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

MPIDR WORKING PAPER WP 2007-020
MAY 2007 (REVISED JANUARY 2008)

# Early traces of the Second Demographic Transition in Bulgaria: A joint analysis of marital and non-marital union formation

Jan M. Hoem (hoem@demogr.mpg.de)
Dora Kostova (kostova@demogr.mpg.de)

Max-Planck-Institut für demografische Forschung
Max Planck Institute for Demographic Research

Direktoren: Prof. James W. Vaupel · Prof. Jan M. Hoem

**Prof. Dr. Jan M. Hoem**

Tel.: +49 (381) 2081-160
Fax: +49 (381) 2081-460
hoem@demogr.mpg.de

JMH/DK, 18 December 2007

**Early traces of the Second Demographic Transition in Bulgaria:**
**A joint analysis of marital and non-marital union formation**

by Jan M. Hoem and Dora Kostova,
Max Planck Institute of Demographic Research, Rostock

**Abstract**: In this paper, we explore trends in entry into the first conjugal union among young women in Bulgaria since 1960, based on data from the national Gender and Generations Survey conducted in 2004. We use an extension of piecewise-constant hazard regression to analyze jointly the transition into a cohabitational union and directly into marriage. This extension allows us to compare relative risks across the two competing transitions, a comparison which is infeasible otherwise. We find, among many other things, that on average women in the Roma sub-population have a considerably higher tendency to start a cohabitation than to start a marriage at each age, *ceteris paribus.* We also find that a pregnancy leads to a dramatic increase in the rate of both kinds of union formation; the increase is by a factor of almost 20 for marriage formation and "only" about half as much for entry into cohabitation, again *ceteris paribus*. The standardized marriage rate has fallen dramatically since the early 1980s; the corresponding rate of entry into cohabitation increased since the early 1960s already and (surprisingly) has fallen moderately toward the end of our period. Cohabitation also tends to last progressively longer in more recent periods. Our findings suggest that in Bulgaria at least some manifestations of the Second Demographic Transition can be detected in the 1980s already, which is earlier than others have claimed for this country before.

***Keywords***: marriage; cohabitation; first union; joint analysis of hazards; Second Demographic Transition; Bulgaria.

## 1. Introduction

As fertility dropped in Central and Eastern Europe after the fall of communism, demographers generally assumed that the new freedoms would induce demographic behavior to adapt to that of Western Europe. Some of the adjustment would be changes in family forms and living arrangements; the expectation was that previously unconventional arrangements, particularly non-marital unions, would take over territory from standard marriage. This change in behavior had already occurred extensively in Western Europe, and a general theory of a Second Demographic Transition had been formulated to cover changes assumed to reflect more individualistic attitudes to life, including attitudes to forms of union formation. Much less attention has been given so far to these issues for Central and Eastern Europe. It seems to have been an open question whether the Second Demographic Transition (SDT) has reached Bulgaria at all, and if it has, whether it all started with the fall of communism. We show that some SDT manifestations actually were well underway by 1990, and that some features go back to the early 1960s.

The vast literature on non-marital cohabitation in the West started with some work about such behavior in Scandinavia (Trost 1978, Trost and Lewin 1978, Finnäs and Hoem 1978, Brunborg 1979, Hoem 1980, Blanc 1983, Hoem and Rennermalm 1985, Bernhardt and Hoem 1985). Papers about other Western countries quickly followed (Brown and Kiernan 1981, Haskey et al. 1986, 1989, Liefbroer 1991, Manting 1996, and many others), and new contributions continue to appear to this day (Reneflot 2006, Steele et al. 2006, Mulder et al. 2006, Liefbroer and Dourlein 2006, Schröder 2006, Musick 2007, and many more). The concept of the Second Demographic Transition was invented by Lesthaeghe and van de Kaa (1986), both of whom have contributed much to its later development also (van de Kaa 1987, 1994, Lesthaeghe 1995, Lesthaeghe and Neidert 2006; see also Raley 2001), and a new summary by Lesthaeghe (2007) will soon appear. There is now an emerging literature about non-marital cohabitation in what used to be the Eastern Block in Europe (Rychtaříková 1994, Kantorova 2004, Spéder 2005, 2006, Koytcheva 2006, Kostova 2007). Some of these refer to signs that the Second Demographic Transition may (or may not) also extend to specific countries in Central and Eastern Europe (Rychtaříková 1999, Zakharov 2005, Gerber and Berman 2005, Sobotka 2008). Koytcheva (2006) has already carried out an extensive similar study based on data collected in the Bulgarian census of 2001 and the MPIDR Social Capital Survey in that country in 2002. She found much the same patterns as we do over the period that she studied (namely the fifteen years 1988-2002, see her Figures 4.20, 4.23, and 4.27), but she claimed that behavioral changes in Bulgaria after 1990 represent a return to pre-communist patters more than a manifestation of the SDT. As we will soon make clear, we reach a different interpretation after a study of data that were unavailable to her; we claim to see early traces of the SDT.

In the present paper we provide a new analysis of changes in union formation over forty-five calendar years (namely, 1960-2004) in Bulgaria, based on data from the first wave of the Bulgarian Gender and Generations Survey (GGS), collected in 2004. The special contributions of our paper are (i) that we have a much longer time perspective than anyone before us, (ii) that we study competing transitions (namely entry into marital and non-marital unions) jointly, in a manner that allows for a direct comparison across the transitions of the relative risks connected with each covariate, and (iii) that we also pursue the conversion of consensual unions into marital unions over the period 1970-2004. The conventional approach is to use hazard-regression analysis to study each competing risk separately, which makes a joint analysis across the two transitions infeasible. With our procedure we can compare the baseline hazards directly and see differential effect of process time (which in our application is age attained) on each transition, standardized for the other covariates. We can also study trends over calendar time across the two competing risks and thus provide a direct check as to whether a drop in marriage is compensated by a counterpart increase in cohabitation, as would be expected under some formulations of SDT theory. Our method is very easy to apply and is close to perfect for initial exploration.

Our procedure reveals that rates of marriage formation have dropped since the 1970s, and strongly since the 1980s. Such a drop would be expected if an SDT were in progress. Rates of entry into a consensual union have increased from the 1960s already, much earlier than what has been expected so far. After the fall of communism, Bulgarian cohabitational entry risks have declined, however, contrary to the continued increase expected by SDT theory. At the same time, rates of conversion of consensual into marital unions have declined considerably, particularly during the first year of the union. Since there has hardly been any disruption of consensual unions at this stage, this means that consensual unions have been maintained progressively longer, a feature that fits well with the SDT description. In summary, we find several behavioral traces of an ongoing SDT in Bulgaria, but a broader picture (like the one we have provided) is needed to see this; it is not enough just to study trends in entry rates for marital and non-marital unions for a rather brief recent period. It is hard to pinpoint a time at which the SDT uniformly started, not that the initiators of the SDT description ever claimed that any uniformity in the timing of SDT elements could be predicted.

The theory of the Second Demographic Transition is fired by fertility developments, and it also covers rates disruption of marital and nonmarital unions, which is predicts will rise. Much is known already about such aspects of behavior in Bulgaria as well as in the rest of Europe, see, e.g., Frejka et al. (2008). By focusing exclusively on union formation, we manage to illuminate an aspect of human behavior that has received much too little attention so far.

## 2. Data

The Gender and Generations Surveys (GGS) is a program of national demographic panel surveys in most countries of Europe and beyond (Vikat et al. 2007). In the first round of each survey complete retrospective childbearing histories as well as complete histories of union formation and disruption are collected. The surveys also have a rudimentary educational history and some information about occupational activity; the plan is to collect more complete educational and employment histories in the second panel wave. We use data for female respondents in the first wave of the Bulgarian GGS; plans are in hand to analyze the second-wave data soon.

The event we study is self-reported first-marriage formation, alternatively the start of a first consensual union. Anecdotal evidence from Bulgaria suggests that a couple will often move together as soon as they are engaged to be married, most likely in the house or apartment of one set of parents. We assume that such a move would normally be recorded as pre-marital cohabitation in our data; unfortunately we cannot consistently distinguish such cohabitations from others. This old-standing tradition seems to have been most common in rural areas, and it seems to have been socially accepted in socialist times as well. There is a general inaccuracy of reporting of starting dates of non-marital unions in Europe (like elsewhere), so we count a union as a direct marriage even if the respondent reported the start of a non-marital union in the same month as the marriage, or one month earlier. The conversion of a consensual union into a registered marriage in the months following union formation turns out to be essential for an understanding of SDT behavior in this country. Observing such conversions allows us to see marriage formation as a stepwise process rather than as a single distinct event for many of our respondents.

In our hazard analysis of initial union formation we use age attained as process time. We have computed it from the date of birth as reported in the interview and grouped it into ages 15-16, 17-18, 19-20, …, 29-30, and 31-34. Other covariates available to us are

- self-reported ethnicity (ethnic Bulgarian, Turk, Roma, other),
- parity-and-pregnancy status (childless not pregnant, childless and pregnant, of parity 1 or more; time-varying covariate),
- the respondent's educational attainment by the time of the interview, and month of attainment, and
- current calendar period at any point in process time; we have used it as a time-varying covariate (grouped into the periods 1960-64, 1965-69, … , 1995-99, 2000-2004).

A small number of records (255) had missing information or were of unacceptable quality otherwise, so we deleted them. The data contain 5610 usable records for women born in 1945 or later, corresponding to a general response rate of 69% (Atanassov et al, 2005). With the ethnicities that our respondents reported, we have recorded 25194 person years of exposure for ethnic Bulgarians, 2005 person years for Turks, 959 person years for Roma, and 200 person years for other ethnic groups. Roma may potentially be particularly problematic, partly because members of any such minority may be reluctant to report themselves as anything else than Bulgarians, and partly because the Roma may report marital and non-marital unions differently from other ethnic groups. The intention was that respondents should report a marriage only if it had been registered by the Bulgarian authorities, but we suspect that the Roma may have reported unions as marriages if they were regarded as such in their own community irrespective of formal status according to the register. Conversely, what they themselves regard as a marriage may have been recorded as cohabitation in our data set. It is a philosophical question which perception of marriage and cohabitation is "right". In any case instead of deleting self-reported ethnicity from our analyses (say) we have retained it to avoid needless compositional effects caused by this covariate.

We have included the parity-and-pregnancy status covariate because we suspect that union-formation behavior may be strongly dependent on the corresponding status. For a non-partnered woman, the arrival of a pregnancy should be highly motivational for union formation.

We also believe that educational attainment may be a determinant of union formation, perhaps with different patterns of entry for marital and cohabitational unions. We do not venture any pre-conceived opinion about how this may work. Highly educated women may have acquired more conventional attitudes to family forms; after all we must expect them to live in a more bourgeois environment. But conversely, they are often seen as the forerunners of new behavioral developments, and if they are, they could be more prone than others to choose non-marital over marital union formation as a first step. In any case, we want to retain the possibility of discovering whatever patterns we can find in the Bulgarian population.

For such a purpose we would also strongly prefer to know a respondent's actual educational attainment at each stage in her life. We do not trust any argument that posits that a woman will behave now according to her educational attainment later in life. A person's current behavior may be influenced by the goals she has set for herself, but we find it hard to accept that the prospect of later attainment will override the effect of current attainment. It is always problematic in an event-history analysis to condition of future outcomes, and there are enough investigations available that warn against such anticipatory analysis.

Unfortunately our data do not contain proper educational histories, so we have had to resort to imputing a time-varying educational covariate from such educational information as is available to us. Using a method suggested by Hoem and Kreyenfeld (2006), we have imputed an educational covariate by "assuming" that the respondent was in education until the date of attainment reported in the interview, and continuously out of education (with the reported level attained) between the date of attainment and the interview. This also "assumes" that the effect of being in education on the intensities of union formation is the same for all levels at which one may take education. We have grouped educational levels as follows:

- low education means primary school, basic and incomplete secondary school;

- middle education means completed secondary school (with final exams),
- high educations means every education higher than secondary, including all levels of university education.

We dislike the fact that this imputed educational covariate is strongly anticipatory, i.e., it features conditioning on the future at each process time before the interview, namely conditioning on not having returned to take more education than the one imputed. We are quite suspicious of what the regression coefficients of such covariates tell us about real behavior and to not want to place much confidence in results based on them. There may be a saving grace in findings (Zabel 2007) that suggest that their biases have been be rather limited in experiments where such constructs are compared with their genuine counterparts, at least in educational systems where there are limited dynamics of return to the educational system once education has been completed. In any case, we have not let educational attainment play any prominent role in this account; it only appears as a (time-varying) control variable.

Since we are interested in the progress of the Second Demographic Transition in real time, we have included calendar period as another time-varying covariate and have avoided using a cohort variable, which is much harder to interpret for our purposes. We have partitioned calendar time into convenient five-year periods, carefully observing that the fall of communism coincides with the intersection of two of our period intervals.

In addition to the variables discussed so far, we have some covariates meant to reflect the respondent’s socialization as she grew up, namely,

- the character of the location where she grew up (urban, rural),
- whether her parents lived together during the greater part of her childhood,
- mother’s educational attainment,
- father’s educational attainment, and
- the number of siblings in the parental home.

These variables are conveniently available to us, and they have of course been included in the data collection because they are commonly assumed to be useful for many analyses, including ours (Vikat et al., 2007). In particular, the educational attainment of each parents is included for reasons similar to those that led us to seek for a representation of the respondent’s own educational attainment. The multidimensional specification of the educational variable should also enable us to study which parent’s attainment is the more important, and whether the respondent’s own attainment (properly measured) overrides the parental influence.

## 3. Method of analysis

In our analyses we have used multiplicative intensity-regression (or proportional-hazard) models with a piecewise-constant baseline hazard. As such this method is well known (see, e.g., Hoem, 1971, 1976, for early descriptions in similar terms), so we only give a rudimentary account here.

Suppose for simplicity that beside age (which we call Factor A) there are only two covariates, namely a fixed factor B and a time-varying factor C, for which we have calendar period primarily in mind. Also suppose that initially we only work with a single decrement, say marriage formation. The effect of age is assumed constant over some given age intervals, and the effect of calendar time is taken as constant over suitably chosen groups of calendar years, called periods. For an individual in age group *i* with level *j* on Factor B we take her transition intensity in period *k* to have the format

$$\mu_{ijk} = a_i b_j c_k \, . \qquad (1)$$

The constants (parameters) $a_1, a_2$, and so on are the effects of age, the parameters $b_1$, $b_2$, and so on are the effects of Factor B, and the constants $c_1$, $c_2$ and so on are the effects of Factor C (calendar time). For normalization, we set $b_{j_0}$ =1 and $c_{k_0}$ =1 for some selected baseline levels $j_0$ and $k_0$. Each parameter $b_j$ then appears as a "relative risk", in that $b_j = \mu_{ijk} / \mu_{ij_0k}$, and the $c_k$ have a similar straightforward interpretation. The parameters $a_i = \mu_{ij_0k_0}$ constitute the baseline hazard.

For all individuals taken together, the likelihood is

$$\Lambda = \exp\{-\sum_i \sum_j \sum_k R_{ijk} a_i b_j c_k\} \prod_i \prod_j \prod_k \{a_i b_j c_k\}^{D_{ijk}} \, ,$$

where (here) the occurrences $D_{ijk}$ are the number of marriages for the factor combination (*i,j,k*) and the exposure $R_{ijk}$ is the corresponding number of person-months observed under risk of marriage. We get the maximum-likelihood estimators $\hat{a}_i$, $\hat{b}_j$, and $\hat{c}_k$ for our parameters by maximizing $\Lambda$ with respect to the three sets of parameters. This is most easily done iteratively by means of some computer program (of which there now exist many; we used one called EvHA that has been developed by colleagues at the MPIDR). Such a program uses the matrix $\boldsymbol{D} = \{D_{ijk}\}$ of occurrences and the matrix $\boldsymbol{R} = \{R_{ijk}\}$ of exposures as input. (Some programs produce these matrices internally from more rudimentary input.) At the cost only of some notational complication these ideas are extended trivially to more than two covariates and to situations where covariates operate in interaction.

**< ---**in about here: **Table 1. Transition to first direct marriage and first cohabitation, Bulgarian women born in 1945-1986, GGS data from 2004.** (Table 1 can now be found just before the diagrams at the end of the text.) **--->**

As a first step we have used a simple proportional-risks model of union formation for marital and non-marital unions separately and have listed the estimated relative risks of union formation in the columns of Table 1. (We display the outcome of more sophisticated analyses later.) Since we have used age attained as our "Factor A", the first panel in Table 1 contains the baseline intensity, given as events per 1000 person-months. The final panel in the table gives the trends over time in the entry risks, and Figure 1 contains a corresponding plot with a finer grid of calendar periods. We see that marriage rates have declined since the early 1980s and that rates of entry into cohabitation have increased after the 1960s but that they actually have declined since the fall of communism. (We return to the latter feature below.) What we cannot see from Figure 1 is whether entry into marriage or into a non-marital union was the more common in any period. Both curves are anchored at the level 1 in 1960-64, and all points on each curve are risks relative to that level. This does not allow for comparisons across the types of union formation.

This can be improved by combining the data for the two risks and analyzing them jointly instead of separately, by which one can get the transition rate at a factor level on one intensity relative to the corresponding factor level on the other intensity. To see how this is done, first extend the above mathematics by introducing an extra subscript, say $\ell$, for the cause of decrement (type of union formed), and get the formula

$$\mu_{ijk\ell} = a_{i\ell} b_{j\ell} c_{k\ell} \tag{2}$$

for the intensity of decrement $\ell$, with $\ell$=1 for entry into a non-marital union, say, and $\ell$=2 for entry into marriage. Corresponding to the two types of decrement there will be two occurrence matrices, $\boldsymbol{D}_1 = \{D_{ijk1}\}$ and $\boldsymbol{D}_2 = \{D_{ijk2}\}$, but there will only be one matrix of exposures, namely the same matrix $\boldsymbol{R}$ as before. Since we operate with two competing risks (namely risks of marital and non-marital union formation, respectively), an individual of course has the same months-exposed-to-risk for both types of transitions, and this holds also in the aggregate. We introduce the combined occurrence and exposures matrices

$$\boldsymbol{D}_* = \begin{pmatrix} \boldsymbol{D}_1 \\ \boldsymbol{D}_2 \end{pmatrix} \text{ and } \boldsymbol{R}_* = \begin{pmatrix} \boldsymbol{R} \\ \boldsymbol{R} \end{pmatrix}.$$

Note that the exposure matrix $\boldsymbol{R}$ appears twice in the definition of $\boldsymbol{R}_*$. Using the combined occurrence and exposure matrices corresponds formally to entering the type of decrement as an extra factor, say Factor D, in the analysis. (Extension to several competing risks is trivial.) A formula like the one for $\mu_{ijk\ell}$ above would then correspond to operating with Factor D in a two-way interaction with each of the Factors A, B, and C. The empirical result would be the same as in separate analyses of the two decrements, with some modification, as follows.

To attain parameter identification when the intensities are analyzed separately, one would let $b_{j_0 1}=1$ and $b_{j_0 2}=1$, presumably for the same baseline B-level $j_0$, and similarly $c_{k_0 1}=c_{k_0 2}=1$ for some C-level $k_0$. For each decrement $\ell$ the parameters $b_{j\ell}$ and $c_{k\ell}$ would appear as relative risks in the corresponding regression.

In joint analysis, it is enough to use one of these normalizations for each factor. Suppose that we let $b_{j_0 1}=1$ and $c_{k_0 1}=1$. With the specification of $\mu_{ijk\ell}$ in (2), our computer program will then produce maximum-likelihood estimates for all the parameters. As we will show in an appendix, the parameters unfortunately will not automatically have an interpretation as relative risks, the way they had in the simpler situation (1). They will have such an interpretation under special circumstances, however.

For illustration now suppose first that

$b_{j1}=b_{j2}$ for all *j* and that also $c_{k1}=c_{k2}$ for all *k*. (3)

In other words, suppose that Factors B and C have the same effects on both decrements. For the moment we do not make a similar assumption for Factor A, so at present we allow each transition intensity to have its own age profile, defined by the two sets of parameters $\{a_{i1}\}$ and $\{a_{i2}\}$. Operating with a model where Factor D appears in interaction with Factor A but not in interaction with Factors B and C would then produce estimates $\{\hat{a}_{i1}\}$, $\{\hat{a}_{i2}\}$, $\{\hat{b}_j\}$, and $\{\hat{c}_k\}$ of the various effects, and we would have exploited the fact that Factors B and C had the same effects for both decrements. In this situation, <u>first</u>

$$\frac{\mu_{ijk\ell}}{\mu_{ij_0k_0\ell}}=b_j c_k\,,$$

which means that the $b_j$ and $c_k$ appear as relative risks in comparisons between intensities for (any) one of the transitions, and <u>second</u>,

$$\frac{\mu_{ijk1}}{\mu_{ijk2}}=\frac{a_{i1}}{a_{i2}}\,,$$

which means that at each age *i* the baseline risk for one transition can be compared direct with the corresponding risk for the other transition. In other words, when Factors B and C have pairwise equal effects on the two risks (as formalized in (3)), then we can make a direct comparison between the two baseline risks. This concludes our discussion of relation (3).

Now suppose alternatively that

$a_{i1} = a_{i2}$ for all age intervals *i*, that also $b_{j1} = b_{j2}$ for all *k*, (4)

but that the $c_{k\ell}$ may depend on the decrement $\ell$ as well as on the level *k* on factor C. To attain identifiable parameters we let $c_{k_0 1} = 1$. We do not need to also let $c_{k_0 2} = 1$ and do not make this assumption at present, as we would do in a separate analysis of the two transition risks. Simple substitution of the parameters into the intensity formulas gives $c_{k\ell} = \mu_{ijk\ell} / \mu_{ijk_0\ell}$, which shows that under these conditions $c_{k\ell}$ can be interpreted as the risk of entry into a union of type $\ell$ when Factor C has level *k* (for any choice of levels *i* and *j* on Factors A and B), relative to the corresponding risk on entry into the same type of union when Factor C has level $k_0$. Similarly, $c_{k1} / c_{k2}$ becomes the risk of entry into a union of type 1 in period *k*, relative to the risk of entry into a union of type 2 in the same period (e.g., the risk of starting a consensual union relative to the risk of starting a marriage), *ceteris paribus*, because $c_{k1} / c_{k2} = \mu_{ijk1} / \mu_{ijk2}$. Thus comparisons across decrements are in order. This kind of relative risk only appears when one uses the technique we have described here. We have found these ideas useful in our empirical analyses (see below).

If the restrictions in (4) on the $\{a_{i\ell}\}$ and the $\{b_{j\ell}\}$ hold only approximately, then joint (comparative) analysis of the two decrements based on the $\{c_{k\ell}\}$ can still be used to produce compact descriptions of patterns in the data. In normal demographic parlance we would say that estimates $\{\hat{c}_{k\ell}\}$ of the effects of Factor C have been standardized with respect to Factors A and B. This is a case of <u>indirect standardization</u>, as explained by Hoem (1987)

Note that the use of the $\{c_{k\ell}\}$ for such compact descriptions does not really use the intensity model $\mu_{ijk\ell} = a_i b_j c_{k\ell}$ implied by (4), where the effects of Factors A and B appear in a multiplicative manner. It is enough to have an intensity formula of the form $\mu_{ijk\ell} = a_{ij} c_{k\ell}$, i.e., Factors A and B may appear in complete interaction, so long as their combined effects $\{a_{ij}\}$ are the same for both decrements. Also, mathematically speaking, Factors B and C appear symmetrically in the arguments above, and it does not really matter that one of them is a fixed covariate and the other time-varying.

As we have explained, the trick above is to introduce the cause of decrement as an extra "factor" and to operate with it in one or more interactions with the "ordinary" factors, which one may also interact with each other. The procedure has been described before in different words by

Gomez de Leon and Potter (1989), by Liefbroer (1991), and by Pierce and Preston (1993), but we have spelt it out for users of piecewise-constant proportional hazards and have clarified conditions under which the usual regression parameters can be interpreted as relative risks. In our appendix we discuss what happens if such conditions are not fulfilled. Judging from the scarcity of papers in demography and other fields that make use of this simple and very useful procedure for the analysis of competing risks, an explanation like ours should have some merit.

## 4. Results

### 4.1. Covariates reflecting the respondent's socialization

We have not found much of an interesting pattern in the effects of the five covariates we had selected to represent the respondent's early socialization (the character of the location where the respondent mostly grew up, whether the respondent's parents stayed together while she grew up, her mother's and father's educational attainment, and her number of siblings). The only systematic feature we can find is some decrease in the intensity of non-marital-union formation as the educational level attained by the respondent's mother increases and a mild reflection of the same pattern for the father's educational level, features which could possibly represent increased social ambitions as the parents' education improved. (The effect can be seen in Table 1 but is not otherwise reported *in extenso* here.) In support of this interpretation we would expect a converse pattern in the intensities of marriage formation, but we see no such features in our data.

The lack of effect of early socialization may have some independent interest, for it represents a break with the theory that led to their inclusion in the first place. This is not important for our main line of argument, however, and we do not to pursue the issue here.

### 4.2. Joint analysis 1: Ethnicity

As a first example of the usefulness of a joint analysis of the two types of union formation we have interacted our ethnicity variable with the cause D of decrement (Factor D). This allows us to pick up any differential effects of ethnicity on the two transition risks. Since there can be similar differential effects of other covariates, we have interacted Factor D with each of them also, much as in our original specification of $\mu_{ijk\ell}$. The ethnicity effects estimated by this exercise are given in Table 2. (These figures should be interpreted as averages standardized for the other covariates.) We see that the Roma have an intensity of starting a non-marital union that is much higher than that of marrying, which is not particularly surprising, but we find it remarkable that we make a similar finding for ethnic Bulgarians as well.

**Table 2. Transition to first direct marriage and first cohabitation, Bulgarian women 1960-2004, by ethnicity. GGS data from 2004**

| Ethnicity | Entry into cohabitation | Entry into marriage | Risk of entry into cohabitation relative to that of marriage |
|---|---|---|---|
| Bulgarian | 1 | 0.75 | 1.33 |
| Turk | 0.91 | 0.92 | 0.99 |
| Roma | 1.11 | 0.77 | 1.44 |
| (Others) | (1.41) | (0.40) | |

### 4.3. Joint analysis 2: The age profile

If we take a three-way interaction between a covariate V, Factor D, and our parity-and-pregnancy variable, we get the effect of V on the two transition risks separately for childless non-pregnant women, childless pregnant women, and mothers. If the covariate V is age, we can plot the age profile of the risks for childless non-pregnant women, as in Figure 2. It is interesting (though not particularly surprising) to see that the profile is the younger for entry into cohabitation.

We have recorded 26958 person years of exposure before any first-union formation for non-pregnant nullipara, 266 person years for pregnant nullipara, and 1133 person years for mothers. Thus the group of women who have no children and are not pregnant have by far the largest exposures, and we refrain from displaying the age profiles for the two much smaller groups which only have some five per cent of the exposures taken together. Concentrating on non-pregnant childless women corresponds to censuring the records for the never-partnered women at the occurrence of a pregnancy.

### 4.4. Joint analysis 3: Time trends in the risks

A three-way interaction between calendar period, Factor D, and the parity-and-pregnancy variable, plotted again for the childless non-pregnant women only, gives a diagram like Figure 3.

We find several features of this diagram remarkable.

-- First, cohabitation has traditions going back at least to the 1960s, as we saw in Figure 1 already. The corresponding entry risk doubled until about 1990, but it has fallen thereafter.

-- Second, the marriage risk was more than twice as high as the entry risk for consensual unions until the early 1980s. (This is where we see the usefulness of comparisons across the transitions.) The risk then declined considerably and has been much the smaller at least since the fall of communism. (It is now only one-fifth of its peak value in the early 1970s.) Its wobbles in the 1960s and 1970 seem to reflect changing family policies under the communist regime.

During periods when abortion was forbidden, marriage must have been a way to secure the legitimacy of an unplanned pregnancy.

In her extensive earlier study on Bulgaria, Koytcheva (2006, e.g. Figure 4.23 and Table B17) also used parity-and-pregnancy status as a control variable and found a risk of entry into a consensual union that increased between 1988 and 2002, which is the period covered by her analysis. To our knowledge, other authors studying demographic behavior in Central and Eastern Europe have mostly been interested in other features of demographic behavior, usually fertility, but to the extend that comparisons are possible, they all report increasing rates of non-marital union formation, not relatively stable and ultimately decreasing rates as in what we have found. (See the national chapters in Frejka et al., 2008.) Any expectation that the fall of state socialism should have sped the progress of the Second Demographic Transition comes to shame in our Figure 3. So long as we confine ourselves to union formation we need to look for other developments. We now turn to trends in rates of conversion of consensual unions to marriage.

## 5. Conversion of consensual unions

We leave our concentration on first entry into consensual and marital unions and turn to what has happened to the consensual unions after they have been formed, following a previous initiative by Koytcheva (2006). Figure 4 contains the rates of conversion of such unions into marriage, by the duration of the cohabitation. (We could only do this meaningfully from around 1970.) After some experimentation we have grouped the curves into periods that highlight the quick entry into marriage before 1990 and the strong subsequent decline in the conversion rates as far as we can follow them. Since there are very few disruptions among these unions at the stage we observe, these features imply considerable lengthening of the period before marriage is contracted. Entry into consensual unions may have slacked off some after 1990, but the non-marital unions that have been formed, have been much more durable than before this crossroad. We interpret this as a clear manifestation of the Second Demographic Transition, boosted by the fall of state socialism.

## 6. Reflections

In this paper we have studied union formation in Bulgaria by means of a method that allows us to address the competing risks of entry into marital and non-marital unions simultaneously. The method is simple and efficient; it has enabled us to quickly outline the main patterns of first-union entry behavior. We have found a striking stability in entry into cohabitation after the early 1980s. Such unions have a long tradition in Bulgaria, perhaps mostly as a precursor of formal marriage, and non-marital partnership seems to keep its place in union-formation behavior to this day, in fact to have become more of a durable initial stage before marriage. By contrast, formal marriage has ceded ground all the time, at least at the initiation of a first union.

A piecewise-constant hazard regression analysis with categorical covariates is the simplest non-trivial type of event-history analysis, and our method is a simple extension of a procedure that has now become commonplace. There are precursors in the literature, but the joint analysis of competing risks is not often encountered; it is our hope that we shall contribute to making it more popular.

We believe that the basic trick of formally including the cause of decrement as an additional (fixed) covariate can be transferred to other methods of hazard regression, indeed this has been shown already by Andersen et al. (1993, p. 495) and Lunn and McNeill (1995) for a Cox model where the baseline hazard is nonparametric.

**Acknowledgements**

We are grateful to Manuela Schicka and Lesia Nedoluzhko for research assistance and to Per Kragh Andersen for providing some of the references to the statistical literature. A conversation with Ron Lesthaeghe inspired some of our reflections concerning the Second Demographic Transition. Comments by members of the journal's editorial staff were very helpful.

**References**

Andersen, P. K, Borgan, Ø., Gill, R. D., and Keiding, N. 1993. *Statistical Models Based on Counting Processes*. Springer-Verlag.

Atanassov, A., Zdravka Toneva, Z., Todorova, S., Kotseva, T., and Kaltchev, Y. 2005. *Final report on the organization and conducting of the study Generation and Gender Survey.* Bulgarian Academy of Sciences, Coordination Research Centre for Social Development and Social Eurointegration.

Bernhardt, E. and Hoem, B. 1985. Cohabitation and social background: Trends observed for Swedish women born between 1936 and 1960. *European Journal of Population* 1 (4); 375-395.

Blanc, A. *The impact of changing family patterns on reproductive behavior: Non-marital cohabitation and fertility in Sweden and Norway*. Princeton, Office of Population Research, doctoral dissertation.

Brown, A. and Kiernan, K. 1981. Cohabitation in Great Britain: Evidence from the General Household Survey. *Population Trends* 25 (1): 4-10.

Brunborg, H. 1979. Cohabitation without marriage in Norway. Central Bureau of Statistics of Norway, *Artikler* 55.

Finnäs, F. and Hoem, J.M. 1978. Cohort trends in cohabitational birth intervals in Denmark, 1975. *Scandinavian Population Studies* 5: 196-210.

Frejka, T. et al. 2008. Levels and trends of fertility in Europe. *Demographic Research*, Special Collection, to appear.

Gerber, T. P, and Berman, D. 2005. Economic crisis or Second Demographic Transition? Trends and correlates of union formation in Russia, 1985-2001. Paper presented to the Annual Meeting of the PAA in Philadelphia and later revised.

Goméz de León, J. and Potter, J.E. 1989. Modelling the inverse association between breastfeeding and contraceptive use. Popul.ation Studies 43, 69-93.

Haskey, J. and Coleman, D. 1986. Cohabitation before marriage: A comparison of information from marriage registration and the General Household Survey. *Population Trends* 43, 15-17.

Haskey, J., and Kiernan, K. 1989. Cohabitation in Great Britain – characteristics and estimated numbers of cohabiting partners. *Population Trends* 58, 23-32.

Hoem, J. M. 1971. Point estimation of forces of transition in demographic models. Journal of the Royal Statistical Society, Series 33 (2): 275-289.

Hoem, J. M. 1976. The statistical theory of demographic rates: A review of current developments (with discussion). Scandinavian Journal of Statistics 3 (4): 169-185.

Hoem, J. M. 1980. Starting age and subsequent birth intervals in cohabitational unions in current Danish cohorts, 1975. *Demography* 17 (3): 279-295.

Hoem, J. M. 1987. Statistical analysis of a multiplicative model and its application to the standardization of vital rates: A review. *International Statistical Review* 55, 119-152.

Hoem, J. M. and Kreyenfeld, M. (2006). Anticipatory analysis and its alternatives in life-course research. Part 1: The role of education in the study of first childbearing. *Demographic Research* 15 (16), 461-484.

Hoem, J. M. and Rennermalm, B. 1985. Modern family initiation in Sweden: Experience of women born between 1936 and 1960. *European Journal of Population* 1: 81-112.

van de Kaa, D J. 1987. Europe's Second Demographic Transition. Population Reference Bureau: *Population Bulletin* 42.

van de Kaa, D. 1994. The second demographic transition revisited: theories and expectations. Pp. 81-126 in *Population and Family in the Low Countries 1993: Late Fertility and Other Current Issues.* Beets, G. C. N., van den Brekel, J. C., Cliquet, R. L., Dooghe, G., and de Jong Gierveld, J. (Eds.): NIDI CBGS Publ., Xwets & Zeitlinger (Amsterdam).

Kantorová, V. 2004. *Family life transitions of young women in a changing society: first union formation and birth of first child in the Czech Republic, 1970-1997*. Doctoral dissertation, Charles University, Prague, and Université de Paris I, Pantheon, Sorbonne. Download from http://www.demogr.mpg.de/?/en/publications/dissertations.htm.

Kostova, D. 2007. The rise of cohabitation in Bulgaria: who are the forerunners of the new family model? Paper presented to the PAA meeting in New York, 30 March 2007.

Koytcheva, E. 2006. *Social-demographic differences of fertility and union formation in Bulgaria before and after the start of the societal transition.* Doctoral dissertation, University of Rostock, Germany. Download from http://www.demogr.mpg.de/?/en/publications/dissertations.htm.

Lesthaeghe, R. and van de Kaa, D J. 1986. Twe demografische transities? Pp 9-24 in *Bevolking: groei en krimp*. Van Loghum Slaterus (Deventer).

Lesthaeghe, R. 1995. The Second Demographic Transition in Western countries: an interpretation. Pp. 17-62 in *Gender and Family Change in Industrialized Countries*, K. O. Mason and A.-M. Jensen (Eds.): Clarendon (Oxford).

Lesthaeghe, R. J. and Neidert, L. 2006. The Second Demographic Transition in the United States: Exception or Textbook Example? *Population and Development Review* 32 (4): 669-698.

Liefbroer, A.C. 1991. The choice between a married or unmarried first union by young adults: a competing risk analysis. *European Journal of Population* 7 (3), 273-298.

Liefbroer, A.C. and Dourleijn, E. 2006. Unmarried cohabitation and union stability: Testing the role of diffusion using data from 16 European countries. *Demography* 43 (2), 203-221.

Lunn, M. and McNeil, D. (1995). Applying Cox regression to competing risks. *Biometrics* 51, 524-532.

Manting, D. 1996. The changing meaning of cohabitation and marriage. *European Sociological Review* 12 (1), 53-65.

Mulder, C.H., Clark, W.A., and Wagner, M. 2006. Resources, Living Arrangements and First Union Formation in the United States, the Netherlands and West Germany. *European Journal of Population* 22 (1): 3–35.

Musick, K. 2007. Cohabitation, nonmarital childbearing, and the marriage process. *Demographic Research* 16 (9): 249-286.

Philipov, D. and Jasilioniene, A. 2007. Union formation and fertility in Bulgaria and Russia: a life table description of recent trends. Download from http://www.demogr.mpg.de/papers/working/wp-2007-005.pdf.

Pierce, D. A. and Preston, D. L. 1993. Joint analysis of site-specific cancer risks for the atomic bomb survivors. *Radiation Research* 134 (2), 134-142.

Raley, K. R. 2001. Increasing fertility in cohabiting unions: Evidence for the second Demographic Transition in the United States? *Demography* 38 (1): 59-66.

Reneflot, A. 2006. A gender perspective on preferences for marriage among cohabiting couples. *Demographic Research* 15 (10): 311-328.

Rychtaříková, J. 1994. Les unions informelles en République Tchèque. *Acta Universitatis Carolinae, Geographica, Supplementum* : 71-85.

Rychtaříková, J. 1999. Is Eastern Europe experiencing a second demographic transition? *Acta Universitatis Carolinae, Geographica* 1: 19-44.

Schröder, C. 2006. Cohabitation in Italy: do parents matter? *Genus* 62 (3-4). 53-85.

Sobotka, T. 2008. The diverse faces of the Second Demographic Transition in Europe. Overview Chapter 6 in Frejka et al., 2008.

Spéder, Z. 2005. The rise of cohabitation as first union and some neglected factors of recent demographic developments in Hungary. *Demográfia* 48: 77-103.

Spéder, Z. 2006. Rudiments of recent fertility decline in Hungary: Postponement, educational differences, and outcomes of changing partnership forms. *Demographic Research* 15 (8): 253-288.

Steele, F., Joshi, H., Kallis, C., and Goldstein, H. 2006. Changing compatibility of cohabitation and childbearing between young British women born in 1958 and 1970. *Population Studies* 60 (2): 137-152.

Trost, J. 1978. A renewed social institution: Non-marital cohabitation. *Acta Sociologica* 21 : 303-315.

Trost, J. and Lewin, B. 1978. Att sambo och gifta sig. Fakta och föreställningar. *Statens offentliga utredningar* 55.

Vikat, A., Spéder, Z., Beets, G., Billari, F.C., Bühler, C., Désesquelles, A., Fokkema, T., Hoem, J.M., MacDonald, A.L., Neyer, G.R., Pailhé, A., Pinnelli, A., Solaz, A. (2007). Generations and Gender Survey (GGS): towards a better understanding of relationships and processes in the life course. *Demographic Research* 17 (14), 389-440.

Zabel, C. 2007. Do imputed educational histories provide satisfactory results in fertility analysis in the West German context? MPIDR Working Paper. Download from http://www.demogr.mpg.de/papers/working/wp-2007-022.pdf.

Zakharov, S. 2005. Recent trends in first marriage in Russia: Retarded second demographic transition. Paper presented to the 25th International Population Conference, Tours, France.

## Appendix

We have shown above how the interpretation of the hazard-regression parameters as relative risks can be maintained in the joint analysis of the competing risks under certain conditions. This is the purposes of our discussion of the consequences of relations (3) and (4). If relations of this nature do not hold, the computer program will still produce estimators $\hat{a}_{i\ell}$, $\hat{b}_{j\ell}$, and $\hat{c}_{k\ell}$ for the basic parameters, but we doubt that in general they can easily be interpreted as relative risks for comparisons across the two transitions. Here is why:

Remember that we normalize the parameters by setting $b_{j_0 1} = c_{k_0 1} = 1$ but that we make no similar normalization of the $b_{j2}$ and $c_{k2}$ in the joint analysis of the two transition risks. Now first note that under (2), in general, $\mu_{ijk\ell} / \mu_{ijk_0\ell} = c_{k\ell} / c_{k_0\ell}$, which we write out as

$$\frac{\mu_{ijk1}}{\mu_{ijk_0 1}} = c_{k1} \text{ for } \ell = 1 \text{ and } \frac{\mu_{ijk2}}{\mu_{ijk_0 2}} = \frac{c_{k2}}{c_{k_0 2}} \text{ for } \ell = 2.$$

To this extent, the $c_{k\ell}$ appear as relative risks for comparisons between factor effects on each risk, and we get similar relations for the $b_{j\ell}$. For comparisons across risks, however, we need to consider quantities like

$$\frac{\mu_{ijk2}}{\mu_{ij_0 k_0 1}} = \frac{a_{i2}}{a_{i1}} b_{j2} c_{k2},$$

and we fail to see how the latter quantities can be given a useful interpretation.

**Table 1.** Transition to first direct marriage and first cohabitation, Bulgarian women born in 1945-86. GGS data from 2004

| | **cohabitation** | **direct marriage** |
|---|---|---|
| **Age** | ***absolute risks per 1000 person months*** | |
| 15-19 | 3.220 | 4.613 |
| 20-24 | 3.958 | 7.448 |
| 25-29 | 2.424 | 3.773 |
| 30-34 | 0.932 | 2.038 |
| **Ethnicity** | ***relative risks*** | |
| Bulgarian | 1 | 1 |
| Turkish | 0.90 | 1.19 |
| Roma | 1.12 | 1.02 |
| Other | 1.39 | 0.53 |
| **Character of region where the respondent grew up** | | |
| Urban | 1 | 1 |
| Rural | 1.17 | 1.12 |
| **Parents lived together when the respondent was 15** | | |
| Yes | 0.80 | 0.93 |
| No | 1 | 1 |
| **Mother's highest level of education** | | |
| High | 1 | 1 |
| Middle | 1.08 | 1.04 |
| Low | 1.40 | 1.14 |
| Don't know | 1.70 | 1.18 |
| **Father's highest level of education** | | |
| High | 1 | 1 |
| Middle | 1.10 | 1.18 |
| Low | 1.18 | 1.04 |
| Don't know | 1.00 | 0.92 |
| **Number of siblings** | | |
| 0 or 1 | 1 | 1 |
| 2+ | 1.07 | 0.98 |
| **Parity/pregnancy status** | | |
| Childless, not pregnant | 1 | 1 |
| Childless, pregnant | 9.64 | 19.92 |
| Parity ≥1 (mother) | 0.53 | 0.53 |
| **Education** | | |
| In education | 0.39 | 0.31 |
| Completed low | 0.95 | 0.71 |
| Completed middle | 1 | 1 |
| Completed high | 1.02 | 1.15 |
| **Calendar year** | | |
| 1960-69 | 1 | 1 |
| 1970-79 | 1.67 | 1.32 |
| 1980-80 | 2.08 | 1.25 |
| 1990-99 | 2.32 | 0.77 |
| 2000-2004 | 1.66 | 0.30 |

**Figure 1. Trends 1960-2004 in standarized risks of union formation, relative to 1960-64, separately for each type of union**

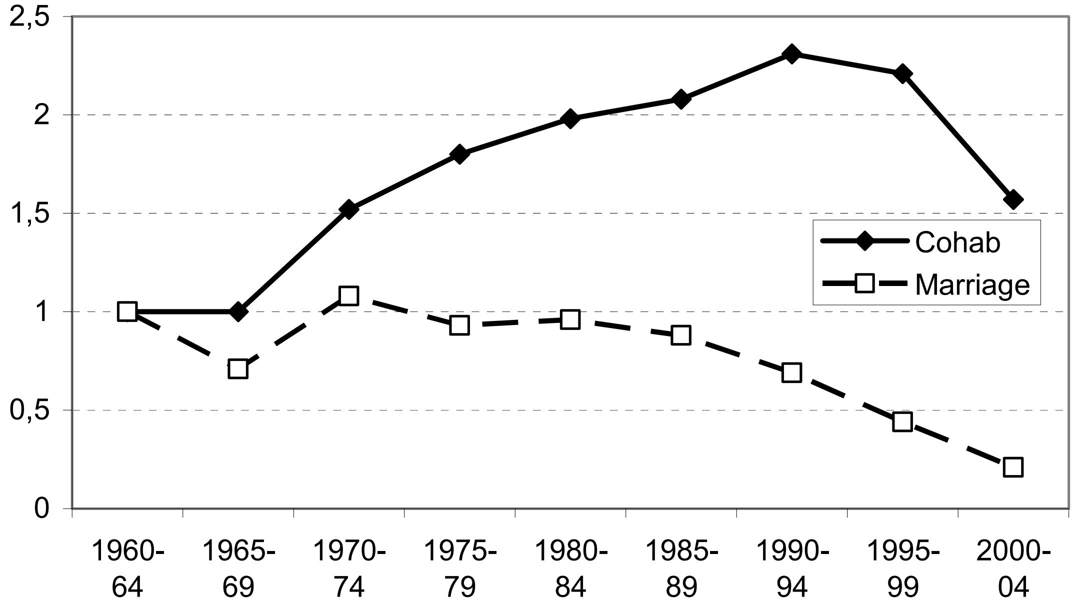


**Figure 2. Baseline age profiles in standardized risks of union formation, jointly for both types. Childless non-pregnant Bulgarian women, per 1000 person months**

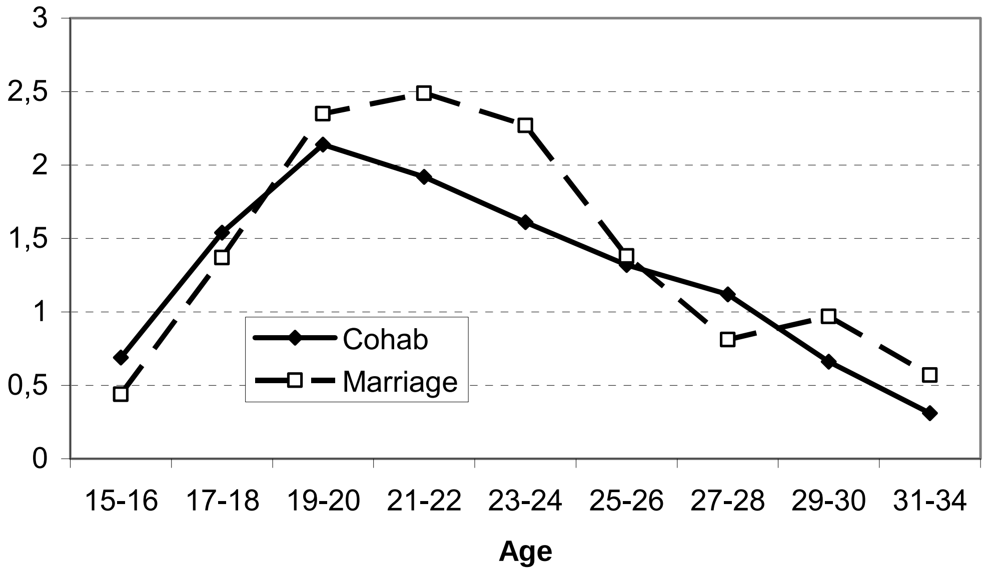
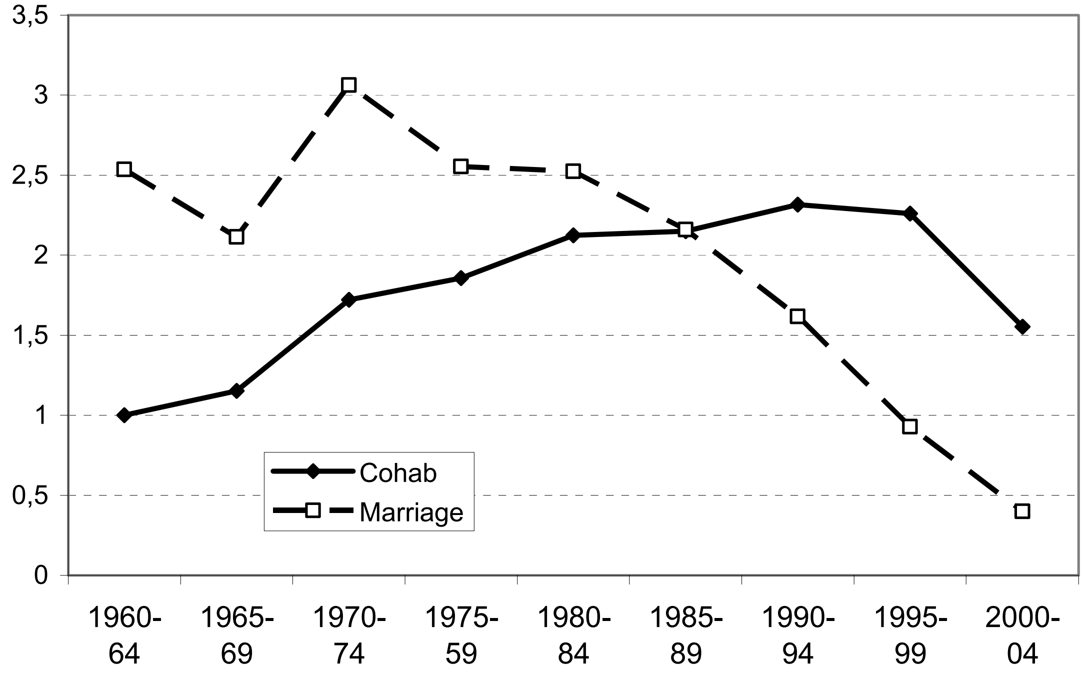


**Figure 3. Trends in relative risks of union formation 1960-2004, by type of union. Childless non-pregnant Bulgarian women**



## Figure 4. Relative rate of transformation of cohabitation into marriage