STOCKHOLM RESEARCH REPORTS IN DEMOGRAPHY

No. 14

THE EFFECTS OF SAMPLING AND NONRESPONSE ON ESTIMATES OF TRANSITION INTENSITIES:

SOME EMPIRICAL RESULTS FROM THE 1981 SWEDISH FERTILITY SURVEY

by

Ingrid Lyberg*

University of Stockholm
Section of Demography
S-106 91 Stockholm
Sweden

ISBN 91-7820-001-6

December 1983

Reprinted March 1991

* Statistics Sweden
S-115 81 Stockholm

Ingrid Lyberg April 15, 1991

The effects of sampling and nonresponse errors on estimates of transition intensities: Some empirical results from the 1981 Swedish Fertility Survey.

ERRATA

Page 19, line 1: substitute μ_{0i} with $\mu_{0i}(u)$.

Page 20, last line: Change to "...$p= 1-q = \exp[-z\mu]$...".

Page 22: Insert " / " before the left bracket in the fourth formula.

Page 25: Substitute $\hat{P}_{ij}(x) = \ldots$ with

$$\hat{P}_{ij}(x) = \sum_{r=0}^{x-1} \frac{\hat{\mu}_{0i}(r)}{\hat{\mu}_0(r)}[\hat{P}_0(r)-\hat{P}_0(r+1)]\left[\hat{\ell}_{ij}(x-r-1|\, r)+\frac{\hat{\mu}ij(x-r-1|\, r)}{\hat{\mu}_i(x-r-1|\, r)}\hat{\ell}_i(x-r-1|\, r)\right]$$

$$-\sum_{r=0}^{x-1} \frac{\hat{\mu}_{ij}(x-r-1|\, r)}{\hat{\mu}_i(x-r-1|\, r)}\hat{w}_i(x,r), \quad .$$

Page 25, last formula: Substitute $\hat{\ell}_0(r-1)$ with $\hat{\ell}_0(r+1)$.

Page 29, 5th line: Change to "(Table 3)".

Page 39, 2nd paragraph, 1st line: Insert "Again those false peaks would hardly give rise to" .

ABSTRACT

By utilizing register data covering the entire population, we have investigated the sampling and nonresponse effects on estimates based on the 1981 Swedish fertility survey. The life histories gathered in this survey are analyzed by means of semi-Markov models. Age- or duration-specific occurrence/exposure rates are used as estimates of the transition intensities in such models. For the evaluation study we have used semi-Markov models describing events contained in the register, namely births and changes of marital status. Occurrence/exposure rates based on data for the entire target sample or only the respondent in the survey are found to be subject to large stochastic errors, especially at high ages or durations where the number of observations becomes rather small. In spite of a fairly high response rate in the survey (87 percent), there exist systematic nonresponse errors, but they are small compared with the stochastic errors which can be estimated. If we disregard single extreme values and use statistical methods to test results based on data for the entire target sample or for the respondents only, we would hardly draw any wrong conclusions about the changes between birth cohorts. Replicated analysis based on data for the entire population, the entire sample, and the respondents show essentially the same pattern.

CONTENTS

1. INTRODUCTION

Life table techniques are often used to describe and analyze demographic phenomena such as fertility, nuptiality and mortality. By means of computed occurrence/exposure rates for consecutive age or duration intervals one can describe the life history of a specific cohort and compare the life histories of different cohorts.

The life histories of an assumed homogenous population (cohort) may be regarded as realizations of a stochastic transition process, for example the transition from the state "parity 0" to the state "parity 1". The occurrence/exposure rates then serve as estimates of the assumed piecewise constant transition intensities in that process. By means of the estimated intensities one can estimate other parameters such as transition probabilities, expected time in various states, etc. (see, e.g., Hoem and Funck Jensen, 1982). The statistical properties of these fundamental estimates of the intensities are well described in the literature (see, e.g., Beyer et al., 1976 and Vaeth. 1977).

Most studies of more complex life histories - not just births and deaths - are based on retrospective observations of an random <u>sample</u> from a population of individuals. In itself, this fact only changes the properties of the estimates via the sample size, provided that the observed life histories can still be regarded as independent outcomes of the same stochastic process as in the total population. This requires a noninformative sampling plan, i.e., independence between life history and inclusion in the sample. For a more detailed discussion of this matter, see Hoem (1983a).

The requirement of a noninformative sampling plan is often met in practice for the target sample. However, sample surveys are usually afflicted by <u>nonresponse</u>, i.e., for various reasons one fails to obtain data for some individuals sampled. The occurrence of missing observations affects the estimates not only by reducing the number of observations and by that increasing the variances, but also by introducing bias into the estimates. The bias arises because there is almost certainly some connection between life history realized and the propensity to respond: in reality the achieved sampling plan becomes informative. This was the case in the 1981 Swedish fertility survey.

In the spring of 1978, Statistics Sweden started the Swedish fertility project to investigate the reasons for Sweden's declining fertility since the middle of the 1960s. The Swedish fertility register was established as a part of this project. For all Swedish women born in 1926-60, it includes information about births, changes of marital status, migration and deaths. So far, three investigations have been based on the fertility register: two studies by Johansson and Finnäs (1983) and one by Quist (1983). In 1981 the Swedish fertility survey was conducted in order to study the connection between fertility and variables not included in the register, such as social background, cohabitational history, and so on. The collected life histories were to be analyzed be means of life-table techniques within the framework of semi-Markov models. Unfortunately, like any survey, this one was subject to nonresponse. There is a connection between response behavior and fertility: the response rate is higher among women with their own children in their households (at the time of the interview) than among childless women, and this difference is largest for the oldest women. Of course, such selective nonresponse may affect the results of a life-history analysis based on data for the respondents only. By means of data from the fertility register, we have been able to study the empirical effects of nonresponse (and sampling) errors on survey estimates of transition intensities, transition probabilities, and test statistics; the results are reported here.

An assessment of our results must take into account the fact that any purely empirical study has its own limitations. The results concern the variables included in both the register and the survey, and they may not be valid for other variables in the survey or for the same variables in other similar surveys. To make inference beyond our empirical findings possible, our investigation has been based on a theoretical nonresponse model presented previously in Lyberg (1983).

Our account proper starts with a presentation of the data used in the nonresponse study: the Swedish fertility survey and the Swedish fertility register. Data have been extracted from the register for the entire target sample of the survey, respondents as well as nonrespondents. This has enabled us to calculate response rates for groups of women with various life histories and to compute separate life-table estimates for the entire target sample and the respondents, respectively. The entire register has been used to compute corresponding population estimates as well. There are some minor discrepancies between the populations covered by the register and by the survey, and the register information about events that occurred before the end of 1960 is not always quite accurate. We believe, however, that these inaccuracies are of no substantive consequence and that the data used are quite appropriate for our purpose, namely describing the effects of nonresponse (and sampling) on estimates of various life table parameters.

In Section 3 we describe the methods used in this study. We have defined two transition models which can be used to analyze data in the fertility register. Occurrence/exposure rates estimate the transition intensities in these models, and we show how transition probabilities are estimated recursively via the estimated intensities. We also describe our theoretical model and some results derived from it. The model concerns traditional competing risks analysis. The relevant parameters are the transition intensities or other model parameters derived from them. We assume a probabilistic response behavior connected with the variables under study. The response probability approach has also been taken by Platek et al. (1978), Andersson (1979), Cassel et al. (1979), Särndal and Hui (1981), Little (1982), Hoem (1983a), and others. None of them, however, has connected a nonresponse model to a competing risks model.

In Section 4 we present the results of our empirical investigation. We start with the response rates. According to our theoretical model, the nonresponse bias of occurrence/exposure rates (estimated intensities) can be expressed as a function of the differences in the response probabilities of individuals who remain in a state and those who leave it during the relevant time period.

By means of response rates (estimated response probabilities) for the decrements and for the survivors, we can estimate an upper limit of the relative nonresponse bias. In general, the response rates are higher for the decrements (from various causes) than for the survivors in various states of our models. For instance, response rates are higher for new mothers and for newly married women than for other women during an age interval or a duration interval. This means that the estimates of the transition intensities are afflicted with a positive nonresponse bias. Usually we do not know the real response rates in a survey for groups of individuals with different life histories. To estimate the size of the nonresponse bias we may then, if possible, use the response rates for various relevant groups according to present status. This would, however, give an overpessimistic picture of the nonresponse bias for estimates of transition intensities for periods far earlier than the time of the interview.

We bring out the size of the nonresponse bias (and the sampling errors) for various estimated transition intensities by comparing estimates based on data for the entire population, for the entire target sample, and for the respondents. As expected, in general the intensity curves based on data for the respondents are higher than those based on data for the entire target sample. However, the differences between estimates based on data for the entire target sample and those for the respondents are very small compared to the stochasitic errors. The estimated standard errors are about 15 times larger for estimates based on the sample than for estimates based on the entire population. Intensity curves which are estimated by means of the sample (or the respondents) are very irregular and sometimes show false peaks and valleys. Nevertheless, when we estimate the standard errors and use statistical methods to test our results, our substantive conclusions come out right.

Fortunately, the life history analyses of the Swedish fertility survey are not based on separate estimates of single age- or duration-specific transition intensities. Instead, we study sequences of such estimates, organized either over ages (or over durations in a state) for selected birth cohorts or over birth cohorts for selected ages (or durations). When we compare such curves to find out differences between birth cohorts or between other population groups, we disregard single extreme values that stand out from surrounding values. This compen-ates for errors due to sampling and bias due to nonresponse. The stochastic errors make our analyses less sharp, but at least their effects can be measured by our statistical methods. The nonresponse bias changes the levels of our curves systematically, but it does not affect them substantially as compared to the stochastic errors.

In the sample data, we can study only the first few years of duration in states like "first-married, parity 0" and "never-married, parity 1", since the number of individuals surviving in these states rapidly diminishes as the duration increases. It is not possible to study either dissolution of marriage (for which the intensities increase but are still very low during the first four years of marriage) or third-birth fertility (which has decreased). Thus, a small sample size permits us to study only the first few phases of women's adult life.

To increase the number of observations in each group, the women have been grouped in the sampling plan into five-year birth cohorts, and to control for age at entry into the states "first-married, parity 0" and "never-married, parity 1" we have used five-year age groups. This grouping means that we cannot detect any real period effects or at which age or for which birth cohort any new pattern started. Furthermore, our model assumption, that the life histories observed for a specific group are independent outcomes of the <u>same</u> stochastic process, becomes less realistic. Among other things, this means that the effect of censoring (at the time of the interview) cannot be ignored. When analyzing the behavior for a five-year birth cohort and five-year age-at-entry group we should remember that the censoring appears at different ages for the single-year birth cohorts.

The main conclusion of our evaluation is that these survey data can be very useful for life history analysis by means of life table techniques. The stochastic errors may be large, but they are measurable, and the effects of nonresponse are likely to be much smaller than in analyses which refer to current circumstances only.

2. DATA

2.1. Data sources

The nonresponse study is based on data about response behavior in the Swedish fertility survey and on life-history data extracted from the Swedish fertility register. By means of these sources we have been able to construct separate data sets which cover

P: the register population, which essentially coincides with the survey population,

S: the entire target sample of the survey, which includes respondents as well as nonrespondents, and

R: the respondents in the survey.

For each data set we have the same information about births and changes of marital status, as obtained from the register. Thus, the sets differ in coverage but not in content.

Of course, our nonresponse study is restricted to the Swedish fertility survey and the variables we have been able to evaluate. The results, however, can give some insight into the effects of nonresponse in fertility surveys in general. The response pattern in the Swedish fertility survey is probably quite representative of such surveys carried out in industrial countries, even if the response rate is comparatively high. The variables evaluated are also, unquestionably, of highest relevance in a fertility survey.

The sources of data are further described below. Table 1 below and Table 2 at the end of Section 3 show the number of cases involved.

2.2. The Swedish fertility survey

In 1981, Statistics Sweden conducted a fertility survey among women born in Sweden 1936-60 and resident in the country at the time of sample selection (February, 1981). Thus the women under study were between 20 and 44 years of age at the end of 1980.

A sample of 4,966 women was drawn from the Swedish Central Population Register by simple random sampling from each of five strata. Each stratum consisted of a five-year birth cohort which constituted a target population of its own. This means that the sampling plan is noninformative.

For the interviews, we used a questionnaire that contained over one hundred questions about childbearing and factors thought to be connected with fertility, such as social background, education, earlier and present occupation, cohabitational and marital history, and so on. The interviews were made by about 200 female members of our regular field staff of interviewers. The median time of an interview was 65 minutes among all women, 50 minutes among single, childless women and about 70 minutes among cohabiting/ married women.

Interviewing took place mainly in March to May, 1981, with some follow-up in June to October, 1981. Great efforts were made to get a high response rate, and interviews were finally achieved with 4 300 respondents (87 %), a comparatively high figure for a fertility survey. The response rate is higher among women with their own children in their households than among women without children and this difference increases with age (see Table 1).

Table 1. Size of population, sample, number of respondents and response rates in the 1981 Swedish fertility survey.

Stratum		Population	Sample	Respondents	Response rates, %		
Year of birth	Age in 1980	n_P	n_S	n_R	All women	With children in their household[1]	Without children in their household[1]
1936-40	40-44	205,719	608	500	82	85	75
1941-45	35-39	271,350	1,180	1,014	86	89	72
1946-50	30-34	283,512	1,186	1,030	87	89	77
1951-55	25-29	249,721	1,187	1,049	88	91	85
1956-60	20-24	245,432	805	707	88	90	87
All	20-44	1,255,734	4,966	4,300	87	89	82

[1] Own children born in 1962 or later and belonging to their households at the time of the sample selection.

Obviously, the nonrespondents constitute a selective group with respect to the presence of children at home, and it is probably even more selective when actual fertility is concerned. This was true in the 1977 Norwegian fertility survey as well. In that survey, the response rates for all women with and without any live births were 86% and 75%, respectively, and the corresponding figures for the oldest women (40 to 44 years of age at the time of the interview) were 80% and 57%, respectively (Statistisk Sentralbyrå, 1981, p 21).

Furthermore, a comparison of educational level between respondents in our fertility survey and in the 1980 Swedish survey of living conditions shows a good correspondence for all our cohorts (strata) except for the oldest one. In fact, the results indicate that the nonresponse rate in the fertility survey is extremely high among less-educated women in the oldest cohort (40-44 years of age) (Arvidsson et al., 1982, p 99).

Thus there is a connection between life history and propensity to respond in the fertility survey, and this connection interacts with age. Therefore, estimates of measures of fertility and its relation to other factors will be afflicted with a bias which may differ between the cohorts. The question is how large the bias is in itself and compared to the random errors, and whether the bias will give rise to any wrong conclusions when the survey results are interpreted.

2.3. The Swedish fertility register

We have used the Swedish fertility register as an external source to describe the nonresponse effects in the fertility survey. In spite of its name, it is not a physical register, but rather a computer system designed to match information from various population registers. Conceptually, however, it is useful to describe it as a register with a specific content. As such it includes data for all women born in 1926-60 who were Swedish citizens at the time of the 1960 population census. In contains retrospective census data about the date of birth of any child born in 1936-60 and living in the household of one of these women in November of 1960. The current marital status of these women and the time (in integer years) of any latest change of status as of that census are also included. Information has been added about any childbearing, migration, change of marital status and death after 1960, as obtained from certificates on vital events collected through the civil registration system.

So far three investigations based on this register have been conducted: two studies by Johansson and Finnäs (1983) and one by Quist (1983). In these studies, register information about vital events up to 1977 has been used. We have used the results from those investigations as population estimates in our study. The transition model of Figure 1 below is a part of the model used by Quist. For this model, aggregate data on the number of cases, occurrences and exposure were available. For the model of Figure 2, unfortunately, such data were not available and we did not calculate them for the nonresponse study. We have used instead some relevant results from the study by Johansson and averaged them to apply to our five-year birth cohorts.

Since data on an unknown number of events that occurred before the end of 1960 (for instance, births before 1943) are missing in the register, the three studies are restricted to women born in 1935 or later. For women born in 1935-45 (who were 15 to 25 years of age at the 1960 census) some data about births and early annulled marriages are missing, but for women born in 1946 or later the information in the register is of very high quality.

For any event in the register we have the exact date (year, month and day) of that event with one exception: the time of changes of marital status before 1961 are recorded in integer calendar years only. In our analysis we have assumed that these events occurred in the middle of the year. This means that the calculated exposure times may be wrong by as much as six months. Furthermore, if a women gave birth to a child and married (or was divorced) in the same calendar year, we may be wrong about which event occurred first. Thus, the results derived in our nonresponse study cannot be used to analyze early life histories for the older cohorts (those born in 1936-45),. We think, however, that they are appropriate for our purposes, namely to describe the effects of nonresponse on estimates of various parameters.

The fertility register and the fertility survey do not cover precisely the same population. First, the register contains data on women who were Swedish citizens on 1 November 1960, while the survey covers women born in Sweden in 1936-60. Thus, women born in November or December 1960 are included in the survey but not in the register, and naturalized women are excluded from the target population of the survey if born outside Sweden. Secondly, the register, but not the survey, includes women who died or emigrated before February, 1981. These discrepancies must be unimportant as regards the nonresponse study and are disregarded here. In fact, we have found no systematic difference between various estimates based on data for the register population and corresponding estimates based on data for the entire target sample.

For our nonresponse study we have used register information about events up to the end of 1980 for women in the target sample of the survey. For the entire register population, however, we have used events only up to the end of 1977. This means that we have different censoring times for the entire population and the target sample, and that we cannot compare results for the two at higher ages for younger cohorts. See Diagram 1 below.

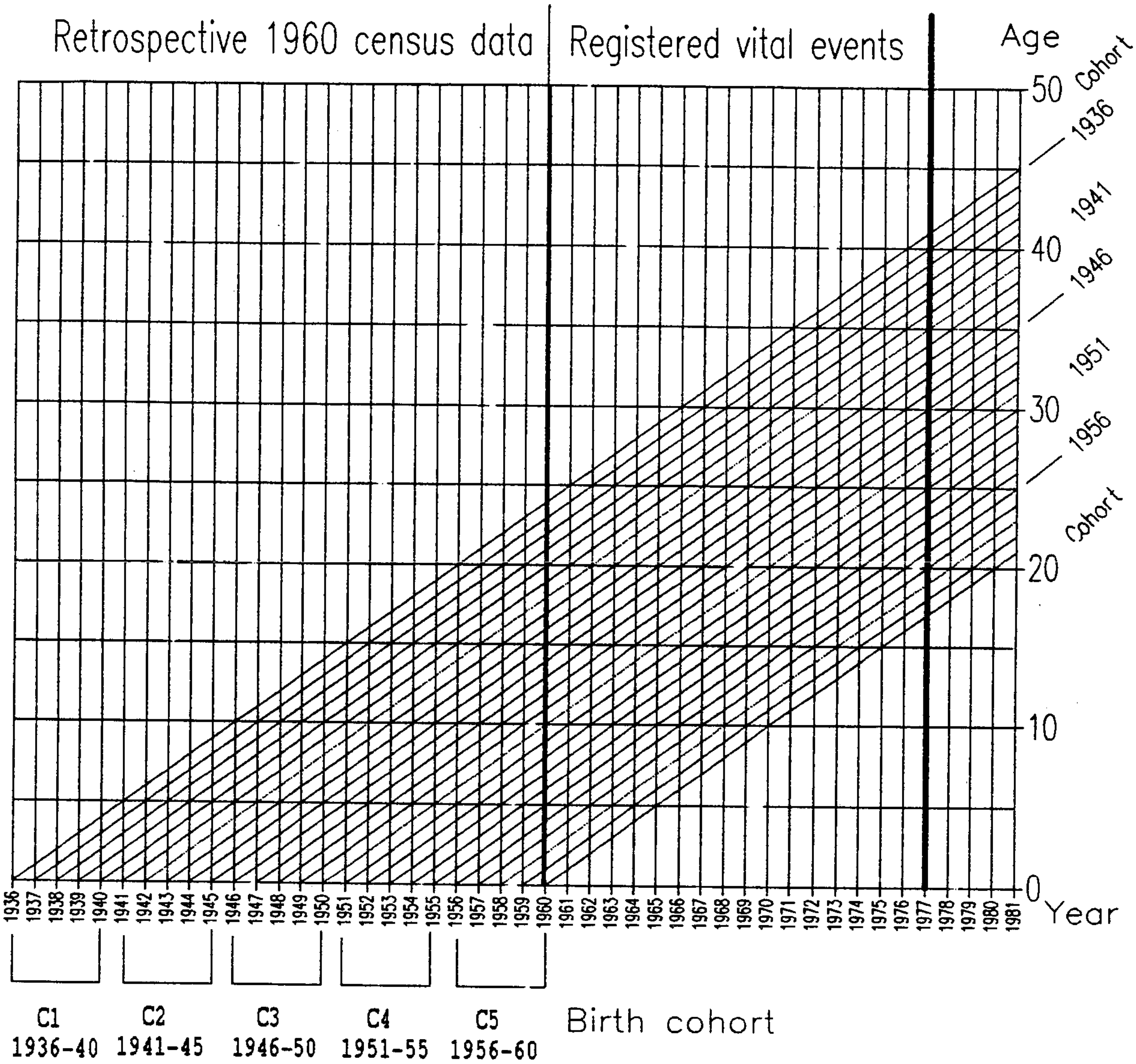


Diagram 1. Lexis diagram of data used in the evaluation study: census data up to the end of 1960 and registered vital events up to the end of 1977 for the entire population and up to the end of 1980 for the sample, respectively.

3. METHODS

3.1 Transition models used in the nonresponse study

The aim of the fertility survey project is not only to analyze the fertility (decline) from a conventional, purely demographic point of view. This could better be done from the register data, and has already been carried out in part in the three studies mentioned above. Instead, the survey aims also to analyze the interaction between fertility and other variables such as cohabitation, social background, education, child-care, and so on, which are not present in the register. The life histories gathered in the fertility survey will be analyzed by means of semi-Markov chain models. For instance, family life histories and fertility life histories can be described by the hierarchical models whose state and transitions diagrams are depicted in Figures 1 and 2, respectively.

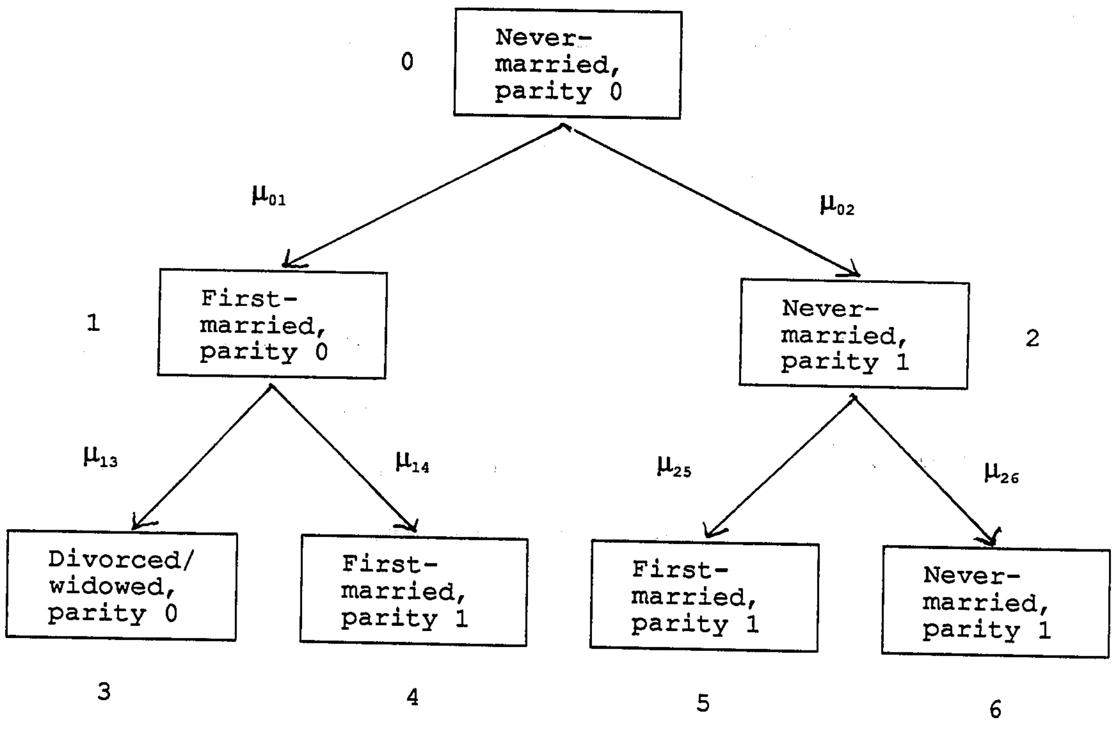


Figure 1. Model of family life histories.

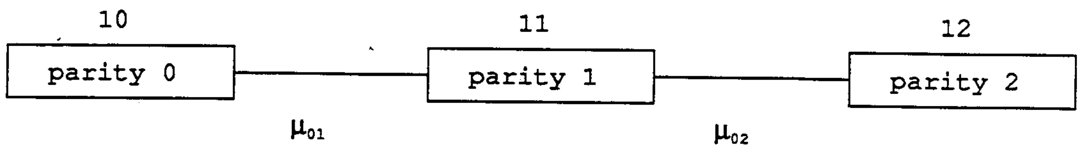


Figure 2. Model of fertility life histories.

The model in Figure 1 was used in one of the register studies (Quist, 1983). In that investigation it was not possible to study the changes in cohabitational incidences, since there is no information about consensual unions in the register. Instead, cohabitational behavior and its interaction with marriage and fertility are studied by means of the survey data. Results of the analysis of the early phases of a woman's adult family life are given in Hoem and Rennermalm (1982).

Model like the one in Figure 1 can be applied to survey data of various groups of women, for instance to women of various social backgrounds. They can also be used to study the interaction between fertility and employment history, and so on.

It would be nice to investigate the effects of nonresponse on such studies directly geared to the main purpose of the fertility survey project. Our opportunities of comparison with a more complete data source are restricted to the variables of the fertility register, however, and we must confine our nonresponse investigation to births and changes of marital status, as in Figure 1 and 2.

Each model has an initial state, State 0 and State 10, respectively, in which every women starts out. The transition intensity functions are defined as

$$\mu_{ij}(s) \quad = \quad \lim_{h\downarrow 0} P_{ij}(s,s+h)/h,$$

where $P_{ij}(s,s+h)$ is the transition probability from State i to State j during the time interval (s,s+h) among individuals who belong to State i at seniority s in that state.

In what follows we refer to the model of Figure 1. We assume that intensities μ_{ij} (i=1,2 and j=3,4,5,6) depend on time s (current duration") and on age t attained at entry into State i. Thus, they are functions, $\mu_{ij}(s|t)$, of two time parameters: current duration and starting age. All women enter State 0 at age 0 and are assumed to remain in that state until age 15. Thus, for transitions out of State 0, the intensities μ_{0i} depend on current duration s=x (or x−15) alone, where x coincides with age.

Most of the life history analyses of the Swedish fertility survey are made in terms of estimated transition intensities (see, e.g., Hoem and Rennermalm, 1982). Other relevant parameters such as transition probabilities, expected sojourn times, and so on, are functions of the intensities and can easily be estimated by means of estimates of those. For future reference we give the formulas for some of these functions.

Define $\mu_i = \sum_{j \neq i} \mu_{ij}$ and let

$$\ell_i(s|t) = \exp\{-\int_0^s \mu_i(u|t)\,du\} \quad \text{for } s \geq 0$$

be the probability of still belonging to the nonabsorbing State i (i=0,1,2) at age/duration s in that state among individuals who entered it at age/duration t in the previous state (where t=0 or 15 for i=0). Also let

$$\ell_{ij}(s|t) = \int_0^s \ell_i(u|t)\,\mu_{ij}(u|t)\,du \quad \text{for } s \geq 0$$

be the probability of entering State j from State i before age/duration s in that state among individuals who entered it at age/duration t in the previous state. Furthermore, let

$$P_i(x) = \int_0^x \ell_0(u)\ \mu_{0i}\ell_i(x-u|u)\ du$$

be the unconditional probability of having entered the nonabsorbing State i (i=1,2) and still belonging to it at age x, and let

$$P_{ij}(x) = \int_0^x \ell_0(u)\ \mu_{0i}(u)\ \ell_{ij}(x-u|u)\ du$$

be the unconditional probability of having entered the nonabsorbing State i (i=1,2) and having left it due to decrement j (j=3,4 and 5,6, respectively) before age x.

By assuming piecewise constant intensities we can compute the formulas above recursively, as will be shown in Section 3.3.

3.2. Nonresponse model

The nonresponse model presented in Lyberg (1983) refers to a competing risks model with right censoring, i.e., a Markov chain with continuous time parameter, one transient state and some finite number K of absorbing states. It is assumed that the transition intensities are constant during some relevant time period. The more complex hierarchical semi-Markov chain models of Figures 1 and 2 can easily be decomposed into a number of such competing risks models if we assume piecewise constant transition intensities.

It is well known that the occurrence/exposure rates (see Section 3.3) are the maximum likelihood estimators of the underlying constant transition intensities, provided that the rates are based on independent, identically distributed life history segments. This requirement is met if the observations constitute a random, noninformative sample from a homogeneous population of stochastically independent individuals (see Hoem, 1983a). In a survey, however, a noninformative sampling plan is likely to be distorted by selective nonresponse.

In the nonresponse model we assume that the response behavior is stochastic with response probabilities which depend on whether decrement has occurred and from what cause during some relevant time period. By means of that model we have investigated the nonresponse effects on the technical bias, variance, variance estimator, and the nonresponse bias of occurrence/exposure rates.

We have also studied two methods of adjusting for the nonresponse bias. We have found that the technical bias (i.e., the bias due to ratio estimation) and the variance in most situations can be treated in the same manner as when there is no nonresponse. The technical bias is mostly insignificant compared to the standard error, and the nonresponse bias of the usual variance estimator (see Section 3.3) is very small for realistic values of response probabilities and intensities. Thus, the most serious effect of nonresponse is probably that it may introduce nonresponse bias in the estimates of the intensities themselves and in other parameters derived from them. With only one cause of decrement the relative nonresponse bias of the occurrence/exposure rate $\hat{\mu}_R$, based on the respondents, becomes

$$rb(\hat{\mu}_R)) = (\mu_R-\mu)/\mu \approx [\gamma_1-\gamma_0]/[\gamma_0+\gamma_1 g(\mu)]$$

$$= [\gamma_1-\gamma]/[\gamma-\gamma_1 h(\mu)],$$

where γ_1 and γ_0 are the response probabilities among the decrements and the survivors, respectively

$$g(\mu) = q/pz\mu - 1 , \tag{3.1}$$

$$h(\mu) = 1 - q/z\mu , \tag{3.2}$$

and z is the censoring time, $\gamma=q\gamma_1+p\gamma_0$ is the overall expected response rate, and p=1-1=exp[-zμ] is the probability of survival.

By means of Taylor expansions of (3.1) and (3.2) it can be shown that

$$\frac{|\gamma_1-\gamma|}{\gamma} < |rb(\hat{\mu}_R)| < \frac{|\gamma_1-\gamma_0|}{\gamma_0}, \tag{3.3}$$

and that

$$|rb(\hat{\mu}_R)| \to |\gamma_1-\gamma|/\gamma \approx 0 \quad \text{when } z\mu \to \infty,$$

and

$$|rb(\hat{\mu}_R)| \to |\gamma_1-\gamma_0|/\gamma_0 \quad \text{when } z\mu \to 0.$$

Clearly, the relative nonresponse bias can be very large for estimates of small intensities, if the response probabilities for decrements and survivors differ greatly. Adjusting for that bias may, however, increase the bias (Lyberg, 1983).

In the life history models of Figures 1 and 2 we have two causes of decrement from every state. The nonresponse bias of the rate for one cause then depends on the response probabilities among both types of decrements and survivors. Let γ_1, γ_2 and γ_0 denote the response probabilities among the two types of decrements and the survivors, respectively, and let μ_1, μ_2 and $\mu=\mu_1+\mu_2$ denote the transition intensities. The relative nonresponse bias of the occurrence/exposure rate $\hat{\mu}_{R1}$, an estimate of μ_1 based on the respondents, then becomes

$$rb(\hat{\mu}_{R1}) = \frac{pz}{t_R}\,[(\gamma_1-\gamma_0) + (\gamma_1-\gamma_2)\,\frac{\mu_2}{\mu}\,g(\mu)], \tag{3.4}$$

where t_R is the expected value of the individual observed exposure time and $g(\mu)$, p and z were defined above. Thus, if $\gamma_1>\gamma_0$, the nonresponse bias of $\hat{\mu}_{R1}$ is probably positive and strengthened if $\gamma_1>\gamma_2$ but reduced if $\gamma_2>\gamma_1$. In the latter case, the bias may even be negative if μ_2 and μ are large.

Of course, the response behavior of an individual depends on her entire life history up to the time of the interview. The response probabilities above are those of individuals who still belong to some initial state at the beginning of some relevant time period, and the are conditional on the outcomes during that period only. Let R=1 if an individual responds and R=0 otherwise, and let I(v) (I(v)=0,1) denote the state held by an individual at seniority v. (For simplicity we deal with the case with one cause of decrement only.) The response probabilities for time period $r<t<r+1$ are then defined as

$$\gamma_{r1} = P[R=1|I(r)=0,I(r+1)=1],$$

and

$$\gamma_{r0} = P[R=1|I(r+1)=0],$$

respectively. Let $\gamma_1(v)$ and $\gamma_0(v)=\gamma_{v0}$ denote the response probabilities among individuals who at seniority v have left the initial state and still belong to it, respectively. Assume that $\gamma_{r1}=\gamma_1(v)$ for some $r\leq v+1$, i.e., that individual who have left the initial state at seniorities r to v have the same response probability. Then

$$\gamma_{r0} = \frac{\ell_0(v)}{\ell_0(r+1)}\gamma_0(v) + \frac{\ell_0(r+1) - \ell_0(v)}{\ell_0(r+1)}\gamma_1(v),$$

and

$$\frac{\gamma_{r1} - \gamma_{r0}}{\gamma_{r0}} = \frac{\gamma_1(v) - \gamma_0(v)}{\gamma_0(v)}\frac{\ell_0(v)}{\ell_0(r+1)}\left[1 + \frac{\gamma_1(v) - \gamma_0(v)}{\gamma_0(v)}\frac{\ell_0(r+1) - \ell_0(v)}{\ell_0(r+1)}\right],$$

which yields

$$\frac{|\gamma_{r1} - \gamma_{r0}|}{\gamma_{r0}} \leq \frac{|\gamma_1(v) - \gamma_0(v)|}{\gamma_0(v)}.$$

It can also easily be shown that $|\gamma_{r1} - \gamma_{r0}|/\gamma_{r0}$ increases as r increases, i.e., the closer the relevant period is to the time of the interview.

3.3. Statistical methods

The response probability for a group of women is estimated by the corresponding response rate obtained in the survey, $\hat{\gamma} = n_R/n_s$, where n_R and n_s are respectively the number of respondents and the number of individuals in the target sample. The variance of such an estimate is estimated by $v(\hat{\gamma}) = \hat{\gamma}(1-\hat{\gamma})/(n_s-1)$.

For each single year of duration we have counted the number of occurrences (OCC) of a particular decrement and computed the exposures (EXP) from data extracted from the Swedish fertility register. (Note that the individual exposures sometimes are based on imputed dates of marriages; see Section 2.3.). The influence of age in States 1 and 2 of Figure 1 and in State 11 of Figure 2 has been treated by forming groups for ages 15-19, 20-24 and 25-29 years of age at entry into these states.

The assumed piecewise constant transition intensities are estimated by occurrence/exposure rates $\hat{\mu}$=OCC/EXP for each five-year birth cohort (and age group at entry) based on data for the entire population (P), for the entire target sample (S), and for the respondents (R). The assumption of piecewise constant intensities means that $\mu(v+s|r+t)$ is estimated by $\hat{\mu}(v|r)$ for integer v and r and for $0<s,\ t\leq 1$.

The nonresponse bias and the relative nonresponse bias of the occurrence/exposure rates are estimated by

$$\hat{b}(\hat{\mu}_R) = \hat{\mu}_R - \hat{\mu}_S, \tag{3.5}$$

and

$$\hat{rb}(\hat{\mu}_R) = (\hat{\mu}_r - \hat{\mu}_S)/\hat{\mu}_S, \tag{3.6}$$

respectively, where $\hat{\mu}_R$ and $\hat{\mu}_S$ are the rates based on data for the respondents and the entire target sample, respectively. The sample rate $\hat{\mu}_S$ is a consistent estimate (predictor) of the corresponding population rate $\hat{\mu}_P$, as well as of the underlying transition intensity μ. This means that (3.5) and (3.6) provide consistent estimates of the nonresponse bias whether one applies a fixed-population approach aiming at predicting $\hat{\mu}_P$ or, as we do, applies a model-based approach aiming at estimating the intensity μ.

The model-based variance of the occurrence/exposure rate $\hat{\mu}$ is estimated by the standard variance estimator $v(\hat{\mu}) = \hat{\mu}/EXP$, whose nonresponse bias is very small in most cases (Lyberg, 1983).

Let two population groups have intensities $\mu^{(1)}$ and $\mu^{(2)}$. To test the hypothesis H_0: $\mu^{(1)} = \mu^{(2)}$, we use the following test statistic:

$$U = \frac{\hat{\mu}^{(1)} - \hat{\mu}^{(2)}}{\{\hat{\mu}^{(1)}/EXP^{(1)} + \hat{\mu}^{(2)}/EXP^{(2)}\}^{\frac{1}{2}}}, \tag{3.7}$$

where the denominator is a consistent estimator of the asymptotic standard error of the numerator, provided that the rates $\hat{\mu}^{(1)}$ and $\hat{\mu}^{(2)}$ are asymptotically independent. Then U is asymptotically normally distributed (0,1) under H_0. (See Aalen and Hoem (1978, Section 4.6), and Hoem and Funck Jensen (1982, Section 4.1)). Schou and Væth (1980) show that the normality approximation is acceptable if the expected number of occurrences exceeds about 10 in each group.

To compare two sets of occurrence/exposure rates concerning the same sequence of age or duration groups we use the test statistic

$$\tilde{U} = \Sigma_x U_x/\sqrt{m}, \tag{3.8}$$

where $\{U_x\}$ are the age- or duration-specific rates formed according to (3.7) and m is the number of age or duration groups involved. In this case we test the hypothesis H_0: $\mu_x^{(1)} = \mu_x^{(2)}$ for all x, against the alternative that $\mu_x^{(1)} < \mu_x^{(2)}$ (or the opposite) for at least some x.

The conditional transition probabilities defined in Section 3.1 are estimated by the following recursive formulas for integer v, r and x, and for $0<s,\ t\leq 1$:

$$\hat{\ell}_i(0|r+t) = 1, \qquad \hat{\ell}_{ij}(0|r+t) = 0,$$

$$\hat{\ell}_i(v+s|r+t) = \hat{\ell}_i(v|r)\ \exp\{-s\hat{\mu}_i(v|r)\},$$

and

$$\hat{\ell}_{ij}(v+s|r+t) = \hat{\ell}_{ij}(v|r) + \frac{\hat{\mu}_{ij}(v|r)}{\hat{\mu}_i(v|r)}\ [\hat{\ell}_i(v|r) - \hat{\ell}_i(v+s|r)],$$

where $\hat{\mu}_{ij}(v|r)$ is the estimate of $\mu_{ij}(v+s|r+t)$ for $0<s,\ t\leq 1$.

The probabilities of belonging to a specific state at age x+t and x are then estimated by

$$\hat{P}_0(x+t) = \hat{\ell}_0(x+t|0) = \hat{\ell}_0(x+t) = \prod_{r=0}^{x-1} \exp\{-\hat{\mu}_0(r)\}\ \exp\{-t\hat{\mu}_0(x)\},$$

$$\hat{P}_i(x) = \sum_{r=0}^{x-1} \int_0^1 \hat{P}_0(r+t)\ \hat{\mu}_{0i}(r)\ \hat{\ell}_i(x-r-t|r)\ dt \equiv \sum_{r=0}^{x-1} \hat{w}_i(x,r),$$

and

$$\hat{P}_{ij}(x) = \sum_{r=0}^{x-1} \frac{\hat{\mu}_{0i}(r)}{\hat{\mu}_0(r)}\ [\hat{P}_0(r) - \hat{P}_0(r+1)] + \frac{\hat{\mu}_{ij}(x-r-1|r)}{\hat{\mu}_i(x-r-1|r)}\ \hat{\ell}_i(x-r-1|r)$$

$$- \sum_{r=1}^{x-1} \frac{\hat{\mu}_{ij}(x-r-1|r)}{\hat{\mu}_i(x-r-1|r)}\ \hat{w}_i(x,r),$$

where

$$\hat{w}_i(x,r) = \frac{\hat{\mu}_{0i}(r)}{\hat{\mu}_0(r) - \hat{\mu}_i(x-r-1|r)}\ [\hat{\ell}_0(r)\ \hat{\ell}_i(x-r|r) - \hat{\ell}_0(r-1)\ \hat{\ell}_i(x-r-1|r)]$$

for i=1,2 and j=3,4 adn j=5,6, respectively.

Our estimates based on the target sample (S) and the respondents (R) are subject to rather large stochastic errors, errors which increase as the number of cases and the exposures decrease. Table 2 below shows the number of cases and the exposures in various states at selected ages/durations. The table refers to women born in 1941-45. The sample size for this cohort was about the same as for the cohorts of women born in

1946-50 and 1951-55, respectively, but about twice and one and a half times as large as for the cohorts of women born in 1936-40 and 1956-60, respectively (see Table 1 above). Thus, for the oldest and youngest cohorts of women the figures are about half and two-thirds, respectively, of those in Table 2.

For the intermediate States 1 and 2 of Figure 1 the number of cases is rapidly reduced, and we can study only the few first years of duration in those states.

Table 2. Number of women and exposures in the states of Figures 1 and 2 at selected ages/durations in the entire population (P), the entire target sample (S), and among the respondents (R), respectively, in the 1981 Swedish fertility survey. Women born in 1941-45.

State	Age at entry	Age/ duration (years)	Number of women			Exposure in women-years		
			P	S	R	P	S	R
0: Never-married, parity 0	0	15	278197	1180	1014	278056	1179	1013
		20	216091	933	807	201330	871	751
		25	72980	332	269	65752	305	249
		30	32060	149	114	30239	141	106
		35	6910	105	72	4954	101	68
1: First-married, parity 0	15-19	0	32706	137	111	17737	74	60
		1	8606	33	27	6635	22	17
		2	5071	14	11	4109	12	9
	20-24	0	120131	502	447	95698	397	351
		1	72667	303	264	59018	240	209
		2	46793	183	161	38097	149	129
		5	15846	71	54	14060	66	49
		10	5716	31	24	4966	30	23
	25-29	0	31686	134	116	26444	109	93
		1	20332	78	65	16268	61	50
		2	12780	48	40	10533	37	31
		5	4607	21	19	3987	19	17
2: Never-married, parity 1	15-19	0	26191	103	89	21450	80	69
		1	17810	67	57	14877	58	50
		2	12287	50	43	10312	41	34
		3	8632	33	27	7371	25	20
	20-24	0	21835	99	91	18077	84	76
		1	15344	74	67	13566	63	56
		2	11955	53	47	10670	48	42
		5	6555	26	22	6082	23	19
	25-29	0	8396	50	40	7594	47	38
		1	7162	45	37	6617	41	33
		2	6035	35	30	5425	27	23
10: Parity 0	0	15		1180	1014		1179	1013
		20		992	858		941	812
		25		513	421		474	387
		30		248	194		233	180
		35		181	176		131	126
11: Parity 1	15-19	0		183	151		182	150
		2		154	127		126	103
		5		50	47		38	36
	20-24	0		475	433		474	432
		2		406	371		355	324
		5		149	135		140	126
		10		84	74		81	71
		15		50	43		41	35
	25-29	0		265	227		265	227
		5		99	83		92	77
		10		53	46		42	35

4. RESULTS

4.1. Response rates

In the 1981 Swedish fertility survey the response rates differed between women with their own children (18 years old or younger) in their households and "childless" women. Among women born in 1941-45 the response rate was 17 percentage points higher in the former as in the latter (Table 1). According to our experience at Statistics Sweden, the response rates in various surveys are also often higher among married people than among others (Lindström, 1983, Chapter 5). Thus, there seems to be a connection between the present family circumstances and the probability to respond in a (fertility) survey.

In life history analysis we are not, however, aiming at describing current circumstances. This means that we should be less concerned about any differences in the response rates between women with different present status than about differences between women with different life histories, i.e., between decrements and survivors, respectively, at selected ages or durations in a state. Of course, there is a connection between the differences in response rates according to present status and differences according to earlier behavior. In general, we would expect the former differences to be larger than the latter, since the group of survivors in any state becomes more homogeneous as age increases, while the response behavior of the decrements probably is rather independent of when they left the initial state (see the end of Section 3.2). This assumption is verified by our empirical findings concerning response rates according to occurrences of first birth. See Table 3 below and Table 5 at the end of this section.

The response rates are higher among new mothers in an age interval than among those who remain childless during the interval, and the differences increase with age. This is true for all cohorts and all selected age intervals with two exceptions for the youngest ages, 15 to 19 (Table 5).

Table 3. Response rates among women born in 1941-45 and still belonging to State 10 of Figure 2 (nullipara) at selected ages, and among those who left that state in selected age intervals. 1981 Swedish fertility survey. Percent.

Ages	Nullipara at the start $\hat{\gamma}$	Nullipara at the end $\hat{\gamma}_0$	New mothers in the age interval $\hat{\gamma}_1$
15-19	86	86	83
20-24	86	82	91
25-29	82	78	86
30-34	78	72	94
15-35	86	71	88

Taking the response rates as estimates of the corresponding response probabilities, we can estimate the limits in (3.3) of the relative nonresponse bias of occurrence/exposure rates. Table 4 shows such estimated limits concerning first birth rates. The figures show that we can expect a moderate nonresponse bias (less than 10 percent) in the estimates of first birth intensities, except for ages over thirty, where the relative bias may be very large (up to 30 percent).

Table 4. Estimated limits of relative nonresponse bias of age-specific first-birth rates (estimated intensities) in the 1981 Swedish Fertility survey. Percent.

Age	Birth cohort				
	1936-40	1941-45	1946-50	1951-55	156-60
15-19	-0, -0	-4, -3	8, 9	-2, -2	5, 6
20-24	1, 1	5, 11	3, 6	2, 3	
25-29	5, 13	4, 10	4, 9	2, 3	
30-34	12, 18	20, 30			

Tables 3 and 4 refer to all women at parity 0, irrespective of marital status. This group is not homogeneous: married and unmarried women show entirely different fertility behavior. (This is the case even in Sweden, where nonmarital cohabitation has become more and more common. See Hoem and Rennermalm, 1982, and Quist, 1983.) As regards the connection between fertility and response behavior, we have not been able to find that it interacts with marital status. See Table 5 concerning first birth and Table 7 concerning second birth. (The tables appear at the end of this section.) The stochastic variations are large, but the connection between occurrence of first birth at different ages and response behavior seems to be similar for never-married and first-married women: the response rates are higher among new mothers than among those still childless in each selected age interval, except for the youngest ones. The results are similar for women of parity 1 at durations of 0 to 4 years since first birth, but the differences between women still of parity 1 and those who attain parity 2 are small.

The response rates are also higher among women who marry as never-married nulliparas in an age interval than among those who remain in the latter state (see Table 6 at the end of this section).

This is probably due to the strong connection between first-marriage and first-birth. The response rates among decrements from state 0 (never-married, parity 0) are about the same no matter whether the decrement is caused by marriage or childbirth. Furthermore, among women who marry as never-married nulliparas at ages 20 to 29, the response rates are higher among those who give birth to their first child at durations of 0 to 4 years of marriage than among those who remain childless (Table 8 at the end of this section).

In general, the response rates are higher among the decrements (i.e., the new mothers and the newly married) than among those who remain in a state during an age interval. The differences in response rates between decrements and survivors do not exceed the corresponding differences according to present status, however, except for the oldest ages. The largest differences we have found concern transitions from State 0 (never-married, parity 0) at ages 30 to 34 for women born in 1941-45: the response rates are 69 percent among those who remain as never-married and childless compared with 96 percent and 94 percent for the newly married and new mothers, respectively. In general, the response rates among survivors in an age interval exceed 75 percent and the response rates among decrements are 85 to 95 percent.

To get rough estimates of how serious the nonresponse bias may be in life history analysis, we can use the response rates for various relevant groups defined by present status. Such rates are often known in practice. We should remember, however, that this probably gives a too pessimistic picture of the nonresponse bias of estimates related to periods far earlier than the time of the interview.

Table 5. Response rates among women who remain in State 10 of Figure 2 (nullipara) at selected ages, and among those who leave that state in selected age intervals, by marital status. 1981 Swedish 1981 Swedish fertility survey. Percent.

Cohort Women born in	Ages	Nullipara at the start of / end of the age interval $\hat{\gamma}$	$\hat{\gamma}_0$	New mothers in the age interval Total $\hat{\gamma}_1$	Married	Never-married	Number of women in the target sample[1]
1936-40	15-19	82	82	82	86	79	514
	20-24	82	82	83	83	82	282
	25-29	82	76	86+[4)]	85	(94)[2)]	123
	30-34	76	73	84	83	..[3)]	85
	35-39	73	72	..	..	..	81
	15-40	82	71	84+	84	83	79
1941-45	15-19	86	86	83	79	86	992
	20-24	86	82	91+	91	92	513
	25-29	82	78	86+	87	80	248
	30-34	78	72	94	94	(94)	181
	15-35	86	71	88+	89	88	174
1945-50	15-19	87	85	93+	97	92	959
	20-24	85	83	88+	89	87	527
	25-29	83	79	87+	87	85	265
	15-30	87	79	89+	89	88	238
1951-55	15-19	88	89	87	(95)	86	1022
	20-24	89	88	90	91	90	633
	15-25	88	86	90+	93++[5)]	89	564
1956-60	15-19	88	87	93	..	92	725
	15-20	88	87	90	87	90	679

1) Still in the state at the end of the age interval.

2), 3) Less than 20 and 10 observations, respectively.

4) Significantly larger than $\hat{\gamma}_0$ according to the standard t-test, one-sided, at the 5 percent significance level.

5) Significantly larger than the corresponding rate among never-married mothers.

Table 6. Response rates among women who remain in State 0 of Figure 1 (never-married, parity 0) at selected ages, and among those who leave that state in selected age intervals, by cause of decrement. 1981 Swedish fertility survey. Percent.

Cohort, women born in	Ages	Women in State 0 at the start of the age interval $\hat{\gamma}$	Women in State 0 at the end of the age interval $\hat{\gamma}_0$	Women leaving State 0 in the age interval, due to marriage $\hat{\gamma}_1$	Women leaving State 0 in the age interval, due to birth $\hat{\gamma}_2$	Number of women in the target sample 1)
1936-40	15-19	82	82	84	79	484
	20-24	82	81	84	82 2)	182
	25-29	81	77	82	(94) 3)	71
	30-34	77	74	(84)	..	48
	35-39	74	74	..	..	38
	15-40	82	72	83	83	36
1941-45	15-19	86	86	82	86	933
	20-24	86	81	89+ 4)	92+	332
	25-29	81	77	86+	80	149
	30-34	77	69	96+	(94)	105
	15-35	86	66	88+	88+	98
1946-50	15-19	87	85	91+	92+	914
	20-24	85	82	88+	87	383
	25-29	82	77	88+	85	177
	15-30	87	76	89+	88+	160
1951-55	15-19	88	88	96+	86	995
	20-24	88	85	93+	90+	517
	15-25	88	84	94++ 5)	89+	442
1956-60	15-19	88	87	93	92	705
	15-20	88	87	89	90	565

1), 2), 3), 4) See Table 5.

5) Significantly larger than $\hat{\gamma}_0$ and $\hat{\gamma}_2$ according to the standard t-test, one-sided, at the 5 percent significance level.

Table 7. Response rates among women who enter State 11 of Figure 2 (parity 1) at ages 20 to 29[6)], in total and among those who at durations 0 to 4 years remain in the state and leave it, respectively. 1981 Swedish fertility survey. Percent.

Cohort Women born in	Total $\hat{\gamma}$	Women still at parity 1 $\hat{\gamma}_0$	Women attaining parity 2 All mothers $\hat{\gamma}_1$	 married mothers	 unmarried mothers	Number of women in the target sample 1)
1936-40	84	81	85	85	(94)[2)]	108
1941-45	89	88	90	90	85	248
1946-50	88	82	91+[4)]	91	89	201
1951-55	90	88	92	92	94	84

1)
2) See Table 5.
3)

6) Ages 20 to 24 for cohorts born in 1951-55.

Table 8. Response rates among women who enter State 1 of Figure 1 (married, parity 0) at ages 20 to 29[6)], in total and among those who at durations 0 to 4 years of marriage remain in the state and leave it, respectively.
1981 Swedish fertility survey. Percent.

Cohort Women born in	Total $\hat{\gamma}$	Women still in the state $\hat{\gamma}_0$	Women leaving the state, due to birth $\hat{\gamma}_1$	 dissolution of marriage $\hat{\gamma}_2$	Number of women in the target sample 1)
1936-40	83	77	84	..3)	52
1941-45	89	80	90+[4)]	..	93
1946-50	88	84	89	(88)[2)]	76
1951-55	93	91	92	100+	34

1)
2)
3) See Tables 5 and 7.
4)
6)

4.2. The size of the nonresponse bias and stochastic errors

For all transitions defined by the models of Figures 1 and 2 we have estimated the transition intensities and the variance of these estimates by means of data for the entire population (P), the entire target sample (S), and the respondents (R) in the 1981 Swedish fertility survey. (P-estimates have been calculated for the model of Figure 1 only.) In this manner we have been able to estimate the nonresponse bias directly, both its sign and how large it is in itself and as compared to the estimated intensity and its standards error. We have also been able to study how much the stochastic errors increase due to sampling and nonresponse.

We have found that the nonresponse bias is about what we would expect from the computed group-specific response rates presented in Table 1 and in the previous section: in general, the first-birth and first-marriage intensities are overestimated due to the nonresponse. The bias is moderate for the youngest ages but may be substantial for the older ages (around 30 and more), where it may amount to 20 to 30 percent of the estimated intensity. See Tables 10 to 12 at the end of this section.

For the intensities of transitions from the intermediate State 1 of Figure 1 and State 11 of Figure 2, the estimated nonresponse bias is in general positive but small. It seldom exceeds 10 percent of the estimated intensity during the first five years of duration in the states. We cannot study longer durations, since the number of cases becomes very small.

The stochastic errors are rather large for age-specific estimates based on data for the target sample (S) and the respondents (R), and they increase rapidly with age. We have found peaks in age-specific intensity curves which are caused by stochastic errors and have been strengthened by the nonresponse bias. In most of these cases, however, the false peak is not found to be significant when tested by the test statistic U in (3.7).

The estimated standard errors are about 15 times as large for estimates based on the target sample (S-estimates) as for estimates based on data for the entire population (P-estimates), while the standard errors of estimates based on the respondents (R-estimates) are only about 10 percent larger than those of S-estimates. Thus, the increase of the stochastic errors caused by the nonresponse is negligible compared with the systematic errors caused by the selective response behaviour.

Figure 3 shows the age-specific first-birth rates among all women (irrespective of marital status) born in 1941-45, based on data for the entire target sample (S) and the respondents (R).

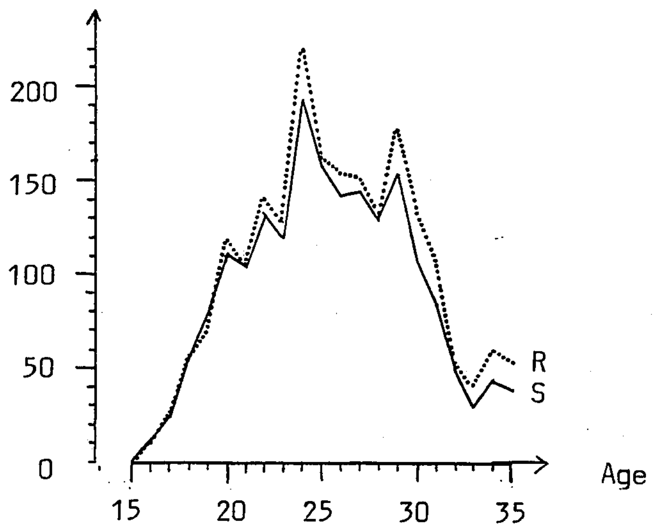


Figure 3. Age-specific first-birth rates based on data for the entire target sample (S) and the respondents (R) in the 1981 Swedish fertility survey.
Women born in 1941-45. Per 1000 women per year.

The curves are very close for the youngest ages (up to 20), but for older ages the R-rates are higher than the S-rates, throughout. The rates reach peaks at ages 24 to 29. The valley at ages 25 to 28 is not genuine, however. According to computations based on data for the register population, there is only one peak, and it is around age 26. See Table 9 below.

Table 9. Estimated intensities of first birth at ages 23 to 30 based on data for the entire population (P), the entire target sample (S), and the respondents (R) in the 1981 Swedish fertility survey. Women born in 1941-45.

Per 1000 women per year.

Age	Estimated intensities			Estimated nonresponse bias
	$\hat{\mu}_P$	$\hat{\mu}_S$	$\hat{\mu}_R$	$b(\hat{\mu}_R)$
23	138	120 ± 27[1)]	128 ± 30	8
24	152	192 ± 36	219 ± 42	27
25	165	158 ± 36	163 ± 40	5
26	168	142 ± 37	154 ± 42	12
27	158	145 ± 40	151 ± 45	7
28	142	129 ± 40	133 ± 45	4
29	125	155 ± 47	177 ± 57	22
30	107	107 ± 42	133 ± 53	26

1) Estimated 95 percent confidence limits.

We have calculated 95 percent confidence limits, although we can hardly assume that the mixed group of married and unmarried women is homogeneous. Nevertheless, the intervals suggest that the stochastic errors are rather large and that we should not put much faith in occasional peaks and crevasses in the diagrams based on data for the target sample or the respondents.

Figure 4 shows the estimated age-specific intensitites of transition out of State 0 of Figure 1 (never-married, parity 0) caused by first birth and first marriage, respectively. Like Figure 3 it refers to women born in 1941-45, but the results are similar for the other birth cohorts (see Table 11 at the end of this section.)
Note that the ordinate unit in the left-hand diagram of Figure 4 is only half that in the right-hand diagram, for the marriage intensities are much larger than the birth intensities among never-married nulliparas. In general both are overestimated if calculated by means of data for the respondents only. The false peak of the estimated first-birth intensities at age 29 for all women (Figure 3) corresponds to the similar false peaks of birth intensities for never-married

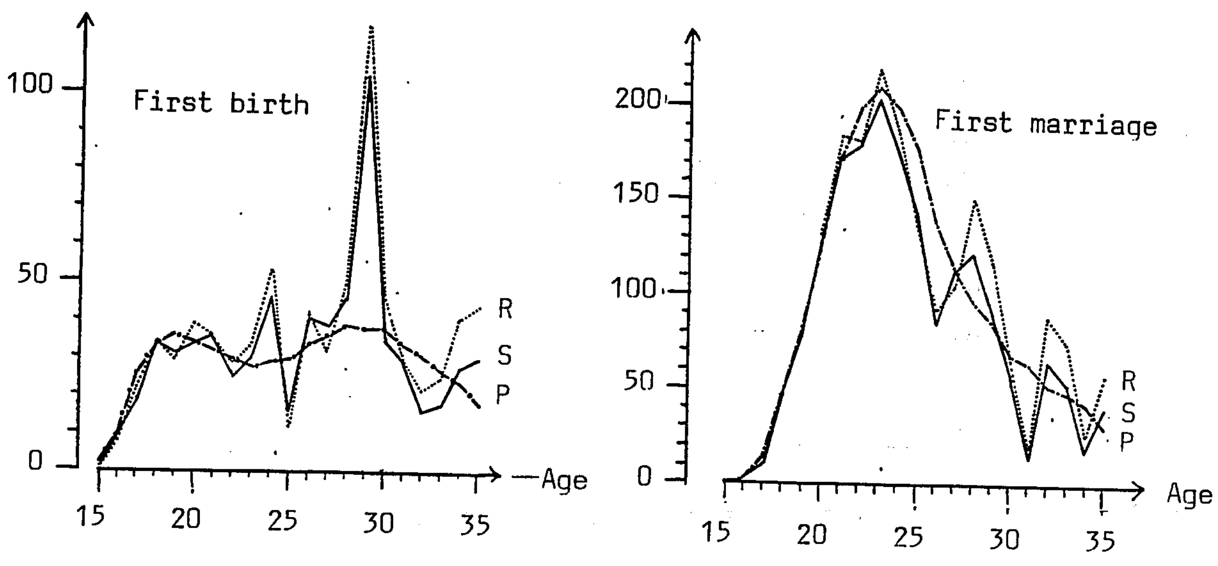


Figure 4. Estimated age-specific intensities of transition out of State 0 of Figure 1 (never-married, parity 0), caused by first birth and first marriage, respectively. Estimates based on data for the entire population (P), the entire target sample (S), and the respondents (R), respectively, in the 1981 Swedish fertility survey. Women born in 1941-45. Per 1000 women per year.

nulliparas at the same age and of their marriage intensities at age 28, respectively.

any wrong conslusions, if the standard errors are calculated and single extreme values are disregarded. See Table 12 at the end of this section.

The first-birth intensities among first-married nulliparas are also in general overestimated by data for the respondents only (see Figure 5 and Table 13 at the end of this section). During the first years of marriage the intensities of dissolution of a childless marriage are small as compared with the marital first-birth intensities, which are very large. (After ten years of childless marriage the first-birth intensities like the dissolution intensities are small, around 20 to 40 per 1000 women per year.) The large first-birth intensities during the first years of marriage imply that the number of cases decreases quickly and the stochastic errors increase very fast. After five years of marriage only 71 women in the target sample were still childless and married among the 502 who were born in 1941-45 and married at ages 20 to 24. During the fourth and later years of childless marriage we have at most ten births per year among women in the target sample. The peaks of first-birth intensities at a duration of three years of marriage based on S-data and R-data are not genuine, and they are not found to be significant according to test statistic U.

Let us look finally at some results concerning second-birth fertility. Figure 6 shows the second-birth rates by duration since first birth for women born in 1941-45 who gave birth to their first child at ages 20 to 24. The second-birth intensities are overestimated by data for the respondents only. This is true for all cohorts with respect to the first five years since first birth if this birth occurred at ages 20 to 24 or 25 to 29 (as far as we have been able to study it).

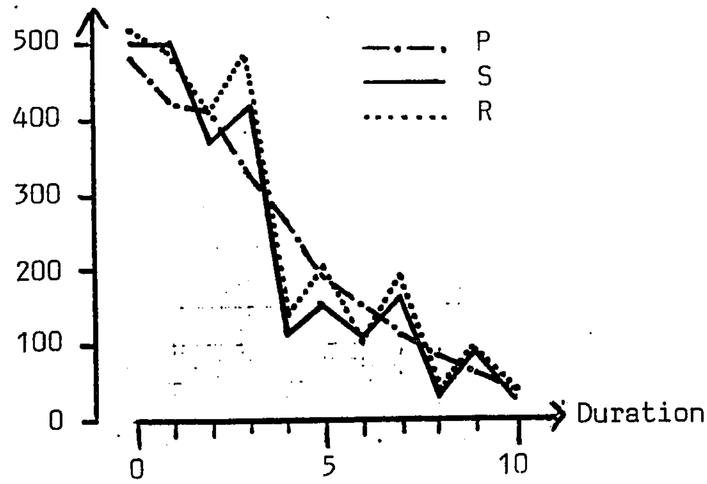


Figure 5. Estimated first-birth intensities among first-married nulliparas, by years of duration of marriage. Estimates based on data for the entire population (P), the entire sample (S), and the respondents (R), respectively, in the 1981 Swedish fertility survey. Women born in 1941-45 and married at ages 20 to 24. Per 1000 women per year.

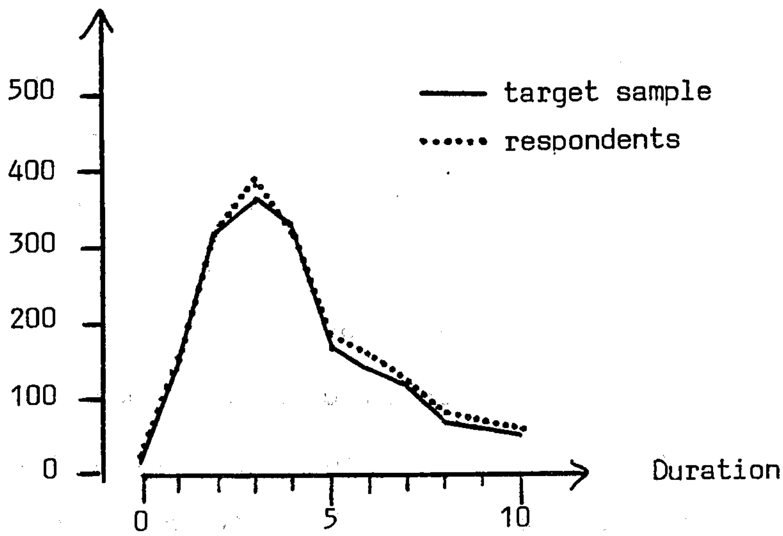


Figure 6. Second-birth rates by duration since first birth based on data for the entire target sample and the respondents, respectively, in the 1981 Swedish fertility survey. Women born in 1941-45 who gave birth to their first child at ages 20 to 24. Per 1000 women per year.

See Table 14 at the end of this section. The nonresponse bias, however, is very small; usually less than five per cent of the estimated intensity.

The second-birth rates for never-married women are subject to rather large stochastic errors. See Table 15 at the end of this section. The rates based on the respondent only seem to overestimate the unmarital second-birth fertility, while the marriage intensitites may be underestimated. Thus, the response probabilities among the decrements from the two causes may have different directions as compared to the response probabilities among the survivors, i.e., $\gamma_1 > \gamma_0 > \gamma_2$. If this is true, the nonresponse bias of the rates is strengthened by the differences between γ_1 and γ_2. See (3.4).

Table 10. First-birth rates for combined five-year age intervals, based on data for the entire population (P), the entire target sample (S), and the respondents (R) in the 1981 Swedish fertility survey. Per 1000 women per year.

Cohort, women born in	Age interval	First-birth rates			Estimated nonresponse bias	
		$\hat{\mu}_P$	$\hat{\mu}_S$	$\hat{\mu}_R$	per 1000 $b(\hat{\mu}_R)$	per cent $rb(\hat{\mu}_R)$
1936-40	15-19	33	32	32	-0	-0
	20-24	120	116	116	0	0
	25-29	157	168	182	14	8
	30-34	75	78	88	11	14
1941-45	15-19	37	33	32	-1	-3
	20-24	125	127	136	9	7
	25-29	151	146	155	9	6
	30-34	.[1)]	66	83	17	26
1946-50	15-19	41	41	44	3	8
	20-24	116	115	120	4	4
	25-29	.	140	149	9	7
1951-55	15-19	30	29	28	-1	-2
	20-24	.	94	97	3	3
1956-60	15-19	.	21	22	1	6

1) Data not available.

Table 11. Rates of transition out of State 0 of Figure 1 (never-married, parity 0) for combined five-year age intervals based on data for the entire population (P), the entire target sample (S), and the respondents (R) in the 1981 Swedish fertility survey. Per 1000 women per year.

Cause of decrement	Cohort, women born in	Age interval	Rates			Estimated non-response bias	
			$\hat{\mu}_P$	$\hat{\mu}_S$	$\hat{\mu}_R$	per 1000 $b(\hat{\mu}_R)$	per cent $rb(\hat{\mu}_R)$
First birth	1936-40	15-19	17	18	19	-1	-4
		20-24	31	34	34	-0	-1
		25-29	27	29	35	5	18
	1941-45	15-19	20	18	19	0	0
		20-24	31	34	37	3	9
		25-29	35	44	44	-0	-1
		30-34	32	26	34	8	29
	1946-50	15-19	27	30	32	2	7
		20-24	47	52	53	1	2
		25-29	61	57	60	3	6
	1951-55	15-19	25	26	25	-1	-3
		20-24	63	57	58	2	3
	1956-60	15-19	17	18	19	1	5
First marriage	1936-40	15-19	26	23	24	0	2
		20-24	167	153	155	1	1
		25-29	166	174	180	6	3
		30-34	58	69	78	8	12
	1941-45	15-19	26	26	25	-1	-5
		20-24	172	166	174	9	5
		25-29	130	113	120	7	7
		30-34	59	42	56	14	32
	1946-50	15-19	22	19	20	1	6
		20-24	119	116	120	4	4
		25-29	106	103	113	10	10
	1951-55	15-19	9	8	9	1	8
		20-24	70	72	76	4	6
	1956-60	15-19	5	7	8	0	6

Table 12. Intensities of transition out of State 0 of Figure 1 (never-married, parity 0) at ages 20 to 30. Estimates based on data for the entire population (P), the entire target sample (S), and the respondents (R) in the 1981 Swedish fertility survey. Women born in 1941-45. Per 1000 women per year.

Cause of decrement	Age	Estimated intensities			Estimated non-response bias	
		$\hat{\mu}_P$	$\hat{\mu}_S$	$\hat{\mu}_R$	per 1000 $b(\hat{\mu}_R)$	per cent $rb(\hat{\mu}_R)$
First birth	20	34 ± 1[1]	34 ± 12	39 ± 14	4	12
	21	32	36	36	-0	-1
	22	30	26	29	3	11
	23	28	30	34	3	11
	24	29 ± 1	46 ± 22	53 ± 26	7	15
	25	30 ± 1	(16 ± 14	12 ± 14	-4	-27)[2]
	26	34	41	42	0	1
	27	36	(39	32	-7	-18)
	28	39	(46 ± 30	50 ± 35	4	10)
	29	38 ± 2	104 ± 50	118 ± 60	14	13
	30	38 ± 2	(36 ± 31	47 ± 41	12	13)
First marriage	20	124 ± 2	129 ± 24	135 ± 26	6	5
	21	173	173	184	11	6
	22	199	178	181	3	2
	23	210	202	219	17	8
	24	199 ± 3	174 ± 43	185 ± 49	12	7
	25	174 ± 3	141 ± 42	137 ± 46	-4	-3
	26	139	87	93	6	7
	27	115	112	105	-6	-6
	28	98	122	151	29	23
	29	86 ± 3	92 ± 47	118 ± 60	26	28
	30	70 ± 3	(57 ± 39	66 ± 49	9	16)

1) Estimated 95 percent confidence limits.

2) Fewer than ten occurrences in the target sample.

Table 13. Rates of transition out of State 1 of Figure 1 (first-married, parity 0), at durations of 0 to 4 years combined, by age at marriage, based on data for the entire population (P), the entire target sample (S), and the respondents (R) in the 1981 Swedish fertility survey. Per 1000 women per year.

Cause of decrement	Cohort, women born in	Age at entry	Rates			Estimated non-response bias	
			$\hat{\mu}_P$	$\hat{\mu}_S$	$\hat{\mu}_R$	per 1000 $b(\hat{\mu}_R)$	per cent $rb(\hat{\mu}_R)$
First birth	1936-40	15-19	779	871	912 ±250[1)]	41	5
		20-24	422	451	456 ± 66	5	1
		25-29	346	317	341 ± 86	24	7
	1941-45	15-19	909	1091	1115 ±211	24	2
		20-24	423	439	462 ± 46	23	5
		25-29	374	439	451 ± 90	12	3
	1946-50	15-19	831	868	989 ±208	121	14
		20-24	428	420	428 ± 51	8	2
	1951-55	15-19	648	587	565 ±187	-22	-4
Dissolution of marriage	1936-40	20-24	6	-	-	-	-
	1941-45	20-24	9	(8	10 ± 7	1	16)[2]
	946-50	20-24	16	14	13 ± 9	-1	-8

1) Estimated 95 percent confidence limits.

2) Fewer than ten occurrences.

Table 14. Second-birth rates at durations of 0 to 4 years combined since first birth, by age at first birth, based on data for the entire target sample (S) and the respondents (R) in the 1981 Swedish fertility survey.

Per 1000 women per year.

Cohort, women born in	Age at first birth	Rates		Estimated non-response bias	
		$\hat{\mu}_S$	$\hat{\mu}_R$	per 1000 $b(\hat{\mu}_R)$	per cent $rb(\hat{\mu}_R)$
1936-40	15-19	188	182 ± 51 [1]	-6	-3
	20-24	223	231 ± 37	9	4
	25-29	198	199 ± 40	1	1
1941-45	15-19	216	225 ± 41	9	4
	20-24	195	196 ± 22	1	1
	25-29	172	175 ± 28	3	1
1946-50	15-19	169	167 ± 29	-2	-1
	20-24	203	208 ± 25	6	3
	25-29 [2]	176	191 ± 30	15	9
1951-55	15-19	164	169 ± 35	5	3
	20-24 [2]	193	200 ± 26	7	3
1956-60	15-19 [2]	156	164 ± 51	8	5

1), 2) See Table 15.

Table 15. Rates of transition out of State 2 of Figure 1 (never-married, parity 1), at durations of 0 to 4 years combined, based on data for the entire population (P), the entire target sample (S), and the respondents (R) in the 1981 Swedish fer-tility survey. Women who gave birth to their first child at ages 20 to 24. Per 1000 women per year.

Cause of decrement	Cohort, women born in	Rates			Estimated non-response bias	
		$\hat{\mu}_P$	$\hat{\mu}_S$	$\hat{\mu}_R$	per 1000 $b(\hat{\mu}_R)$	per cent $rb(\hat{\mu}_R)$
Second birth	1936-40	92	85	93 ± 57 [1]	8	9
	1941-45	65	80	86 ± 38	6	8
	1946-50	90	110	109 ± 31	-1	-1
	1951-55 [2]	70	114	120 ± 30	6	5
First marriage	1936-40	248	246	231 ± 91	-15	-6
	1941-45	196	209	224 ± 61	15	7
	1946-50	137	119	114 ± 32	-5	-4
	1951-55 [2]	150	136	133 ± 31	-3	-2

1) Estimated 95 percent confidence limits.

2) Censored at the end of 1980.

4.3. The effect of nonresponse and stochastic errors on analyses and conclusions

In life history analysis based on estimated intensities we look at intensity curves and compare them for different birth cohorts or periods. Single extreme values are disregarded. Stochastic errors make our curves more irregular and our tests less powerful, i.e., they decrease our chances of discovering genuine differences (at any given significance level). Systematic errors, like nonresponse bias, change the levels of our curves and could give rise ton wrong conclusions. If the nonresponse bias has the same direction and is of about the same size in curves which are compared, however, the nonresponse bias need not be so harmful.

As shown in the previous section, the nonresponse bias is positive in general for the estimated transition intensities in the models of Figures 1 and 2. (Exceptions are some intensities for the youngest ages and for intensities of transition out of State 2 (never-married, parity 1) caused by marriage.) Furthermore, the nonresponse bias seldom exceeds 5 to 10 percent of the estimated intensity. Therefore, we can expect the nonresponse bias not to disturb our analyses too much. This is also verified when we conduct replicated analysis of the model of Figure 1 based on data for the entire population, the entire target sample , and the respondents. In fact, our findings are the same, but of course less sharp for data based on the target sample and the respondents than for results based on the entire population.

Sweden has had decreasing fertility and nuptiality during the two most recent decades. Quist (1983) has analyzed the interaction between these phenomena by means of a model similar to that of Figure 1. He found that the decreasing nuptiality is an intermediate factor leading to the decreasing fertility. The birth intensities for married women have been almost constant over birth cohorts. For never-married women the birth intensities have increased, but not enough to compensate for the decreased nuptiality (see Table 11 above). From survey data, Hoem and Rennermalm (1983) found that the first-birth intensities have been almost constant for never-married single women and for never-married cohabiting women, taken separately.)

Figure 7 below shows the estimated probabilities of belonging to each of the states of Figure 1 at age 25 for two of the birth cohorts in the survey: women born in 1941-45 (C2) and in 1951-55 (C4). (Figure 16 at the end of this section shows the same results for all our cohorts in the survey and for ages 20 and 30 as well.) The figures outside the boxes show the estimated probabilities of leaving State 0 before age 25, by cause. Note that States 3 to 6 are absorbing. The figures in the boxes for those states therefore refer to the probabilities of entering the states before age 25.

Comparing the cohorts 1941-45 (C2) and 1951-55 (C4) we see that 49 (39+10) per cent in the former group and 27 (17+10) per cent in the latter have married and given birth to a child at age 25. In both cohorts 10 percentage points have given birth to the first child before marriage. Of course, this does not mean that nonmarital fertility has been stable while the marital fertility has decreased, as we can easily see by analysing the figures in boxes 1 and 2 and the figures outside the boxes.

The results of Figures 7 and 16 are based on data for the entire population. Fortunately, we would attain almost the same results from data based on the entire target sample or on the respondents only (see Table 16 below and Table 19 at the end of this section.)

The estimated ~~probabilities~~ describe the effect of changing marriage and fertility behavior rather than the changes in themselves.

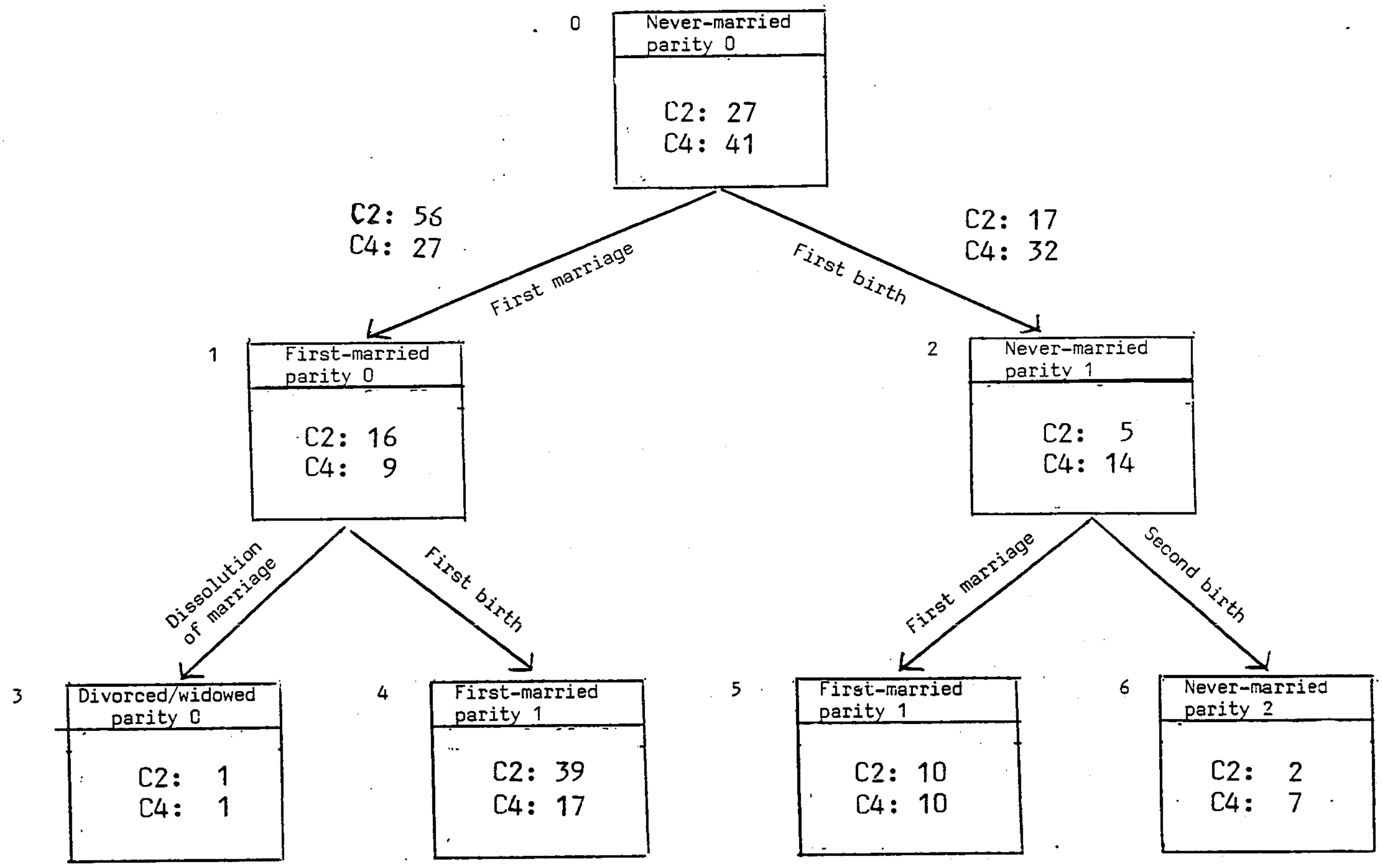


Figure 7. Probabilities of belonging to each of the states of Figure 1 at age 25 (within boxes), and of leaving State 0 before the same age, by cause (outside boxes). Estimates based on data for the entire register population. Women born in 1941-45 (C2) and in 1951-55 (C4), respectively. Percent.

Table 16. Probabilities of belonging to each of the states of Figure 1 at age 25. Estimates based on data for the entire population (P), on the entire target sample (S), and on the respondents (R) in the 1981 Swedish fertility survey.

Women born in 1941-45 (C2) and in 1951-55 (C4). Percent.

State	Women born in					
	1941-45			1951-55		
	P	S	R	P	S	R
0: Never-married, parity 0	27	28	27	41	44	42
1: First-married, parity 0	16	15	14	9	9	9
2: Never-married, parity 1	5	5	5	14	12	12
3: Divorced/widowed, parity 0	1	0	0	1	1	1
4: First-married, parity 1	39	40	41	17	17	17
5: First-married, parity 1	10	10	10	10	9	9
6: Never-married, parity 2	3	3	3	7	8	9

A quicker and surer way to describe and establish any changes in the behavior is to analyze the estimated transition intensities directly. Figure 8 below shows the estimated intensities of transition from State 0 of Figure 1 (never-married, parity 0), caused by first birth and marriage, respectively.

The curves based on data for the entire population show that the first-birth intensities have increased over the cohorts, while the marriage intensities have decreased strongly. The curves are smooth and we hardly need use any statistical test to confirm these conclusions. The curves based on data for the target sample and for the respondents are strongly influenced by stochastic errors, and we may want to test the significance of the differences between them. The results of such tests are shown in Table 17 below. To illustrate these results we show the rates of transition out of State 0 for combined five-year age intervals (from Table 11 above), although we know that the underlying intensities are far from constant in such large intervals.

As expected, almost every difference, no matter how small, is found to be significant at the 5 percent level ($|\tilde{U}| > 1.96$) for estimates based on data for the entire population. When these differences are substantial (see the population curves in Figure 8 and 9), we also often found them to be significant (at the 5 percent significance level) for data based on the target sample or on the respondents. By means of data for the respondents (and target sample) we can establish that the first-birth intensities for never-married women increase from cohort C2 to C3 at the ages 15 to 24, but the increase at the ages 25 to 29 is not found to be significant. Two other substantial increases of nonmarital fertility are found not to be significant in the same way: the increase from C1 to C2 at ages 25 to 29 and the increase from C3 to C4 at ages 20 to 24. In the first case the increase is found to be significant by

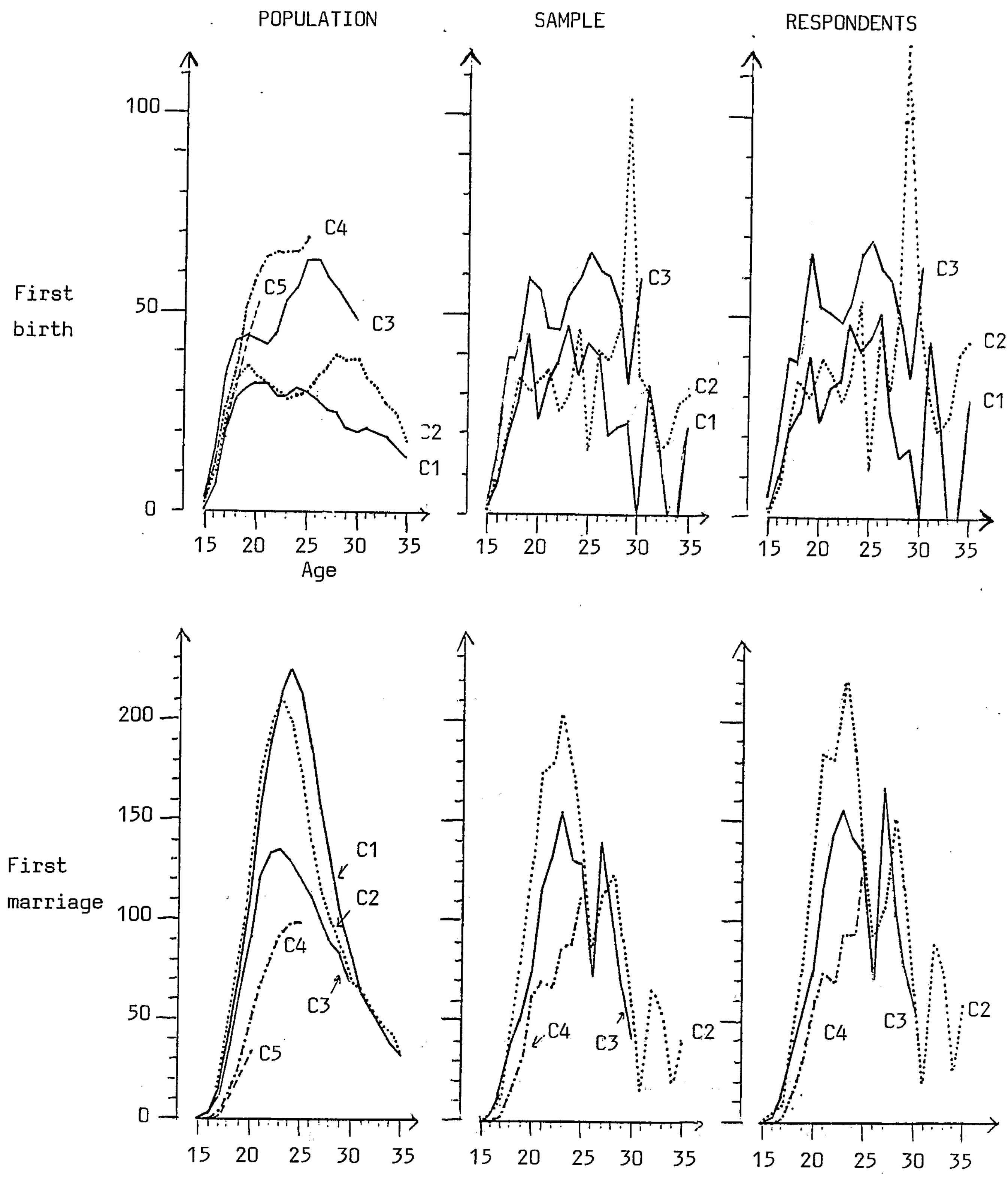


Figure 8. Cohort comparisons of intensities of first birth and first marriage, respectively, among never-married nulliparas. Calculations based on data for the entire population, the entire target sample, and the respondents in the 1981 Swedish fertility survey, by birth cohort 1936-40 (C1), 1941-45 (C2), 1946-50 (C3), 1951-55 (C4), and 1956-60 (C5).

Per 1000 women per year.

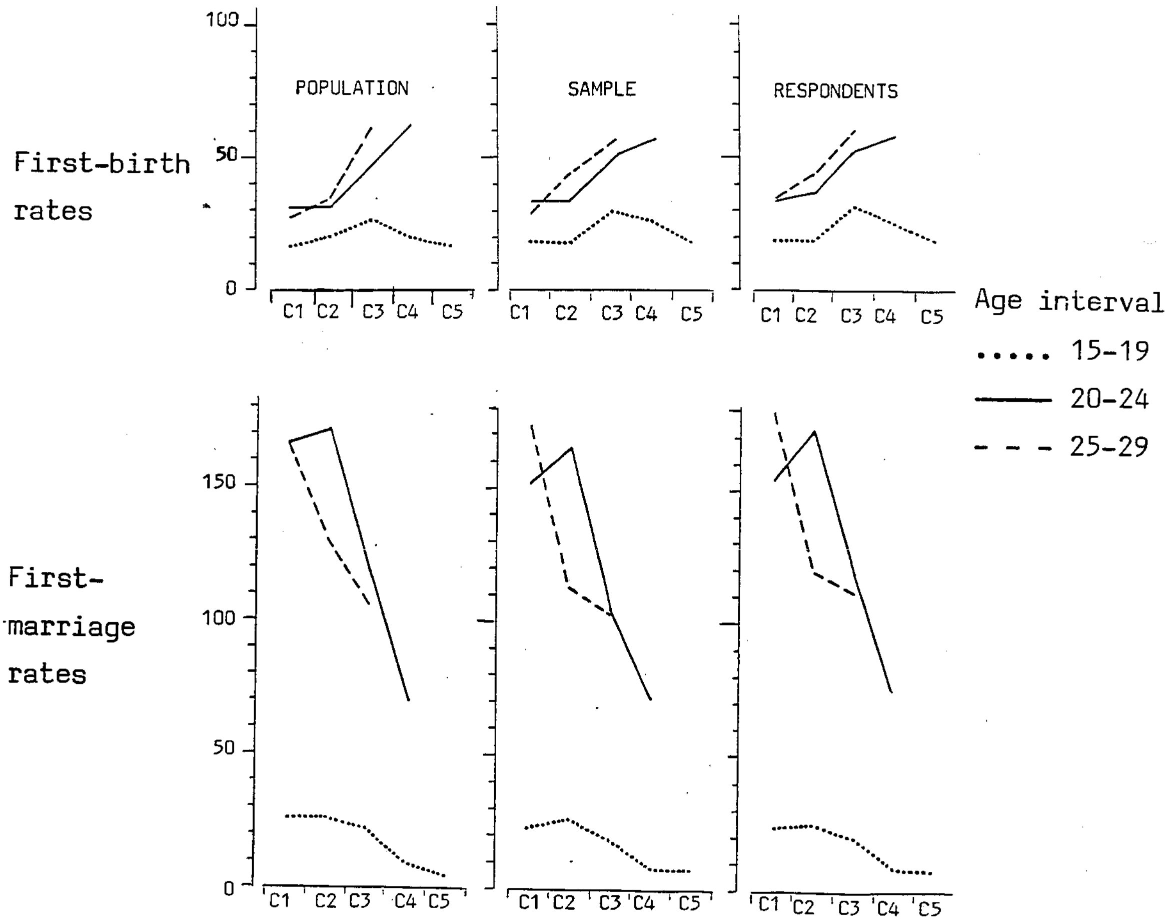


Figure 9. Rates of transition out of State 0 (never-married, parity 0) for combined five-year age intervals, by cause and birth cohort. Calculations based on data for the entire population, the entire target sample, and the respondents in the 1981 Swedish fertility survey. Per 1000 women per year.

Table 17. Cohort comparisons of intensities of first birth and first marriage among never-married, nulliparous women (State 0 of Figure 1). Test statistic $\tilde{U}$ in (3.8) combined for sequences of five years of age, based on data for the entire population (P), the entire target sample (S), and the respondents (R) in the 1981 Swedish fertility survey.

Age	Data	First birth				First marriage			
		C2-C1 1)	C3-C2	C4-C3	C5-C4	C2-C1	C3-C2	C4-C3	C5-C4
15-19	P	19.8	35.7	-11.4	-26.2	0.5	-18.1	-74.8	-16.5
	S	-0.4	4.1	-2.1	-1.8	0.8	-2.0	-5.3	-0.7
	R	-0.3	4.4	-2.8	-1.2	0.5	-1.2	-5.1	-0.7
20-24	P	0.9	48.4	41.1		7.0	-86.5	-87.3	
	S	-0.1	3.4	0.7		1.1	-5.3	-5.9	
	R	0.4	2.8	0.7		1.5	-5.1	-5.4	
25-29	P	15.7	39.6			-28.3	-23.5		
	S	2.0	1.4			-2.6	-0.9		
	R	1.2	1.6			-2.1	-0.7		
30-34	P	16.3				-1.0			
	S	2.3				-1.1			
	R	2.2				-0.7			

1) Test statistic for the difference between cohort of women born in 1941-45 (C2) and cohort of women born in 1936-40 (C1). See notations of cohorts in Figure 8.

means of data for the entire sample but not for the respondents. Nevertheless, we can establish the decrease of unmarital fertility and nuptiality among the teenagers from cohort 1946-50 (C3) to cohort 1951-55 (C4) (i.e., in the late 1960s), which may be explained by the legalization of the pill in 1965.

As regards the marriage intensities among women of parity 0, the data for the respondents (and for the target sample) also show the right picture, but two substantial decreases are not found to be significant: the decrease from cohort 1941-45 (C2) to cohort 1946-50 (C3) at ages 15 to 19 and at ages 25 to 29. In the former case we would have been able to establish the decrease if there had been no nonresponse.

By means of data for our survey respondents (and target sample) we obtain the same conclusions as Quist (1983) did in his investigation based on the entire fertility register. Among never-married nulliparas born in 1936-60, the marriage intensities have decreased over the birth cohorts, while the birth intensities have increased with one exception: at ages 15 to 19, the birth intensities began to decrease from the cohort 1946-50 to the cohort 1951-55. Of course, the results based on data for the respondents (and for the target sample) are less sharp than those based on the entire register population. In no case, however, would sampling errors or nonresponse bias have induced us to assert any false differences and changes, provided that we test our results properly.

Let us now study what happens when the women leave the maiden status due to marriage or childbirth, i.e., when they have entered State 1 or State 2 of Figure 1. Unfortunately, we can only study the few first years in these states, since the number of observations in the sample decreases rapidly (see Table 2). Furthermore, our data are subject to censoring at the end of 1977 and 1980 for the population and for the sample, respectively (see Diagram 1 above).

Figure 10 shows the birth intensities during the <u>first year</u> of marriage, by age at marriage, for each of our five five-year birth cohorts. We might expect the first-birth fertility among first-married women to increase over the cohorts as an effect of the decreasing first-marriage intensities, i.e., we might expect that among the younger cohorts only those women marry who plan childbirth in the near future. This is not the case, however, In fact, among the teenage wives the first-birth intensities

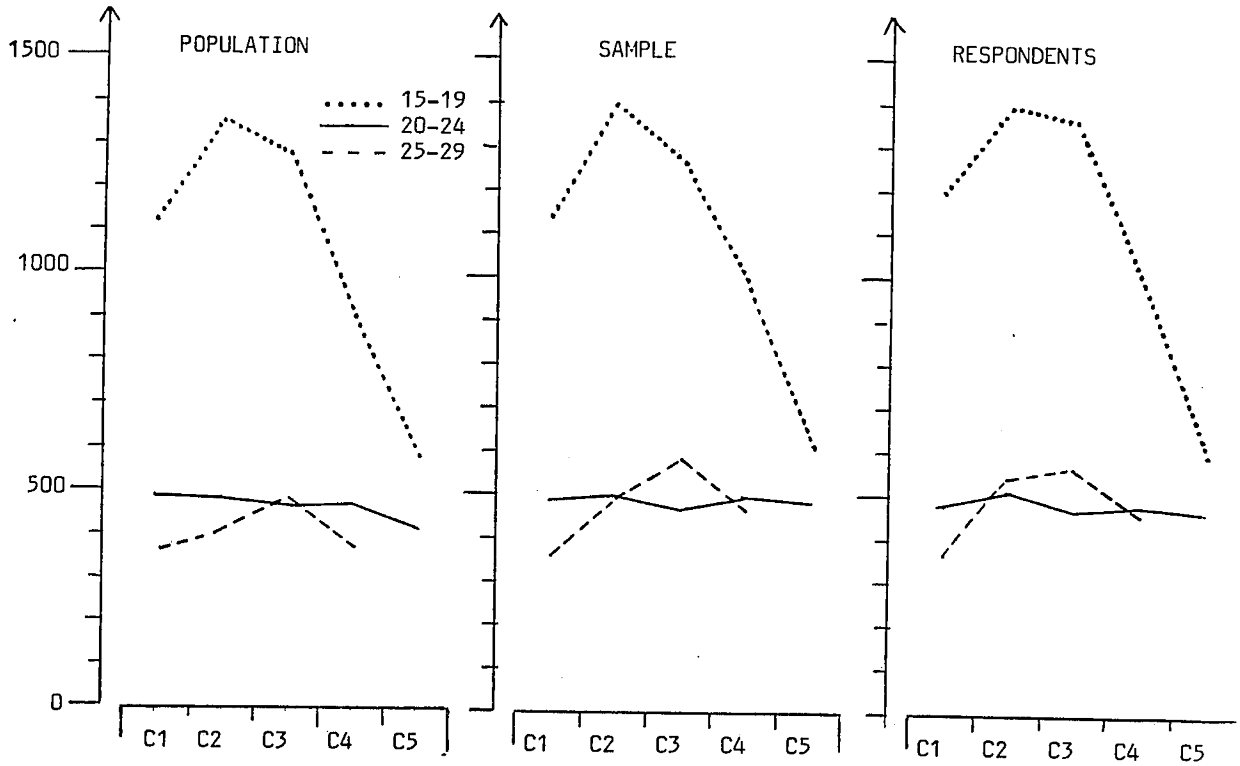


Figure 10. First-birth intensities during the first year of first marriage among women who married at ages 15-19, 20-24, and 25-29, respectively, by birth cohort. Estimates based on data for the entire population, on the entire target sample, and on the respondents in the 1981 Swedish fertility survey. Per 1000 women per year.

during the first year of marriage have decreased strongly from cohort 1946-50 (C3) over cohort 1951-55 (C4) to cohort 1956-60 (C5).

Thus at the ages 15 to 19 we have a decrease over the cohorts in nuptiality as well as marital and nonmarital fertility. By means of data for the respondents (and target sample) we can establish that these decreases started for women born in 1951 to 1955 (C4) or earlier. By estimating the intensities for single one-year birth cohorts, which is possible for the entire register population, we can see that the decreases have been uninterrupted (with few exceptions) from the cohort of 1946 to the one of 1958 (Figure 11). Such detailed analysis, however, is not possible from data for the sample only.

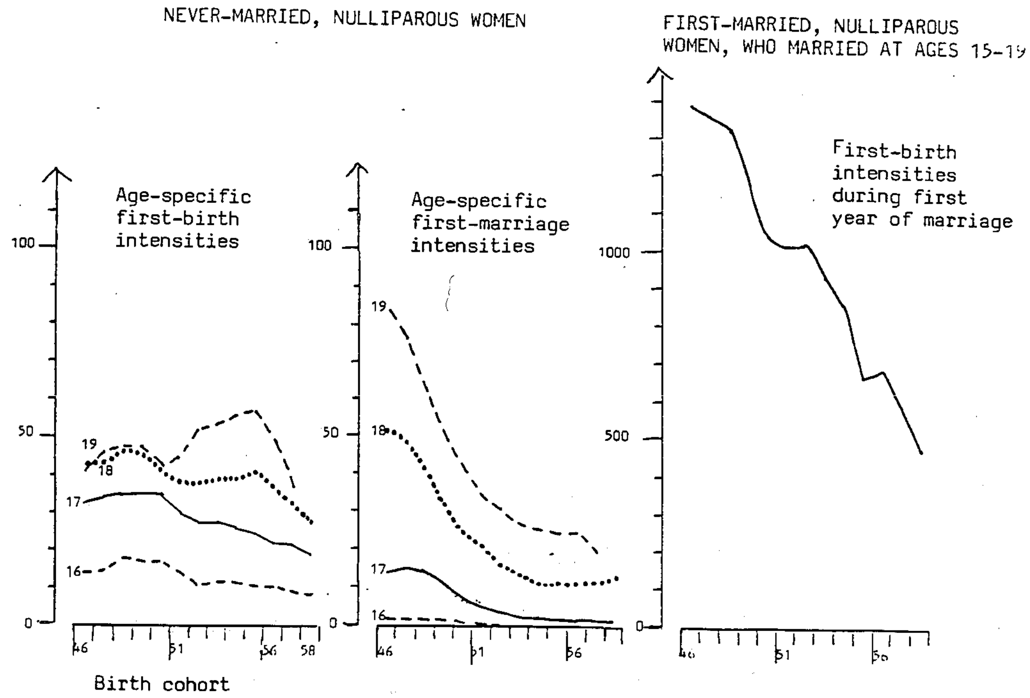


Figure 11. Age-specific first-birth and first-marriage intensities at ages 16-19 among never-married, nulliparous women, and first-birth intensities during the first year of marriage among women who married at ages 15-19, by single-year birth cohort. Estimates based on data for the entire register population. Per 1000 women per year. Note that the ordinate unit in the rightmost diagram is ten times those in the left diagrams.

Thus from data for the target sample (and for the respondents) we can get at least a rough picture of the decreasing propensity of teenagers to raise a family, which can be analyzed much more sharply from data for the entire register population.

The marriage rates among teenagers have been small for all cohorts studied. (In Sweden a women cannot marry without special permission before the age of eighteen.) Therefore, the large decline in fertility and nuptiality in Sweden during the two latest decades can be explained only to a small degree by a new pattern among the youngest women.

For first-married nulliparas who married at ages 20 to 24, the first-birth intensity during the first year of marriage has been stable over the cohorts (Figure 10). Neither sampling errors nor nonresponse biases distort that conclusion. (As regards women who married at ages 25 to 29 we can draw no conclusions, since our data are subject to censoring that occurs at different times for the population and for the sample. See Diagram 1). According to the population estimates of Figure 12 below, the first-birth intensities are stable over the cohorts (as far as we can study it) at a duration of 0, 1, 2, 3 years of marriage, respectively, among women who married at ages 20 to 24.

From estimates based on the sample and on the respondents we arrive at the same conclusion for the first two years of marriage. At higher durations the stochastic errors become very large. From cohort 1946-50 (C3) to cohort 1951-55 (C4) there seems to be an increase at a duration of two years but a decrease at a duration of three years. The increase at two years is not significant: the test statistic U in (3.7) is 0.7 and 1.5 for the sample and for the respondents, respectively. The decrease at three years, however, is found to be significant: the test statistic U is -2.1 and -2.4 for the sample and for the respondents, respectively. This is an example of a false significance, which we expect to get in five percent of the tests.

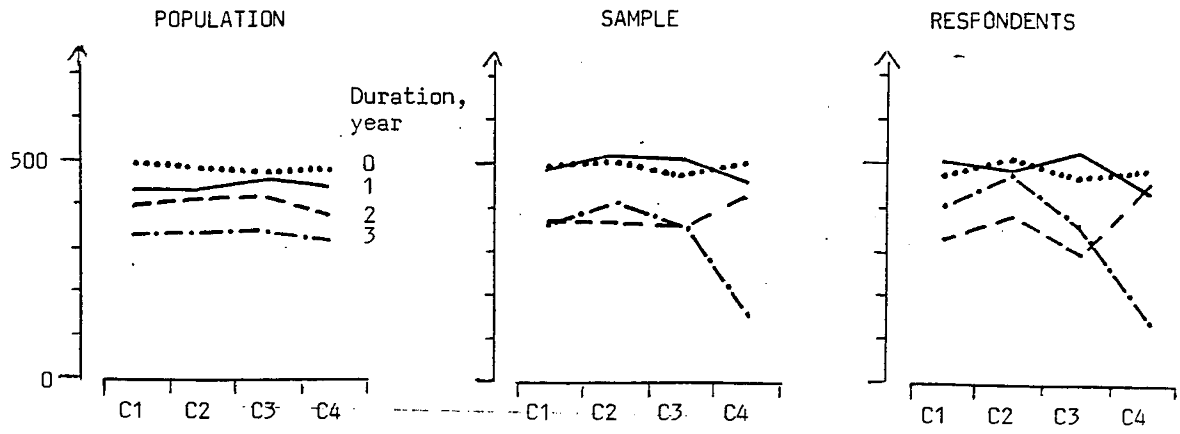


Figure 12. Duration -specific first-birth intensities among first-married women who married at ages 20 to 24, for durations of 0 to 3 years of marriage, by birth cohort. Estimates based on data for the entire population, for the entire sample, and for the respondents in the 1981 Swedish fertility survey. Per 1000 women per year.

Our small sample size hardly permits us to analyze the intensities of marriage dissolution, which are very small during the first years of marriage. It is only at four to five years of marriage that divorce (rarely widowhood for the ages studied) becomes a tangible "competitor" to childbirth among first-married, nulliparous women (see Figure 13).

One should note that the results in Figure 13 (and in Figures 10 and 12) are based on censored data, and that the censoring appears earlier for the population data than for the sample data. If our model is true, i.e., if within each five-year birth cohort and age at entry

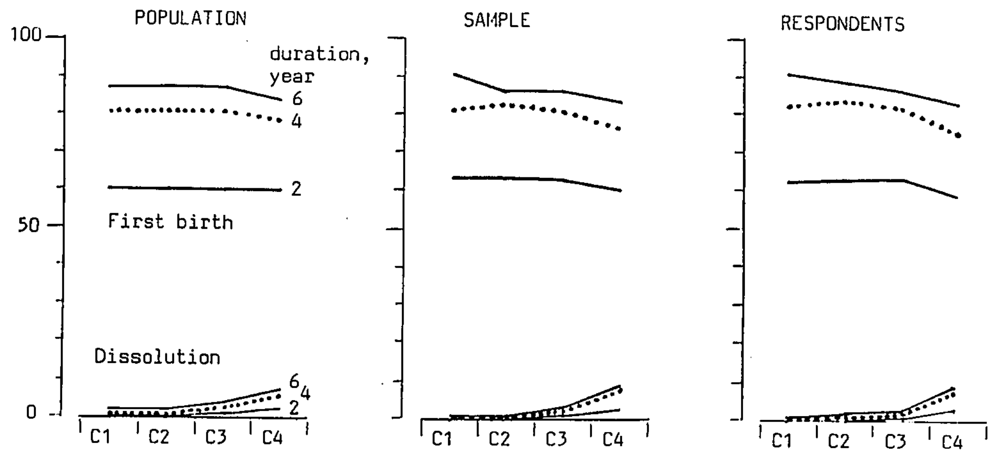


Figure 13. Probabilities of leaving State 1 of Figure 1 (first-married, parity 0) within 2, 4, and 6 years of marriage, respectively, by cause and by birth cohort, among women who married at ages 20 to 24. Estimates based on the entire population, on the entire target sample, and on the respondents in the 1981 Swedish fertility survey. Percent.

group the women are homogeneous regarding the phenomena studied, then the censoring has no effect on the estimates. Our model is not quite realistic, however. We should expect changes to appear gradually over ages (at entry) and over birth cohorts rather than stepwise over five-year intervals. Due to censoring the estimates for the later five-year birth cohorts and longer durations are based on data for earlier birth cohorts and younger ages (at entry) for the population than for the sample (and for the respondents) (see Diagram 1). Surely, this censoring is selective and affects the estimates. We have not, however, investigated this problem further here.

Finally we study the never-married women who give birth to a child, i.e., who enter State 2 of Figure 1. According to the upper diagrams of Figure 8 and 9, the first-birth fertility among never-married women has increased over the birth cohorts at age 20 and upwards. Hoem and Rennermalm (1982) have shown that this is explained by an increasing exposure to the risk of conception in consensual unions. In fact, among never-married women, the first-birth fertility has been remarkably stable taking single women and cohabiting women separately. Single women have had a very low fertility in all cohorts studied: the first-birth rates are around 20 to 30 per 1000 women per year at ages around 20. Among cohabiting women the first-birth rates are 100 to 200 per 1000 women per year at durations 7 to 60 months in the consensual union combined. (For married women, the rates are higher at the younger ages.)

A large proportion of the respondents in the survey who reported a first birth as never-married were in fact cohabiting at the time of the childbirth, almost half for the three earliest cohorts (C1, C2, and C3) and a good eighty percent for the two latest cohorts (C4 and C5). This means that "never-married women of parity 1" is a mixed group of single and cohabiting women in which the proportion of the latter group increases over the birth cohorts. Changes in fertility and nuptiality among "never-married women of parity 1" must therefore be interpreted as an effect of changes in earlier phases of the women's adult family life.

Figure 14 shows the marriage rates among never-married women of parity 1 during the <u>first</u> year since first birth, by age at that birth. The marriage rates have decreased strongly for never-married women of parity 1, aged 15-19 or 20-24 at first birth, during the first year since that birth. Data for the entire population, for the entire target sample and for the respondents all show the same pattern.

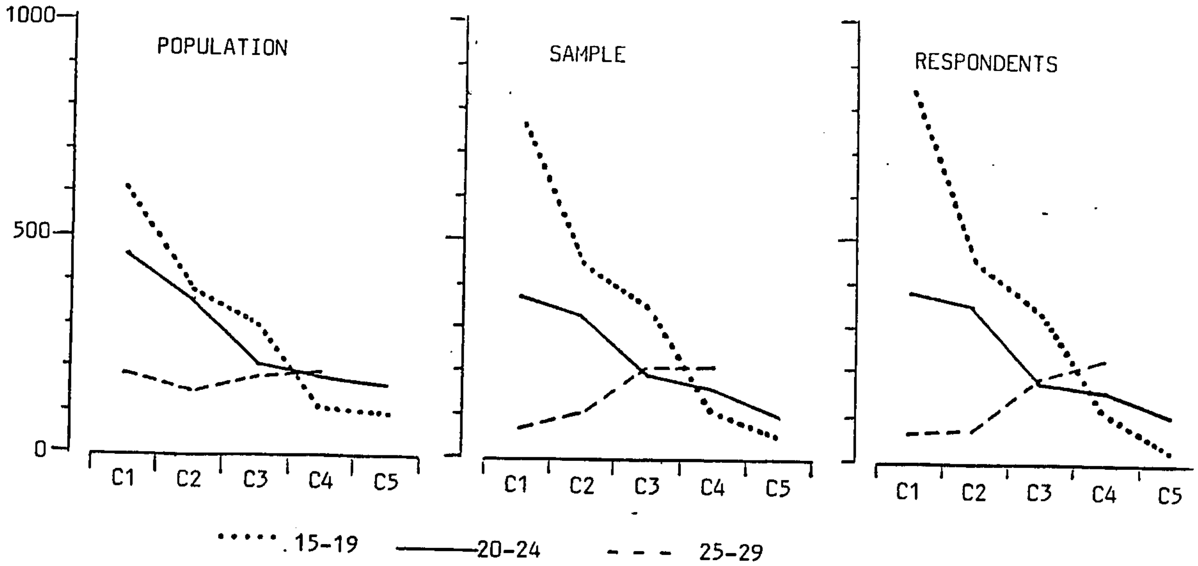


Figure 14 . First-marriages rates for never-married women of parity 1 during the first year since first birth, by age at that birth. Calculations based on data for the entire population, for the entire target sample, and for the respondents in the 1982 Swedish fertility survey.
Per 1000 women per year.

Due to sampling errors, the differences between women aged 15 to 19 and women aged 20 to 24 at first birth are exaggerated for the earlier cohorts C1, C2, and C3. Calculated (model based) 95 percent confidence intervals for these differences show that they are "significant" for cohorts C1 and C3 only, according to data for the sample and the respondents. See Table 18.

Our model assumption of a homogeneous population group is clearly not fulfilled regarding the mixed group of "never-married women of parity 1". Nevertheless, the calculated confidence intervals, as well as the test statistic U, show the size of the stochastic errors, and may help us to judge how much we can believe in our findings.

The rates for the sample and the respondents shown in Figure 14 are based on very small numbers of women and exposures (see Table 2).

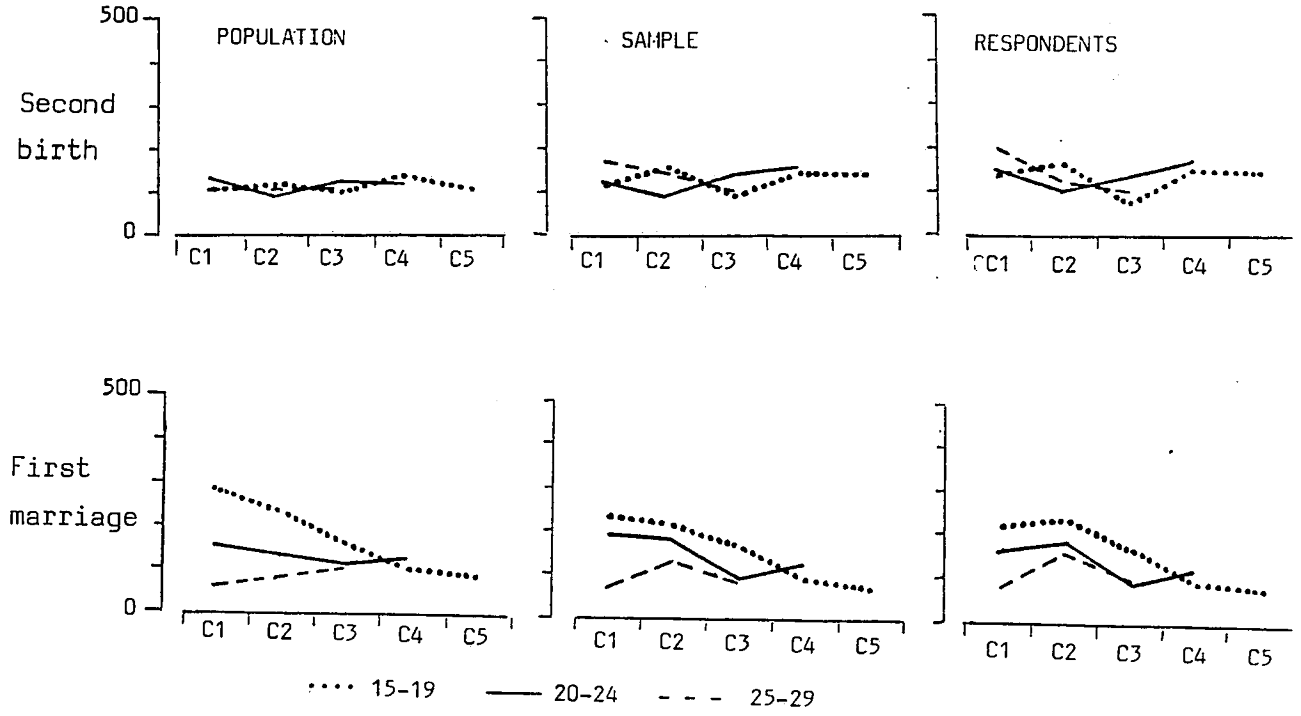
Table 18. Differences in marriage intensities for never.married women of parity 1 during the first year since first birth between women aged 15 to 19 and 20 to 24, respectively, at first birth. Calculated model based 95 percent confidence interval estimates based on data for the entire population (P), for the entire target sample (S), and for the respondents (R) in the 1981 Swedish fertility survey.
Per 1000 women per year.

Birth cohort	P	S	R
C1: 1936-40	154 ± 18 *	402 ± 340 *	473 ± 401 *
C2: 1941-45	34 ± 6 *	152 ± 188	148 ± 202 *
C3: 1946-50	88 ± 8 *	156 ± 121 *	165 ± 126 *
C4: 1951-55	-71 ± 6 *	-50 ± 80 *	-48 ± 86
C5: 1956-60	-70 ± 17 *	-65 ± 80 *	-74 ± 83

* The interval does not include 0.

In fact, most of the rates are based on exposure of less than 100 woman-years. For the cohort born in 1936-40 (C1) the exposures for the respondents are only 27, 37, and 24 woman-years for age groups 15-19, 20-24, and 25-29, respectively, and slightly more for the target sample. Knowing this, the extent to which we get the right picture of the trends is surprising. Furthermore, the nonresponse bias does not distort that picture.

Figure 15 shows the rates of transition out of State 2 of Figure 1 (never-married, parity 1) at durations 1 to 3 years since first birth combined. The marriage rates are about half of those for the first year since first birth, and they are slightly decreasing over the birth cohorts for the younger mothers. Once again, we would draw the same conclusions from results based on the sample and the respondents as we would from results based on the entire population.



<u>Figure 15</u>. Rates of transition out of State 2 of Figure 1 (never-married, parity 1), at durations 1 to 3 years in the state combined, by cause and by age at first childbirth. Calculations based on data for the entire population, for the entire target sample, and for the respondents in the 1981 Swedish fertility survey. Per 1000 women per year.

To sum up our result, we have analyzed the transitions depicted in the model of Figure 1 using data for the entire population, for the entire target sample, and for the respondents in the 1981 Swedish fertility survey. The main results are always the same for the three data sets used: at younger ages (15-19) the propensity to form a family has always been small and has become even smaller from cohort 1946-50 onwards. At greater ages (20-29) the marriage intensities among never-married nulliparous women have decreased substantially while the first-birth intensities have increased slightly.

Among first-married women of parity 0 the first-birth intensities have been stable during the first three years of marriage (except for the youngest ages). Never-married women of parity 1 show decreasing nuptiality over the cohorts, at least during the first year since the childbirth. The second-birth fertility among those women, however, seems to be stable during the first three years since childbirth.

Very often the stochastic errors for estimates based on the sample (and on the respondents) are large and may result in false peaks or trends or in exaggeration of real trends. If we are careful and disregard single extreme values and compute the standard errors, however, we would hardly draw seriously wrong conclusions on the basis of the sample estimates. Furthermore, we have found no cases where the nonresponse bias change any of the main results.

The effect of changed behavior regarding fertility and nuptiality is a change (over the birth cohorts) of the probabilities of belonging to the various states of Figure 1 at given ages. Estimates of such probabilities are shown in Figure 16 below. These results are based on data for the entire register population, but results based on data for the sample or for the respondents are almost exactly the same (Table 19). Note that we cannot follow first-married nulliparas up to the age of 30 due to lack of observations in the sample. Due to nonresponse, the estimated probability of remaining a never-married nullipara at age 25 and upwards is <u>underestimated</u> by one to two percentage points.

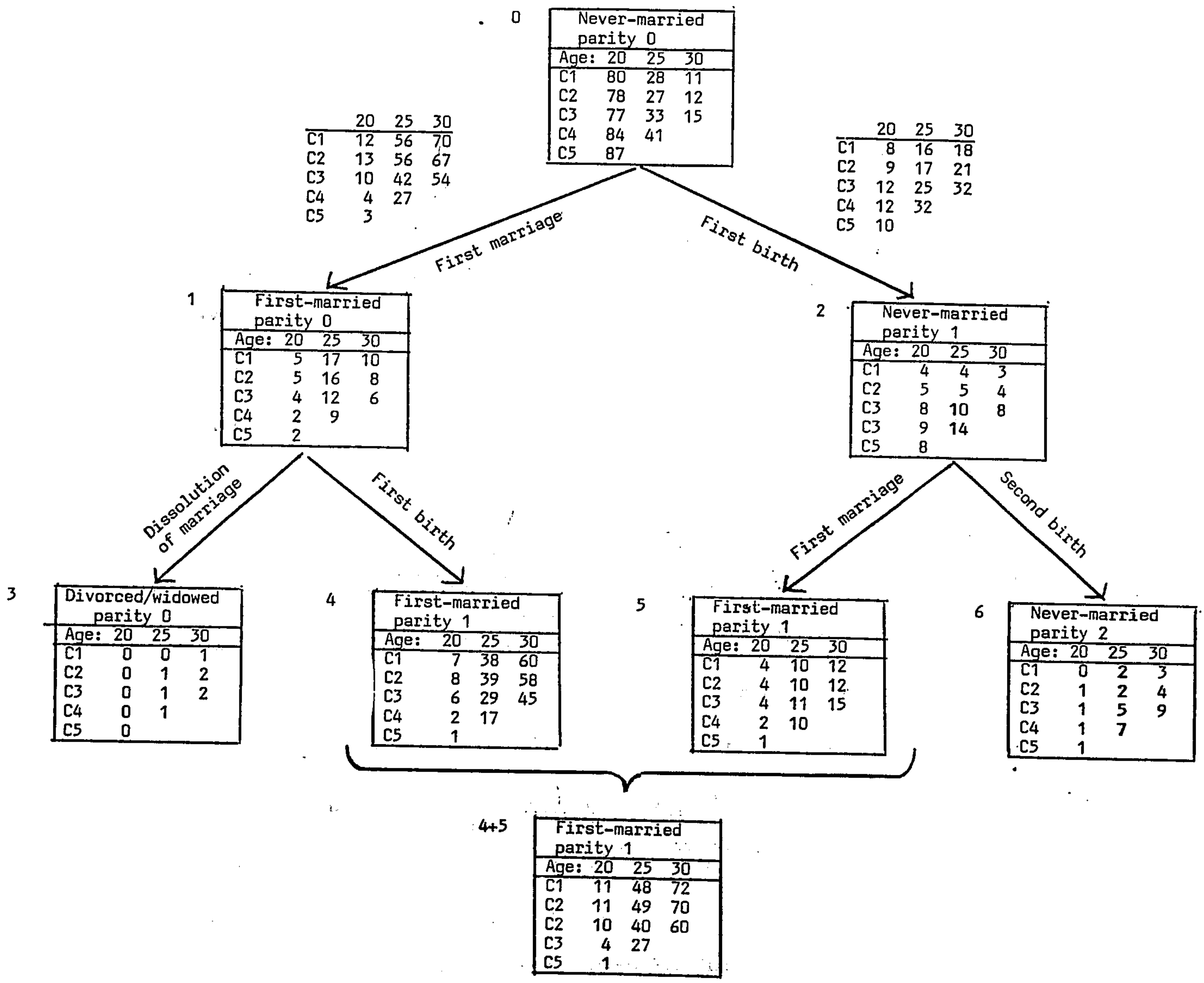


Figure 16. Probabilities of belonging to each of the states of Figure 1 at ages 20, 25, and 30 (within boxes) and of leaving State 0, (never-married, parity 0) before the same ages, by cause (outside boxes). Estimates based on data for the entire Swedish fertility register, by birth cohort: 1936-40 (C1) 1941-45 (C2), 1946-50 (C3), 1951-55 (C4), and 1956-60 (C5). Percent.

Table 19. Probabilities of belonging to the various states of Figure 1 at ages 20, 25, and 30. Estimates based on data for the entire population (P), the entire target sample (S), and the respondents (R) in the 1981 Swedish fertility survey. Percent.

Cohort, women born in	Age		Never-márried, parity 0	First marriage entered as nullipara (State 1) Divorced/ widowed, parity 0	First-married, parity 0	First-married, parity 1	First child born as unmarried (State 2) First-married, parity 1	Never-married, parity 1	Never-married, parity 2
			0	3	1	4	5	2	6
C1:1936-40	20	P	80	-	5	7	4	4	0
		S	80	-	5	6	5	4	0
		R	80	-	5	7	5	4	0
	25	P	28	0	17	38	10	4	2
		S	30	-	15	37	11	5	2
		R	30	-	15	37	11	5	2
	30	P	11	1	10	60	12	3	3
		S	12	1	9	58	13	4	4
		R	11	.. 1)	..	..	13	3	4
C2:1941-45	20	P	78	0	5	8	4	5	1
		S	79	-	4	8	3	5	0
		R	80	-	4	7	4	5	0
	25	P	27	1	16	39	10	5	3
		S	28	0	15	40	10	5	3
		R	27	0	14	41	10	5	3
	30	P	12	2	8	58	12	4	4
		S	13	..	..	..	12	5	5
		R	11	..	..	..	13	5	5
C3:1946-50	20	P	77	0	4	6	4	8	1
		S	77	0	3	6	3	9	1
		R	76	-	3	6	4	9	1
	25	P	33	1	12	29	11	10	5
		S	32	0	12	27	13	10	6
		R	31	0	12	28	13	10	5
	30	P	15	2	6	45	15	8	9
		S	15	2	6	43	17	7	10
		R	13	2	6	44	18	7	10
C4:1951-55	20	P	84	0	2	2	2	9	1
		S	84	0	2	2	2	10	1
		R	84	0	2	2	2	9	1
	25	P	41	1	9	17	10	14	7
		S	44	1	9	17	9	12	8
		R	42	1	9	17	9	12	9
C5:1956-60	20	P	87	0	2	1	1	8	1
		S	88	0	2	1	1	7	1
		R	87	0	2	2	1	7	1

1) The group of women entering State 1 at ages 15-19 is empty before a duration of 15 years.

5. CONCLUSIONS

Life table techniques are useful for analyzing life histories. This is true even when the number of individuals under study is as small as in our survey (500 to 1000 respondents in the separate cohorts). Due to random variation the estimated intensity curves often become very irregular. However, a genuine difference between curves will be detected in most cases, at least if the difference is substantial, and only genuine differences are likely to become firmly established. Of course, the small sample size makes our analysis less sharp than an analysis based on the entire population. For example, it is not possible to follow the women to higher ages or longer durations in a state in the sample data.

In order to base our estimates on reasonable numbers of observations we have grouped the individuals. First, they have been grouped naturally into five-year birth cohorts by our sampling plan. Secondly, we have used five-year age intervals to control for starting age in the intermediate states "first-married, parity 0" (State 1) and "never-married, parity 1" (State 2). As a consequence of this aggregation, we cannot find out whether there is any periodic effect or for which (if any) birth cohort a new pattern of behavior started, and at what age. For example, it has not been possible to study any sharp effects of new legislation: the legalization of the pill (1964), the more liberal laws for marriage/divorce (1974) and abortions (1975), and so on, nor have we been able to compare women who marry (or give birth to a child) before the lowest legal marriage age (18) with those who start at later ages. Another consequence of the groupings is that our model assumption, that the life histories observed are outcomes of the same stochastic process, becomes less realistic. Among other things this implies that the effect of censoring cannot be ignored.

A primary finding of this investigation is that the existence of nonresponse would not distort substantive analysis. Granted, the estimated transition intensities in general are afflicted by some positive nonresponse bias. However, the fact that the bias has the same sign and about the same size for all intensities reduces its potential harmfulness.

It is hoped that similar results will be valid for variables other than those we have investigated in our survey and in other surveys with similar circumstances. Nevertheless, it is legitimate to ask to what extent our encouraging findings can be generalized beyond this particular survey and the particular variables investigated. In response to this question, we first note that the stochastic errors of various estimates can always be measured an any survey by statistical methods, at least when the sample is selected in a probabilistic and noninformative manner. We have applied a model-based approach when estimating standard errors and testing differences between occurrence/exposure rates. Thus, our inference concerns model parameters, and we have assumed that the life histories observed for a group of women are independent outcomes of the same stochastic process. We have also assumed that the model transition intensities are piecewise constant over integer years of ages and integer years of duration in a state. These assumptions are not quite realistic, because we have grouped birth cohorts and starting ages in a state. We would expect the intensities to change gradually rather than stepwise across birth cohorts and ages. Nevertheless, when we compare calculated standard errors and test statistics based on the sample (and on the respondents) with corresponding results based on the entire population, the differences are as we would expect according to the model. Thus, our model-based approach seems applicable although its assumptions are not quite realistic. It is a particularly simplifying aspect of our sample, however, that it was selected with equal probabilities within each five-year birth cohort. With unequal selection probabilities the model-based approach may be less tenable if the model assumptions cannot be maintained. In such a case, we could either apply a design-based approach aiming at predicting and comparing population rates (see Cassel, 1982 and Hoem, 1983b) or be more careful to have the model assumptions fulfilled, e.g., by restricting analysis to groups which really are homogeneous. In fact, equal probability selection was chosen for the survey in order to avoid problems of this nature.

On the other hand, any nonresponse bias can never be measured completely by means of data for the respondents only. Whether results from our empirical investigation may be valid for other variables and in other surveys depends on whether the pattern of response probabilities is similar to ours. The nonresponse bias can be expressed as a function of the differences in response probabilities for the decrements (from various causes) and for the survivors during the relevant time period, as defined in our theoretical model. According to the model we expect the differences in response probabilities (and thereby the nonresponse bias) to increase with age or with duration in a state. This is also verified in our empirical findings, and it should hold true in all life table analyses where the response behavior depends on whether decrement occurs (and on its cause), but not on when it happens. In these situations, the nonresponse bias for estimates referring to current status at interview is larger in absolute value than the bias of corresponding estimates of intensities referring to earlier periods. Thus, the former can be used to estimate an upper limit of the latter.

It is often possible in practice to estimate the response probabilities for (or the nonresponse bias of estimates of proportions belonging to) various groups defined by current status. Lindström (1983) has carried out an extensive investigation of the nonresponse errors in various surveys of living conditions in Sweden. He found that the nonresponse bias of estimated proportions (referring to present status at interview) very seldom exceeded two percentage points. The largest bias found amounted to five percentage points. For the surveys in his investigation, our fertility survey was among those most subject to nonresponse bias. It is true that the nonresponse rate (13 percent) is comparatively low in the fertility survey, but the nonrespondents constitute a selective group regarding nuptiality and fertility. The bias of the estimated proportion of women with their own children in their households at the time of the interview amounted to between two and three percentage points. Nevertheless, we have found that the life history analyses regarding these factors would not be distorted in our survey. This promises well for the effects of nonresponse on life history analysis in other surveys.

Of course, the nonresponse in our survey may be more selective if we study factors other than just nuptiality and fertility. For example, we may expect the differences in response probabilities between women with and without children to be largest for low-educated women (a group with high fertility), which would imply that the differences in fertility between educational groups would be over-estimated. Similarly, nonresponse may be more problematic in other surveys and for other relevant variables. The existence of nonresponse can never be completely ignored. Even so, our results strongly suggest that nonresponse effects may be systematically smaller in life history analysis than they are in more conventional analyses or descriptions of circumstances at the time of the interview.

ACKNOWLEDGEMENTS

I am very grateful to Jan M. Hoem for his great interest, painstaking advice and useful suggestions. I also want to thank Lars Nordin for excellent programming.

REFERENCES

Aalen, O.O. and J.M. Hoem (1978). Random time changes for multivariate counting processes. Scand. Actuarial J. 1978(2), 81-101.

Andersson, H. (1979). On nonresponse and response probabilities. Scand. J. Statist. 6, 107-112.

Arvidsson, A. et al. (1982). Kvinnor och barn. Intervjuer med kvinnor om familj och arbete. Statistics Sweden, Information i prognosfrågor, IPF 1982:4. (In Swedish.)

Beyer, J.E., M. Keiding and W. Simonsen (1976). The exact behaviour of the maximum likelihood estimator in the pure birth process and the pure death process. Scand. J. Statist. 3, 61-72.

Cassel, C.M., C.E. Särndal, J.H. Wretman (1979). Some uses of statistical models in connection with the nonresponse problem. Proceedings, Symposium on Incomplete Data, August 1979. Washington D.C.: National Academy of Sciences.

Cassel, C.M. (1982). Skattning av genomsnittliga intensiteten i en redovisningsgrupp då man tar hänsyn till både modellaspekter och urvalsaspekter. Statistics Sweden, Methodological Problems in Statistics on Individuals and Households, MIHS No 31. (In Swedish.)

Hoem, J.H. (1983a).Weighting, misclassification, and other issues in the analysis of survey samples of life histories. University of Stockholm, Department of Statistics, Stockholm Research Reports in Demography, No 11.

Hoem, J.H. (1983b). Balancing bias in vital rates due to an informative sampling plan. Statistical Review, 5.

Hoem, J.M., and U. Funck Jensen (1982). Multistate life table methodology: A probabilistic critique. Pp. 155-264 in K.C. Land and A. Rogers (eds.), Multidimensional Mathematical Demography. New York etc.: Academic Press.

Hoem, J.M., and B. Rennermalm (1982). Cohabitation, marriage, and first birth among never-married Swedish women in cohorts born 1936-60. University of Stockholm, Department of Statistics, Stockholm Research Reports in Demography, No. 8.

Johansson, L. and F. Finnäs (1983). Fertility of Swedish women born in 1927-1960. Statistics Sweden, Urval No. 14.

Lindström, H. (1983). Nonresponse errors in sample surveys. A discussion of causes and effects of nonresponse in sample surveys of living conditions. Statistics Sweden. Urval No. 16.

Little, R.J.A. (1982). Models for nonresponse in sample surveys. J. Amer. Statist. Ass. 77, 237-250.

Lyberg, I. (1983). Nonresponse effects on survey estimates in the analysis of competing exponential risks. University of Stockholm, Department of Statistics, Stockholm Research Reports in Demography, No. 12.

Platek, R., M.P. Singh, and V. Tremblay (1978). Adjustments for nonresponse in surveys. In N.K. Namboodiri (ed.), Survey Sampling and Measurement. New York etc.: Academic Press. Also in Survey Methodology Journal (Statistics Canada, 1977) **3**(2), 1-24.

Qvist, J. (1983). Fruktsamhets- och civilståndsutvecklingen i Sverige för kohorterna 1930-1960. Statistics Sweden, Forecasting Institute, 1982-01-28. (In Swedish.)

Schou, G. and M. Væth (1980). A small sample study of occurrence/exposure rates for rare events. Scand. Actuarial J. 1980(4), 209-225.

Statistisk Sentralbyrå (1981). Fruktbarhetsundersøkelse 1977. Fertility survey 1977. Norges offisielle statistikk B 197.

Särndal, C.E., and Tak-kee Hui (1981). Estimation for nonresponse situations: To what extent must we rely on models? Pp. 227-46 in D. Krewski, R. Platek, and J.N.K. Rao (eds.), Current Topics in Survey Sampling. New York etc.: Academic Press.

Væth, M. (1977). A note on the sampling distribution of the maximum likelihood estimators in a competing exponential risks model. <u>Scand Actuarial J</u>. 1977(2), 81-87.