Discussion Paper

Central Bureau of Statistics, P.B. 8131 Dep, 0033 Oslo 1, Norway

No. 42 February 1989

UNOBSERVED HETEROGENEITY IN MODELS OF MARRIAGE DISSOLUTION

By

Rolf Aaberge
Øystein Kravdal
Tom Wennemo

ABSTRACT

The goal of this paper is to examine the impact of unobserved heterogeneity when analysing the determinants of marriage dissolution. In the present analysis the parameter estimates of the explanatory variables appear to be insensitive to the omission of unobservables. The parameter estimates of the baseline hazard, however, are sensitive. When the unobserved heterogeneity is taken into consideration, the divorce risks increase steadily with duration. This supports the view that the declining hazard found in most studies of marital instability is due to a selection mechanism. Our analysis also demonstrates that the unobservables account for a considerable amount of the population variation in divorce propensity compared to the amount accounted for by the observed covariates.

1. INTRODUCTION

During the last three decades, when the lifelong marital commitment has gradually lost some of its attraction in a large part of the world, several scholars have devoted their attention to studies of sociodemographic divorce differentials. A variety of factors have proved to be closely correlated with the propensity of marriage dissolution: age at marriage, number of children, the children's age, timing of first birth relative to marriage, age difference between the spouses, education of husband and wife, place of residence, occupation, income and employment status of both spouses, religious denomination, church attendance and other socioeconomic or family background characteristics (see e.g., Becker et al., 1977; Bumpass and Sweet, 1972; Castro and Bumpass, 1987; Cherlin, 1977; Hoem and Hoem, 1988; Menken et al., 1981; Teachman, 1983; Trussell et al., 1988)

To our knowledge all these variables have never been analysed simultaneously. A subset of the observed demographic and socioeconomic characteristics are usually included in the analysis, and it is implicitly assumed that these observed covariates include all systematic sources of variation in divorce propensity. Even if all the variables referred to had been included, however, it would not be possible to capture all heterogenity in the population. Human decision-making and behaviour is, of course, far too complex to be completely explained by a set of standard socioeconomic and demographic factors. A considerable amount of variation in marital instability is likely to be caused by variables that are not observed or can hardly be observed. It is, for instance, difficult to obtain sufficient measurements of factors like social environment, normative barriers associated with a marital break-up, or the thoroughness of the search for a suitable mate. In most analyses of divorce, it is therefore reason to believe that important explanatory factors are omitted. As is well known, failure to adequately control for unobservables can produce severe bias in the parameter estimates of the included covariates, as well as create a misleading impression of duration dependence. The last problem may arise because of a dynamic selection process. The idea is that married couples have different divorce intensities, and that those with high intensities are selected out of the marriage cohort over time, leaving those with low intensities behind. In this way one gets a selection of married couples that are more robust against divorce than the rest. Consequently, failure to control for heterogeneity may lead the researcher to misinterpret the observed duration dependence as true duration dependence at the individual level.

Our approach is to estimate hazard models based on data for complete marriage cohorts of Norwegian women born after 1935. The standard procedure of controlling for unobserved heterogeneity in hazard models is to assume a functional form for the duration distribution given observed and unobserved covariates and a parametric functional form for the distribution of unobservables (see, e.g. Heckman and Willis (1977) and Harris (1982)). In this paper we apply a more flexible strategy proposed by Heckman and Singer (1982, 1984), where the distribution of unobservables is approximated by a multinomial distribution. Then the estimation problem consists of fitting mixing densities to data. This strategy has formerly been employed by Heckman and Singer (1984) for analysing durations of unemployment, by Trussel and Richards (1985) in a study of child mortality and second births, and by Montgomery (1987), who studied marriage formation and home ownership.

The primary goal of the present study is to examine the importance of a number of measured covariates of marriage dissolution in Norway compared to the importance of unobservables. Our focus is on judicial divorce in first marriage. To our knowledge, neither divorce nor separation has been analysed in the light of unobserved heterogeneity yet.

Secondly, we want to find out whether the duration dependence at the individual marriage level - controlling for unobservables - differs from that found on the basis of standard methods.

Our third goal is to test the sensitivity of the covariate effects estimates to the omission of unobservables.

Finally, we include a discussion of the importance of first birth timing, which is a divorce determinant we have previously paid particular attention to (Kravdal, 1988). Our intention is to see whether we can gain improved insight into the correlation between first birth timing and divorce when the unobserved heterogeneity is accounted for by the Heckman-Singer procedure.

2. SPECIFICATION OF THE STATISTICAL MODEL

Our basic framework is the standard proportional hazard model for single spell data assuming time invariant covariates. More precisely, the hazard function h is assumed to be of the form

$$\ln h(t|x,v) = \ln\psi(t) + \beta x + v \qquad (1)$$

where ψ is the baseline hazard, x is a vector of observed covariates and v is a variable that summarizes the effect of the omitted covariates. Then the distribution function of t given x and v, $F(t|x,v)$, satisfies

$$F(t|x,v) = 1 - \exp(-e^{\beta x+v} \int_0^t \psi(u)du) \qquad (2)$$

For reasons that will be explained in the next section, we have defined the duration observations to be from the beginning of the third year of marriage (t=0) and until divorce or the censoring time. Censoring may be due to death, emigration, widowhood, or having reached the end of the maximum observation time in our data set, which is 31 December 1984.

There are two distinct sources of variation in divorce propensity. The first is due to the fact that the spouses are making decisions under uncertainty. This is the reason why we have chosen a stochastic framework and have modelled instantaneous divorce probabilities.

The second source of variation is due to individual differences in attitudes or preferences and in the environment the spouses face. This variation is partly caught by the included observed covariates and the remainder by unobserved covariates. This fact represents the motivation for including the heterogeneity component v into the model.

In our analysis we assumed that v either was equal to zero or multinomial distributed. Then the distribution $F(t|x)$ takes the following form

$$F(t|x) = \sum_{j=1}^{s} q_j[1-\exp(-e^{\theta_j+\beta x} \int_0^t \psi_j(u)du)] \qquad (3)$$

where $\sum_{j=1}^{s} q_j = 1$, and v is assumed multinomial with s cells and support points at locations $\theta_1, \theta_2 ..., \theta_s$. The standard model without unobservables (v=0) is included in (3) and emerges for s=1.

Model (3) specifies that a randomly selected married couple has the probability q_j of belonging to group j where the marriage duration distribution is given by

$$F(t|x) = 1 - \exp(-e^{\theta_j+\beta x} \int_0^t \psi_j(u)du) \qquad (4)$$

In all but one of the model alternatives we have assumed that ψ_j is independent on j. This means that the intensity in one group($j=j_1$) is a constant multiple of the intensity in another group ($j=j_2$).

We have confined our study to situations where s is fixed equal to 2 or 3.

In the present analysis the baseline hazard ψ is assumed to be on the two following forms,

$$\psi(t) = c\lambda^c t^{c-1}(1+\lambda^c t^c)^{-1} \qquad (5)$$

or

$\psi(t)$ is a step function with 8 steps.

The step function parameters a_1 - a_8 correspond to intervals 0-1 year, 1-2 years, 2-3 years, 3-4 years, 4-5 years, 5-9 years, 9-11 years, 11-14 years. (By mistake step number 6 has been too wide, but this does not affect the main empirical conclusions that we draw in this paper.)

The specification in (5) means that the duration distribution is log logistic. The hazard function is monotone decreasing from infinity if $c < 1$ and is monotone decreasing from λ if $c = 1$. If $c > 1$ the hazard increases from 0 to a maximum at $t = \lambda^{-1}(c-1)^{1/c}$ and decreases thereafter.

As λ becomes very small, the log-logistic can be approximated by $\psi=kct^{c-1}$ where k is a constant. The function ct^{c-1} is known as the Weibull hazard. In some of the estimated models the log-logistic hazard degenerates to a Weibull hazard.

The models are estimated by a maximum likelihood procedure. Let us assume that we have N women in our sample. Furthermore, it is assumed that the individuals are independent and that they have the probabilities F(t) of becoming divorced before time t. If N1 women divorce at time t_i, i=1,2,....,N1,respectively, and N-N1 women are censored (as married) at time z_i, i=N1+1,....,N, respectively, the likelihood L is given by

$$L = \prod_{i=1}^{i=N1} f(t_i) \cdot \prod_{i=N-N1}^{i=N} (1-F(z_i)) \qquad (6)$$

where f(t) is the derivative of F(t) with respect to t.

The maximum likelihood is found by applying an optimization routine in the NAG-library. This routine is based on a modified Newton-Rapson algoritm. Several initial values were tried, and they all gave the same parameter estimates. This is no guarantee, of course, that the global (and not only a local) maximum is reached. Standard deviations of the parameters are estimated by inverting the Hessian matrix (consisting of partial derivatives of the likelihood with respect to the parameters).

3. DATA

3.1 Individual life histories

Our analysis is based on individual marriage and birth histories for complete cohorts of Norwegian women born after 1935. The life histories are derived from the Central Population Register. Data from the three censuses in 1960, 1970 and 1980 are added. This gives information on the socioeconomic characteristics of the women, their husbands and their parents as well as place of residence and religious denomination. This data source provides us with an excellent opportunity for detailed studies of some divorce determinants that are supposed to be important.

The Central Population Register was established in October 1964. Everyone who has lived in Norway for some period after 1960 has been assigned an identification number and is included in the register. No one is ever removed from the register, but there are codes showing whether a person is a resident, emigrated or dead. Each individual is represented by a data record containing purely demographic information like marital status, mother's and father's identification number etc. The register is continuously updated, but the history of codes, reflecting changes of status, is kept for analytical purposes. Therefore, it has been possible to construct individual histories - dating back to 1964 - of changes in marital status.

For those born after the register was established, the parents' identification numbers are included on the birth certificates and entered into the register when the newborn are entered. For older birth cohorts the parents' identification numbers were included for children who lived with their parents when the Population Census 1970 took place. This implies that the records contain identification numbers of mothers and fathers for most children born after 1953.

Reorganizing this material, we have been able to create a file with birth and marriage histories for all Norwegian women born after 1935. These histories are updated to 1984. The file also contains the identification number of the spouses as well as the fathers of the children.

Since very few women born after 1935 had a child before 1953, we have almost complete birth histories up to 1984 for the women in our data file. The marriage histories are complete from 1964. We have information about year of marriage for a large proportion of the women who married earlier, but to study the effects of a factor like first birth timing we need the exact date. Consequently, we have confined ourselves to marriages after 1964.

In the analysis presented in this paper we examine first marriages that were recently contracted when the 1970 census took place. Therefore, we have based this study on data for the 1968-1970 marriage cohorts. After some further limitations (to be described below) a population of 51000 women remained. 12 per cent cent of these marriages were dissolved by divorce before 31 December 1984.

3.2 Covariates included in the models

All covariates included in our models are constant during the observation period. In principle it would be possible, though much more complicated, to have time-varying covariates as well. It is well known from Norway (Kravdal, 1988) and other countries that the number of children and the age of the children have a considerable impact on divorce risks. Particularly the couples who are childless at a given duration of marriage, have a high propensity of subsequent divorce. Therefore, the childless are excluded in our calculations. We have focused on couples who are still married after two years, and who at that time have at least one child. It is not distinguished between those who at two years duration or at a later stage in the marital life course have one child and those who have more children. This will probably not affect the shape of the estimated hazard. By including the age of the child, however, the hazard might be depressed with increasing duration. Kravdal and Noack (1988) found that there was a sharper reduction of the intensity after 10 years of marriage if this variable was included than if it was left out. Stated otherwise: when the intensity does not show a more clear decline than we have observed in simple estimates of intensity by duration, it is because the children are getting older.

A covariate that is of particular interest for us, is the timing of the first birth. There are four important categories for this variable:

1: Women who have got a child before marriage, and who have married another man than the father of the child.
2: Women who have got a child before marriage, and who have subsequently married the father of the child.
3: Women who have got their first child during 0-6 months of marriage.
4: Women who have got their first child during 7-24 months of marriage.

The distinction between the first two groups is based on the personal identification numbers of husbands and fathers in the life histories. In our analysis group 1, which is quite small and has a very high divorce rate, is left out.

Two other important variables are the age at marriage and the time at marriage. The inclusion of the latter variable needs some explanation:

The census data contains variables describing the situation as of 1 November 1970. For women married 1968-1970 we know , for instance, the education and the place of residence as of 1 November 1970. At that time some women in the actual marriage cohorts have been married for 22 months, others for only a day, and those marrying in November or December 1970 are still spinsters. We have information about the woman as well as her parents and her husband. However, if she is separated at the time of the census or (of course) if she is not yet married, the characteristics of the husband are missing. For that reason we have left out couples not registered as married at the time of the census. Consequently the population under study consists of <u>women who have married January 1968 - October 1970, and who are not separated 1 November 1970</u>. Because we want to assess the effect of first birth timing and exclude the currently childless, it is most convenient to observe the women (those who are still married by then) <u>from their second marriage anniversary</u>, i.e., January 1970 - October 1972. (This corresponds to t=0.) For those married in the beginning of 1968, there will be very few divorces in the fourth year of marriage (from the beginning of 1971), as those separated 1 November 1970 are excluded. For those marrying in, 1970 there will be more divorces during the fourth year (1973), as the few couples removed due to separation 1 November 1970 have already divorced. This selection problem is only relevant in the first years of the marriage, when divorce rates are very small. It is not likely that the parameter estimates, being based on observations up to as late as 1984, will be affected. Nevertheless, time at marriage (in months since January 1970) is included as a covariate. This is also because we cannot rule out from the outset that there is <u>real</u> effect of even such a small span in historical time, as divorce risks have escalated during the last decades. It appears, however, that this variable has a small effect on the divorce rates. The difference in divorce intensity between January 1968 and October 1970 is not more than 20 per cent in any of the models estimated in the paper. Furthermore, the parameter estimates of the other covariates are not perceptibly changed when the time-at-marriage variable is excluded.

We will briefly describe the other variables that we have used and the categories. With one exception, which is clearly pointed out, they all refer to the situation as of 1 November 1970.

Place of residence: group 1: Non-rural parts of Eastern Norway
group 2: Rural parts of Eastern Norway, non-rural parts of Southern, Western, Middle and Northern Norway
group 3: Rural parts of Southern, Western, Middle and Northern Norway
(Norway consists of 5 main regions: Eastern, Southern, Western, Middle and Northern Norway.)

Woman's education and
husband's education : number of years at school (minimum 9, maximum 18)

Woman's educational activity and
husband's educational activity : student/not student

Her parents education (1 November 1960): group 1: missing, or not living with parents 1960
group 2: low
group 3: medium
group 4: high

Woman's taxable income per year and
husband's taxable income per year : 0-272000 (average: 7000 for women and 27000 for husbands)

Woman's occupation: group 1: no occupation
group 2: pedagogical, medical work
group 3: clerical work, sale, commerce
group 4: agriculture, forestry, fishing
group 5: industry, craft
group 6: hotel and restaurant work, charwork
group 7: technical work
group 8: post, telecommunication
group 9: other occupations

Husband's occupation: group 1: no occupation
group 2: pedagogical, medical work
group 3: clerical work, sale, commerce
group 4: agriculture, forestry, fishing
group 5: industry, craft

group 6: hotel and restaurant work, charwork
group 7: technical work
group 8: post, telecommunication
group 9: sea transport
group 10: other transport
group 11: military work
group 12: other occupations

"Other occupations" includes fairly small groups with widely different divorce rates: different kinds of service work, (buisiness) administration, police, firemen, religious work, juridical work, artists etc.

Woman's and husband's religious denomination:

group 1: both spouses members of the Norwegian Church
group 2: both spouses members of another religious society
group 3: none of them member of a religious society
group 4: other combinations

Age at marriage, time at marriage, education and income are treated as continuous covariates. These variables enter the model measured on the log scale and the log scale squared level. All other covariates are dummy variables.

4. EMPIRICAL RESULTS

The answers to the questions posed in the introduction are known to be quite sensitive to the functional form imposed on the hazard (Trussell and Richards, 1985). Accordingly, we believe that we will be on safest ground with a rather flexible structure. As stated in the methods section our primary choice has been the log-logistic function, which in our situation is either monotonically increasing within the relevant duration interval or rises to a maximum and then decreases. This functional form is also applied by Trussell and Richards (1985) and Montgomery (1987).

An increasing intensity during the first years of the marriage seems reasonable, as it, according to Norwegian law, takes at least one year between separation and judicial divorce. Besides, it takes time to build up a marital conflict and to make the final decision to split up. It is found in Norway, as well as in several other countries, that the divorce intensities (disregarding the possible effects of unobserved heterogeneity) level out or decline after 5-10 years. The log-logistic hazard will be able to fit such a pattern. There exists, however, very little theoretical guidance in specification of functional forms, but most authors emphasize that, in general, investment in marital specific capital (e.g., a common social network) should make the intensity decrease with duration.

We have also estimated models based on a semiparametric hazard where the baseline hazard takes the form of a step function. As our data set contains as much as 51000 observations, there is no need to economize very much with the parameters, so we have chosen 8 parameters in this second baseline hazard.

Since our data material comprises a very large number of individuals we have had the opportunity to include several covariates in our models and divide into a fairly large number of levels within each covariate. The estimated effects of the covariates may therefore be of some interest for those engaged in studies of divorce differentials. We will, however, refrain from commenting on these differentials, as our main concern in the current paper is the insight gained by controlling for unmeasured heterogeneity.

4.1 Estimated models

In the remaining part of this paper we refer to the following 5 models:

Equal baseline models:

1. Log-logistic function, standard model without unobserved heterogeneity
   a. 3 covariates
   b. 5 covariates
   c. 14 covariates
2. Log-logistic function (in one case degenerating to a Weibull), Heckman-Singer model with unobserved heterogeneity
   a. 3 covariates
   b. 5 covariates
   c. 14 covariates
3. Step function, standard model without unobserved heterogeneity
4. Step function, Heckman-Singer model with unobserved heterogeneity

Unequal baseline models:

5. Log-logistic and Weibull function, Heckman-Singer model with unobserved heterogeneity
   14 covariates

4.2 The Heckman-Singer approach with log-logistic baseline hazard

The estimates for models with a log-logistic baseline hazard are given in tables 1, 2 and 3. In model 1a, which is a standard model, only a few important demographic variables (age and time at marriage and first birth timing) are included as covariates. It appears that when we estimate a Heckman-Singer model with two support points (model 2a), -2 log L (L is the likelihood) is reduced by 37. This is a fairly large improvement of the model fit with only two additional parameters. The relative importance of the other covariates and the ranking of the categories does not differ very much from one model to the other, but the shapes of the hazards differ markedly. While there is a maximum at about 6 years duration in the standard model, the hazard is monotonically increasing for each of the two sub-populations according to the Heckman-Singer model. One sub-population, the high-risk group, comprises 10 per cent of the women, and the intensity in this group is 27 times higher than in the low-risk group (exp (3.3) = 27.1).

==> tab 1

When more covariates are added to the standard model, -2 log L is reduced by 554 (compare model 1c in table 2 and model 1a in table 1). This might seem to be a very large number, but one should keep in mind that as much as 37 parameters are added. We will not comment on the estimated covariate effects, but merely point out that the effect of first birth timing, in which we have taken a particular interest, actually is slightly increased when the other covariates are included. The high divorce risk among women with a birth prior to marriage is not explained by their socioeconomic characteristics.

==> tab 2

A Heckman-Singer model with two support points (model 2c) gives a further reduction of -2 log L (55 with 2 parameters). This means that although we have improved the model fit considerably by including several covariates, there is still room for further improvement. In terms of reduction of -2 log L the improvement is even better than the one obtained with a smaller number of covariates. Stated otherwise, the inclusion of more observables have not diminished the importance of taking the unobserved heterogeneity into account.

As noticed when we compared model 1a and model 2a, there are no important discrepancies between the corresponding parameter estimates of the two models. The variables have the same relative importance in models 1c and 2c, but the differences in divorce intensity between the different levels of the variables have increased. For example, the divorce intensity at age 18 is 3.3 times higher than at age 30 in model 1c and 5.0 times higher in model 2c. The divorce risk increases slightly with the woman's educational attainment, the difference being 7 per cent between elementary education and a master level, according to model 1c. Model 2c yields 8 per cent as a difference. There is a negative relationship between husband's education and the propensity of marital break-up. By comparing the same educational levels as above, we find that the divorce intensity is reduced by 40 per cent according to model 1c and

Table 1. Parameter estimates with standard errors in models based on log-logistic hazards

		Model 1a Standard model		Model 2a Heckman-Singer model (two support points)	
AGE AT MARRIAGE (years)	log	-31.737	(3.235)	-48.476	(4.673)
	$\log^2$	4.673	(0.527)	7.176	(0.754)
TIMING OF FIRST BIRTH	Before marriage	0		0	
	Within 0-6 months	-0.326	(0.044)	-0.446	(0.065)
	Within 7-24 months	-0.658	(0.047)	-0.912	(0.072)
TIME AT MARRIAGE (months/12)	log	842.290	(998.638)	1279.590	(1091.917)
	$\log^2$	-98.627	(117.738)	- 149.757	(128.735)
c		1.776	(0.046)	1.750	(0.041)
$\lambda \cdot 10^3$		11.669	(n.c.)	4.421	(0.912)
$\lambda^c \cdot 10^3$		0.369	(0.042)		
constant		-1746.909	(2117.632)	-2563.650	(2315.349)
q_1		-		0.099	(0.012)
Θ_1		-		3.303	(0.133)
-2 log L		50 273		50 236	

n.c.= not calculated

Table 2. Parameter estimates with standard errors in models based on log-logistic hazards

		Model 1c Standard model		Model 2c Heckman-Singer model (two support points)	
AGE AT MARRIAGE (years)	log	-29.529	(3.359)	-45.845	(4.574)
	$\log^2$	4.323	(0.547)	6.778	(0.738)
TIMING OF FIRST BIRTH	Before marriage	0		0	
	Within 0-6 months	-0.407	(0.045)	-0.542	(0.064)
	Within 7-24 months	-0.712	(0.049)	-0.953	(0.071)
TIME AT MARRIAGE (months/12)	log	1748.312	(1010.787)	2513.563	(1071.117)
	$\log^2$	-205.899	(119.179)	-295.914	(126.293)
PLACE OF RESIDENCE	Eastern Norway non-rural	0		0	
	Eastern Norway rural, remaining non-rural	-0.447	(0.027)	-0.614	(0.039)
	Remaining rural	-0.897	(0.045)	-1.207	(0.064)
LENGTH OF EDUCATION (years·10)	log	1.941	(6.709)	3.629	(8.505)
	$\log^2$	-0.190	(0.717)	-0.361	(0.907)
EDUCATIONAL ACTIVITY	not studying	0		0	
	studying	0.286	(0.073)	0.392	(0.100)
LENGTH OF EDUCATION FOR HUSBAND (years·10)	log	-1.872	(4.322)	-4.444	(5.473)
	$\log^2$	0.123	(0.458)	0.369	(0.578)
EDUCATIONAL ACTIVITY FOR HUSBAND	not studying	0		0	
	studying	-0.080	(0.053)	-0.128	(0.071)
INCOME (/1000+1)	log	0.087	(0.046)	0.095	(0.061)
	$\log^2$	-0.017	(0.015)	-0.015	(0.020)
INCOME FOR HUSBAND (/1000+1)	log	0.223	(0.056)	0.280	(0.076)
	$\log^2$	-0.087	(0.013)	-0.115	(0.017)
OCCUPATION	no occupation, unknown	0		0	
	pedagogical work, medical work	-0.174	(0.064)	-0.214	(0.083)
	clerical work, sales work	-0.028	(0.047)	-0.045	(0.063)
	agriculture etc.	-0.361	(0.180)	-0.380	(0.206)
	industry, craft etc.	0.097	(0.061)	0.152	(0.084)
	hotel, restaurant work, charwork	0.188	(0.060)	0.250	(0.084)
	technical work	0.021	(0.155)	0.069	(0.200)
	post, telecommunication	-0.015	(0.098)	-0.065	(0.132)
	other occupations	0.186	(0.091)	0.259	(0.128)

(cont.)

Table 2 cont.

OCCUPATION FOR HUSBAND	no occupation, unknown .	0		0	
	pedagogical work, medical work	-0.054	(0.111)	-0.026	(0.147)
	clerical work, sales work	-0.131	(0.091)	-0.154	(0.124)
	agriculture etc.	-0.509	(0.113)	-0.648	(0.148)
	industry, craft etc. ...	-0.137	(0.086)	-0.162	(0.119)
	hotel, restaurant work, charwork	0.369	(0.107)	0.531	(0.151)
	sea transport	0.449	(0.105)	0.659	(0.149)
	other transport	-0.037	(0.093)	-0.042	(0.128)
	technical work	-0.228	(0.108)	-0.263	(0.144)
	military work	-0.190	(0.094)	-0.235	(0.130)
	post, telecommunication	-0.022	(0.143)	0.031	(0.192)
	other occupations	0.144	(0.099)	0.207	(0.136)
PARENTS' EDUCATION	unknown, not living with parents	0		0	
	low	-0.176	(0.065)	-0.231	(0.088)
	medium	-0.102	(0.080)	-0.161	(0.109)
	high	0.342	(0.100)	0.441	(0.137)
COUPLE'S RELIGION	both Norwegian Church ..	0		0	
	both other rel. society	-0.685	(0.159)	-0.872	(0.187)
	none of them member of rel. society	0.665	(0.127)	0.888	(0.182)
	other combinations	0.236	(0.049)	0.341	(0.069)
c		1.773	(0.046)	1.750	(0.040)
$\lambda \cdot 10^3$		11.376	(n.c.)	4.931	(0.839)
$\lambda^c \cdot 10^3$		0.357	(0.040)		
constant		-3661.484	(2143.094)	-5258.126	(2270.672)
q_1		-		0.137	(0.016)
Θ_1		-		3.087	(0.120)
-2 log L		49 719		49 664	

n.c.= not calculated

45 per cent according to model 2c. The same structure is revealed for the categorical covariates, as is more clearly seen directly from the table. For instance, the divorce risk for women with a premarital birth relative to women who get their first child as married, is increased from 2.0 in model 1c to 2.6 in model 2c. Likewise, the risk in non-rural parts of Eastern Norway relative to the rural parts of Southern, Western, Middle and Northern Norway is increased from 2.5 to 3.4. Such a magnification of the covariate effects is also found by Trussell and Richards (1985) in their study of second births, where they used a log-logistic hazard and two support points.

In figure 1 we have plotted the predicted intensities for a particular group of women in order to illustrate the different shapes of the hazards, as well as the difference in intensity between the high-risk group and the low-risk group. The intensities for other categories of women have the same profile, but are generally smaller or larger.

==> fig 1

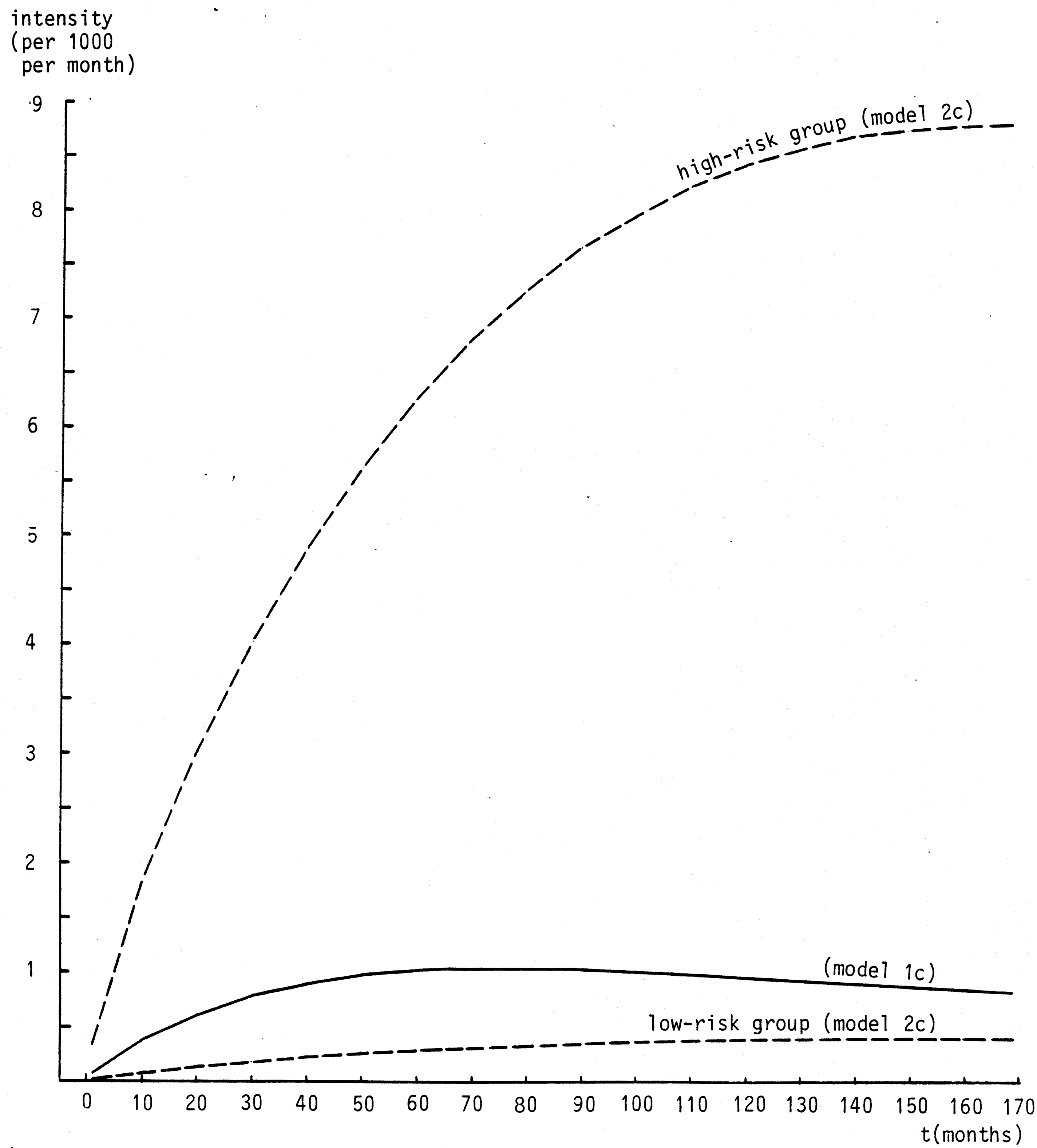
It is clearly seen that the baseline hazard has a different shape when the Heckman-Singer procedure is used (model 2c) than when we use a standard model (model 1c). The profiles resemble those found in model 1a and 2a. When we use a Heckman-Singer model, we get a hazard that rises with increasing duration, though with a gradually smaller steepness. According to the standard model, however, the hazard has a maximum at about 6 years duration. The shape of this hazard is illustrated more clearly in figure 2.

==> fig 2

It also turns out that the proportions of women in the high- and low-risk groups in model 2c and the difference between the intensities of these two groups are quite close to that found for model 2a. About 14 per cent of the women are in the group with a very high divorce risk. Actually, the risk is 22 times higher than for the remaining 86 per cent of the marriage cohort. In other words, we are fairly close to a socalled "mover-stayer" situation. A small group experience that a very large proportion divorce, while the majority have quite few marital breakdowns. Among the female population selected when drawing figure 1, the proportion of divorce predicted according to model 1c is 13.7 per cent, while model 2c gives 67.1 per cent in the high-risk group and 5.0 per cent in the low-risk group. (These proportions are the contributions to F(t) in equation (3) from the two sub-populations.)

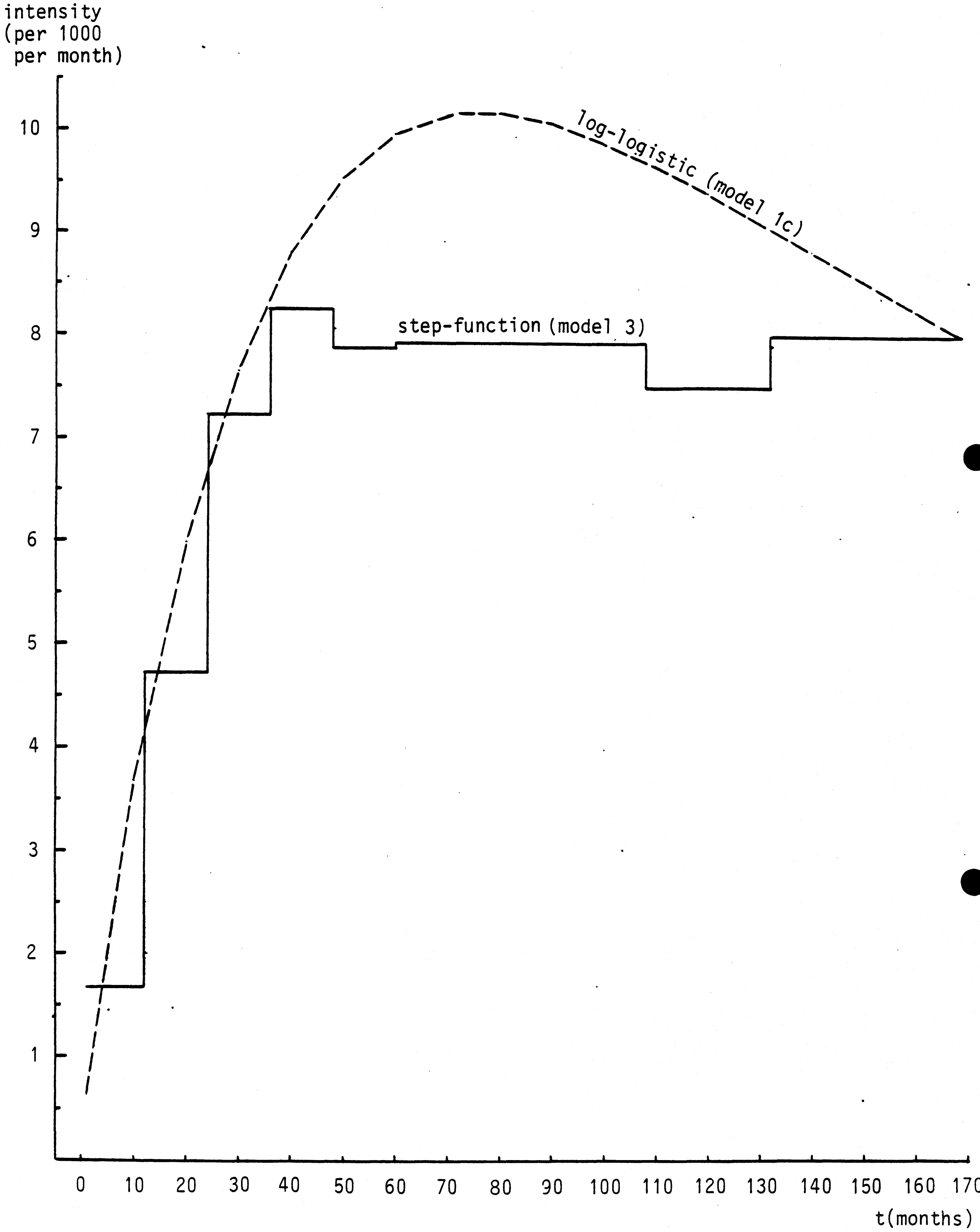
Employing a more flexible model (model 5) with different c and λ for each of the two subgroups (i.e. c_1, λ_2 for j=1 and c_2, λ_2 for j=2 in equation (3)) we obtain a further increase in goodness of fit by a reduction of 10 measured by -2 log L. The estimates are given in table 3, and predicted intensities are plottet in figure 3. We get a much larger high-risk group than in the equal baseline intensity model; 29 per cent of the women are in the high-risk group.

Figure 1. Predicted intensities[1)] based on model 1c and model 2c



1) For women who marry at age 22 in the middle of 1968, who have their first child more than 7 months after marriage, who have income 5000, who have a husband with income 40000, who have 9 years of schooling, and who have husbands with 9 years of schooling. With respect to the other covariates the women are assumed to belong to the baseline categories.

Figure 2. Shape of hazard[1)] according to model 1c and 3



1) Adjusted with a multiplicative factor so that both curves end in the same point.

Furthermore, the intensity profile for this group shows a clear maximum. For the low-risk group, however, the intensity increases with duration. With a log-logistic function for the low-risk group it was difficult to find a maximum of the likelihood. However, it was obvious that λ would be very small for the low-risk group, so we substituted with a Weibull function: $\psi(t) = k \cdot c \cdot t^{c-1}$, which is a good approximatiation to the log-logistic as λ becomes small. The constant k was considered as a part of the constant term in the β-vector. The resulting maximum likelihood turned out to be virtually identical to the one we obtained with a log-logistic hazard when we terminated the non-converging search process.

==> tab 3

==> fig 3

According to model 5 the intensity for the high-risk group relative to the low-risk group depends on duration. The relative intensity is very high in the first stage of marriage, but is only 17 at 10 years and 6 at 14 years. The average value (obtained by comparing integrals under the intensity curves) is 22, which is equal to the result we found with the equal baseline intensity model (model 2c). The corresponding divorce probabilities, which are 67 and 5 per cent according to model 2c, are 39 and 2 per cent according to model 5. In other words, we are even closer to a "mover-stayer" situation.

We have also estimated equal baseline models with three support points (see model 2c, table 3). The reduction of -2 log L is not negligible: 13 with 2 additional parameters. According to the estimated model, a small group of married couples (about 6 per cent) are found to have a very high intensity, a larger group (about 23 per cent) a medium intensity, and the remaining women have very stable marriages. We cannot rule out the possibility that a further increase to four support points would have given even greater improvement of the model fit, but due to the very long CPU-times required to estimate our most complex models, we have not given priority to an eleboration of this matter.

4.3 The Heckman-Singer approach with semiparametric baseline hazard

It is interesting to note that we obtain almost the same results when we employ a semiparametric baseline hazard. Particularly the parameter effects for the covariates and the size and relative intensities of the high-risk and low-risk groups (see table 4) are very close to those estimated in models where we assumed a log-logistic baseline hazard (see table 2). The shape of the hazards do not differ much either. In figure 2 we have shown the shapes according to the standard models 1c and 3. (As we are interested in the shape rather than the absolute size, the scale is not important. For convenience the curves are ajusted with a multiplicative factor so that they end in the same point.) It appears that the semiparametric

Table 3. Parameter estimates with standard errors in models based on log-logistic and Weibull hazards

		Model 2c Heckman-Singer model[1] (three support points)		Model 5 Heckman-Singer model with unequal baselines[2] (two support points)	
AGE AT MARRIAGE (years)	log	-53.253	(5.342)	-44.504	(4.317)
	log²	7.884	(0.859)	6.616	(0.695)
TIMING OF FIRST BIRTH	Before marriage	0		0	
	Within 0-6 months	-0.621	(0.074)	-0.527	(0.061)
	Within 7-24 months	-1.092	(0.083)	-0.905	(0.067)
TIME AT MARRIAGE (months/12)	log	2091.983	(1201.016)	1933.770	(1132.484)
	log²	-246.108	(141.610)	-227.477	(133.533)
PLACE OF RESIDENCE	Eastern Norway non-rural	0		0	
	Eastern Norway rural, remaining non-rural	-0.697	(0.047)	-0.575	(0.037)
	Remaining rural	-1.367	(0.076)	-1.113	(0.059)
LENGTH OF EDUCATION (years·10)	log	3.653	(9.392)	2.456	(7.896)
	log²	-0.365	(1.001)	-0.243	(0.842)
EDUCATIONAL ACTIVITY	not studying	0		0	
	studying	0.469	(0.121)	0.416	(0.099)
LENGTH OF EDUCATION FOR HUSBAND (years·10)	log	-6.097	(6.231)	-2.729	(5.086)
	log²	0.531	(0.658)	0.200	(0.538)
EDUCATIONAL ACTIVITY FOR HUSBAND	not studying	0		0	
	studying	-0.167	(0.081)	-0.101	(0.066)
INCOME (/1000+1)	log	0.133	(0.070)	0.091	(0.056)
	log²	-0.025	(0.022)	-0.016	(0.018)
INCOME FOR HUSBAND (/1000+1)	log	0.307	(0.086)	0.270	(0.071)
	log²	-0.130	(0.020)	-0.108	(0.016)

(cont.)

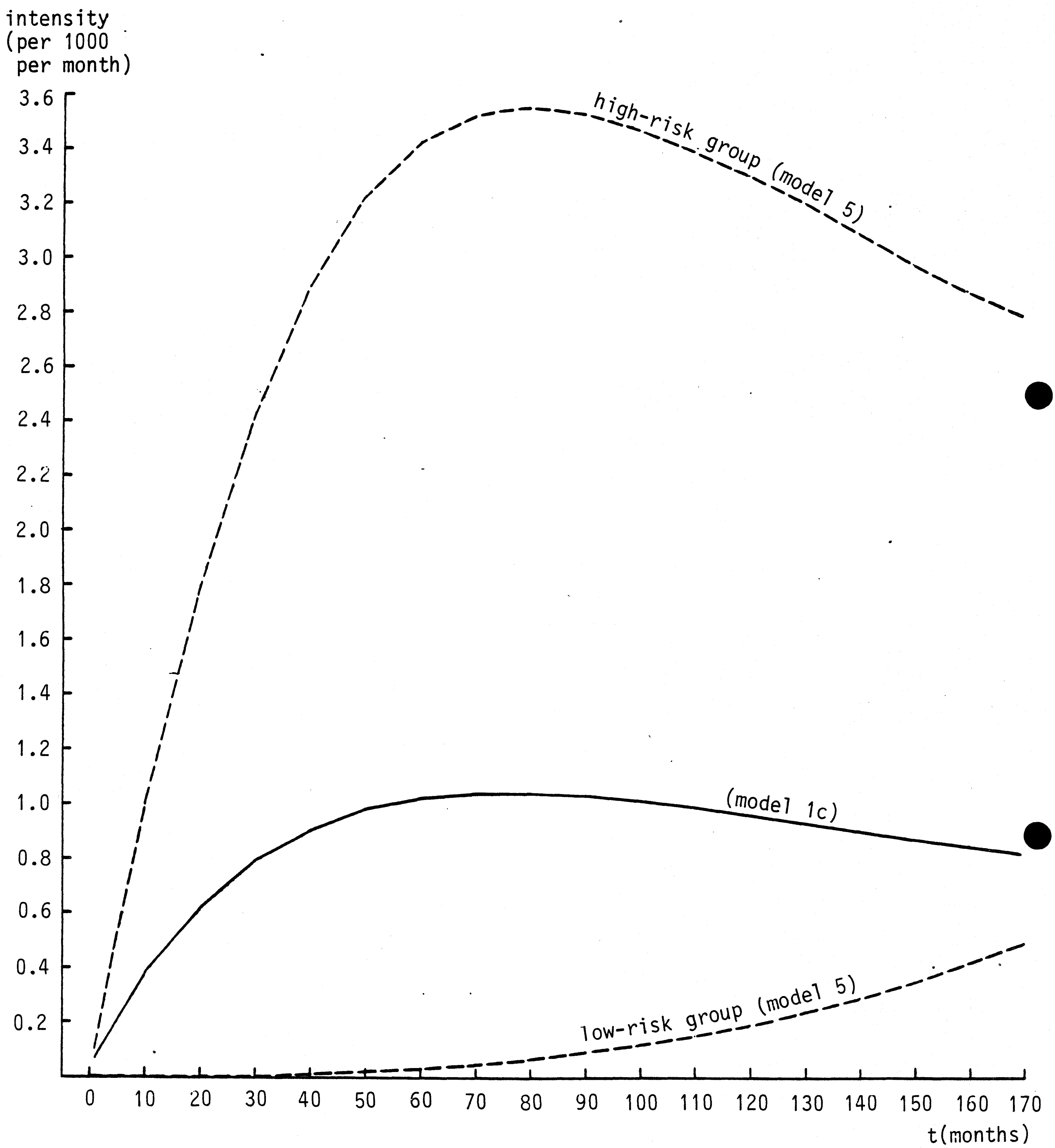
[1]) log-logistic hazards

[2]) log-logistic hazard (defined by c_1 λ_1) and Weibull hazard (defined by c_2)

Table 3 cont.

OCCUPATION	no occupation, unknown .	0		0	
	pedagogical work, medical work	-0.234	(0.094)	-0.182	(0.077)
	clerical work, sales work	-0.055	(0.072)	-0.024	(0.058)
	agriculture etc.	-0.413	(0.228)	-0.360	(0.195)
	industry, craft etc. ...	0.188	(0.095)	0.160	(0.079)
	hotel, restaurant work, charwork	0.302	(0.097)	0.236	(0.079)
	technical work	0.035	(0.228)	0.073	(0.190)
	post, telecommunication	-0.092	(0.147)	-0.062	(0.119)
	other occupations	0.280	(0.144)	0.283	(0.119)
OCCUPATION FOR HUSBAND	no occupation, unknown .	0		0	
	pedagogical work, medical work	-0.032	(0.164)	0.000	(0.135)
	clerical work, sales work	-0.173	(0.140)	-0.110	(0.113)
	agriculture etc.	-0.769	(0.168)	-0.553	(0.136)
	industry, craft etc. ...	-0.198	(0.134)	-0.116	(0.108)
	hotel, restaurant work, charwork	0.616	(0.174)	0.553	(0.141)
	sea transport	0.737	(0.166)	0.643	(0.137)
	other transport	-0.058	(0.145)	0.003	(0.118)
	technical work	-0.293	(0.162)	-0.209	(0.132)
	military work	-0.254	(0.147)	-0.172	(0.119)
	post, telecommunication	0.039	(0.214)	0.006	(0.177)
	other occupations	0.262	(0.152)	0.228	(0.124)
PARENTS' EDUCATION	unknown, not living with parents	0		0	
	low	-0.236	(0.100)	-0.169	(0.081)
	medium	-0.148	(0.123)	-0.091	(0.100)
	high	0.545	(0.158)	0.440	(0.126)
COUPLE'S RELIGION	both Norwegian Church ..	0		0	
	both other rel. society	-0.988	(0.217)	-0.802	(0.177)
	none of them member of rel. society	1.192	(0.227)	0.944	(0.189)
	other combinations	0.381	(0.078)	0.313	(0.065)
c_1		1.858	(0.048)	1.918	(0.054)
c_2		-		3.702	(1.009)
$\lambda_1 \cdot 10^3$		3.281	(0.872)	11.930	(1.100)
constant		-4349.796	(2546.413)	-4050.440	(2400.647)
q_1		0.056	(0.010)	0.289	(0.037)
q_2		0.230	(0.042)	-	
Θ_1		5.191	(0.327)	21.627	(5.297)
Θ_2		3.010	(0.209)	-	
-2 log L		49 651		49 654	

Figure 3. Predicted intensities[1)] based on model 1c and model 5



1) see note 1 figure 1

hazard has an almost constant level from 6 years duration, while there is a decline for the log-logistic hazard. Based on a log-logistic hazard it is actually impossible to obtain a sudden change from a steep rise to a virtually constant level. Figure 4 shows a similar plot for the shapes obtained for the Heckman-Singer models 2c and 4. In this case the two hazards have almost identical shapes.

==> tab 4

==> fig 4

4.4 Brief discussion of the duration dependence

The commonly observed divorce profile - with a steep rise in the initial stage of the marriage and a subsequent reduction or levelling out - is, according to our results, explained primarily by a selection mechanism. Apparently, the decline of the observed hazard in standard models is due to a very quick thinning-out of the most divorce-prone women (the high-risk group in for instance model 2c). It is somewhat surprising that according to the most flexible model (model 5) the divorce risk within the low-risk group keeps rising over the entire 14 years interval that we study. The longer the marriage has lasted, the more likely is the subsequent break-up. A similar profile was found for both high- and low-risk groups according to model 2c. This pattern could be consistent with an hypothesis that the spouses little by little "wear each other out".

Traditional theories contrast in a very marked way with our empirical suggestions that marriages are becoming more and more weak. Having the empirically observed hazard in mind, it has often been argued among demographers that investment in socalled marital specific capital (see e.g., Becker et al., 1977) tends to reduce divorce rates by increasing duration of marriage. Sticking to the economist concepts, we might say that this capital includes "goods" which are not so easily enjoyed when the marriage is broken, and which may be more valuable relative to other "goods" the longer the marriage has lasted. Of crucial importance is, for instance, contact with children, housing facilities that have gradually been built up through joint efforts, and the social network which has been created around the nuclear family.

In this connection we would like to repeat what we briefly mentioned previously: If we include number of children, and in particular the age of the youngest child, as time-varying covariates, we might have forced the hazard to decrease or level out after some years of marriage.

4.5 Importance of unobservables compared to the observed variables

It seems reasonable to give some consideration to the widely differing prevalence of divorce in the high-risk group compared to the low-risk group. The divorce risk in the high-risk group relative to the low-risk group is as large as 22 in model 2c (table 2), while a

Table 4. Parameter estimates with standard errors in models based on step-function hazards

		Model 3 Standard model		Model 4 Heckman-Singer model (two support points)	
AGE AT MARRIAGE (years)	log	-29.536	(3.359)	-44.945	(4.442)
	log²	4.324	(0.547)	6.662	(0.716)
TIMING OF FIRST BIRTH	Before marriage	0		0	
	Within 0-6 months	-0.407	(0.045)	-0.525	(0.062)
	Within 7-24 months	-0.713	(0.049)	-0.917	(0.068)
TIME AT MARRIAGE (months/12)	log	1661.892	(973.509)	2308.690	(1056.972)
	log²	-195.525	(114.784)	271.769	(124.625)
PLACE OF RESIDENCE	Eastern Norway non-rural	0		0	
	Eastern Norway rural, remaining non-rural	-0.447	(0.027)	-0.592	(0.038)
	Remaining rural	0.898	(0.045)	-1.159	(0.061)
LENGTH OF EDUCATION (years·10)	log	1.966	(6.710)	3.075	(8.188)
	log²	-0.193	(0.717)	-0.305	(0.873)
EDUCATIONAL ACTIVITY	not studying	0		0	
	studying	0.287	(0.073)	0.396	(0.097)
LENGTH OF EDUCATION FOR HUSBAND (years·10)	log	-1.868	(4.322)	-3.868	(5.269)
	log²	0.123	(0.458)	0.314	(0.557)
EDUCATIONAL ACTIVITY FOR HUSBAND	not studying	0		0	
	studying	-0.080	(0.053)	-0.125	(0.068)
INCOME (/1000+1)	log	0.087	(0.046)	0.093	(0.058)
	log²	-0.017	(0.015)	-0.015	(0.019)
INCOME FOR HUSBAND (/1000+1)	log	0.224	(0.056)	0.274	(0.073)
	log²	-0.087	(0.013)	-0.112	(0.017)
OCCUPATION	no occupation, unknown	0		0	
	pedagogical work, medical work	-0.174	(0.064)	-0.202	(0.080)
	clerical work, sales work	-0.028	(0.047)	-0.041	(0.060)
	agriculture etc.	-0.361	(0.180)	-0.376	(0.201)
	industry, craft etc.	0.097	(0.061)	0.149	(0.081)
	hotel, restaurant work, charwork	0.188	(0.060)	0.241	(0.081)
	technical work	0.021	(0.155)	0.065	(0.195)
	post, telecommunication	-0.015	(0.098)	-0.063	(0.125)
	other occupations	0.185	(0.091)	0.241	(0.123)

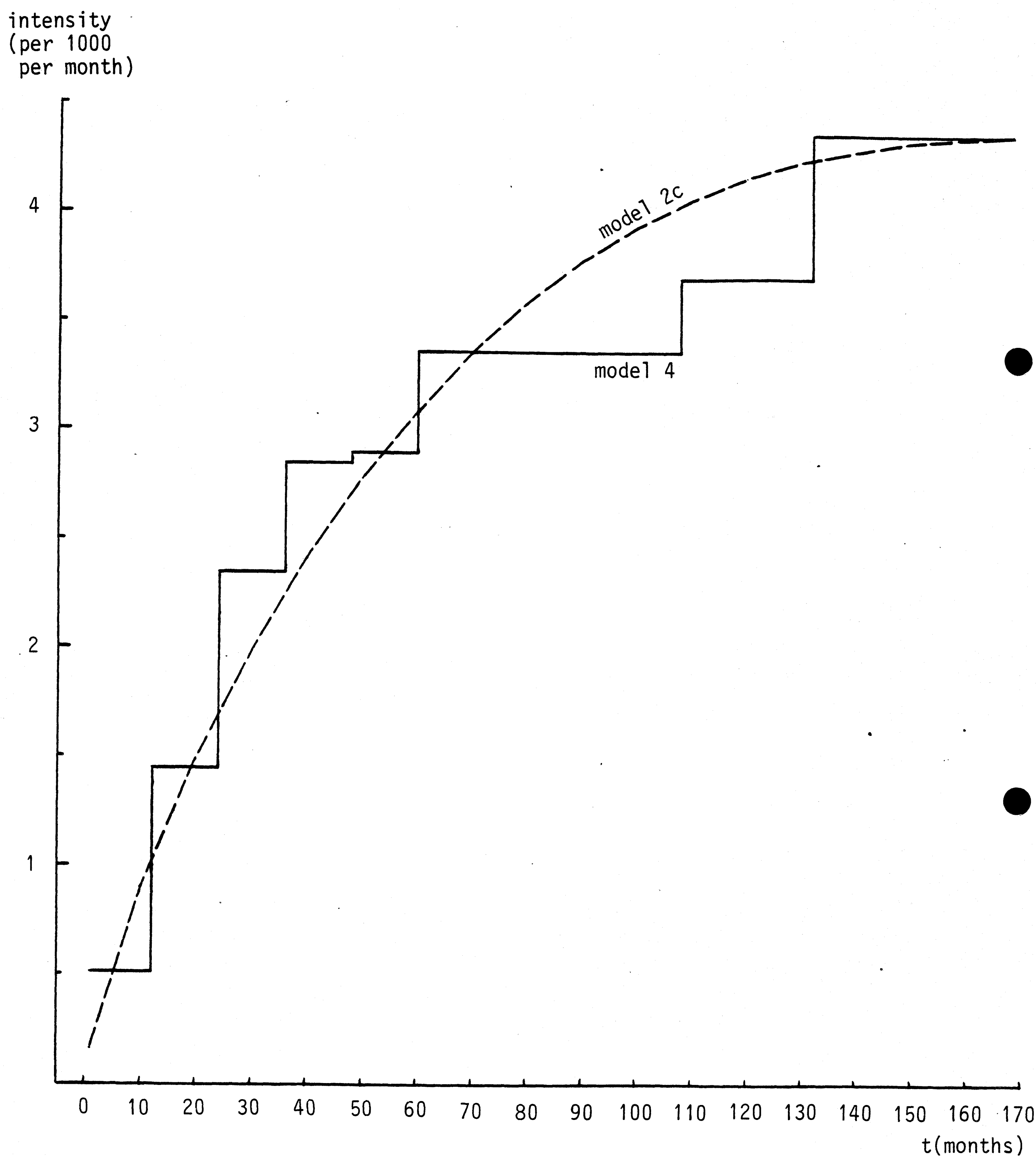
(cont.)

Table 4 cont.

OCCUPATION FOR HUSBAND	no occupation, unknown .	0		0	
	pedagogical work, medical work	-0.054	(0.111)	-0.027	(0.141)
	clerical work, sales work	-0.131	(0.091)	-0.143	(0.119)
	agriculture etc.	-0.509	(0.113)	.-0.613	(0.142)
	industry, craft etc. ...	-0.137	(0.086)	-0.153	(0.113)
	hotel, restaurant work, charwork	0.369	(0.107)	0.531	(0.145)
	sea transport	0.449	(0.105)	0.629	(0.143)
	other transport	-0.037	(0.093)	-0.038	(0.123)
	technical work	-0.228	(0.108)	-0.247	(0.138)
	military work	-0.198	(0.094)	-0.218	(0.124)
	post, telecommunication	-0.022	(0.143)	0.036	(0.185)
	other occupations	0.144	(0.099)	0.203	(0.130)
PARENTS' EDUCATION	unknown, not living with parents	0		0	
	low	-0.176	(0.065)	-0.208	(0.085)
	medium	-0.102	(0.080)	-0.133	(0.105)
	high	0.342	(0.100)	0.444	(0.132)
COUPLE'S RELIGION	both Norwegian Church ..	0		0	
	both other rel. society	-0.685	(0.159)	-0.836	(0.182)
	none of them member of rel. society	0.666	(0.127)	-0.897	(0.178)
	other combinations	0.236	(0.049)	0.324	(0.067)
constant		-3477.267	(2064.077)	-4826.083	(2240.838)
step function parameters:					
$a_1 \cdot 10^3$		1.791	(0.166)	0.985	(0.113)
a_2		5.023	(0.308)	2.836	(0.253)
a_3		7.696	(0.407)	4.579	(0.359)
a_4		8.778	(0.447)	5.564	(0.397)
a_5		8.368	(0.436)	5.647	(0.379)
a_6		8.431	(0.311)	6.556	(0.298)
a_7		7.963	(0.351)	7.214	(0.326)
a_8 [1])		8.500		8.500	
q_1		-		0.164	(0.023)
Θ_1		-		2.902	(0.115)
-2 log L		49 712		49 659	

[1]) fixed to 8.500

Figure 4. Shape of hazard[1] according to model 2c and 4



1) see note 1 figure 2

relative risk of about 3 must be considered large for the effects of the observed variables. For instance, the risk of those with a premarital birth is 2.6 relative to those who have a more traditional birth-marriage-sequence. Women living in the non-rural areas of Eastern Norway and who have a child when they marry have a risk which is 8.7 times higher than that of women from rural districts in, say, Western Norway who get their first child subsequent to marriage. Even this large difference between two fairly extreme groups is smaller than the difference between the high-risk group and the low-risk group. Also when we apply group specific baseline intensities the influence of unobservables is important compared to the effects of observed covariates.

4.6 The impact of first birth timing on divorce when unobserved heterogeneity is taken into consideration

We now turn to the models estimated in order to throw some light on the correlation between first birth timing and divorce. As already referred to, it is well documented that there is a close link between timing of first birth relative to marriage and propensity of marital breakdown (see e.g., Billy et al., 1986; Christensen and Meissner, 1953; McLaughlin et al., 1986; Morgan and Rindfuss, 1985; O'Connell and Rogers, 1984; Teachman, 1983). These relations were confirmed in a recent study based on individual life histories extracted from the Central Population Register of Norway (Kravdal 1988). In particular, it was found that the high divorce risks among women who have got a child prior to marriage, is not confined to those who have married another man than the father of their child. Apparently, also the women who have got a premarital child with their husband, run a much higher risk of marital breakdown than those who have got their first baby in-wedlock.

Several explanations are advanced to account for the high divorce propensity of women who have a child when they marry or who are pregnant. One of the theories that have attracted a good deal of attention, claims that the marriages contracted subsequent to childbearing or conception are quickly arranged in order to "legitimize" the baby. According to, for instance, Furstenberg (1979), "such haste may produce less compatible matches".

It has also been argued that there is an association between first birth timing and some socioeconomic characteristics in the beginning of the marriage or later in life. O'Connell and Rogers (1984) have found that - controlling for age at marriage - couples who have a child before marriage or who were pregnant at marriage, are more economically disadvantaged than others. They believe that this also affects their divorce rates. In this connection we want to add that any effect of income or other socioeconomic characteristics may have a direct effect, but may also work in a more spurious way. If it is found that women with a premarital birth often have low education, the reason may be that an unplanned pregnancy has disrupted the educational carreer, and that this in turn affects the marital stability. Alternatively, there may be a tendency for women with low education to have a baby out-of-wedlock as well as to experience a divorce later in life.

Our model estimates have shown that inclusion of the couple's educational level at marriage, the educational level of their parents and several other covariates does not diminish the relative divorce risk of women who have had a child prior to marriage. This has led us to hypothesize that there may be a certain spuriousness due to variables that we have not included. We hold the view that the couple's attitudes, and in particular their family values, may affect both their probability of having a child prior to marriage and their inclination to dissolve an unhappy marriage. Such an explanation has received some attention in the literature (e.g., Furstenberg 1979), but perhaps less than it might deserve.

Within the analytical framework we have chosen, we will never be able to confirm such an hypothesis, but we hope at least to find some support for it. Employing the Heckman-Singer procedure with a mixture of two survival time distributions, we are able to split the individuals into two groups: a high-risk group with women who are very divorce-prone, and a low-risk group with women who are likely to experience very stable marriages. Whether this heterogeneity should be explained by attitudes or other individual characteristics will, of course, remain a matter of speculation. If we estimate separate models for women with a premarital birth and for women who have had their first child more than 7 months after marriage, we can identify differences in the relative size of the high-risk group. According to our hypothesis we might expect that there are more women with "modern" values among those with a premarital birth, and that this could be reflected in a larger proportion in the high-risk group.

For simplicity we have assumed the same baseline intensity for both the high- and low-risk group.

It was shown in table 1 and table 2, and has already been commented on, that when more covariates were included, the effect of first birth timing actually increased (comparison of models 1a and 1c). An estimated Heckman-Singer model with two support points (model 2c), led to a further increase. In this heterogeneity model, however, we have forced the probability mass over the two support points to be equal for all groups, while we stated above that we expected to detect a larger proportion of couples selected for high divorce risks among women with a birth prior to marriage than among the others. Therefore, we have also estimated separate models for women who had their first child before marriage (in the remaining part of the paper referred to simply as group 1) and those who had their first child more than 7 months after marriage (referred to as group 2). Since the former group consists of no more than 3400 women, we had to impose some restrictions with respect to the number of covariates and categories. Table 5 shows the results obtained when only a subset of the covariates and the categories are included. This is called model 2b.

==> tab 5

Table 5. Parameter estimates with standard errors for two separate sub-populations, using a Heckman-Singer strategy with two support points and a log-logistic or Weibull hazard (model 2b)

		Group 1[1]		Group 2[2]	
AGE AT MARRIAGE (years)	log	-16.101	(13.209)	-80.450	(9.536)
	$\log^2$	2.082	(2.129)	12.043	(1.518)
PLACE OF RESIDENCE	Eastern Norway non-rural	0		0	
	Eastern Norway rural, remaining non-rural	-0.651	(0.145)	-0.684	(0.073)
	Remaining rural	-1.479	(0.205)	-1.496	(0.124)
LENGTH OF EDUCATION	Elementary	0		0	
	More than elementary	0.195	(0.147)	0.117	(0.072)
LENGTH OF EDUCATION FOR HUSBAND	Elementary	0		0	
	More than elementary	-0.118	(0.133)	-0.327	(0.072)
PARENTS' EDUCATION	unknown, not living with parents	0		0	
	Elementary	-0.487	(0.227)	-0.034	(0.159)
	More than elementary	-0.818	(0.381)	0.326	(0.181)
c		1.757	(0.129)	1.838	(0.077)
$\lambda \cdot 10^3$		-		4.442	(1.468)
constant		19.658	(20.541)	131.442	(14.945)
q_1		0.109	(0.030)	0.088	(0.017)
Θ_1		3.203	(0.322)	3.602	(0.212)
-2 log L		4 583		15 494	

[1]) women with a premarital birth. Weibull hazard

[2]) women who have had their first child more than 7 months after marriage. Log-logistic hazard

The results for group 1 revealed that the parameter λ would be very small. It was difficult, however, to find the exact value corresponding to a maximum of the likelihood. This was solved by substituting with a Weibull function.

As previously supposed, there is a larger proportion of high-risk women in group 1 than in group 2, but the difference is rather small. In group 1 the proportion is 11 per cent, while it is 9 per cent in group 2 (In a group consisting of women who were pregnant at marriage the proportion was 13.) We may conclude that only a small part of the difference that has previously been found between divorce risks for women in group 1 and women in group 2, is due to a larger group of high-risk women in the former group. An interesting question that we will try to answer is whether there are large differences in divorce risks between group 1 and group 2 among the high-risk women regarded as a separate population, and among the low-risk women as well.

In order to answer this question we start with an illustration (see figure 5) of divorce differentials for a certain group of women (married at age 22, elementary education both for the woman, her parents and her husband, and living in the non-rural parts of Eastern Norway). The intensities are predicted according to table 5.

==> fig 5

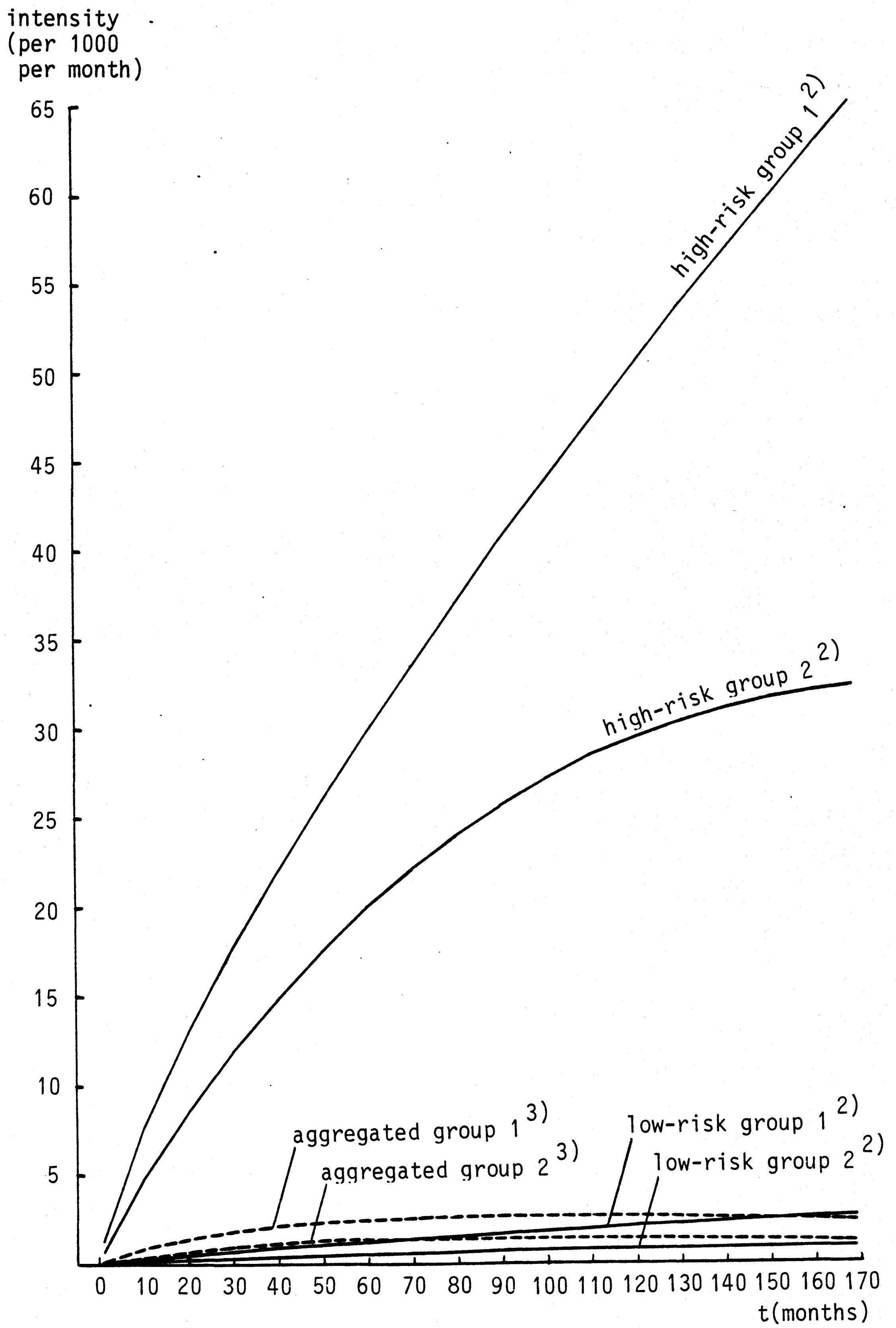
To compare the intensities when there is non-proportionality between group 1 and group 2, we have preferred to calculate the integral A of the intensity over the entire duration interval. For high-risk women in group 1 this integral is 6.32, and for low-risk women it is 0.26. The corresponding values for women in group 2 are 3.75 and 0.10. (The partial divorce probabilities ($F = 1- e^{-A}$) for these four sub-populations are 0.998, 0.227, 0.977 and 0.097, respectively.) This means that within the high-risk group the A-value for group 1 is 1.68 times that of group 2, whereas this quotient within the low-risk group is 2.51.

Now, we shall compare these results with the results from a standard model. The parameter estimates are shown in table 6, and predictions are plotted for a group of women in figure 5 (dotted lines). This is, of course, the same group as for the other curves in figure 5.

==> tab 6

Based on this model we get A-values equal to 0.37 for group 1 and 0.20 for group 2 (and partial probabilities 31.1 and 18.2 per cent). This yields an A-quotient of 1.86, which is higher than 1.68 obtained for the high-risk group as a separate population, and lower than 2.51 obtained for the low-risk group. These figures, of course, would vary with the value of age at marriage, education etc., but results not tabulated show that we get an acceptable picture of the

Figure 5. Predicted intensities for women[1] with a premarital birth (group 1) and women who had their first child more than 7 months after marriage (group 2)



1) Married at age 22, living in the non-rural parts of Østlandet, elementary education both for the woman, her parents and her husband.
2) Based on the model estimates given in table 5.
3) Based on the model estimates given in table 6.

Table 6. Parameter estimates with standard errors for two separate sub-populations, using a standard model and a log-logistic hazard (model 1b)

		Group 1[1]		Group 2[2]	
AGE AT MARRIAGE (years)	log	-14.906	(9.066)	-52.403	(6.629)
	$\log^2$	2.060	(1.463)	7.804	(1.065)
PLACE OF RESIDENCE	Eastern Norway non-rural	0		0	
	Eastern Norway rural, remaining non-rural	-0.431	(0.092)	-0.489	(0.049)
	Remaining rural	-1.016	(0.125)	-1.096	(0.087)
LENGTH OF EDUCATION	Elementary	0		0	
	More than elementary	0.192	(0.099)	0.099	(0.051)
LENGTH OF EDUCATION FOR HUSBAND	Elementary	0		0	
	More than elementary	-0.096	(0.090)	-0.215	(0.050)
PARENTS' EDUCATION	unknown, not living with parents	0		0	
	Elementary	-0.277	(0.167)	-0.006	(0.125)
	More than elementary	-0.534	(0.260)	0.235	(0.137)
c		1.705	(0.138)	1.849	(0.088)
$\lambda \cdot 10^3$		8.598	(2.429)	10.802	(1.353)
constant		25.394	(13.995)	85.014	(10.283)
-2 log L		4 585		15 513	

[1]) women with a premarital birth
[2]) women who have had their first child more than 7 months after marriage

general situation. The A-quotient in the standard model has an average of 2.59 (not weighted), which agrees fairly well with differences in β-estimates around 0.75 between these two birth-marriage sequences. Most quotients attain values between 1.5 and 3.5, and the standard deviation is 1. Within the low-risk sub-population the A-quotient is generally larger than obtained on the basis of the standard model. Within the high-risk population it is more difficult to draw a clear conclusion, but the average (not weighted) is 2.48, and this is slightly smaller than 2.59.

We believe that the value of such arguments, whether they are based on A-quotients or probability (F(t)) quotients, which would have been an acceptable alternative, should not be over-emphasized. However, it seems reasonable to conclude that within the high- and low-risk populations there is an effect of first birth timing not very different from that found on aggregate level (i.e, with standard models). Neither inclusion of several covariates nor controlling for unobserved heterogeneity reduces the large effect of first birth timing on divorce propensity.

5. SUMMARY

In this article we have applied a Heckman-Singer type of duration models. Our focus has been on judicial divorce in first marriage.

The analysis is based on individual birth and marriage histories extracted from the Central Population Register of Norway. These biographies are linked to socio-demographic characteristics recorded in the Population Censuses 1960, 1970 and 1980. The model estimates are based on data for all Norwegian women married in the period 1968-1970. Such a large data set allows us to include several covariates that are assumed to be important determinants, and gives the opportunity to apply flexible functional forms for the hazard.

It appears that even with 14 covariates included in the model, there is still a considerable amount of unexplained variation in divorce intensity. Applying two support points we find that there is a much higher divorce intensity in one group than in the other. This difference in intensity appears to be large compared to the effects of observed variables. However, the difference in intensity as well as the size of the high- and low-risk groups depend, of course, on the model specification. Imposing a multiplicative structure on the hazards, we find that about 14 per cent of the population are members of the high-risk group, where the divorce intensity is 22 times higher than that for the remaining 86 per cent of the population. In other words, one part of the population are very divorce prone, while the majority have characteristics that make marital break-up quite unlikely. We are apparently fairly close to a socalled "mover-stayer" situation. If we, however, introduce group-specific baseline hazards, we find that 29 per cent of the population are divorce-prone. The intensity for the high-risk group is considerably lower than that found in the model with equal baseline hazards.

Another result that we consider important is that within the low-risk group the divorce intensity rises with marital duration. Within the high-risk group, however, the intensity declines after about 6 years of marriage. When we assume equal baseline hazards, the profile of the low-risk group dominates, so that there is a monotone increase both within the low- and high-risk group.

The finding that the divorce propensity tends to increase steadily with marital duration seems to be in contrast with the theories contending that investment in marital specific capital tends to depress divorce rates with increasing duration. On the contrary, the results may indicate that the spouses gradually "wear each other out", and that the usually observed hazard which emerges in standard models, to a large extent is due to a selection mechanism, where the divorce-prone are thinned out during the marital life course.

Furthermore, it appears that the covariate effects estimates are rather insensitive to the omission of unobservables. Truly, all effects are somewhat larger in the Heckman-Singer model than in a standard model, but the relative importance of the covariates and the ranking of the importance of the categories is unchanged.

Our hypothesis that the proportion of divorce-prone might be larger among women with a premarital birth than among those who conceived their first child in a marital union, was not confirmed empirically. In both these groups there are about 10 per cent high-risk women, and among the high-risk women first birth timing has virtually the same effect as we found in standard models where unobserved heterogeneity was not taken into account. In other words, even when we include several explanatory variables and control for unobserved heterogeneity there remains considerable variations in divorce intensity between women with a premarital birth and those with a first child conceived in a marital union.

The results presented in this paper demonstrates that the inclusion of unobserved heterogeneity by means of the Heckman-Singer approach seems to be a promising avenue for further research on marital stability - in the sense that new empirical evidence which could challenge existing dissolution theories, might be produced.

REFERENCES

Becker, G.S., E.M. Landes and R.T. Michael 1977. An economic analysis of marital instability. Journal of Political Economy 85 (6) : 1141-1187.

Billy, J.O.G., N.S. Landale and S.T. McLaughlin 1986. The effect of marital status at first birth on marital dissolution among adolescent mothers. Demography 23 (3) : 329-349.

Bumpass, L.L. and J.A. Sweet 1972. Differentials in marital stability : 1970. American Sociological Review 37: 754-766.

Castro, T. and L. Bumpass 1987. Recent differentials in marital dissolution. CDE Working Paper 87-20. University of Wisconsin-Madison.

Cherlin, A.J. 1977. The net effect of children on marital dissolution. Demography 14 (3) : 265-272.

Christensen, H.T. and H.H. Meissner 1953. Studies in child spacing: Premarital pregnancy as a factor in divorce. American Sociological Review 18 : 641-644

Furstenberg, F.F. Jr. 1979. Premarital pregnancy and marital instability. In G. Levinger and O.C. Moles (eds): Divorce and Separation. Basic Books. New York.

Harris, J. 1982. Prenatal medical care and infant mortality. In V. Fuchs (ed): Economic Aspects of Health. University of Chicago Press. Chicago.

Heckman,J. and Willis 1977. A beta logistic model for the analysis of sequential labor force participation of married women. Journal of Political Economy 85 : 27-58.

Heckman,J. and B. Singer 1982. The identification problem in econometric models for duration data. In H. Hildenbrand (ed.): Advances in Econometrics. Cambridge University Press.

Heckman,J. and B. Singer 1984. A method for minimizing the impact of distributional assumptions in econometric models for duration data. Econometrica 52 (2).

Hoem, B. and J.M. Hoem 1988. Dissolution in Sweden: The break-up of conjugal unions to Swedish women born 1936-60. Stockholm Research Reports in Demography 45. University of Stockholm.

Kravdal, Ø. 1988. The impact of first birth timing on divorce: New evidence from a longitudinal analysis based on the Central Population Register of Norway. To appear in European Journal of Population.

Kravdal Ø. and T. Noack 1988. Skilsmisser i Norge 1965-1985. En demografisk analyse. (in Norwegian). Rapporter 88/6. Central Bureau of Statistics. Oslo-Kongsvinger.

McLaughlin, S.D, W. Grady, J. Billy, N. Landale and L. Winges 1986. The effect of the sequencing of marriage and the first birth during adolescence. Family Planning Perspectives 18 (1) : 12-18.

Menken J., J. Trussell, D. Stempel and O. Babakol 1981. Proportional hazard life table models: an illustrative example of socio-demographic influences on marriage dissolution in the United States. Demography 18 (2) : 181-200.

Montgomery, M. 1988. Household formation and home ownership in France. Presented at IUSSP Seminar on Event History Analysis, Paris, March 1988.

Morgan, S.P. and R.R. Rindfuss 1985. Marital disruption: structural and temporal dimensions. American Journal of Sociology 90 (5) : 1055-1077.

O'Connell, M. and C. Rogers 1984. Out-of-wedlock births, premarital pregnancies and their effect on family formation and dissolution. Family Planning Perspectives 16 (4) : 157-162.

Teachman, J.D. 1983. Early marriage, premarital fertility and marital dissolution. Journal of Family Issues 4 81) : 105-126.

Trussell, J. and T. Richards 1985. Correcting for unmeasured heterogeneity in hazard models using the Heckman-Singer procedure. In N. Tuma (ed): Sociological Methodology 1985. Jossey-Base: San Fransisco.

Trussell, J., G. Rodriguez and B. Vaughan 1988. Union dissolution in Sweden. Presented at IUSSP Seminar on Event History Analysis, Paris, March 1988.

ISSUED IN THE SERIES DISCUSSION PAPER

No.20 K. H. Alfsen, D. A. Hanson, and S. Glomsrød: Direct and Indirect Effects of reducing SO_2 Emissions: Experimental Calculations of the MSG-4E Model.

No.21 J. K. Dagsvik: Econometric Analysis of Labor Supply in a Life Cycle Context with Uncertainty.

No.22 K. A. Brekke, E. Gjelsvik, B. H. Vatne: A Dynamic Supply Side Game Applied to the European Gas Market.

No.23 S. Bartlett, J. K. Dagsvik, Ø. Olsen and S. Strøm: Fuel Choice and the Demand for Natural Gas in Western European Households.

No.24 J. K. Dagsvik and R. Aaberge: Stochastic Properties and Functional Forms in Life Cycle Models for Transitions into and out of Employment.

No.25 T. J. Klette: Taxing or Subsidising an Exporting Industry.

No.26 K. J. Berger, O. Bjerkholt and Ø. Olsen: What are the Options for non-OPEC Producing Countries.

No.27 A. Aaheim: Depletion of Large Gas Fields with Thin Oil Layers and Uncertain Stocks.

No.28 J. K. Dagsvik: A Modification of Heckman's Two Stage Estimation Procedure that is Applicable when the Budget Set is Convex.

No.29 K. Berger, Å. Cappelen and I. Svendsen: Investment Booms in an Oil Economy - The Norwegian Case.

No.30 A. Rygh Swensen: Estimating Change in a Proportion by Combining Measurements from a True and a Fallible Classifier.

No.31 J.K. Dagsvik: The Continuous Generalized Extreme Value Model with Special Reference to Static Models of Labor Supply.

No.32 K. Berger, M. Hoel, S. Holden and Ø. Olsen: The Oil Market as an Oligopoly.

No.33 I.A.K. Anderson, J.K. Dagsvik, S. Strøm and T. Wennemo: Non-Convex Budget Set, Hours Restrictions and Labor Supply in Sweden.

No.34 E. Holmøy and Ø. Olsen: A Note on Myopic Decision Rules in the Neoclassical Theory of Producer Behaviour, 1988.

No.35 E. Biørn and H. Olsen: Production - Demand Adjustment in Norwegian Manufacturing: A Quarterly Error Correction Model, 1988.

No.36 J. K. Dagsvik and S. Strøm: A Labor Supply Model for Married Couples with Non-Convex Budget Sets and Latent Rationing, 1988.

No.37 T. Skoglund and A. Stokka: Problems of Linking Single-Region and Multiregional Economic Models, 1988.

No.38 T. J. Klette: The Norwegian Aluminium industry, Electricity prices and Welfare,1988

No.39 I. Aslaksen, O. Bjerkholt and K. A. Brekke: Optimal Sequencing of Hydroelectric and Thermal Power Generation under Energy Price Uncertainty and Demand Fluctuations, 1988.

No.40 O. Bjerkholt and K.A. Brekke: Optimal Starting and Stopping Rules for Resource Depletion when Price is Exogenous and Stochastic, 1988.

No.41 J. Aasness, E. Biørn and T. Skjerpen: Engel Functions, Panel Data and Latent Variables, 1988.

No.42 R. Aaberge, Ø. Kravdal and T. Wennemo: Unobserved Heterogeneity in Models of Marriage Dissolution.