WP. 13
27 September 2007

ENGLISH ONLY

UNITED NATIONS STATISTICAL COMMISSION and ECONOMIC COMMISSION FOR EUROPE

STATISTICAL OFFICE OF THE EUROPEAN COMMUNITIES (EUROSTAT)

CONFERENCE OF EUROPEAN STATISTICIANS

Joint Eurostat/UNECE Work Session on Demographic Projections
(Bucharest, 10-12 October 2007)

Agenda item 7: Household projections

TOWARDS A DYNAMIC HOUSEHOLD PROJECTION MODEL FOR THE NETHERLANDS

Invited Paper

Submitted by Statistics Netherlands][1]

I. Introduction

1. Statistics Netherlands biennially produces a set of long-term demographic projections. These projections are made in a three stage process. First, the national population forecasts are made, which project the population by age, gender and country of origin. Next, the household projections are computed. These project the population by household position and marital status, as well as the number of households by type and composition. Finally, the regional population and household projections are made. The regional projections are produced jointly by Statistics Netherlands and the Netherlands Institute for Spatial Research (RPB) and have a forecast horizon of twenty years. The national projections currently extent until 2050.

2. Compared to the population forecasts, household projections at Statistics Netherlands have a relatively short history. The first ones were published in 1992 (De Beer, 1992). The projection model was, in the terminology of Kuijsten and Vossen (1988) a mixture of a static and a dynamic macrodemographic model. Transitions between marital statuses were modelled and in that sense the model was dynamic. Transitions between household positions were not modelled. In this sense, the model was static. Since the unit of analysis was subpopulations rather than individual persons, it was a macro model. The projection method was purely demographic. No socio-economic or other non-demographic determinants were taken into account.

3. Over the years, a number of adjustments have been made to the original model. In 1993, the model was extended to include projections of the composition of households by size and by

[1] Prepared by Coen van Duin (cdin@cbs.nl).

number of dependent children in the household (De Jong, 1994). In 2001, a cohort-perspective was introduced into the static part of the model (De Jong, 2001). In the latest round of household projections, an additional (static) sub model has been added which projects the household positions for first and second generation immigrants by country of origin (Van Duin, 2007). Despite these adjustments, the present household model remains a hybrid, with static and dynamic elements. For the new round of projections early 2009, Statistics Netherlands aims to develop a projection method where transitions between household positions are also modelled. This article discusses some issues in the development of the new model. In section 2, the current model is described in some more detail. Section 3 describes the register-based household data on which the new projections will be based. Section 4 outlines the preliminary choices for the new projection method. Section 5 focuses on a particular problem in the formulation of assumptions for the transition probabilities: the effect of increasing (healthy) life expectancy on pair dissolution among the elderly.

II. The current household projection method

4. Statistics Netherlands' current household projection model consists of four sub models, which are run consecutively.

5. In the first stage, the population by marital status is projected using a macro simulation model. The input for this model are marriage probabilities, divorce risks, birth- and death rates and net migration numbers by marital status. The population by marital status is computed by generating marital status transitions, deaths and births in the start-of-year population, and by applying net migration numbers, in order to obtain the end-of-year population Small corrections are then made to impose consistency with the total population by age and gender from the national population projection. Also, consistency relations between marital status transitions for men and women are enforced.

6. Next, the population by household position and marital status is projected using a static macro model. This static model in essence distributes the projected population by age, gender and marital status over the household positions. Persons with the household positions "single", "single parent" or "head of household of type Other" are reference persons in their household. Women with the household position "Living with a partner" are also counted as reference persons. The number of households by type is calculated from the number of reference persons. In the third stage the household composition by size and number of dependent children are projected using a dynamic macro model (De Jong, 1994). The transitions included in this model are birth of children and children leaving the parental household.

7. The fourth sub model is again static. It projects the number of first and second generation immigrants by household position. The projection is based on assumptions about how the age and gender-distributions of the household positions for the immigrant groups will change compared to those for the total population (convergence or divergence). Consistency with the results and assumptions of the household projection for the total population is imposed.

8. The multi-stage structure of the projections may seem overly complicated. Partly, it is a result of the fact that different parts of the model were developed at different times. It is less time consuming and saver to add a new sub model at the end of the projection-chain than to change a well-functioning existing model. A disadvantage of this kind of staged process is that information generated at the later stages is not used at the earlier stages. For instance: the influence of a changing household structure on the number of births is not explicitly modelled in

the current method, since population projections are made first and household projections second.

9. An advantage of the multi-stage structure is that the most reliable projections are made first and provide a reference frame for the projection of more uncertain quantities. Even though household-developments have an influence on the number of births, fertility behaviour is usually more predictable than household formation. This makes it unlikely that explicitly modelling the influence of expected trends in household composition on the number of births will improve the fertility forecast.

10. Another advantage of the multi-stage approach is that optimal use can be made of the available data. The marital status model uses the long time series that are available on marital status transitions. Similarly long time series exist on births by parity and age of the mother. These data are exploited in the macro simulation model which projects households by size and number of children. Reliable data on transitions between household positions was, until recently, scarce. This is the main reason why a static projection model for household positions continues to be used.

III. Register-based household data

11. In the period that the first household forecasts at Statistics Netherlands were made, data concerning household positions was only available from surveys. The projection of 1992 used data from the Household Demand Survey (WBO). This survey, held every four years, collected information from some 50 thousand households. Among other variables, it contained information on household positions at the time of the survey and on the household position a year earlier. Later projections used data from the Labour Force Survey, for which data was collected every year from 100 thousand persons.

12. In the early 1990s, data from the Household Demand Survey was used for dynamic household projections in the Netherlands. These were the first projection with the micro simulation model PRIMOS (Heida, 1991) and a number of demographic scenario's (one of which, termed "the realistic scenario", could be considered a projection) with the macro model LIPRO (Imhof and Keilman, 1991). Both of these dynamic projections used the retrospective data from the survey to estimate transition probabilities. Not each transition probability in the model could be estimated directly from the data in the survey, which made it necessary to use a number of simplifying assumptions. Sampling errors in the data were another problem for these early dynamic projections.

13. The introduction of the electronic municipal population register at the end of 1994 meant that a near-integral source of information became available. The data in the electronic register go back to the year 1995, which means that, at the moment, time series over a period of more than a decade are available. Since 2001, Statistics Netherlands publishes household statistics based on the information in this register.

14. Unlike the housing demand surveys, the population register does not contain the variable "household position". It does contain data on address, age, gender, marital status and family relations (spouse, parents, children) of registered persons. This data is used by Statistics Netherlands to construct a households register (Harmsen, 2003). First, persons are grouped together by address. Next, a derivation scheme is used to determine which groups of persons at

the address constitute a household, and what the household positions of the persons in these groups are.

15. The derivation scheme is partly deterministic. If only one person is registered as living at an address, the household position is determined to be “single”. If a married couple is living at the same address, their household positions are determined to be “partner in a married couple”. If two unrelated, unmarried persons move to a new address at the same date, they are classified as cohabiting. Using these types of decision rules a household position can be determined for about 93% of the cases.

16. In about 7% of the cases, various household positions are possible given the data in the register. In these cases, a probabilistic imputation scheme is used. This scheme takes into account the age difference of the people living at the address, their gender, their marital status, possible family relations between the people at the address and whether the address is in a rural or urban area.. The parameters of the imputation model are estimated from survey data.

17. For each new year, the probabilistic imputations are performed independently of those for the previous year. As a result, artificial transitions between household positions would be found if data from consecutive years were compared at the level of individual persons. In one year, two persons living at the same address might have been classified as cohabiting, while in the next year, the positions for the same two people at the same address would be imputed as single. The probabilistic fluctuations caused by the derivation scheme only lightly affect transversal quantities such as total number of persons per household position per year, because the fluctuations average out. They do however limit the usefulness of the data for longitudinal analyses.

18. To overcome this problem, a scheme is used where the imputed household positions in previous years are made consistent with those from the most recent year (Witvliet, 2002). The philosophy behind this procedure is that the data from the most recent year will be the most accurate, since it is based on the largest amount of information. For instance, if two persons that are living at the same address get married, it is very likely that they were cohabiting at this address before they got married. If their household positions were imputed as “single” in the previous year, this should be corrected retrospectively.

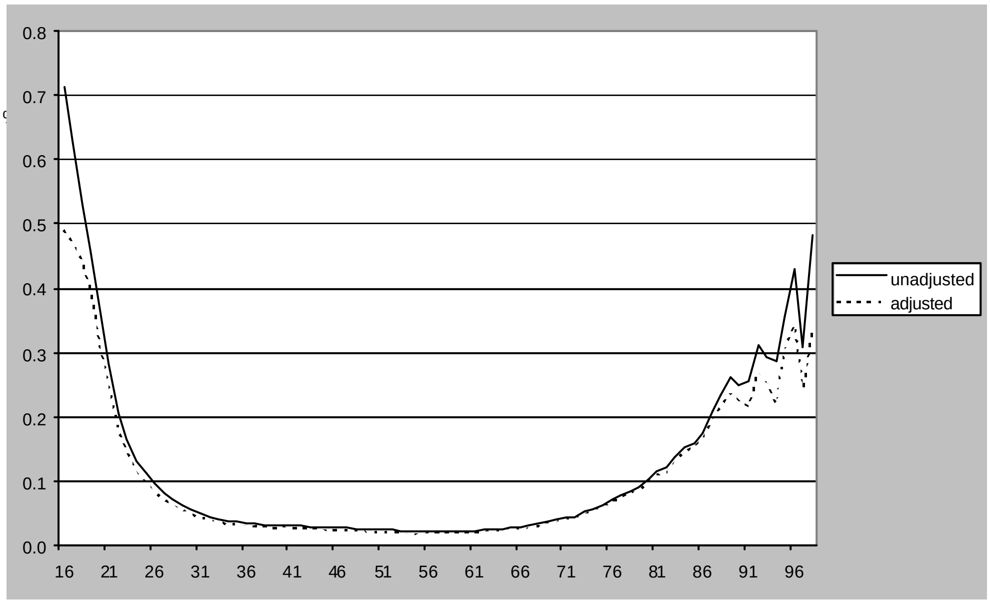
19. After the household positions for a new year have been derived, the imputed positions for the previous year are adjusted to conform with those for the new year. The adjustments introduce a small bias, since imputation errors where cohabiting people are imputed as singles are more likely to be discovered (for instance if they marry, or if they move together to a new address) than errors where two single people are imputed as cohabiting. Therefore, a second round of adjustments is applied which restores the original shares of the various household positions in the population. This second round of adjustments also assures that the adjusted numbers of persons per household position for the previous year differ little from the numbers that were derived, and published, a year earlier.

20. After the imputations in the previous year have been adjusted, those in the year before that are adjusted in the same way. This procedure is repeated until the first year in the register is reached. It has been shown that the method is stable and that the number of adjustments does not grow as one adjusts data from years that are further in the past.

21. Figure 1 shows the yearly exit probability from the state Living with partner for women in 2005, calculated from unadjusted and adjusted household register data. The adjustments reduce the estimated number of transitions at the young and old ages. At intermediate ages, the adjusted and unadjusted transition probabilities differ very little.

22. The register-based household data provide a rich source of information on household transitions. A large number of background variables is available, such as age, gender, country of birth (of registered persons and their parents), date of immigration (for immigrants), number and age of children, etc.. Because of the large mass of data, analyses at quite detailed levels are feasible. Also, the coupling between the household- and the population register allows us to investigate the relation between household transitions and demographic events, such as migration, births and deaths. Since part of the data is based on imputation rather than direct observation, the imputation procedure that was used should always be taken into account when interpreting the results of analyses on this data set.

***Figure 1**: Yearly exit probabilities from the state Living with partner (women, 2005)*



23. The register-based household data have been used by Statistics Netherlands and external users for a wide variety of demographic analyses. The regional projection models PRIMOS (ABF research) and PEARL (Statistics Netherlands/ Netherlands Institute for Spatial Research) also use it. The data is also employed in a micro simulation model that is being developed as part of the MicMac project (De Beer, 2006).

IV. Outlines of the new projection method

24. Statistics Netherlands aims to develop a fully dynamic household projection method which exploits the longitudinal information in the household register. One advantage of such a method is that it yields projections for flows between household positions. With the development of the regional projection model PEARL, producing national projections on such flows has become urgent. Since PEARL models household transitions, it needs national projections for these flows on which to base the time-development of the household transitions at the regional level.

25. The new projection method will be based on a dynamic macro model. It will therefore model transitions between demographic states on the basis of transition probabilities. These transition probabilities can depend on a number of background variables. Certainly, they should depend on age and gender, possibly on other variables such as ethnicity or number of children. There are no plans to include non-demographic variables, such as for instance level of education, into the projection model.

26. Since we aim to produce the same output as the current projection model, the states that are distinguished in the model should preferably correspond at least to those for which projections are now made. That means that we require a distinction by seven household positions, crossed with four marital states. The household positions distinguished in the current projection model are

- Single
- Living with partner
- Living in parental home (dependent child)
- Single parent
- Head of private household of type “other” (meaning: not a single person household, a household headed by two partners or a single parent household)
- Other member of private household
- Living in institution

27. The four marital states are Never married, Married, Widowed and Divorced. As some of the 28 combinations of household position and marital status are very rare, there will be room to reduce the number of states somewhat. On the other hand, it may be necessary to include additional states in order to improve the performance of the model. The expected developments in the household structure for various subgroups in the population should be taken into account when defining the states in the projection model (Prinz, 1994).

28. Apart from a projection of the number persons by household position and marital status, Statistics Netherlands also publishes projections about the number of households by size, by number of children and by age of the youngest child living in the household. These projections are made using the third component of the current projection model: the dynamical model for household composition. This model distinguishes females by age, parity (0,1,2,3,4+ children) and by the ages of the eldest and youngest child still living in the household. Introducing all this additional background information into the dynamic projection model for household positions would very much enlarge the state space. At present, we therefore aim for the new model to replace only sub models 1 and 2: the dynamical model for marital status projections and the static model for the projection of household position by marital state. In time, it should be able to replace sub model 4 as well (household positions of first- and second generation immigrants). Since children are attributed to households in a separate procedure, it is not necessary to split the position Dependent child into sub states for different household types, as was done in the LIPRO household projection for the Netherlands of 1991 (Imhoff, 1991).

29. For the technical implementation of the model, there are several options. In the research stage, we will probably be using simple prototypes programmed in excel. For the longer term, it seems more attractive to use an established model such as the LIPRO model of the Netherlands Interdisciplinary Demographic Institute, or an adjusted version of PEARL, rather than reinventing the wheel. Because of its flexibility, the LIPRO model will probably also be used in parts of the development process.

V. Time dependence of pair dissolution among the elderly

30. Once a model has been specified and implemented, a next challenge is to specify the set of "most probable" transition probabilities on which to base the projection. Typically, assumptions about the time development of the transition probabilities used in a projection will be quite conservative. One will assume that the probabilities will remain as they are, that a decreasing or increasing trend in (a summary measure of) the transition probabilities will continue for some more years and than stabilize, or that the transition probabilities will return from the current value to an average level over the past (say) ten years.

31. Death rates are an exception to this rule. Most population projections assume a reduction of the death rates right up to the projection horizon. If the deceased person is living with a partner, his or her death will have generated a household transition of the partner. This means that, especially at the higher ages, it may be insufficient to use a no-change assumption for the exit probabilities from the state Living with a partner.

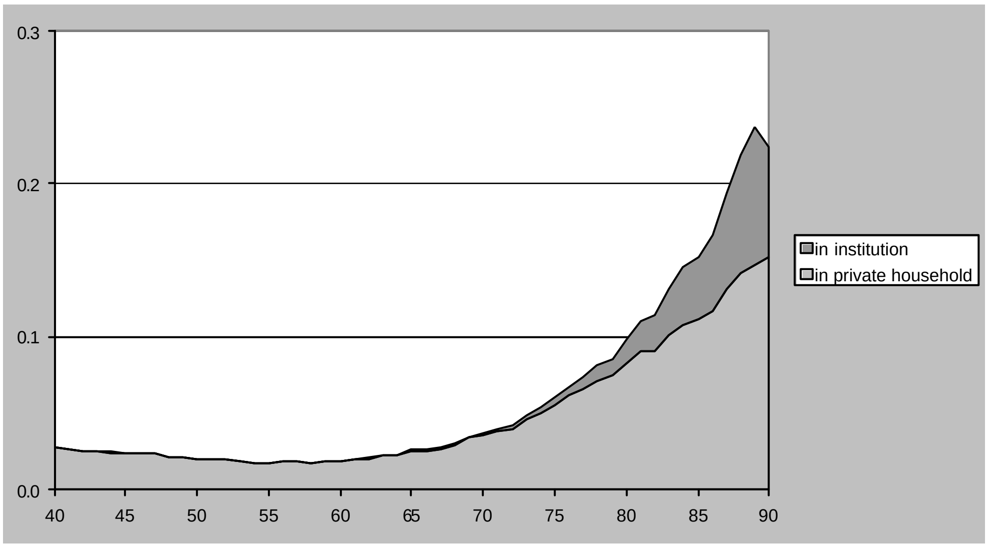
32. Another case where a no-change assumption is likely to be insufficient is the transition from living in a private household to living in an institution. In the Netherlands, there is a strong trend for people to remain living independently rather than moving to a retirement home up to increasingly higher ages. There are good reasons to assume that this trend will continue. An increase in healthy life expectancy will mean that people will become dependent on institutional care only at higher ages. Also, with an increase in the population of the elderly, there will be financial pressures to enable this group to continue living outside an institution for as long as possible.

33. As a first analysis into the time-dependence of transition probabilities, we look at the effect of expected changes in death rates for men and in the risks for men to move to an institution on the separation risks for women.

34. It could be argued that we do not need to introduce time-dependence into the a priori exit probabilities as a result of the partner dying or moving to an institution. After all, the outlined problem is one of consistency between transitions for men and women. If the projection model imposes consistency between the number of partnership dissolutions between men and women (as LIPRO does), then a decrease in the death/institutionalization probability for men will result in a decrease in the posteriori pair dissolution probabilities for women, and vice versa. A problem with this approach is that the changes in death and institutionalization rates are very age-specific, while most consistency algorithms, also those implemented in LIPRO, only impose consistency between numbers of transitions at an aggregated level. In other words: a reduction of the probability for men to die at age 80 would lead to an adjustment in the probability for women to lose their partner at age 20.

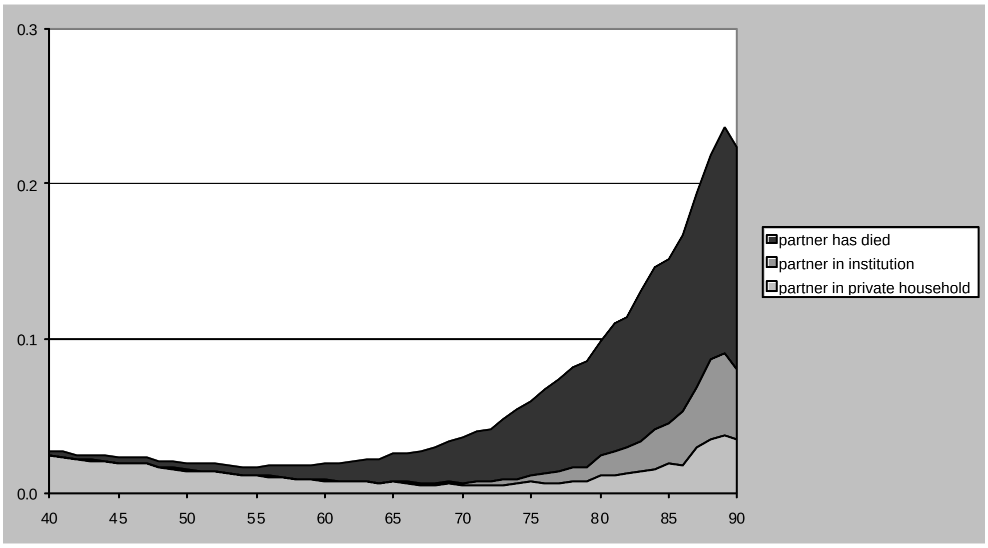
35. Figure 2 shows the yearly exit probabilities for women from the state Living with a partner for the year 2005, distinguished by the household position at the end of the year (institution or private household). Between the ages of 40 and 55, the exit probabilities for women from this state decrease with age, as a result of a decreasing risk of separation. Above the age of 55, the exit risk starts to rise. For ages of 80 years and older, a large part of this increase is due to women moving into an institution.

Figure 2*: yearly exit probabilities for women from the state Living with a partner by target state (2005)*



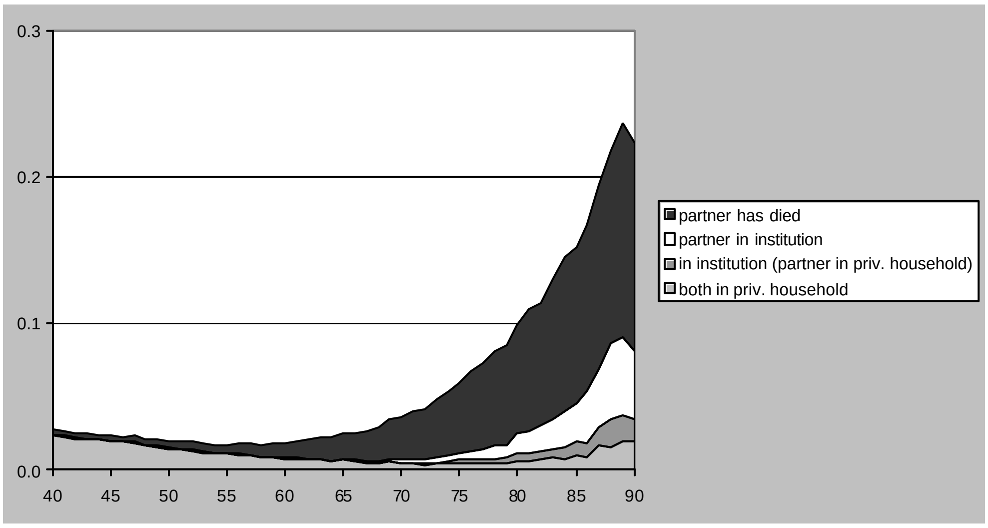
36. Figure 3 again shows the yearly exit probabilities for women from the state Living with a partner, but now distinguished by the state of the partner at the end of the year (dead, in an institution or in a private household). It can be seen that the rise of the exit probabilities between the ages of 55 and 75 can be almost completely attributed to an increase in the risk of the partner dying. At ages of 75 and over, an increase in the institutionalization risk of the partner also begins to play a role.

Figure 3*: yearly exit probabilities for women from the state Living with a partner by target state of the partner (2005)*



37. In figure 4, the exit probabilities are distinguished simultaneously by the end-of-year state of the women and of their partners. It can be seen that nearly the entire increase in the exit probabilities can be explained in terms of the partner dying, the partner entering an institution, or the woman herself entering one. Clearly, the expected changes in these processes could have a large effect on the household positions of the elderly.

Figure 4: *yearly exit probabilities for women from the state Living with a partner by target state of the woman and her partner (2005)*



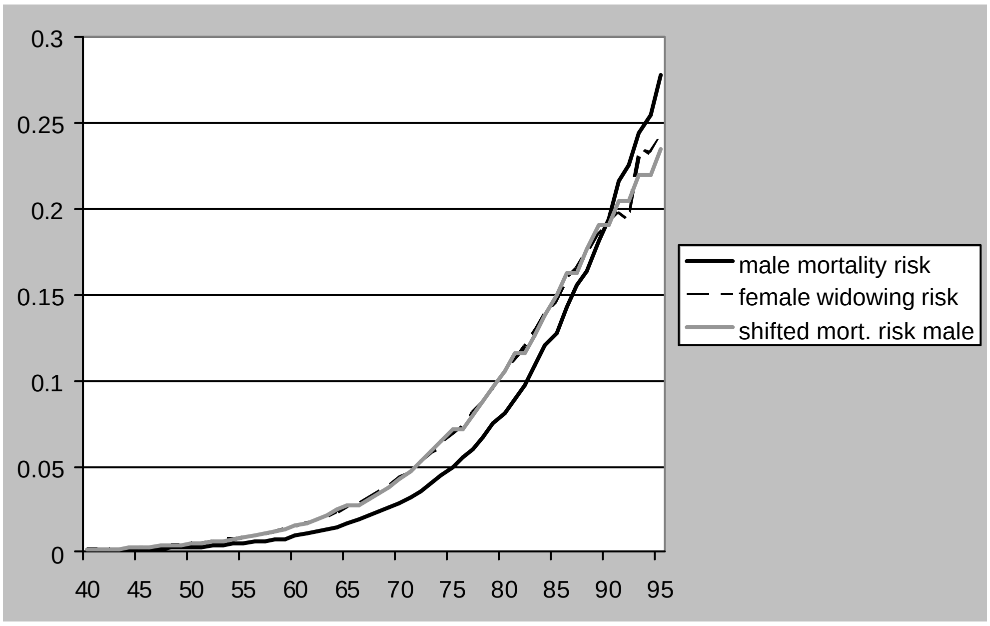
38. In order to quantify this effect, we can adapt a procedure that is being used to estimate the widowing risk in the dynamic marital status projection model. The following estimate is used

$p_{wid.}(x, female, married) \approx p_{death}(x + \Delta_x, male, married)$,

where $p_{wid.}$ is the age specific yearly risk to be widowed and p_{death} the age specific yearly risk to die, x is the age of the women and $?_x$ is an age-dependent shift which transforms the death-risks curve for married men into a widowing-risks curve for married women.

39. Figure 5 shows the age-specific widowing risk for women, and the shifted and unshifted mortality risk for married men. The risks are calculated for the period 1995-1999. The shifted mortality curve for men provides quite a good estimate for the widowing curve for women.

Figure 5: *Female widowing risk and mortality risk of married males, 1995-2000*

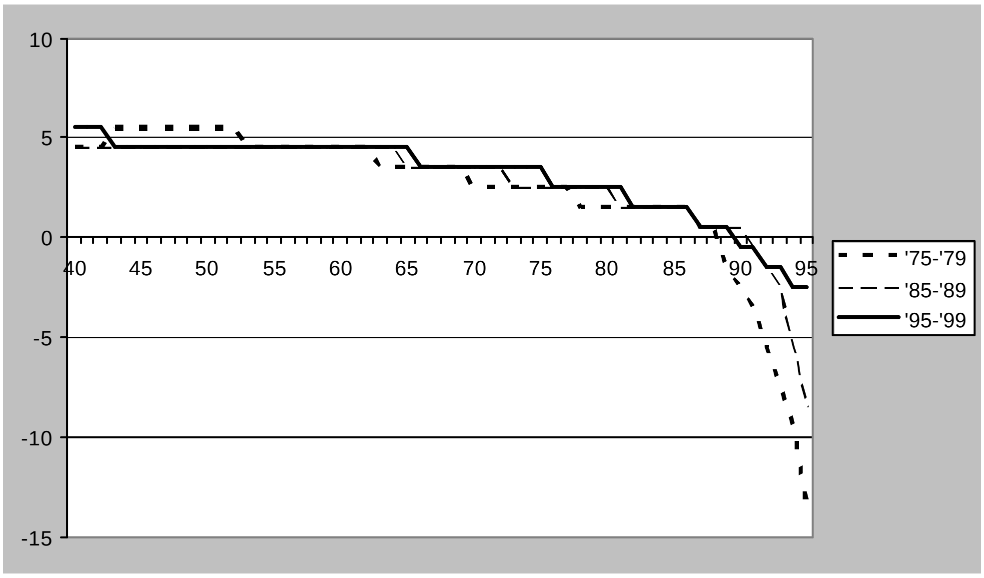


40. Figure 6 shows the shift parameter $?_x$ calculated on data from the periods 1975-1979, 1985-1989 and 1995-1999. For middle-aged women, the age shift is between 4 and 5 years. In

the Netherlands, the average age difference between the male and female partner in a marriage is 2-3 years. However, middle aged women with relatively older partners are more likely to become widowed, which is why the age shift is larger than the age difference for marrying couples. At the oldest ages, the shift parameter becomes negative. This is due to selection: a woman who is still not widowed in her 90s is more likely to have a younger husband than a married woman of 40.

41. There is some time dependence in the shift parameters which can be explained by the selection effect becoming weaker as the life expectancy of (married) men increases. Upton the age of 85, the time dependence of the shift parameter is weak. Much of the effect of reduced mortality for men on the widowing risks for women can therefore already be captured in a model that uses time independent shift parameters.

Figure 6*: Age-shift parameters calculated from data from three periods*



42. Using age-shift parameters, we can relate, in a similar way, the mortality risks of men that are living with a partner (married or unmarried) to the exit probability for women from this state because of the death of their partner. We can do the same for the risk of men entering an institution. We label the event that a partner dies "event 1" and the event that the partner moves to an institution "event 2". The probabilities of these events for women can be estimated as

$$p_1(x,\ female, living\ with\ partner) \approx p_{death}(x+\Delta^{(1)}{}_x, male, living\ with\ partner)$$

$$p_2(x,\ female, living\ with\ partner) \approx p_{institution}(x+\Delta^{(2)}{}_x, male, living\ with\ partner)$$

and we can estimate the time development for the female separation risks as

$$p_1(x,\ female, living\ with\ partner, j_1) \approx$$

$$p_1(x,\ female, living\ with\ partner, j_0)\frac{p_{death}(x+\Delta^{(1)}{}_x, male, j_1)}{p_{death}(x+\Delta^{(1)}{}_x, male, j_0)},$$

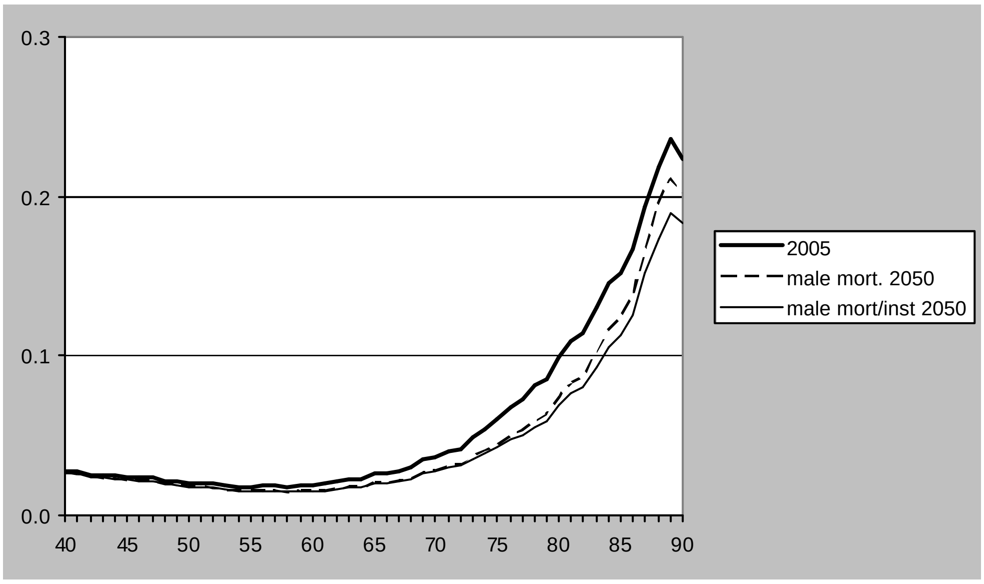
$$p_2(x,\ female, living\ with\ partner, j_1) \approx$$

$$p_2(x,\ female, living\ with\ partner, j_0)\frac{p_{institution}(x+\Delta^{(2)}{}_x, male, j_1)}{p_{institution}(x+\Delta^{(2)}{}_x, male, j_0)}.$$

43. Here, we have made an additional assumption that the relative change in p_{death} and $p_{institution}$ for men living with a partner is the same as for men not distinguished by household position.

44. To get a feeling for the size of these effects, we calculate the exit probability from the state Living with partner for women based on adjusted risks for these two events. We adjust the risks for event 1 by using the mortality risks for men in 2005 and the projected risks in 2050. As an approximation, we use the age-shift parameters for widowing. The risks for event 2 are adjusted by assuming that the risk for men to move to an institution is reduced by 40%. This is roughly in line with the assumptions for the proportion of persons living in an institution in 2050 in the latest household projections. Since the same reduction factor is applied to all ages, the age-shift parameters for event 2 have no influence on the results. Figure 7 shows the adjusted exit probabilities

***Figure 7**: Exit probability from state Living with partner for women, 2005, with adjusted risks for event 1 (dashed line) and for events 1 and 2 (thin solid line).*



45. The assumed trends in mortality for men and in the age at which men move to an institution would result in a shift, to higher ages, of the separation risk-curve for women with some 3-4 years. This is not a small effect. Therefore, it seems necessary to take into account the age specific relation between these transitions for men and the separation transitions for women.

VI. References

Kuijsten, A. and Vossen, A. (1988), *Introduction* in Modelling Household Formation and Dissolution, by Keilman, Kuijsten and Vossen, pages 3-12,Clarendon Press-Oxford

Beer, J. de, Jong, A. de, Hoorn, W. van and Latten, J. (1992), National household forecasts 1992 (in Dutch), Maandstatistiek van de Bevolking 40 (7) , pages 28-46, Centraal Bureau voor de Statistiek

Beer, J. de and Gaag, N. van der (2006), *Using MicMac to project living arrangements: an illustration of biographic projections,* presented at joint Eurostat/UNECE worksession of demographic projections, 2006

Harmsen, C. and Israëls, A., *Register based household statistics,* Paper for European Population Conference 2003, 26-30 August 2003.

Jong, A. de (1994), *A macrosimulation model for projecting households by size,* Working paper no. 34, ECE/Eurostat joint worksession on demographic projections, 1-4 June 1994

Jong, A. de and Beer, J. de (2001) *The household projection model* (in Dutch), Maandstatistiek van de Bevolking 49 (7) , pages 16-20, Centraal Bureau voor de Statistiek

Imhoff, E. van and Keillman, N., (1991). *LIPRO 2.0: An application of a dynamic demographic projection model to household structure in the Netherlands*, Swets and Zeitlinger, Amsterdam

Prinz, C. (1994) *Problem formulation, model selection and data quality –some issues in the context of living arrangement projections*, Working paper no. 32, ECE/Eurostat joint worksession on demographic projections, 1-4 June 1994

Van Duin (2007), to be published in Bevolkingstrends 55(3) (in Dutch)

Witvliet, H. (2002), *Household transitions –corrected counts and transition probabilities,* (in Dutch), Statistics Netherlands, Internal paper

*****