



**Karolinska  
Institutet**

# **sttde: a time-dependent and post- estimation within time-interval command**

Hugo Sjöqvist  
[hugo.sjoqvist@ki.se](mailto:hugo.sjoqvist@ki.se)

Nicola Orsini  
[nicola.orsini@ki.se](mailto:nicola.orsini@ki.se)

# Exposure data of income (Year\_1990-Year\_2015) collected over 25 years, with record of the death of the individual

Start	End	Death	Year_1990	Year_1991	.....	Year_2014	Year_2015
1991	2013	0	3423	5431		3456	6641
2005	2014	1	5321	2341		4321	.
1998	2015	0	.	.		5464	5431
1990	1995	0	342	453		224	156
1995	2015	1	.	7585		4257	1335
2001	2014	1	21235	4523		4568	.
2008	2009	1	23854	124		.	.

## Other possible research questions

- Applied example 1.
  - What is the association between suicide and being a parent, and how does it change during the adult's lifespan?
  
- Applied example 2.
  - How long does it take for immigrants from specific regions to utilize the Swedish health care system to the same length, or more as a Swedish native?

# The assumption of proportionality and time-dependence, in survival analysis

- A common setting in survival analysis is that exposure is recorded at baseline and then it is assumed to stay constant.
- However, exposures and covariates may change over time. Ignoring this fact may lead to misleading and uninformative results (eg. Yearly income at migration)
- Time-dependent survival analysis, in short, is to divide the individual's follow-up at the time-point when the exposure changes.

## What is sttde?

- The program has several functions
  1. Re-structure and manage the data to fit a time-dependent analysis.
  2. Use previous information of the user's survival analysis input to estimate a general time-dependent regression model.
  3. If chosen by the user, estimate within time-intervals of the study period to be able to examine the change of the estimates over time.
  4. Receive graphically the exposure's estimates over time for a clearer understanding (with Stata's own graphical `twoway` layout, making the user able to design them as he/she wishes).
- The program also functions by ignoring step 1, granting the user an informative picture of the behavior of a non-time-dependent exposure (recorded at baseline) over time.

# Possible structure of the data

start	stop	birth_date	died	imi	male	income_1991	income_1992	income_1993	income_1994
01jan2011	31dec2015	15aug1975	0	1	0	.	.	.	.
01jan1991	31dec2015	15jun1946	0	0	0	0	4	4	4
01jan1991	31dec2015	15feb1958	0	0	0	2	2	2	2
01jan1991	22nov2005	15dec1917	1	0	0	1	1	1	1
14jun2005	31dec2015	15jun1987	0	0	1	.	.	.	.
01jan1991	06apr1999	15nov1914	1	1	1	2	2	1	1
01jan1991	31dec2015	15apr1935	0	0	1	4	4	4	4
01jan1991	17aug2000	15mar1916	1	0	1	0	0	0	0
15jul2007	31dec2015	15jul1989	0	0	0	.	.	.	.
01jan1991	14jul2003	15jul1913	1	0	1	0	0	1	1
01jan1991	31dec2015	15mar1971	0	0	0	1	2	3	3
01jan2012	31dec2015	15oct1960	0	1	1	.	.	.	.
01jan1991	15feb2010	15jun1914	1	0	0	1	1	1	1
01jan1991	31dec2015	15feb1969	0	0	0	3	2	3	3
15feb2012	31dec2015	15feb1994	0	0	0	.	.	.	.
01jan2010	31dec2015	15oct1970	0	1	1	.	.	.	.
14may1996	31dec2015	15may1978	0	0	1	.	.	.	.
01jan1991	21nov1993	15mar1913	1	0	1	3	4	4	4

## Code to execute program

Minimum required code

```
. sttde income_1991-income_2015
```

```
. sttde income_1991-income_2015, streg(cox) type(ib2.) eform adjustfor(i.imi i.male) tb(16(2)40)
```

**streg:** the regression model to be used – default is cox regression

**type:** continuous, categorical (with or without chosen ref.) – default is continuous

**eform:** the Efron method to handle tied failures – Breslow is default

**adjustfor:** covariates or non-time depending exposures to be included in the model

**tb:** specification of the within-time estimation – excluding will yield only the overall estimate

# Code to execute program

```
. stset stop, scale(365.25) id(lopnr) origin(birth_date) failure(died) enter(start)
```

```

            id:  lopnr
      failure event:  died != 0 & died < .
obs. time interval:  (stop[_n-1], stop]
enter on or after:  time start
exit on or before:  failure
  t for analysis:  (time-origin)/365.25
      origin:  time birth_date
```

---

```

105,212  total observations
         3  ignored because never entered
   3,106  observations end on or before enter()
```

---

```

102,103  observations remaining, representing
102,103  subjects
   22,948  failures in single-failure-per-subject data
1,733,754  total analysis time at risk and under observation
                                     at risk from t =          0
                                     earliest observed entry t = 17.99863
                                     last observed exit t = 109.8097
```

```
. sttde income_1991-income_2015, type(ib2.) eform adjustfor(i.imi i.male) tb(16(2)40)
```



# Hazard ratios within intervals of the time-scale (age)

```
. sttde income_1991-income_2015, streg(cox) type(ib2.) eform adjustfor(i.imi i.male) tb(16(2)40)
```

```
-----  
Time interval estimation (95% conf. int.)  
-----
```

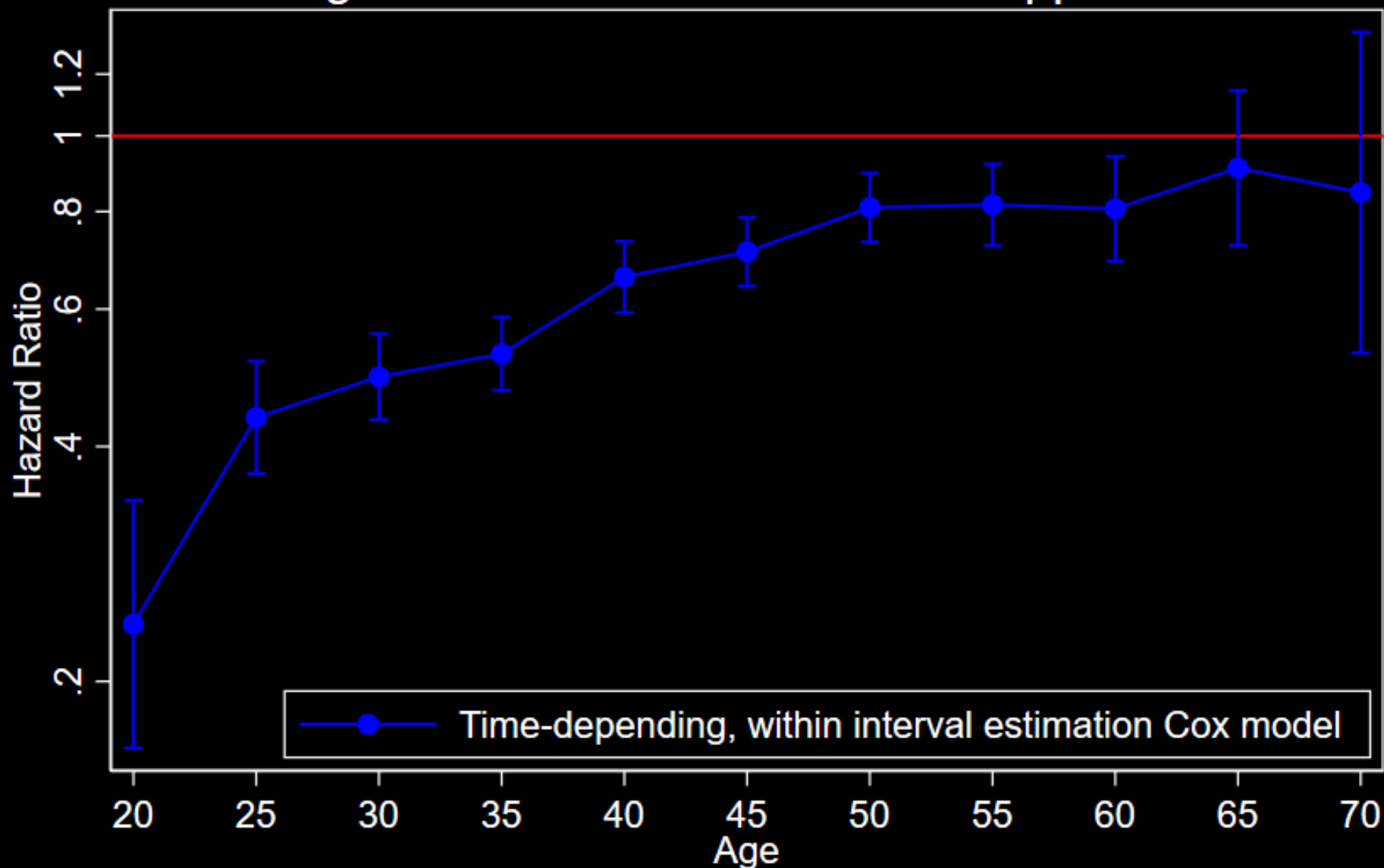
Variable	18-20	20-22	38-40	40-110
income				
0	5.617 (1.665; 18.953)	1.794 (0.458; 7.026)	1.507 (0.704; 3.226)	0.734 (0.705; 0.763)
1	2.684 (0.712; 10.120)	2.834 (0.752; 10.688)	0.979 (0.424; 2.258)	0.949 (0.916; 0.984)
2	1	1	1	1
3	1.077 (0.217; 5.338)	0.994 (0.201; 4.925)	0.474 (0.165; 1.365)	0.946 (0.902; 0.992)
4	0.376 (0.039; 3.620)	0.712 (0.119; 4.268)	0.693 (0.278; 1.726)	0.809 (0.768; 0.852)
imi				
1	0.431 (0.164; 1.133)	1.111 (0.389; 3.174)	0.724 (0.320; 1.635)	1.239 (1.181; 1.299)
male				
1	2.293 (1.151; 4.566)	2.665 (1.103; 6.442)	2.296 (1.287; 4.094)	1.590 (1.548; 1.633)

# Applied example 1: The association between children and parents' suicide\*

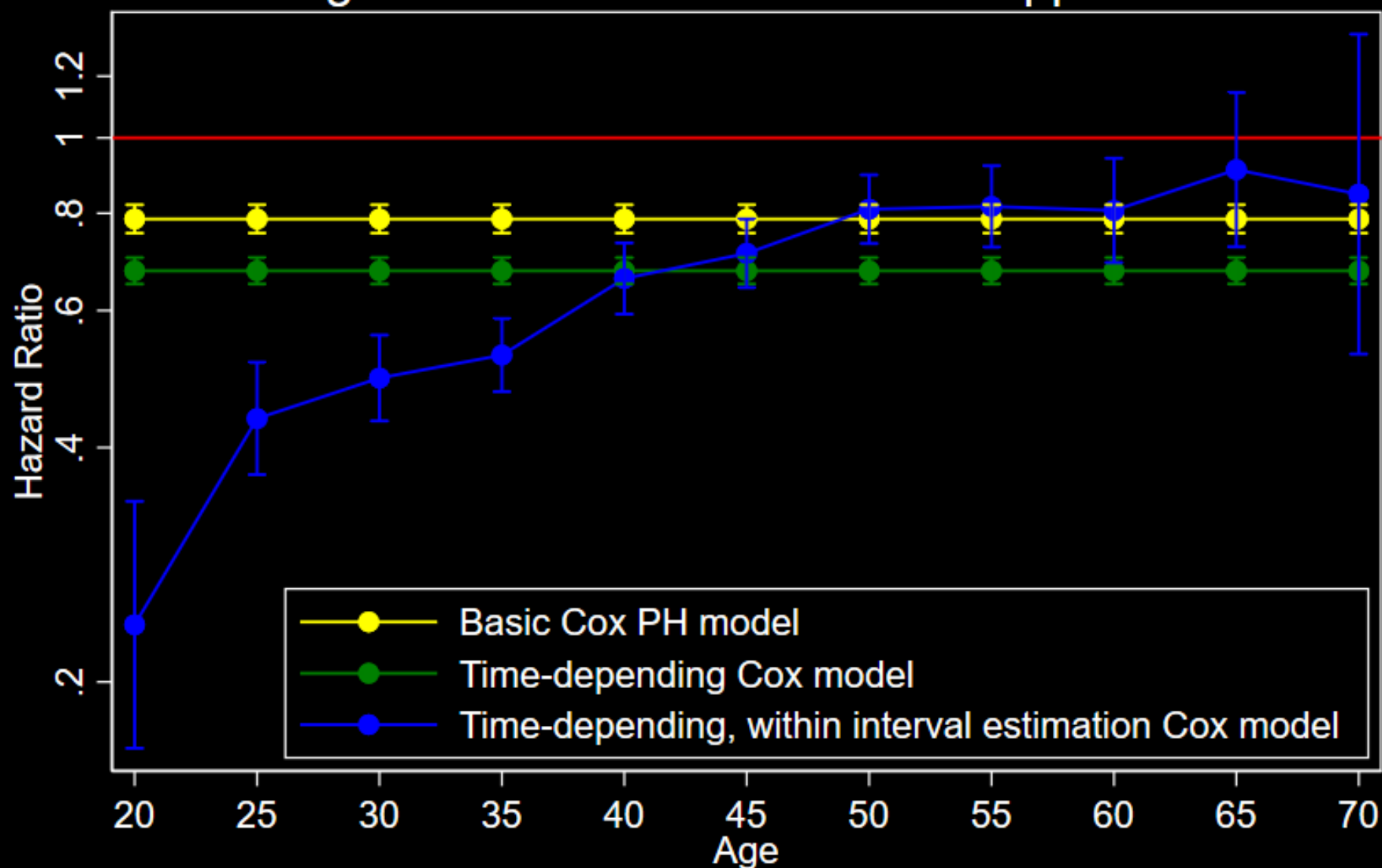
- A Swedish population study measuring the association the number of children have with the risk of parents taking their own life at a later date.
- The men and women entered the study at earliest the year of 1990 or at the age of 20, whichever occurred last. They were followed until 2011, or at loss of follow-up.
- Due to the young age of most individuals, the number of children recorded at the baseline was few – *leading us to treating the exposure as time-dependent.*

\*"Parenthood and lower risk of suicide in women and men" (in manuscript). Main author: Alma Söderberg Wallin, Karolinska Institutet ([alma.sorberg.wallin@ki.se](mailto:alma.sorberg.wallin@ki.se))

# Application 1: Risk of suicide for being a parent, using different survival estimation approaches



# Application 1: Risk of suicide for being a parent, using different survival estimation approaches

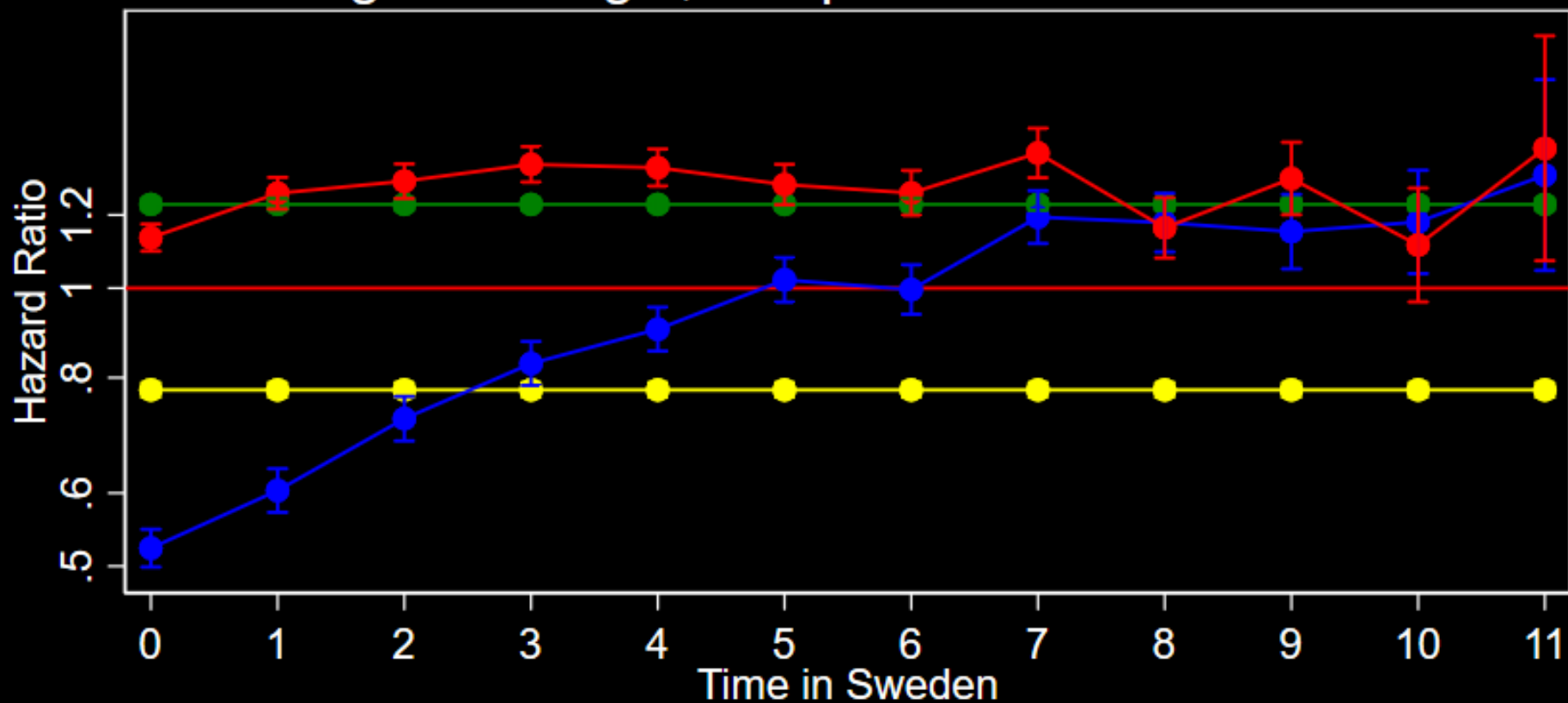


## Applied example 2: Estimation of time to psychiatric care use among migrants, compared to Swedish born citizens\*

- Sweden has, to a certain extent, free healthcare and people in need of psychiatric care is always greatly encourage to contact and utilize it.
- It is not surprising to assume that migrants being granted citizenship in Sweden does not fully understand or grasp the healthcare system, to the same extent as people born in the country, at the start.
- This study examined the association between migration status and utilizing psychiatric healthcare in Sweden, compared to Swedish born. *The exposure here is **not** treated as time-dependent.*

\*"Psychiatric care use among migrants to Sweden compared with natives – a longitudinal cohort study of 5,682,485 people in Sweden" (in manuscript). Main author: Anna-Clara Hollander, Karolinska Institutet ([anna-clara.hollander@ki.se](mailto:anna-clara.hollander@ki.se))

## Application 2: Psychiatric care use in regards to region of origin, compared to Swedish born



- PH Cox: Middle-East and N. Africa
- Within interval: Middle-East and N. Africa
- PH Cox: Eastern Europe
- Within interval: Eastern Europe, RUS & Baltics

## Future work

- Keep designing it so it can handle as many “stset” approaches as possible.
- Calibrate the program so the computational speed decreases.
- Adding more options to the `sttde` command.