



<http://www.diva-portal.org>

Preprint

This is the submitted version of a paper presented at *41st Annual Meeting of the Population Association of America: Denver, April 26-28, 2018.*

Citation for the original published paper:

Broström, G. (2018)

The effect of socioeconomic status on marital fertility during the demographic transition, northern Sweden 1821--1950: Efficient data analysis with process point of view

In:

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-173128>

The effect of socioeconomic status on marital
fertility during the demographic transition,
northern Sweden 1821–1950: Efficient data analysis
with process point of view¹²

Göran Broström
Centre for Demography and Ageing Research
Umeå University, Umeå, Sweden

19 Apriln 2018

¹Paper presented at the 41st Annual Meeting of the Population Association of America: Denver, April 26–28, 2018.

²email: goran.brostrom@umu.se

Abstract

The impact of the family's socioeconomic status at marriage on later child births during the demographic transition (1821–1950) is studied. It is found that the fertility decline starts in the upper classes in the decades prior to 1900. The farmers are characterized by relatively high fertility in high ages even in the end of the study period. A stopping behavior seems to dominate a spacing one, especially along cohorts. The effect of declining infant mortality over time is minor. We show how efficient statistical modeling leads to easy and fast estimation of rather complicated data. The tools are statistical sufficiency and data reduction combined with models for fertility behavior, stopping and spacing.

1 Introduction

This research question is not new. Mason (1997) discussed contemporary explanation models of the fertility transition and argued that there cannot be one explanation alone. More recent studies of Swedish circumstances were given by Dribe and Scalone (2014) and Bengtsson and Dribe (2014). One perhaps convincing theory is that the fertility decline was a necessary result of the mortality decline. For a detailed and initiated overview of the research field, see the aforementioned papers and also two thoughtful papers by Ní Bhrolcháin (Ní Bhrolcháin, 1992, 2011).

2 The study area

Our data cover the time period 1821–1950, see Figure 6, and the geographical area defined by the Skellefteå and Umeå towns with surrounding rural parishes. Where in Sweden? See map in Figure 1.



Figure 1: The Skellefteå and Umeå regions in Sweden.

The Skellefteå region was one big parish in the eighteenth century. It successively split up in smaller parishes.

The development of population size in the two regions is shown in Figure 2.

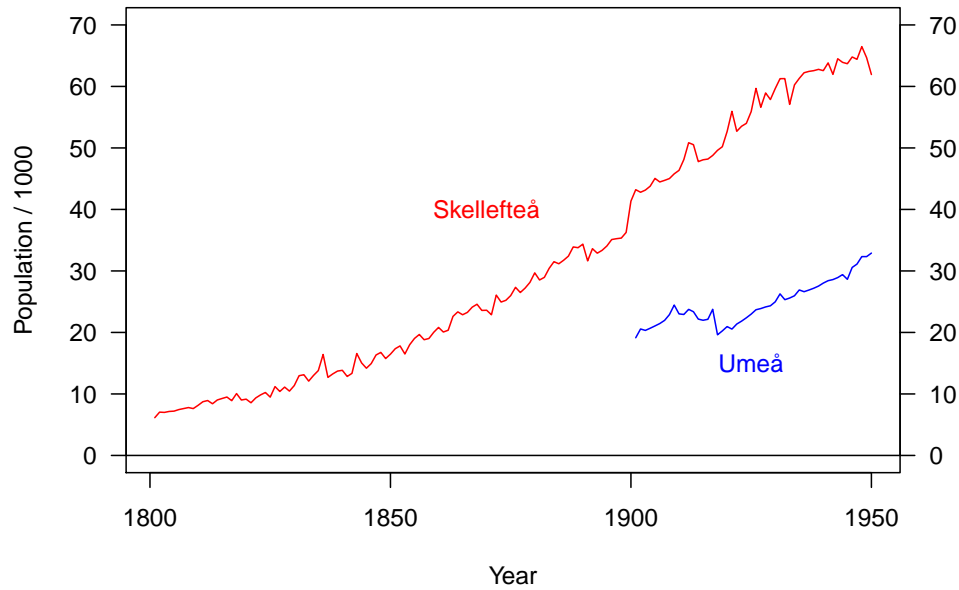


Figure 2: Population size by year and region.

2.1 Skellefteå

At the outset of the study the region consisted of one large rural parish, Skellefteå. By 1900 three new parishes had been detached into separate units, but their populations are still included in the study. The region was large, both in terms of area and of population. With an area of about 1700 square miles, Skellefteå was considerably larger than most rural parishes in Sweden. It was considered a one-day journey to travel from the northern to the southern border, and a ride from the coast to the more remote and sparsely populated parts of the parish in the west could take even longer, especially in wintertime. The main part of the population was, however,

concentrated in the coastal area and in river valleys. In the early 19th century the population size was around 6900, and it increased rapidly during the first half of the century. By 1850 it had reached to about 17000 and at the turn of the century it had further doubled. Despite the large increase in population, which was mainly the result of a high natural growth, the population density on the whole remained low (Alm Stenflo, 1994).

Skellefteå was during the studied period a rural area with a mixed economy, based on animal husbandry, forestry and sidelines such as tar and saltpeter production. By the mid-19th century export of tar and lumber became an increasingly important part of the economy. The majority of the farmers in the region were smallholders and there were no large estates. Some small sawmills were established early in the century, but before 1900, industrialization had little impact on the local economy. In 1835, approximately 85 percent of the population made their living from farming. Although the distribution of economic resources was more uniform than in several other Swedish regions, the social stratification became more pronounced throughout the 19th century. The increasing proletarianization was mainly a consequence of rapid population growth. The number of farming households remained fairly stable, while the number of landless households increased. The socio-economic development was also influenced by two devastating subsistence crises in the region, in the 1830s and in the 1860s (Engberg, 2005).

Infant mortality was comparatively low. Fertility was high, not only by Swedish standards, but also in European comparison and there are no indications of family planning. Total fertility fluctuated around five children per woman and, although fertility did decline during the nineteenth century, the actual fertility transition occurred late in the district (Alm Stenflo, 1994; Coale and Watkins, 1986). The rate of illegitimacy was low in comparison with many other parts of Northern Sweden, where frequent pre-nuptial conceptions and illegitimate births were common. The illegitimacy rate fluctuated between three and six per cent during the nineteenth century (Alm Stenflo, 1994).

2.2 Umeå

The Umeå region is one of the newest in the collection of regions digitized at the Demographic Data Base, Umeå University. The covered time period is shorter than the one for the Skellefteå region in that registration starts

around January 1, 1901.

Umeå is older and substantially larger than Skellefteå as a town: The birth year of the town Umeå is generally recognized as 1622. Umeå became early a center for administration and education, but Umeå University is not a fact before 1963.

Early in the twentieth century, two military regiments were placed in Umeå, and they remained there during the first half of the century. The population size as defined in our data sets was 19000 on January 1, 1900, 33000 on January 1, 1950 and 104000 when the end of the century.

The surrounding rural area is similar in demography and economy to the corresponding parts of the Skellefteå region.

3 Models

The statistical modeling is conveniently done in the framework of counting (birth) processes, but there are still choices to contemplate: Choice of time scale, and how to handle the dynamic aspects of the counting process. In most of the earlier attempts to investigate changes in *stopping* and *spacing* behavior over time, models for *birth intervals* have been employed (Bengtsson and Dribe, 2006; Dribe and Scalone, 2014; Bengtsson and Dribe, 2014; Bras, 2014), but in my opinion this approach is too indirect. Instead I argue for the *counting process* approach, where each married woman contributes an age interval (or a union of such intervals), starting at age of marriage or age 20, whichever comes *last*, and ends with the dissolution of marriage or age 50, whichever comes *first*. We also allow intervals to start and end with migration events. In the so defined interval, birth time (age) points are recorded.

An obvious alternative for studying marital fertility is to start the clock at the *date of marriage*, but it complicates things in some ways, one being that it will be difficult (but not impossible) to apply the Coale-Trussell model (Coale and Trussell, 1978). Therefore, in the following, the basic time scale is *age*.

3.1 Fertility as a counting process

We define

$$\{N_i(t), m_i \leq t \leq 50\}, \quad i \in \text{mother's ID}$$

as the *number of births before and including age t and after marriage at age m_i* for woman No.

i. By age we mean *exact age*, measured by a precision of a day, but with *time unit year*. Thus $\{N_i, i = 1, 2, \dots\}$ are *counting processes* with jumps of size one (except for multiple births) at the age of the mother at deliveries. As an example, see Figure 3 for the marital fertility history of mother No. 233: She married at age 25, died at age 46, and in between she had seven births. See (Aalen et al., 2008) for details on the counting process theory.

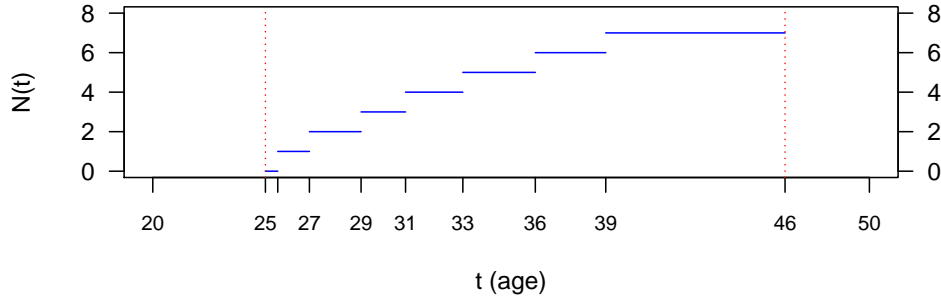


Figure 3: Birth history as a counting process, woman with id 233.

$N(t)$ is a *right-continuous* function, meaning essentially that at jump points, the value is the larger of the two possible values. $N(t-)$ denotes its *left-continuous* counterpart, that is, it is defined as the number of births up to but *not including* age t .

3.2 The Coale-Trussell (CT) model

We will assume a piecewise constant hazard model on the age span 20–50, with jumps every fifth year. We apply the restriction proposed by Coale (1971) and Coale and Trussell (1978):

$$\lambda(t) = Mn(t)e^{mv(t)}, \quad 20 \leq t < 50,$$

where the piecewise constant functions n, v are given by

	(20-25]	(25-30]	(30-35]	(35-40]	(40-45]	(45-50]
$n(t)$	0.46	0.431	0.395	0.322	0.167	0.024
$v(t)$	0.00	-0.279	-0.667	-1.042	-1.414	-1.671

The parameter m governs the form of the (cumulative) hazard function, while M measures level. Figure 4 shows two cases, $(m, M) = (0, 1)$ and $(m, M) = (1, 1.57)$. These two cases generate the same expected number of births over a full reproductive period (20–50), that is, the same *Total Marital Fertility Rate* (TMFR). The first is an example of *spacing*, but no *stopping*, while the second is a case of *stopping* (and eventually *spacing*).

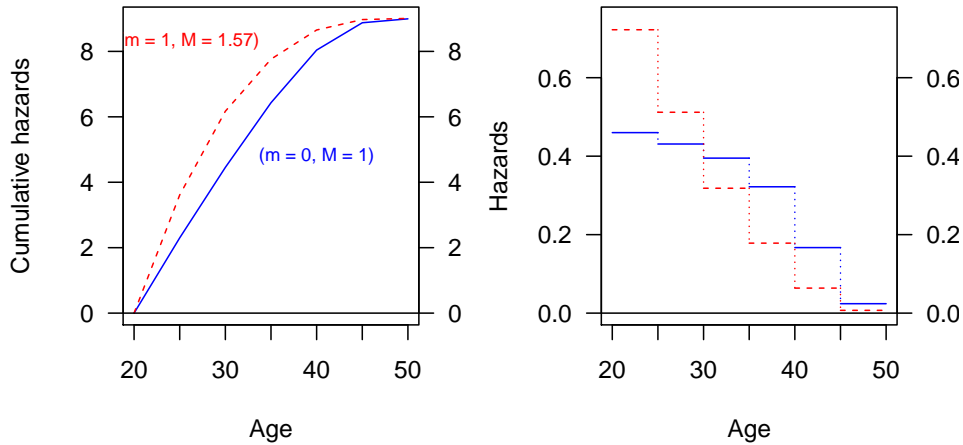


Figure 4: The Coale-Trussell (CT) distribution, two examples: $(m = 0, M = 1)$ is solid, $(m = 1, M = 1.57)$ is dashed.

The parameter M measures level, but note that keeping M fixed and increasing m *lowers* the TMFR.

How realistic is the CT model? We compare its fits for two populations, one with an m close to zero, and one with m close to one, and we compare it to a piecewise constant hazard (PCH) model with the same interval partitioning as in the CT but with freely varying levels. See Figure 5.

The fit seems to be satisfactory, even if a formal test would reject the CT model in favor of the PCH.

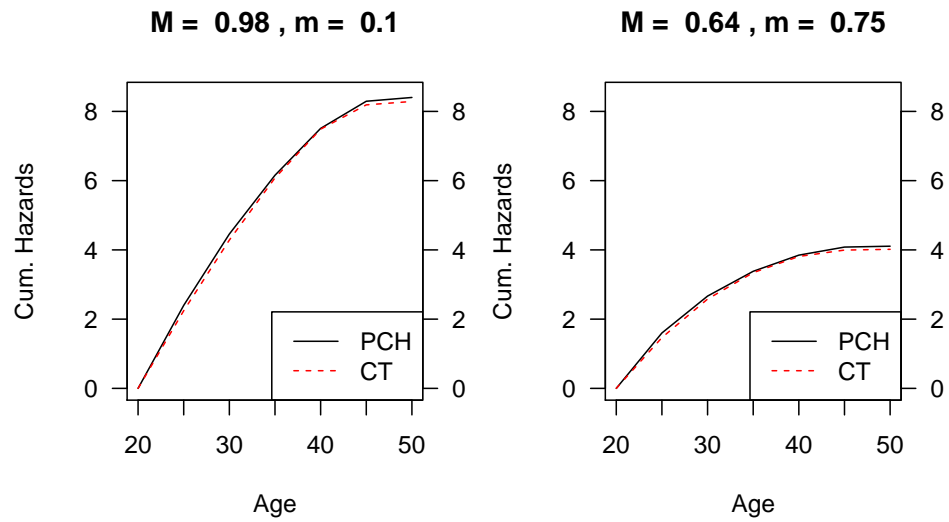


Figure 5: Evaluation of the Coale-Trussell (CT) model, compared to a piecewise constant (PCH) model. Left panel from the time period 1851–80, the right panel from 1931–40. Non-farmers.

4 Data

Data come from The Demographic Data Base, Centre for Demographic and Ageing Research (CEDAR) at Umeå University, Sweden, and cover the Skellefteå region, situated in the north of Sweden (Westberg et al., 2016), and the Umeå region, situated 100 km south of Skellefteå. All women, ever married and born between 1 January 1801 and 31 December 1935, are included in the data set and followed throughout their fertility period, but not after December 31, 1950. For each woman, the exact timing of her deliveries (if any) after the onset of the first marriage and after age 20 are recorded.

Figure 6 shows the sampling frame. The parallelogram is the basis for *cohort* analysis, and the rectangle correspondingly for *period* analysis.

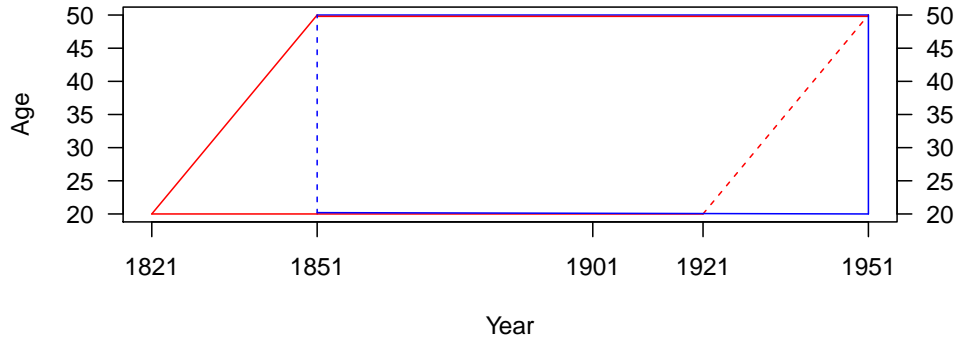


Figure 6: The sampling frame (solid lines) by married woman's age.

4.1 Variables

The following variables are part of our analysis.

- **region** Birth region, Skellefteå or Umeå.
- **urban** TRUE if birth in one of the towns Skellefteå and Umeå, otherwise FALSE.
- **parish** one of *Umeå (town)*, *Umeå lands (rural)*, *Skellefteå (town)*, *Skellefteå lands (rural)*, *Byske*, *Bureå*, *Norsjö*, *Jörn*, and *Malå*.

- **birthdate** A real number indicating mother's birth date.
- **ch.birthdate** The birth date of the child.
- **deathdate** The death date of the child.
- **enter** Mother's age at previous child, marriage, first seen, or age 20.
- **exit** mother's age at birth of current child or date last seen or age 50.
- **event** Indicator of a birth at **exit**.
- **event1** Indicator of a birth at **exit** that survives age one.
- **lopnr** Birth order among *observed* births between ages 20 and 50.
- **antrec** Total number of records for current woman.
- **hisclass** Hisclass with the following collapsed categories.
 - **upper** The cream of the society.
 - **farmer** Farmers, by far the largest group. Decreases in relative size over time.
 - **lower** Not so skilled workers.
- **cohort** Birth cohort of mother.
- **period** Calendar time period.
- **marAge** Age at marriage.
- **age** Mother's age at child birth grouped in five-year-intervals.
- **exposure** Interval length.
- **ind.death** Did the born child die before age one?
- **n** Vector of constants from the Coale-Trussell model.
- **v** Vector of constants from the Coale-Trussell model.

Central covariates are family **hisclass**, mother's age at marriage, and mother's **birthdate**. The rows corresponding to woman No. 233 in our data file are, with selected columns,

	parish	hisclass	birthdate	enter	exit	event	inf.death	marAge
1	Norsjö	3	1834-05-21	24.715	25.232	TRUE	FALSE	(20-25]
2	Norsjö	3	1834-05-21	25.232	26.557	TRUE	FALSE	(20-25]
3	Norsjö	3	1834-05-21	26.557	28.720	TRUE	FALSE	(20-25]
4	Norsjö	3	1834-05-21	28.720	30.582	TRUE	FALSE	(20-25]
5	Norsjö	3	1834-05-21	30.582	32.980	TRUE	FALSE	(20-25]
6	Norsjö	3	1834-05-21	32.980	36.014	TRUE	FALSE	(20-25]
7	Norsjö	3	1834-05-21	36.014	38.962	TRUE	FALSE	(20-25]
8	Norsjö	3	1834-05-21	38.962	46.468	FALSE	FALSE	(20-25]

Woman No. 233 got married at age 24 (years), was followed until age 46.468 when she died, and she had seven births. The last interval (almost) always

ends with no birth (`event = FALSE`): Women are followed until age 50, end of marriage, or lost to follow-up, whichever comes first. The first interval starts at age at marriage, or age at first observation as married, or age 20 and married, whichever comes *last*. Note that childless marriages are included.

Figure 7 shows the distribution (per cent) over socioeconomic status and time periods of exposure time for married women in the sample. The socioeconomic status is measured by the occupation of the head of the household (a man). The big shift is that the fraction farmers is more than halved during this century. All the other classes are increasing in size.

4.2 Sufficiency and data reduction

Since all or interesting covariates are, or can be transformed into, categorical variables, *aggregating* the variables `exposure`, `event`, and `event1` over all possible combinations of values of the other covariates creates multidimensional tables that can be stored in a compact way. The table of explanatory variables is

	period	cohort	marAge	hisclass	parish	region	urban
1	1821-1850	1821-1840	(15,25]	upper	Skellefteå.land	ske	FALSE
2	1821-1850	1821-1840	(15,25]	upper	Skellefteå.land	ske	FALSE
3	1821-1850	1821-1840	(15,25]	upper	Skellefteå.land	ske	FALSE
4	1821-1850	1821-1840	(15,25]	upper	Skellefteå.land	ske	FALSE
5	1821-1850	1821-1840	(15,25]	upper	Skellefteå.land	ske	FALSE
6	1821-1850	1821-1840	(15,25]	upper	Skellefteå.land	ske	FALSE

and the table of variables more or less connected to the response is

	age	event	event1	exposure	N	V
1	(20-25]	34	26	84.886178	0.460	0.000
2	(25-30]	60	51	141.807487	0.431	-0.279
3	(30-35]	21	19	85.802864	0.395	-0.667
4	(35-40]	7	6	42.824514	0.322	-1.042
5	(40-45]	2	2	23.241963	0.167	-1.414
6	(45-50]	1	1	7.923488	0.024	-1.671

In Figure 8 is shown how the *age-period-cohort* plane is split up for each combination of the levels of the explanatory variables.

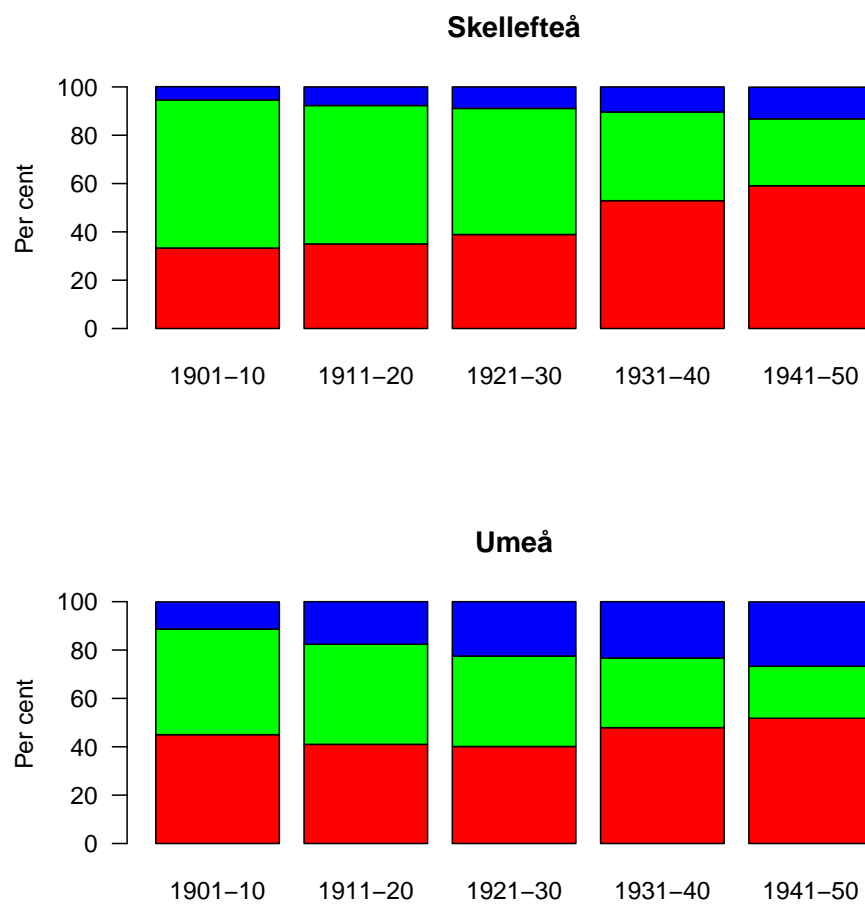


Figure 7: Exposure by social class and time period. From top to bottom: Upper class, Farmer, Lower class.

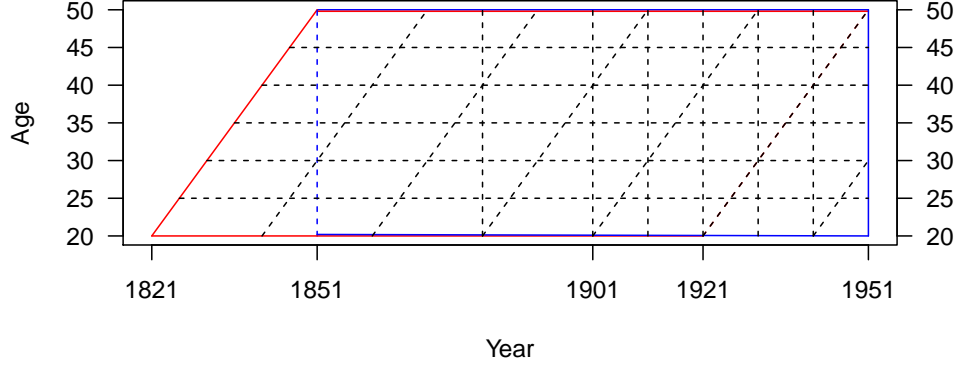


Figure 8: Lexis split. In each sub-area the number of births and total exposure time are calculated.

There are 225,792 unique combinations of the levels of the explanatory variables and the age-period-cohort combinations, but many of them are either structural or random zeros, so the total number of observed combinations is only 6228. The original data frame contains 300,579 records.

Thus, the analyses of our problem gain hugely in speed by using the aggregated tables, and further, it makes it very easy to implement the Coale-Trussell model. The worries are: Are we losing information? Will our results be biased?

The answers to these questions are: Given that the model we are fitting, and given the original data, the tables we create form *sufficient statistics*. This means that, given our tables, the extra information in the original data is pure noise. It will *not* improve our results.

That is the good news. The back side is that the “pure noise” can be used to judge model fit. If an analysis shows that it is *not* pure noise, but perhaps is related to a (subset of) our data, then we are in trouble. So, the advice is: perform all the estimations and testing in the tabular world, and evaluate model fit with the original data. It saves a huge amount of time compared to doing everything with the original sample.

In our specific case, with the piecewise constant hazards model, we can

utilize the fact that the likelihood function that is given by a survival analysis is exactly the same as if we interpret the number of births in each of our cells as following a *Poisson* distribution and write down the resulting likelihood function. But note carefully that this does *not* imply that the numbers of births really follow a Poisson distribution (they do not). Thus, we can utilize ordinary software for Poisson regression and get very efficient and fast parameter estimation. The (small) price we have to pay is that for instance estimates of the baseline hazard function are not coming automatically but require some thought.

These ideas are not new, see e.g. Broström (1985). Long before software for survival analysis (read: Cox regression) was commonly available, in the late 1970s, Jan M. Hoem taught us how to do this with software analyzing log-linear models for contingency tables, see Hoem (1987), Hoem and Mureşan (2011) and Laird and Olivier (1981). This is interesting: Because there were no good statistical software for survival analysis around 1980, innovative methods like utilising the fact that the log-linear models work outside their so-called natural domain were discovered. When later new software appeared for direct survival analysis these methods were forgotten (more or less). Now they are found useful again, but for a different reason: We are analyzing larger and larger data sets, and raw calculation ties up too much time and other resources, so by thinking and remembering the theory about sufficient statistics, we can do things much faster and with much smaller, and anonymous, data sets (tables).

5 Analysis, all births

5.1 Stopping and spacing

We consider the development of a stopping behavior both by period and by cohort. It is measured by the 'm' of Coale and Trussell (1978). Similarly, the spacing behavior is measured by their 'M'. We show the development by period and by cohort, in both cases for each `hisclass`, adjusted for `marriage age` and `parish`.

The results are only shown graphically, and the reason is that there are many three-way interactions involved, see the table below.


```

Model:
event ~ offset(log(exposure * N)) + V * (marAge * period + marAge *
      parish + period * parish)
      Df Deviance   AIC    LRT Pr(>Chi)
<none>          6919.2 24224
V:marAge:period 21   6960.8 24224 41.524 0.004826 **
V:marAge:parish 24   6961.4 24218 42.128 0.012485 *
V:period:parish 47   6992.3 24203 73.034 0.008844 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

The variable V has the regression coefficient m , Coale-Trussell's "little m ". This model contains no less than 240 estimated coefficients: You do not want to see them in print, and I certainly do not want to show them. Instead, in what follows, the results are graphically reported. The development over time may be driven by time periods or cohorts (or both).

5.1.1 By period

From Figure 8 is seen that there are two time periods in the nineteenth century, 1851-1880 and 1881-1900, and five ten-year-long periods in the twentieth century. For each period, and for each `hisclass`, a Poisson regression model is fitted as follows (exemplified with the period 1931-1940 and `hisclass = 3`):

```

fit <- glm(event ~ offset(log(N * exposure)) + V + urban + marAge,
  data = kTab, family = poisson,
  subset = (hisclass == "farmer" & period == "1931-1940"))

```

	Estimate	Std. Error	p-value
log(M)	-0.3716	0.0263	0.0000
V	0.4987	0.0349	0.0000
urban	-0.2223	0.1182	0.0513
marAge			0.0000
(20,25]	0		
(25,30]	0.1534	0.0319	
(30,35]	0.2745	0.0515	
(35,50]	0.4675	0.0957	

From this we get that $\hat{M} = \exp(-0.3716) = 0.690$ and $\hat{m} = 0.499$, that is, strong signs of both spacing (small \mathbb{M}) and stopping (large \mathbb{m}). Note that $M = \exp(\text{Intercept})$ is evaluated for the reference categories, *rural* residence and *marriage age* (15,25].

This is then repeated for all hisclasses and all periods, and the results are shown in Figure 9, where it is clear that the **upper** class is introducing a *stopping* behavior already in the late nineteenth century (lower panel) and also leading in lowering the overall level, or *spacing*. Farmers are slow in both respects.

5.1.2 By cohort

Then we show the development by cohort and hisclass, see Figure 10. Now the spacing behavior disappears completely, while the stopping behavior is a little more accentuated compared to the period analysis.

An explanation to this may be that the CT model fits best to cohort data, and when looking at period data, a mixture of cohorts in different stages of stopping is resulting.

5.2 TMFR

We repeat the analyses with *TMFR* replacing (M, m) .

Figure 11 shows the TMFR by hisclass and period. The elite and the middle class are slightly ahead in the fertility decline, while farmers and unskilled workers lags behind. There are no big differences at all in the last period, 1941–1950.

The cohort analysis, on the other hand, gives a weak differential between hisclasses, even if the elite and the middle class are slightly ahead, see Figure 12.

A drawback with the TMFR is that it gives an overly optimistic view of the marital fertility and number of births per woman (Hoem and Mureşan, 2011). With the cohort data a radical solution is available: Count the number of births in a cohort and relate that number to the total number of women contributing to the cohort. The result is seen in Figure 13.

As expected, the overall levels are much lower than the TMFR, but what stands out is the farmers' relatively high level. An partial explanation to

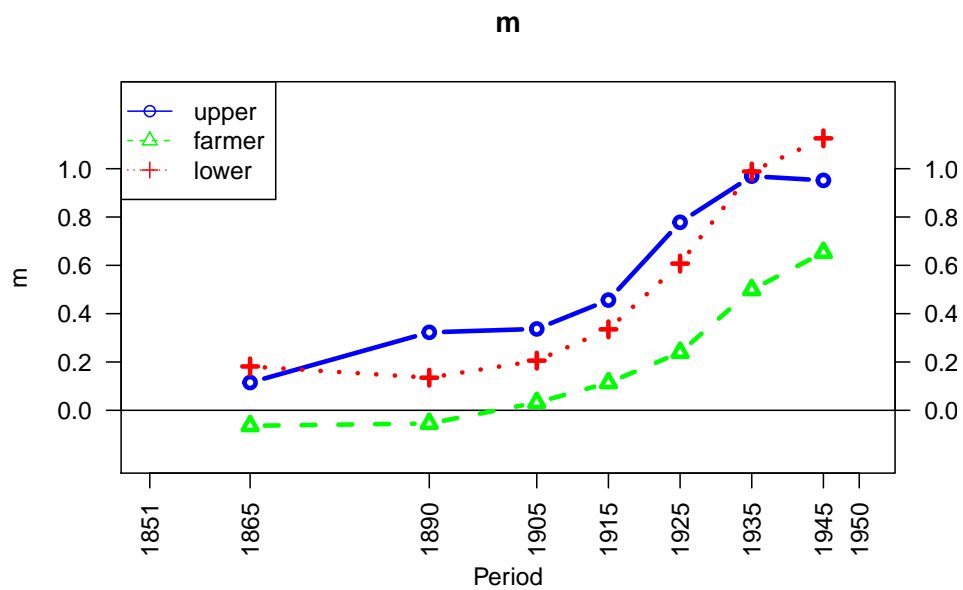
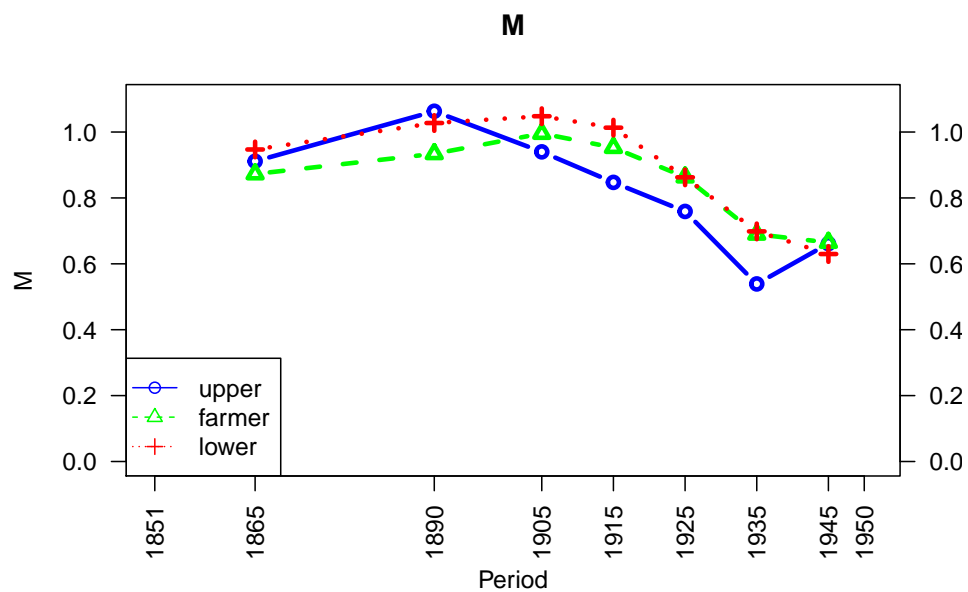


Figure 9: Coale-Trussell's M and m by hisclass and period.

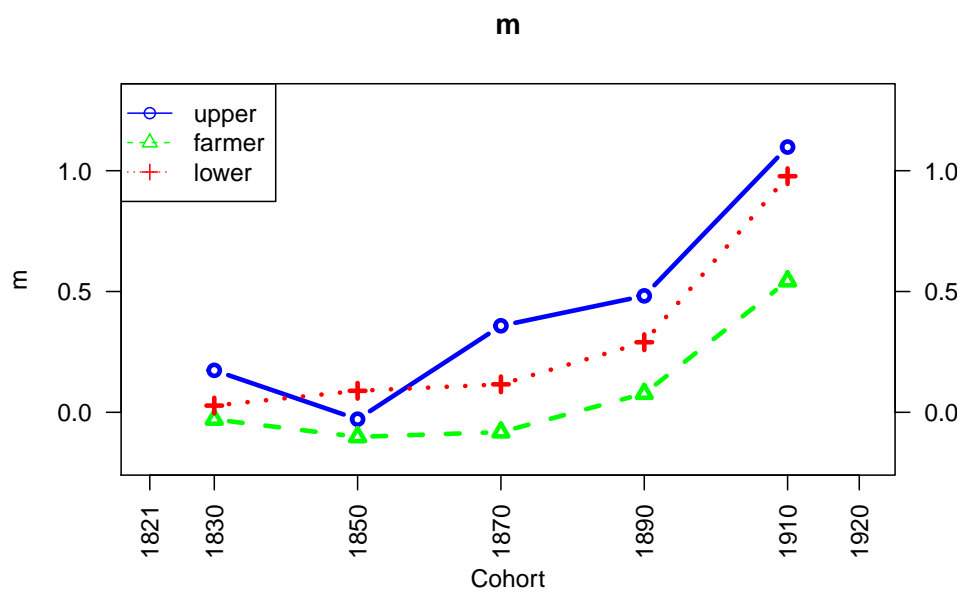
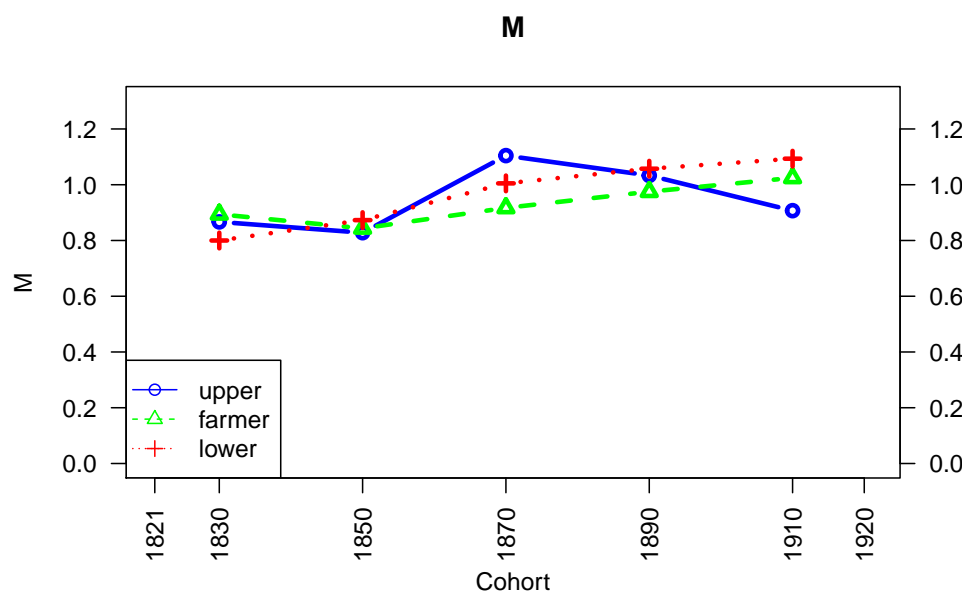


Figure 10: Coale-Trussell's M and m by hisclass and cohort.

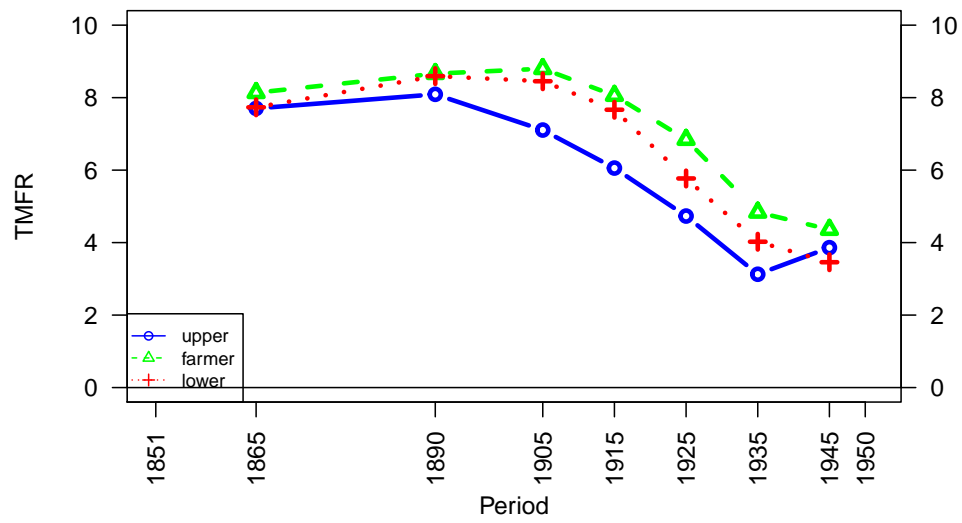


Figure 11: TMFR by hisclass and period.

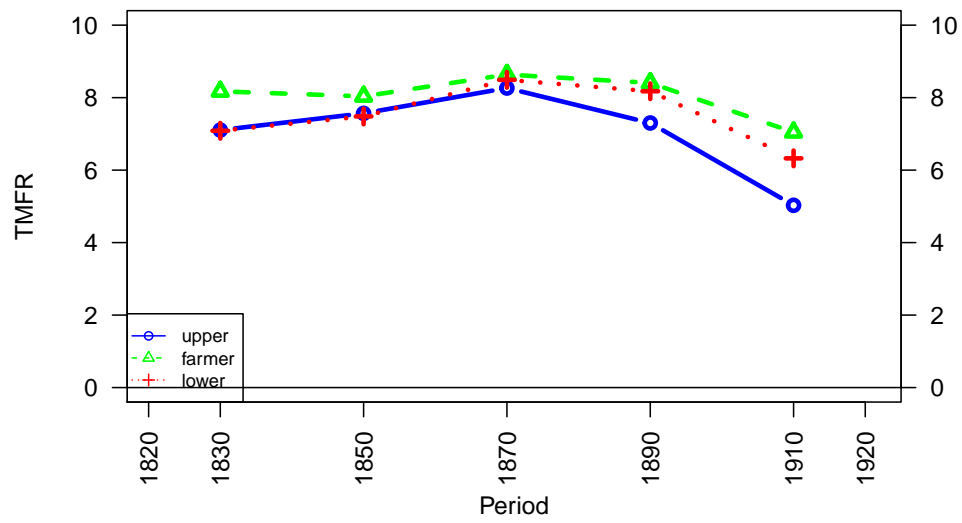


Figure 12: TMFR by hisclass and cohort.

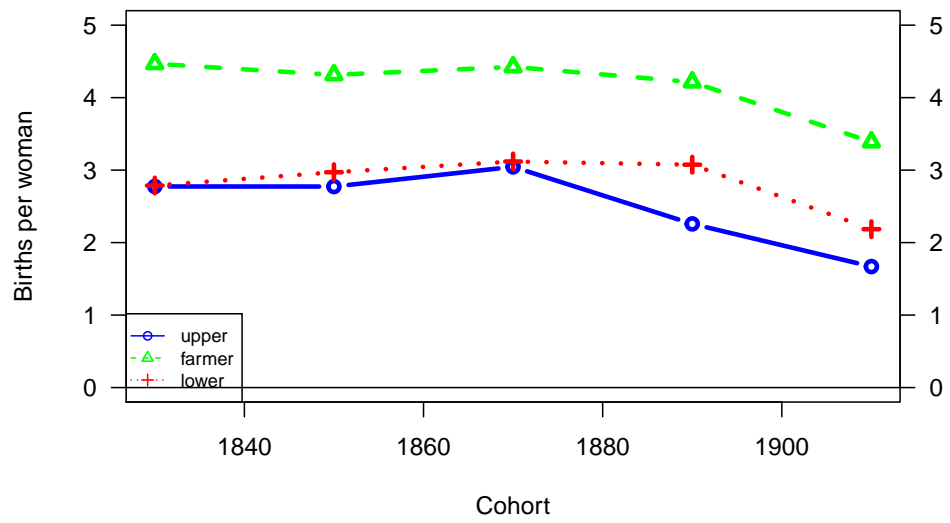


Figure 13: Number of births per woman by hisclass and cohort.

this may be that the average fertile period length per married woman is 18.2 years for farmers, but only 14.7 years for non-farmers. And that may be explained by the fact that farmers on average are more stationary than others.

6 Analysis, surviving births

We define (for now) a *surviving birth* as a birth of a child that survives at least one year after birth. Technically, in the counting process framework, the time point of such birth is *not a stopping time* (we need to wait one year after a birth to know if it can be classified as a success), but it is a minor problem in this context. If we instead define an event time as the time when an infant turns into a child (at the first birthday), we have a stopping time, and subtracting one does not change that fact; we can regard it as a change of time scale. Of course, it may happen that it is impossible to decide if a birth results in a child one year later, but it is quite rare in our data, and we chose to ignore it: such a birth is regarded as a survivor.

The Ccoale-Trussell model will once again be used in the analysis, which is a repetition of the one in the previous section, but with the new version of a birth. If the infant death is independent of parity, we would expect that m would be the same but M would be reduced.

With the new data set, where `event` is replaced by `event1` (indicating a birth that survives for one year), we get the results presented in the rest of this section.

6.1 Stopping

First, we consider the development of a stopping behavior both by period and by cohort, see Figure~14.

Then we repeat the cohort approach with surviving births, see Figure 15.

Compared to the corresponding figures with all births, nothing much changes, which would imply that a stopping behavior later in time is not a consequence of decreasing infant mortality.

The TMFR for survivors are shown in Figure 16. Here we can see slightly lower levels than in the case with all births, which is quite natural since the TMFR directly measures level in contrast to Coale-Trussell's m .

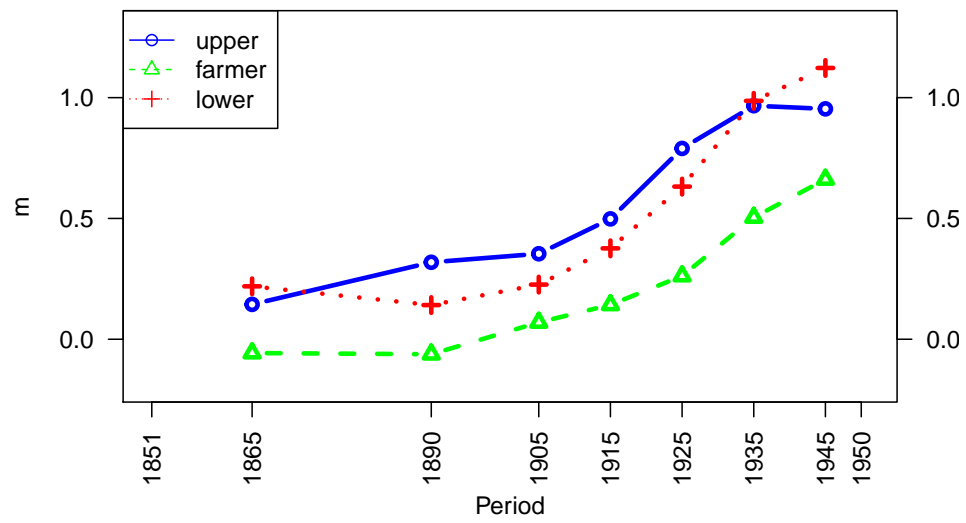


Figure 14: Coale-Trussell's m by hisclass and period, surviving births.

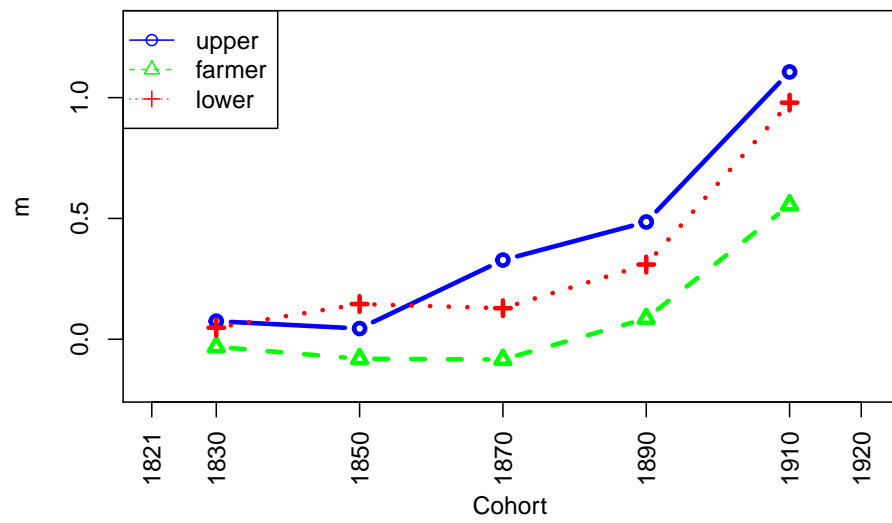


Figure 15: Coale-Trussell's m by hisclass and cohort, surviving births.

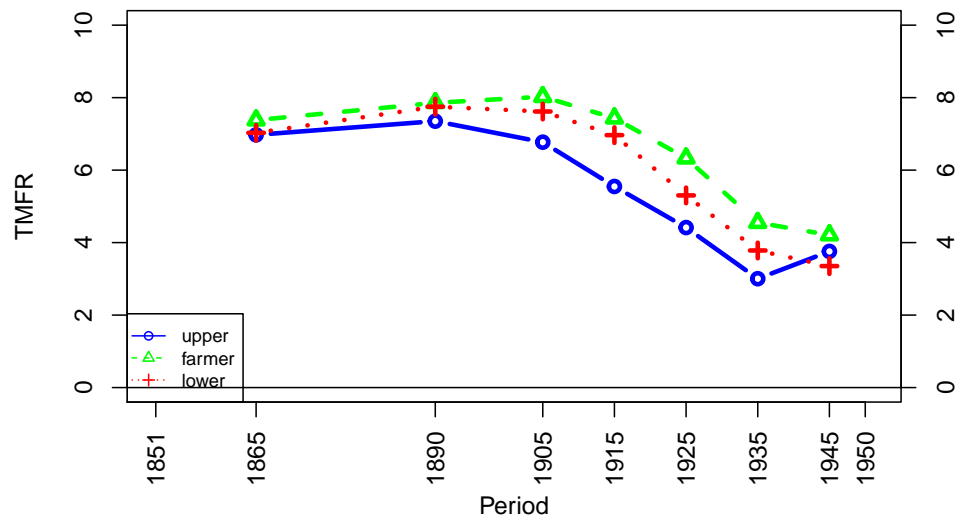


Figure 16: TMFR by hisclass and period, surviving births.

Finally, we show the expected number of surviving infants per woman in Figure~17. Nothing spectacular, just slightly lower levels compared to the case with all infant births per woman.

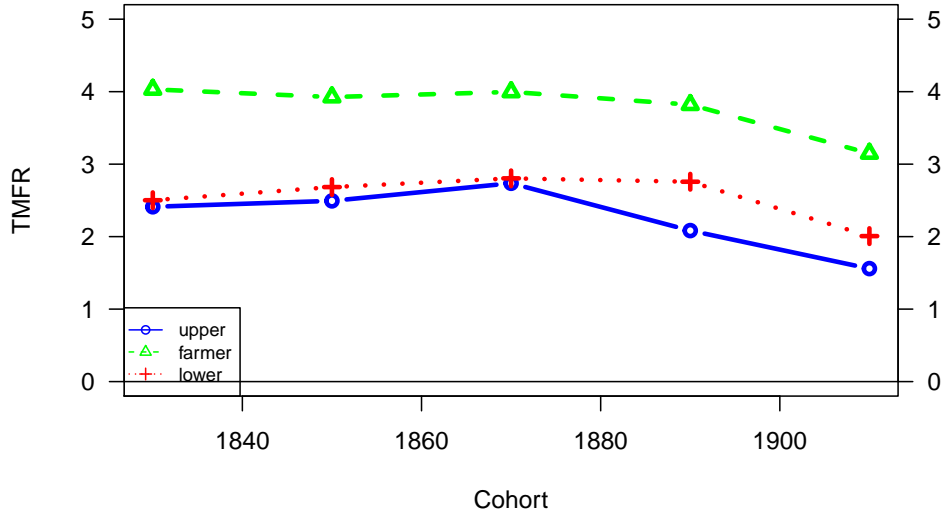


Figure 17: Number of surviving births per woman by hisclass and cohort.

7 Conclusion

The fertility decline starts in the upper classes in the decades prior to the turn of the century. The farming part of the population is characterized by a relatively high fertility in high ages until the end of the study period. It is also by far the largest group in the society, near 45 per cent are in the farming category on average over the study period. Towards the end of the period this share is decreasing markedly.

The question of whether it is spacing or stopping that dominates the behavioral explanation of the fertility decline over time is not definitively settled, but the stopping measure, Coale-Trussell’s “little m ” shows a more dramatic change over time than the corresponding spacing measure “big M ”. This is especially true in the cohort setting.

The effect of the declining infant and child mortality on the stopping or spacing behavior in fertility is minor.

A topic for future research is the interplay between mother’s age and duration in marriage. It is intuitively very reasonable to use “time in marriage” as the basic time scale instead of “age”, but how that would effect the spacing and stopping models is still in the open. One option is to utilize a two-dimensional basic time scale.

This work was done in the **R** environment for statistical analysis (R Core Team, 2017), mainly with the package **eha** (Broström, 2017, 2012).

References

- Aalen, O., Borgan, O., and Gjessing, H. (2008). *Survival and Event History Analysis, A Process Point of View*. Springer, New York.
- Alm Stenflo, G. (1994). *Demographic description of the Skellefteå and Sundsvall regions during the 19th century*. Demographic Data Base, Umeå University, Umeå, Sweden.
- Bengtsson, T. and Dribe, M. (2006). Deliberate control in a natural fertility population: Southern Sweden, 1766–1864. *Demography*, 43:727–746.
- Bengtsson, T. and Dribe, M. (2014). The historical fertility transition at the micro level: Southern Sweden 1815–1939. *Demographic Research*, 30:493–534.
- Bras, H. (2014). Structural and diffusion effects in the Dutch fertility transition, 1870–1940. *Demographic Research*, 30:151–186.
- Broström, G. (1985). Practical aspects of the estimation of the parameters in Coale’s model for marital fertility. *Demography*, 22:625–631.
- Broström, G. (2012). *Event History Analysis with R*. Chapman & Hall/CRC, Boca Raton.
- Broström, G. (2017). *eha: Event History Analysis*. R package version 2.4-6.
- Coale, A. J. (1971). Age patterns of marriage. *Population Studies*, 25:193–214.

- Coale, A. J. and Trussell, T. J. (1978). Technical note: Finding the two parameters that specify a model schedule for marital fertility. *Population Index*, 44:203–213.
- Coale, A. J. and Watkins, S. C. (1986). *The decline of fertility in Europe*. Princeton University Press.
- Dribe, M. and Scalone, F. (2014). Social class and net fertility before, during, and after the demographic transition: A micro-level analysis of Sweden 1880–1970. *Demographic Research*, 30:429–464.
- Engberg, E. (2005). I fattiga omständigheter. Fattigvårdens former och understödstagare i Skellefteå socken under 1800-talet. Umeå: Demographic Data Base. Swedish.
- Hoem, J. M. (1987). Statistical analysis of a multiplicative model and its application to the standardisation of vital rates. *International Statistical Review*, 55:119–152.
- Hoem, J. M. and Mureşan, C. (2011). The total fertility rate and its extensions. *European Journal of Population*, 27:295–312.
- Laird, N. and Olivier, D. (1981). Covariance analysis of censored survival data using log-linear analysis techniques. *Journal of the American Statistical Association*, 76:231–240.
- Mason, K. O. (1997). Explaining fertility transitions. *Demography*, 34:443–454.
- Ní Bhrolcháin, M. (1992). Period paramount? A critique of the cohort approach to fertility. *Population and Development Review*, 18:599–629.
- Ní Bhrolcháin, M. (2011). Tempo and the TFR. *Demography*, 48:841–861.
- R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Westberg, A., Engberg, E., and Edvinsson, S. (2016). A unique source for innovative longitudinal research: The Poplink database. *Historical Life Course Studies*, 3:20–31.