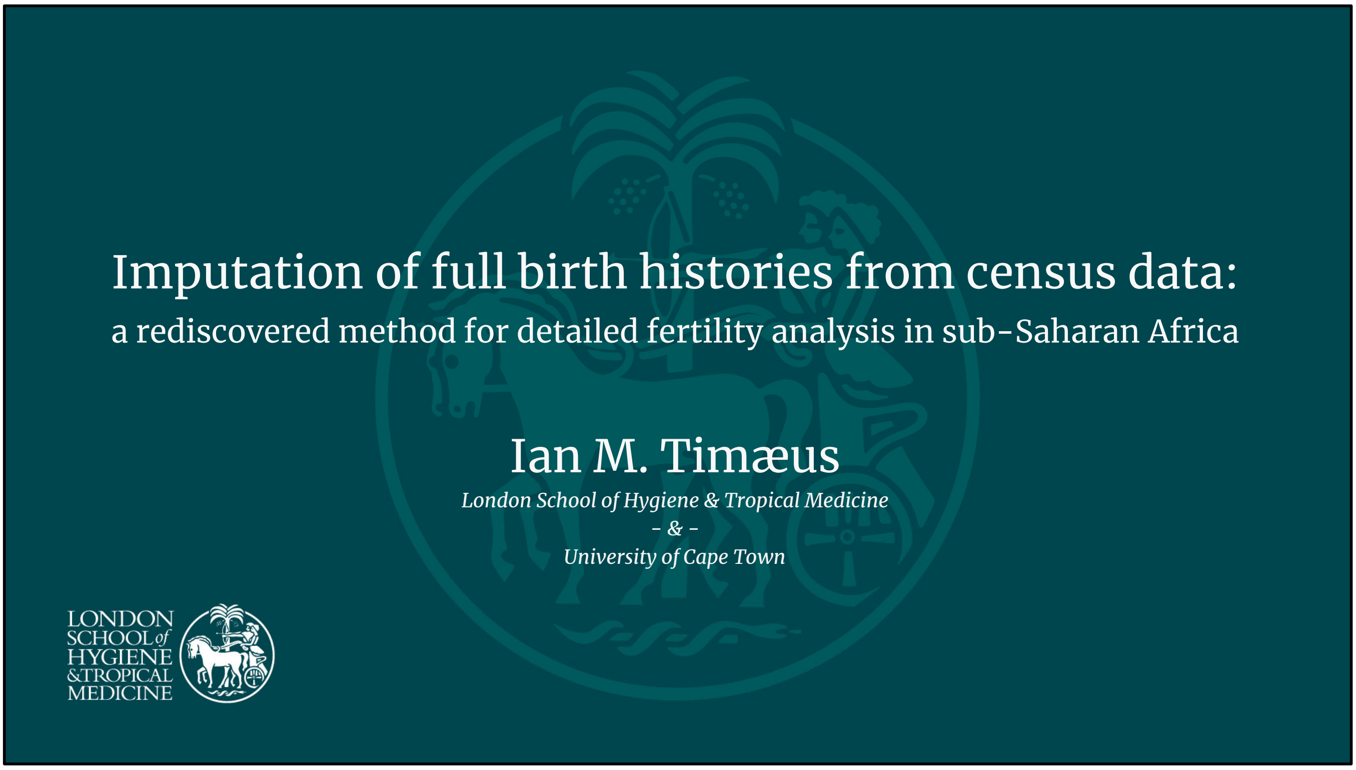
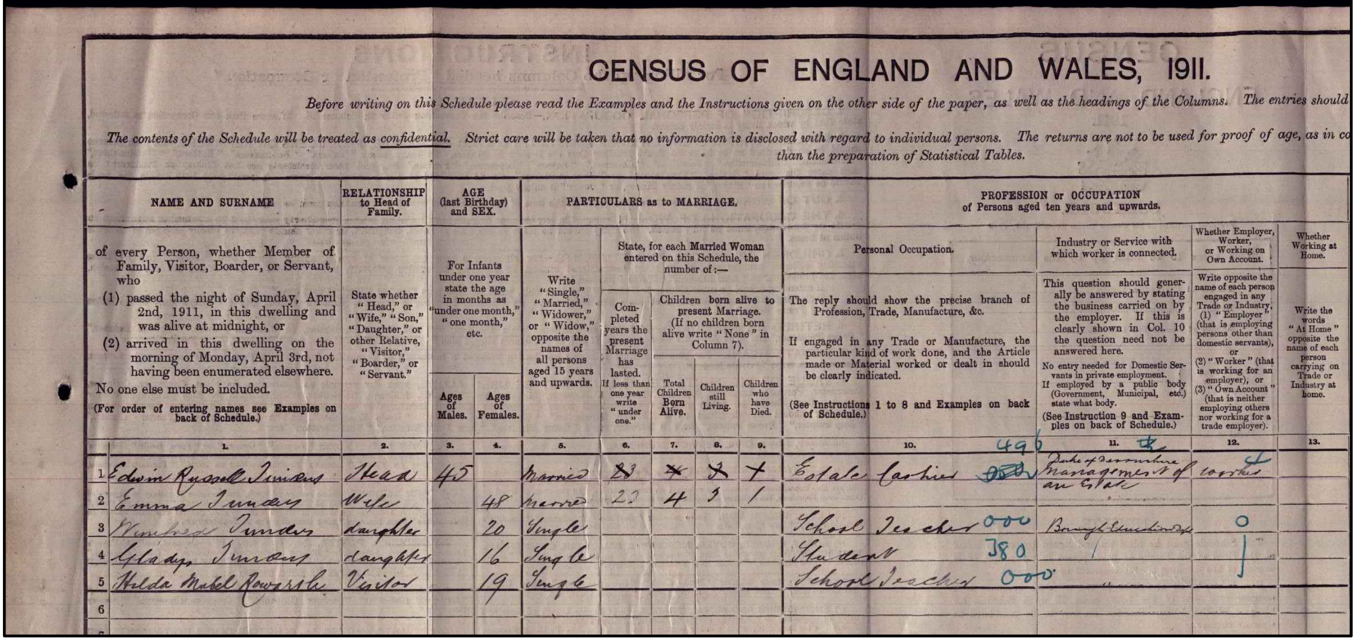


Hallo. What I want to discuss today is a 25-year-old method for making detailed fertility estimates from a census age distribution that most of you will never have heard of. First, I am going to describe the method and then I am going to illustrate its potential using some results for four Eastern and Southern African countries.

The method was first proposed by Luther and Cho in 1998 and can be thought of as an elaboration of the Own-Children Method of estimating fertility. The key idea is to supplement information on the ages of women's own children with information from the summary birth history questions about how many children each woman has borne in total and how many of her children have since died.

Summary birth history questions on fertility

LONDON SCHOOL of HYGIENE & TROPICAL MEDICINE


CENSUS OF ENGLAND AND WALES, 1911.

Before writing on this Schedule please read the Examples and the Instructions given on the other side of the paper, as well as the headings of the Columns. The entries should

The contents of the Schedule will be treated as confidential. Strict care will be taken that no information is disclosed with regard to individual persons. The returns are not to be used for proof of age, as in co than the preparation of Statistical Tables.

NAME AND SURNAME	RELATIONSHIP to Head of Family.	AGE (last Birthday) and SEX.		PARTICULARS as to MARRIAGE.					PROFESSION or OCCUPATION of Persons aged ten years and upwards.			
of every Person, whether Member of Family, Visitor, Boarder, or Servant, who (1) passed the night of Sunday, April 2nd, 1911, in this dwelling and was alive at midnight, or (2) arrived in this dwelling on the morning of Monday, April 3rd, not having been enumerated elsewhere. No one else must be included. (For order of entering names see Examples on back of Schedule.)	State whether "Head," or "Wife," "Son," "Daughter," or other Relative, "Visitor," "Boarder," or "Servant."	For Infants under one year state the age in months as "under one month," "one month," etc. Ages of Males.	Ages of Females.	Write "Single," "Married," "Widower," or "Widow," opposite the names of all persons aged 15 years and upwards.	State, for each Married Woman entered on this Schedule, the number of:— Completed years the present Marriage has lasted. If less than one year write "under one."	Children born alive to present Marriage. (If no children born alive write "None" in Column 7). Total Children Born Alive.	Children still Living.	Children who have Died.	Personal Occupation. The reply should show the precise branch of Profession, Trade, Manufacture, &c. If engaged in any Trade or Manufacture, the particular kind of work done, and the Article made or Material worked or dealt in should be clearly indicated. (See Instructions 1 to 8 and Examples on back of Schedule.)	Industry or Service with which worker is connected. This question should generally be answered by stating the business carried on by the employer. If this is clearly shown in Col. 10 the question need not be answered here. No entry needed for Domestic Servants in private employment. If employed by a public body (Government, Municipal, etc.) state what body. (See Instruction 9 and Examples on back of Schedule.)	Whether Employer, Worker, or Working on Own Account. Write opposite the name of each person engaged in any Trade or Industry, (1) "Employer" (that is employing persons other than domestic servants), or (2) "Worker" (that is working for an employer), or (3) "Own Account" (that is neither employing others nor working for a trade employer).	Whether Working at Home. Write the words "At Home" opposite the name of each person carrying on Trade or Industry at home.
1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.
1 Edwin Russell Timaeus	Head	45		Married	23	4	3	1	Estate Cashier	Management of an Estate	Worker	
2 Emma Timaeus	Wife		48	Married	22	4	3	1				
3 Winifred Timaeus	daughter		20	Single					School Teacher			
4 Gladys Timaeus	daughter		16	Single					Student			
5 Hilda Mabel Rowarth	Visitor		19	Single					School Teacher			
6												

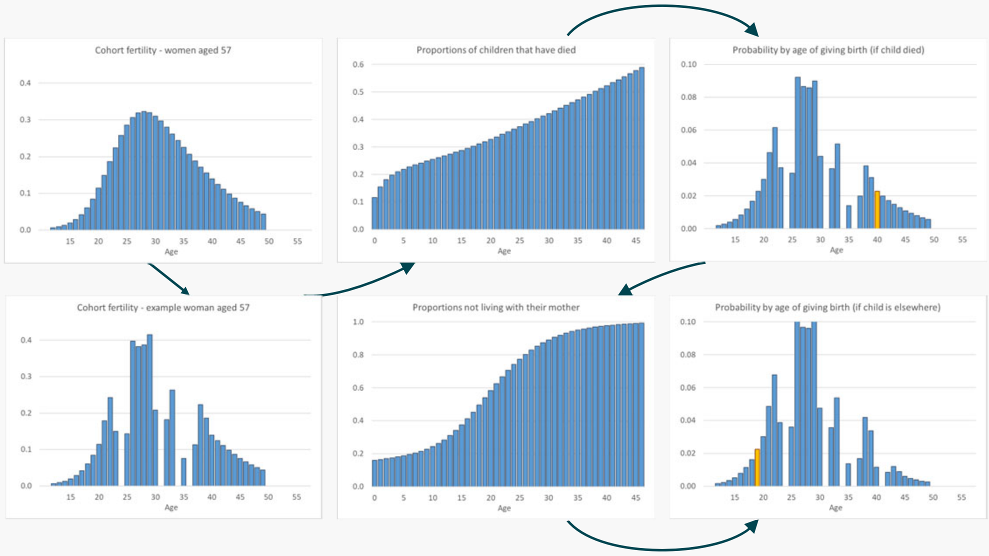
For example, if we inspect this census schedule, we find Edwin Timaeus living with his 48-year-old wife, Emma. Looking at the next columns, they have been married 22 years and Emma has had 4 children, 3 of whom are still living and one of whom has died.

Her two daughters are listed on the following rows of the form. Winnie was born 20 years earlier when her mother was 28 and Gladys 16 years earlier when her mother was 32. Thus, Emma had two children whose ages we don't know: one who was dead and another, my grandfather, who was living elsewhere.

By working through the women in a census one by one, imputing an age to each dead or absent child, one can reconstruct complete individual-level birth histories for the women.

Imputing ages to dead and absent children

(Woman aged 57, own-children aged 21, 23, 29 & 33, 1 absent, 1 dead)



Key reference: Luther, N.Y. and Cho, L-J. Reconstruction of birth histories from census and household survey data. *Population Studies*, 1988, 42, 451-72

Luther and Cho's implementation of this idea began by estimating an age-specific fertility distribution for the population using the Own-Children Method. For each woman, they then cut notches out this fertility distribution surrounding her age at the times when each of her own children was born.

To impute an age to a dead child, they multiplied the notched fertility distribution by children's cumulative probability of having died by age and normalised to produce a probability distribution of ages at which the woman could have given birth to the child. They then randomly imputed an age to the child drawn from that distribution.

Having done this, the next step is to cut an additional notch out of the fertility distribution around this child's imputed age at birth and repeat the process for women's other missing children. The procedure for absent children is similar, except that one multiplies the fertility distribution by the age-specific proportions of children living apart from their mothers in the census to produce the distribution from which to randomly impute the child an age.

Refining the Luther & Cho method

Limitations of the Luther & Cho method:

- women's birth distributions differ, especially if fertility has been changing
- their variance is smaller than that of the population's fertility distribution

Implications: the method attenuates differentials and exaggerates intervals

Solution:

- impute ages to dead/absent children using fertility distributions from the OCM (i.e. the Luther & Cho method)
- fit a regression model of age-specific fertility to the imputed full birth histories (FBHs)
- re-impute the FBHs using the predicted fertility for each sub-group of women
- iteratively **re-fit** the model of age-specific fertility and **re-impute** the missing ages

Model differential child mortality using a regression-based variant of the Brass method

My version of the method extends it to allow for heterogeneity in fertility within the population. Specifically, after imputing an initial set of full birth histories to women using the Own-Children Method, I model fertility in the histories using explanatory variables to capture differentials between groups in the age pattern of fertility. For example, for this study, I modelled differential fertility by region and women's level of schooling.

This allows me to re-impute ages to the dead and absent children using fertility distributions that differ between sub-groups of women. One can then go through the process again. About 5 iterations usually produces a stable set of estimates.

Modelling fertility & mortality

Log-linear rates model with Poisson errors

Fit an Age-Period-Cohort model to the rates, using cubic splines (N.B. omit the linear effect of cohort as it is collinear with age and period)

Also model:

- Spatial variation (e.g. province or district)
- Socioeconomic factors (e.g. woman's schooling)
- Family size – women with above and below average children ever-born for their age
- Interactions of these variables with the linear splines for age and period

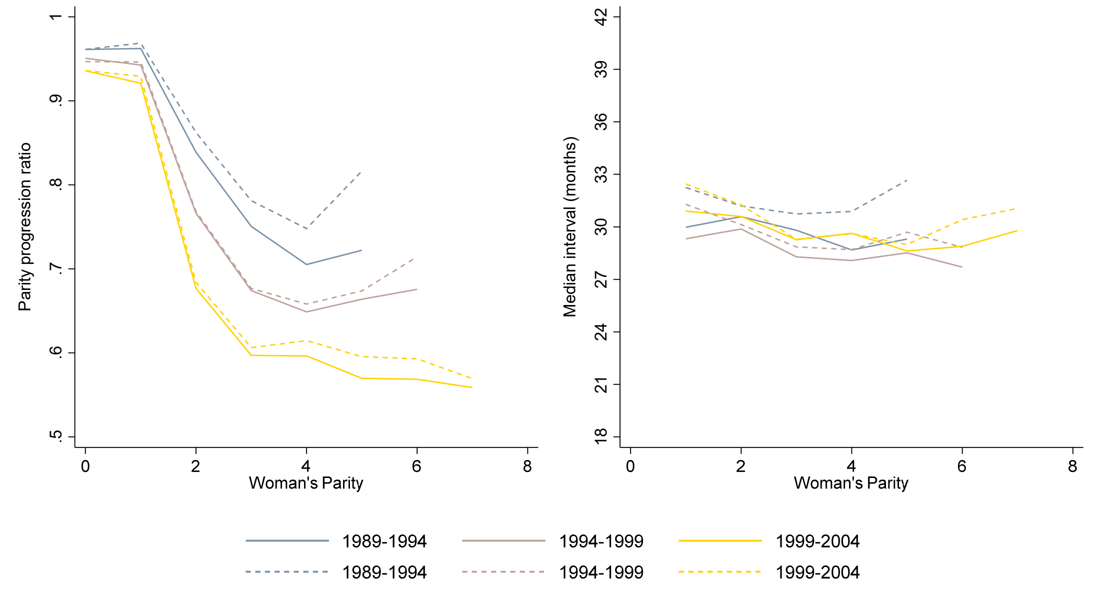
The age-specific mortality of children is modelled using a regression-based variant of Brass's indirect method and the summary birth history data

I modelled age-specific fertility in the birth histories using Poisson regression for rates and fitted an Age-Period-Cohort model to the rates using cubic splines but omitted the linear effect of cohort, which is collinear with age and period.

The estimates of child mortality come from a regression-based variant of Brass's indirect method of estimation applied to the summary birth history data and I also included covariates in this model.

Comparison of iterated and reported estimates

2006 National Family Health Survey of India



This figure compares estimates imputed in this way with those calculated directly from the full birth histories collected by the 2005-6 *National Family & Health Survey* of India. The solid lines were calculated using reported dates of birth. They show estimates of parity progression, on the left, and the median duration of closed birth intervals on the right. The different colours represent three successive five-year periods prior to the collection of the data. The dotted lines were produced by discarding the reported ages of dead and absent children and imputing ages to them using my elaboration of Luther and Cho's method.

The two sets of estimates agree remarkably well with each other. The imputed dataset somewhat overestimates the higher-order progression ratios and intervals in the most distant period, suggesting that rather too many children were imputed to be aged 12-16. Apart from this, the analysis based on the SBH data produces almost identical results to an analysis of the FBHs.

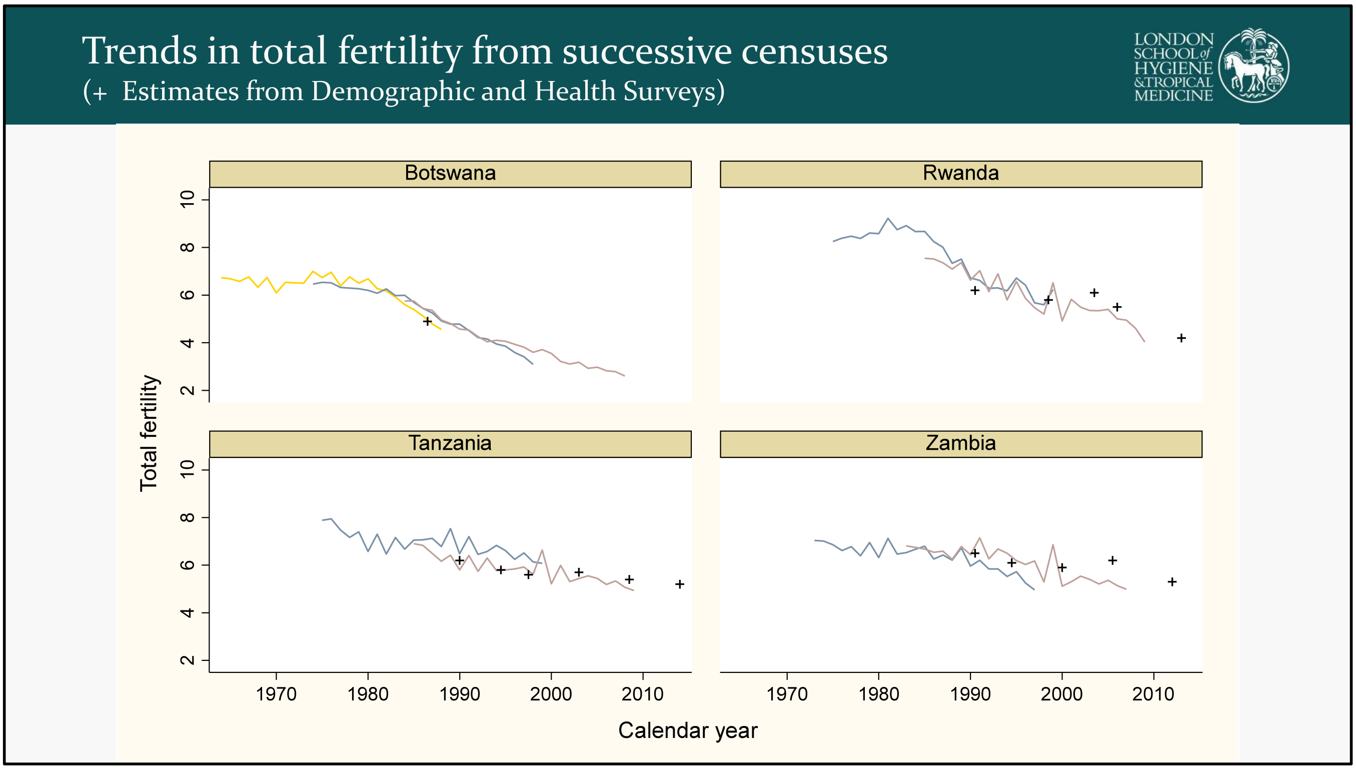
Luther and Cho's method may have dropped off the radar of those studying fertility because they both retired soon after developing the technique. However, it is a method whose time has now come for two reasons. First, thanks to IPUMS, a large number of census microdata samples are now readily available to researchers, many of which include summary birth history data. Second, most demographers now own personal computers that can readily process sample census files of many thousands or millions of records.

IPUMS sample census files

Country	Year	Women (15-79)	Own-children	Dead children	Absent children
Botswana	1991	38,919	86,784	14,665	56,596
	2001	54,607	111,141	16,545	68,651
	2011	68,792	133,961	17,269	77,195
Rwanda	2002	250,290	646,373	219,348	184,244
	2012	319,787	837,935	178,465	242,509
Tanzania	2002	1,089,573	2,654,317	718,778	1,542,961
	2012	1,299,036	3,195,754	626,988	1,858,693
Zambia	2000	271,942	660,400	167,907	254,446
	2010	367,280	975,881	186,420	293,821

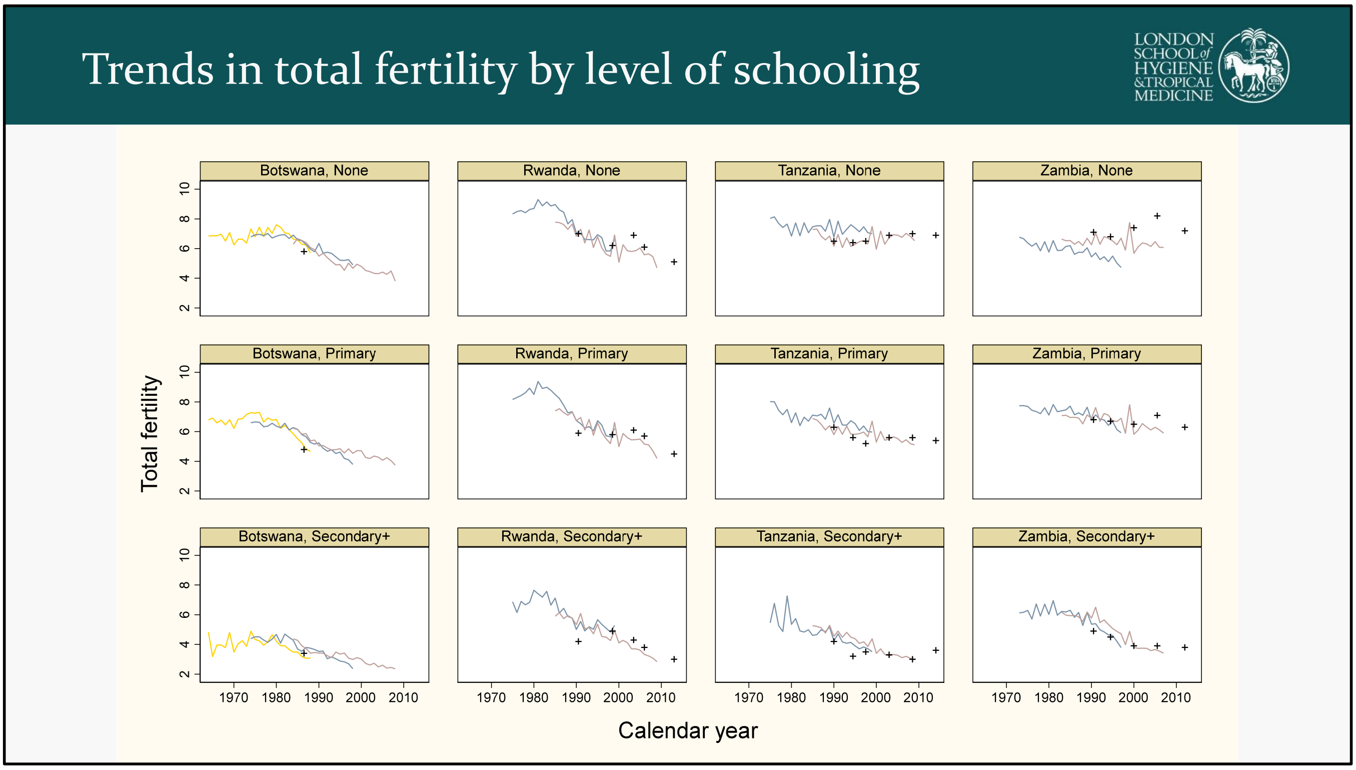
Perhaps some of you are wondering why I am trying to reconstruct women's full birth histories from census data when DHS surveys collect the real thing. The simple answer is sample size! The 2010 DHS in Tanzania collected birth histories from 10 thousand women – the 2012 census sample contains information on 1.1 million women aged <50, more than 100 times as many. Sample census data can provide detailed annual estimates, estimates for small groups such as highly-educated women, and estimates for individual districts and cities. DHS surveys can't be used for any of these purposes.

In my remaining time, I want to illustrate the value of this method using sample census data that I obtained from IPUMS for four African countries – Botswana, Rwanda, Tanzania and Zambia. Each country has deposited data from at least two censuses with IPUMS. This allows me to cross-check the estimates from two different censuses for overlapping dates.



Here I am presenting the series of total fertility rates computed from the different censuses in each country, together with estimates of total fertility from the Demographic and Health Surveys conducted in these countries, indicated by plus signs. During periods when the estimates from the different censuses overlap, they are highly consistent. Moreover, the census-based estimates agree closely with those from the DHS.

Note that the early onset of fertility transition in Botswana in the early 1980s is evident. Rwanda had higher pre-transitional fertility than the other countries but is also the country where fertility has dropped fastest.

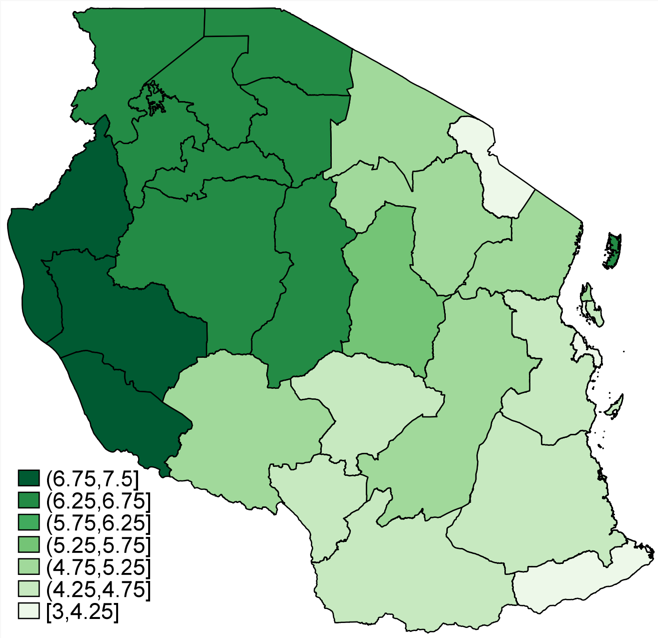


Even after splitting up the women by their level of schooling, the estimates from different censuses remain highly consistent with each other and estimates from the DHS. Only among uneducated women in Zambia is the trend in fertility unclear.

This graph reveals numerous noteworthy differences between the countries. For example, fertility is less differentiated according to women's education in Rwanda than elsewhere. Also, whereas fertility among the small group of women with secondary education began to drop in Tanzania at about the same time as in Botswana and Rwanda, educated women's fertility only began to drop in Zambia a decade later.

Total fertility by region

Tanzania, 2007

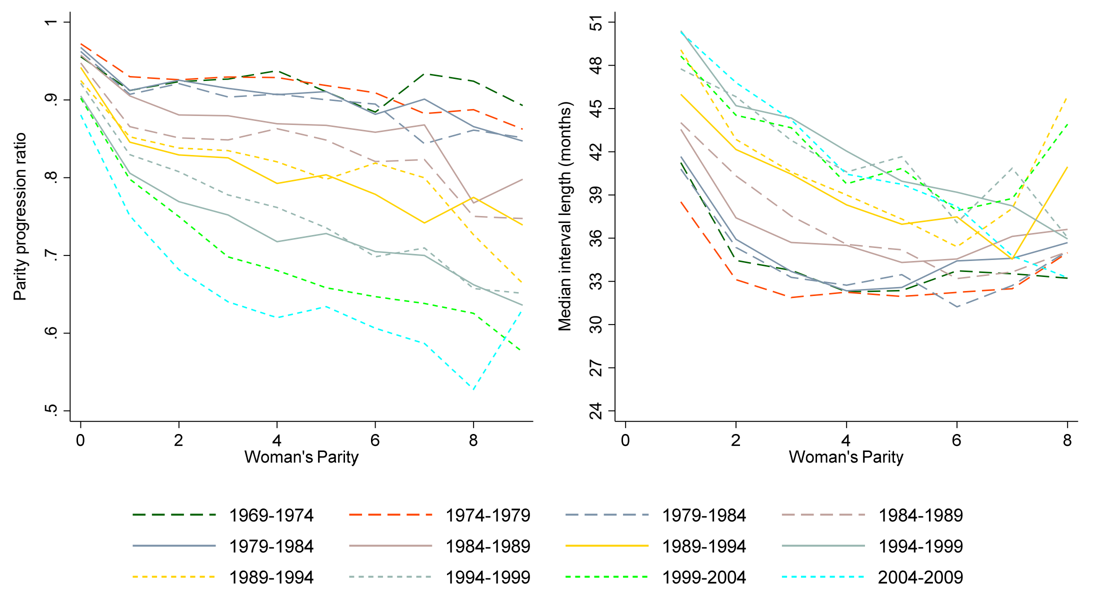


Here I've mapped regional differentials in total fertility in Tanzania. StatCompiler will draw you this map but how the DHS measure total fertility using samples for regions of only a few hundred women is a mystery to me. Certainly, their map looks rather different from mine.

These estimates for about 2007 reveal a clear regional divide within Tanzania: total fertility remained above 6.25 in the whole of the north-western half of the country but had dropped to at least a child less than this in the rest of the country.

Parity progression and median intervals

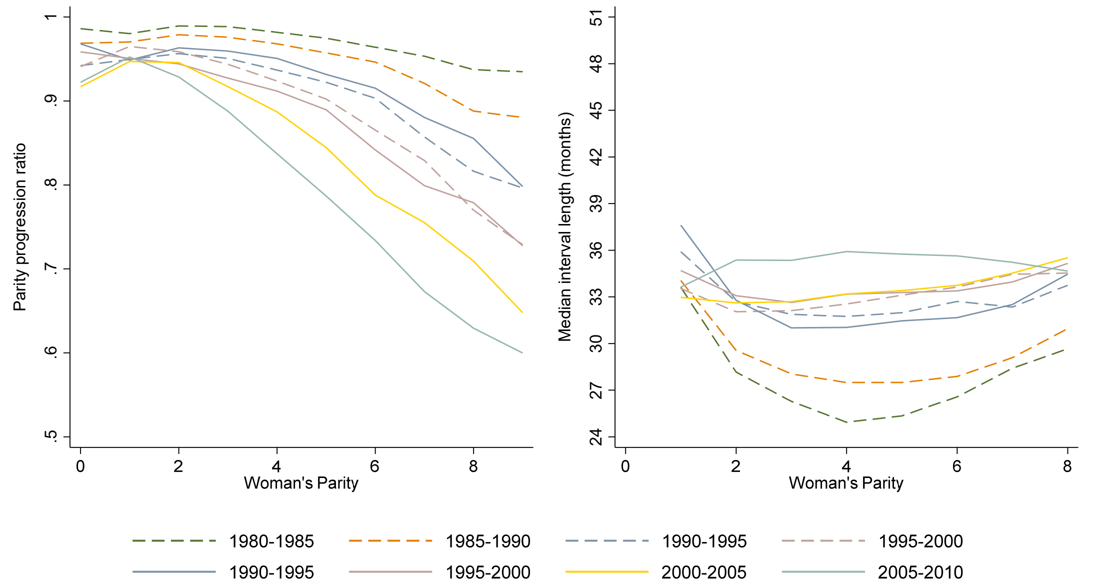
Botswana



This figure is comparable to the one for India. It shows parity progression and median interval length across the different parities for the 3 censuses in Botswana. Lines of the same colour are estimates for the same 5-year period from two different censuses. All these pairs of sets of estimates agree rather well.

There has been a big drop in parity progression. However, while birth intervals hardly changed at all over time in India, in Botswana they lengthened a lot between 1985 and 2000, particularly at the lower parities, before stabilising again. There are also intriguing differences in the pattern of decline in parity progression between Botswana and India. A small but striking increase in the proportion of women remaining childless or having only one child has occurred in Botswana, whereas there was no evidence of this in India.

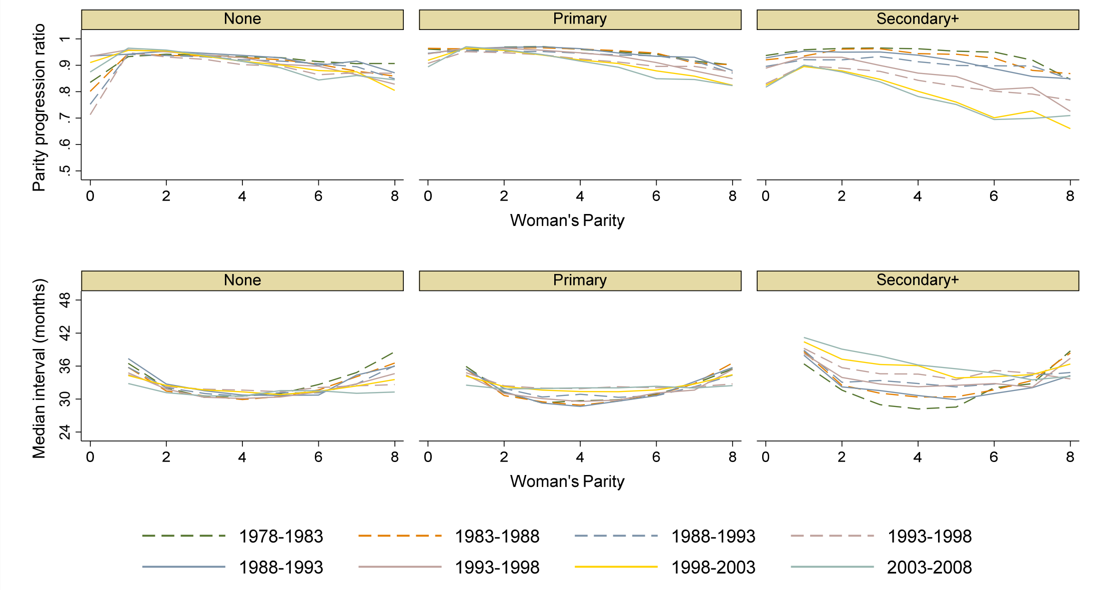
Parity progression and median intervals
Rwanda



This is Rwanda, where only two census samples are available. They again yield rather consistent series of estimates. The country has also seen a large decline in parity progression. It differs from Botswana, however, in that no evidence exists that women are limiting their families to just a few children. Instead, the more children a woman has had, the more likely she is to stop childbearing.

Birth intervals were very short in Rwanda before the transition began, explaining why fertility was exceptionally high. They lengthened until the mid-1990s, but then stabilized for a decade. Even in 2007, intervals were a bit longer than they had been in Botswana 25 years earlier.

Parity progression and median intervals by schooling
Zambia



Lastly, these estimates for Zambia examine parity progression and birth intervals by women's level of schooling. Fertility has only dropped significantly among women with secondary schooling. Moreover, for this group, the drop in parity progression was concentrated in the 1990s and stalled in the first years of this century. Note that, as in Rwanda, the sets of ratios have fanned out over time, with the drop in progression being larger the higher the parity of the women.

Conclusions

It is possible to reconstruct full birth histories from the household rosters of censuses that asked the SBH questions on children ever-born and surviving

Because census samples are large, they are valuable for birth interval analysis and for the study of fertility in small and geographically-disaggregated population sub-groups

India conforms to the conventional stereotype of demographic transition as driven by parity-specific family size limitation; the length of birth intervals has hardly changed

Fertility transition in sub-Saharan Africa is following diverse paths:

- Botswana has adopted family size limitation, but also prolonged postponement of births
- Women in Rwanda and Tanzania are curtailing childbearing, but not yet limiting their families to 2 or 3 children – the propensity to stop childbearing increases with family size
- Fertility transition in Zambia had yet to spread beyond women with secondary education

To conclude, it is possible to reconstruct full birth histories from the rosters of censuses that asked the SBH questions on children ever-born and surviving.

Because census samples are large, they are valuable for birth interval analysis and for the study of fertility in small and geographically-disaggregated population sub-groups.

India conforms to the conventional stereotype of the demographic transition as driven by parity-specific family size limitation; the length of birth intervals has hardly changed.

In contrast, fertility transition in sub-Saharan Africa is following diverse paths. Botswana is characterised by family size limitation, but also by lengthy postponement of births. Women in Rwanda are curtailing childbearing but not yet limiting their families to 2 or 3 children – the propensity to stop childbearing increases with family size. Even in 2010, transition in Zambia had yet to spread beyond women with secondary education.

Lastly, if anyone listening is interested in the potential of this method, there are many more sample census files from all parts of the world available from IPUMS that include summary birth history data.