Bayesian Forecasting of Cohort Fertility

Carl Schmertmann	Florida State Univ
Emilio Zagheni	MPIDR
Joshua Goldstein	MPIDR
Mikko Myrskylä	MPIDR

"It's difficult to make predictions… especially about the future."

-- Yogi Berra??

-- Niels Bohr??

-- Winston Churchill??

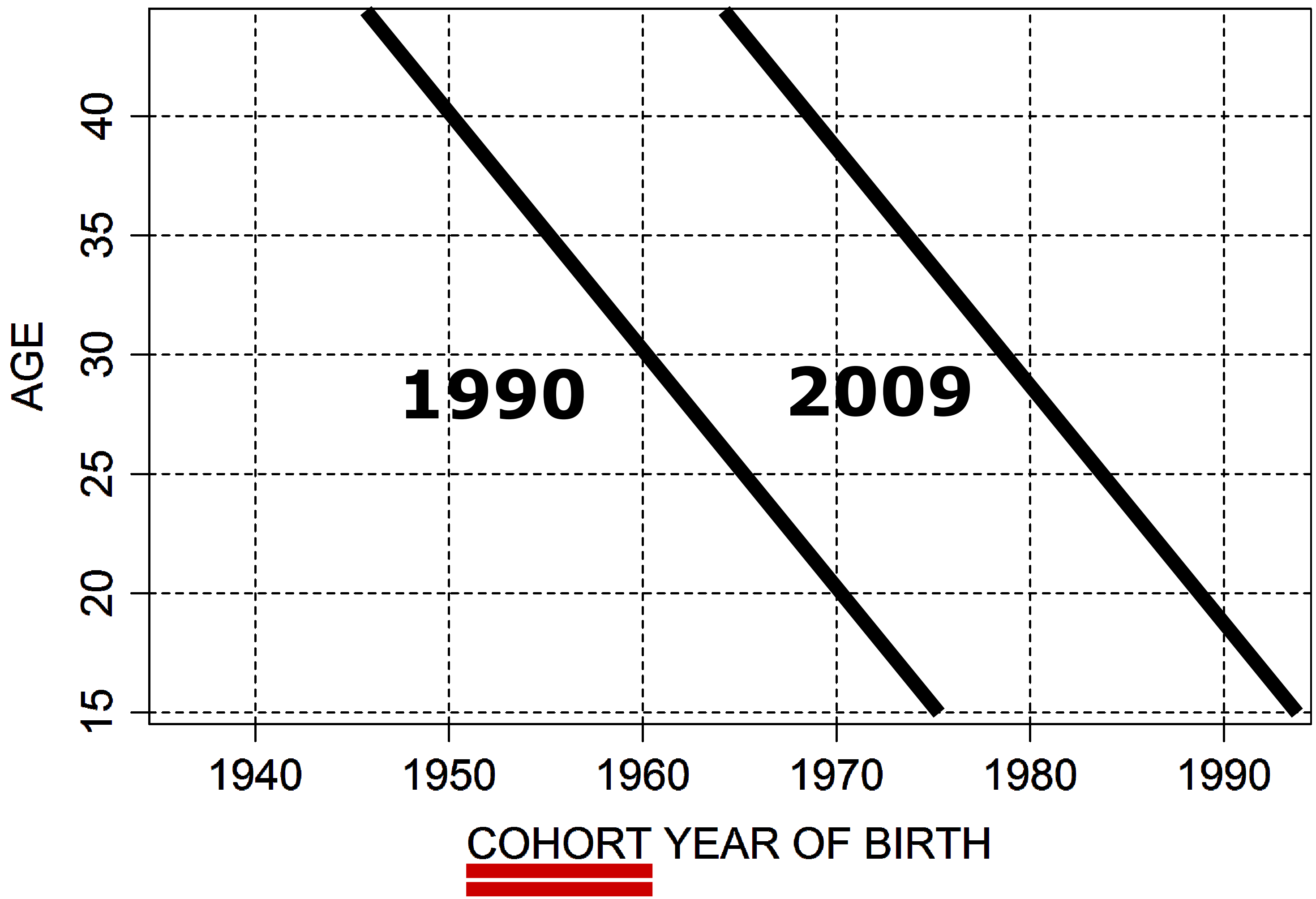
Motivation: CFR Rebound?

- PTFR up recently in many countries…
- … but CFR may be our real interest
- Change in period measures from...
  - Increasing cohort levels? (CFR up)
  - Decelerating postponement? (CFR ???)
  - Both?

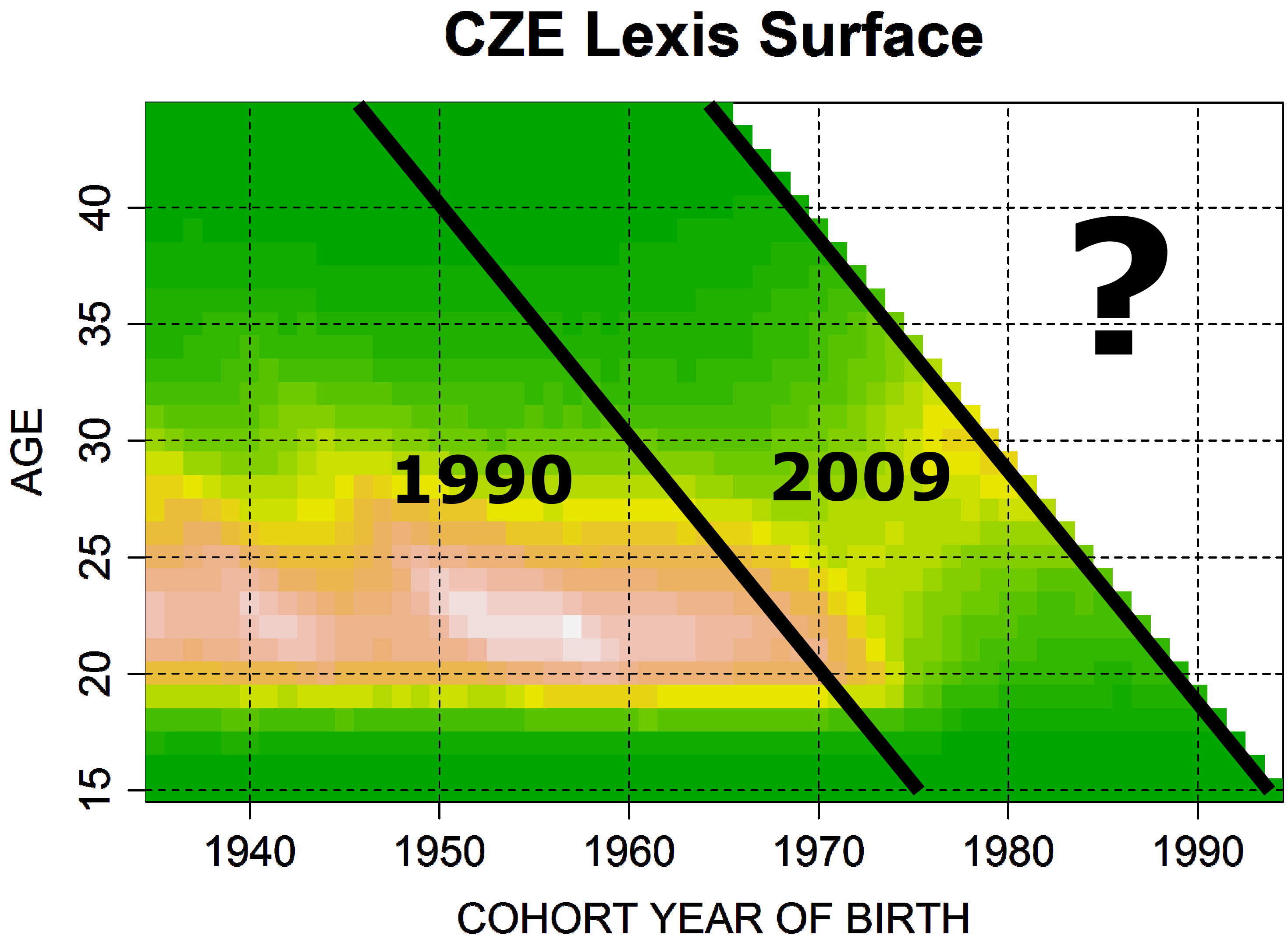
Objectives

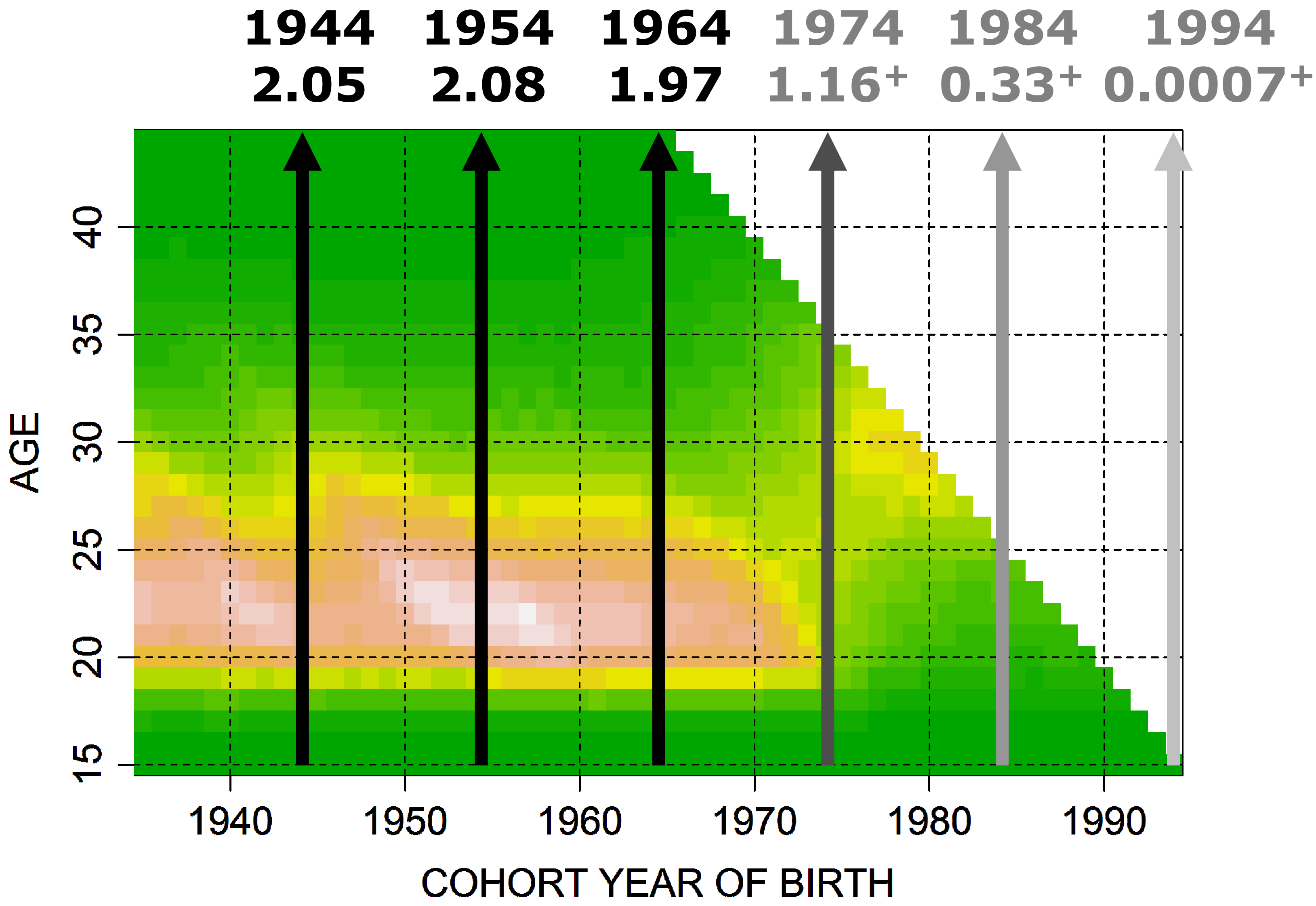
- Forecast completed CFR for cohorts already 15 but not yet 45
- Build a procedure that automatically includes uncertainty estimates
- Use historical data (**HFD**) to
  - design model
  - calibrate uncertainty

Cohort Lexis Surface



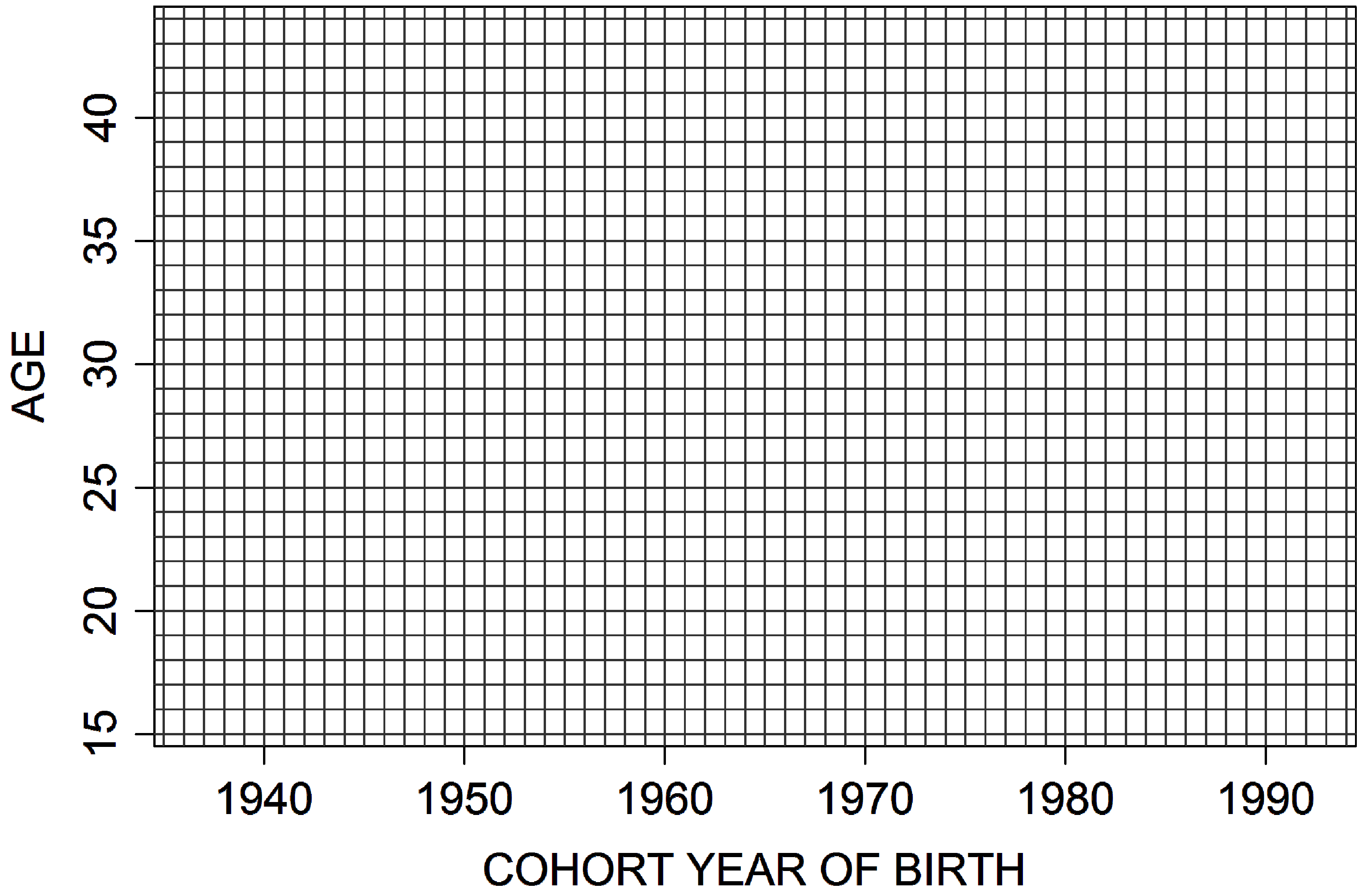
SOME HFD DATA...





1944
2.05
1954
2.08
1964
1.97
1974
1.16+
1984
0.33+
1994
0.0007+
AGE
15
20
25
30
35
40
1940
1950
1960
1970
1980
1990
COHORT YEAR OF BIRTH

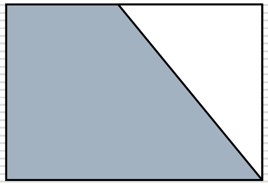
Our Strategy, inspired by Girosi & King (2008):
Model the entire surface non-parametrically
(i.e., one parameter per age-cohort cell)



Bayesian Model

- $\theta = \{ \theta_{ac} \}$ = **Surface**
  (e.g., 30x60 = 1800 parameters)

- {Estimates of some θ_{ac} } = **Data**
  (e.g., 1365 f_{ac} estimates from HFD)



- **L(Data | Surface)** … std likelihood

- f (**Surface**) … prior distribution

Do demographers know anything about fertility?

In the absence of data, are some surfaces more likely than others?

Bayesian Model

$$P(\,\theta \mid \mathbf{Data}) \;\; \underline{\alpha} \;\; L(\,\mathbf{Data} \mid \theta) \cdot f(\theta)$$

Posterior
How likely are alternative surfaces θ, given our observations?

Likelihood
How likely are our observations for alternative surfaces θ?

Prior
How likely are alternative surfaces θ, before we see data?

… informed by **HFD**

$$P(\,\theta \mid \textbf{Data, HFD}) \;\underline{\alpha}\; L(\,\textbf{Data} \mid \theta) \cdot \underline{f\,(\theta \mid \textbf{HFD})}$$

Prior

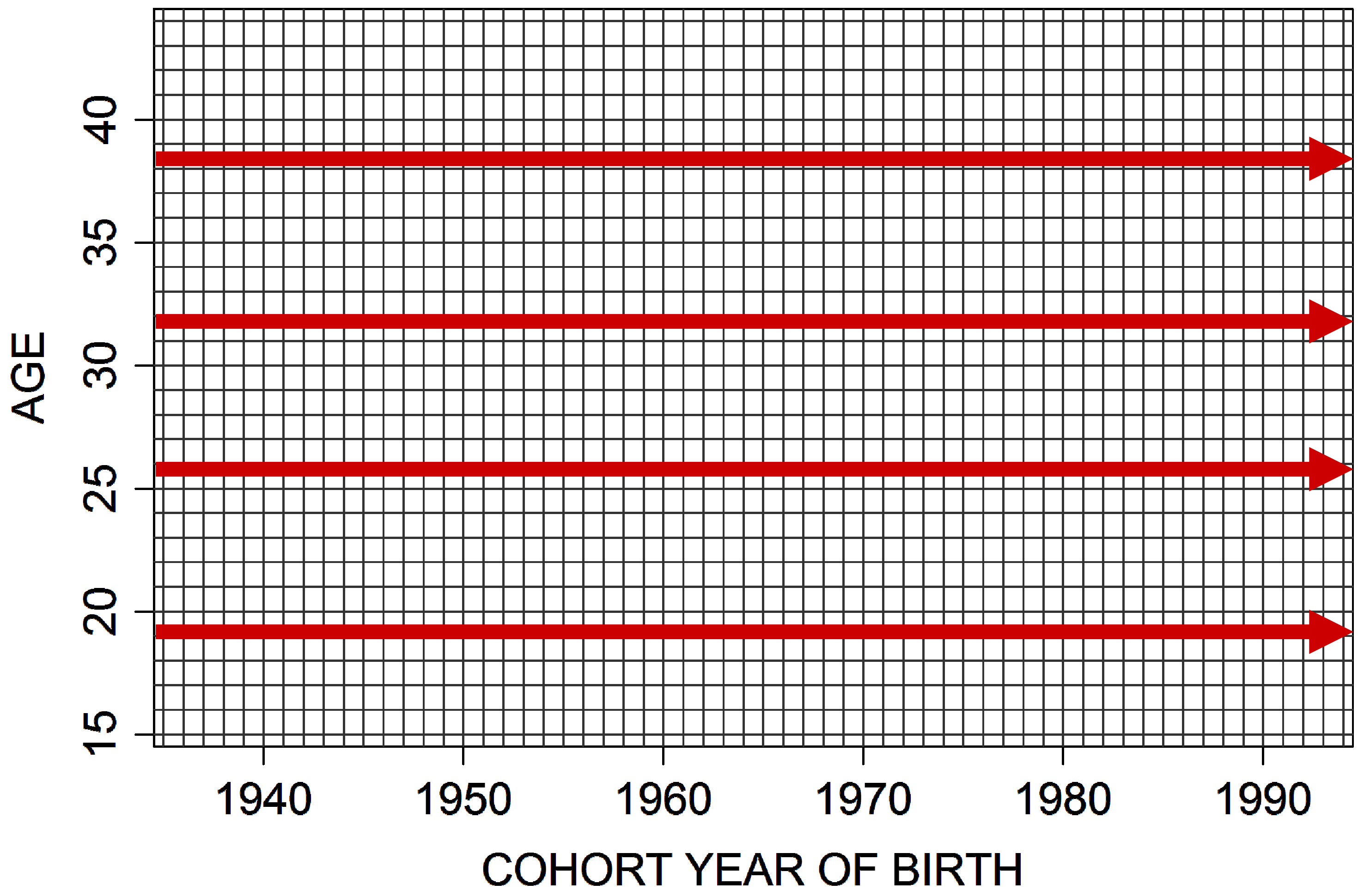
How common are alternative types of surfaces θ in the HFD?

NORMAL Posterior ← **Proper NORMAL** **Improper NORMAL with quadratic penalties in θ**

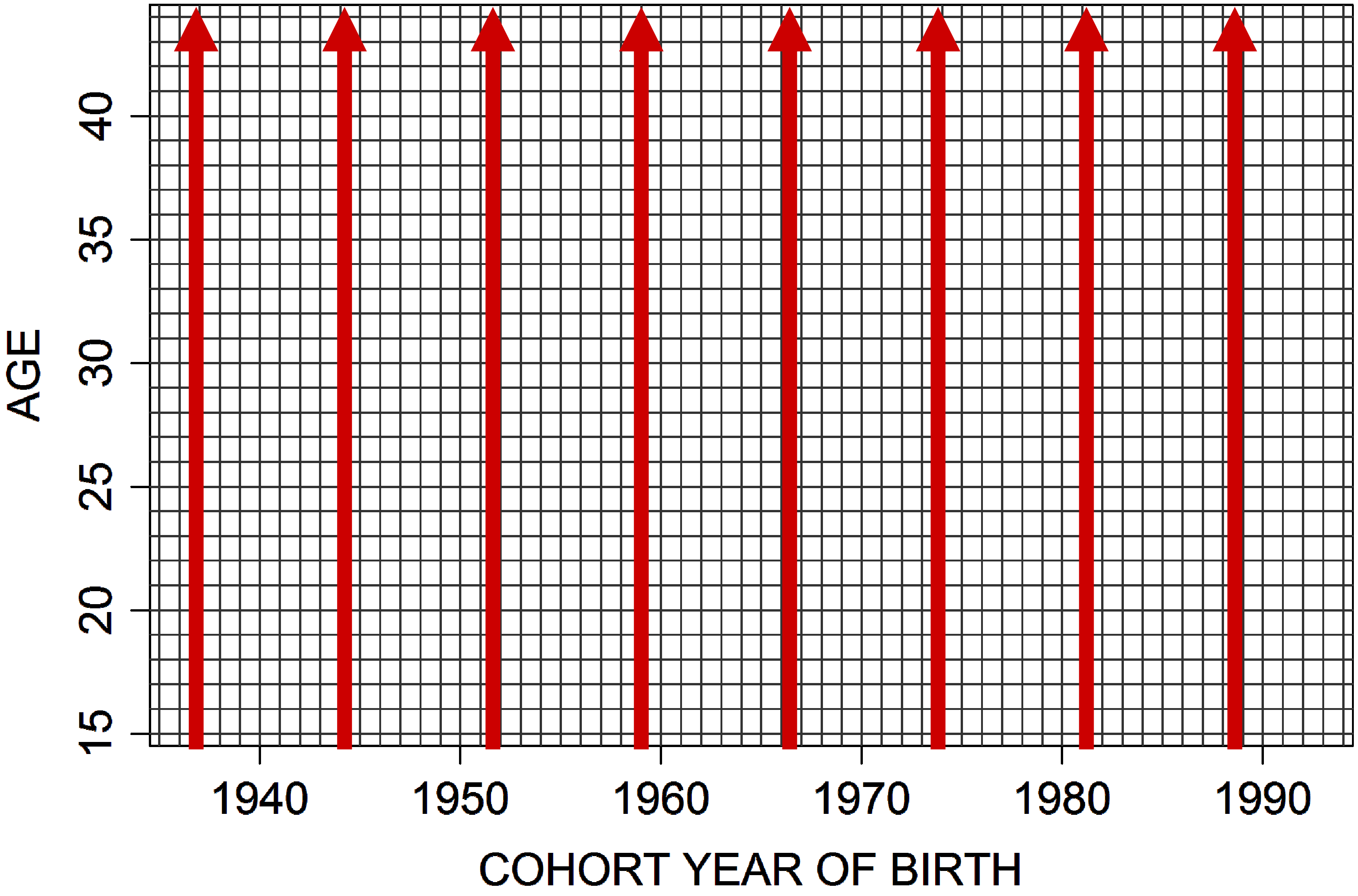
Bayesian Forecast Results

- (Closed-form) posterior mean vector and covariance matrix describe
  - Best-guess fit to observations
  - Best-guess forecasts
  - Uncertainty
  - Means and variances of **CFR**s and **ASFR**s

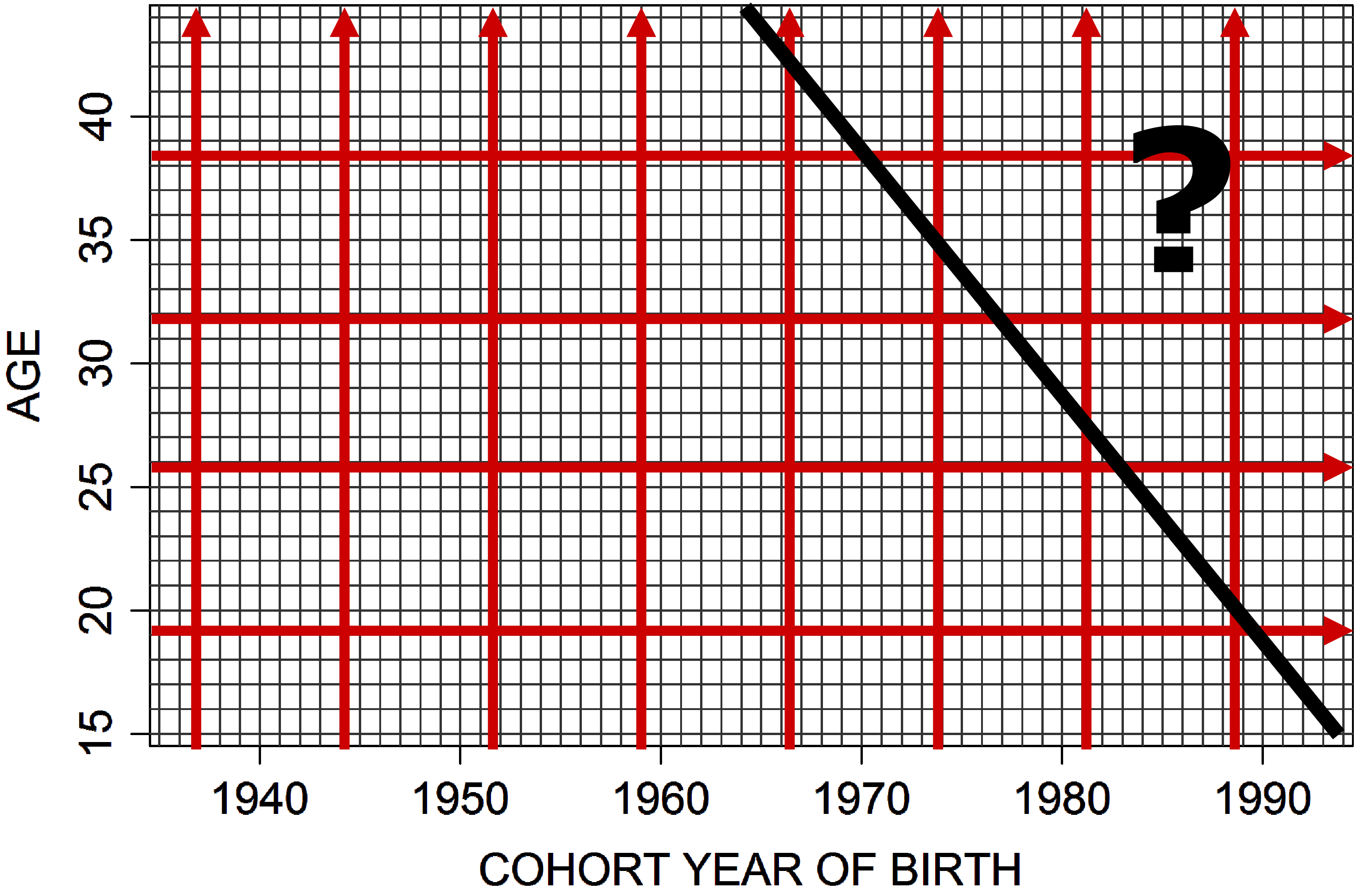
PRIOR #1:
How smooth is a time series likely to be at a given age?
... find out from HFD



PRIOR #2:
What are typical shapes of cohort schedules?
… find out from HFD



COMBINED PRIORS:
What are likely/unlikely Lexis surfaces? **f(θ | HFD)**



Using the HFD to build/calibrate

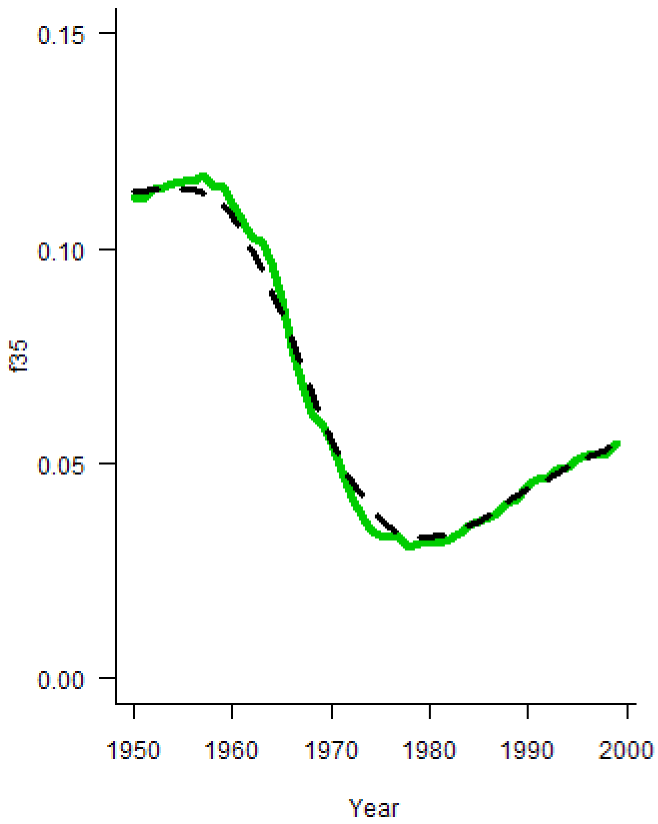
- Define squared-error penalties for each prior (high penalty → low prior prob.)

- Calibrate penalty weights to HFD data

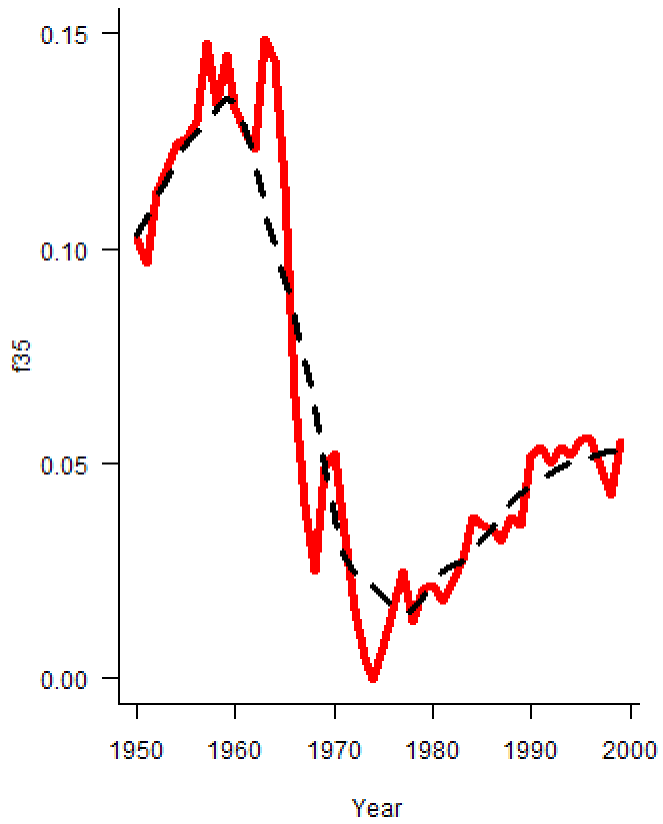
  → "time series as wiggly as in HFD"
  → "cohort shapes regular as in HFD"

PRIOR #1: Time Series are locally linear

MORE LIKELY

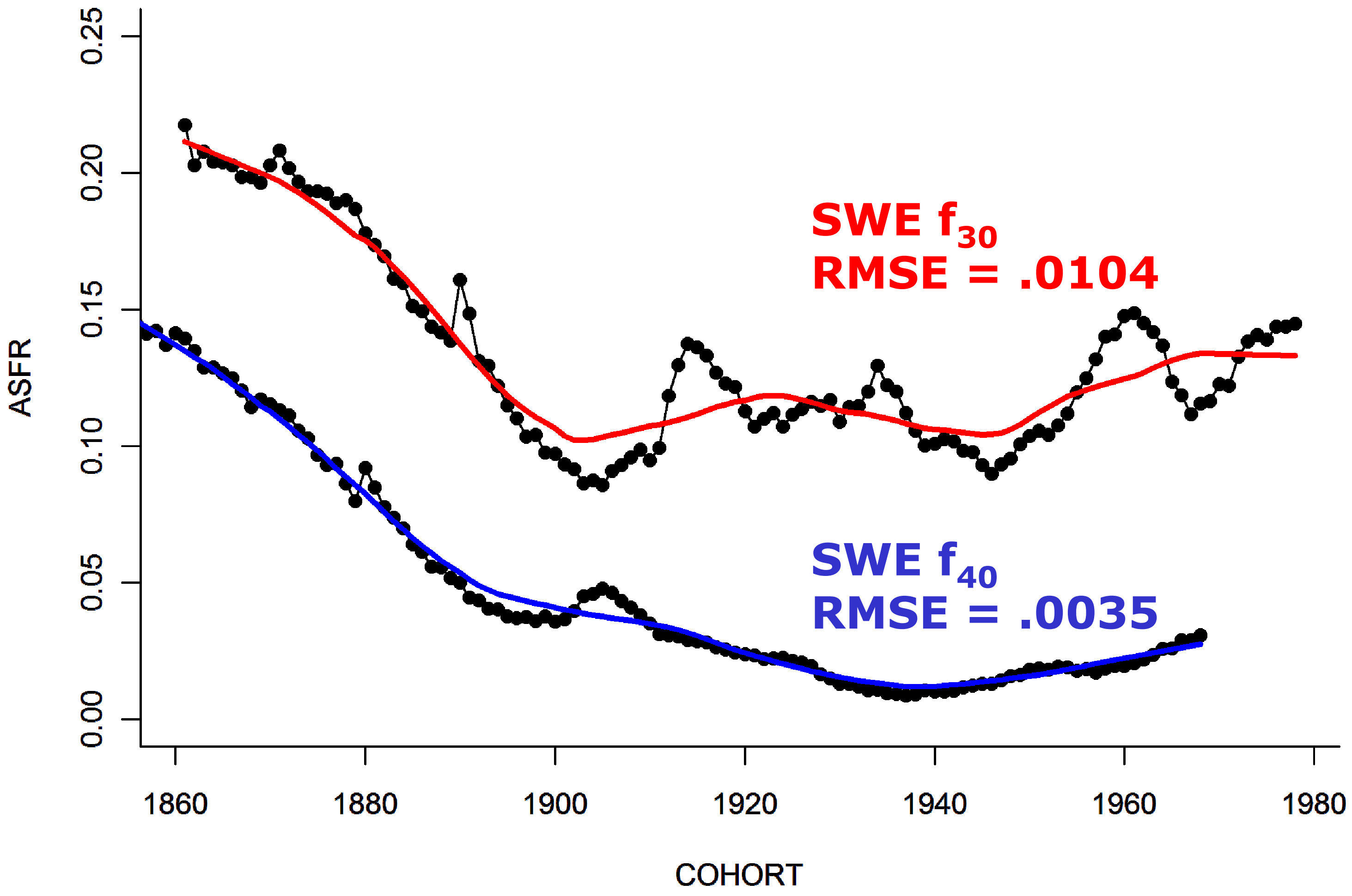


LESS LIKELY



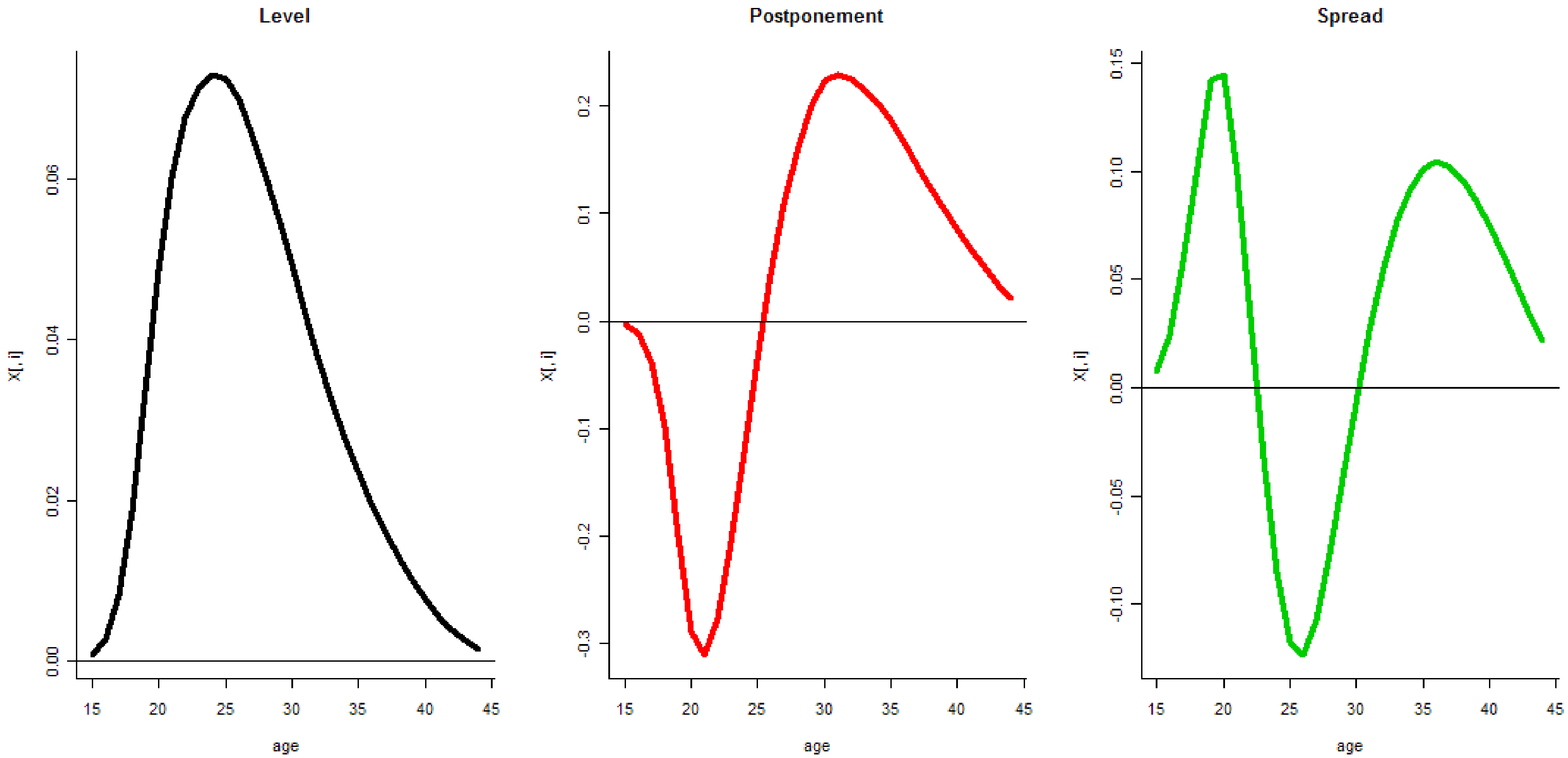
$$\boxed{w_a \, \epsilon_a' \epsilon_a \; = \; w_a \, \phi_a' \, (I - S)' \, (I - S) \, \phi_a \qquad a = 1 \ldots A}$$

PRIOR #1: Time Series are locally linear



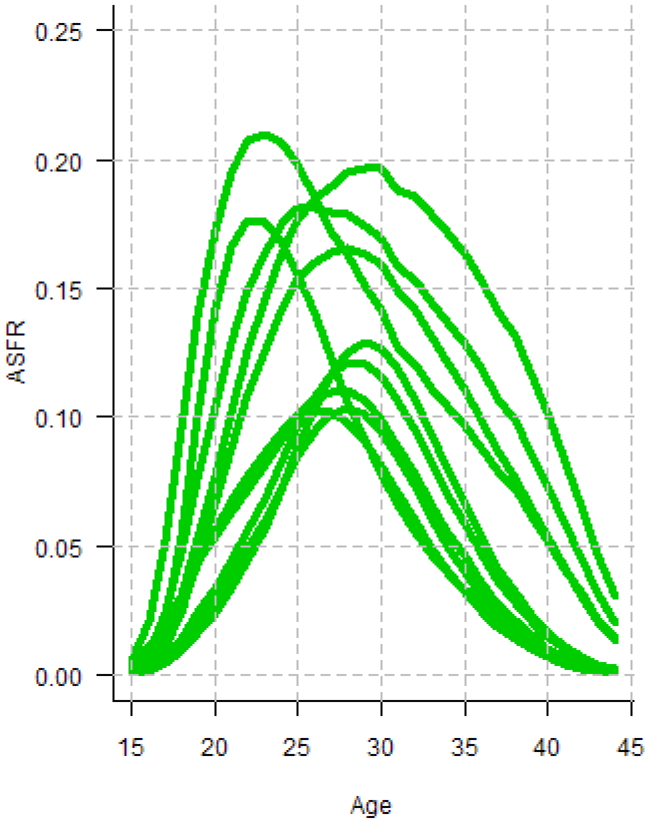
Calibrate prior so that E_f[RMSE(θ)] = avg RMSE(θ) in HFD

PRIOR #2: Cohort schedules are well approx. by SVD components from **HFD**

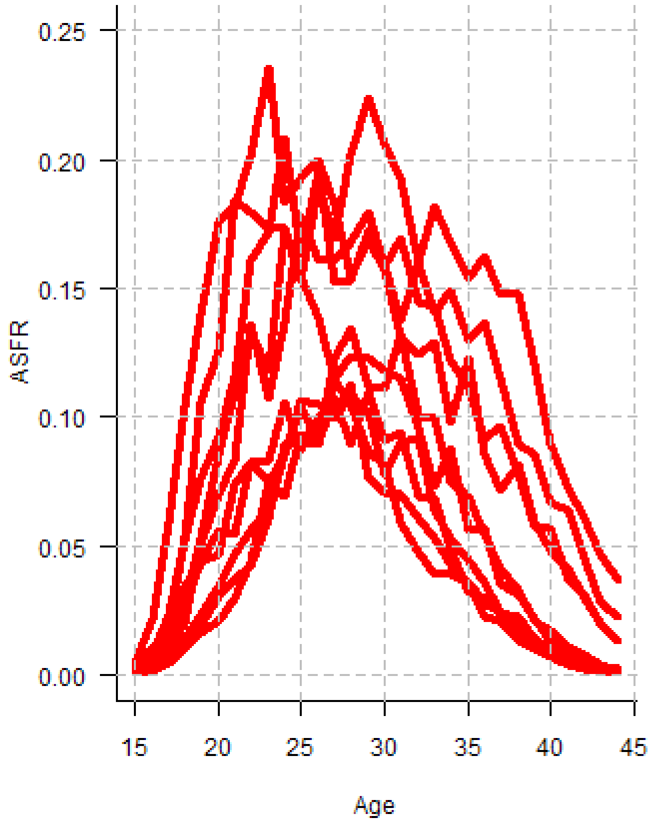


PRIOR #2: Cohort schedules are well approx. by SVD components from **HFD**

MORE LIKELY

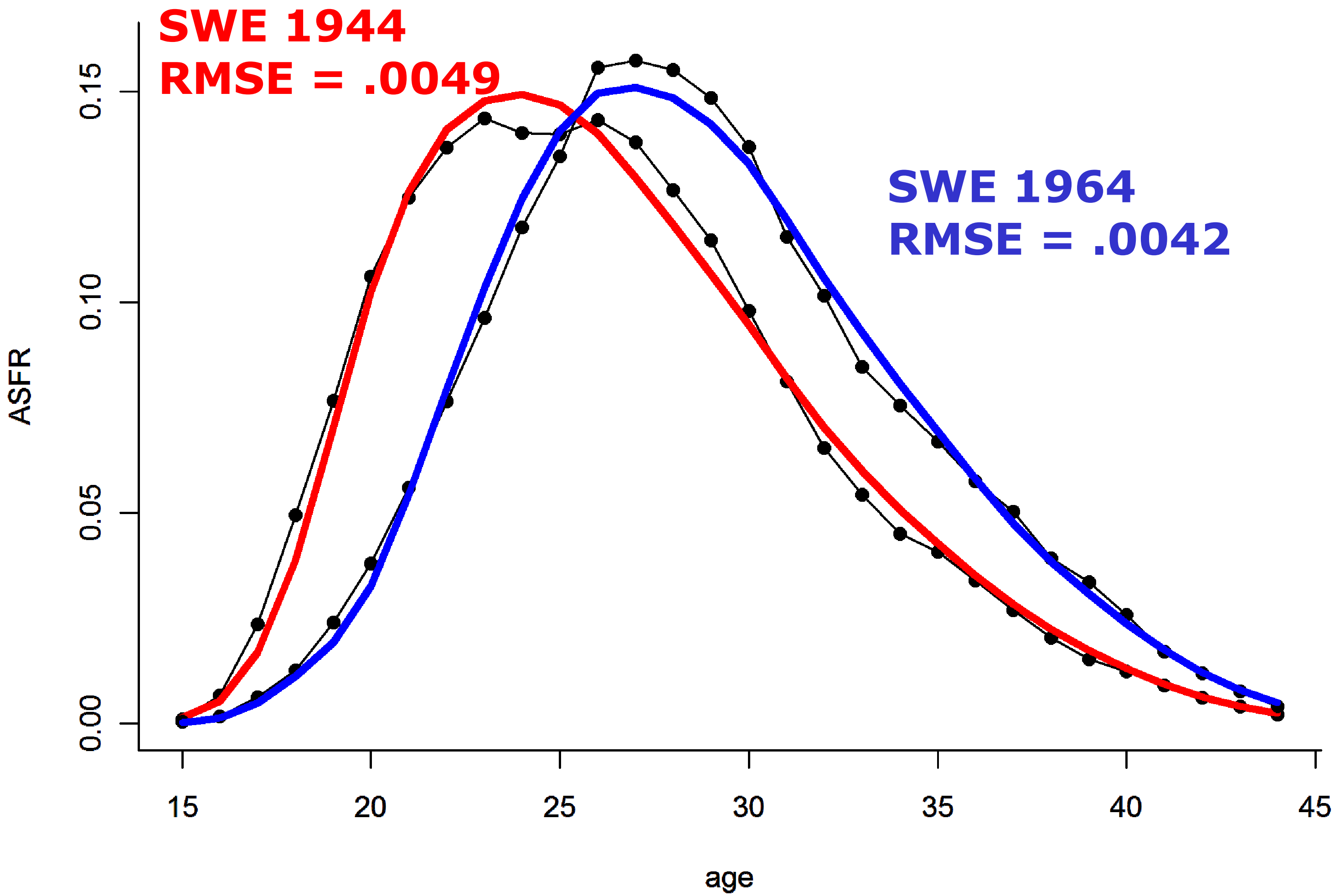


LESS LIKELY

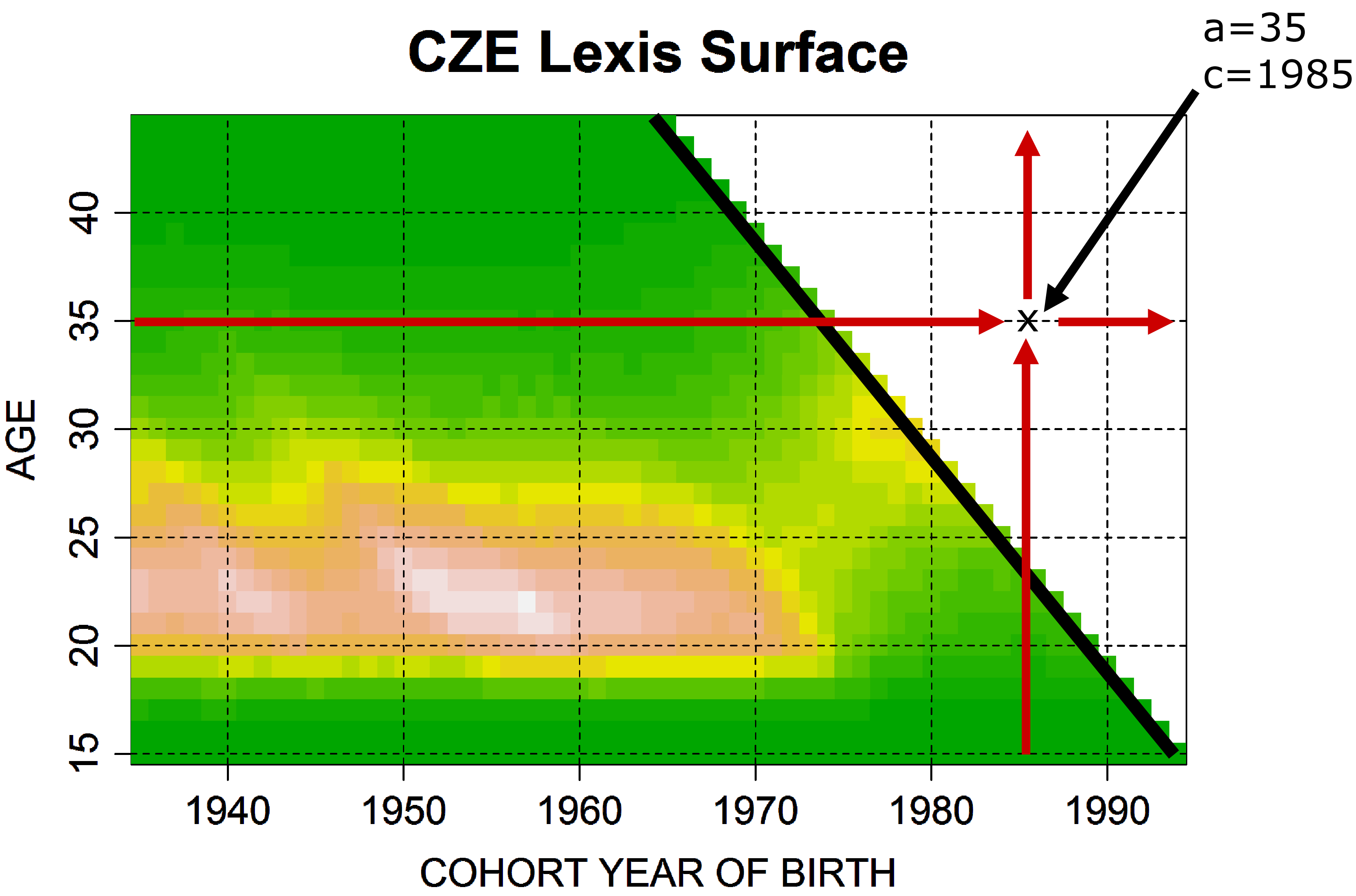


$$w_1 \, \varepsilon_t' \varepsilon_t \; = \; w_1 \, (\phi_t - \bar{H})' \, M \, (\phi_t - \bar{H}) \qquad\qquad t = 1 ... T$$

PRIOR #2: Cohort schedules are well approx. by SVD components from **HFD**



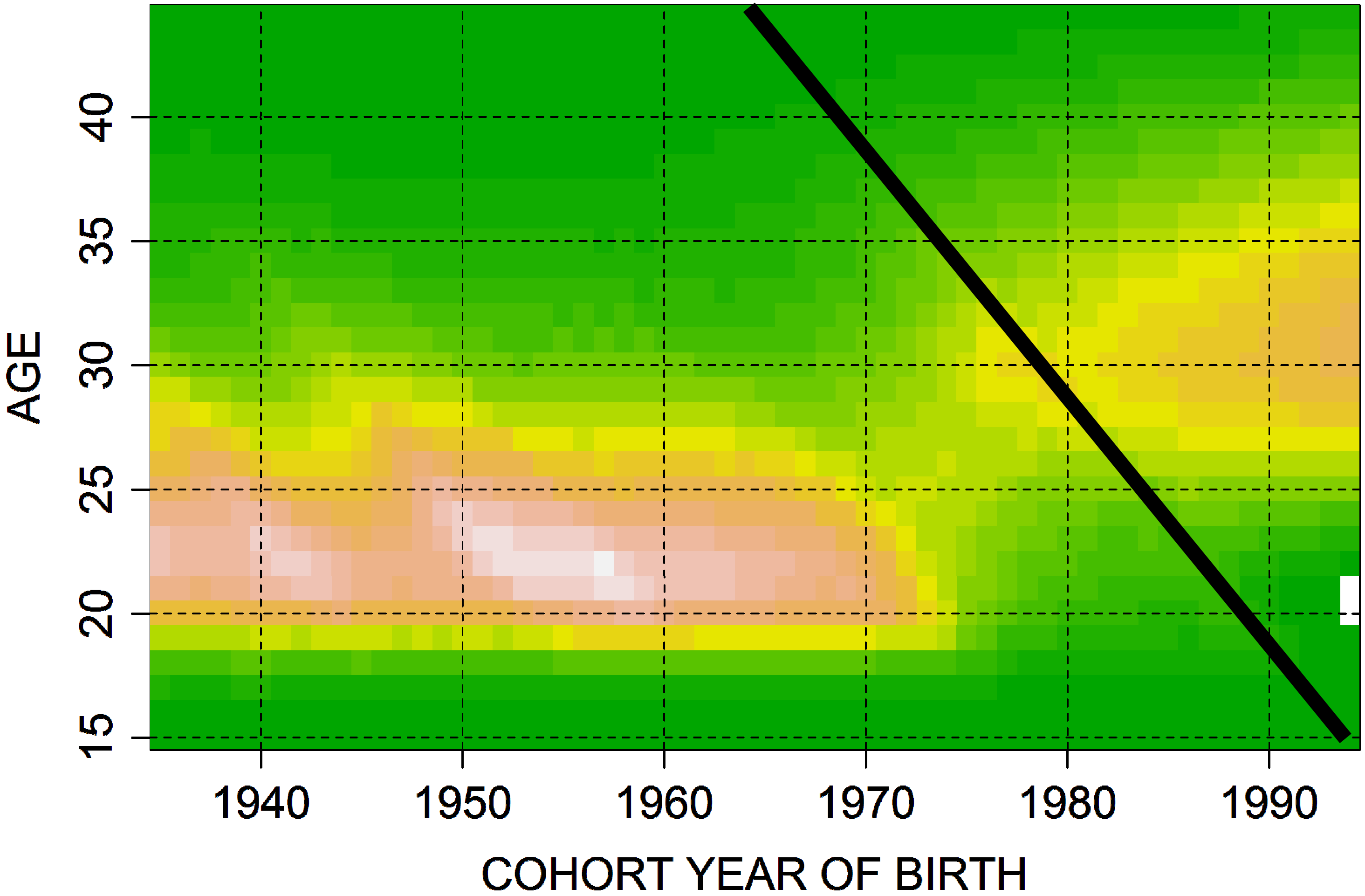
Calibrate prior so that E_f[RMSE(θ)] = avg RMSE(θ) in HFD



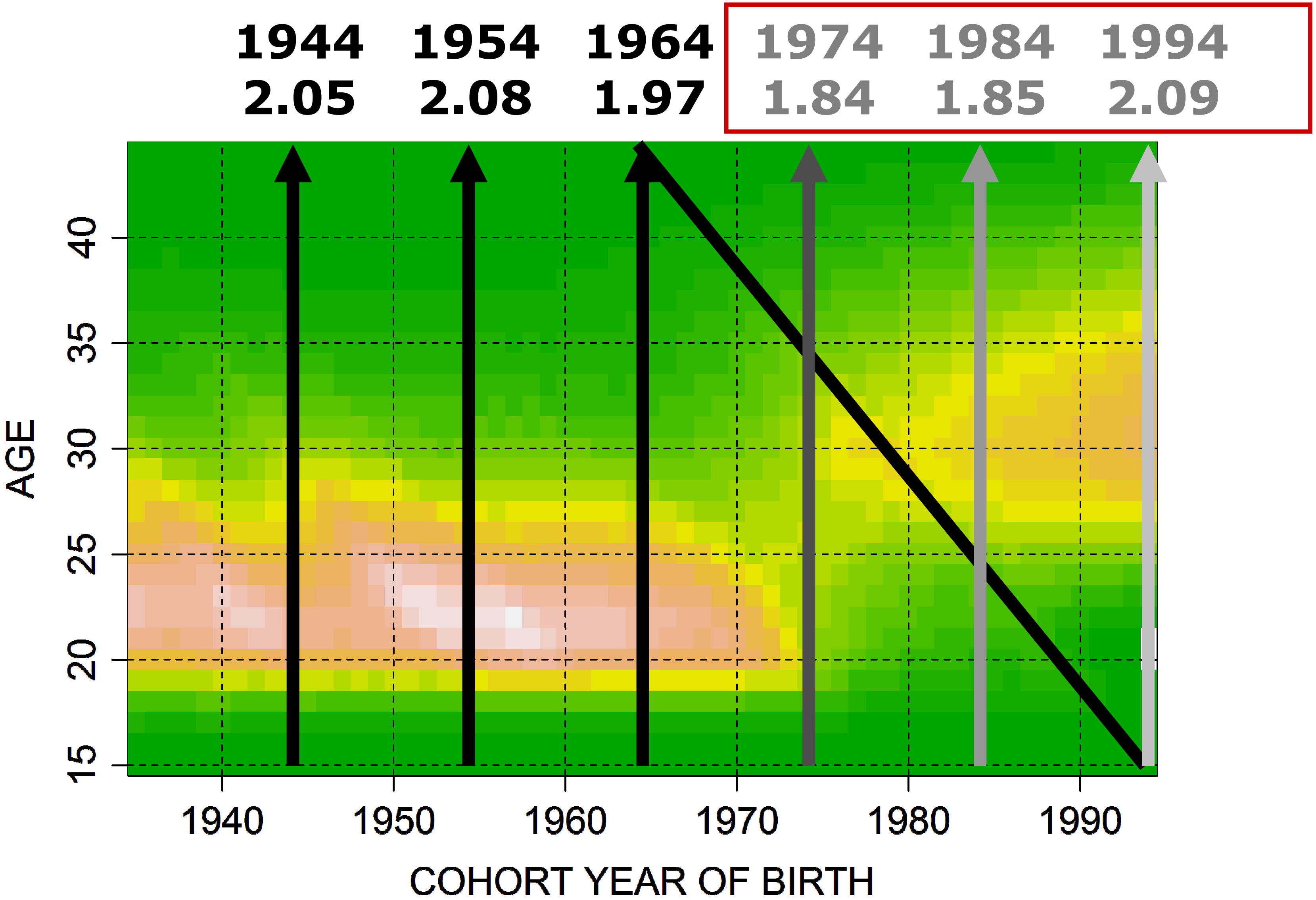
CZE Lexis Surface
a=35
c=1985
AGE
15
20
25
30
35
40
1940
1950
1960
1970
1980
1990
COHORT YEAR OF BIRTH

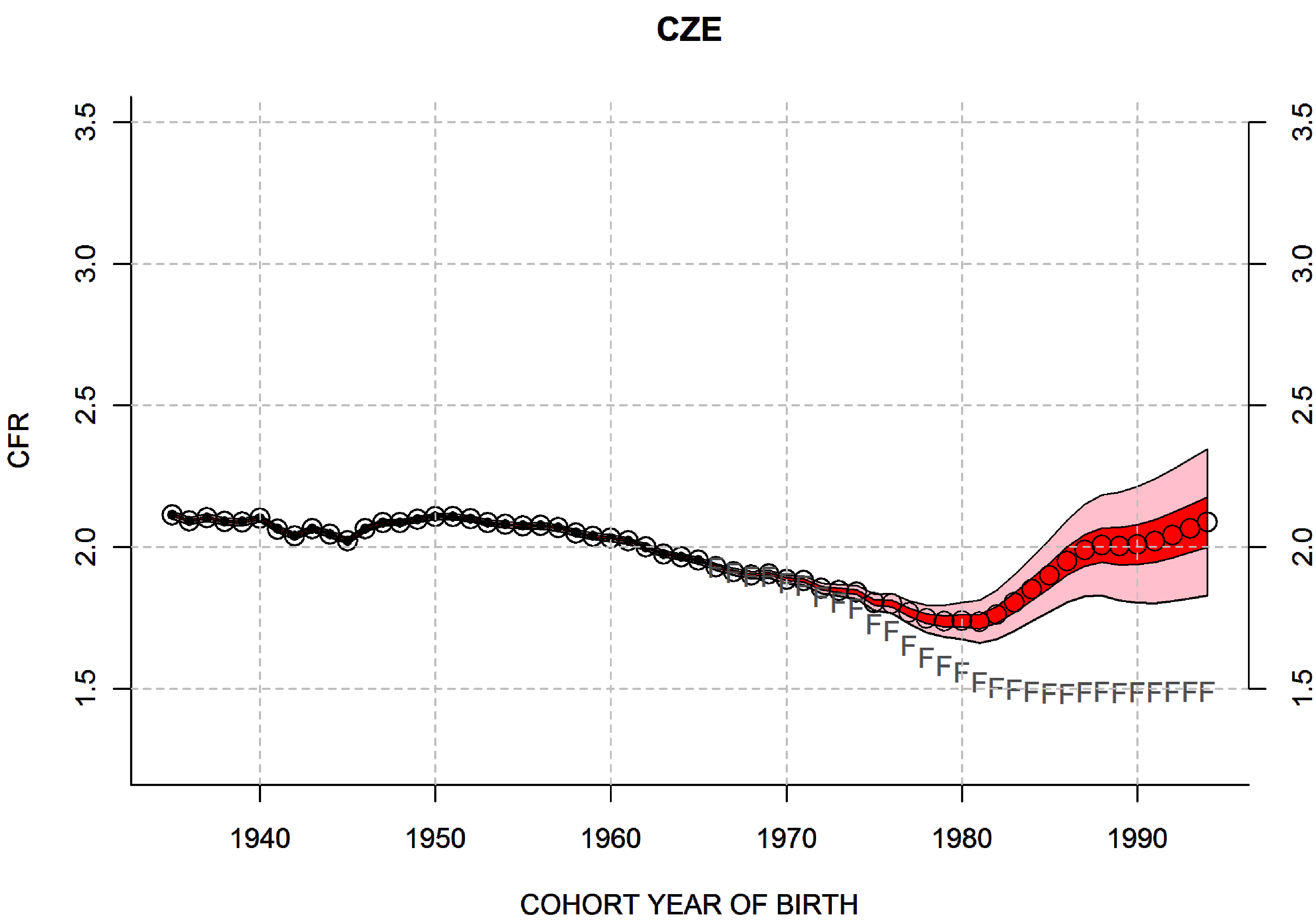
Maximum a posteriori surface

(+ sd and covariances of cell estimates)

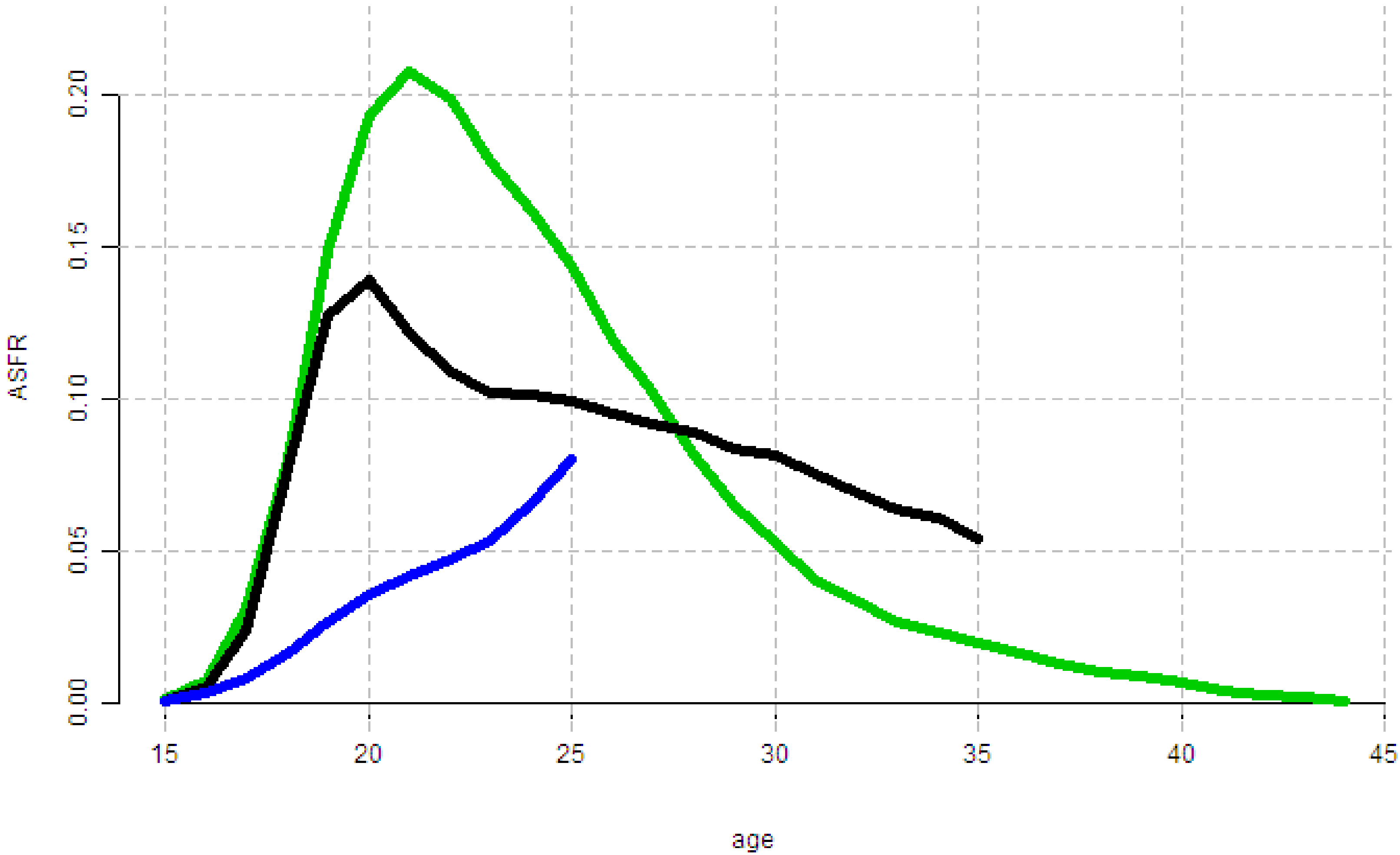


θ is uncertain → std errors

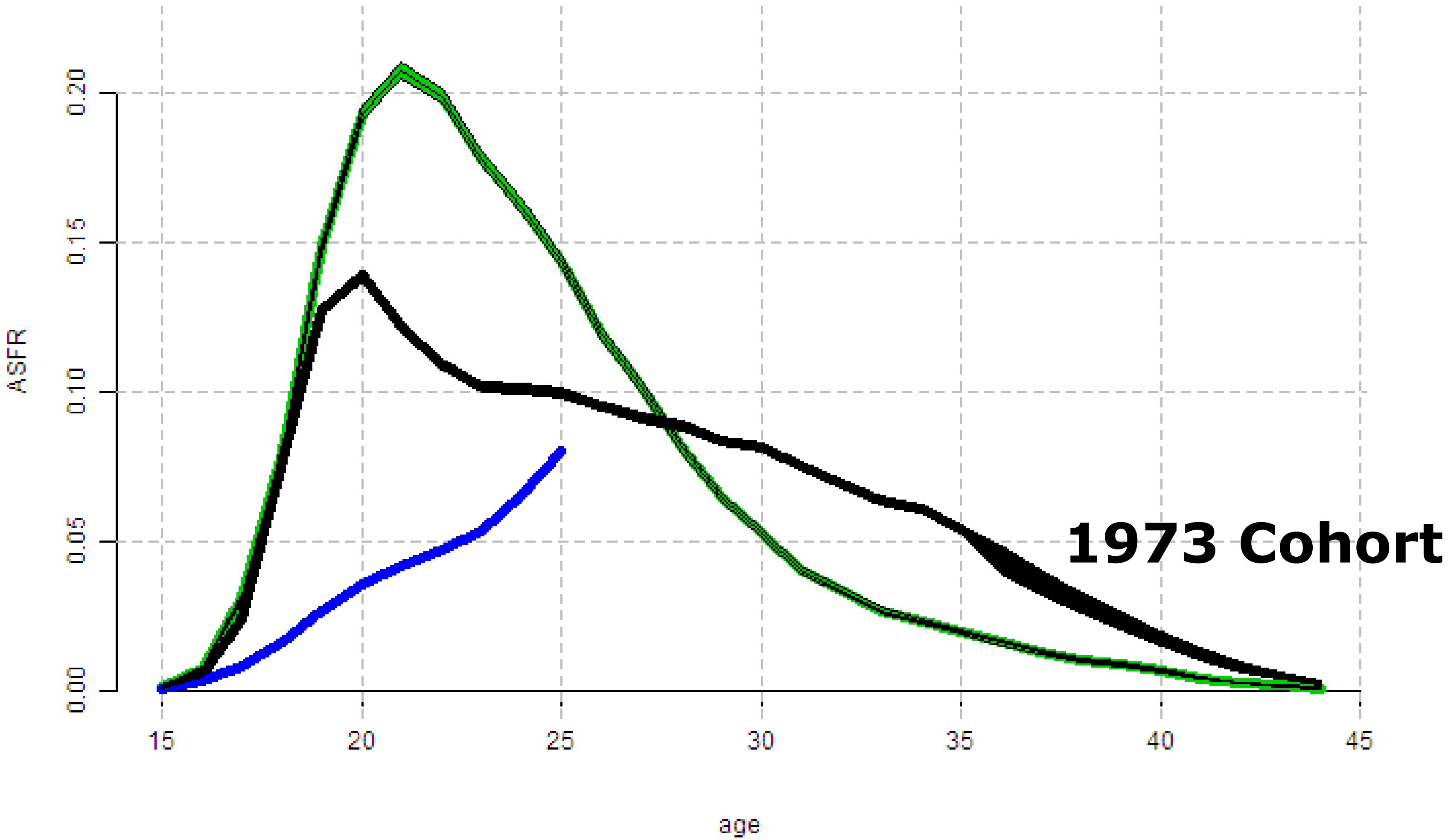



CZE
CFR
COHORT YEAR OF BIRTH
1940
1950
1960
1970
1980
1990
1.5
2.0
2.5
3.0
3.5

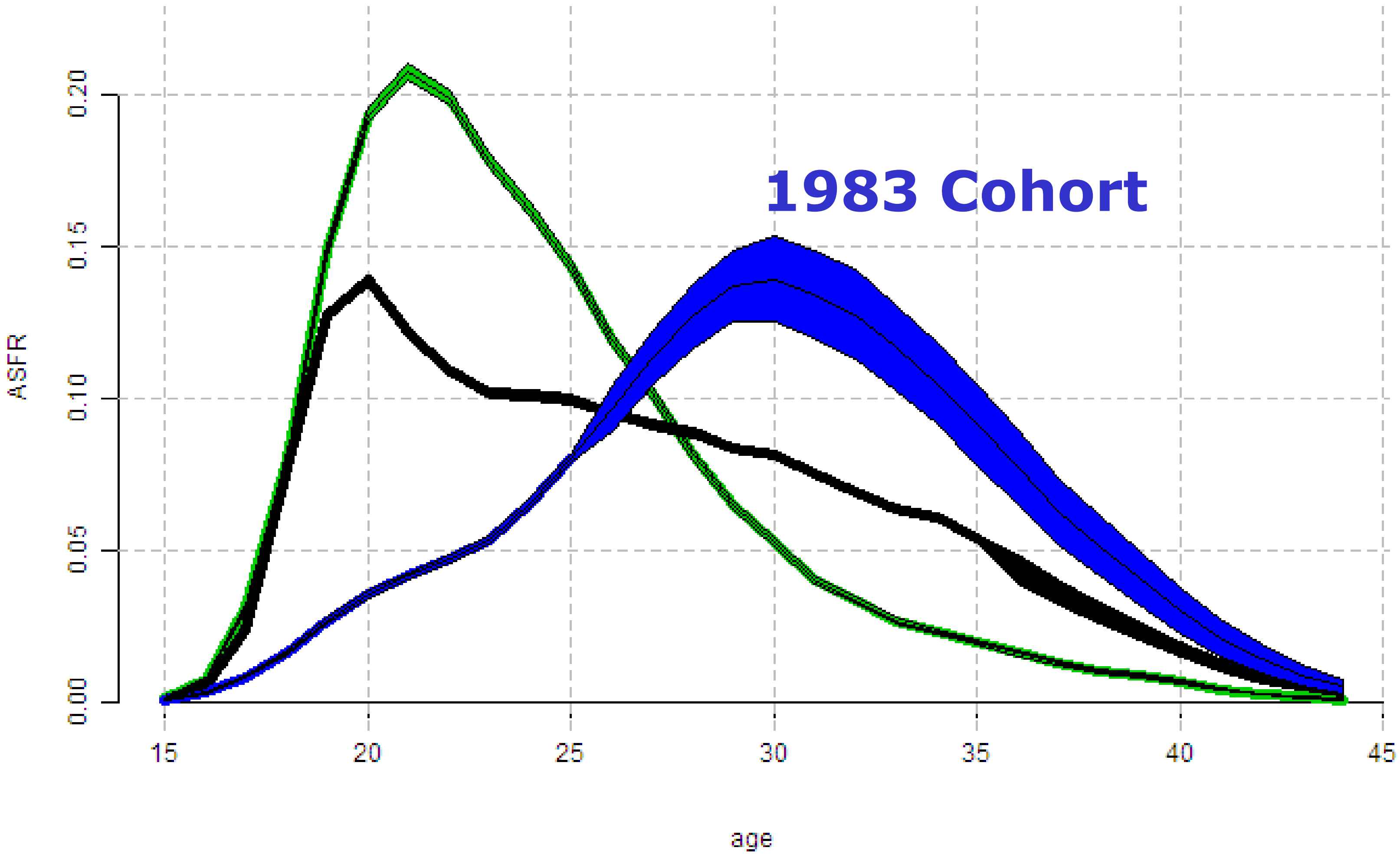
Bayesian Forecast Results:
CZE Cohort Schedules



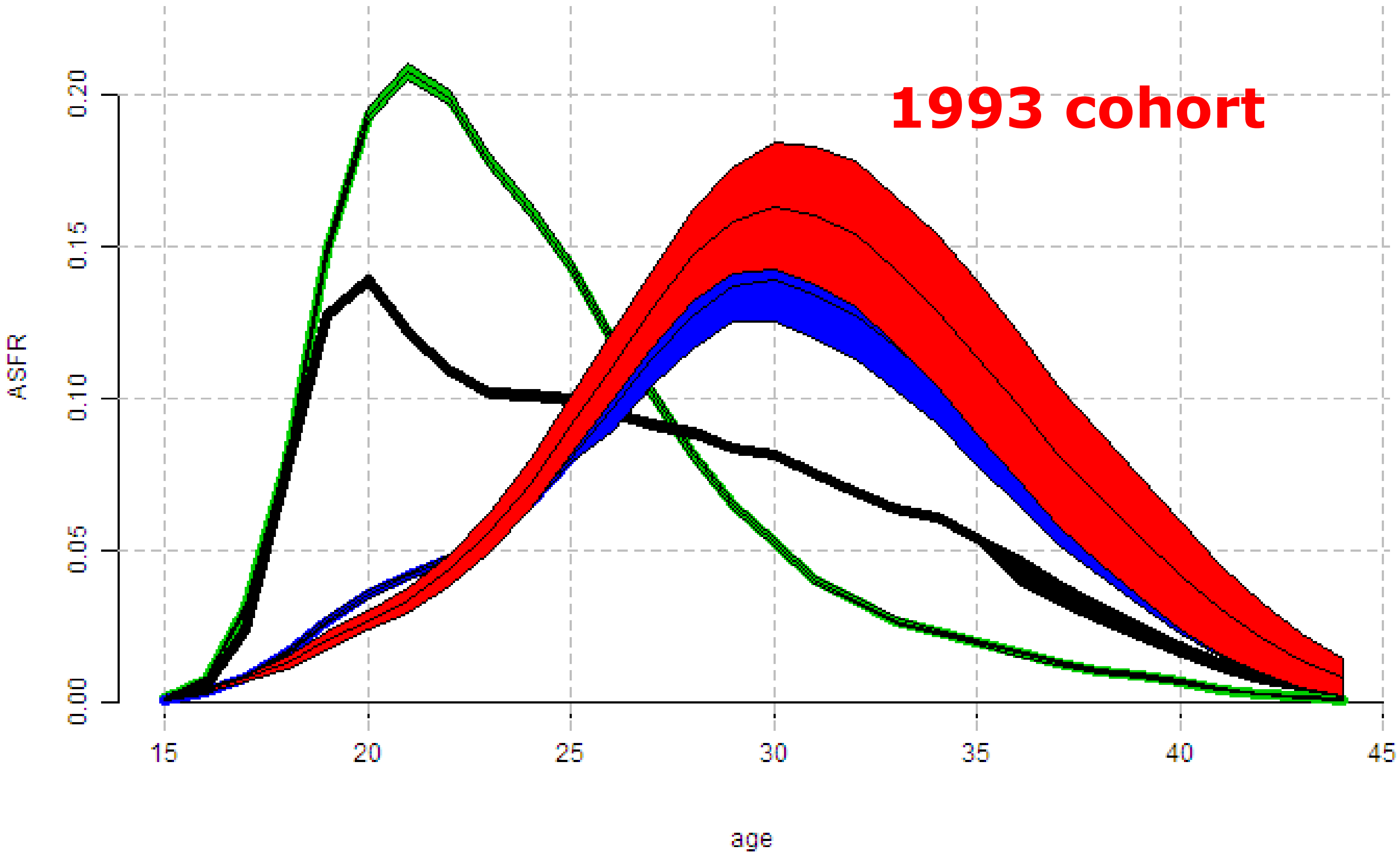
Bayesian Forecast Results:
CZE Cohort Schedules



Bayesian Forecast Results: CZE Cohort Schedules



Bayesian Forecast Results: CZE Cohort Schedules



LESSONS LEARNED

- Incorporating qualitative information about Lexis surfaces into a forecast is feasible
- HFD data are valuable for building priors that describe qualitative features of fertility surfaces (smoothness, shapes, etc.)
- Uncertainty estimates **still need work**
- Cohort CFR seems likely to rise, at least a little, in many low-fertility countries

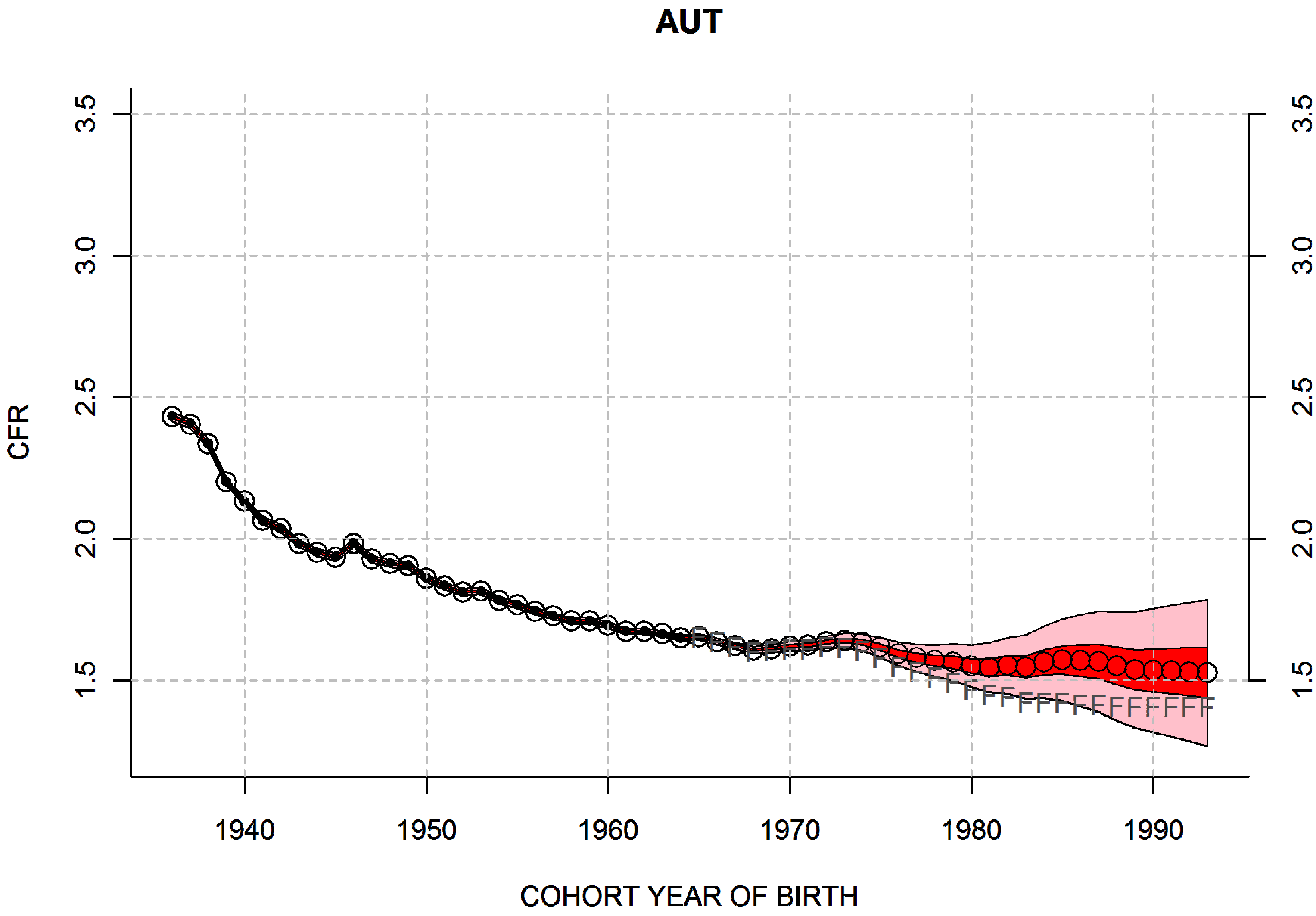
Coming soon...

- Forecasts for other countries (esp. Southern Europe)
- Evaluate uncertainty estimates: Simulate “forecasts” made in 1985, 1990,… and compare to later observations
- Experiments with (much more flexible but much slower) MCMC estimation methods

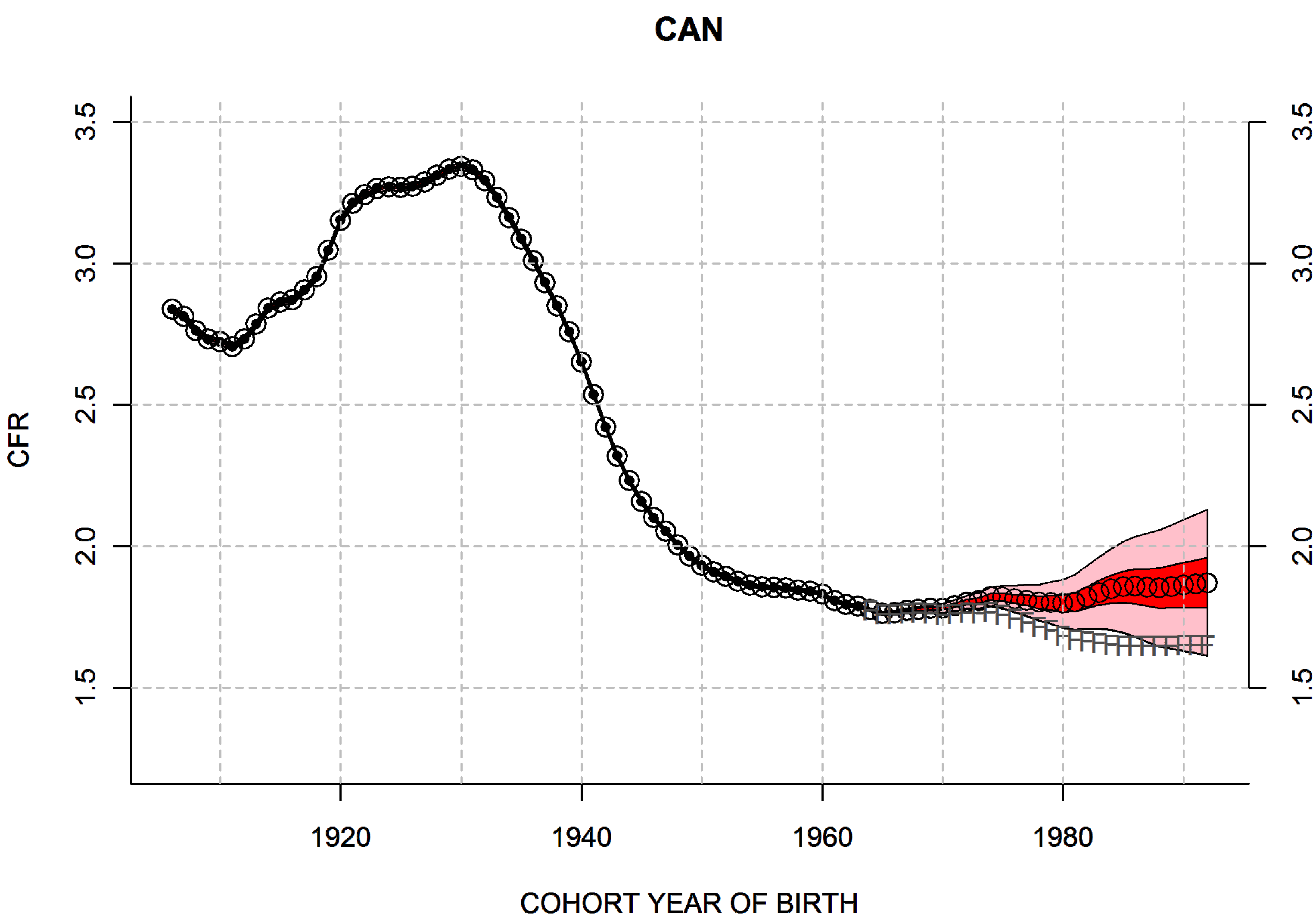
Thanks!

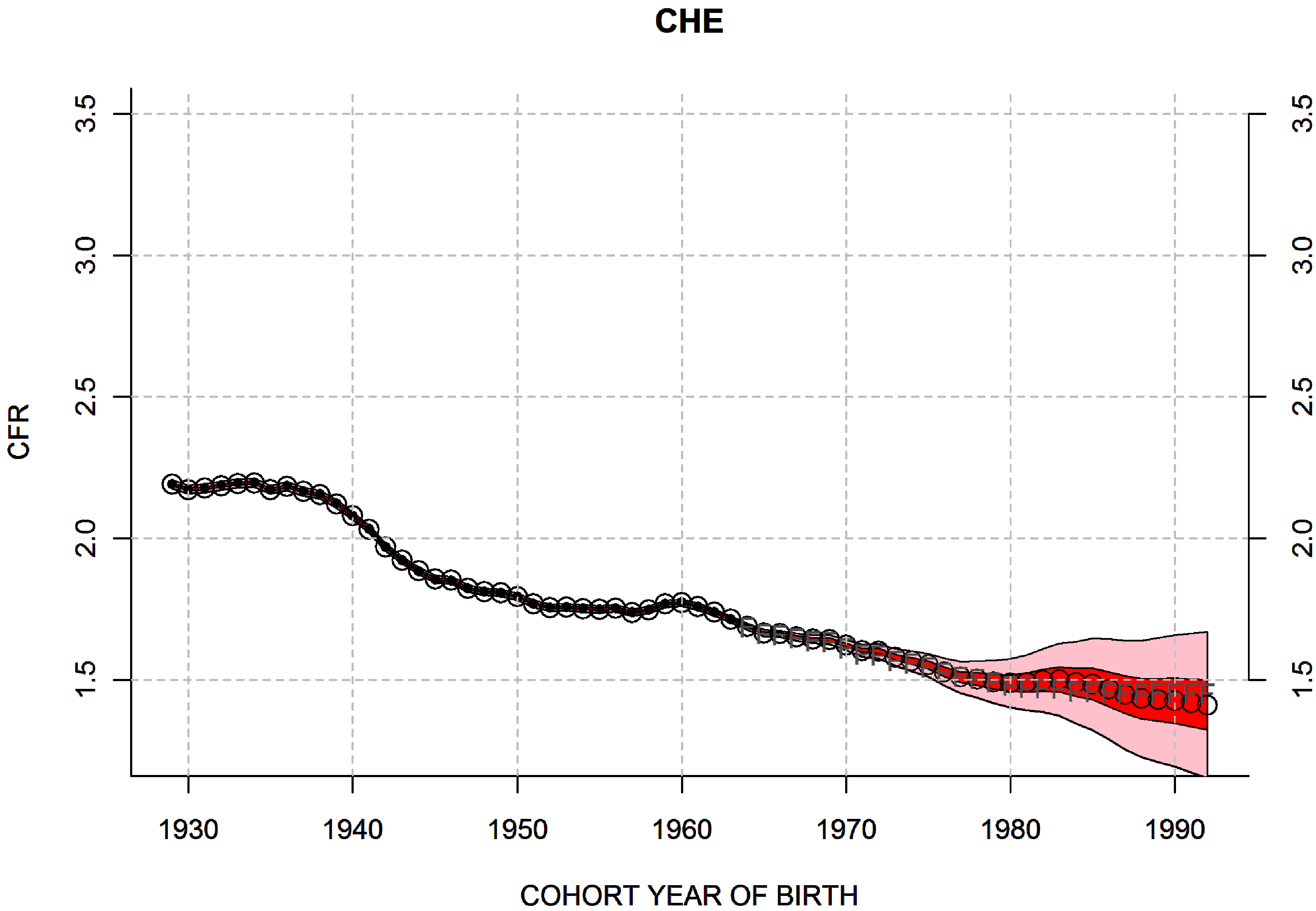
Vielen Dank!

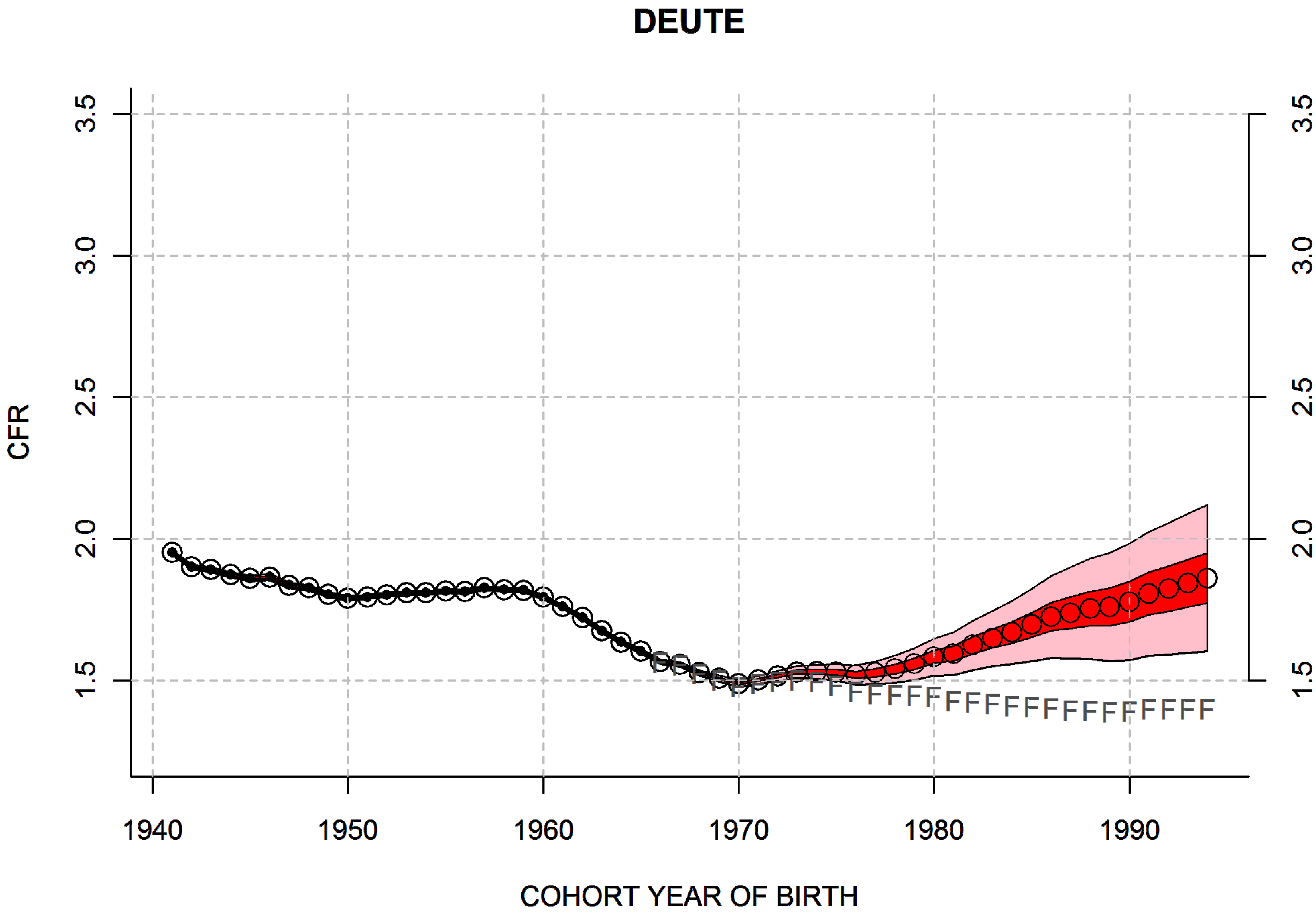
Extra stuff...



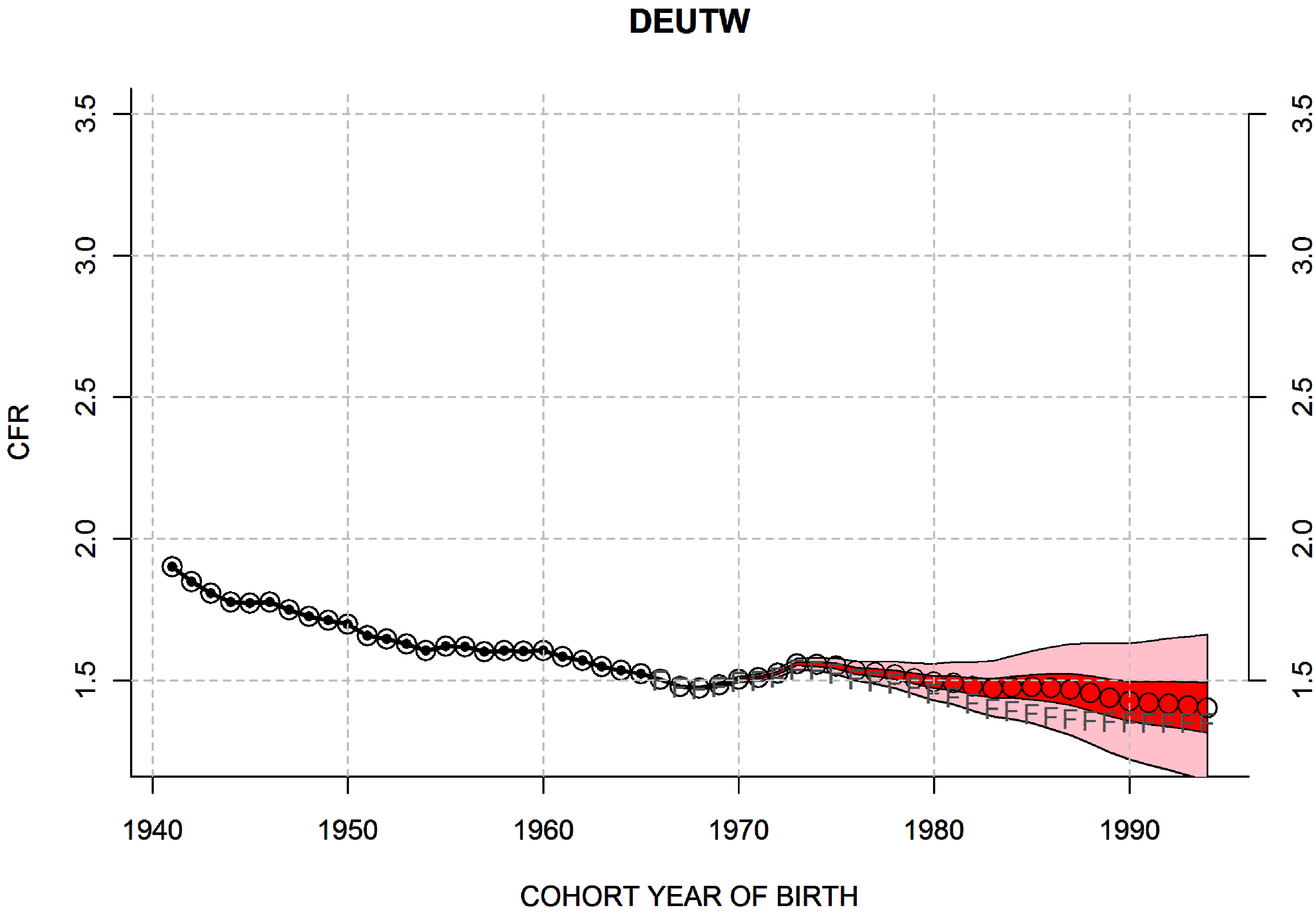
AUT
CFR
COHORT YEAR OF BIRTH
1940
1950
1960
1970
1980
1990
1.5
2.0
2.5
3.0
3.5

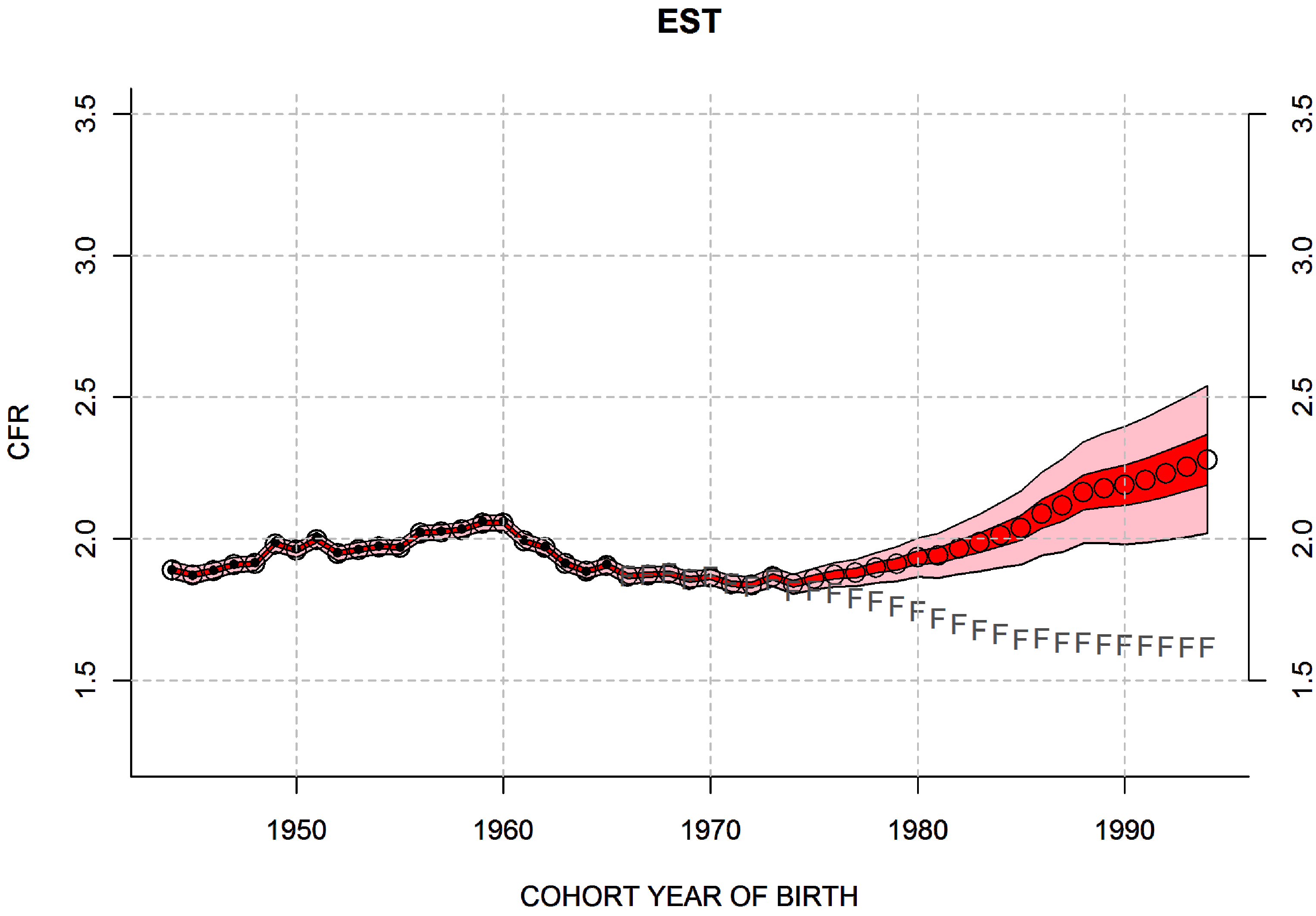

CAN
CFR
3.5
3.0
2.5
2.0
1.5
1920
1940
1960
1980
COHORT YEAR OF BIRTH


CHE
CFR
COHORT YEAR OF BIRTH
1930
1940
1950
1960
1970
1980
1990
1.5
2.0
2.5
3.0
3.5

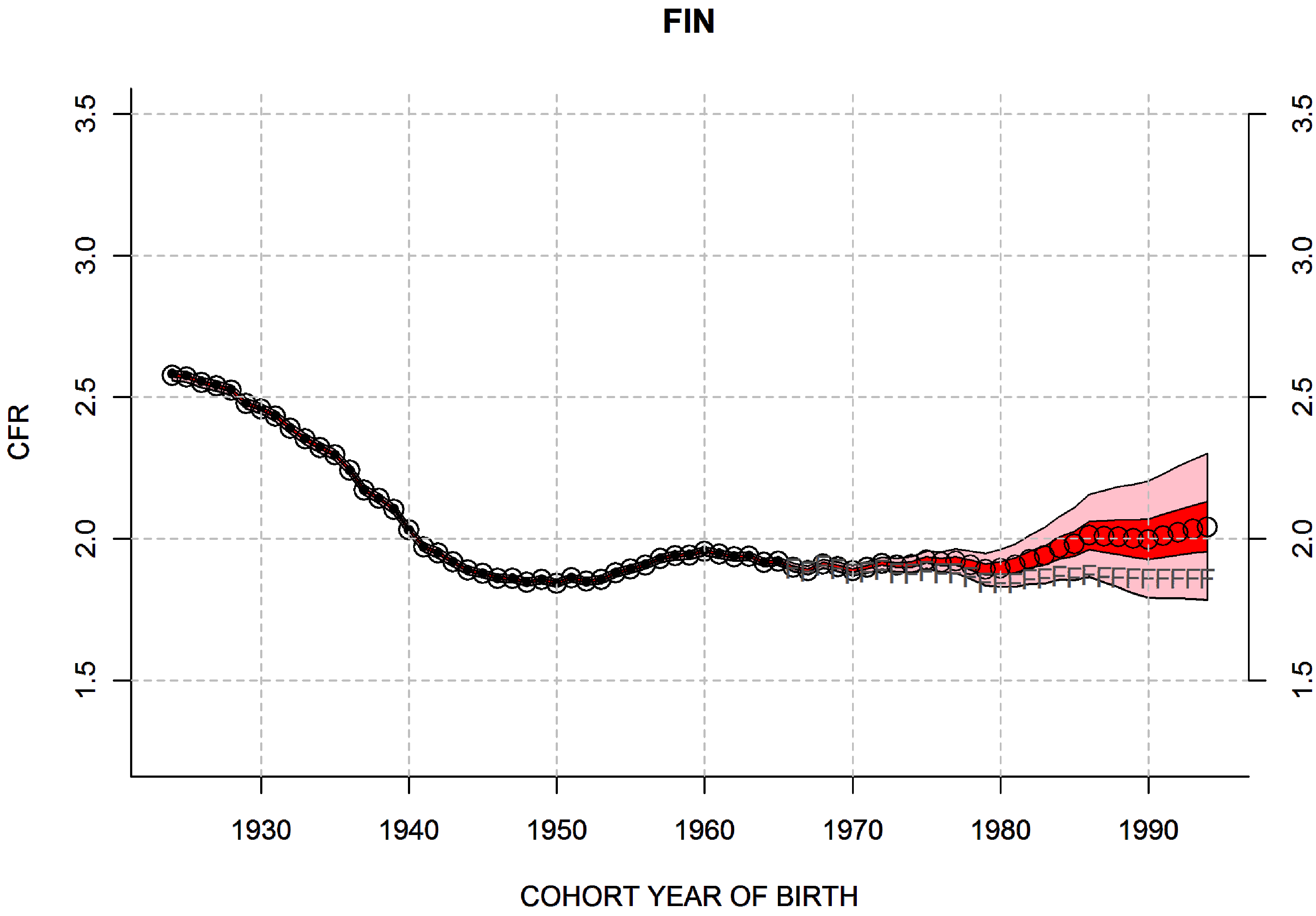


DEUTE
CFR
3.5
3.0
2.5
2.0
1.5
1940
1950
1960
1970
1980
1990
COHORT YEAR OF BIRTH

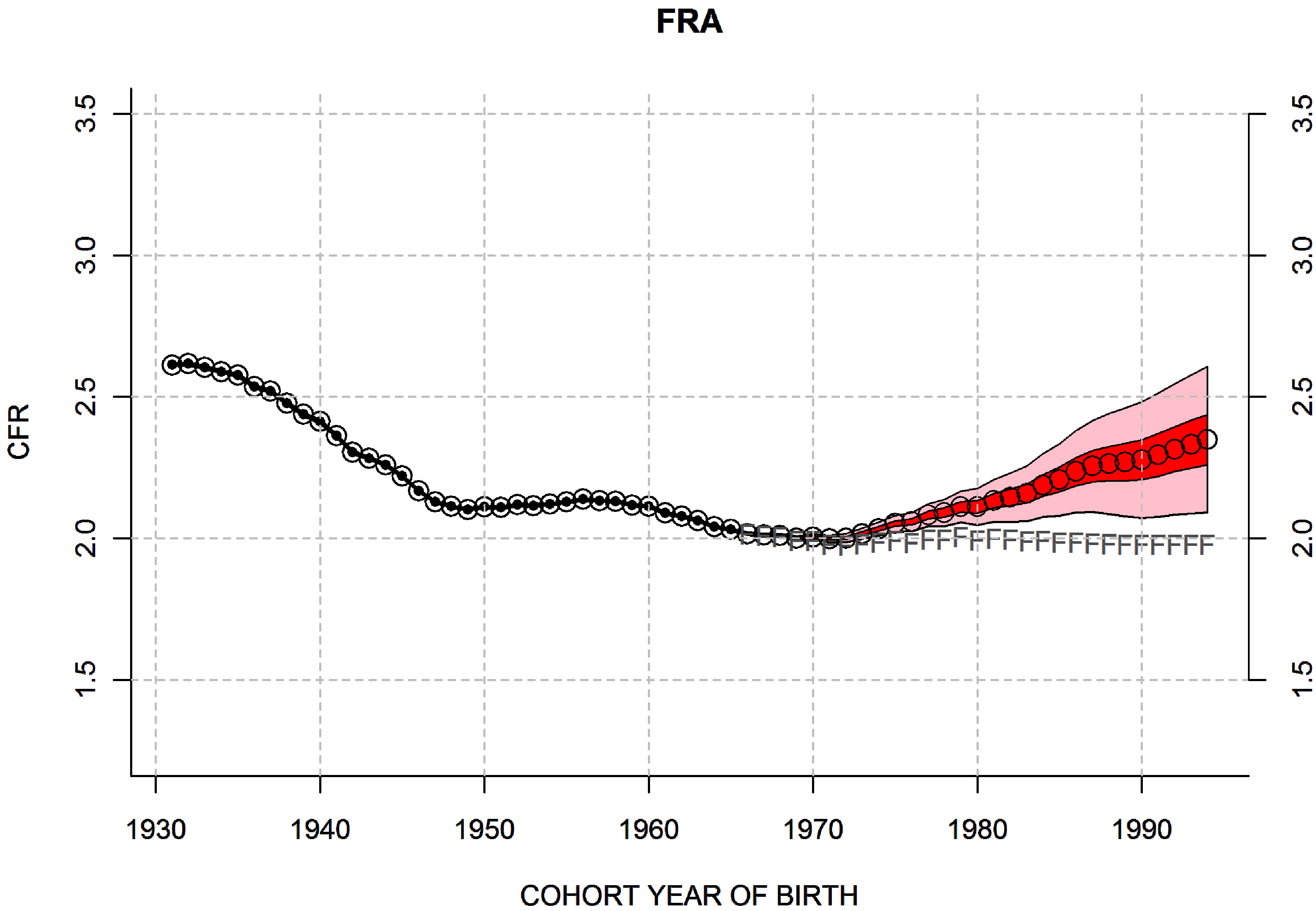

DEUTW
CFR
3.5
3.0
2.5
2.0
1.5
1940
1950
1960
1970
1980
1990
COHORT YEAR OF BIRTH

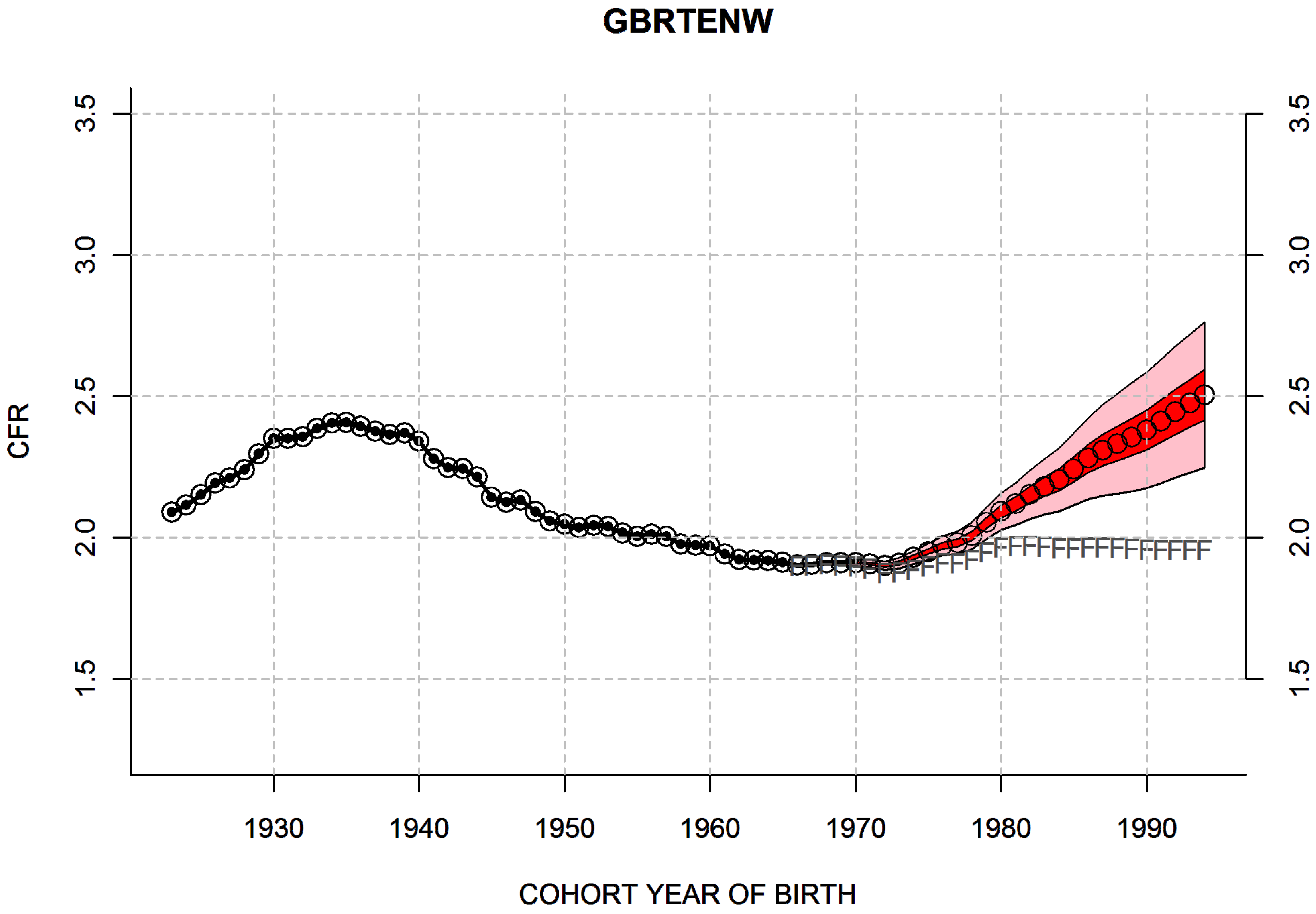


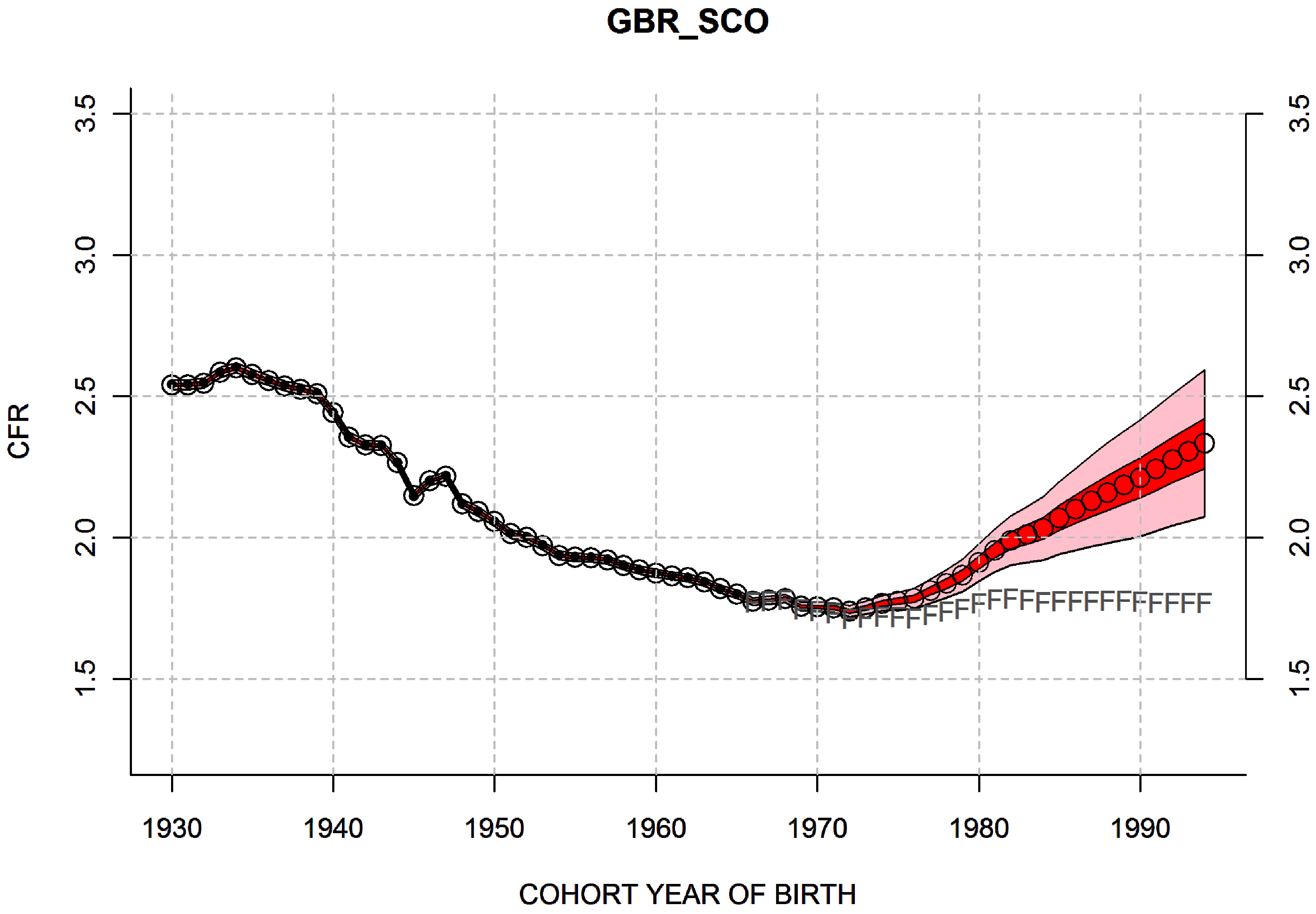
EST
CFR
3.5
3.0
2.5
2.0
1.5
1950
1960
1970
1980
1990
COHORT YEAR OF BIRTH



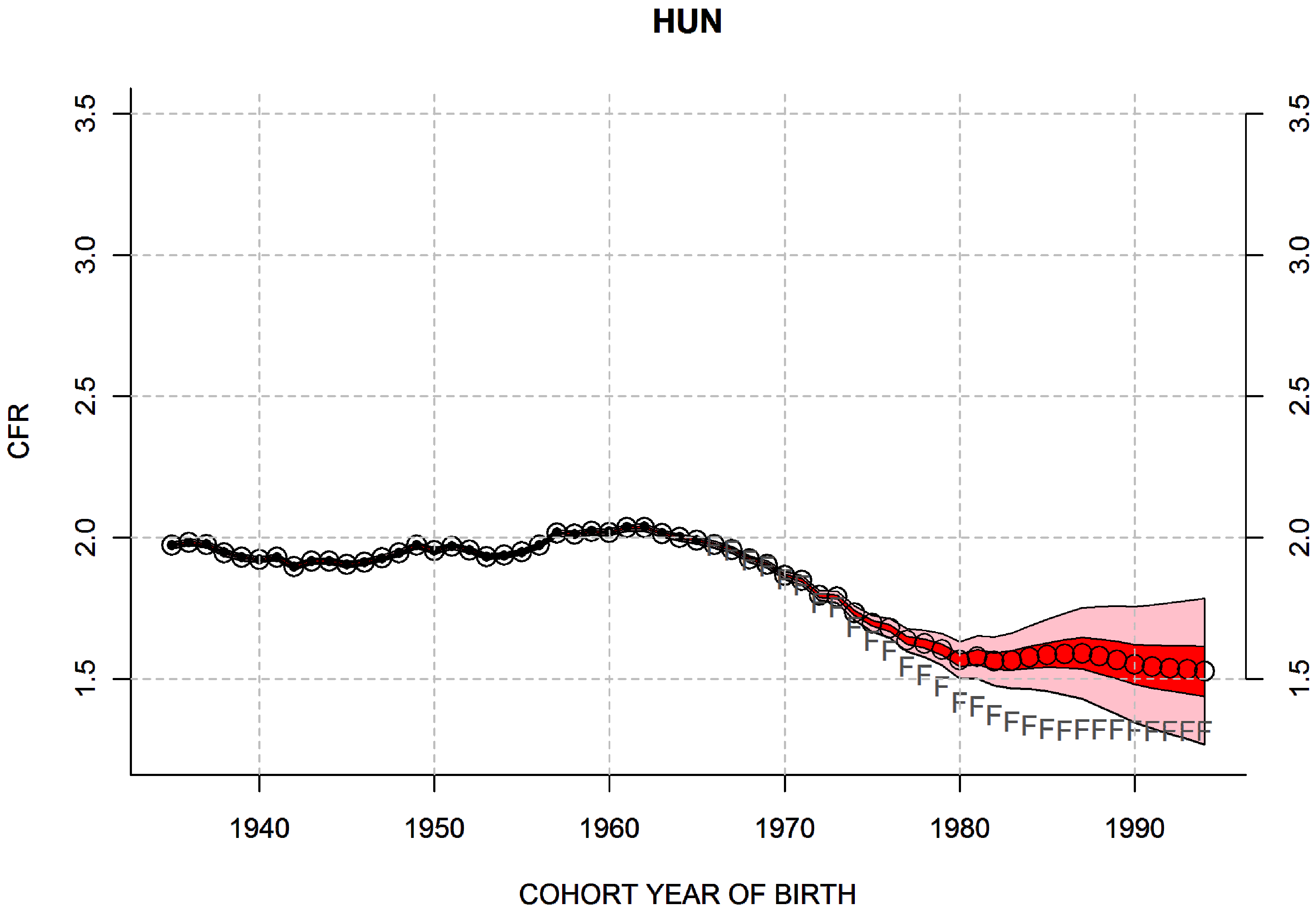
FIN
CFR
1.5
2.0
2.5
3.0
3.5
1930
1940
1950
1960
1970
1980
1990
COHORT YEAR OF BIRTH

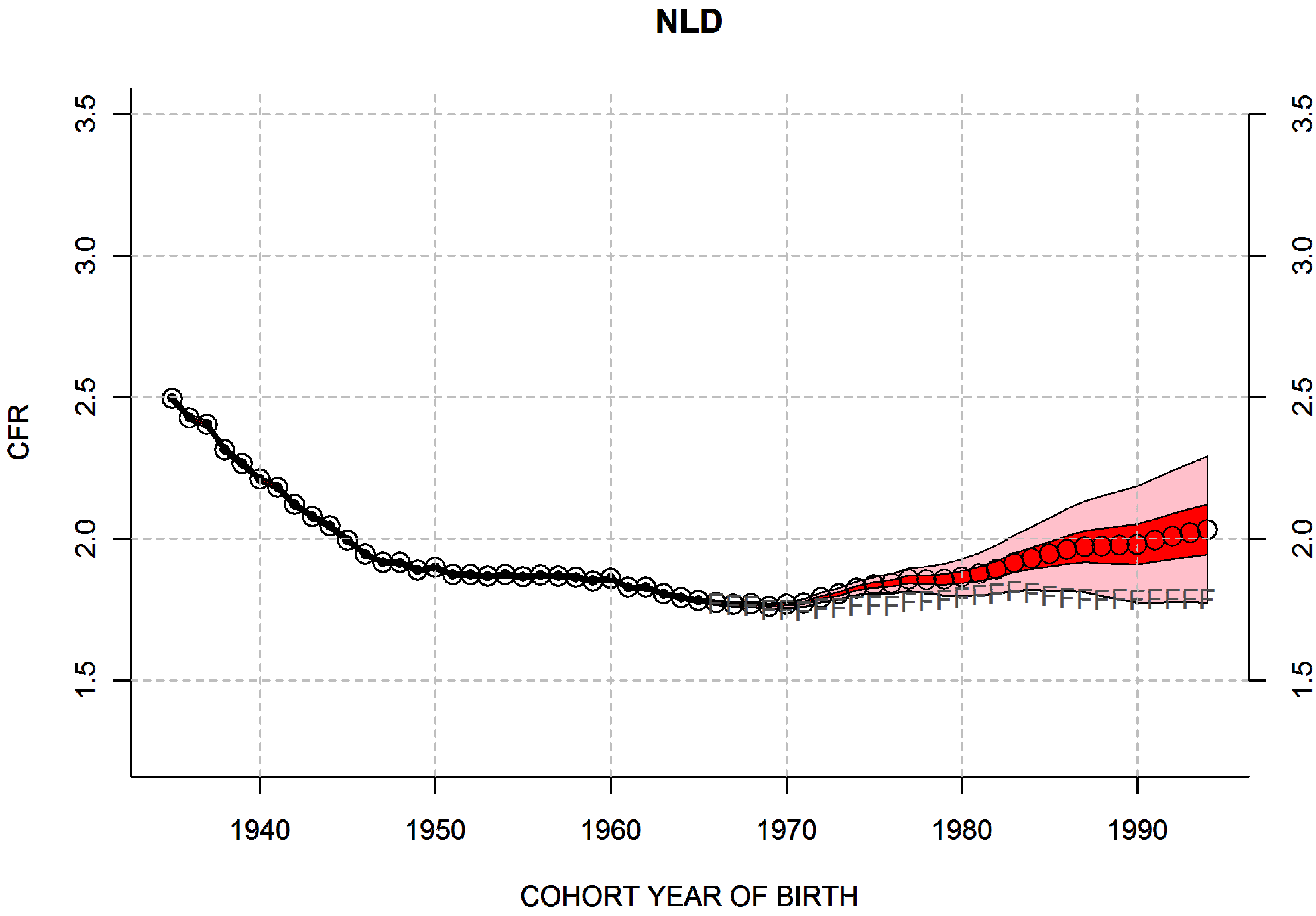

FRA
CFR
3.5
3.0
2.5
2.0
1.5
1930
1940
1950
1960
1970
1980
1990
COHORT YEAR OF BIRTH

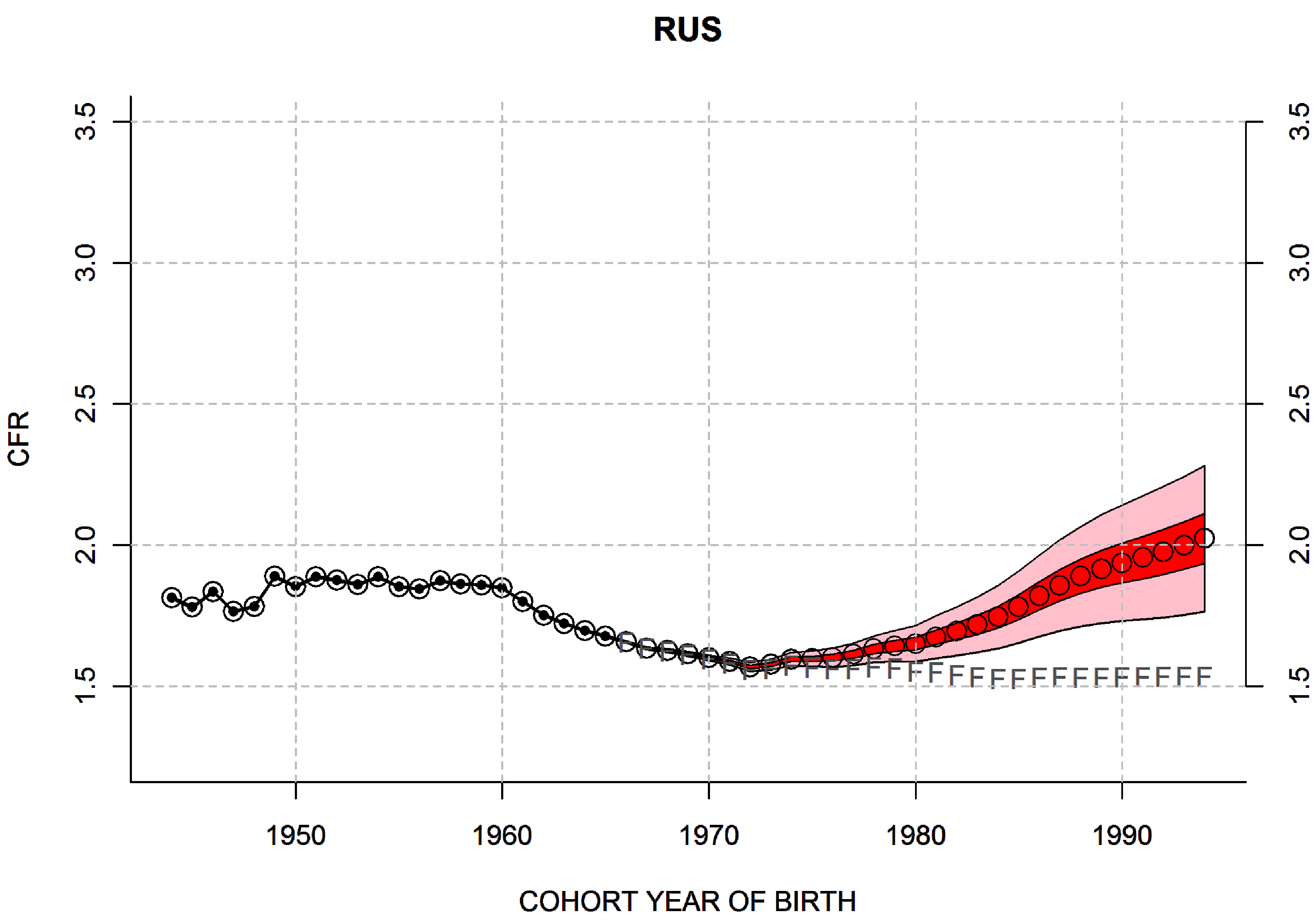

GBRTENW
CFR
COHORT YEAR OF BIRTH
1930
1940
1950
1960
1970
1980
1990
1.5
2.0
2.5
3.0
3.5

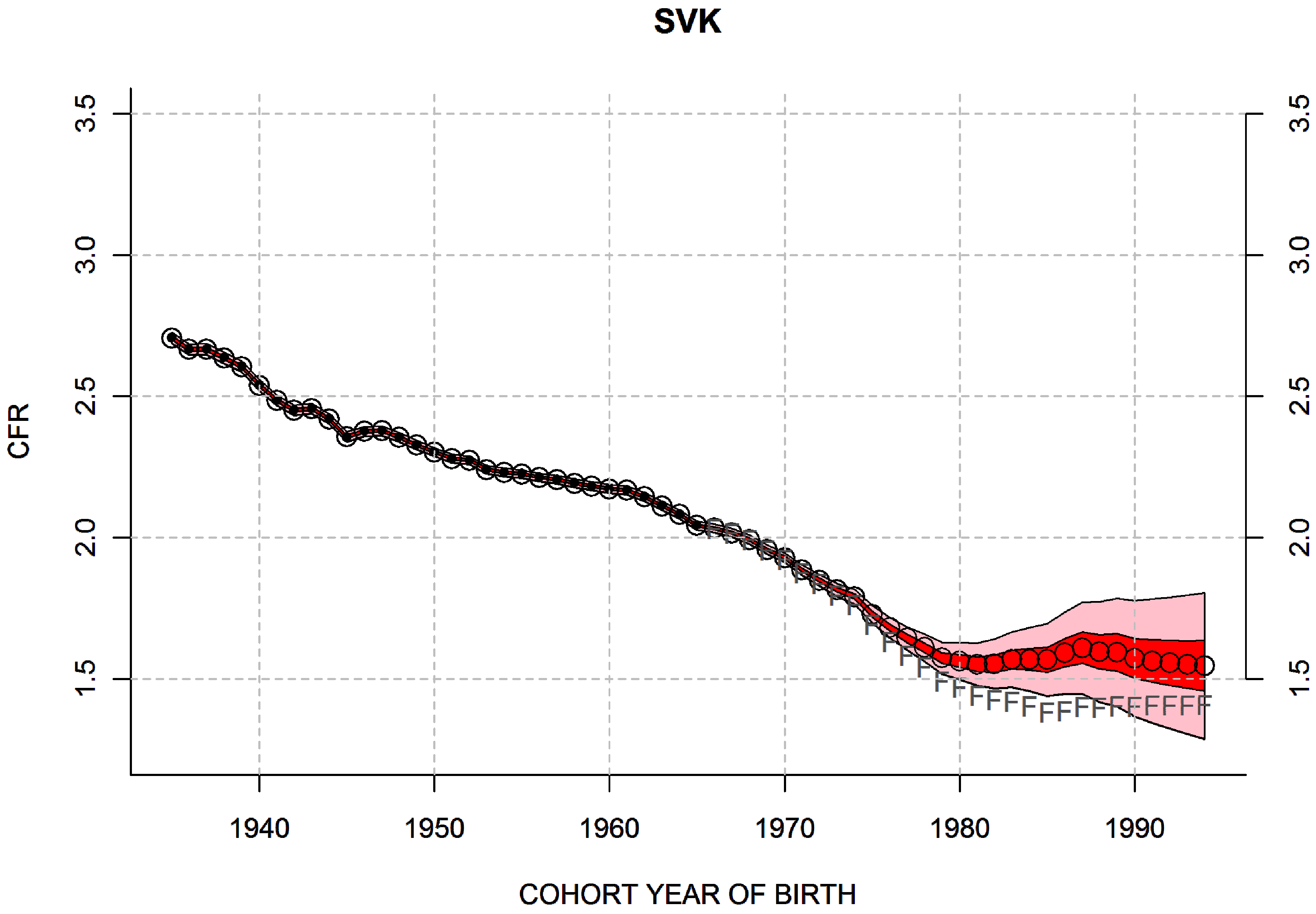


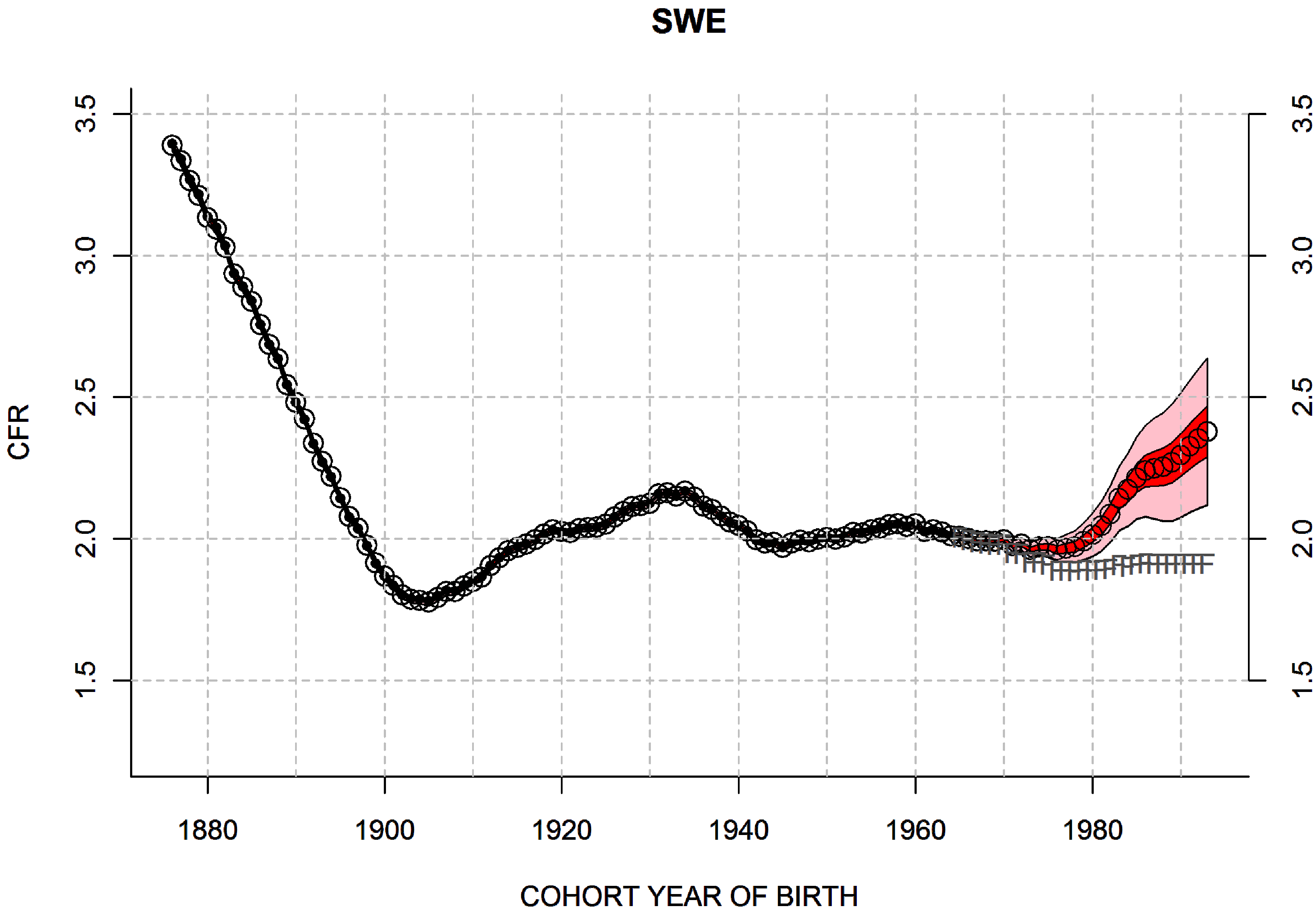
GBR_SCO
CFR
COHORT YEAR OF BIRTH
1930
1940
1950
1960
1970
1980
1990
1.5
2.0
2.5
3.0
3.5

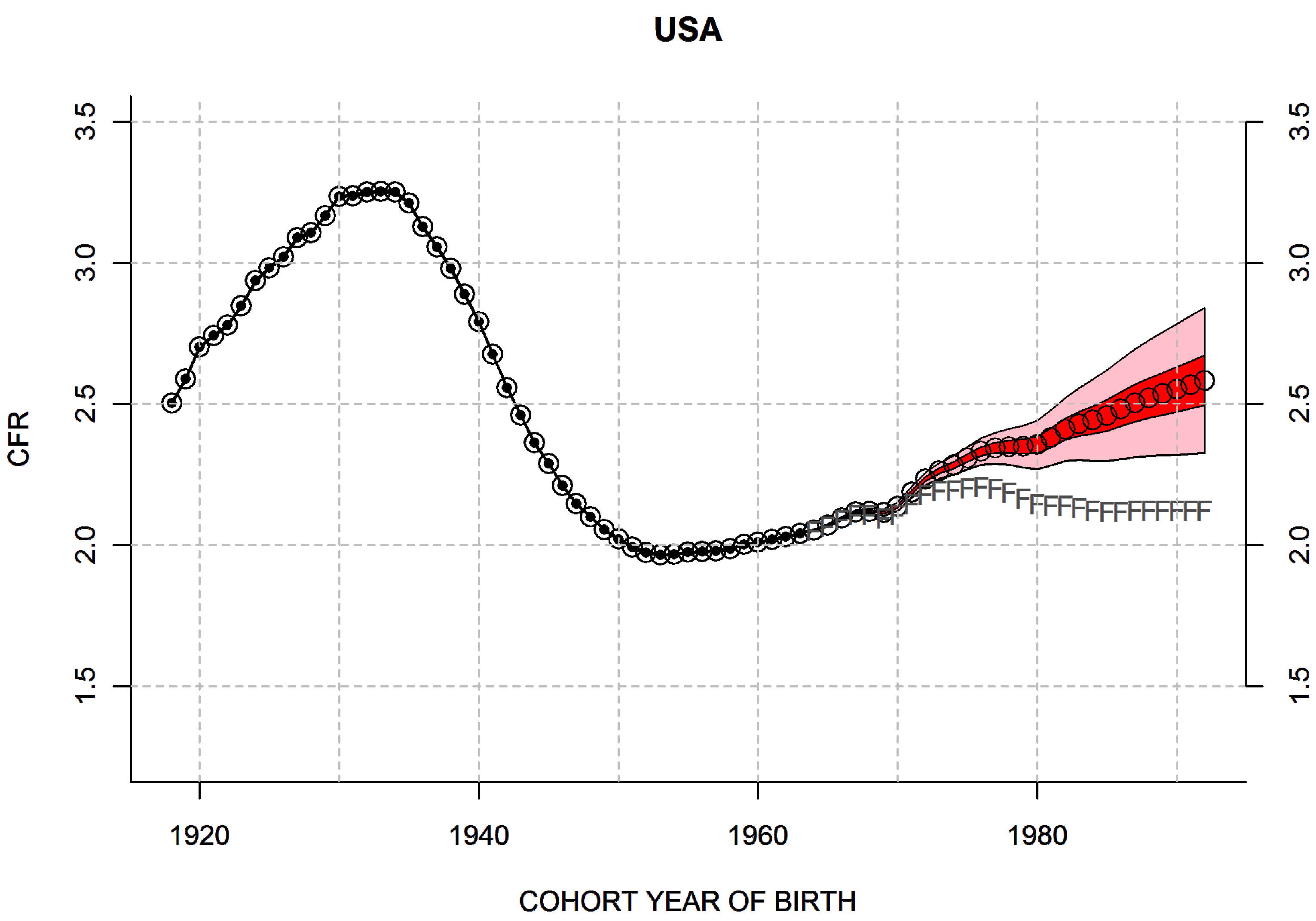

HUN
CFR
3.5
3.0
2.5
2.0
1.5
1940
1950
1960
1970
1980
1990
COHORT YEAR OF BIRTH


NLD
CFR
COHORT YEAR OF BIRTH
1940
1950
1960
1970
1980
1990
1.5
2.0
2.5
3.0
3.5

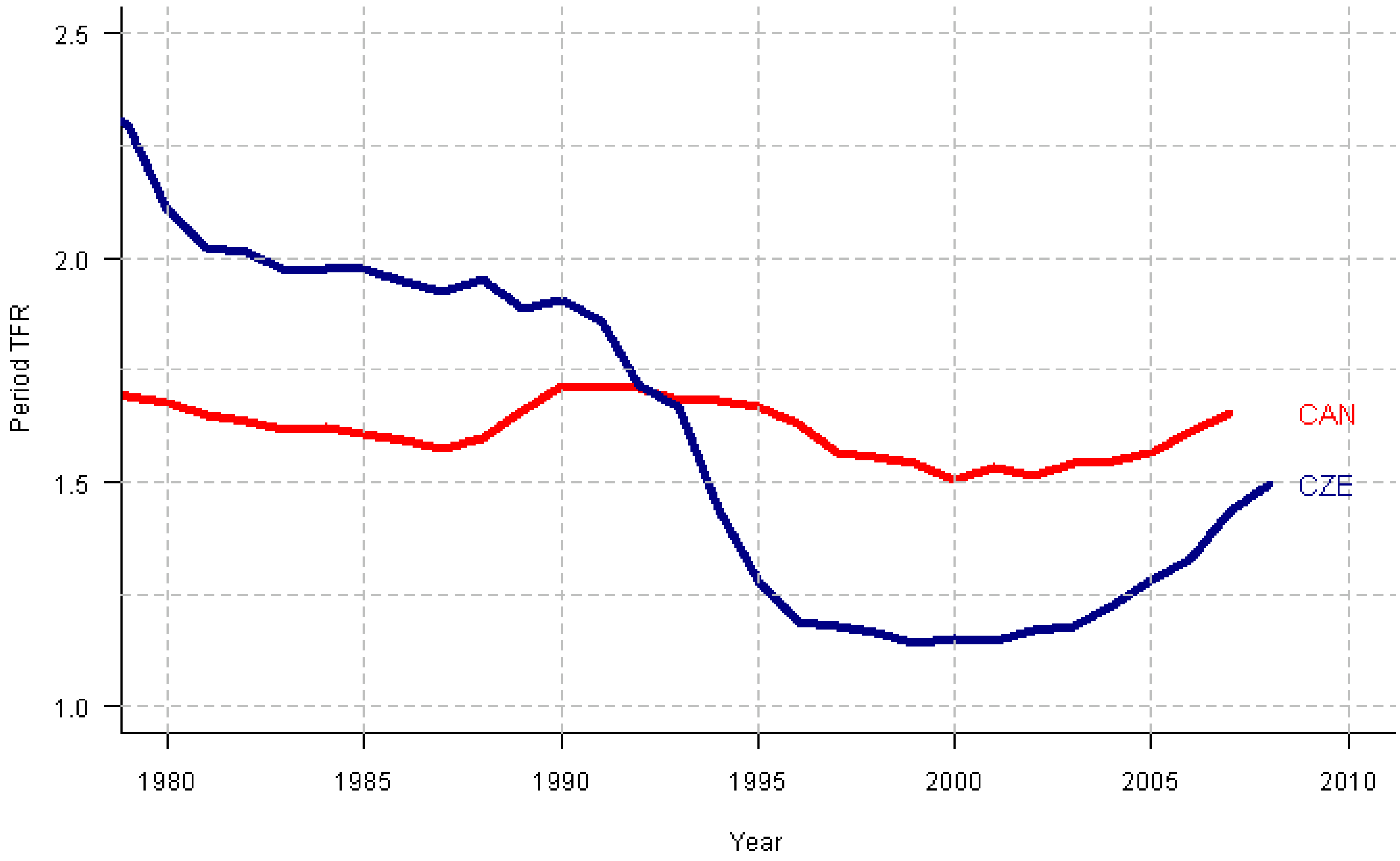

RUS
CFR
3.5
3.0
2.5
2.0
1.5
1950
1960
1970
1980
1990
COHORT YEAR OF BIRTH


SVK
CFR
COHORT YEAR OF BIRTH
1.5
2.0
2.5
3.0
3.5
1940
1950
1960
1970
1980
1990

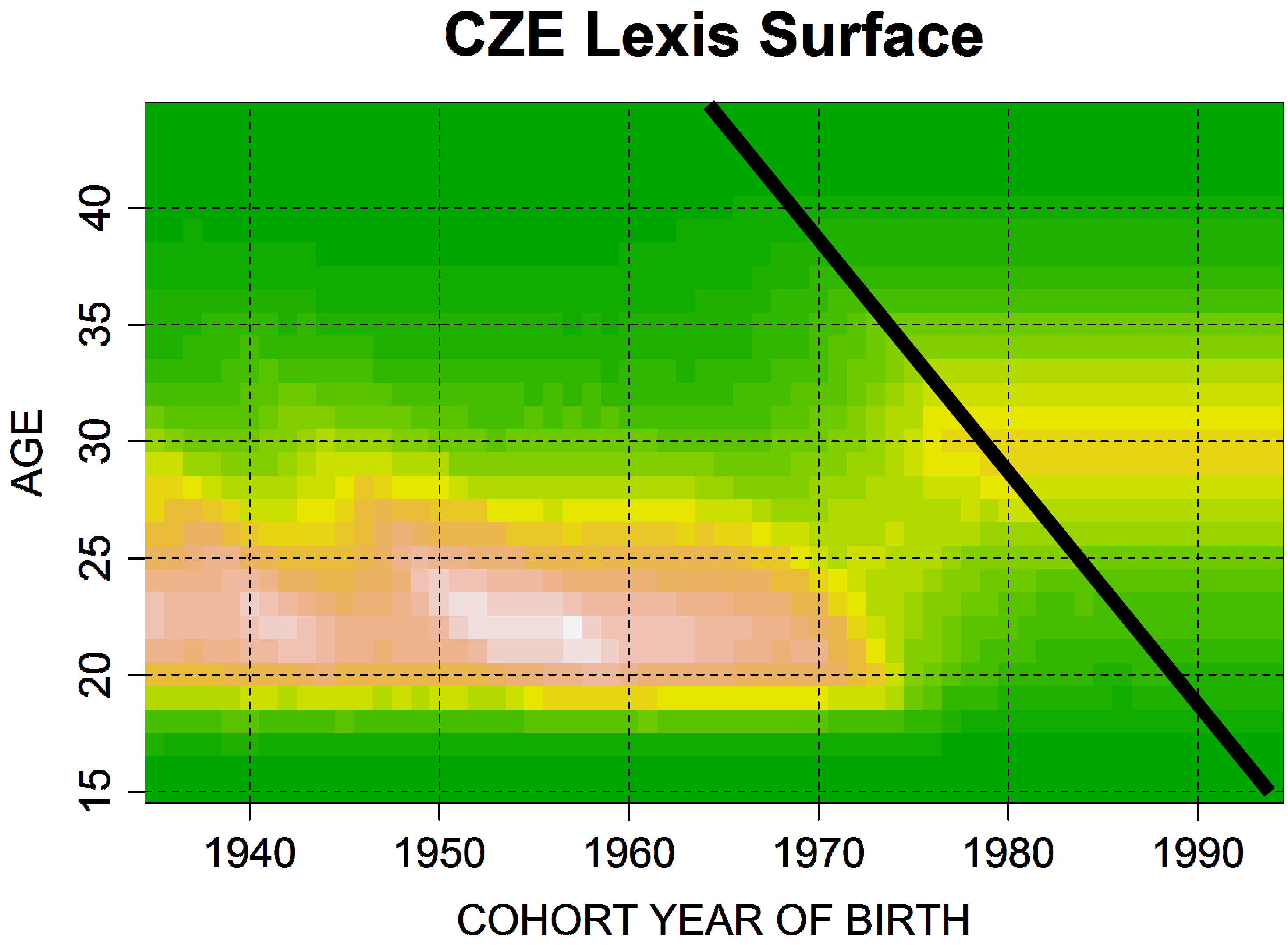

SWE
CFR
3.5
3.0
2.5
2.0
1.5
1880
1900
1920
1940
1960
1980
COHORT YEAR OF BIRTH

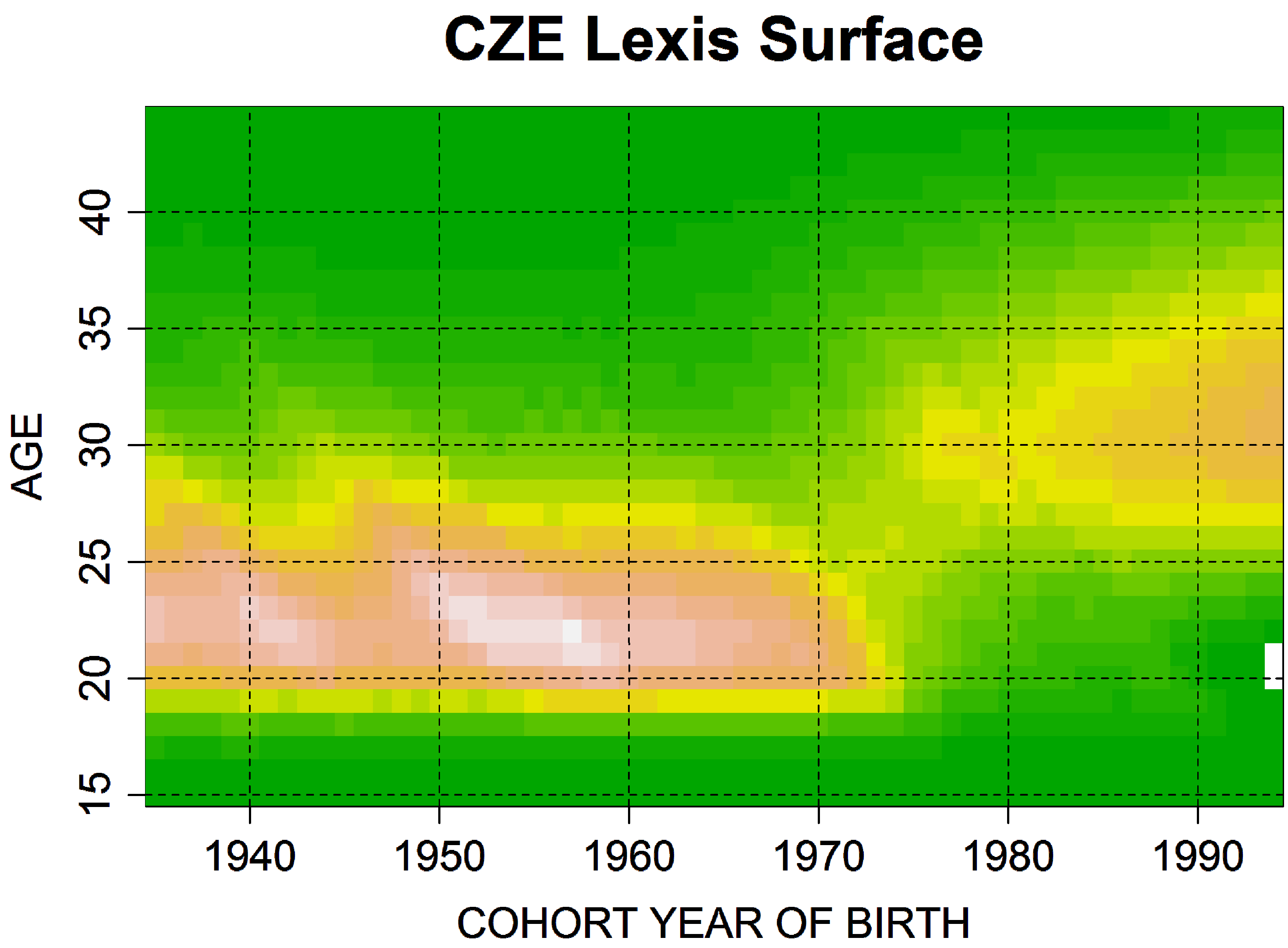

USA
CFR
COHORT YEAR OF BIRTH
1920
1940
1960
1980
1.5
2.0
2.5
3.0
3.5

Period TFR 1980-2008
Canada and Czech Rep



Freeze Rate Surface




CZE Lexis Surface
AGE
15
20
25
30
35
40
1940
1950
1960
1970
1980
1990
COHORT YEAR OF BIRTH

Forecast Methods

- Extrapolate time series f(a,t)
  - Freeze Rates: no change in future

    $$f(a, \text{NOW}+h) = f(a, \text{NOW})$$

  - Freeze Slopes: change continues linearly

    $$f(a, \text{NOW}+h) = f(a, \text{NOW}) + h * \text{slope}_{\text{NOW}}$$

Forecast Methods

- Model/Extrapolate cohort schedules
  - Li & Zheng (2003)
    - SVD decomposition of complete cohort data
    - Cohort schedules modeled as
      $f_{coh} = (\text{mean vec}) + k_{coh} * (1^{st}\ \text{princ. comp.})$
    - Estimate each k_{coh} from partial cohort history
  - Myrskylä & Goldstein (2010)
    - Parametric models for cohort schedules

Forecast Methods

ADVANTAGES for CFR estimation:

Freeze Rate	Avoids "nonsensical" forecasts
Freeze Slope	Utilizes recent trends
Cohort Model	Focuses on correct dimension

DISADVANTAGES for CFR estimation:

Freeze Rate	Ignores recent trends
Freeze Slope	Possible "nonsensical" forecasts
Cohort Model	Treats cohorts as independent (does not 'borrow strength' across demographic dimensions)