Max-Planck-Institut für demografische Forschung
Max Planck Institute for Demographic Research
Konrad-Zuse-Strasse 1 · D-18057 Rostock · GERMANY
Tel +49 (0) 3 81 20 81 - 0; Fax +49 (0) 3 81 20 81 - 202;
http://www.demogr.mpg.de

MPIDR WORKING PAPER WP 2006-006
MARCH 2006

Anticipatory analysis and its alternatives in life-course research. Part 1: Education and first childbearing

Jan M. Hoem (hoem@demogr.mpg.de)
Michaela Kreyenfeld (kreyenfeld@demogr.mpg.de)

Anticipatory analysis and its alternatives in life-course research. Part 1: Education and first childbearing

Reflections by Jan M. Hoem and Michaela Kreyenfeld

Abstract: Procedures that seek to explain current behavior by future outcomes (anticipatory analysis) constitute a widespread but problematic approach in life-course analysis because they disturb the role of time and the temporal order of events. Nevertheless the practice is often used, not least because it easily produces useful summary measures like the median age at first childbearing and the per cent permanently childless in various educational groups, defined by ultimate attainment. We use an empirical example to demonstrate the issues involved and to propose an alternative "non-anticipatory" research strategy, which, however, does not equally easily provide summary measures.

Keywords: anticipatory analysis, conditioning on the future, fertility by educational attainment

1 Introduction

Time and the *temporal order of events* play a decisive role for our understanding of behavioral processes that evolve over time in an individual's life. The topic of this paper is anticipatory analysis, which is any approach where one attempts to explain *past or current* behavior by *future* outcomes, in other words by conditioning on the future. It is important to understand the function and outcome of such a practice, for it remains quite popular. Here are a couple of typical examples that have appeared recently in the best of demographic journals:

(i) In a paper in *Demography* concerned with first-birth rates for women above age 30, Martin (2000) analyzed complete fertility histories from the U.S. Current Population Survey using educational attainment measured at the date of interview as an explanatory variable. The analysis is anticipatory because the educational outcome is known only at the end of the periods for which fertility behavior is recorded. The practice is ubiquitous and a literature review is hardly necessary.

(ii) In a paper in the *Journal of Marriage and the Family*, Corijn, Liefbroer and Gierveld (1996) also study entry into motherhood. One of their regressors is religious affiliation measured at the date of interview. Because religious affiliation is not necessarily fixed over the life-course, their analysis is anticipatory, because religious affiliation is used as a determinant of fertility behavior before the interviews. De Wit and Ravanera (1998) followed the same practice in a similar study, as did Hoem and Hoem (1989).

A considerable literature warns against the use of an anticipatory approach (Hoem 1996; Kravdal 2004), but researchers vary in their attitudes. An advantage of anticipatory analysis is that it sometimes easily provides descriptive summary measures of demographic behavior (like the median ages at first birth and the percentage ultimately childless by ultimate educational attainment, as in Figure 1 below), while this can be much harder with a non-anticipatory approach. Such summary measures can be useful to layman and professional alike, because they encapsulate important consequences of the transition rates that are so popular among life-history analysts. We therefore address the following general questions: Is anticipatory analysis really harmful in situations like those in these examples? Is it not rather a research tool that can be used to discover patterns of social processes, patterns that might be hard to reveal otherwise? In particular, can conditioning on the future be an acceptable research strategy when educational histories are not available but educational attainment at the time of interview is? Is an anticipatory approach harmful when it is used for causal inference but still acceptable for descriptive purposes? Or is the outcome of some anticipatory analysis deceptively and misleadingly simple, and are such procedures a total malpractice that violates basic principles of statistical methodology, perhaps by regularly producing biased results? By extension, we ask which strategies are available to avoid anticipatory analysis.

The authors of these reflections have found a need to discuss such issues extensively with each other. The purpose of the present text is to share our considerations with others and to display various possible procedures of analysis. For those who like to know where the road leads to, let us note at the outset that we have found the easy descriptions produced by an anticipatory analysis enticing but potentially deceptive, in that they may give a seriously biased and overly simplified impression of the patterns of real behavior. We offer an alternative procedure that is not anticipatory and not subject to the same flaws. It is an elementary extension of ordinary life-table theory. It exploits a particularly simple representation of educational-and-childbearing histories where all that is known is the educational level attained at the time of interview and the age at which was attained, from which we impute a rudimentary educational history. This type of data occurs often in practice, and the procedure we present works most straightforwardly where the educational system is quite rigid. It can be generalized to situations where more complete histories have been collected and where the educational system is more flexible, but that is not part of our account here.

Of course the procedure we propose builds on a simplification of reality too, but at least it has the advantage of representing education and childbearing explicitly as two interacting processes. We trust that the simplification does not in itself produce distortions that lead readers to a new set of misunderstandings. We illuminate these considerations by working through an empirical example based on real data in the sections that follow, and we do the same for the connection between marriage formation and childbearing in our companion paper. We believe that the examples have some independent interest in their own right. Unfortunately, the procedure we propose does not provide anything like a median age at first childbearing or a percentage ultimately childless by educational attainment, except by conditioning on the ultimate level of the latter, thus returning to the anticipatority we set out to eliminate in the first place.

2 Education and fertility

2.1 Anticipatory indicators of the impact of educational attainment on fertility

2.1.1 Cross-sectional fertility indicators

The connection between education and family dynamics has been discussed intensively in demographic, economic, and sociological work. (See, e.g., Hoem 1986, De Wit and Ravanera 1998, Blossfeld and Huinink 1991, Liefbroer and Corijn 1999, Santow and Bracher 2001.) The standard reasoning is that more highly educated women spend longer periods of their lives in education; when they enter the labor market, they earn higher wages, are more work-oriented, and enter more challenging employment careers. All of these factors are thought to work toward postponing family formation and increase childlessness. Here are some immediate questions: How old are university graduates when they have their first child? How many of them remain childless throughout their lives? And how is their behavior in comparison to other women?

To answer such questions it would be useful to have easily accessible summary indicators. Some obvious examples are the mean or median age at first birth and the fraction permanently childless among the women in each educational subgroup. Such indicators are used frequently in demographic research (see, e.g., Rindfuss, Morgan and Offutt 1996; Björklund 2006). They are also of major public interest. A high percentage childless (for instance) in any group suggests a strong incompatibility between work and family life in that group. In the recent public debate in Germany, the published finding that university graduates have particularly high levels of childlessness has found strong resonance among politicians and the public (Bernd 2005). Since such indicators are easily picked up by a wider audience, they are possibly better suited to promote political action than complex indicators derived from more sophisticated analytical strategies.

In Table 1, we display some cross-sectional fertility indicators. The data for this and all subsequent analysis come from the German Family and Fertility Survey (FFS), conducted in 1992. We have selected West-German women of ages 30-39 at the time of the interview and have grouped them into three categories according to the highest educational level they have attained *by the time of the interview*, namely women with (i) a university degree, (ii) a vocational-training certificate, and (iii) none of these attainments.[1] The table shows a strong association between recorded educational attainment and fertility. According to the table, university graduates were the older at first birth, they were much more likely than others to remain childless, and on average they gave birth to a smaller total number of children than other women.

Table 1: Cross-sectional fertility indicators by woman's educational level

	No degree or certificate	Vocational certificate	University degree
Mean values			
Mean age first birth	23.09	25.39	28.17
Number of children (distribution in per cent)			
Childless	24.18	29.83	48.28
One child	19.28	28.23	14.48
Two children	34.64	33.83	24.83
Three and more children	21.90	8.11	12.41
Total	100.0	100.0	100.0
Mean total number of children	1.54	1.20	1.01

Source: German FFS 1992 (our own estimates).

2.1.2 Reflections

According to Table 1, childlessness at ages 30-39 was radically more common among university graduates than among other educational groups. A straightforward and common explanation is that highly educated women are the more career-oriented, and that they remain childless to a large extent because work and family life are not easily compatible in Germany.

1 What we have called university degree includes 'Fachhochschulabschluss' and 'Hochschulabschluss'. Vocational certificate includes ‚Lehre', ‚Meister', Techniker', 'Fachschulabschluß'. We do not consider primary- and secondary-school degrees separately in our analysis. If a woman receives an "Abitur" but does not get a vocational certificate or a university degree, we classify her as having "no degree". This procedure seems reasonable if one takes into account the allocation principles of the German labor market. Formal qualifications, like university degrees or vocational training certificates, are more important for wages and labor market positions than years of primary and secondary schooling (Shavit and Müller 1998).

This is probably true, but we have a number of reservations to simply basing the argument for it on statistics like those in Table 1.

First, the interpretation just mentioned is plainly wrong to the extent that *causality* works in the opposite direction. Suppose that a woman must discontinue her university studies because she has a child. For her it is not (lack of) career orientation that makes her have fewer children; quite on the contrary her childbearing limits her educational choices. If this pattern is common, the table would only provide limited insight into the causal relationship between education and first birth, not least because the *temporal order* of events has not been used when Table 1 was produced. An event can be the cause of another event only if it happens first. To interpret Table 1 in a causal manner, one must assume that educational choice precedes family formation, and we just argued that this need not be the case. Let us take another example: Suppose a woman completes some vocational training at age 20 and has a child at 21. At age 28, she goes back to take more education, and she receives a university degree at age 32. Her fertility choice was probably made before she even contemplated going to the university. Nevertheless she would be classified as a university graduate in Table 1. Again, this amounts to a time-sequence reversal.

A fertility indicator computed separately by final educational level in some sense assumes that education is a fixed trait of the individual. How sensible this is, depends on the structure and flexibility of the educational system. During communism (in East Germany) educational choices were largely made at an early age and rarely revised. If education is completed regularly before childbearing begins, a causal interpretation of fertility by final level of education may be meaningful, because then it does not matter when educational attainment is measured. The more that people pursue extended or multiple educational careers and the more they re-train at later ages, the less meaningful it is to use education as a fixed characteristics of an individual, because such re-orientation takes time and is likely to stretch into the childbearing period. The only alternative apparent to us that makes the anticipatory procedure meaningful is to see ultimate educational attainment as revealing a lifetime plan which guides the individual's behavior until completion and which therefore is a fixed characteristic. We are skeptical of such a teleological interpretation.

Second, a related problem arises from the fact that some women are *in education on the date of interview*. At ages 30-39, as in Table 1, those still enrolled were most likely undergoing university education. Since in Germany very few of them can have had a vocational certificate and since they had not earned a university-level degree yet, they were coded as not having any degree or certificate. This does not seem to be a particularly good solution. Prospective university graduates (who just have not finished by the time of the interview) will behave differently from women who have completed their education at a lower level or have dropped out of education without having earned a degree or certificate. One could omit from the analysis women who are in education, but this solution has its own problems. It biases the results because women who are under education at interview surely include those who postpone fertility longer than others.

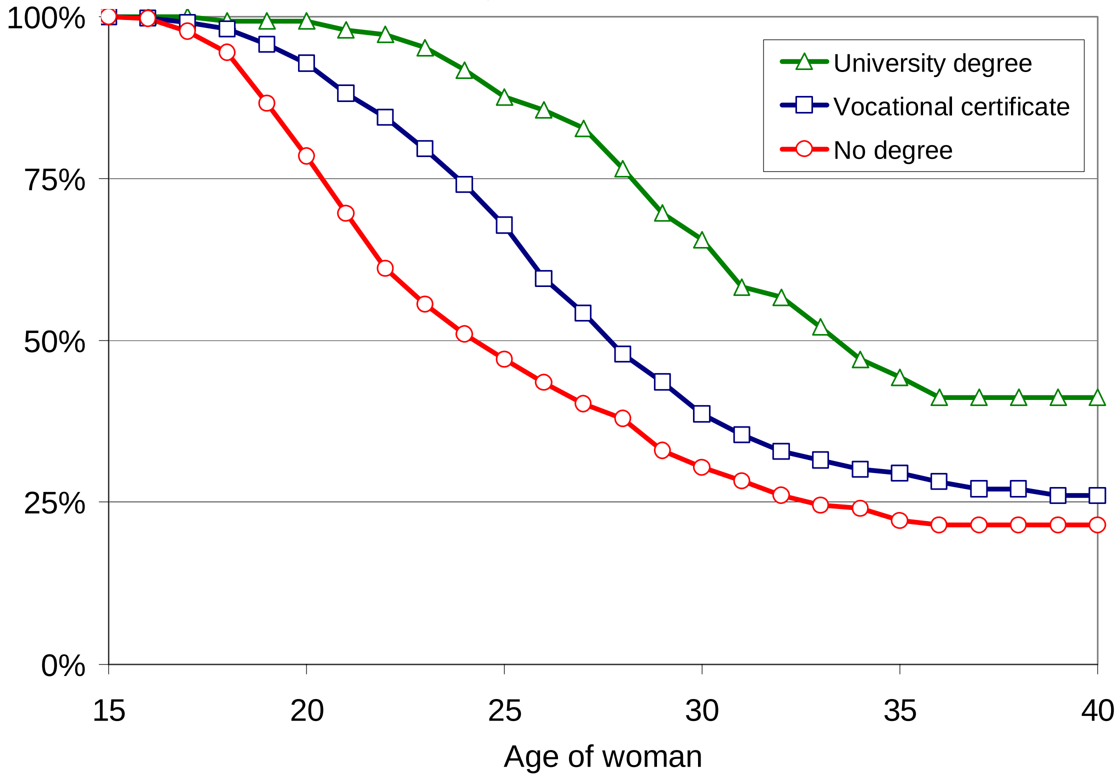
A *third* issue is that women aged 30 to 39 still are in the reproductive age span. Some of them might have children later than the interview. Therefore, Table 1 does not really provide estimates of completed fertility, the way the interpretation assumes. Less biased indicators can be calculated based on data sets which include older cohorts. Women at ages 45-60 are hardly reproductive any longer. For them there will be no underestimation of ultimate fertility caused by childbearing after the interview. However, since we deal with retrospective (survey) data, to the extent that mortality is differential by educational attainment, there will be a slight bias due to selection by virtue of survival. The older the cohort is, the stronger this selection becomes. Waiting until a cohort has completed fertility also means that one will mainly describe a historical development. For women aged 45-60 at interview, a retrospective fertility study mainly reflects childbearing behavior some twenty to thirty years ago on average. More up-to-date fertility indicators would certainly be preferable for those interested in current trends.

Survival analysis has been devised to account for censoring and to allow us to analyze the fertility of cohorts who are still in their reproductive years. A summary statistic like the *median age at first birth* can be derived from survival curves. We now turn to this possibility.

2.1.3 Survival curves by final level of education

Figure 1 shows survival curves for time to first birth, by level of education attained at interview for our 30-to-39-year-olds. These survival curves explicitly take censoring of the main event (childbearing) into account, but they too treat education as a fixed personal trait. In principle, our respondents came under the risk of childbearing at age 15. Everything that happened before this age is fixed for the first-birth process. (For example, the woman's own place and year of birth trivially are fixed factors.) We recapitulate that this is not so for educational attainment; this factor varies over the life-course. At age 15, none of the respondents has a vocational certificate or a university degree yet. On average, a vocational certificate is earned at age 19 in this data set, a university degree at age 28. When respondents are classified throughout their life histories (as far as we have observed them) according to educational attainment at interview, their educational level is essentially wrongly coded during life segments before they attain that "final" level. For instance, the first-birth survival curve for university graduates provides estimates for the fraction childless at ages 15 to19, but at such ages no university degree has been earned.

Figure 1: Kaplan-Meier survival curves for the arrival of the first child, by final level of education



Note: Computations based on data for West German women aged 30 to 39 at interview.
Source: German FFS 1992 (our own estimates)

We derive the following Table 2 from the diagram. A comparison with corresponding entries in Table 1 shows considerable adjustment of the figures for university graduates but only smaller changes for those with a lower educational attainment. We get these changes because Table 2 catches the women at an age on average five years later than in Table 1. (Note that for those with a university degree the median age in Table 2 is more than five years higher than the corresponding mean in Table 1.)

Table 2: Median age at first birth and per cent childless at age 40, by educational attainment at interview, based on Figure 1.

	No degree or certificate	Vocational certificate	University degree
Median age at first birth[2]	24.00	27.67	33.67
Childlessness at age 40, in per cent	21.50	26.03	41.21

Source: German Family and Fertility Survey 1992 (our own estimates).

One might feel that summary statistics like these, particularly Table 2 and the curves in Figure 1, are useful in describing the association between educational attainment and first childbearing. The statistics are easy to compute and to apparently understand. However, there is reason to be more skeptical. It is a major problem that while in reality the two lifetime processes develop dynamically in interaction with each other, the procedure just used treats them asymmetrically and handles educational attainment as if it were fixed throughout (at its final level) and only lets childbearing develop dynamically. A regular event-history analysis using educational attainment as a time-varying covariate is safer, particularly in that it minimizes the risk of estimation bias. To get a closer look at these issues, we now turn to the latter option.

2.2 Dynamic modeling of education

2.2.1 An event-history model incorporating current educational level and enrolment in education as a dynamic process

A major advantage of an event-history approach is that it makes it possible to consider education as a dynamic determinant of the behavior in focus, in our case first childbearing (see

[2] The median age at first birth was calculated as the smallest survival time values for which the Kaplan-Meier survival function is less than or equal to 50 per cent.

e.g., Hoem 1986, De Wit and Ravanera 1998, Blossfeld and Huinink 1991, Liefbroer and Corijn 1999, Santow and Bracher 2001). An essential requirement is that the data contain information about the respondent's educational attainment and any current educational enrolment for each month during the period of observation. Unfortunately, such detail is not always available, and the German FFS is a case in point. In that survey, respondents were "only" asked to report the highest educational level they had attained at the time of interview and the month and year in which they completed that education. They could choose between nine different educational levels, which we have regrouped into the three categories mentioned above (university degree completed, vocational certificate earned, and none of these). We have also constructed a (time-varying) binary variable that we hope will indicate periods in and out of education reasonably faithfully. We coded the respondents as being in education all the time before they attained the level reported in the interview. After the date of completion reported, we coded them as out of education. For respondents who had never attained a university-level degree or a vocational certificate, no real completion date was reported, and we have imputed a drop-out-date from education for each member of this rather heterogeneous group[3] and have coded her as in education until the drop-out date. Respondents who reported a vocational certificate as their highest educational attainment, were coded as being "in education" until the completion of the certificate. Respondents who reported a university-level education, were coded as being "in education" until completion of the university degree. It is obvious that this practice gives a simplified representation of reality. It does not account for more complex and diverse educational histories. Cases are not adequately considered where people receive multiple degrees and where they resume education after periods of employment. There may also be other, less obvious types of miscoding. However, the way education was surveyed in the German FFS, we do not have much choice.[4] At least our procedure has the merit of simplicity. It should also be sufficiently accurate for our methodological purpose.

---

3 To impute the drop-out-date, we proceed as follows. For most cases, we know the date when the woman completed primary and/or secondary school and also the date of labor market entry. If the first date was missing, we imputed June of the year in which she attained age 16. If the second date was missing, we imputed the month in which she attained age 20 unless this was earlier than the reported school-leaving age, in which case we used the latter as an imputed labor-market entry age also. For each woman we have then assigned a random drop-out date from education between these two dates, whether imputed or recorded. Remember that after she left school, a woman could go on to take some education which she did not complete.

4 In fact, our quandary can serve as a warning to data collectors who believe that they can get away with the bare-bones information about educational histories used in the German FFS and many other similar surveys.

With educational histories imputed as just described, we have fitted an event-history model to the data. Our process time is the age of the woman, used in the interval from age 15 to age 39, which we have partitioned by cut-points at ages 20, 25, and 30. The baseline hazard is essentially modeled as a function which is piecewise constant over the resulting four intervals.[5] Educational level and educational activity were entered together as a combination factor, in that we combined educational activity and educational level into a single time-varying covariate with the values indicated in the head of Table 3 below. Let us call it current educational attainment to underline that education is accounted for in a dynamic way. We cannot use a straightforward hazard model where the effects of these two covariates (age and current educational attainment) are represented in a multiplicative manner, and we have included them in interaction. Since no respondent can reach the highest educational level at a very young age, some combinations of age and current educational attainment are impossible in practice, as is indicated by the minus signs in one corner of Table 3. Figure 2 contains a plot of the absolute risks against age for the four columns in that table. It shows how the first-birth risks vary by age and educational attainment. Note how there is not monotonic dependence between educational attainment and childbearing risk across all ages.

Table 3: First-birth risk by age and current educational attainment, per 1000 woman-months

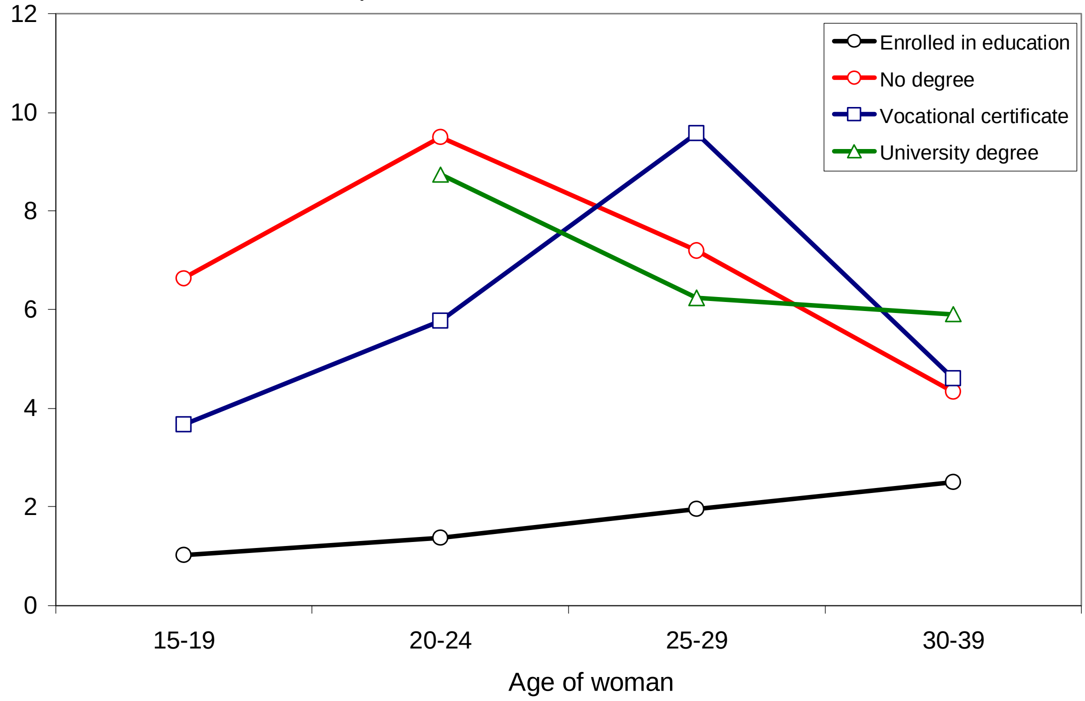
k	Enrolled in education 0	Not enrolled, no degree or certificate 1	Not enrolled, vocational certificate 2	Not enrolled, university degree 3
15-19	1.02	6.64	3.67	--
20-24	1.38	9.50	5.77	8.74
25-29	1.96	7.20	9.57	6.24
30-39	2.50	4.34	4.60	5.91

Notes: The sample comprises West German women aged 30 to 39 at the time of interview.
Source: German Family and Fertility Survey 1992 (our own estimates).

---

5 The last interval covers ages 30-39.

Figure 2: First-birth risk by age and current educational level, per 1000 woman-months



Notes: Data from Table 3
Source: German Family and Fertility Survey 1992 (our own estimates).

2.2.2 Survival curves by current educational attainment

In some sense the results of Section 2.2.1 represents the answer to the substantive questions we have asked. A wider audience may find the consequences of hazard curves such as those in Figure 2 rather inaccessible, however, and the professional would also find some summary measures useful as indicators of what curves like these mean for the age at childbearing and the per cent permanently childless. It may be easier to interpret what the curves mean if we convert them to a format similar to the survival curves in Figure 1. One possibility is then to provide survival curves by current educational level. To do so, we proceed as follows.

For k = 0, 1, 2, 3, let $\varphi_k(x)$ be the first-birth hazard for a respondent whose current educational attainment at (exact) age x is k, and for k = 1, 2, 3 let the corresponding single-intensity survival function be

$$\ell_k(x) = \exp\{-\int_{15}^{x} \varphi_k(s)ds\} \text{ for } x \geq 15.$$

(The value of *k* is given in the heading of each column in Table 3.) We can then use estimates $\hat{\varphi}_k(x)$ such as those in Table 3 to produce corresponding estimates for the survival functions. For illustration, let us calculate the points of the survival function for the ages 20, 25, 30 and 40 for women with a vocational certificate (*k*=2). Since there are sixty months in each five-year interval and since the items in the table are given per 1000 person-months, we get

$$\hat{\ell}_2(15) = \exp\left\{-\frac{60\times 3.67}{1000}\right\} = 0.80;$$

$$\hat{\ell}_2(20) = \hat{\ell}_2(21)\times \exp\left\{-\frac{60\times 5.77}{1000}\right\} = 0.57;$$

$$\hat{\ell}_2(30) = \hat{\ell}_2(25)\times \exp\left\{-\frac{60\times 9.57}{1000}\right\} = 0.32; \text{ and}$$

$$\hat{\ell}_2(40) = \hat{\ell}_2(30)\times \exp\left\{-\frac{120\times 4.60}{1000}\right\} = 0.18.$$

After similar computations for the various $\hat{\ell}_k(x)$ for *k*=1 and *k*=3, we can draw survival curves like those in Figure 3,[6] from which we derive the mean ages at first birth and the percentages childless in Table 4 corresponding to the values in Table 2. A comparison between Tables 2 and 4 shows that the dynamic modeling of educational attainment using single-intensity functions gives a picture of the role of education that is completely different from what we got by conditioning on educational attainment at interview. In particular, by this account the behavior of women with a university degree is far less radically different from other women that what the anticipatory analysis indicated. According to our (single-intensity) dynamic analysis, "only" twenty per cent of university-educated women were childless at age 40 (instead of 41% as estimated by the anticipatory analysis and even 48% as estimated in the descriptive analysis of Table 1). Their median age at first birth is just over 27, which is some six years lower than what the anticipatory analysis gave.

---

6 We could have computed $\hat{\ell}_2(x)$ for a finer grid of ages *x*, but the five-and-ten-year grid just described and the linear interpolations between the points on the grid in the diagrams should suffice for our current purposes.

Figure 3: Survival curves for the arrival of the first child, by current educational level

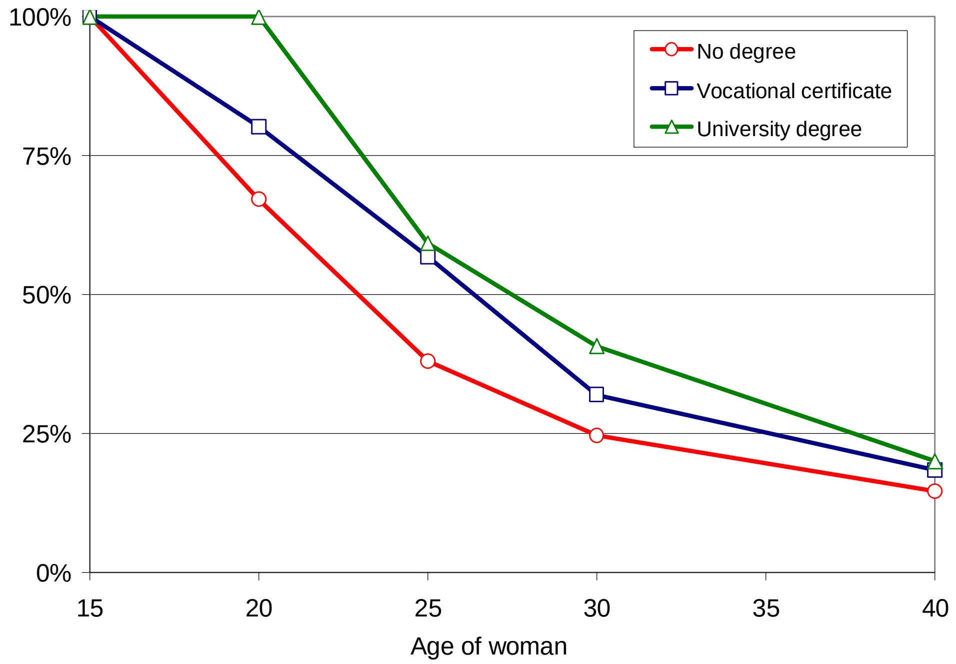


Table 4: Median age at first birth and per cent childless at age 40, by current educational level, based on Figure 3.

	No degree	Vocational certificate	University degree
Median age at first birth	22.94	26.36	27.49
Childless at age 40, in per cent	14.65	18.39	20.03

Notes: The sample comprises West German women aged 30 to 39 at the time of interview.
Source: German Family and Fertility Survey 1992 (our own estimates).

2.3 Accounting for the interrelation between childbearing and educational attainment

It would be neat if we could stop here and say that our dynamic analysis has proved irrevocably that the anticipatory analysis gives terribly biased results and that the truth is quite different from what the anticipatory analysis shows. Unfortunately, things are not quite so simple. Our results in Section 2.2 do not immediately represent a “truth” that anticipatory analysis can be compared with. It is important to note that the single-intensity survival functions $\ell_k(x)$ are constructs that must be interpreted with considerable care themselves, for they do not take into account the dynamic interaction between childbearing and educational attainment. Both educational progress and first childbearing are dynamic processes, and we need to take them both into consideration at the same time. This can be done as follows.

The first-birth intensities $\varphi_k(x)$ (k=0, 1, 2, 3) are picked from a model that incorporates both processes. A simple representation of this model is given in Figure 4, where the boxes represent life-course statuses that individuals can move between and the arrows reflect direct transitions that individuals can make. The functions associated with the arrows are corresponding transition intensities (or hazards). The intensities $\gamma_1(x)$, $\gamma_2(x)$, and $\gamma_3(x)$ are age-specific rates at which childless individuals change educational status for each age *x*. Thus $\gamma_1(x)$ is the rate at which they leave the educational system without formally completing either a vocational certificate or a university-level degree, while $\gamma_2(x)$ and $\gamma_3(x)$ are the rates at which they leave the educational system with a vocational certificate or a complete university degree, respectively. Because of the character of the FFS data at our disposal, we have needed to simplify central features of the German educational system, and the peculiarities of our representation are reflected as follows:

(1) Individuals remain in the state marked START ("no child, enrolled, ed = 0") as long as they are enrolled in education and until they enter motherhood or else complete a certificate or degree.

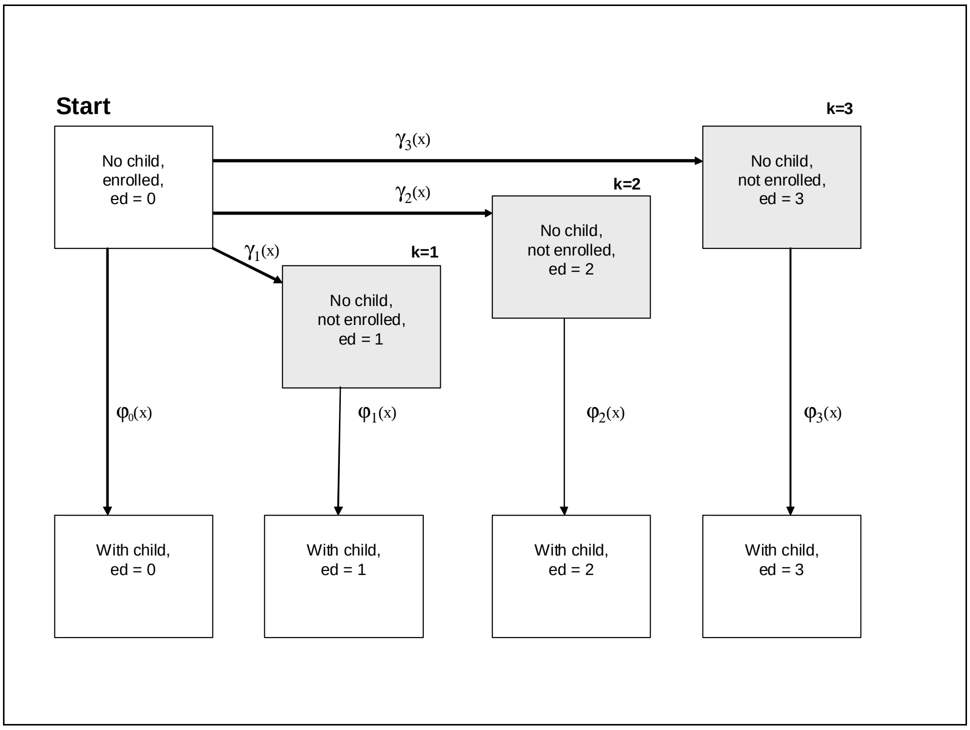
(2) Once an individual has left the educational system, there is no return.

(3) If an individual leaves the educational system without a vocational certificate or a university degree before entering motherhood, she moves to educational level 1 and remains there forever after.

(4) If they do not drop out, enrolled individuals can complete their education by acquiring a vocational certificate (which means that they go to educational level 2) or by completing university studies (educational level 3).

(5) One does not go through ed=2 as an intermediary step towards reaching ed=3.

Figure 4: Status-and-transitions diagram for education and first childbearing.



Note that this specification makes $\gamma_1(x)$, $\gamma_2(x)$, $\gamma_3(x)$, and $\varphi_0(x)$ the intensities (or hazards) of competing risks of transition out of the state marked START in Figure 4, while $\varphi_1(x)$, $\varphi_2(x)$, and $\varphi_3(x)$ are intensities of the only possible transition out of their respective states.[7] If we let *k*=0 represent the "educational attainment" corresponding to being enrolled in education and having no child (i.e., location in state START), then the survival function corresponding to this situation would be

$$\ell_0(x) = \exp\{-\int_{15}^{x} [\varphi_0(s) + \sum_{j=1}^{3} \gamma_j(s)]ds\} \text{ for } x \geq 15.$$

This would be the probability of not leaving the status "*No child, enrolled, ed=0*" (the state marked START) before age *x*.

For *k* = 1, 2, 3, $\ell_k(x+t)/\ell_k(x) = \exp\{-\int_0^t \varphi_k(x+s)ds\}$ is the probability that an individual will remain in the status marked "*No child, not enrolled, ed=k*" until age *x+t*, given that she has reached that status by age *x*. Both of these exponential formulas are derived in the

[7] We could have let the three latter transition intensities depend on time since educational attainment (i.e., time since entry into current state), but this must be a needless refinement at the present stage of analysis.

same manner as when we compute a normal life-table survival probability by forming

$${}_t p_x = \ell_{x+t} / \ell_x = \exp\{-\int_0^t \mu_{x+s} ds\}$$ when the force of mortality is μ_x.

The probability of having become a mother and also having reached $ed=k$ (i.e., one of the lowermost states in Figure 4) by age x is

$$\pi_0(x) = \int_{15}^{x} \ell_0(s)\varphi_0(s)ds \text{ for } k = 0,$$

$$\pi_k(x) = \int_{15}^{x} \ell_0(s)\gamma_k(s)[1-\frac{\ell_k(x)}{\ell_k(s)}]\, ds \text{ for } k = 1,2,3.$$

For some empirical values, see Table 5. (The columns of Table 5 are estimates of $\pi_0(x)+\pi_1(x)$, $\pi_2(x)$, $\pi_3(x)$, and their sum, respectively, for the various ages x indicated.)

Table 5: Probability of having a child and having an educational attainment, by age attained (multiplied by 1000)

k	Having a child, no certificate or degree 1	Having a child, vocational certificate 2	Having a child, university-level degree 3	Having a child, all educational attainments together
Age				
15	0	0	0	0.000
20	72	38	0	0.111
25	155	177	16	0.348
30	194	368	40	0.601
40	220	476	78	0.774
Fraction completing at given educational level, with or without a child	$\Pi_{0+}\Pi_1$=0.259	Π_2=0.624	Π_3=0.116	$\Sigma\ \Pi_k$=1
Of which fraction childless, in percent	0.038	0.148	0.038	0.225

Notes: The sample comprises West German women aged 30 to 39 at the time of interview.
Source: German Family and Fertility Survey 1992 (our own estimates).

Table 5 contains a considerable amount of information about the moves individuals have made in the two dimensions we operate in (educational attainment and first childbearing). Among other features we see that about one-quarter of the respondents in our cohort ended up without a vocational certificate or university-level degree by age 40, that some 15 per cent ended up childless and with a vocational certificate, while about one-fourth as many ended up at childless but with a university-level degree at age 40. The latter is about the same fraction that ended up childless and without any education at those two higher levels.

We have not been able to devise a measure similar to the median ages at first birth by educational attainment in Tables 2 and 4, except by appealing once more to an anticipatory procedure. We do the latter as follows.

Let $\Pi_0 = \pi_0(40) = \int_{15}^{40} \ell_0(s)\varphi_0(s)ds$ and let $\Pi_k = \int_{15}^{40} \ell_0(s)\gamma_k(s)ds$ for k=1, 2, 3.

Then each Π_k is the probability of ever leaving the state marked START in Figure 4 along an easily identified arrow leading out of that state. Since very few women attain a vocational certificate or a university degree after entry into motherhood in our data, Π_2 and Π_3 essentially are the probabilities of reaching educational level 2 and 3, respectively, by age 40, and $\Pi_0 + \Pi_1$ is the corresponding probability of remaining at a lower educational level. The conditional probability that a woman became a mother before age x, given that she reached educational level k by age 40, is therefore $[\pi_0(x) + \pi_1(x)]/[\Pi_0 + \Pi_1]$ for k=1, and it is $\pi_k(x)/\Pi_k$ for k=2 or 3. All of these conditional probabilities can be estimated from our data and plotted in the form of survival curves[8] as in Figure 5, from which we can derive Table 6.

[8] The curves plotted for k=2 and 3 are for the functions 1-$\pi_k(x)/\Pi_k$, and the curve marked “no education” is for 1-$[\pi_0(x) + \pi_1(x)]/[\Pi_0 + \Pi_1]$.

Figure 5: Survival curves for the arrival of the first child,
accounting for the interrelation between childbearing and educational attainment

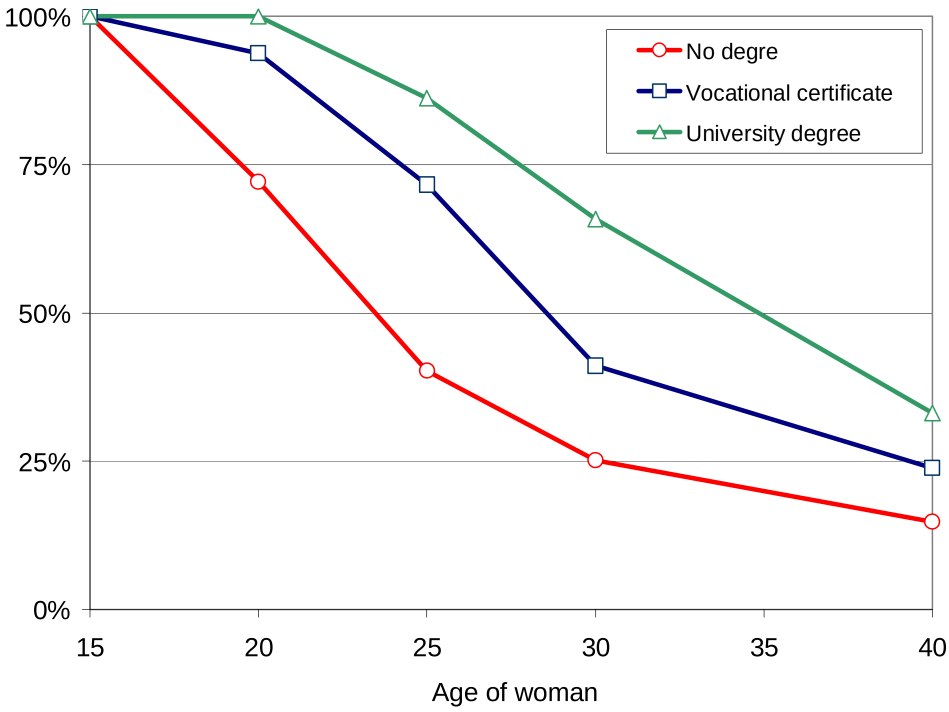


Table 6: Median age at first birth and per cent childless at age 40,
accounting for the interrelation between education and fertility, based on Figure 5.

	No degree	Vocational certificate	University degree
Median age at first birth	23.47	28.54	32.42
Childlessness at age 40 (per cent)	14.74	23.79	33.10

Source: German Family and Fertility Survey 1992 (our own estimates).

In the reflections above, we have described four different ways of producing median age at first birth and per cent permanently childless by educational attainment. To provide a summary of our findings, we list the main traits of our previous tables (Table 7). We see that the mean ages at childbirth computed according to the ideas of the present section are pretty close to (but not identical with) those computed by organizing the data according to educational attainment observed at interview. However, our approach provides vastly different levels of final childlessness by educational level. The main advantage of the new measures is that they better reflect the dynamics of the interaction between education and first childbearing. Crudely plotting survival curves by ultimate educational attainment, as in Figure 1, misses out on the interaction between the two individual-level processes involved.

Table 7: Median age at first birth and per cent childless at age 40, computed according to the four procedures described in this paper

	No degree	Vocational certificate	University degree
Median age at first birth			
Survival curves by final level of education	24.00	27.67	33.67
Survival curves by current educational attainment	22.94	26.36	27.49
Accounting for interrelation between education and fertility	23.47	28.54	32.42
Childlessness at age 40 (per cent)			
Survival curves by final level of education	21.50	26.03	41.21
Survival curves by current educational attainment	14.65	18.39	20.03
Accounting for interrelation between education and fertility	14.74	23.79	33.10

Source: German Family and Fertility Survey 1992 (our own estimates).

3 Conclusion

How harmful is it to display survival curves to first birth by final level of education, as in Figure 1? Problems evolve from the fact that educational participation is hardly ever completed before the respondent enters the risk period of first birth. Respondents come under the risk of first childbearing roughly at age 15, an age well before they earn their certificates and degrees (if they continue to be enrolled in education long enough). Educational attainment is therefore a time-varying factor in the first-birth process. Summary indicators by final level of education do not consider the dynamic nature of this interplay in the life-course of individuals. How harmful fertility indicators by final level of education are, may depend on how flexible the educational system is. The easier re-entry into enrollment is and the more people can have multiple educational careers and can re-train at later ages, the more problematic is treating education as a fixed characteristics.

One apparent solution that we consider is to model education dynamically in an event-history formulation which explicitly takes into account that educational attainment changes over the life-course. However, there are problems with a dynamic modeling of education in the analysis of first births too. In particular, we only managed to get summary measures like a median age at first childbearing and a per cent ultimately childless by conditioning on final educational outcome, albeit in a setting that fully exploits the dynamic interaction between the two dimensions. The "totally clean solutions" we can offer that do not represent conditioning on the future, explicitly or indirectly, have the weakness that they do not provide measures of centrality or of ultimate childlessness. They rely on descriptions of behavior in the form of intensity curves. This may be satisfactory to professionals, but some of their consequences are probably hidden, even to sophisticates. Further development is a matter for future research.

Acknowledgement

The authors wish to thank the Bundesinstitut für Bevölkerungswissenschaft (BiB) for its permission to use the FFS data on which parts of this study are based. For valuable comments, we thank Kerrie Nelson.

References

Bernd, Ulrich (2005): Warum habt ihr Angst vor mir? Akademiker bekommen immer weniger Kinder. Sie sind keine Egoisten. Ihnen fehlt auch nicht das Geld. Sie haben bloß keinen Mut. *Die Zeit* 08/2005.

Björklund, Anders (2006): Does family policy affect fertility? *Journal of Population Economics* (forthcoming).

Blossfeld, Hans-Peter/ Huinink, Johannes (1991): Human capital investment or norms of role transition? How women's schooling and career affect the process of family formation, *American Journal of Sociology* 97(1): 143-168.

De Wit, Margaret L./ Ravanera, Zenaida R. (1998): The changing impact of women's educational attainment on the timing of births in Canada. *Canadian Studies in Population* 25 (1): 45-67.

Hoem, Britta/ Hoem, Jan M. (1989): The impact of women's employment on second and third births in modern Sweden. *Population Studies*, 43: 47-67.

Hoem, Jan (1996): The harmfulness or harmlessness of using an anticipatory regressor: how dangerous is it to use education achieved as of 1990 in the analysis of divorce risks in earlier years? *Yearbook of Population Research in Finland* 33: 34-43.

Hoem, Jan (1986): The impact of education on modern family- union initiation, *European Journal of Population* 2(2): 113-133.

Kravdal, Øystein (2004): An illustration of the problems caused by incomplete education histories in fertility analyses, *Demographic Research* (Special Collection) 3, 135-154.

Liefbroer, Aart/ Corijn, Martine (1999): Who, what, and when? Specifying the impact of educational attainment and labour force participation on family formation. *European Journal of Population* 15 (1): 45-75.

Martin, Steven P.(2000): Diverging fertility among U.S. women who delay childbearing past age 30. Demography 37: 523-533.

Rindfuss, Ronald.R./ Morgan, Philip S./ Offutt, Kate, (1996): Education and the changing age pattern of American fertility: 1963-1989. *Demography* 33 (3): 277-290.

Santow, Gigi/ Bracher, Michael (2001): Deferment of the first birth and fluctuating fertility in Sweden. *European Journal of Population* 17 (4): 343-363

Shavit, Yossi / Müller, Walter (1998) (eds.): *From School to Work. A Comparative Study of Educational Qualifications and Occupational Destinations*, Oxford: Clarendon: 1-48.

Appendix

The following computations lead to the values in Table 5:

The first-birth risks φ_k and the educational-attainment risks γ_k are piecewise constant. The intervals of constancy are mostly five years long, but the last interval is ten years long. For $x \geq 15$ we have defined

$$\ell_0(x) = \exp\{-\int_{15}^{x} [\varphi_0(s) + \sum_{j=1}^{3} \gamma_j(s)]ds\},\ \pi_0(x) = \int_{15}^{x} \ell_0(s)\varphi_0(s)ds, \text{ and}$$

$$\pi_k(x) = \int_{15}^{x} \ell_0(s)\gamma_k(s)ds - \ell_k(x)\int_{15}^{x} \gamma_k(s)\frac{\ell_0(s)}{\ell_k(s)}ds \text{ for } k = 1,2,3.$$

We note that $\ell_0(15) = 1$ and that $\pi_k(15) = 0$ for k=0, 1, 2, 3. To compute the values of these various functions for *x* = 20, 25, 30, and 40, we first introduce

$$\sigma(x) = \varphi_0(x) + \gamma_1(x) + \gamma_2(x) + \gamma_3(x) \text{ and}$$

$$_{\Delta}p_0(x) = \ell_0(x+\Delta)/\ell_0(x) = \exp\{-\Delta\sigma(x)\},$$

which we need for Δ =5 when *x*=15, 20, and 25, and for Δ =10 when *x*=30. Once the $_{\Delta}p_0(x)$ have been computed, we can compute the $\ell_0(x)$ recursively by the formula $\ell_0(x+\Delta) = \ell_0(x)\,_{\Delta}p_0(x)$ for Δ =5 or Δ =10 in the usual manner.

To compute $\pi_0(x)$, let $\delta_0(x) = \int_{x}^{x+5} \ell_0(s)\varphi_0(s)ds$ and note that

$$\delta_0(x) = \varphi_0(x)\ell_0(x)\int_{x}^{x+5} \exp\{-\int_{x}^{s} \sigma(x)du\}ds = \varphi_0(x)\ell_o(x)[1-\exp\{-5\sigma(x)\}]/\sigma(x).$$

Then $\pi_0(x+5) = \pi_0(x) + \delta_0(x)$ for $x = 15, 20, 25,$ while for *x*=40 we get, correspondingly, $\pi_0(40) = \pi_0(30) + \varphi_0(30)\ell_0(30)[1-\exp\{-10\sigma(30)]/\sigma(30)$. Note that $\varphi_0(30)$ and $\sigma(30)$ are the intensity values that are taken as constant between ages 30 and 40.

To compute $\pi_k(x)$ for k=1, 2, 3, let $\Gamma_k(x) = \int_{15}^{x} l_0(s)\gamma_k(s)ds$ and let

$$\delta_k(x) = \int_{x}^{x+5} \ell_0(s)\gamma_k(s)ds = \gamma_k(x)\ell_0(x)[1-e^{-5\sigma(x)}]/\sigma(x).$$

Then $\Gamma_k(15)=0$, $\Gamma_k(x+5)=\Gamma_k(x)+\delta_k(x)$ for $x=15,20,25,$ while

$$\Gamma_k(40)=\Gamma_k(30)+\gamma_k(30)\ell_0(30)[1-\exp\{-10\sigma(30)\}]/\sigma(30).$$

Similarly, let $\Lambda_k(x)=\int_{15}^{x}\gamma_k(s)\frac{\ell_0(s)}{\ell_k(s)}ds$ and $\lambda_k(x)=\int_{x}^{x+5}\gamma_k(s)\frac{\ell_0(s)}{\ell_k(s)}ds$.

As long as $\sigma(x)>\varphi_k(x)$, we get

$$\lambda_k(x)=\gamma_k(x)\frac{\ell_0(x)}{\ell_k(x)}\int_{x}^{x+5}\frac{\exp\{-\int_x^s\sigma(x)du}{\exp\{-\int_x^s\varphi_k(x)du}ds =$$

$$\gamma_k(x)\frac{\ell_0(x)}{\ell_k(x)}\int_{x}^{x+5}\exp\{-(s-x)[\sigma(x)-\varphi_k(x)]\}ds =$$

$$\gamma_k(x)\frac{\ell_0(x)}{\ell_k(x)}\langle 1-\exp\{-5[\sigma(x)-\varphi_k(x)]\}\rangle/[\sigma(x)-\varphi_k(x)].$$

Then $\Lambda_k(15)=0$ and $\Lambda_k(x+5)=\Lambda_k(x)+\lambda_k(x)$ for $x=15$, 20 25, while similarly

$$\Lambda_k(40)=\Lambda_k(30)+\gamma_k(30)\frac{\ell_0(30)}{\ell_k(30)}\frac{1-\exp\{-10[\sigma(30)-\varphi_k(30)]\}}{\sigma(30)-\varphi_k(30)}.$$

Finally, $\pi_k(x)=\Gamma_k(x)-\ell_k(x)\Lambda(x)$ for x=20, 25, 30, and 40.

To compute the items of Table 5, we need to (i) convert the values of the estimates $\varphi_k(x)$ in Table 3, which are given in terms of 1000 woman-months, to corresponding values for woman-years through multiplication by 12/1000, (ii) convert the values of the estimates $\hat{\gamma}_k(x)$ correspondingly, and (iii) insert the results in the formulas that we have just derived. The values of the $\hat{\gamma}_k(x)$ are a side issue in this paper and we have not listed them.