Fertility estimates from full birth histories and HDSS

BRUNO SCHOUMAKER

UCLOUVAIN - BELGIUM

Outline

- Data and methods
- Single year ASFRs
  - Confidence intervals, consistency across surveys
  - Smoothing
- Under-15 fertility
  - Confidence intervals
  - Consistency across surveys and reconstructing trends

Data and methods

104 countries, 424 surveys, 10 HDSS

- World Fertility Surveys (WFS)
- Standard Demographic and Health Surveys (DHS)
- Aids Indicators Surveys (AIS), Malaria Indicators Surveys (MIS), Interim DHS surveys, Special DHS surveys
- Multiple Indicators Cluster Surveys (MICS)
- Reproductive Health Surveys (RHS)
- PAPFAM and PAPCHILD
- Selected HDSS data (Senegal, Burkina Faso, Ethiopia, Malawi, Bangladesh) – more will be included

Surveys by program

Survey type	Total number of surveys	Number of surveys processed	Harmonization
Standard DHS	297	261	Not needed
MIS	30	27	Not needed
AIS	11	5	Variable names
Other DHS	27	12	Variable names, file structure
WFS	43	35	Variable names, file structure, all-women factors
MICS	49	37	Variable names, file structure
PAPFAM/ PAPCHILD	32	23	Variable names, all-women factors
RHS	41	24	Variable names, file structure, date imputations
TOTAL	530	424	

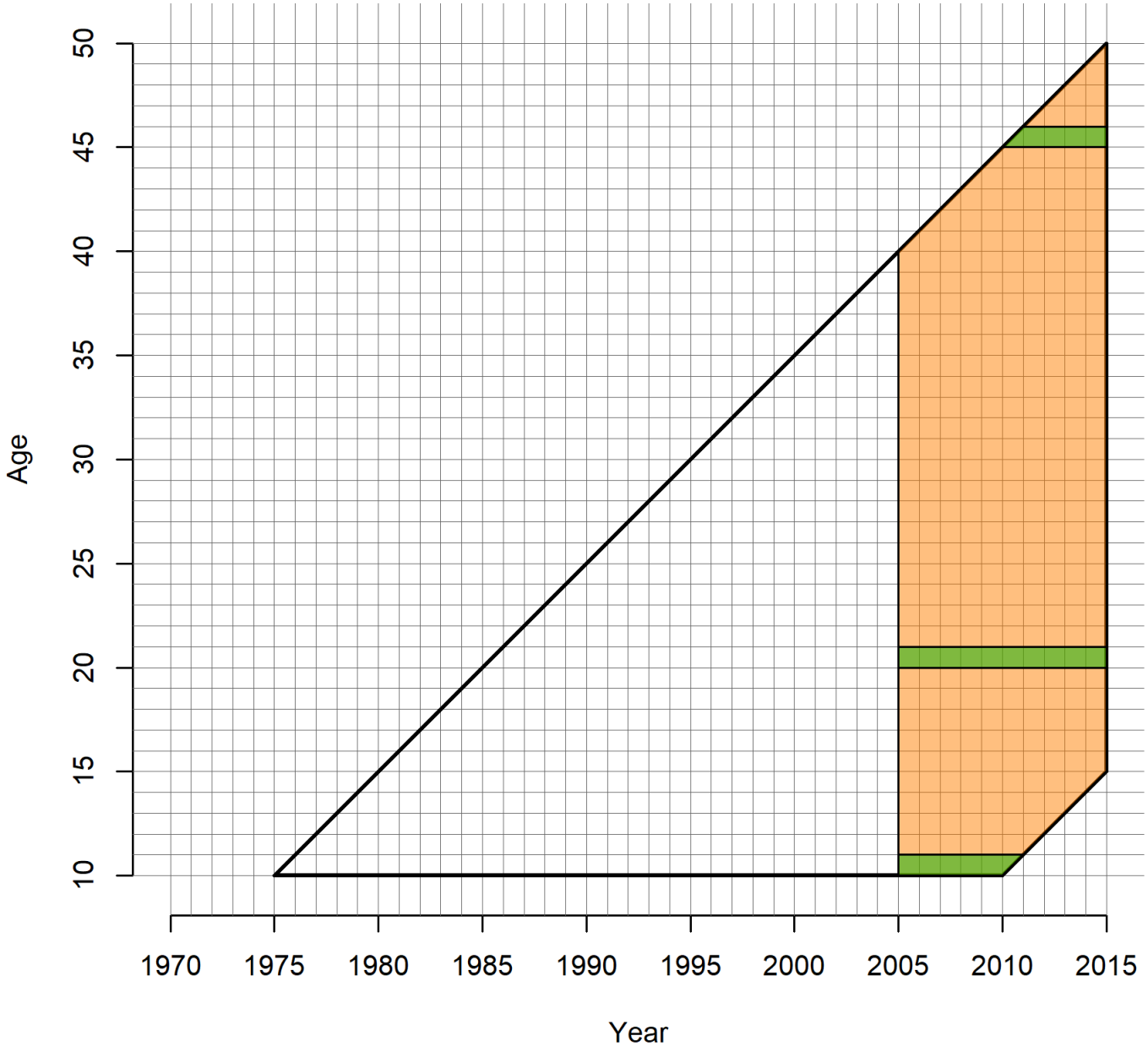
Methods

- Data harmonization, computation of all-women factors if necessary, imputation of dates if necessary
- Computation of births and exposure by age and year (or period), fertility rates, and standard errors
- Under-15 fertility : DHS « Lexis approach » (Pullum et al. 2018)
  - Use of "factors *f* that will inflate the observed exposure *e* according to the geometry of the Lexis" (Pullum et al., 2018, p.10)
- Several *Stata* commands developed
  - adof11 : computation of under-15 fertility from birth histories
  - singa11: computation of single year ASFR from birth histories
  - tfr2 and itfr : computation of ASFR and TFRs from surveys and HDSS
- Two-stage sample design taken into account

Single year ASFRs

FROM SURVEY DATA

Single year ASFRs – 10 years preceding the survey



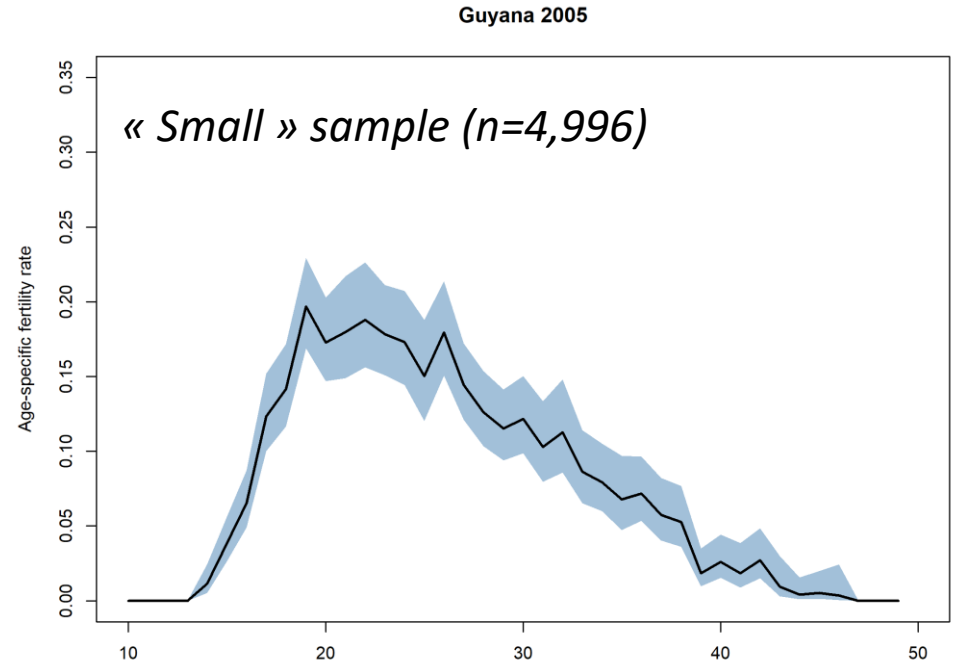
Database

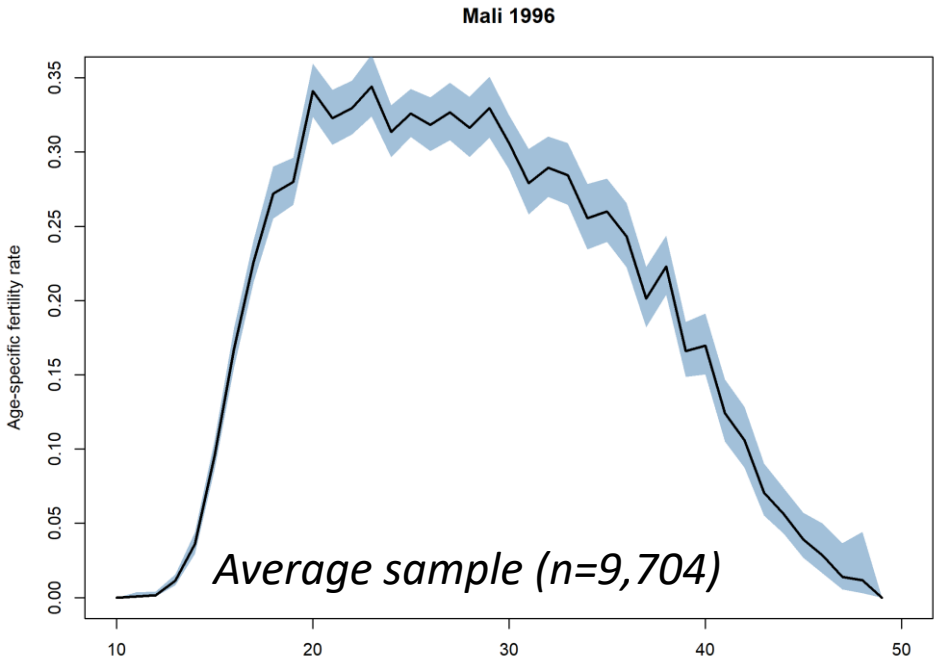
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	LocID	ISO3	ISO2	LocName	RegName	DataCatalo	CatalogNa	CatalogSh	datafile	ageg	per	window	centry	rate	lower	upper	se_lograte	events	exposure	
2	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	10	0-9	10	2006	0.0000	0.0000	0.0000	0.1058	0	21979	
3	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	11	0-9	10	2006	0.0000	0.0000	0.0000	0.1031	0	23709	
4	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	12	0-9	10	2006	0.0015	0.0009	0.0024	0.2530	34	22979	
5	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	13	0-9	10	2006	0.0053	0.0038	0.0073	0.1652	113	21560	
6	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	14	0-9	10	2006	0.0158	0.0124	0.0203	0.1253	317	19988	
7	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	15	0-9	10	2006	0.0370	0.0318	0.0432	0.0781	714	19277	
8	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	16	0-9	10	2006	0.0721	0.0647	0.0803	0.0552	1297	17986	
9	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	17	0-9	10	2006	0.1155	0.1052	0.1268	0.0477	1939	16786	
10	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	18	0-9	10	2006	0.1581	0.1490	0.1677	0.0301	2427	15354	
11	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	19	0-9	10	2006	0.2150	0.2050	0.2254	0.0241	3042	14153	
12	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	20	0-9	10	2006	0.2630	0.2510	0.2755	0.0237	3447	13109	
13	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	21	0-9	10	2006	0.2908	0.2791	0.3030	0.0210	3566	12264	
14	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	22	0-9	10	2006	0.3284	0.3148	0.3425	0.0216	3757	11441	
15	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	23	0-9	10	2006	0.3154	0.3037	0.3277	0.0194	3400	10780	
16	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	24	0-9	10	2006	0.3260	0.3114	0.3413	0.0234	3356	10296	
17	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	25	0-9	10	2006	0.3222	0.3103	0.3346	0.0192	3135	9731	
18	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	26	0-9	10	2006	0.3298	0.3147	0.3456	0.0239	3059	9275	
19	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	27	0-9	10	2006	0.3224	0.3068	0.3389	0.0254	2884	8944	
20	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	28	0-9	10	2006	0.2929	0.2787	0.3078	0.0254	2500	8537	
21	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	29	0-9	10	2006	0.3074	0.2928	0.3228	0.0249	2565	8342	
22	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	30	0-9	10	2006	0.2923	0.2780	0.3074	0.0256	2364	8086	
23	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	31	0-9	10	2006	0.2798	0.2675	0.2927	0.0229	2222	7942	
24	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	32	0-9	10	2006	0.2642	0.2508	0.2784	0.0266	2047	7747	
25	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	33	0-9	10	2006	0.2473	0.2356	0.2595	0.0246	1867	7552	
26	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	34	0-9	10	2006	0.2187	0.2072	0.2309	0.0277	1616	7387	
27	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	35	0-9	10	2006	0.2074	0.1951	0.2204	0.0311	1449	6987	
28	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	36	0-9	10	2006	0.1944	0.1830	0.2065	0.0309	1294	6658	

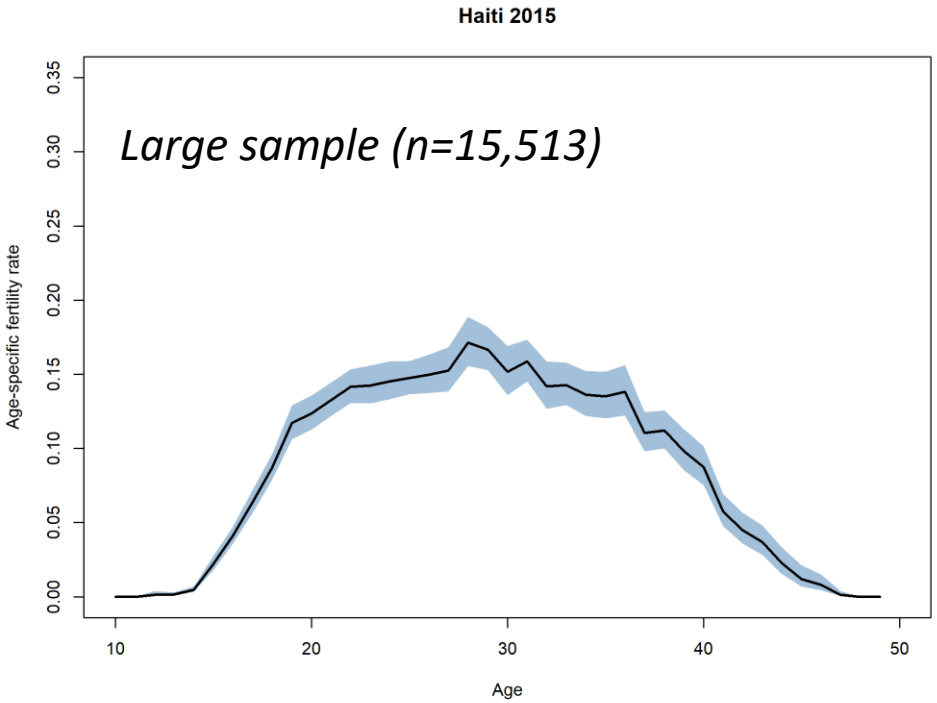
Sampling errors

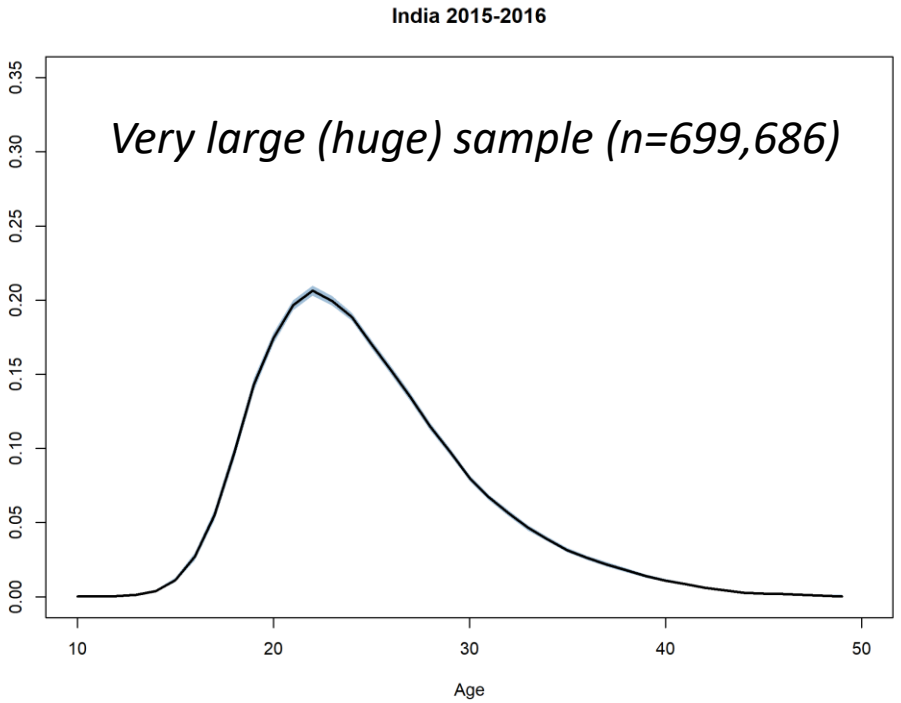
95% CONFIDENCE INTERVALS

10-YEAR ESTIMATES

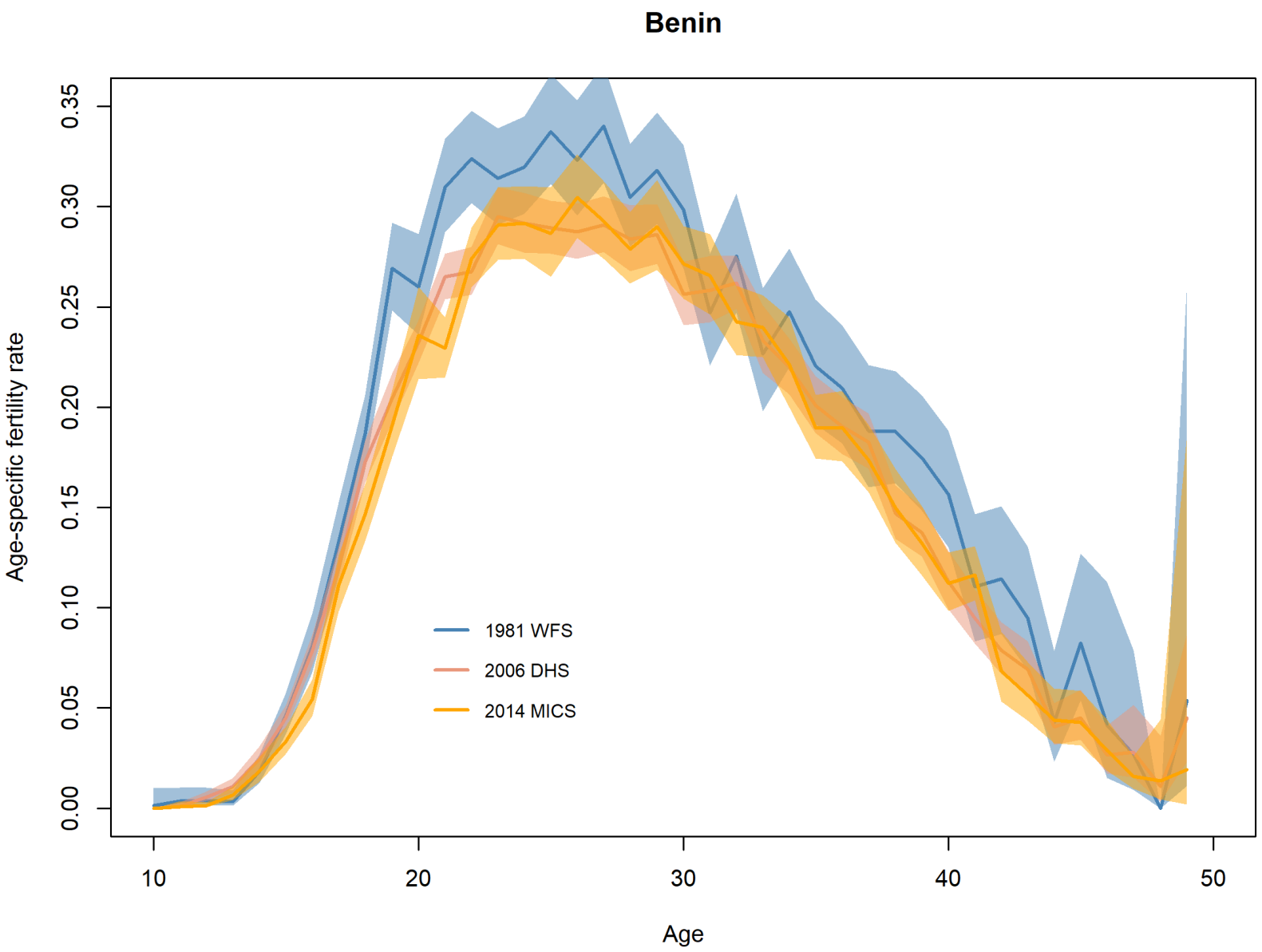

Guyana 2005
« Small » sample (n=4,996)
Age-specific fertility rate
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
10
20
30
40
50

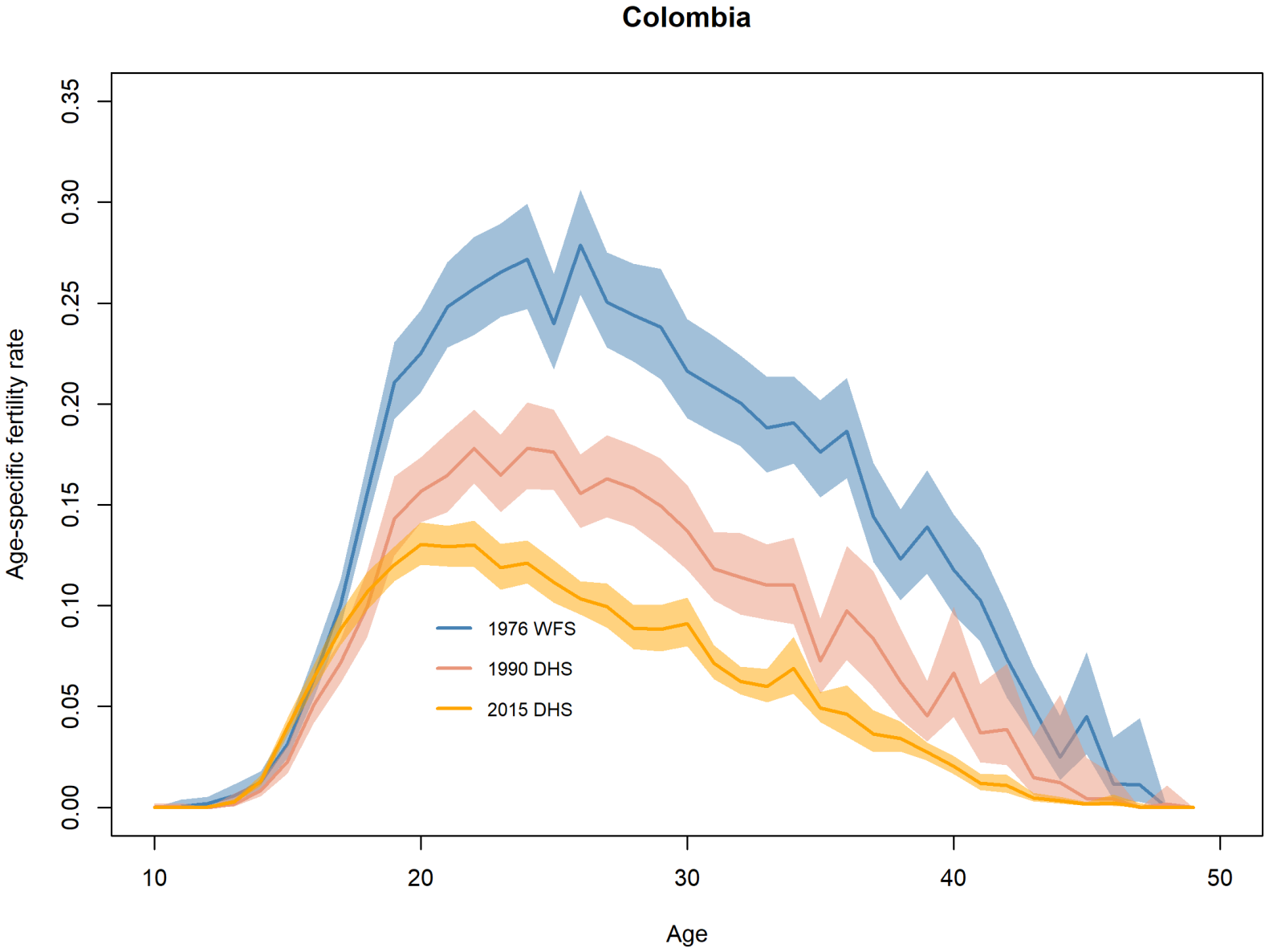

Mali 1996
Average sample (n=9,704)
Age-specific fertility rate
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
10
20
30
40
50

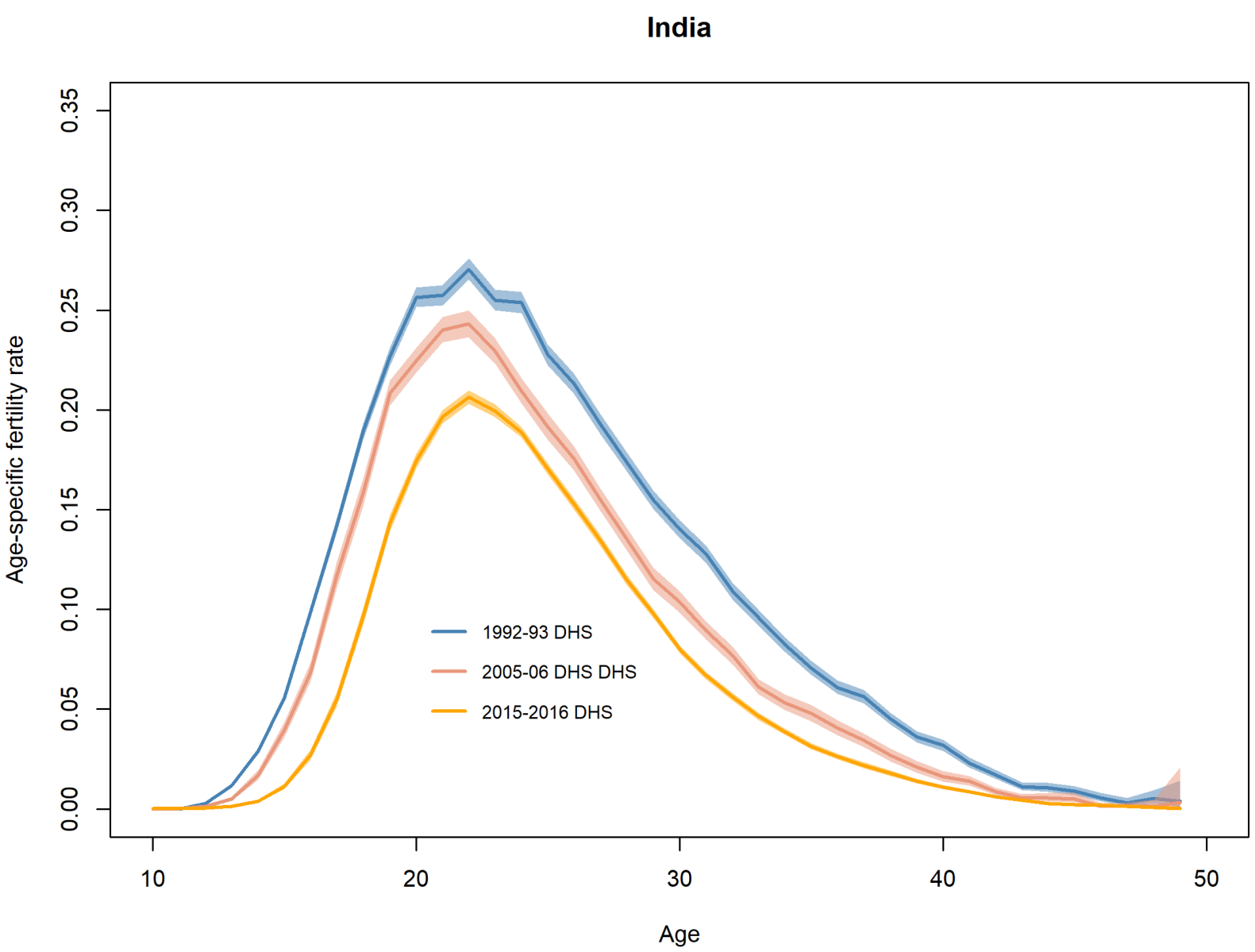

Haiti 2015
Large sample (n=15,513)
Age-specific fertility rate
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
10
20
30
40
50
Age

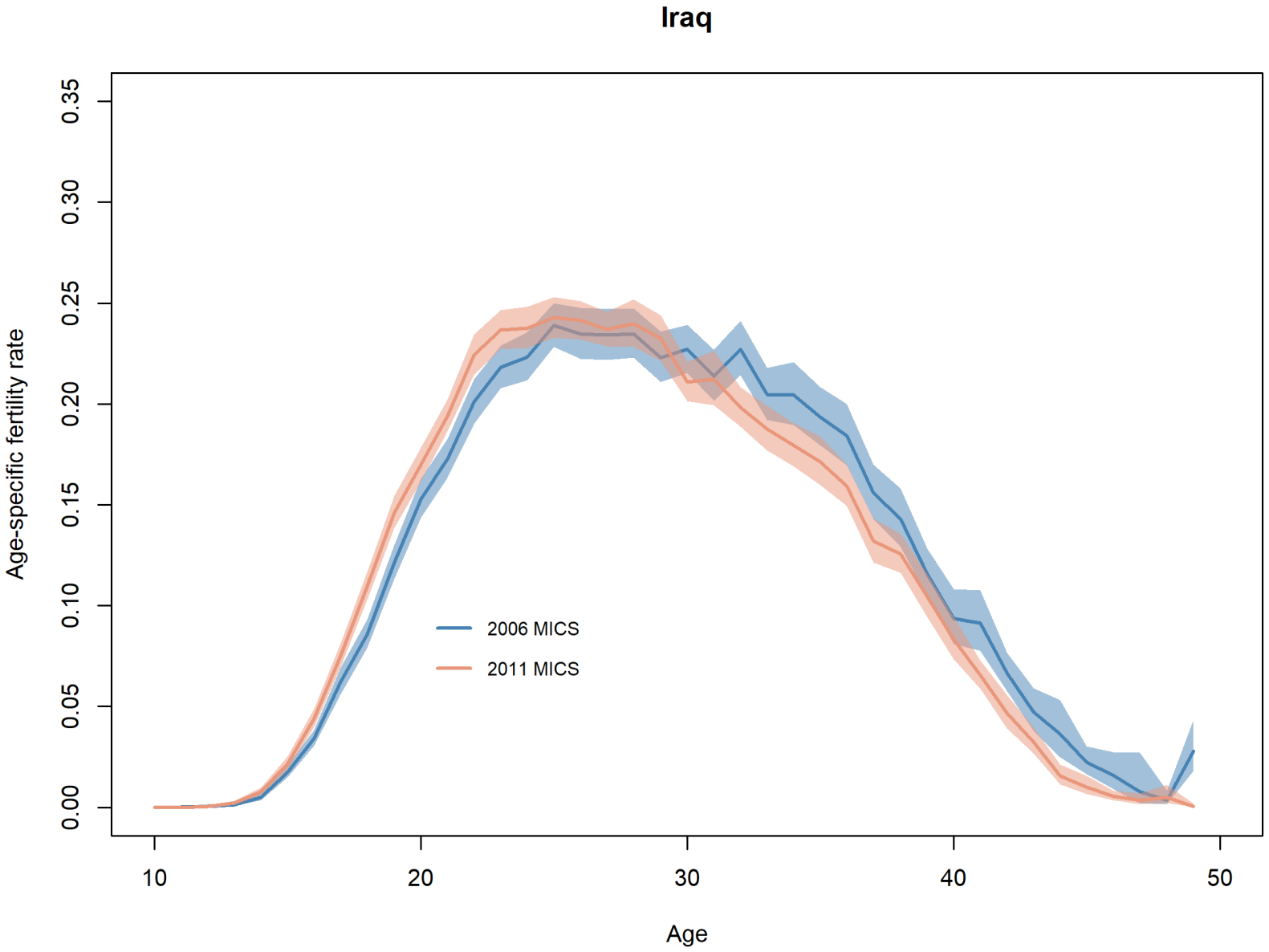

India 2015-2016
Very large (huge) sample (n=699,686)
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
10
20
30
40
50
Age

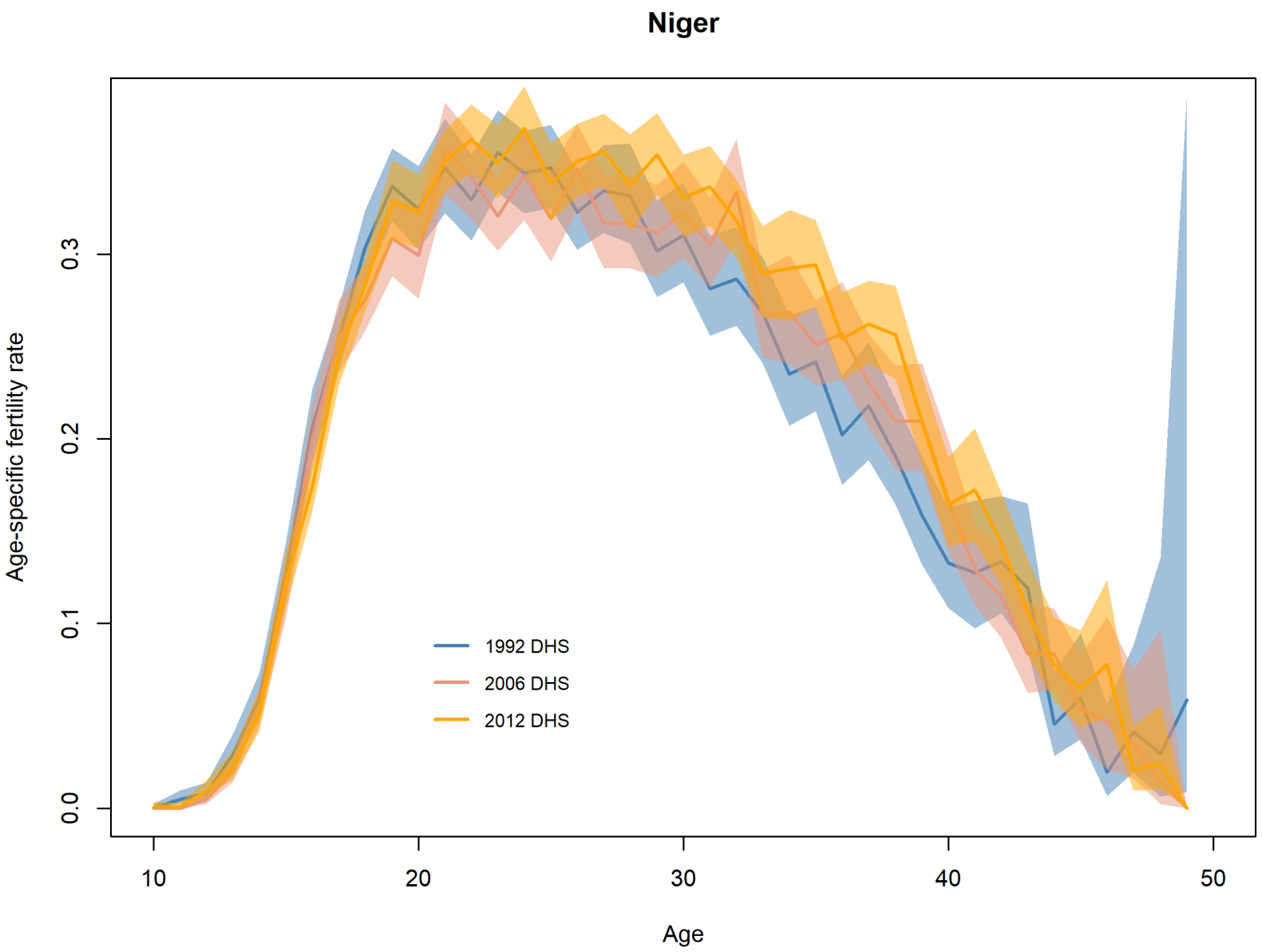
Consistency of age patterns within countries

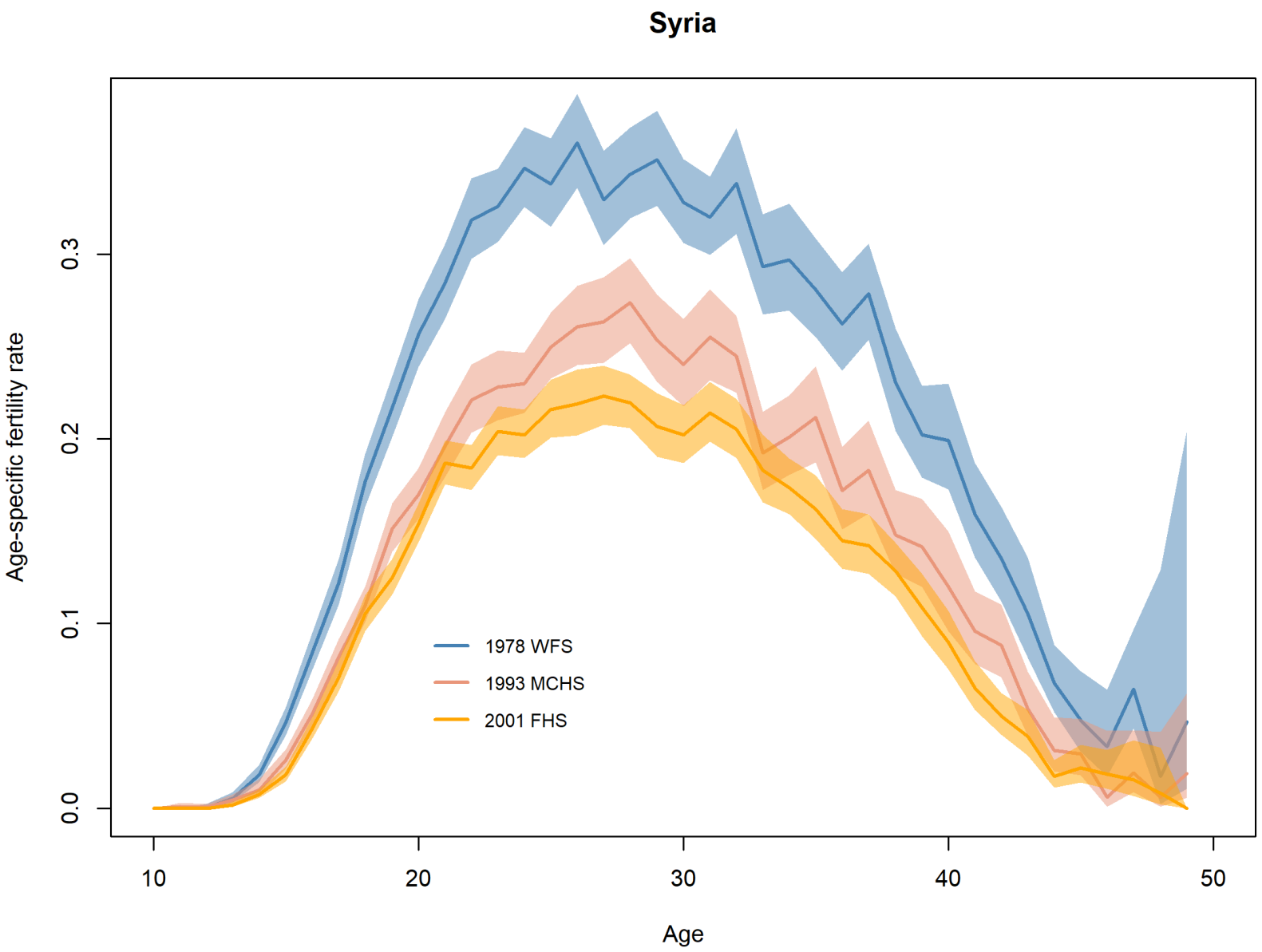

Benin
Age-specific fertility rate
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
10
20
30
40
50
Age
1981 WFS
2006 DHS
2014 MICS


Colombia
Age-specific fertility rate
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
10
20
30
40
50
Age
1976 WFS
1990 DHS
2015 DHS


India
Age-specific fertility rate
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
10
20
30
40
50
Age
1992-93 DHS
2005-06 DHS DHS
2015-2016 DHS

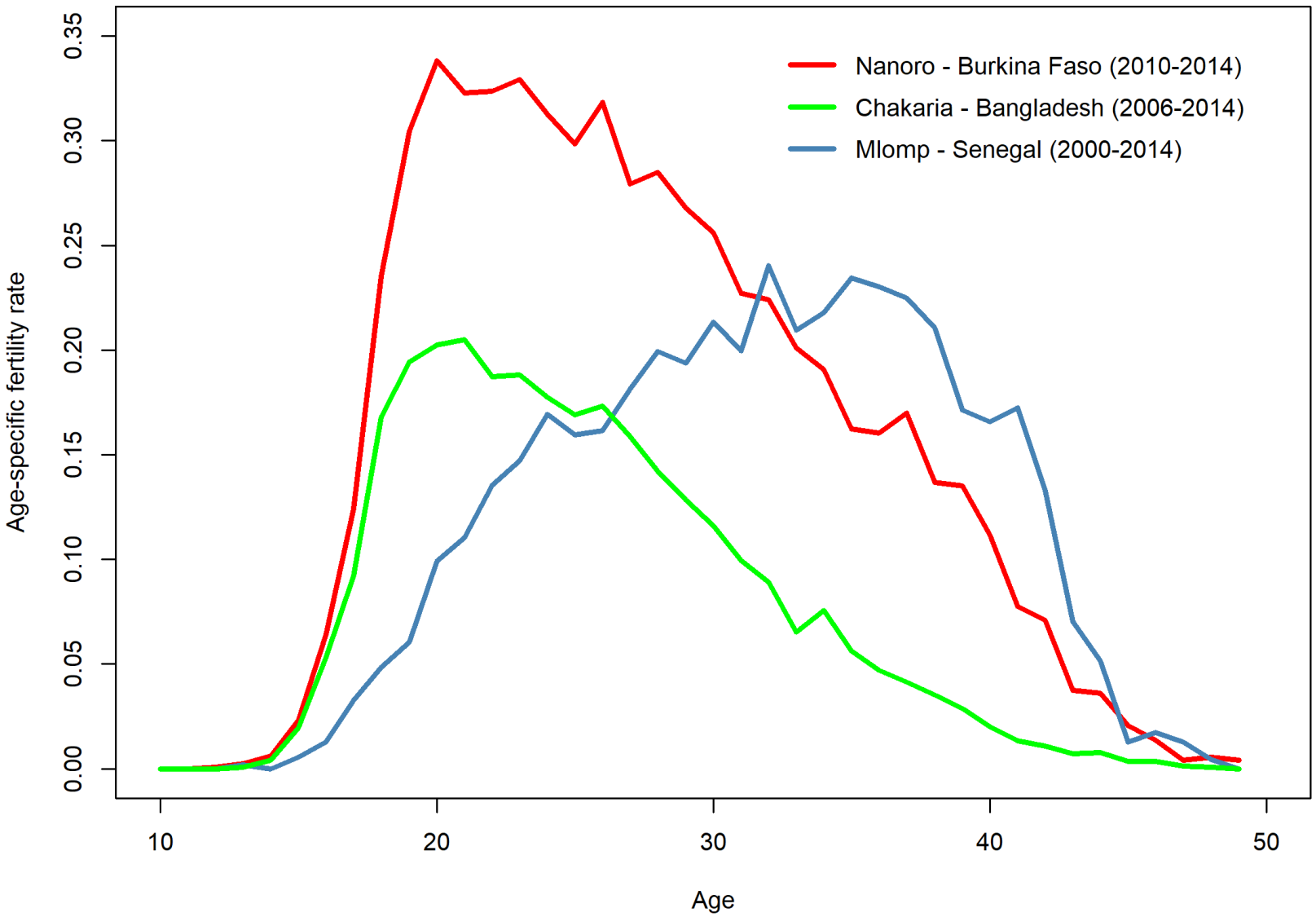

Iraq
Age-specific fertility rate
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
10
20
30
40
50
Age
2006 MICS
2011 MICS


Niger
Age-specific fertility rate
0.0
0.1
0.2
0.3
10
20
30
40
50
Age
1992 DHS
2006 DHS
2012 DHS


Syria
Age-specific fertility rate
0.0
0.1
0.2
0.3
10
20
30
40
50
Age
1978 WFS
1993 MCHS
2001 FHS

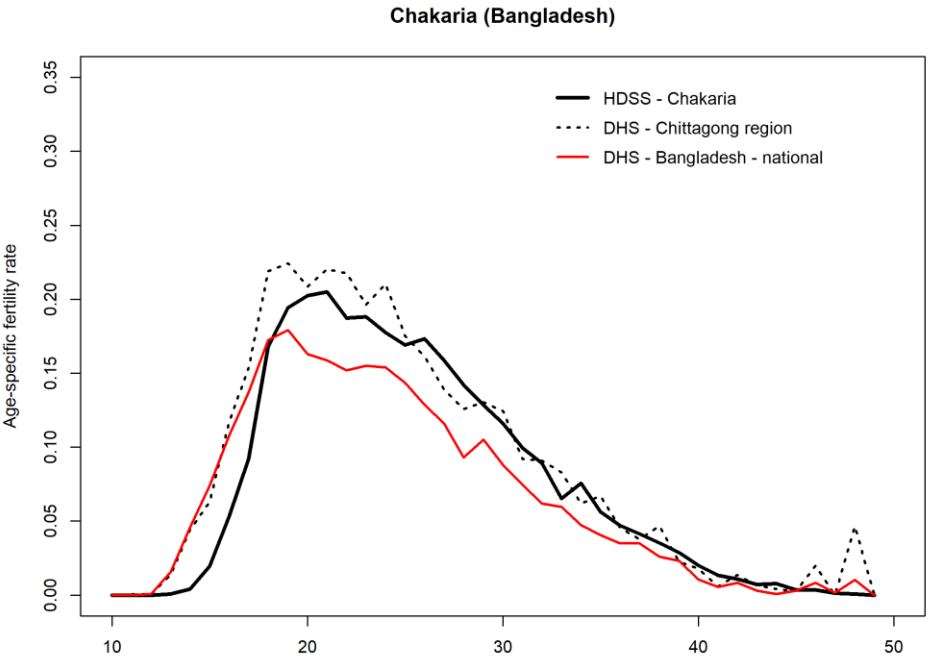
Single year ASFRs

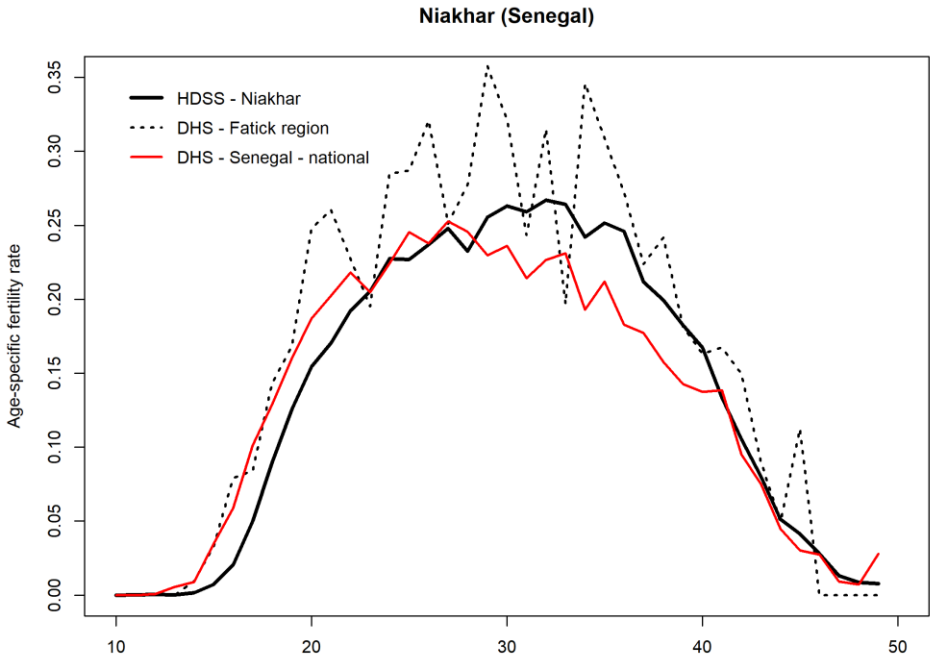
FROM SELECTED HDSS

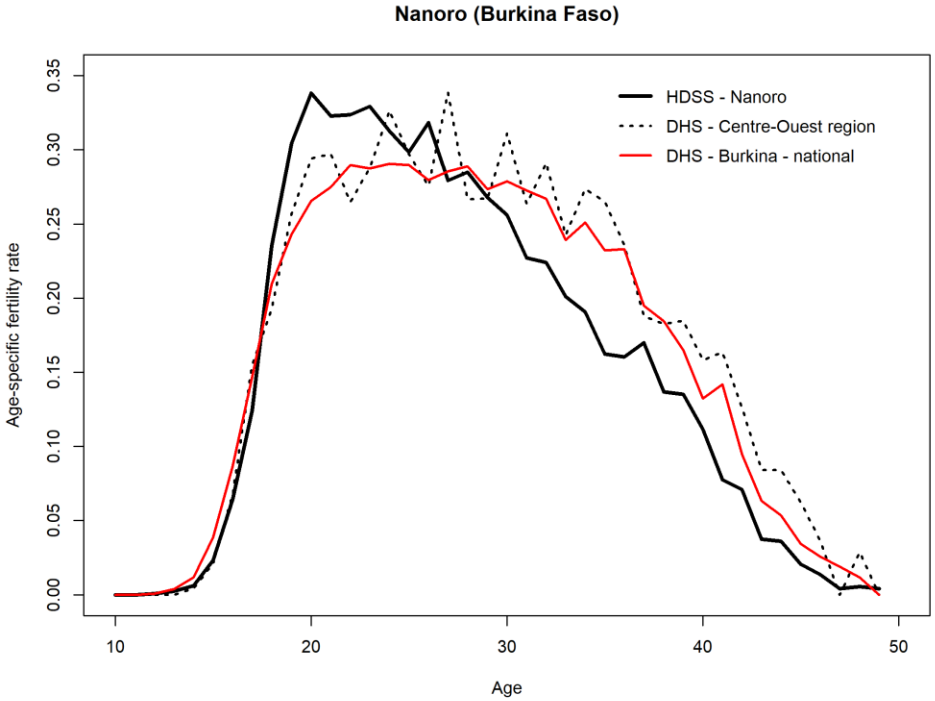

Selected HDSS
Nanoro - Burkina Faso (2010-2014)
Chakaria - Bangladesh (2006-2014)
Mlomp - Senegal (2000-2014)
Age-specific fertility rate
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
10
20
30
40
50
Age

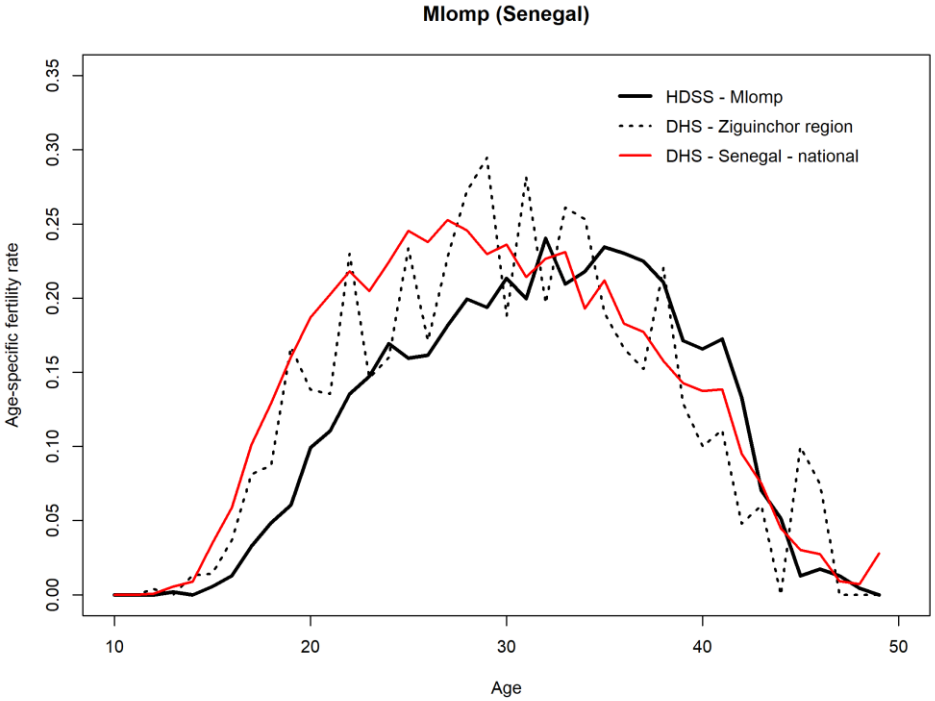
Comparisons HDSS – DHS

IN SELECTED HDSS, DHS REGION AND DHS COUNTRY


Chakaria (Bangladesh)
HDSS - Chakaria
DHS - Chittagong region
DHS - Bangladesh - national
Age-specific fertility rate
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
10
20
30
40
50


Niakhar (Senegal)
HDSS - Niakhar
DHS - Fatick region
DHS - Senegal - national
Age-specific fertility rate
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
10
20
30
40
50


Nanoro (Burkina Faso)
HDSS - Nanoro
DHS - Centre-Ouest region
DHS - Burkina - national
Age-specific fertility rate
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
10
20
30
40
50
Age

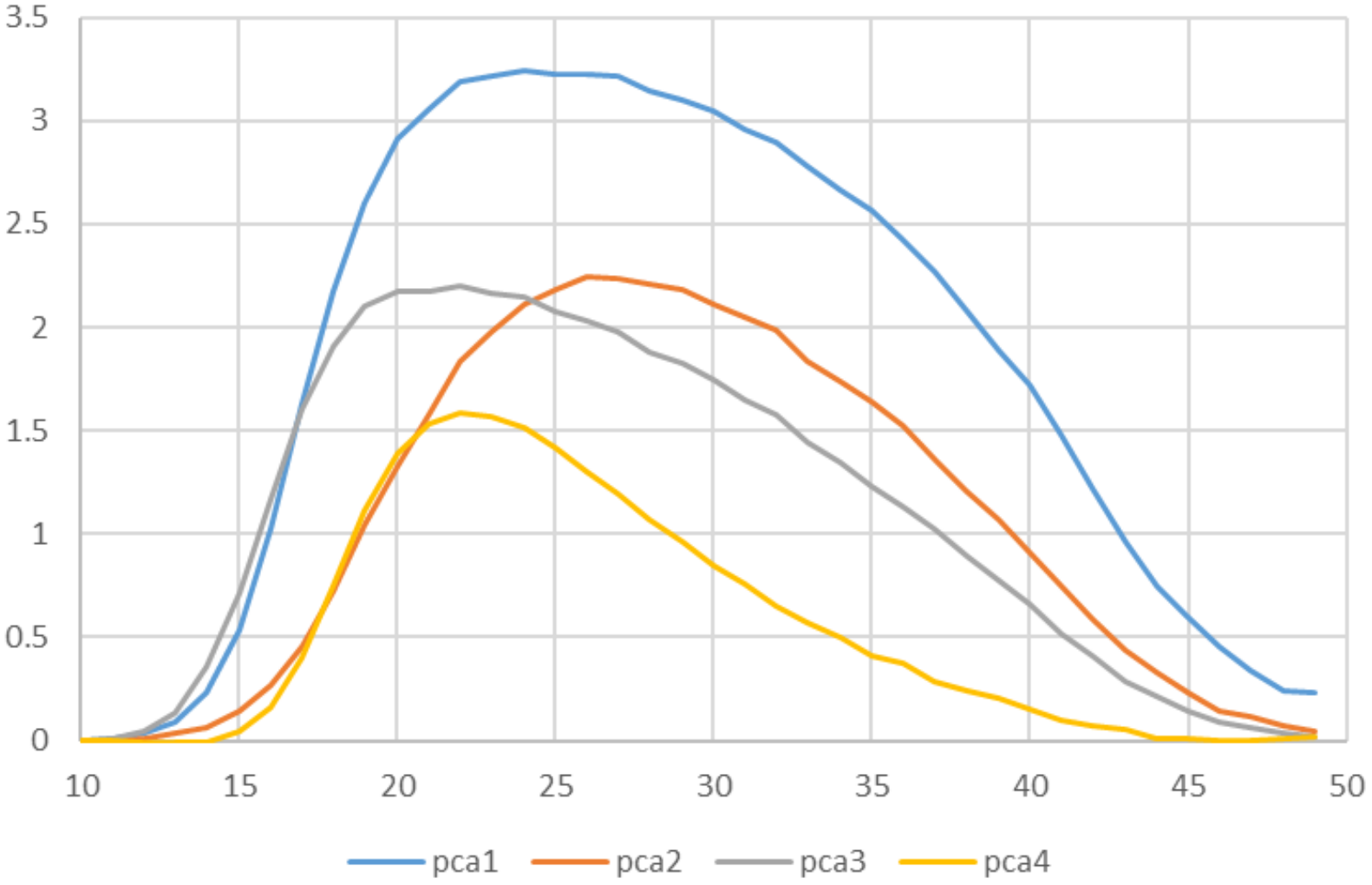

Mlomp (Senegal)
HDSS - Mlomp
DHS - Ziguinchor region
DHS - Senegal - national
Age-specific fertility rate
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
10
20
30
40
50
Age

Modelling and smoothing

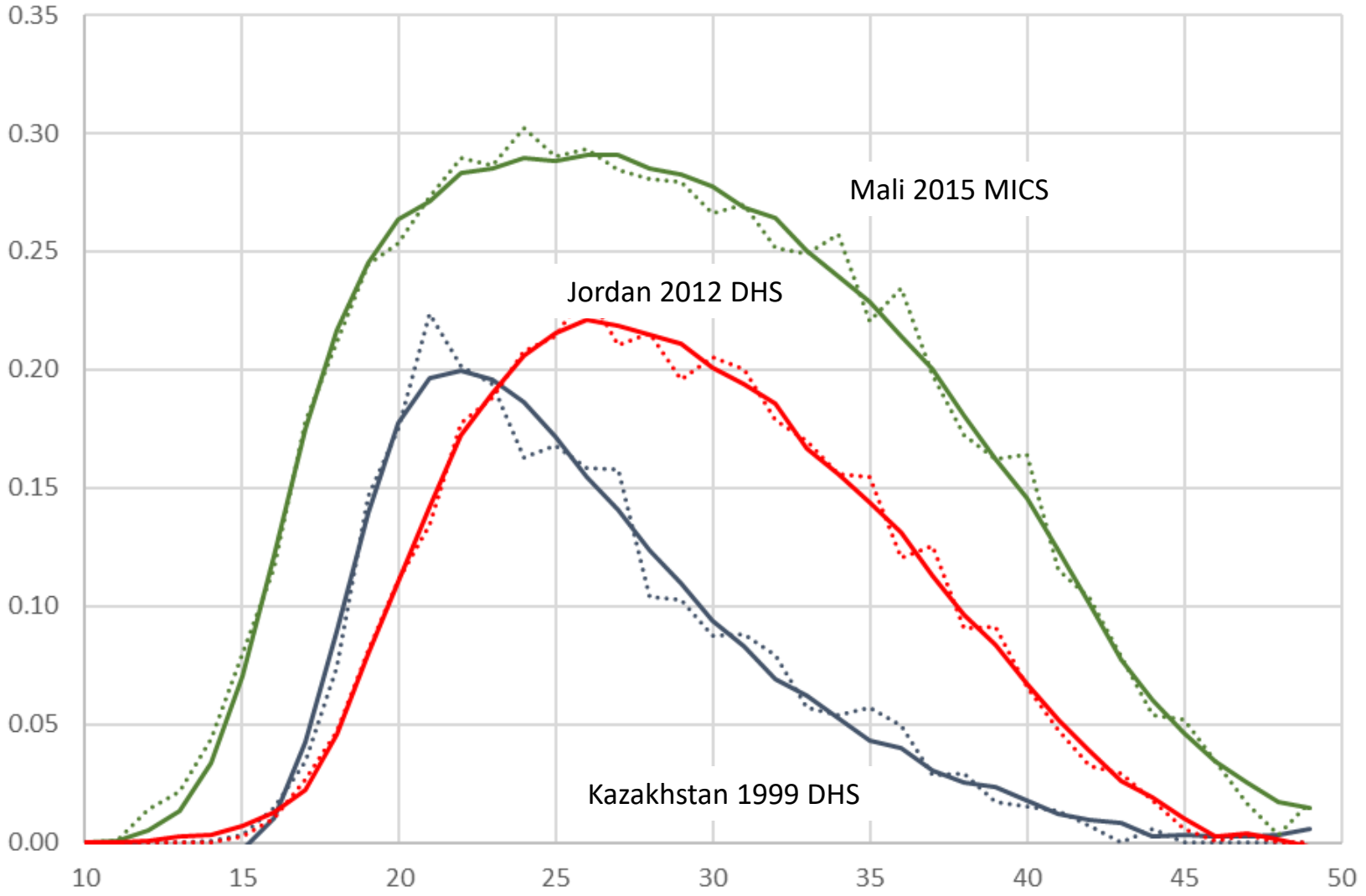
Principal component analysis of fertility rates

- Work very much in progress – inspired by Pantazis and Clark (2018)
- Around 400 series of single year age fertility rates
- 4 components represent 99% of the variance
- For each survey, weights for each of the 4 components
- Smoothed rates obtained as the linear combination of the 4 components and survey-specific weights

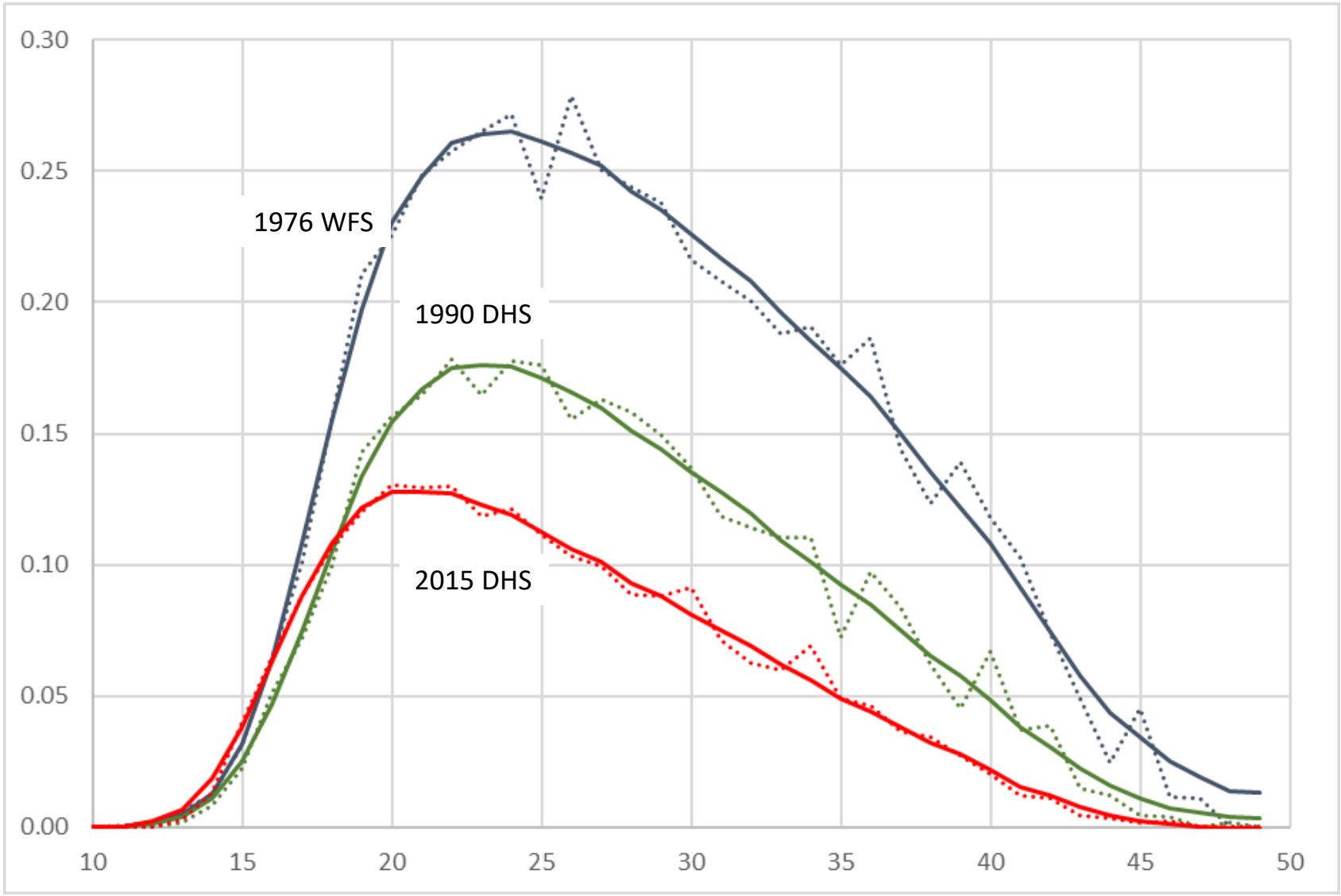
4 components (rotated)



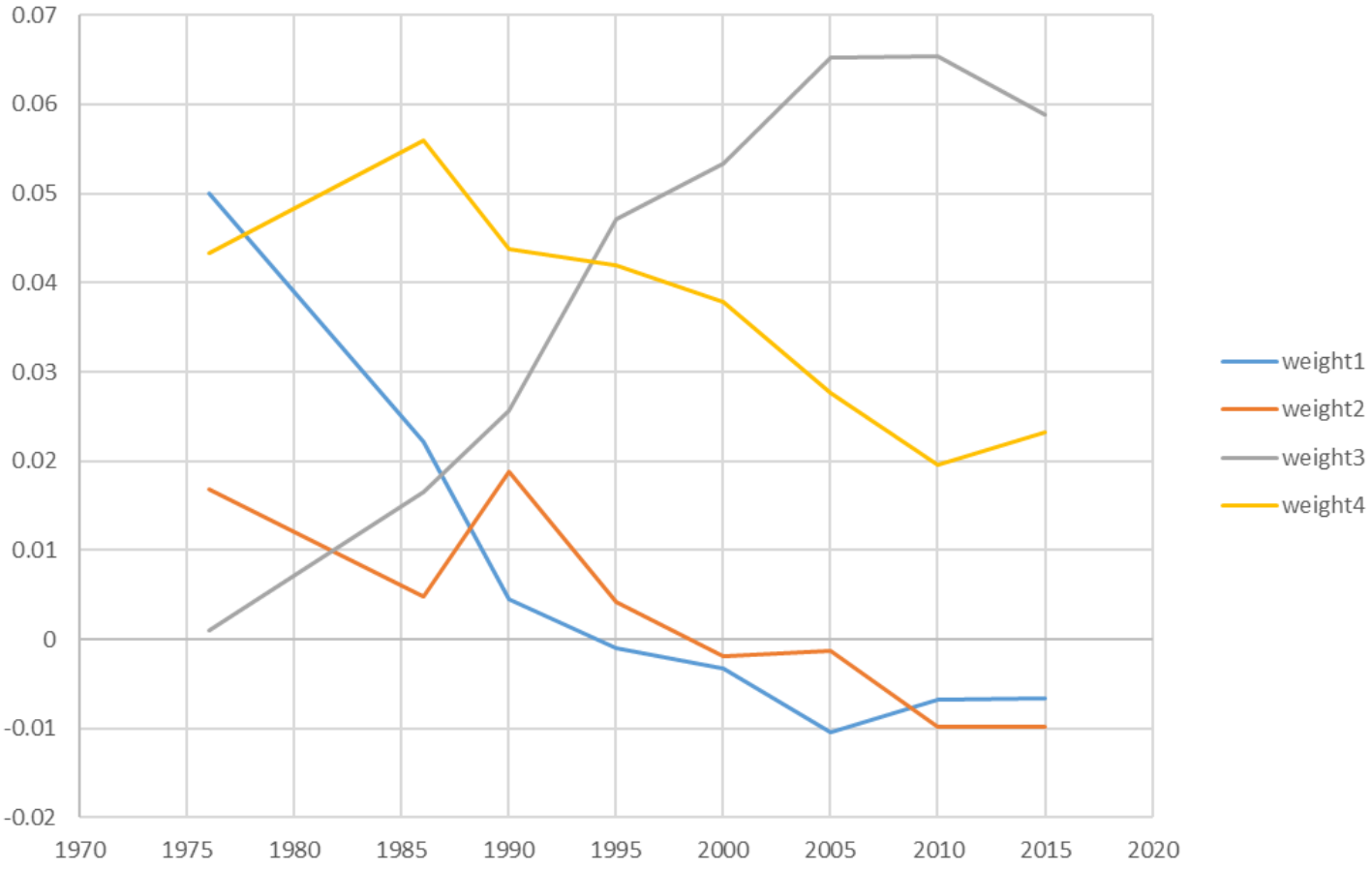
Observed and PCA smoothed



Colombia – observed and PCA smoothed



Trends in weights - Colombia



Key results

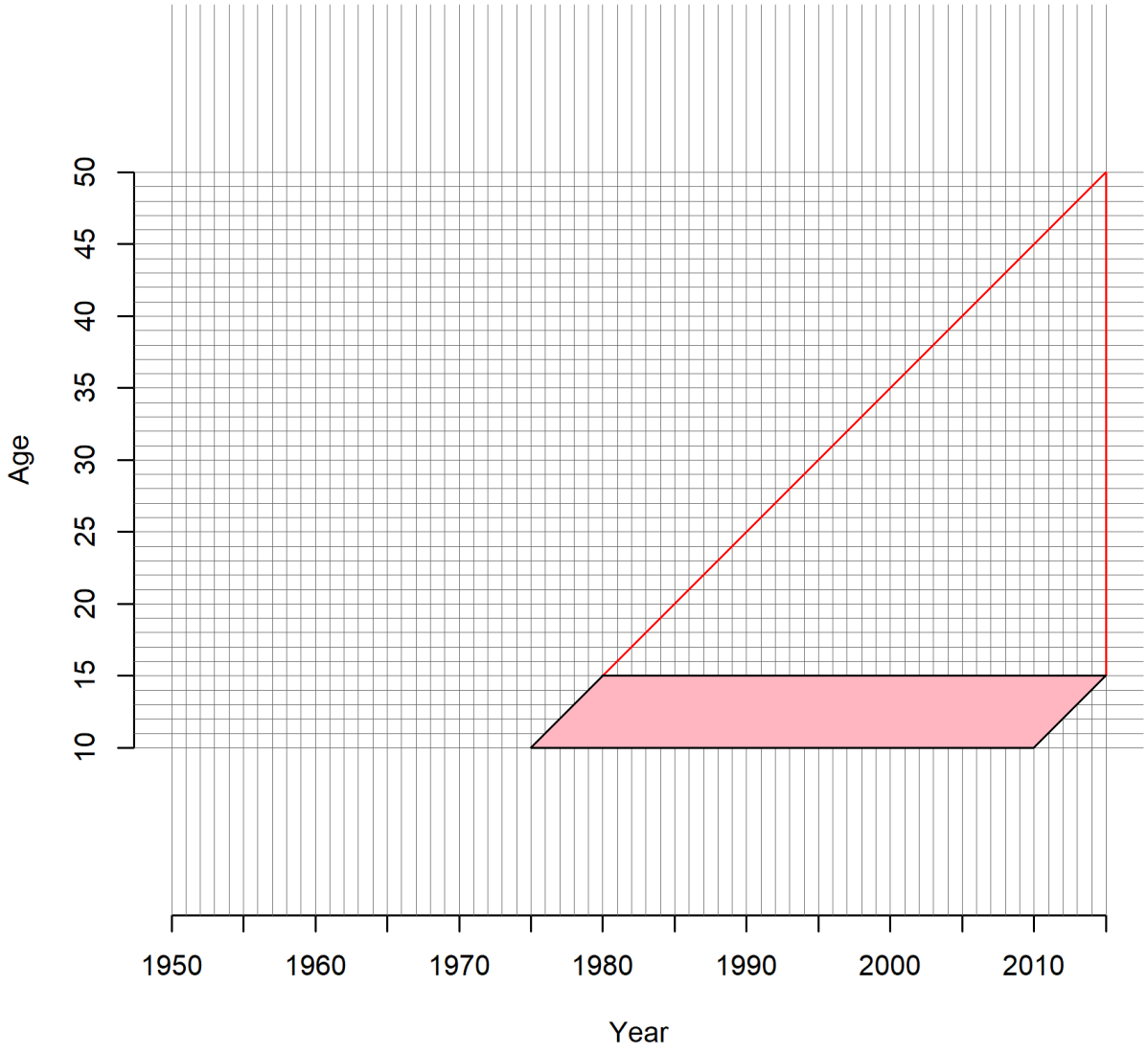
- A lot of trustworthy information on single year ASFRS in developing countries
- Age patterns are highly consistent across surveys within countries
- Good consistency with HDSS data
- Large variety of age patterns
- ASFRs available from age 12 (almost no births below 12)
- Smoothing is necessary
  - PCA (Pantazis & Clark, 2018) ? Splines, moving averages ?
- PCA could be used to reconstruct past age patterns
- Censuses could be exploited too

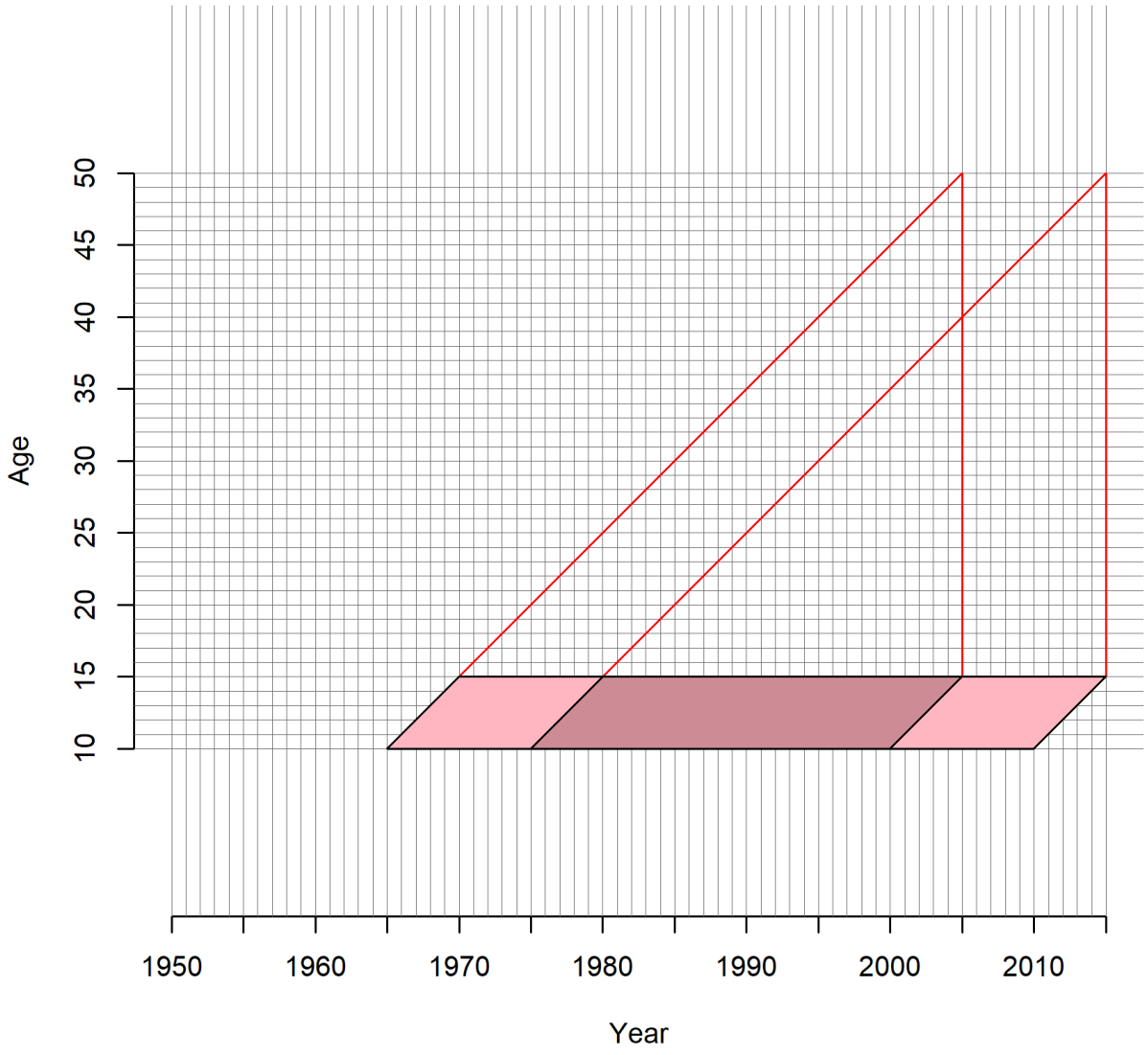
Trends in under-15 fertility

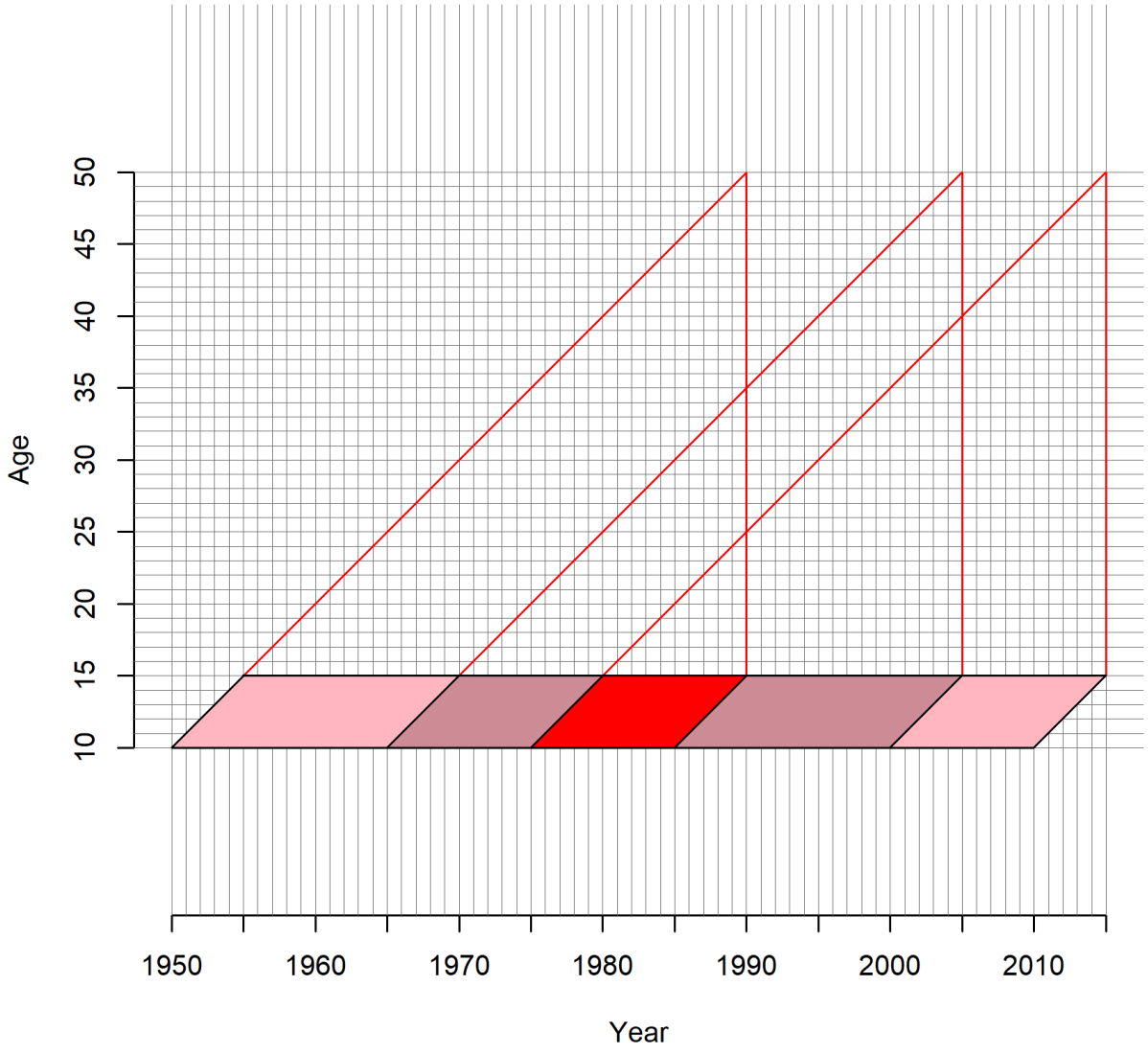
FROM SURVEYS

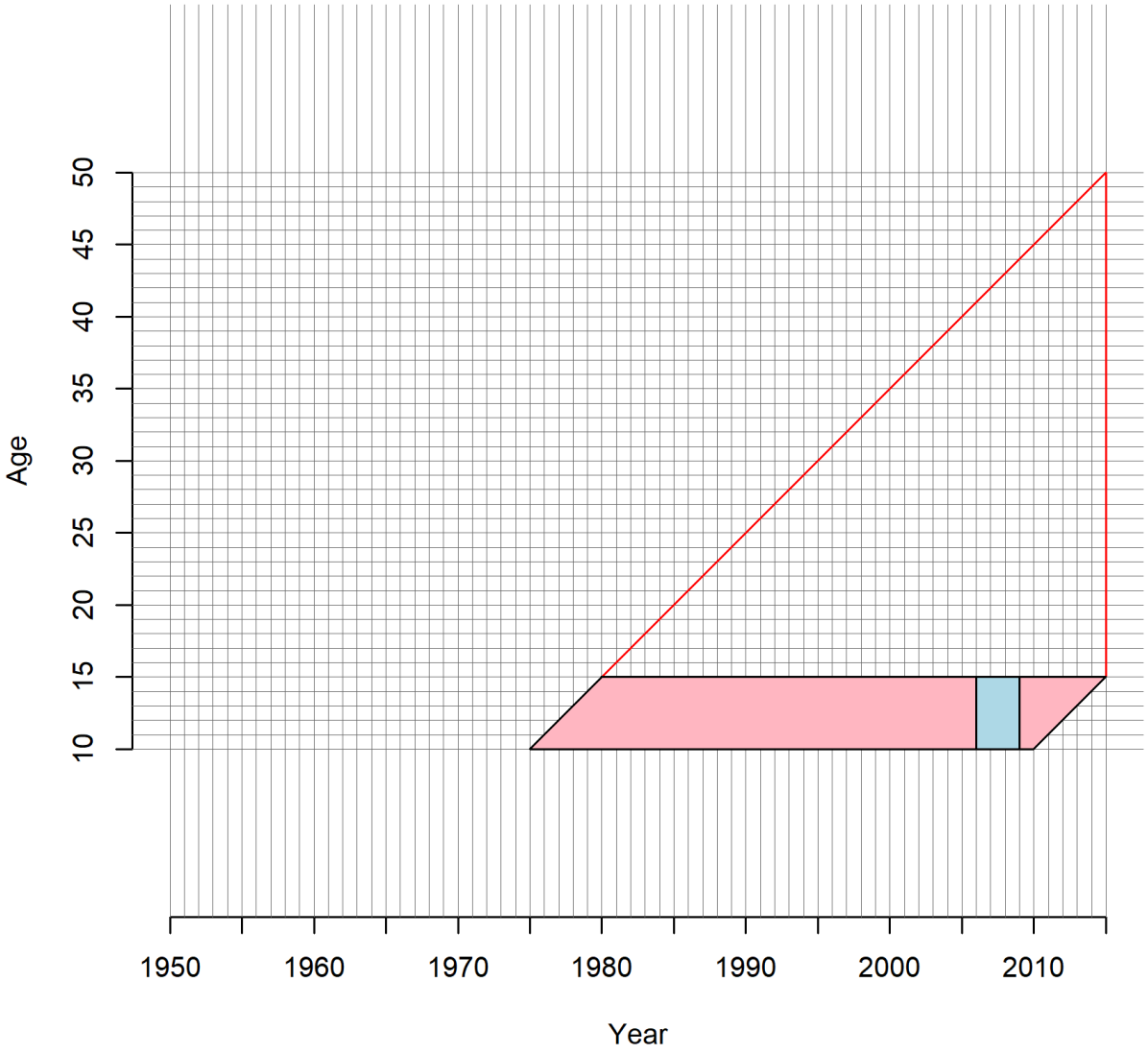
Available estimates of 10-14 fertility rates - developing countries

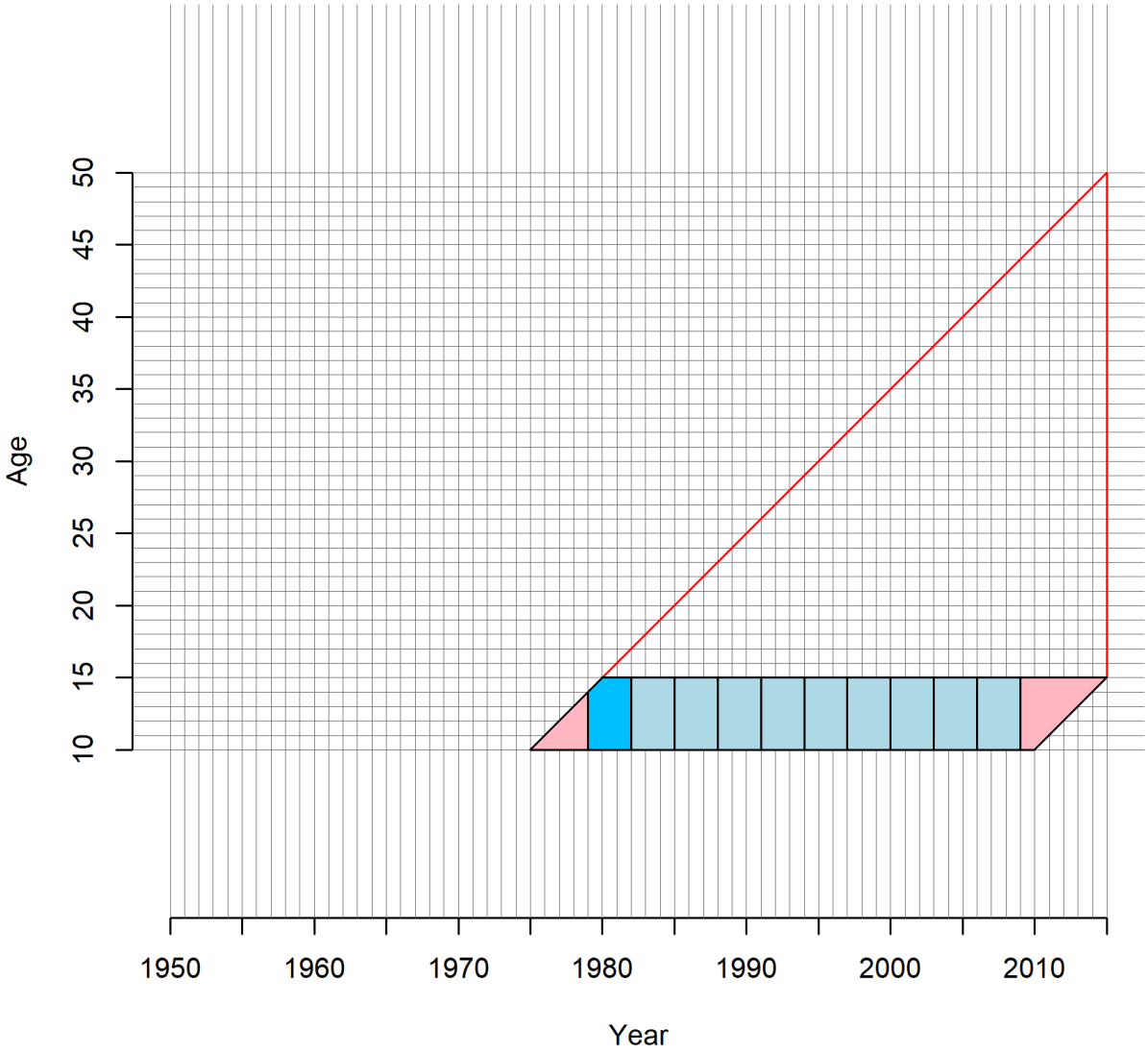
- Demographic and Health Surveys (DHS) Reports and StatCompiler
  - Estimates for the 3 years preceding the survey
    - Recent addition to StatCompiler : recent estimates from all the DHS
- United Nations Demographic Yearbooks, Human fertility database (HFD) and Human fertility collection (HFC)
  - Rates often not available (only numbers of births) and age-groups not well defined
  - Limited data in developing countries
- Global Burden of Disease (GBD) (2018)
  - "To our knowledge, our analysis provides the *first annual time series of fertility rates* in these age groups. [10-14 and 15-19]"
    - Largely relies on modelling

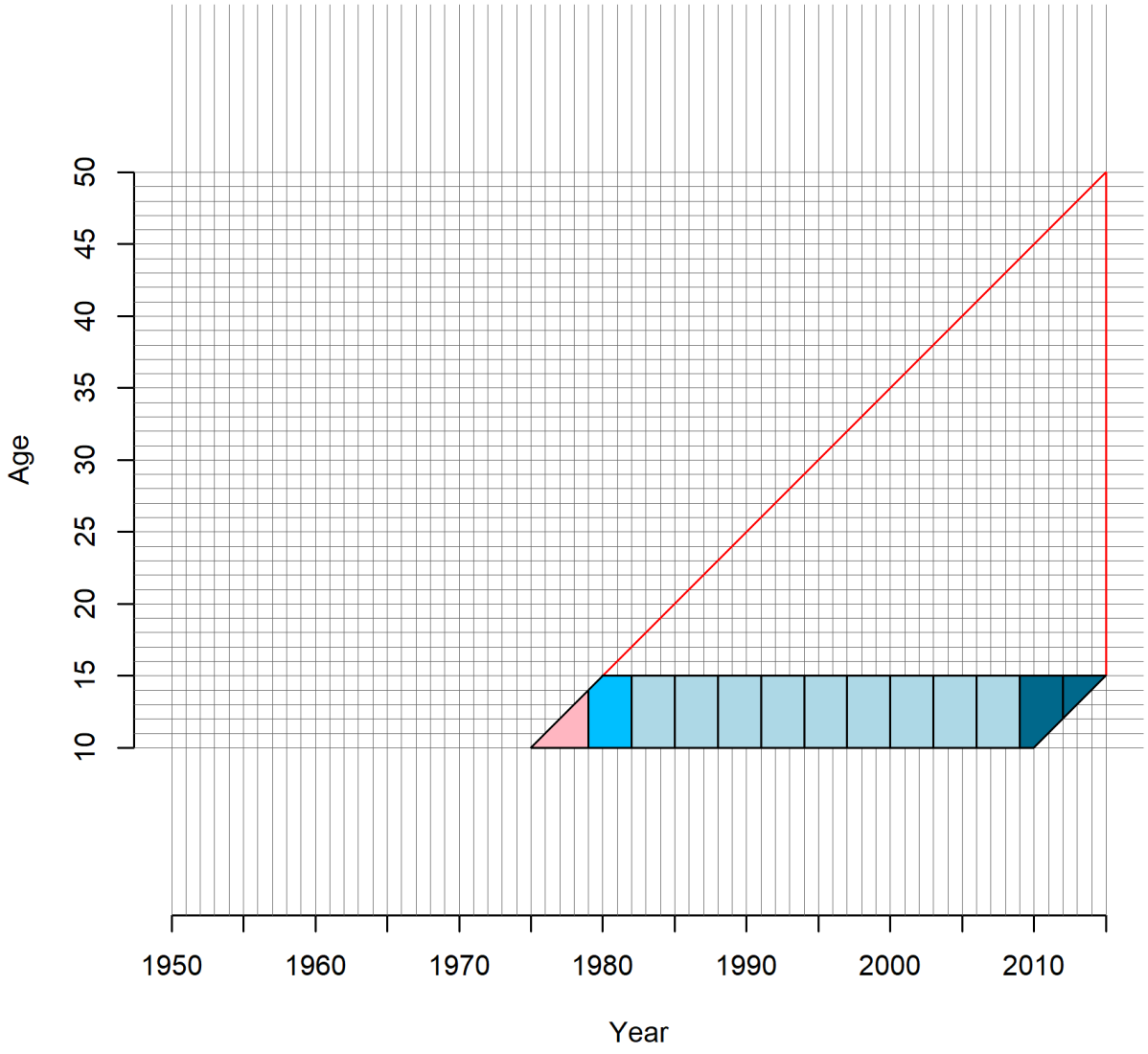

Age
10
15
20
25
30
35
40
45
50
1950
1960
1970
1980
1990
2000
2010
Year


Age
10
15
20
25
30
35
40
45
50
1950
1960
1970
1980
1990
2000
2010
Year


Age
10
15
20
25
30
35
40
45
50
1950
1960
1970
1980
1990
2000
2010
Year


Age
10
15
20
25
30
35
40
45
50
1950
1960
1970
1980
1990
2000
2010
Year


Age
10
15
20
25
30
35
40
45
50
1950
1960
1970
1980
1990
2000
2010
Year



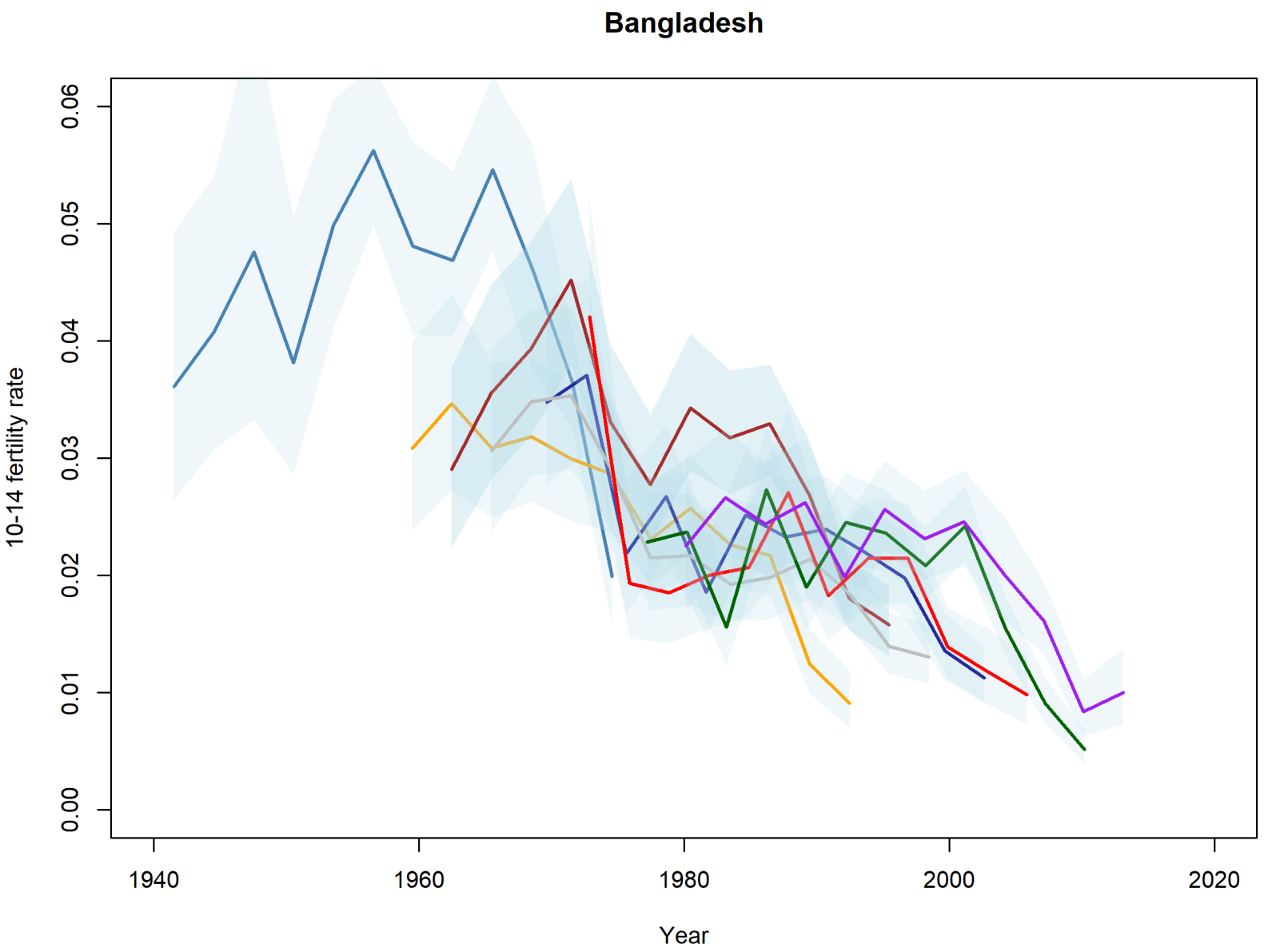
DHS ‘Lexis Approach’

Main potential issues

- Women's age misreporting
  - Women with a child may be 'aged' artificially -> underestimation
- Birth omissions
  - Especially if children died -> underestimation
- Misreporting of birth dates
  - Recent births may be displaced -> underestimation of recent estimates
  - Dates (and age at) early births may be misreported -> over or underestimation of past estimates
- Sample selection of survivors
  - If mortality higher among childbearing adolescents -> underestimation
- Sampling errors

Database

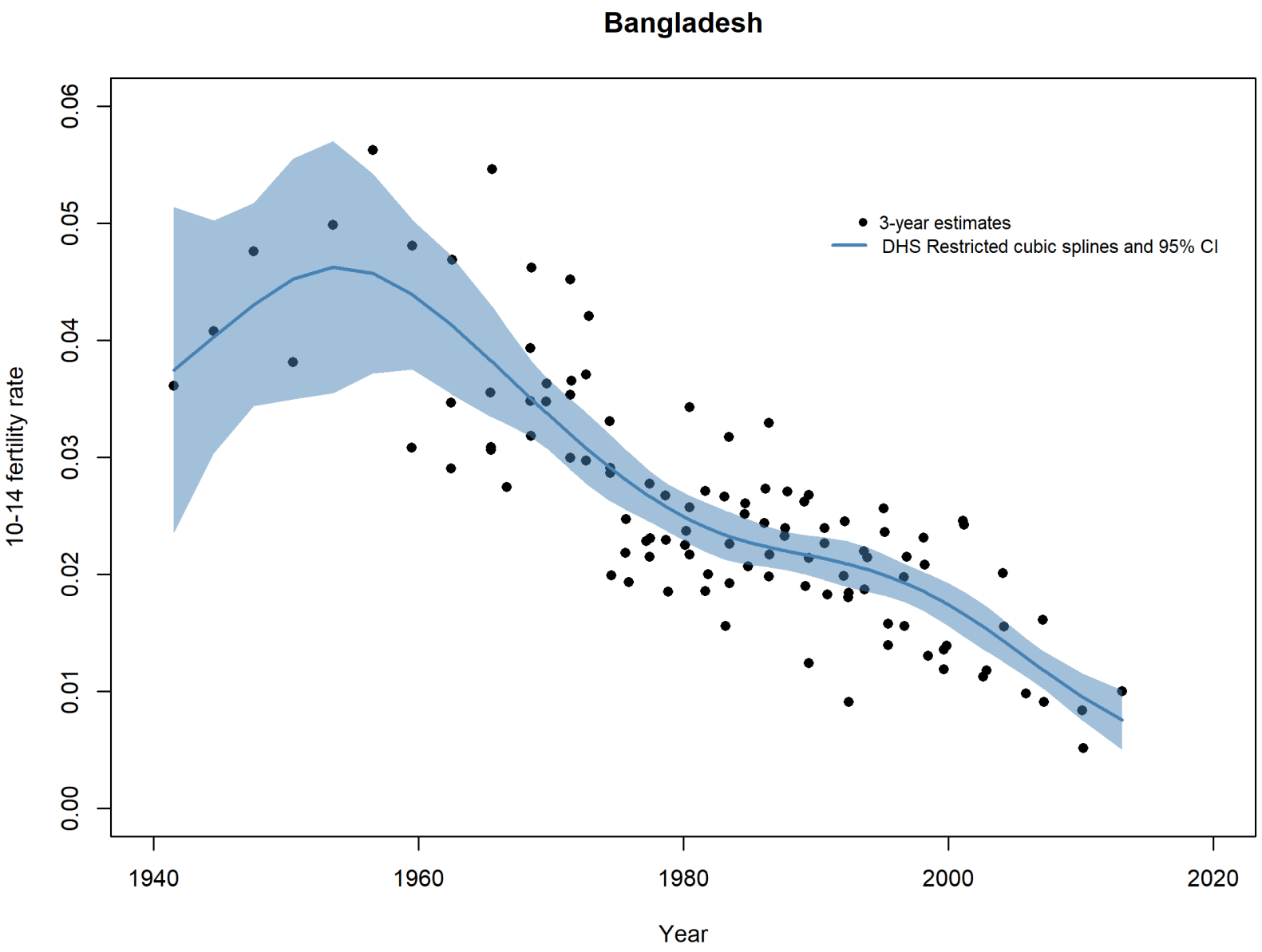
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	LocID	ISO3	ISO2	LocName	RegName	DataCatalc	CatalogNa	CatalogSh	datafile	ageg	per	window	centry	rate	lower	upper	se_lograte	events	exposure	
2	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	10-14	33-35	3	1976	0.0083	0.0060	0.0114	0.1647	60	7260	
3	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	10-14	30-32	3	1979	0.0122	0.0097	0.0154	0.1175	99	8059	
4	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	10-14	27-29	3	1982	0.0081	0.0062	0.0105	0.1332	86	10622	
5	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	10-14	24-26	3	1985	0.0143	0.0117	0.0174	0.1002	176	12343	
6	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	10-14	21-23	3	1988	0.0106	0.0084	0.0133	0.1165	126	11911	
7	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	10-14	18-20	3	1991	0.0154	0.0130	0.0184	0.0886	195	12619	
8	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	10-14	15-17	3	1994	0.0164	0.0139	0.0194	0.0843	267	16250	
9	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	10-14	12-14	3	1997	0.0139	0.0117	0.0165	0.0883	261	18832	
10	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	10-14	9-11	3	2000	0.0094	0.0077	0.0113	0.0960	209	22339	
11	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	10-14	6-8	3	2003	0.0081	0.0068	0.0096	0.0866	247	30593	
12	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	10-14	3-5	3	2006	0.0043	0.0034	0.0054	0.1199	136	31897	
13	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	10-14	0-2	3	2009	0.0008	0.0004	0.0014	0.3237	9	12132	
14	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	10-14	30-34	5	1978	0.0104	0.0086	0.0127	0.0992	140	13468	
15	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	10-14	25-29	5	1983	0.0126	0.0106	0.0150	0.0880	237	18801	
16	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	10-14	20-24	5	1988	0.0122	0.0102	0.0145	0.0899	243	20009	
17	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	10-14	15-19	5	1993	0.0148	0.0128	0.0171	0.0737	369	24933	
18	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	10-14	10-14	5	1998	0.0131	0.0112	0.0152	0.0780	432	33104	
19	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	10-14	5-9	5	2003	0.0077	0.0066	0.0089	0.0752	388	50662	
20	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	10-14	0-4	5	2008	0.0014	0.0010	0.0018	0.1527	44	32026	
21	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	15-19	27-29	3	1982	0.1145	0.1046	0.1254	0.0464	912	7967	
22	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	15-19	24-26	3	1985	0.1457	0.1350	0.1574	0.0392	1274	8742	
23	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	15-19	21-23	3	1988	0.1508	0.1404	0.1619	0.0362	1717	11388	
24	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	15-19	18-20	3	1991	0.1882	0.1776	0.1995	0.0296	2336	12409	
25	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	15-19	15-17	3	1994	0.2077	0.1958	0.2204	0.0301	2445	11770	
26	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	15-19	12-14	3	1997	0.1903	0.1800	0.2011	0.0283	2619	13767	
27	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	15-19	9-11	3	2000	0.1471	0.1391	0.1556	0.0287	2543	17286	
28	4	AFG	AF	Afghanista	Southern A	5045	Afghanista	2010 DHS	afir66fl.dta	15-19	6-8	3	2003	0.1529	0.1452	0.1610	0.0263	3006	19659	



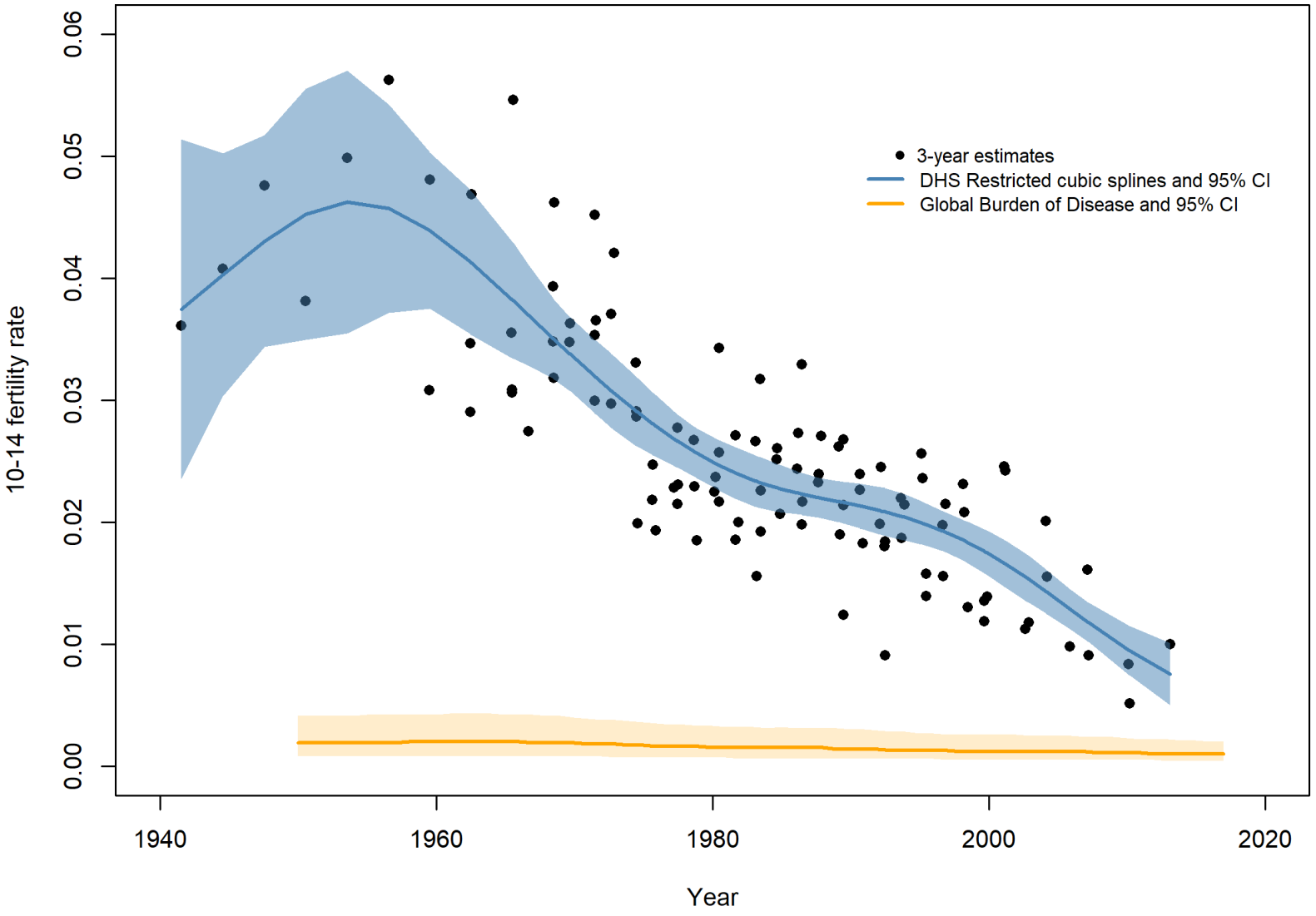
9 surveys in Bangladesh, covering more than 70 years.

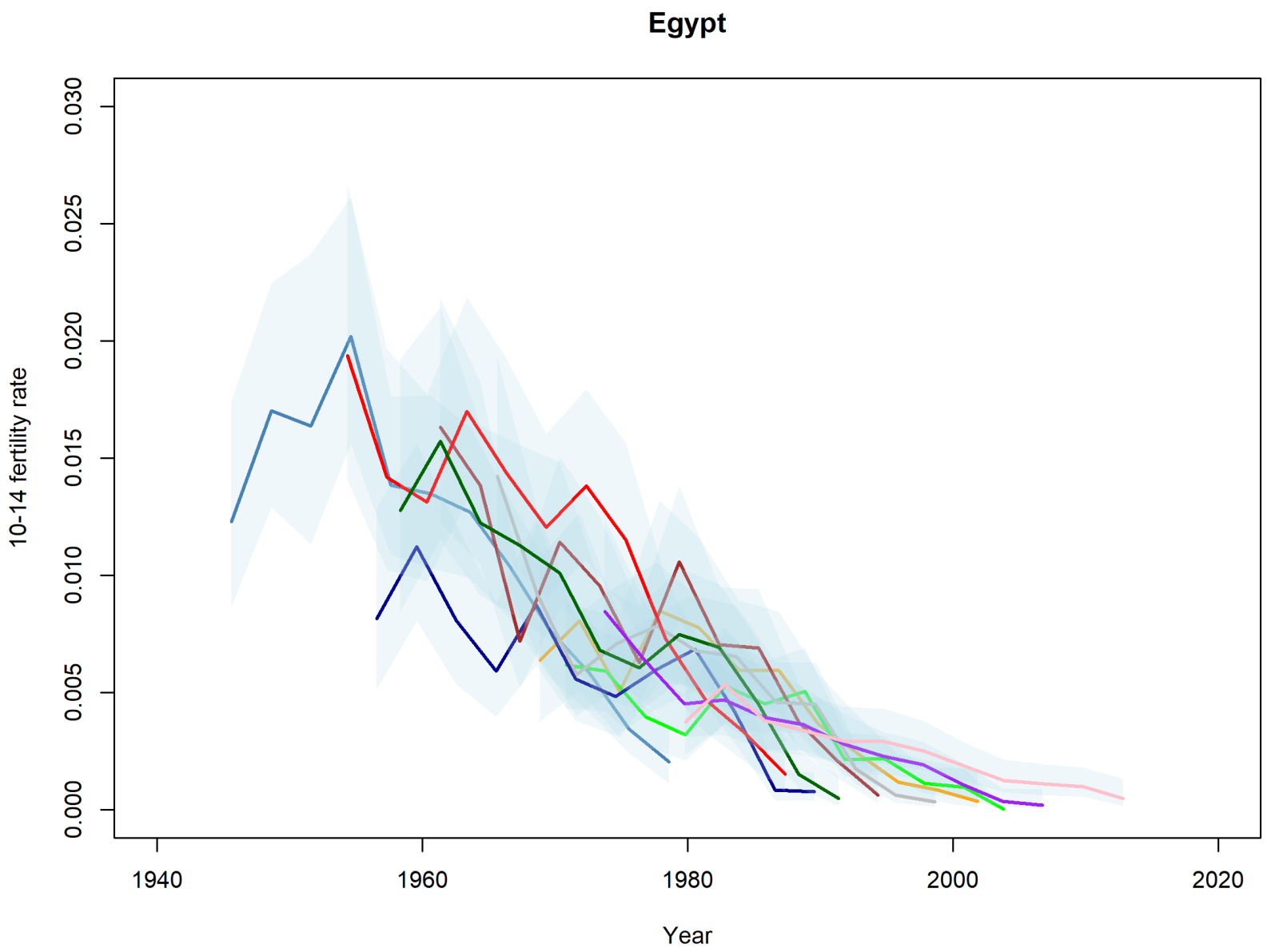
Method

- Pooled surveys (events and exposure)
- Negative binomial regression with restricted cubic splines (knots spaced by 10 years)
  - Log rate model

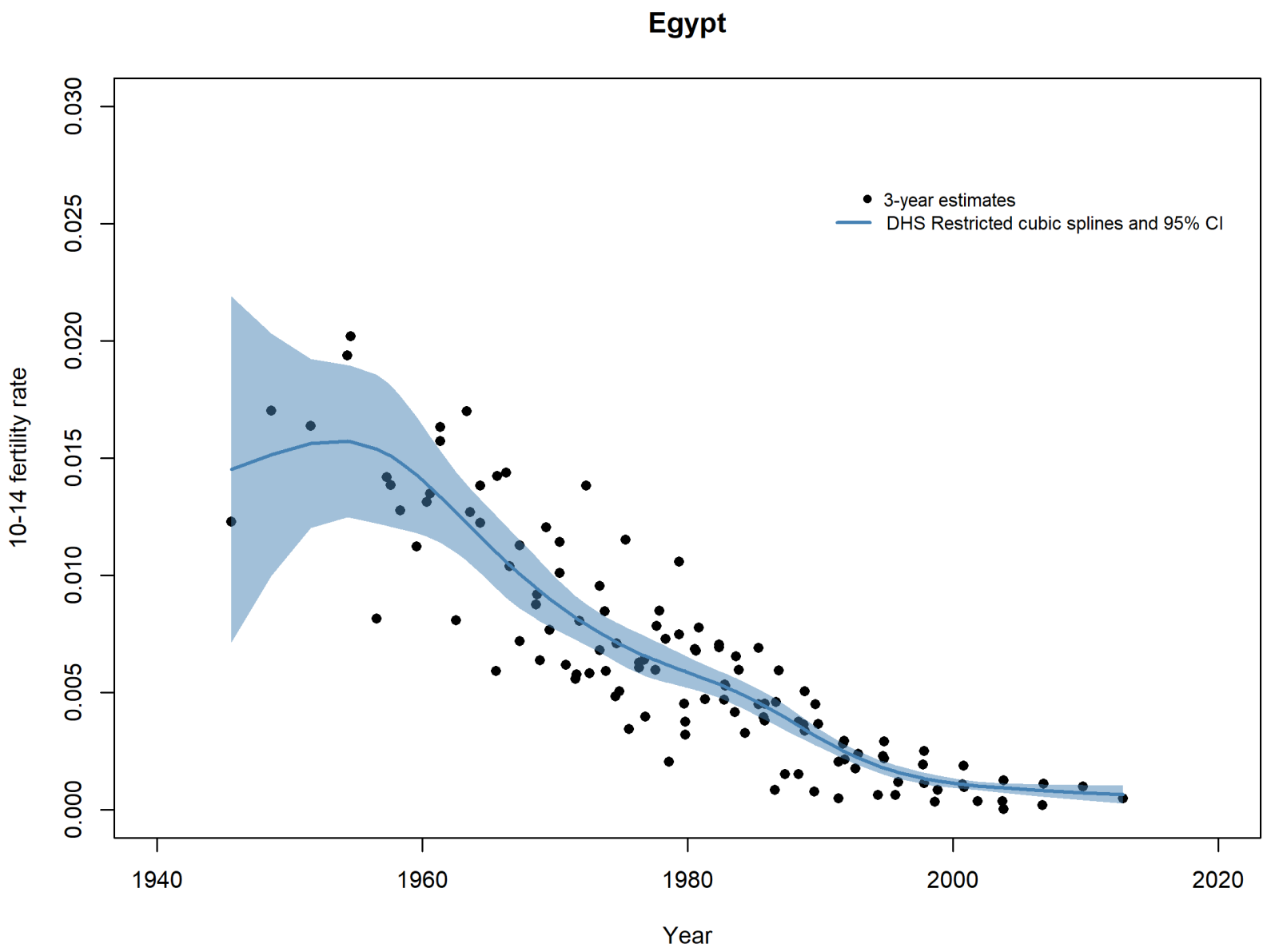


Pooled surveys (events and exposure). Negative binomial regression with restricted cubic splines (knots spaced by 10 years)

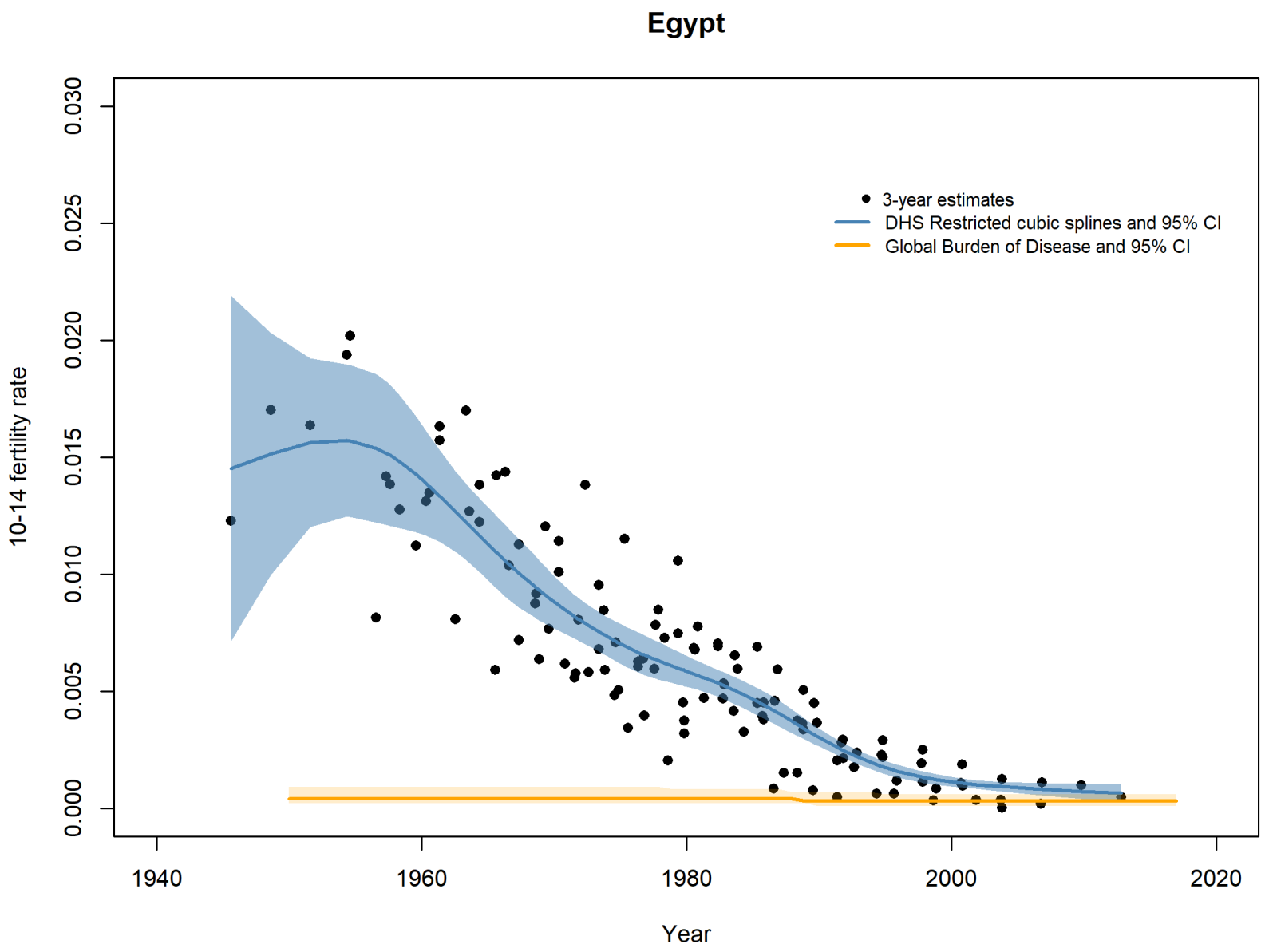

Bangladesh
10-14 fertility rate
0.00
0.01
0.02
0.03
0.04
0.05
0.06
1940
1960
1980
2000
2020
Year
3-year estimates
DHS Restricted cubic splines and 95% CI
Global Burden of Disease and 95% CI

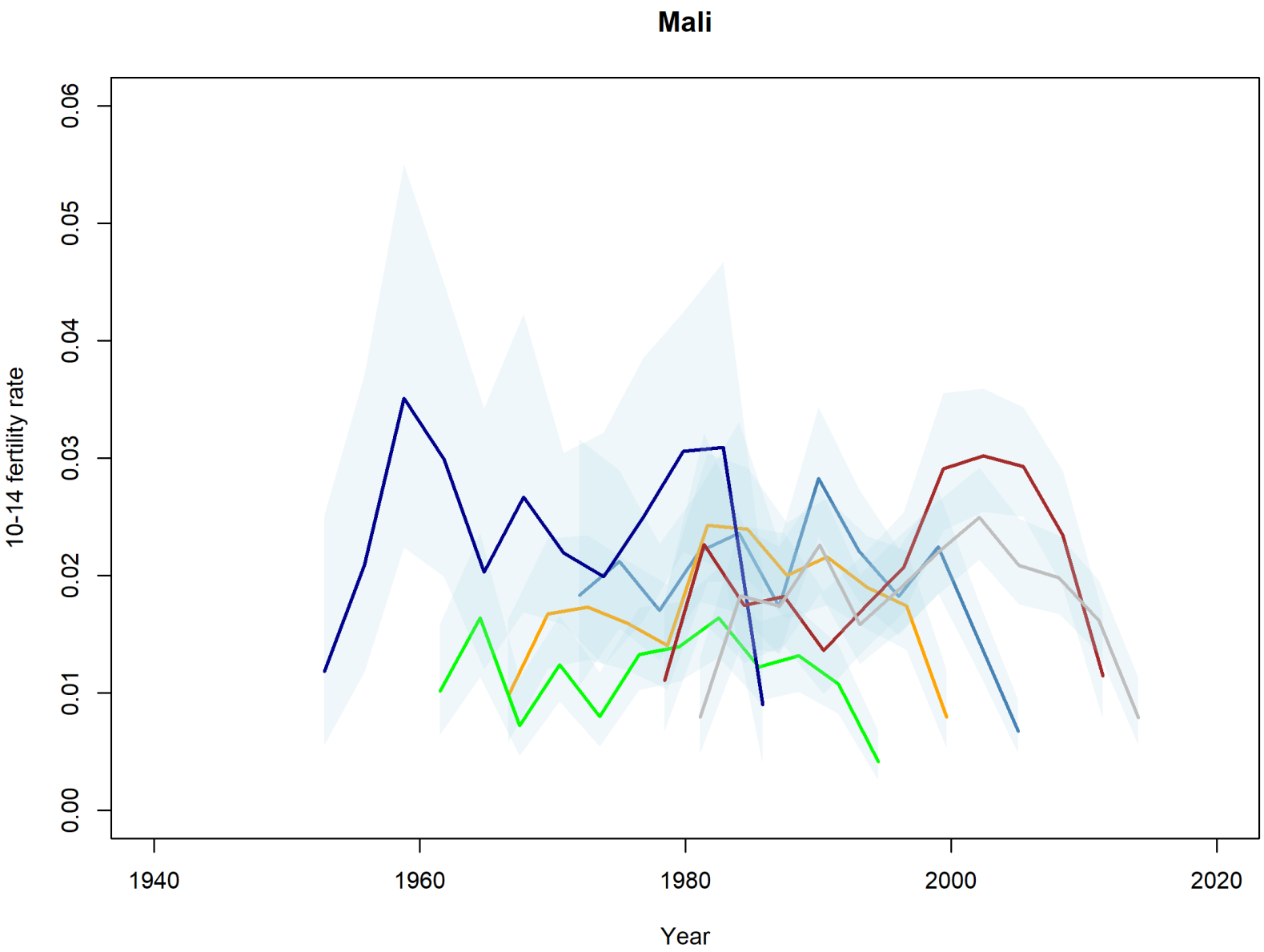


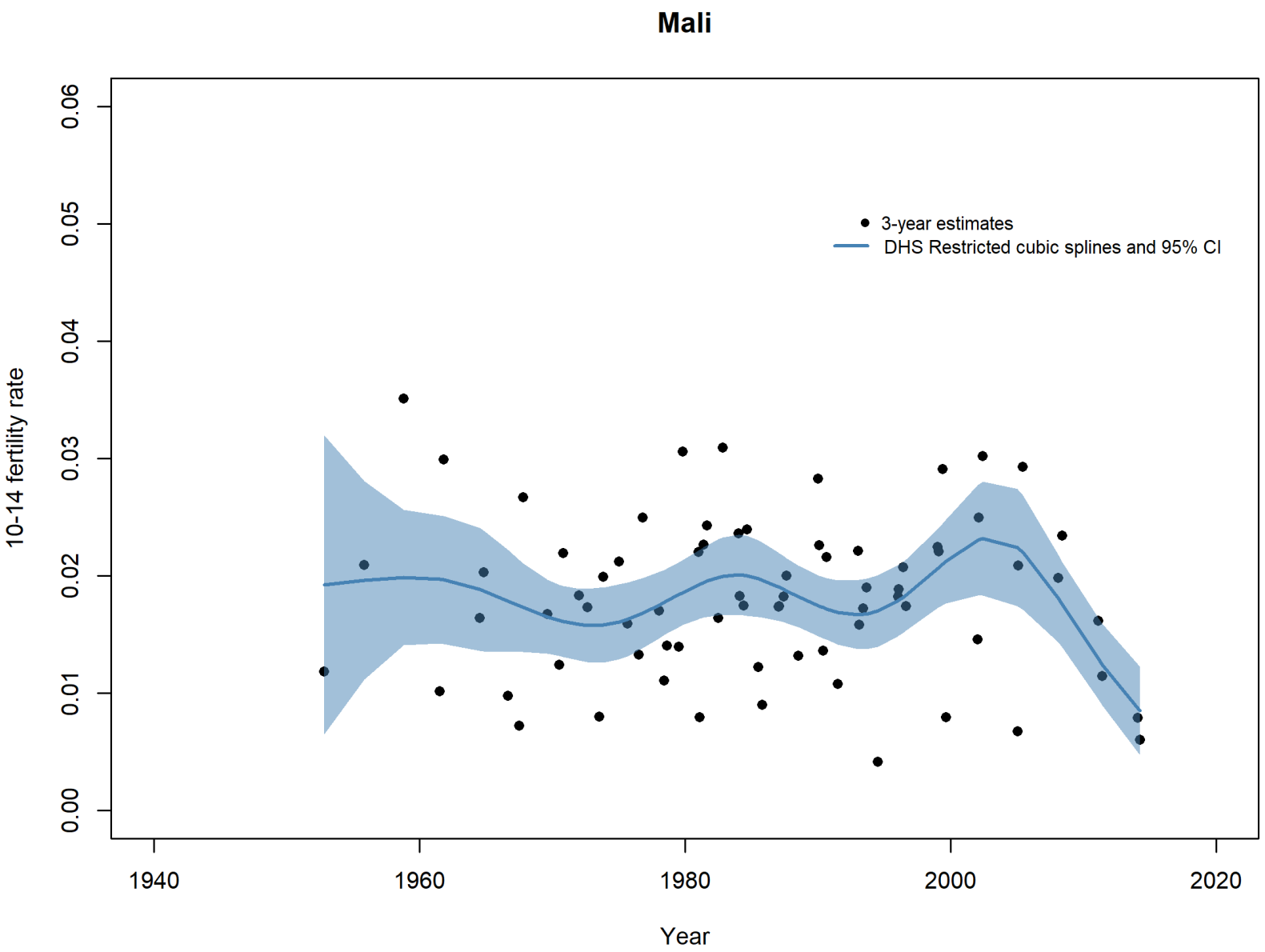
10 surveys in Egypt, covering close to 70 years.



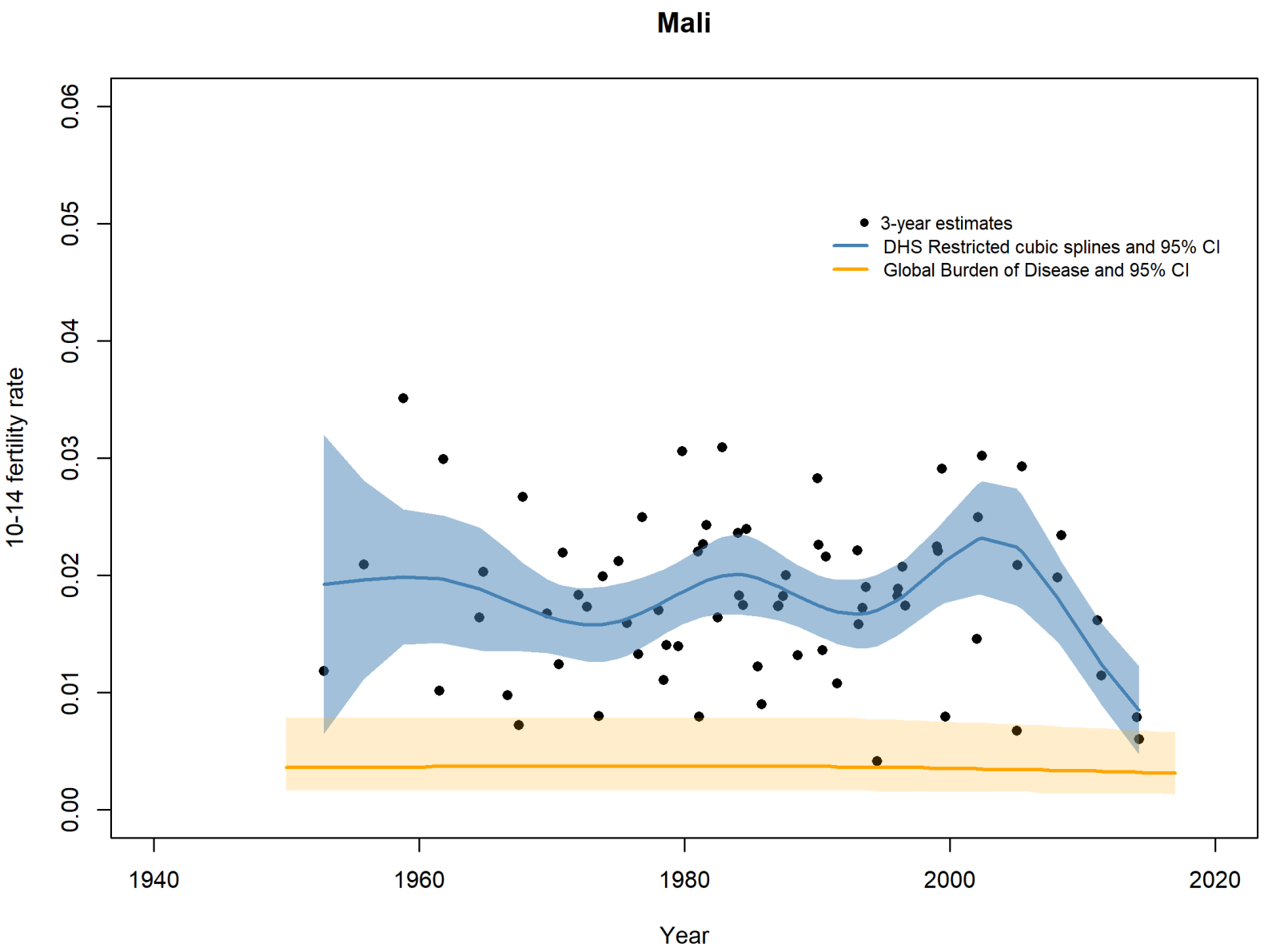
Pooled surveys (events and exposure). Negative binomial regression with restricted cubic splines (knots spaced by 10 years)


Egypt
10-14 fertility rate
0.000
0.005
0.010
0.015
0.020
0.025
0.030
1940
1960
1980
2000
2020
Year
3-year estimates
DHS Restricted cubic splines and 95% CI
Global Burden of Disease and 95% CI

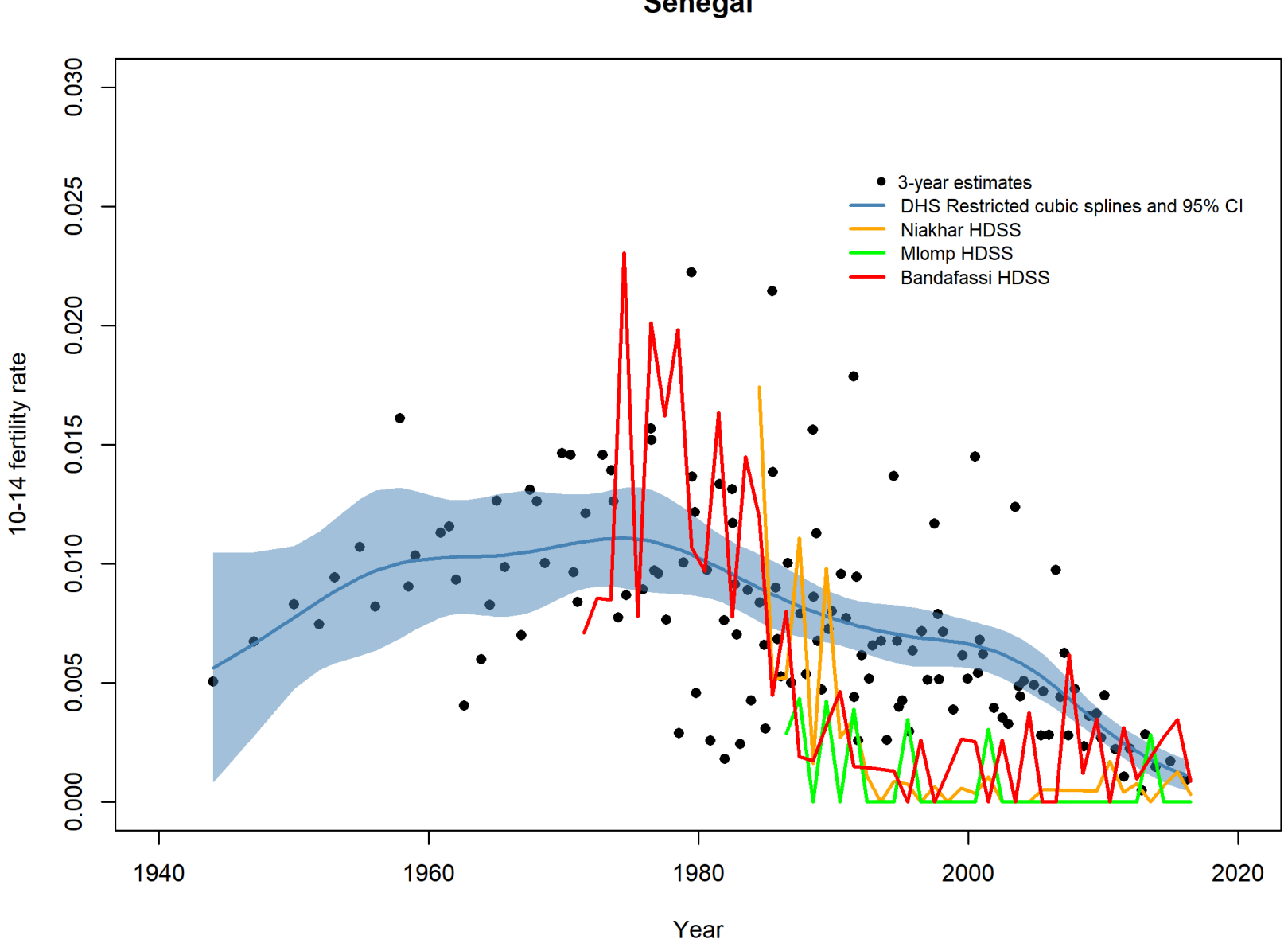

Mali
10-14 fertility rate
0.00
0.01
0.02
0.03
0.04
0.05
0.06
1940
1960
1980
2000
2020
Year



Pooled surveys (events and exposure). Negative binomial regression with restricted cubic splines (knots spaced by 10 years)


Mali
10-14 fertility rate
0.00
0.01
0.02
0.03
0.04
0.05
0.06
1940
1960
1980
2000
2020
Year
3-year estimates
DHS Restricted cubic splines and 95% CI
Global Burden of Disease and 95% CI

Under-15 fertility in HDSS


Senegal
10-14 fertility rate
0.000
0.005
0.010
0.015
0.020
0.025
0.030
1940
1960
1980
2000
2020
Year
3-year estimates
DHS Restricted cubic splines and 95% CI
Niakhar HDSS
Mlomp HDSS
Bandafassi HDSS

Key results

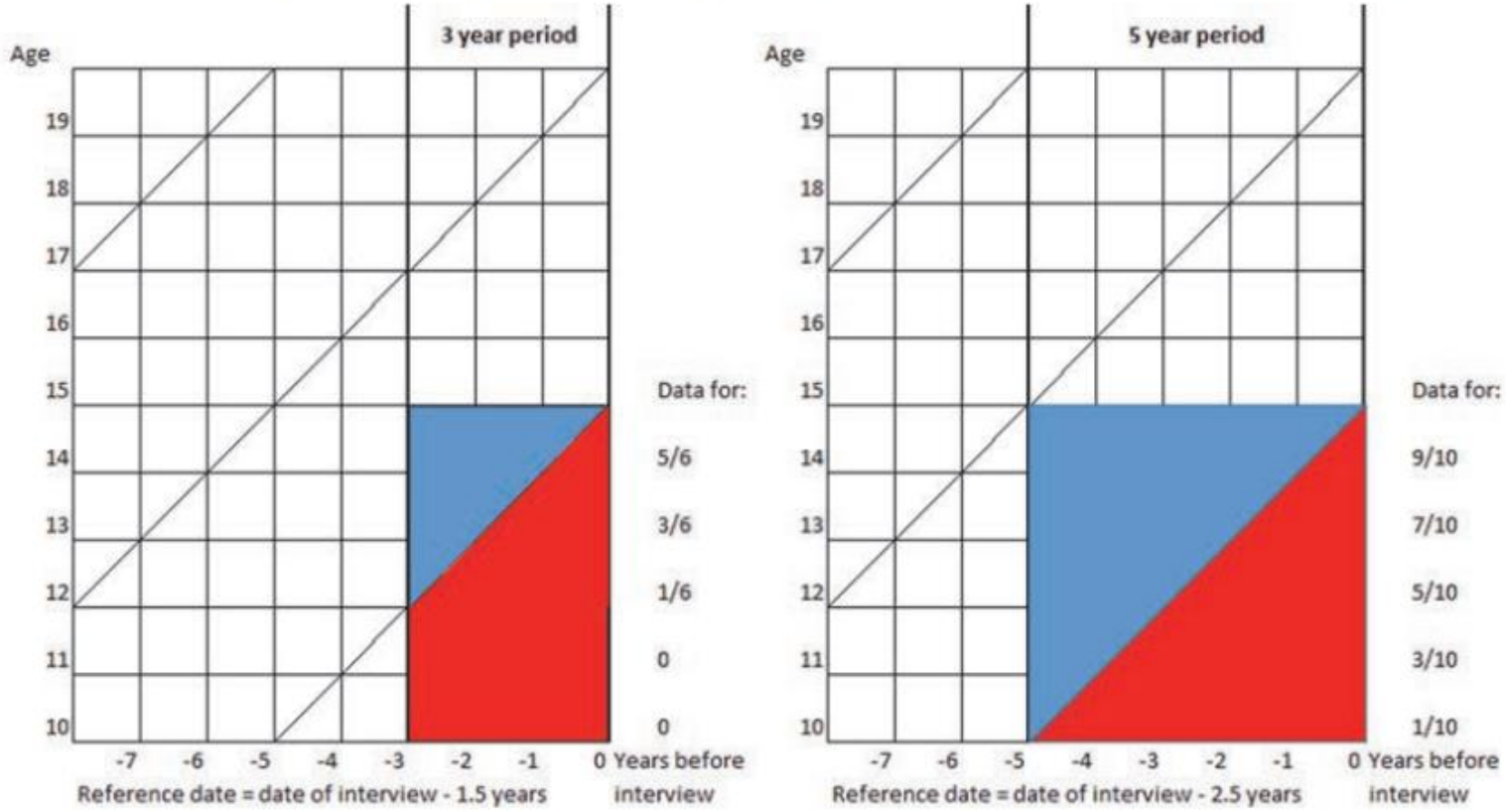
- Possible to estimate long-term trends in under-15 fertility
  - Data on young adolescents fertility widely available and largely untapped
  - Fairly consistent estimates across surveys
  - Reasonable confidence intervals
  - Seems consistent with HDSS, but more comparisons needed
- Compared to existing estimates
  - Longer trends than published DHS
  - Much larger set of countries than in Human Fertility Collection
  - Much more realistic estimates than GBD estimates
- Very limited impact on TFR (usually << 1%)

Thank you

Appendix

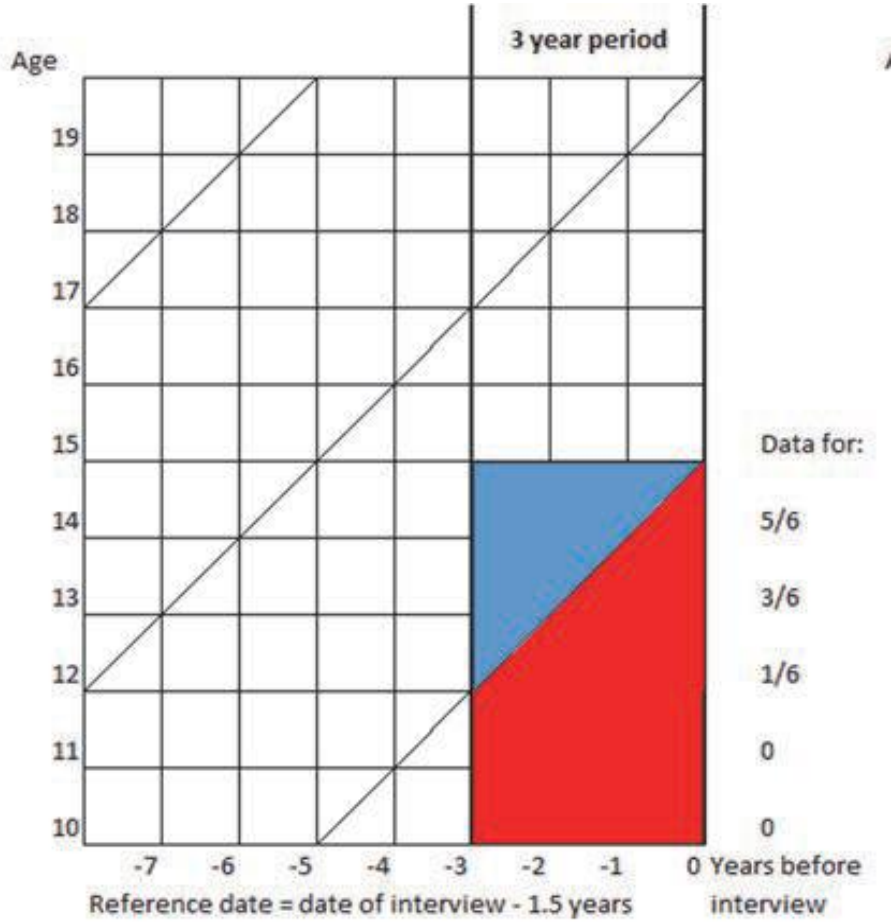
Recent estimates : truncated data

Figure 2.1 **Lexis diagrams showing, in blue and red, the areas for which births and exposure are required for the fertility rates for age 10-14 during the 3 years and 5 years before the survey**



Source : Pullum, Thomas W., Trevor Croft, and Kerry L. D. MacQuarrie. 2018. Methods to Estimate Under-15 Fertility Using Demographic and Health Surveys Data. DHS Methodological Reports No. 23. Rockville, Maryland, USA: ICF.

The Lexis Approach



$$W_{12} = \frac{\frac{6}{1} * E_{12}}{D}$$

$$W_{13} = \frac{\frac{6}{3} * E_{13}}{D}$$

$$W_{14} = \frac{\frac{6}{5} * E_{14}}{D}$$

$$D = \frac{6}{1} * E_{12} + \frac{6}{3} * E_{13} + \frac{6}{5} * E_{14}$$

$$R1_{Lexis} = W_{12} r_{12} + W_{13} r_{13} + W_{14}\, r_{14}$$

$$R2_{Lexis} = 0.6 * R1_{Lexis}$$