tag{3}$$

- Forecast the CDD specific at each age. Assuming that the CDD within each age-group are independent, the model can read as

$$A2_{Ind}: \qquad clr(d^x_{t,c} \ominus \alpha^x_c) = \beta^x_c \kappa^x_t + \epsilon^x_{t,c}. \tag{4}$$

Approach 2 (A2) with coherence between age-specific CDD

Assuming that the CDD within each age-group are not independent, and are driven by a similar trend, the second step of the approach is

$$A2_{Coh}: \qquad clr(d^x_{t,c} \ominus \alpha^x_c) = B_c K_t + \beta^x_c \kappa^x_t + \epsilon^x_{t,c} \tag{5}$$

Approach 3 (A3): Forecast the ADD within each cause

2 steps:

1. Forecast mortality by cause only, using a standard CoDA model:

$$clr(d_{t,c} \ominus \alpha_c) = \beta_c \kappa_t + \epsilon_{t,c}. \tag{6}$$

2. Forecast the ADD specific to each cause. Assuming that the ADD within each cause are independent, the model can read as

$$A3_{Ind}: \qquad clr(d^c_{t,x} \ominus \alpha^c_x) = \beta^c_x \kappa^c_t + \epsilon^c_{t,x}. \tag{7}$$

Approach 3 (A3) with coherence between cause-specific ADD

Assuming that the ADD are not independent between causes, the model is:

$$A3_{Coh}: \qquad clr(d^c_{t,x} \ominus \alpha^c_x) = B_x K_t + \beta^c_x \kappa^c_t + \epsilon^c_{t,x}. \tag{8}$$

Approach 3 (A3)

When forecasting the CDD, the time trends are assumed to follow (1) a random-walk with drift ($A3_{Ind,RW}$, $A3_{Coh,RW}$) and (2) a Holt-damped trend ($A3_{Ind,HD}$, $A3_{Coh,HD}$), so that the causes of death distribution eventually reach a constant.

Data

Multiple-decrement life tables were calculated from the Human Mortality Database (2020) life tables and WHO causes-of-death data.

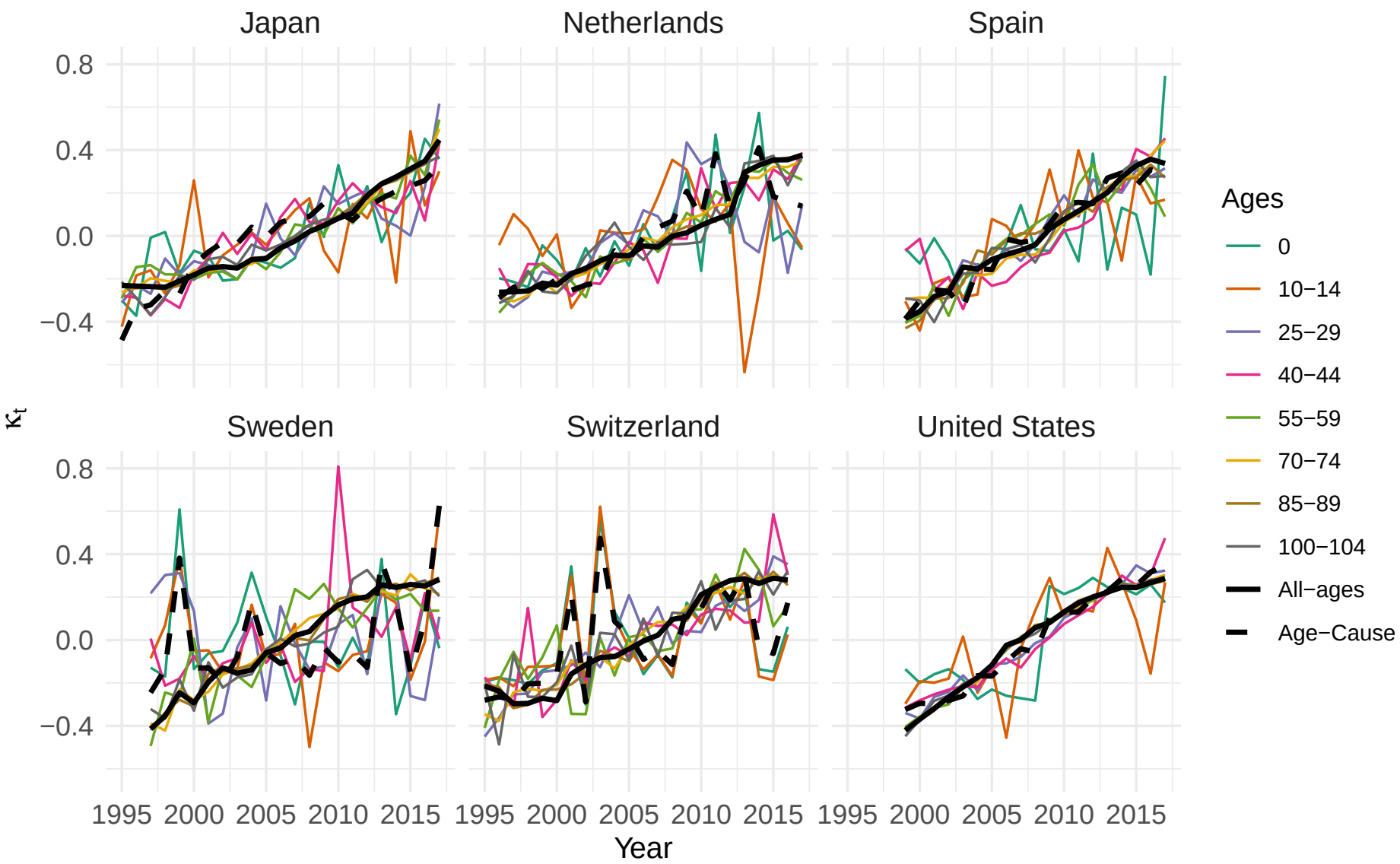
Countries: Japan (1995–2017), the Netherlands (1996–2017), Spain (1999-2017), Sweden (1997–2017), Switzerland (1995–2016) and the United States (1999–2017)

Sex: Females and males

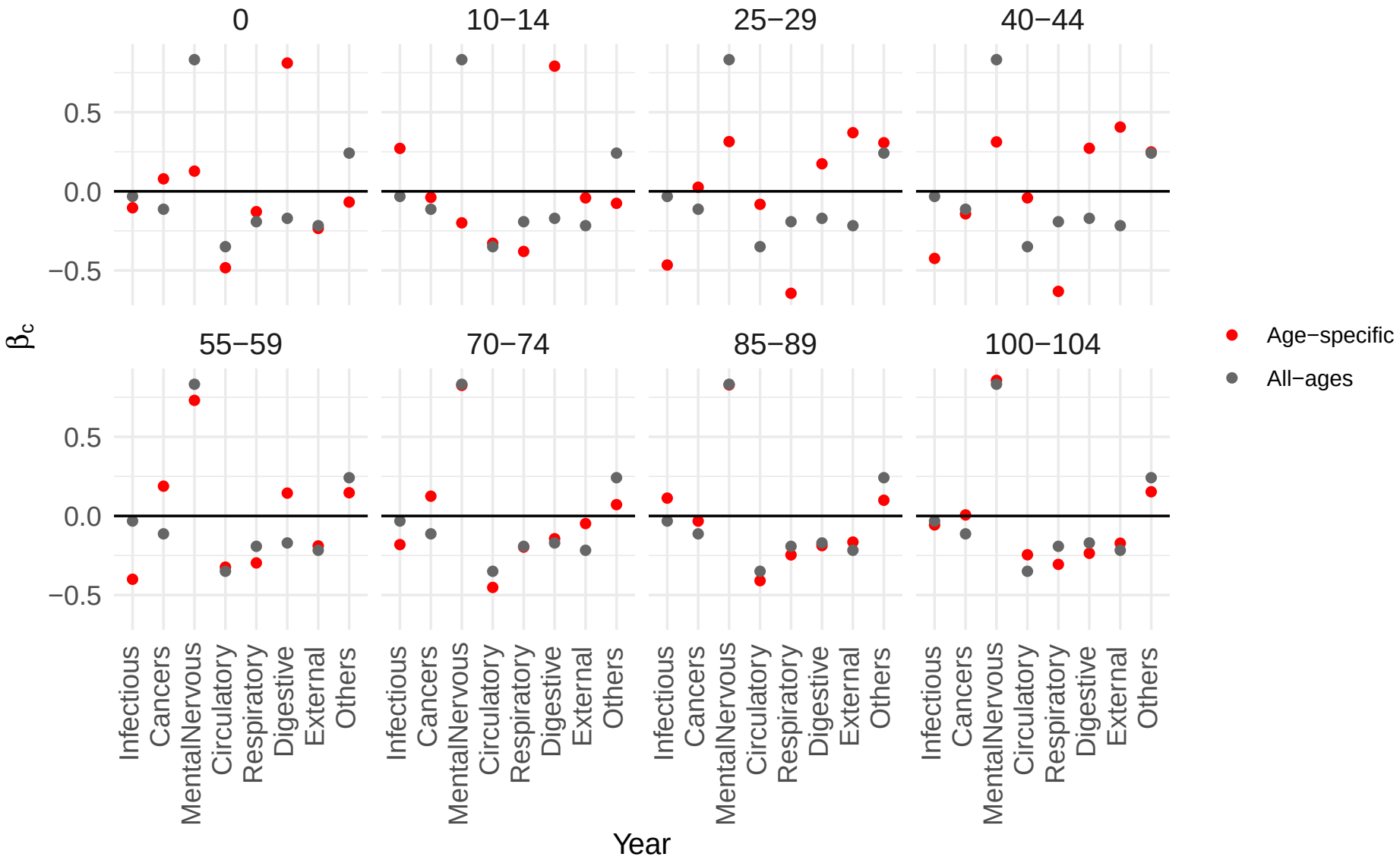
Ages: 0-110 by 5-year age-groups.

Causes: Infectious and parasitic diseases; Cancers; Mental and behavioral disorders and diseases of the nervous system; Diseases of the circulatory system; Diseases of the respiratory system; Diseases of the digestive system; External causes; Others.

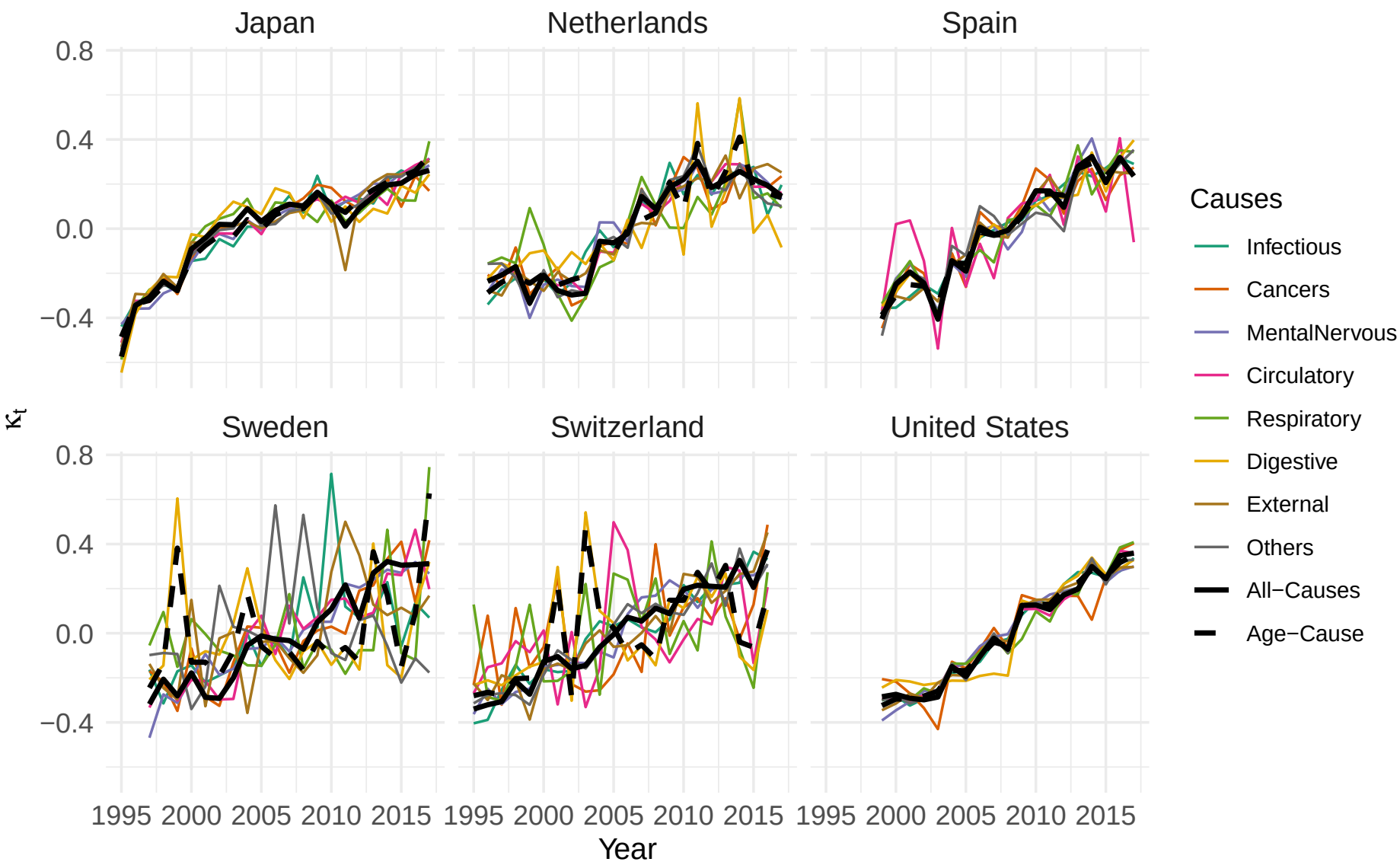
Parameters: Age-specific κ_t, Females



Parameters: Age-specific β_c, Japan, Females



Parameters: Cause-specific κ_t, Females



Parameters: Cause-specific β_x, Japan, Females

Out-of-sample analysis, Females

	$A1_{1T}$	$A1_{MT}$	$A2_{Ind}$	$A2_{Coh}$	$A3_{Ind,RW}$	$A3_{Ind,HD}$	$A3_{Coh,RW}$	$A3_{Coh,HD}$
	Causes-of-death distribution							
Japan	0.329	0.336	0.321	**0.302**	0.341	0.327	0.341	0.327
Netherlands	0.390	0.386	**0.376**	0.380	0.389	0.417	0.389	0.417
Sweden	0.276	**0.199**	0.292	0.325	0.288	0.231	0.288	0.231
Switzerland	0.228	**0.216**	0.217	0.278	0.230	0.239	0.230	0.239
Spain	**0.179**	0.197	0.180	0.241	0.189	0.188	0.189	0.188
United States	0.196	0.209	0.205	0.204	0.203	**0.135**	0.203	**0.135**
Mean	0.266	0.257	0.265	0.288	0.273	**0.256**	0.273	**0.256**
		(2)	(3)			(1)		(1)
	Age-at-death distribution							
Japan	0.801	0.803	**0.773**	**0.773**	0.813	0.821	0.884	0.886
Netherlands	0.886	0.837	0.837	0.837	0.821	**0.820**	1.019	1.017
Sweden	0.868	**0.808**	0.819	0.819	0.831	0.826	1.177	1.174
Switzerland	0.792	0.754	**0.739**	**0.739**	0.757	0.757	0.981	0.980
Spain	**0.627**	0.596	0.632	0.632	0.599	0.597	0.804	0.802
United States	0.427	0.434	**0.391**	**0.391**	0.430	0.429	0.524	0.522
Mean	0.734	0.705	**0.699**	**0.699**	0.709	0.708	0.898	0.897
		(2)	(1)	(1)		(3)		

Out-of-sample analysis, Females

	$A1_{1T}$	$A1_{MT}$	$A2_{Ind}$	$A2_{Coh}$	$A3_{Ind,RW}$	$A3_{Ind,HD}$	$A3_{Coh,RW}$	$A3_{Coh,HD}$
	Age-and-cause-of-death distribution							
Japan	3.787	3.803	3.832	**3.774**	3.814	3.817	3.929	3.982
Netherlands	7.105	7.198	7.183	7.352	7.178	7.240	**6.606**	6.634
Sweden	8.387	8.869	8.513	**8.346**	8.946	8.970	8.811	8.767
Switzerland	7.891	7.857	7.765	**7.331**	7.916	7.937	7.476	7.492
Spain	4.068	4.156	4.064	**4.027**	4.194	4.221	4.181	4.140
United States	**2.211**	2.356	2.228	2.272	2.288	2.310	2.508	2.384
Mean	5.575	5.697	5.598	**5.517**	5.723	5.749	5.585	5.566
	(3)			(1)				(2)
	Life expectancy at birth							
Japan	0.512	**0.508**	0.537	0.537	0.516	0.520	0.625	0.628
Netherlands	0.294	0.227	0.207	0.207	**0.193**	0.194	0.313	0.313
Sweden	0.476	**0.154**	0.206	0.206	0.255	0.252	0.329	0.329
Switzerland	0.419	0.215	**0.188**	**0.188**	0.293	0.294	0.199	0.200
Spain	0.308	0.240	0.287	0.287	0.236	**0.235**	0.304	0.307
United States	0.597	0.541	**0.381**	**0.381**	0.526	0.449	0.525	0.465
Mean	0.434	0.314	**0.301**	**0.301**	0.337	0.324	0.383	0.374
		(2)	(1)	(1)		(3)		

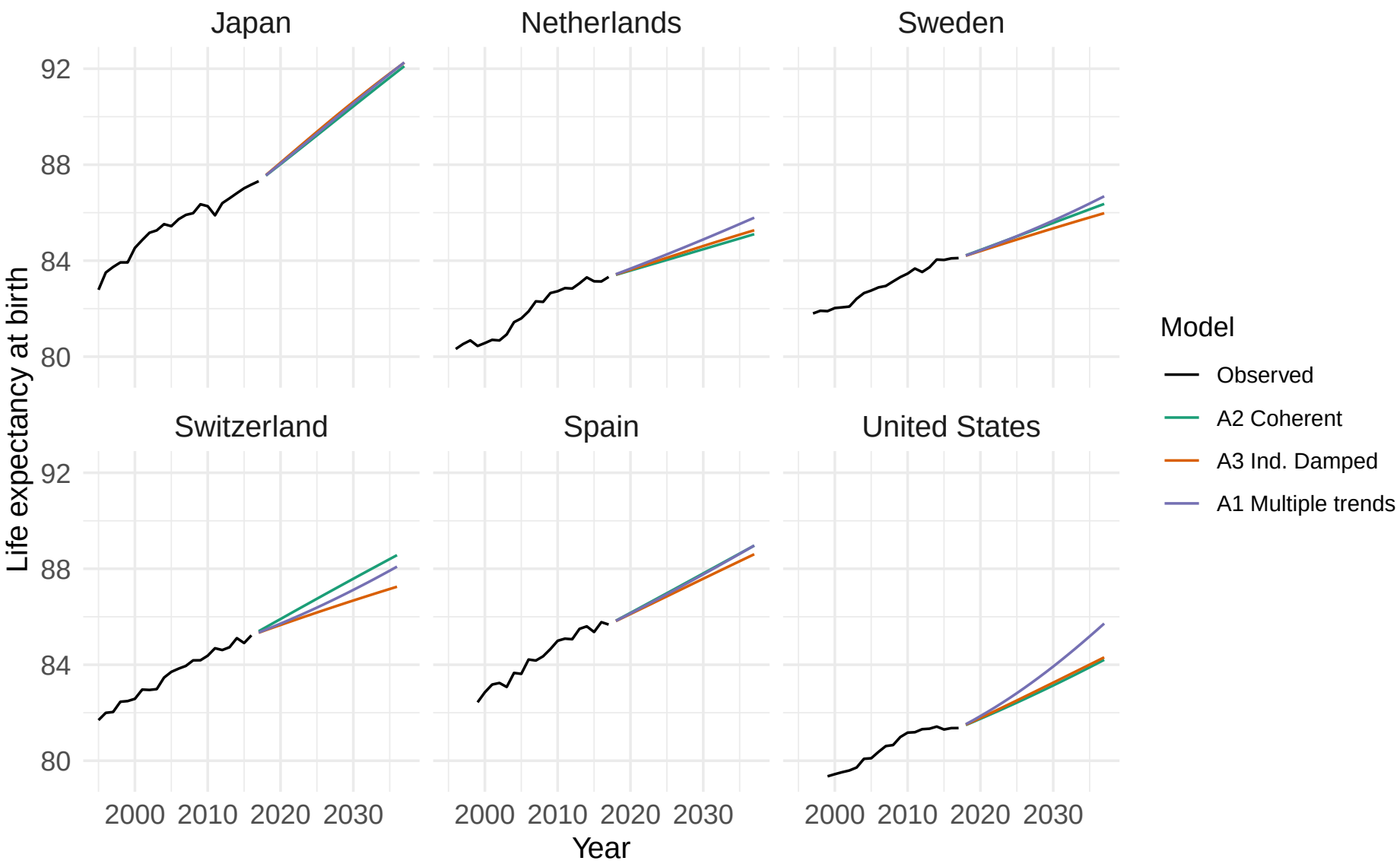
Out-of-sample analysis, Males

	$A1_{1T}$	$A1_{MT}$	$A2_{Ind}$	$A2_{Coh}$	$A3_{Ind,RW}$	$A3_{Ind,HD}$	$A3_{Coh,RW}$	$A3_{Coh,HD}$
	Causes-of-death distribution							
Japan	0.292	0.293	0.280	**0.241**	0.287	0.246	0.287	0.246
Netherlands	0.391	0.401	**0.389**	0.415	0.412	0.443	0.412	0.443
Sweden	0.275	0.244	0.290	0.341	0.277	**0.232**	0.277	**0.232**
Switzerland	0.179	**0.173**	0.198	0.318	0.241	0.217	0.241	0.217
Spain	0.183	0.191	0.179	0.251	0.188	**0.175**	0.188	**0.175**
United States	0.195	0.207	0.197	0.199	0.177	**0.123**	0.177	**0.123**
Mean	0.252	0.252	0.256	0.294	0.264	**0.239**	0.264	**0.239**
	(3)	(2)				(1)		(1)
	Age-at-death distribution							
Japan	0.680	0.686	**0.670**	**0.670**	0.689	0.687	0.878	0.871
Netherlands	0.728	**0.716**	0.730	0.730	0.718	0.721	0.898	0.898
Sweden	0.769	**0.721**	0.757	0.757	0.728	0.726	1.005	0.999
Switzerland	0.741	0.711	0.792	0.792	**0.703**	0.705	0.948	0.950
Spain	**0.587**	0.594	0.599	0.599	0.590	0.589	0.760	0.752
United States	0.649	0.645	**0.619**	**0.619**	0.639	0.641	0.704	0.708
Mean	0.692	0.679	0.694	0.694	**0.678**	0.678	0.866	0.863
		(3)			(1)	(2)		

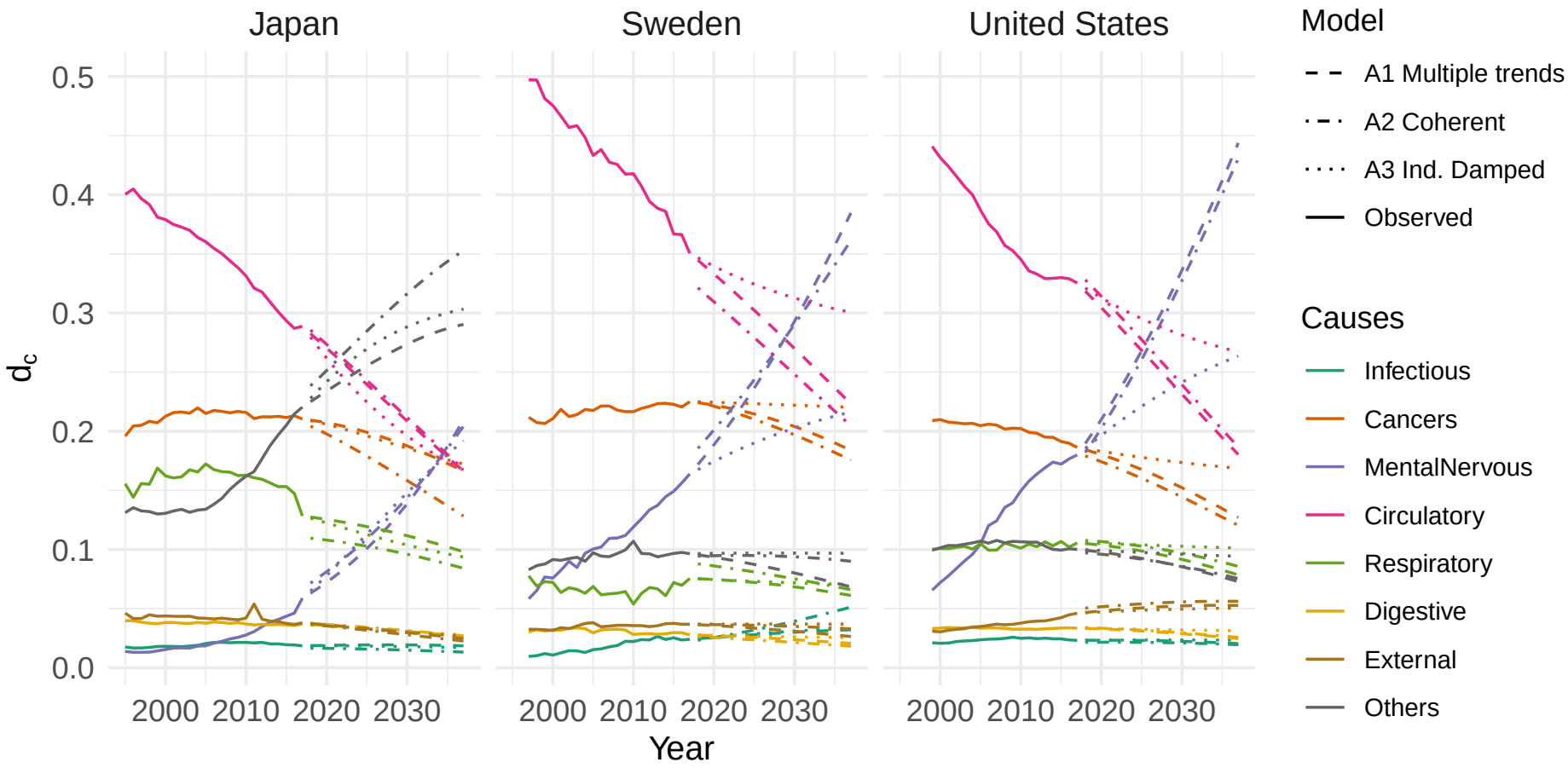
Out-of-sample analysis, Males

	$A1_{1T}$	$A1_{MT}$	$A2_{Ind}$	$A2_{Coh}$	$A3_{Ind,RW}$	$A3_{Ind,HD}$	$A3_{Coh,RW}$	$A3_{Coh,HD}$
	Age-and-cause-of-death distribution							
Japan	3.335	3.351	3.353	**3.270**	3.344	3.300	3.656	3.663
Netherlands	5.398	5.548	5.548	**5.148**	5.601	5.667	5.354	5.397
Sweden	7.058	7.343	7.014	**7.013**	7.435	7.397	7.195	7.064
Switzerland	10.692	11.263	10.716	10.359	11.434	11.420	10.083	**10.065**
Spain	4.276	4.444	4.278	**3.971**	4.454	4.443	4.232	4.211
United States	2.504	2.517	2.522	**2.486**	2.512	2.566	2.673	2.576
Mean	5.544	5.744	5.572	**5.374**	5.797	5.799	5.532	5.496
				(1)			(3)	(2)
	Life expectancy at birth							
Japan	**0.263**	0.270	0.290	0.290	0.273	0.268	0.478	0.475
Netherlands	0.381	0.225	0.372	0.372	**0.209**	0.211	0.380	0.378
Sweden	0.717	0.311	**0.098**	**0.098**	0.400	0.384	0.186	0.174
Switzerland	**0.227**	0.295	0.364	0.364	0.276	0.280	0.266	0.268
Spain	0.312	0.252	0.311	0.311	**0.236**	0.240	0.341	0.331
United States	0.859	0.833	0.709	0.709	0.771	0.739	0.711	**0.691**
Mean	0.459	0.364	0.357	0.357	0.361	**0.354**	0.393	0.386
			(2)	(2)	(3)	(1)		

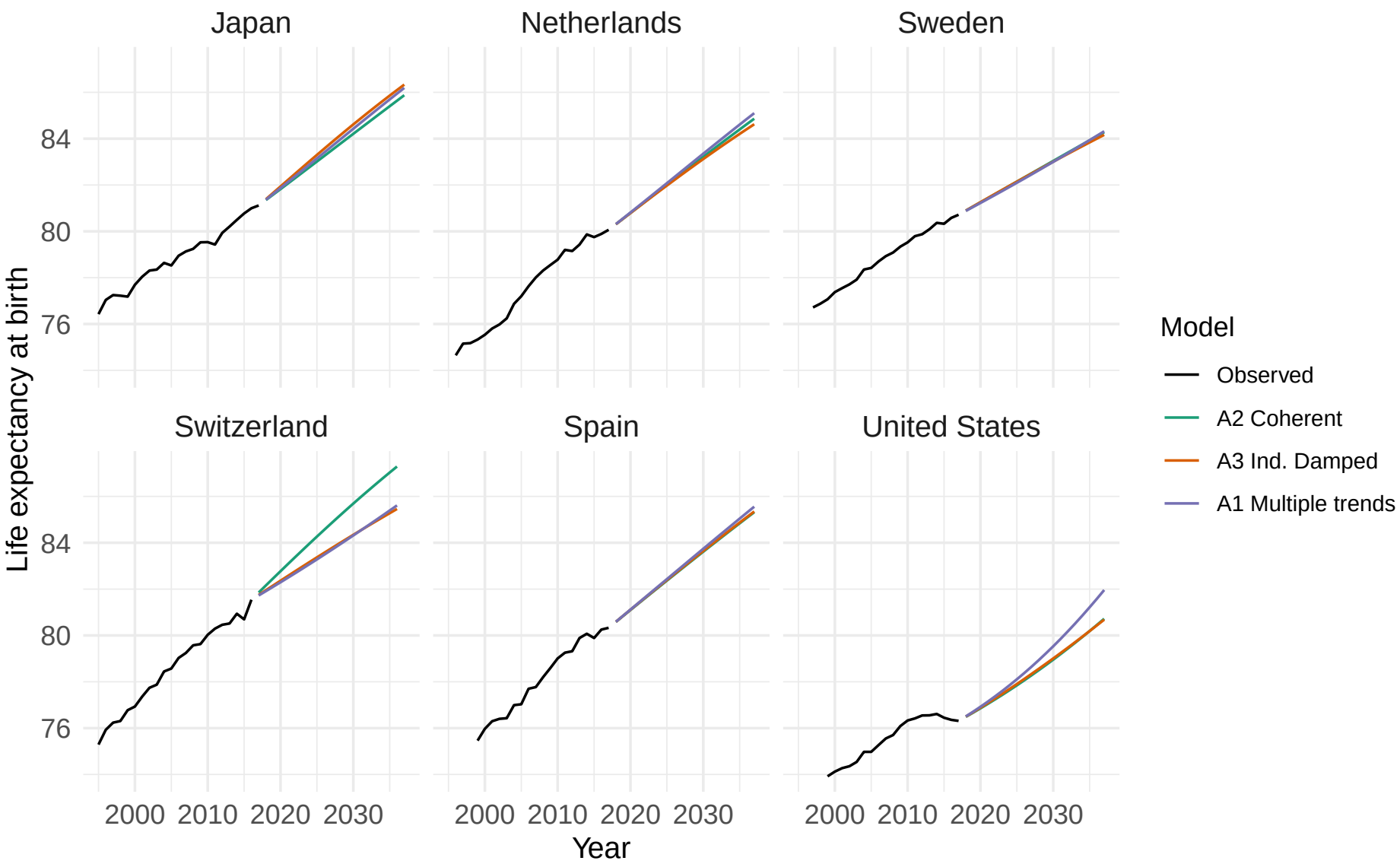
Forecast 20 years ahead, Females



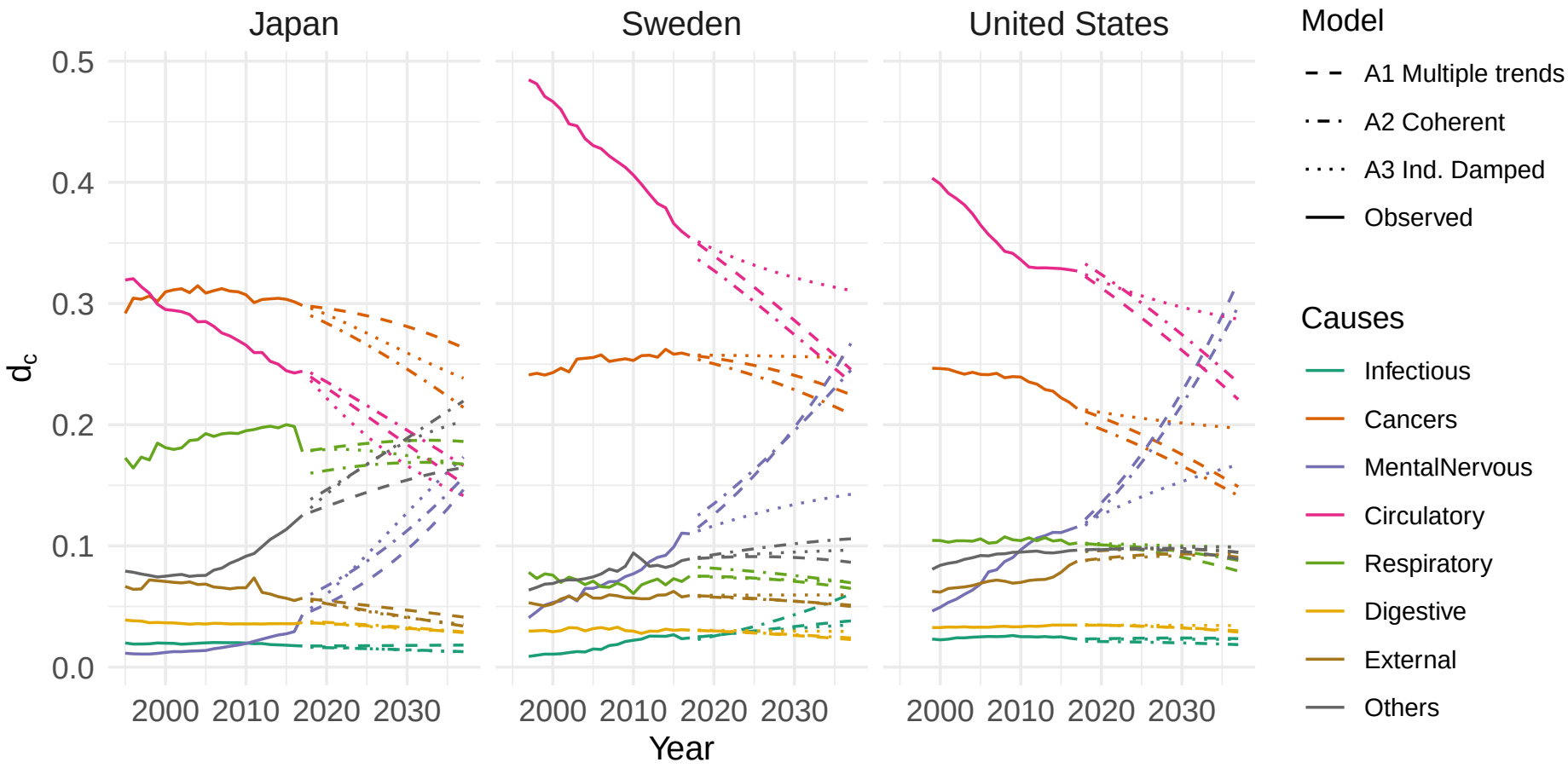
Forecast 20 years ahead, Females



Forecast 20 years ahead, Males



Forecast 20 years ahead, Males



Discussion

1. Inherent pessimism: No systematic pessimism;
2. Can lead to unrealistic trends: Approach A3 allows one to control unlimited transfer towards one cause;
3. Modifications to the International Classification of Diseases (ICD) create discontinuities over time;
4. Inconsistent with an all-cause forecast: Approach A2 controls for this;
5. Trajectories of causes of death are considered to be independent, yet in reality they are interconnected ...

How do they consider dependance between ages and causes?

1. Approach A1: dependence between ages and causes is considered.
2. Approach A2: dependence between age-groups and dependence between causes by age are considered. CDD between age-groups are considered independent, unless a coherent approach is used.
3. Approach A3: dependence between causes and dependence between ages by cause are considered. ADD between causes are considered independent, unless a coherent approach is used.

Discussion

- Results are sensitive to the populations and time-period.
- With approach A2, the use of different models to forecast the all-cause ADD will lead to different accuracy.
- Which approach to use?
- Can approach A2 be modified to control for unlimited transfer towards one cause?
- Use of weight.

Conclusion

There are many ways to forecast mortality by age and cause in a coherent manners. The model selection remains, however, a complicated task.

Merci!
Thank you!

mpbergeron@sdu.dk