Forward and backward prevalences: changes in living arrangements observed with the US Health and retirement surveys

REVES 2016. Vienna

Nicolas Brouard Yao-Chi Shih

National Institute for Demographic Studies (INED) and University of Massachusetts, Boston/Taipei

REVES Vienna, 8 June 2016

IMaCh, for Interpolated Markov Chain, is a software which provides period Health Expectancy (HE) estimated from a cross-longitudinal survey. Information required is only the age and status (Healthy, Unhealthy or Death) at each interview. Because of different time exposures between successive interviews, the program maximises the likelihood of the sample based on a multinomial logistic model of a monthly transition probability (Lièvre et al., 2003).

Cross-subsectional participation "rates":

- 1949 Wolfbein S. L., "The length of working life", Population Studies, 3, p. 286-294.

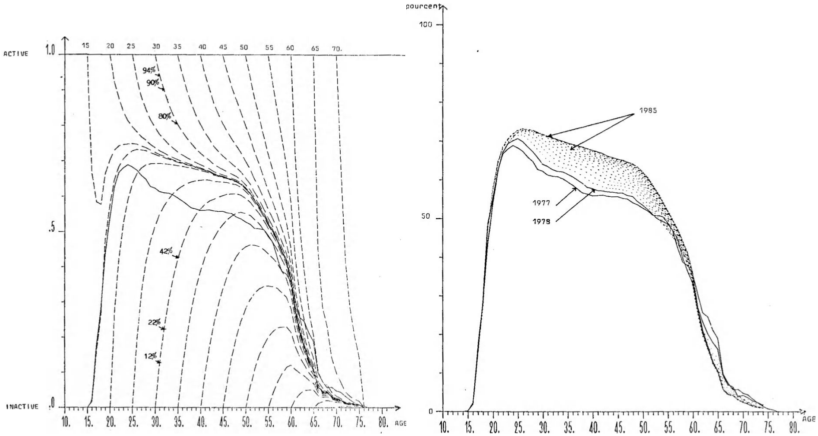
Labor force participation "rates" deduced from flows:

- 1976 Hoem J. M., Fong M., "A Markov chain model of working life tables ", Laboratory of actuarial mathematics, University of Copenhagen.
- 1980 Brouard, N. Espérance de vie active, reprises d'activité féminine : un modèle. Revue économique, 31(6):1260-1287 (Brouard, 1980)

Age $x+1$ 1978	ACTIVE	INACTIVE	
Age x 1977			
ACTIVES	N_{11}	N_{12}	$N_{1.}$
INACTIVES	N_{21}	N_{22}	$N_{2.}$
	$N_{.1}$	$N_{.2}$	$N_{..}$

We can compute the age specific probability of exiting the labor force in one year $\hat{c}_x = \frac{N_{12}}{N_{1.}}$ as well as [re]entering the labor force $\hat{a}_x = \frac{N_{21}}{N_{2.}}$
And by multiplying the age specific matrices P_x, we can compute the probability to be out of the labor force after n years for somebody in the labor force at age x as well as the probability to be in the labor force after n years for somebody out of the labor force at age x, etc:

$$P_x = \begin{pmatrix} 1-\hat{c}_x & \hat{c}_x \\ \hat{a}_x & 1-\hat{a}_x \end{pmatrix} \quad {}_nP_x = P_x P_{x+1} \cdots P_{x+n-1}$$



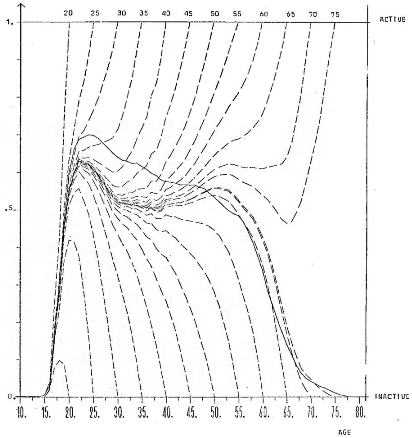
Cross-sectional vs period (stable) participation ratios. Forward projection of female participation rates (based on flows between two French labor force surveys 1977-78). Children do no more force women to leave the labor force.

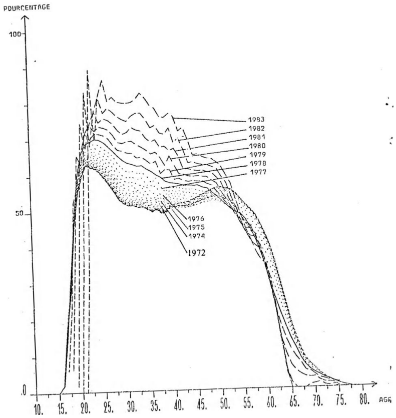
Age $x+1$ 1978	ACTIVE	INACTIVE	
Age x 1977			
ACTIVES	N_{11}	N_{12}	$N_{1.}$
INACTIVES	N_{21}	N_{22}	$N_{2.}$
	$N_{.1}$	$N_{.2}$	$N_{..}$

If we compute *backward* probability $b^{12} = \frac{N_{12}}{N_{.2}}$ defined as the probability to be inactive at age $x-1$ knowing that we will be active at age x and $b^{21} = \frac{N_{21}}{N_{.1}}$ as the probability to be active at age $x-1$ knowing that will be inactive at age x, we can make backward convergent projections.

$$B_x = \begin{pmatrix} 1-b_x^{21} & b_x^{12} \\ b_x^{21} & 1-b_x^{12} \end{pmatrix} \quad {}_nB'_x = B'_x B'_{x-1} \cdots B'_{x-n+1}$$

and because of weak ergodicity we get convergence in the past (as well as divergence in the future).





Backward projection of female participation rates (France 1977-78).
Back to the situation before 1968.

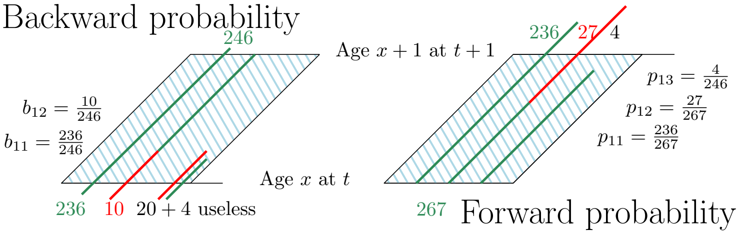


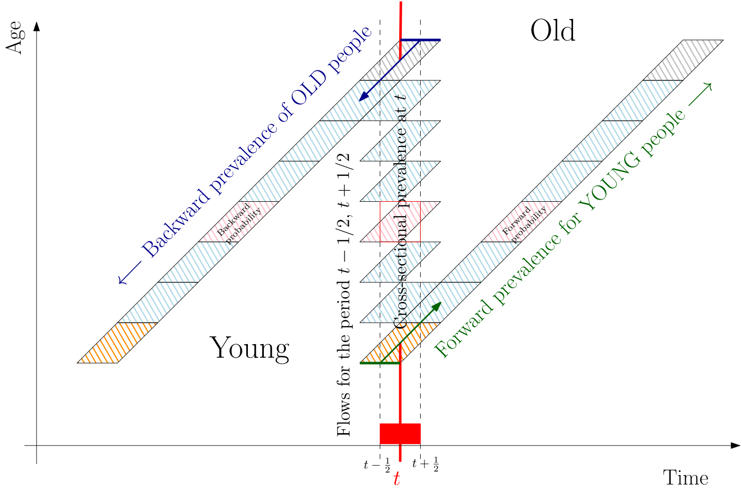
Figure: Lexis diagram of forward and backward approaches with DIFFERENTIAL MORTALITY. Green for active, red for inactive, black for death.

$$\begin{array}{lc} & \\ \text{Active} & 267 \\ \text{Inactive} & 280 \\ \text{Total} & 547 \end{array} \quad \begin{array}{ccc} \text{Act.} & \text{In.} & \text{Dead} \\ \end{array}$$

		Act.	In.	Dead
Active	267	236	27	4
Inactive	280	10	250	20
Total	547	246	277	24

We can see on Figure the classical forward approach of a Lexis diagram showing the future of a cohort of people active at age x with their corresponding statuses at age $x+1$: still active, inactive and dead. And the backward approach where we are looking at the probability to be active a year before, knowing the status this year.

The probability to be active today knowing that we will be active next year is different from the probability to be active next year knowing that we are active today.


Age
Old
← Backward prevalence of OLD people
Backward probability
Flows for the period $t-1/2$, $t+1/2$
Cross-sectional prevalence at t
Forward probability
Forward prevalence for YOUNG people →
Young
$t-\frac{1}{2}$
$t+\frac{1}{2}$
t
Time

Using a cross-longitudinal survey, it is now classical to estimate the age-specific cross-sectional prevalences of disability (from the first wave of the suvey) and to compare this curve with the "period" prevalences (computed from the age-specific flows observed between two waves). This period prevalence, or forward period prevalence, corresponds to the prevalence which will be observed in a younger cohort if the transitions rates between the various disability states as well as mortality, remained constant in the future. For example, a lower period prevalence of disability at old ages compared to the corresponding cross-sectional prevalence indicates that disability is currently declining.

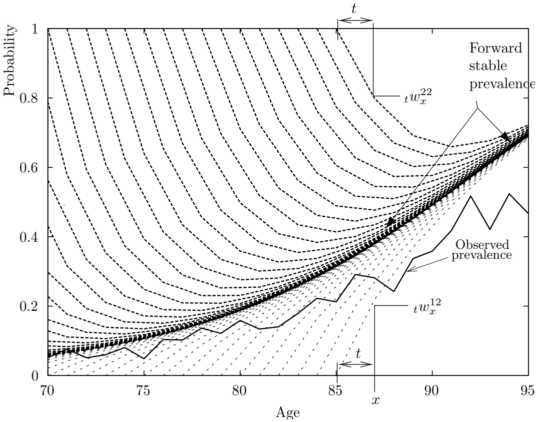


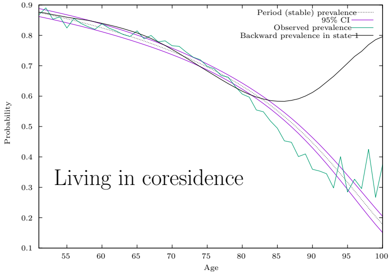
Figure: Forward prevalence of disability at age x according to the original health status (1 for disability-free and 2 for disabled) of the cohort at age $x-t$. The stable forward prevalence of disability is calculated from the changes in individual health status from 1984 to 1986 (LSOA). It is compared with cross-sectional prevalence observed at first interview in 1984 (LSOA). Institutionalized people were not surveyed at first round.

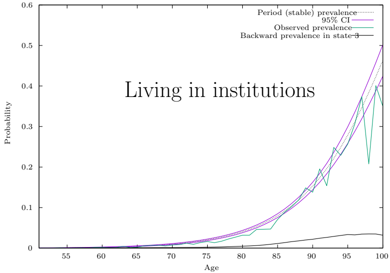
Here, we want to emphasize this result by introducing the age-specific backward prevalence which is the prevalence which would have been observed in the past at younger ages, in an cohort reaching an old age today. We can prove a convergence to a unique curve independent of the health status of the older cohort and therefore we can call it the “backward prevalence”. In comparison to the forward prevalence which depends only of the transition rates, the backward prevalence depends also of the cross-sectional prevalence. With the ageing process and changing conditions, we can expect that the cross-sectional prevalence is *in between* the backward and forward prevalences.

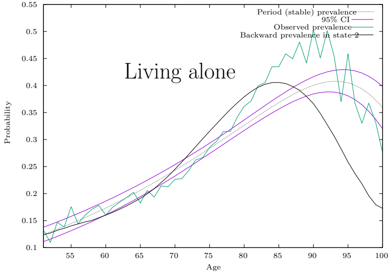
We explore the properties of this proposed index using the latest US Health and Retirement Survey with eight waves (1998-2012). Also, we are interested not only in Health Expectancy but in how the living arrangements are changing in relation with the disability statuses while a person is aging(Shih, 2016).
As we implemented the calculation of the backward prevalences in the beta version 0.99 of our IMaCh software, we have the opportunity to use three different living arrangement statuses, "coresidence", "alone" and "in institution" and potentially to use the disability statuses, also measured at each wave of the survey, as a determinant variable.

- Biennial Living Arrangements 1998 to 2012, US HRS ;
- Three live states: 1 co-residence, 2 alone, 3 in institutions.
- The 4th state is death.
- 8 waves;

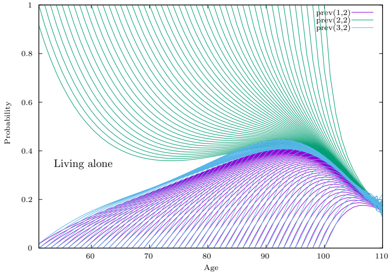
First results showing a decline in coresidence and a higher mortality in institution are presented.

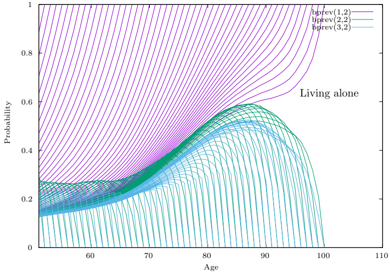


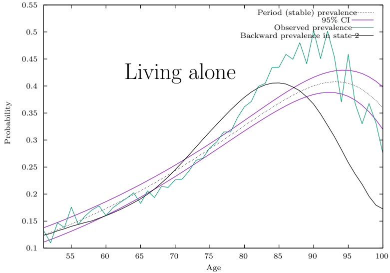




- 3 graphs: Coresidence -> Alone -> Institutions ?
- 3 curves per graph: Cross-sectional prevalence in between forward (with 95% CI) and backward prevalences.

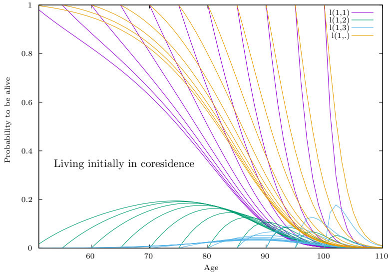


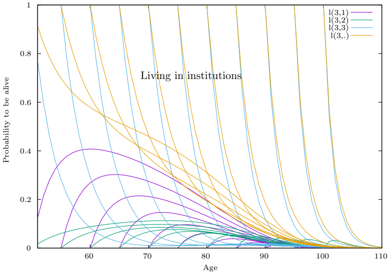


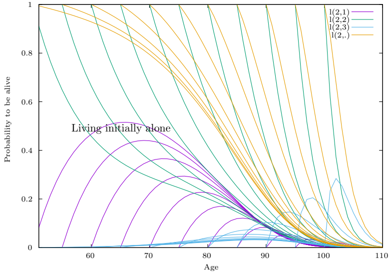


Focus on Living alone shift:

- Convergence to forward and backward prevalence of living alone (HRS).
- Shift to older ages (backward, cross-sectional, forward).







- Higher mortality for people initially in institutions compared to being alone or in coresidence.
- Controlling for disability and chronic diseases? Constant versus variable covariates in IMaCh (code is written in forthcoming version but convergence problem encountered and other priorities).
- Differential by gender are huge (see Yao Chi PHD thesis). Major patterns of women's living arrangement sequences are more diverse than those for men.

Parameter file of version 0.99:

```
#IMaCh 0.99r6
#Number of iterations & function calls = 22 & 8063, -2 Log likelihood = 114377.710692913606
#title=1st_example datafile=HRS/RAND_HRS/randostata/randhrsoV2.txt lastobs=19819 firstpass=1
      lastpass=8
title=1st_example datafile=randhrsoV3cwgt.txt lastobs=19723 firstpass=1 lastpass=8
ftol=1.000000e-09 stepm=12 ncovcol=1 nqv=1 ntv=2 nqtv=1 nlstate=3 ndeath=1 maxwav=8 mle=1 weight
      =1
model=1+age+.
# Parameters nlstate*nlstate*ncov a12*1 + b12 * age + ...
12    -5.7019604     0.0312616
...
# agemin agemax for life expectancy, bage fage (if mle==0 ie no data nor Max likelihood).
agemin=51 agemax=110 bage=50 fage=100 estepm=12 ftolpl=6.000000e-04
begin-prev-date=1/1/1998 end-prev-date=31/12/2012 mov_average=0
pop_based=1
prevforecast=1 starting-proj-date=1/1/1998 final-proj-date=1/1/2020 mobil_average=0
backcast=1 starting-back-date=1/1/1998 final-back-date=1/1/1980 mobil_average=1
```

Datafile of version 0.99:

```
# datafile=HRS/RAND_HRS/randostata/randhrsoV2.txt
# RAND HRS O 2015
# Biennial Living Arrangements 1998 to 2012, US HRS
# hhidpn sex iadl adlw r4wtresp   bmoyr  dmoyr  imoyr4 livarnb4 imoyr5 livarnb5 imoyr6 livarnb6
      imoyr7 livarnb7 imoyr8 livarnb8 imoyr9 livarnb9 imoyr10 livarnb10 imoyr11 livarnb11
000002010 1   0  0   3121   10/1934 11/2001 6/1998   1   3/2000   1   99/9999   4
     99/9999   4   99/9999   4   99/9999   4   99/9999   4   99/9999   4
000003010 0   0  0   3686   1/1936  99/9999 4/1998   1   3/2000   1   4/2002   1
     7/2004   1   3/2006   1   2/2008   1   9/2010   1   7/2012   1
```

References I

Brouard, N. (1980). Espérance de vie active, reprises d'activité féminine : un modèle. *Revue économique*, 31(6):1260–1287. (numéro sur la démographie économique).

Lièvre, A., Brouard, N., and Heathcote, C. (2003). The Estimation of Health Expectancies from Cross-longitudinal studies. *Mathematical Population Studies*, 10(4):221–248.

Shih, Y.-C. (2016). *Living alone and subsequent living arrangement transitions among older Americans*. PhD thesis, University of Massachusetts Boston, Boston, MA.

Thank you for your attention!

Appendix

$$\begin{bmatrix} N_{1.} & 0 & 0 \\ 0 & N_{2.} & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} N_{11} & N_{12} & N_{13} \\ N_{21} & N_{22} & N_{23} \\ 0 & 0 & 0 \end{bmatrix}$$

$$\begin{array}{cc} N_{1.} & \begin{bmatrix} N_{11} & N_{12} & N_{13} \\ N_{21} & N_{22} & N_{23} \end{bmatrix} \\ N_{2.} & \\ N_{..} & \begin{matrix} N_{.1} & N_{.2} & N_{.3} \end{matrix} \end{array}$$

$$\begin{aligned} \begin{bmatrix} w_1 & 0 & 0 \\ 0 & w_2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ 0 & 0 & 1 \end{bmatrix} &= \begin{bmatrix} w_1 p_{11} & w_1 p_{12} & w_1 p_{13} \\ w_2 p_{21} & w_2 p_{22} & w_2 p_{23} \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \frac{N_{1.}}{N_{..}} \frac{N_{11}}{N1.} & \frac{N_{1.}}{N_{..}} \frac{N_{12}}{N1.} & w_1 p_{13} \\ w_2 p_{21} & w_2 p_{22} & w_2 p_{23} \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} N_{11} & N_{12} & N_{13} \\ N_{21} & N_{22} & N_{23} \\ 0 & 0 & 0 \end{bmatrix} / N_{..} \end{aligned}$$

$$b_{ij} = \frac{N_{ij}}{N.j} = \frac{w_i p_{ij}}{\sum_i w_i p_{ij}} = \frac{N_{i.}}{N_{..}} \frac{N_{ij}}{N_{i.}} / \sum_i$$

and then we multiply the matrices to get the backward prevalence limit!!!!

$$B_{x+1} = \begin{bmatrix} \frac{w_1 p_{11}}{w_1 p_{11} + w_2 p_{21}} & \frac{w_1 p_{12}}{w_1 p_{12} + w_2 p_{22}} \\ \frac{w_2 p_{21}}{w_1 p_{11} + w_2 p_{21}} & \frac{w_2 p_{22}}{w_1 p_{12} + w_2 p_{22}} \end{bmatrix}$$

$$= \begin{bmatrix} w_1 & 0 & 0 \\ 0 & w_2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{w_1 p_{11} + w_2 p_{21}} & 0 & 0 \\ 0 & \frac{1}{w_1 p_{12} + w_2 p_{22}} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$B_{x+1} = \mathrm{Diag}(w_x) P_x \, \mathrm{Diag}(\sum_i w_x^i p_x^{ij})$$

$$ {}_nB_x = B_{x-(n-1)} B_{x-(n-2)} \cdots B_{x-1} B_x $$

to be compared with

$$ {}_nP_x = P_x P_{x+1} \cdots P_{x+n-1} $$

In order to higlight the period prevalence at exact age x as the limit when $n \to \infty$, it can be rewritten

$$ \begin{aligned} {}_nP_{x-n} &= P_{x-n} P_{x-(n-1)} \cdots P_{x-1} \\ &= P_{x-n} \cdot {}_{n-1}P_{x-(n-1)} \end{aligned} $$