University of Copenhagen, Dept. of Economics
5 Øster Farimagsgade, Building 26, Room 26.3.39,
DK-1353 Copenhagen K, Denmark
Phone +45 35 32 32 61 (Office)
      +45 43 90 84 72 (Home)
      +45 21 97 88 95 (Cell phone)
E-mail HansOluf.Hansen@econ.ku.dk
home page http://www.econ.ku.dk/personal/usihoh/
Ftp-server: ftp://ftp.ibt.ku.dk/usihoh/

Survivorship impacts of latent biological heterogeneity and period elements in human mortality

A simulation and projection study

By

Hans Oluf Hansen

(Work in progress)

Introduction

- Efforts of separating age, period and cohort effects in demography dates back well over a century (Hobcraft et al. 1982).

- In my view, an important reason of this apparent lack of success is tacit assumption of homogeneity on statistical analysis.

- For example, a birth cohort is a grouping of individuals characterized by a wide range of personal traits, most of them unobservable; but nonetheless real and socially important; and one of them being genetically defined frailty or health.

- How do we separate the endogenous effects of biological frailty and aging by attrition of cells from exogenous impact on survivorship?

- What are the selection effects of heterogeneity related to birth cohort and time-dependent exogenous effects?

- And what are the overall impacts of heterogeneity for selection in survivorship?

Purpose

This presentation emphasizes four issues:

1. A hazard model aiming at uncovering and separating hidden endogenous biological from latent exogenous effects in mortality across individual life courses.

2. Interpretation of heterogeneity and selection in cohort-based population mortality.

3. Analysis of the structure and change of exogenous mortality in historical and projected mortality as anticipated by Statistics Sweden and DREAM/Statistics Denmark.

Facts

- Human existence is contingent on interacting biological and environmental factors

- The vast majority of human live births are *monozygotic*, and thereby endowed with a *personal frailty* on conception.

- Man decays with *age* by *biological attrition*.

- From conception to death human existence is subject not only to *endogenous biological factors* but also to *exogenous environmental influence*.

- *Frailty on conception* constitutes individual difference in suitability to survive in presence of prevailing environmental conditions.

- By the principles of natural selection and survival of the fittest, depending on *biological attrition* and environmental conditions, the mean and variance of the distribution of frailties among survivors reduce over the life course as the birth cohort becomes less heterogeneous with age.

Problem

- Difference in personal frailty and selection distort empirical, non-homogeneous risk sets of cohort survivors. This may lead to biased inference about mortality change.

- By *selection*, I understand the natural or artificial process, which prunes a non-homogeneous birth generation of individuals less suited to endure prevailing environmental conditions.

- Rational reflection and evaluation of hypotheses regarding latent biological heterogeneity and exogenous influence upon human survivorship calls for statistical modeling.

- As biological frailty is *personal* the modeling must be at the level of individuals

Model

I propose a multivariate proportional hazard model of individual mortality as follows.

$$m_i\left(x, z_i, t\right) = z_i m_s\left(x, z = 1\right)\varphi\left(x, t\right)$$

With:

- i denoting personal ID;
- z_i indicating individual frailty;
- $m_s(x, z = 1)$ representing biological attrition in terms of a baseline hazard shared by men and women; *s* denoting *standard*.
- and $\varphi(x, t)$ signifying a shared conglomerate of environmental conditions with bearing upon survivorship at time t.

Computation in the model requires fixing or normalization of one or several effects, here distribution of frailty z_i and baseline $m_s(x, z = 1)$. The exogenous effect $\varphi(x, t)$ is established iteratively by stochastic simulation until achievement of statistical equivalence between empirical and model-based population mortality.

Estimation in the model

- Non-observability does not imply non-existence.
- I assume the frailty distribution and biological attrition by age to be known and non-differential by sex and fixed in time. This implies that the exogenous effect will differ by sex.
- Empirical data on life courses is available at population level.
- The non-parametric representation of biological attrition is known up to a factor.
- Latency and non-homogeneity of the risk set rules out data-based estimation e.g. using ML-approaches.
- As empirical data estimation is not an option, hypothesizing about the nature and importance of latent factors calls for stochastic micro simulation of individual life times using empirical population mortality as a target (benchmark) of the simulation.
- For lack of adequate information I truncate the simulated life courses by age 94.

- The latent baseline and the gamma parameters used in the application to follow have been uncovered by extensive experimenting based on elected samples reflecting extreme variation in empirical cohort mortality across the demographic transition in Iceland, Denmark, Sweden, and Japan.

- The results suggest that the baseline and the gamma parameters could be fixed in time (Hansen 2008, 2011).

- For further details, cf. Hansen (2013, circulated background paper).

Hypotheses

- Human control aspires, more or less successfully, to reduce mortality by decreasing $\varphi(x,t)$

- If $\varphi(x,t)$ is high then mortality is high. We would then expect selection to be high.

- Conversely, if $\varphi(x,t)$ is low, for example by successful human intervention, then I would expect selection to be low.

- At advanced ages people tend to die from ill-health associated with biological decay of cells.

I now proceed to examine the hypotheses using Swedish survivor experience as an example

Data

- Stochastic simulation in the proposed micro-model illustrates consequences in terms of heterogeneity and selection in historical as well as projected macro-scopic mortality.

- Hence, to obtain a long time series I extend historical data with informed official projections of population mortality.

- By the existence of cross-sectional national time series of annual occurrence-exposure data dating back to 1751, Sweden offers unique opportunities for scientific mortality research.

- The data is free and publicly available on annual basis thru Statistics Sweden and Human Mortality Database (HMD) from 1751 to date.

- Statistics Sweden, in my view, is leading in mortality analysis and medium and long-term prediction of national population mortality.

- The data bank of Statistics Sweden offers easy and non-bureaucratic access to mortality projections covering 2012 to 2110.

Stochastic micro simulation

- The available empirical and expected population mortality is defined on the state space {alive; dead}with random variates *frailty* and *waiting time to death or truncation* (the simple life model).

- Exogenous period influence upon mortality is uncovered iteratively by simulation. I shall leave some interesting associated technical challenges aside in this presentation.

- The number of simulated life times equals empirical or expected annual numbers of live births by sex.

- Iteration is brought a halt when statistical equivalence between population mortality based on the simulated life times and empirical mortality is achieved.

Uncovered latent lifetimes

- Stochastic micro-simulation in the model produces a set of individual life times truncated by age 94.

- The basic characteristics of a simulated life time are *birth year, sex, frailty*, and *exact age at death*.

- As *year of death* is equal to *year-of-birth* plus *age-at-death* the derived model-based population mortality may be studied not only under a *cohort-based*, but also under a *cross-sectional observational plan*.

Some inference on latent selection in human survivorship

1. Overall description of empirical and officially anticipated mortality.

   - As the model application is fully parameterized, aggregates based on the simulated micro-registers should comply statistically with empirical or officially predicted population mortality under both a cohort-based and a cross-sectional observational plan.

   - This is demonstrated graphically by figure 1.

Figure 1

Expected length of life at birth $e[0,94[$ truncated at the ninety-fourth birthday.
By birth year of cohort and year of observation of cross-sections of male cohorts.
Empirical mortality 1751-2010 and projected mortality 2011-2110 of males by
Statistics Sweden (fully drawn curves)
Modeled cohort-based and cross-sectional mortality according to the frailty and
selection model (cross markers)
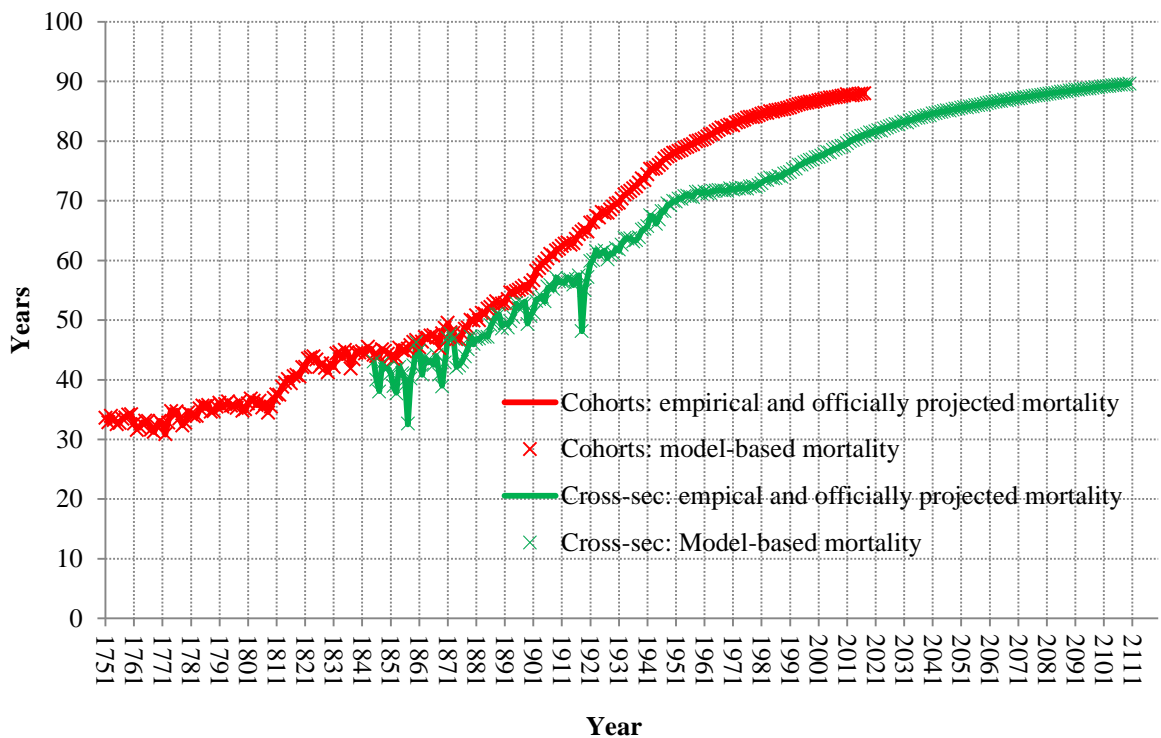
2. Heterogeneity of population mortality

- To study diversity in survivorship I group by fractile intervals based on the frailty generating gamma probability distribution shared by men and women. The elected fractile intervals are in per cent. To illustrate historical and expected change I focus on the cohorts born 1900, 1958, and 2017.

- Aggregated mortality has fallen dramatically across the generations born in 1900, 1958, and is expected to continue so regarding the generation to be born in 2017

- People with low frailties live longer than people with high frailties.

- The gains in life expectancy are much greater among persons with high frailties than among persons with low frailties.

- Females have lower mortality than males; however, males are catching up and are expected to have acquired statistically about the same expected length of life as the females to be born in 2017.

- Danish mortality development follows suit but is systematically somehow lower than Swedish mortality of all three birth cohorts (table 1).

Table 1. Expected length of life at birth before the ninety fourth birthday among elected Swedish and Danish birth cohorts before the ninety fourth birthday.

Cohort	Sex	Population	Frailty group								
			All	1	2	3	4	5	6	7	8
1900	F	DK	61.55	88.21	82.81	75.39	65.30	53.16	42.28	33.21	24.85
		SE	61.76	88.28	82.77	76.20	65.72	53.49	41.54	33.53	24.73
	M	DK	56.62	84.95	78.16	70.17	59.85	48.32	37.39	29.02	21.48
		SE	56.56	88.09	80.48	72.03	61.86	50.32	39.42	31.50	22.61
1958	F	DK	81.93	92.47	90.91	88.44	83.97	78.40	73.02	68.34	63.91
		SE	83.82	92.95	91.36	89.31	85.73	81.08	75.88	71.28	67.39
	M	DK	77.76	91.95	89.18	85.59	80.10	73.31	66.92	60.83	56.56
		SE	81.87	92.88	91.56	89.19	85.04	78.68	72.83	68.26	62.38
2017	F	DK	89.43	93.57	93.08	92.18	90.52	88.05	85.35	83.10	81.18
		SE	89.55	93.57	93.04	92.31	90.61	88.47	85.42	83.69	80.69
	M	DK	88.82	93.64	92.81	91.76	89.80	87.58	84.47	81.82	79.86
		SE	89.30	93.66	93.09	92.30	90.74	88.31	85.29	82.56	78.98

Source. Computions based on micro registers by stochastic simulation on assumption of fixed, non-differential biological parameters.

Table 2. Cohort size and number of censorings at age 94 by cohort, sex, country, and frailty group. Elected Danish and Swedish cohorts born in 1900, 1958, and 2017.

Coh-ort	Sex	Popul-ation	Frailty group								
			All	1	2	3	4	5	6	7	8
Birth cohorts											
1900	F	DK	35,087	2,082	2,169	6,070	9,693	8,698	4,302	1,174	899
		SE	67,299	3,989	4,106	11,700	18,654	16,437	8,326	2,404	1,683
	M	DK	37,042	2,179	2,262	6,535	10,193	9,074	4,544	1,318	937
		SE	70,840	3,481	3,567	10,444	17,872	17,851	10,586	3,505	3,534
1958	F	DK	36,225	2,131	2,211	6,371	10,005	8,941	4,376	1,256	934
		SE	50,988	2,963	3,054	9,087	14,100	12,513	6,237	1,729	1,305
	M	DK	38,456	2,239	2,284	6,875	10,534	9,565	4,683	1,283	993
		SE	54,514	2,730	2,647	8,135	13,553	13,807	8,187	2,728	2,727
2017	F	DK	30,559	1,821	1,805	5,555	8,346	7,480	3,749	1,022	781
		SE	63,305	3,757	3,789	10,982	17,421	15,640	7,861	2,191	1,664
	M	DK	29,103	1,747	1,704	5,160	8,125	7,023	3,538	1,030	776
		SE	60,290	2,989	3,016	9,153	14,958	15,371	8,900	2,976	2,927
Censorings by age 94											
1900	F	DK	1,329	908	283	130	8				
		SE	1,912	1,496	307	103	6				
	M	DK	268	266	2						
		SE	336	332	4						
1958	F	DK	8,099	1,764	1,445	2,597	1,833	424	36		
		SE	11,999	2,451	2,002	3,976	2,794	718	55	3	
	M	DK	4,775	1,616	1,012	1,540	563	44			
		SE	6,677	2,024	1,320	2,155	1,018	155	5		
2017	F	DK	14,671	1,700	1,546	4,017	4,394	2,332	577	79	26
		SE	30,770	3,505	3,211	8,029	9,445	5,101	1,238	181	60
	M	DK	11,086	1,601	1,336	3,231	3,234	1,434	229	16	5
		SE	21,937	2,715	2,415	5,953	6,506	3,472	746	91	39

Source. Counts based on micro registers by stochastic simulation on assumption of fixed, non-differential biological parameters.

3. Selection in population mortality

Selection in population mortality may be illustrated:

- By computing mean frailty and the associated standard deviation at inter-valued ages over the life course [0, 94[(figure 2)
- By plotting frailty against age at death (figure 3)

Figures 2-3 illustrate that selection has been postponed to still older ages in the cohorts born in 1900 and 1958, not least among groups #5-8 with frailties above mode in the gamma distribution.

Selection among frailty groupings #1-4 is expected to become postponed to very high ages or almost suspended in the generation to be born in 2017.

Figure 2. Estimated mean frailty and standard deviation among survivors at integer-valued ages Swedish men born 1900, 1958, and 2017
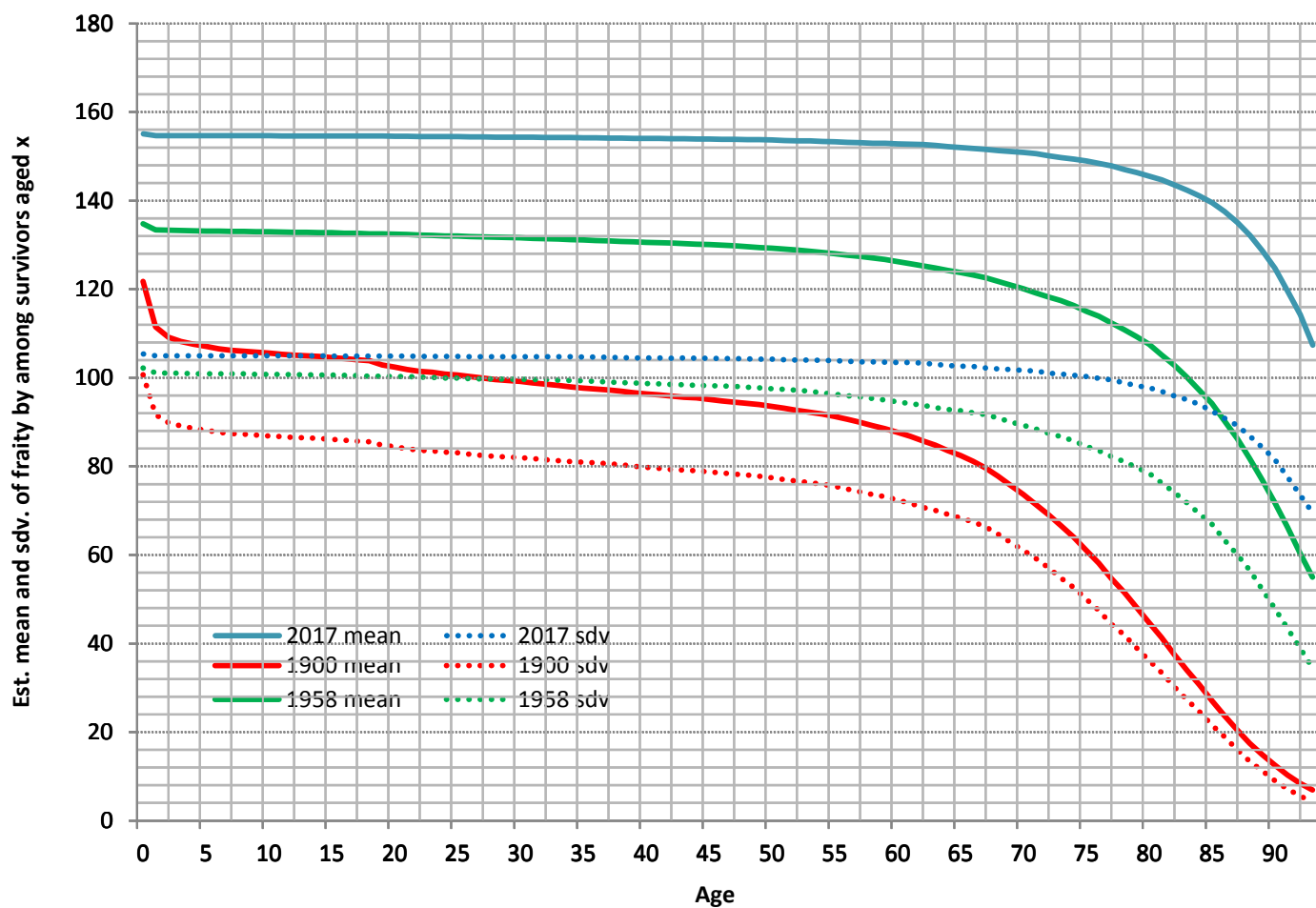
Figure 3. Personal frailty plotted against individual age at death. Swedish male cohorts born 1900, 1958, and 2017. With indication of fractile intervals (red horizontal lines)
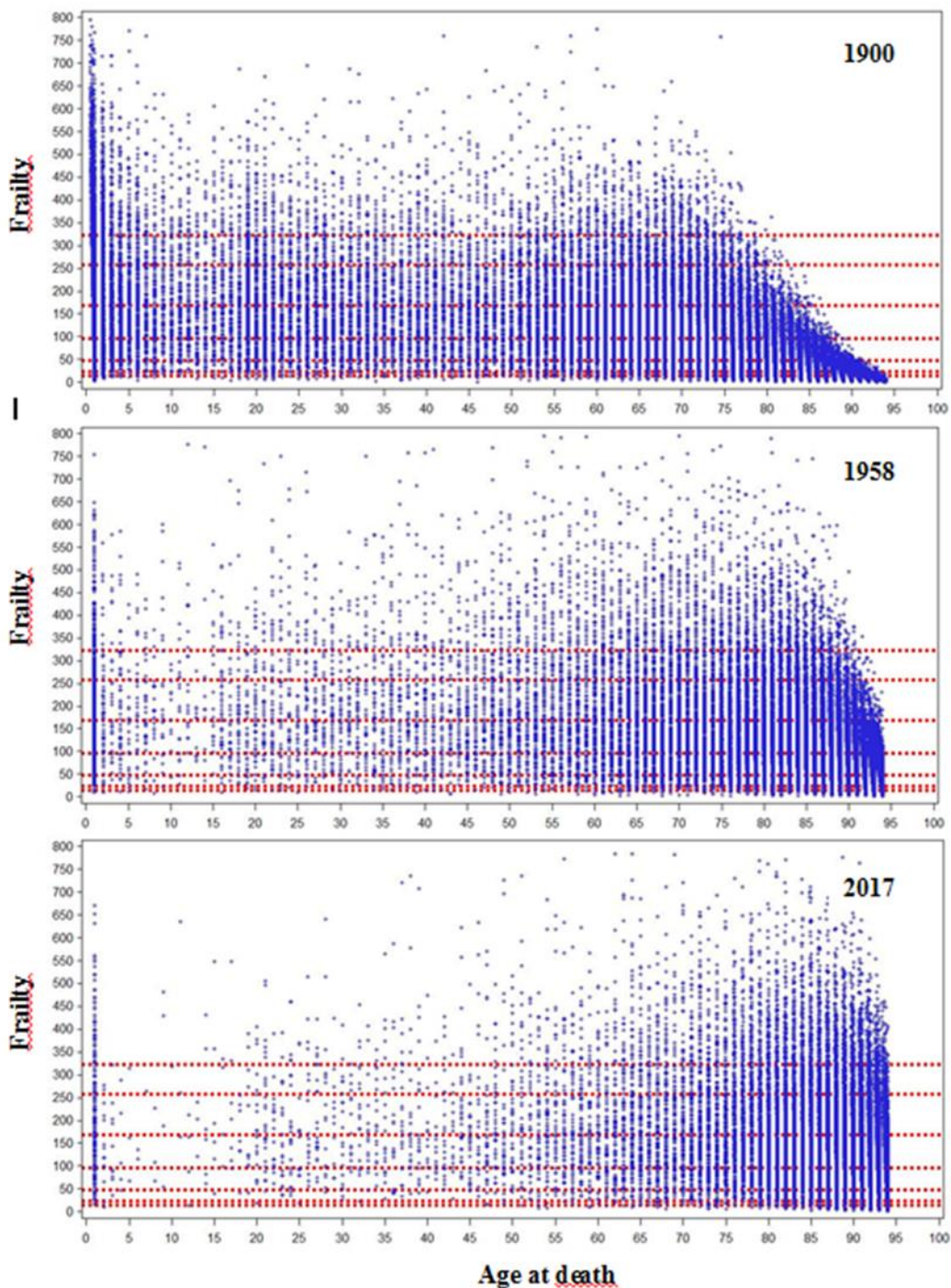
4. Statistical analysis of the exogenous effect and prediction of population mortality

These issues cannot be treated due to the time limits of this presentation; it will be dealt with elsewhere.

General conclusions

This project has shown that:

- Separating endogenous biological effects from exogenous environmental effects in the simple life model is perfectly possible under a coherent multiplicative hypothesis in the framework of a hazard model

- All effects are latent (non-observable). The model-based experiment suggests that non-homogeneity and selection plays a much greater role in human survivorship than hitherto acknowledged.

- This insight has great consequences for measurement and interpretation of current population dynamics and for the quality of population projections and related economic and social forecasting.

- It will be technically though not necessarily practically feasible, to generalize the approach adopted in this study to a multidimensional state space. Existence of appropriate cohort-based empirical benchmarks is crucial for meaningful calibrating of such modeling. For an example of stochastic micro-simulation in a multidimensional state space with duration dependency cf. Hansen (2000).

Thank you for your attention!