Mathematical Population Studies
1997, Vol. 6(3), pp. 187–216
Reprints available directly from the publisher
Photocopying permitted by license only

A MULTI-DIMENSIONAL MODEL FOR PROJECTING FAMILY HOUSEHOLDS – WITH AN ILLUSTRATIVE NUMERICAL APPLICATION

ZENG YI[a,b,*], JAMES W. VAUPEL[b] and WANG ZHENGLIAN[b]

[a]*Institute of Population Research at Peking University, China;*
[b]*Max Planck Institute for Demographic Research, Eschenstrasse 3, D-18057, Rostock, Germany*

This paper develops a multi-dimensional model for projecting households and population. The model is constructed to ensure consistency between the demographic events occurring to males and females as well as to parents and children. The model permits projection of characteristics of households, their members, and population structure, using data that are usually available from conventional sources. Unlike the traditional headship-rate method, our model can closely link the projected households with demographic rates. The model includes both nuclear and three-generation households, so that it can be used for countries where nuclear households are dominant and for countries where nuclear and three-generation households are both important. The illustrative application to China, although brief, provides some policy-relevant information about future trends of Chinese household size, structure, and the age and sex distribution of the population, with a focus on the elderly.

KEY WORDS: Household; Projection; Consistency; Scenarios; Policy; Ageing

INTRODUCTION

In almost all countries, the size and structure of family households are changing. Populations are ageing in most countries as a result of lower fertility and increasing life expectancy. Age patterns of childbearing are changing, with more people having their first child at older ages. In some countries, marriage rates are declining and divorce rates are rising. People are living longer, so that an increasing number of middle-age workers have living children, parents, and even grandparents. The gap between male and female life expectancy is increasing, leaving more widows. Increased mobility is leading children to move to areas distant from their parents. These factors, in various combinations and strengths for different populations, are yielding new patterns and distributions of family household structures.

* Corresponding author. We have been working on the first version of the user-friendly computer software with a manual associated with our model described in this paper. Scholars who are interested in applying this model and software may write to: Wang Zhenglian, Max Planck Institute for Demographic Research, Eschenstrasse 3, D-18057, Rostock, Germany.

There is an important interaction between changes in household structure and the health status of elderly. Living alone or without family ties can cause or worsen ill health and disability. Furthermore, family members often supply support to the elderly who are ill or disabled. In the absence of such support, need for nursing homes, social services, and health-care services will increase. Health-care costs and social services provided to the elderly now account for over 10% of GNP in many countries, including the United States. As the proportion elderly grows, these costs grow as well. For this and other reasons, projections of household structure are clearly of considerable interest to planners and policy analysts in governmental, business, nonprofit and academic organizations.

How will demographic changes alter the number and proportion of different kinds of family households, including the single-parent family, the three-generation family, the family household consisting of only one or two elderly people without children, etc.? How many elderly persons will need assistance, but will not have children, spouse and other close relatives to provide it? How many middle-age persons will have to care for both elderly parents and young children? This article is aimed at developing a multi-dimensional model for projecting family households that can be used to address questions such as these.

A brief review of the three kinds of major models for family household projection is presented in the next section. This section is not intended to be an exhaustive literature review: it simply provides some background information related to the construction of our model. We then discuss our model in detail. An illustrative example of the application focusing on China is presented in the last section.

MODELS FOR FAMILY HOUSEHOLD PROJECTION

Demographers mainly use three kinds of models to project household structure: microsimulation, macrosimulation and headship-rate.

Microsimulation models have major advantages in studying the variability of individuals and households and their distributions (Hammel *et al.*, 1976; Wachter, 1987; Smith, 1987; Nelissen, 1991). However, in a large population in which households are classified by a relatively large number of characteristics, the size of the representative sample to be used as the starting point of a projection should also be large. For example, a sample of one per cent of the population of China and the United States consists of about 12 and 2.6 million persons respectively. To simulate so many persons one by one would take very substantial computing power and time. Another problem is that the census usually asks simple questions which cannot provide enough data for the microsimulation to model detailed characteristics of individuals. Hammel, Wachter and their colleagues handled this problem by starting their simulations with a pre-simulation for a few decades before the beginning year of their projections. Using a manageable sample for this pre-simulation they were able to approximate the family, household and kinship distribution of the beginning year of the projection, and then simulate it forward. This approach is not as good as directly using the population, family household, and kinship distribution obtained from a census or a large survey, for two reasons. First, the simulated family and population distribution at the beginning year of the projection may not be accurately the same as the census or survey enumeration. Second, the procedure demands additional detailed data for a few decades before the beginning year of the projection, and such data may not be available.

Despite their widespread use, headship-rate methods suffer several serious shortcomings. The head is an arbitrary choice and a vague one that varies from area to area and may change over time: this creates great difficulties for projection (Murphy, 1991). Trends in headship rates are not easy to model (Mason and Racelis, 1992, p. 510). Often the information produced by projections using the headship-rate method is inadequate for planning purposes (Bell and Cooper, 1990). Above all, the major disadvantage of the headship method is the unclear link to underlying demographic events: it is very difficult to incorporate demographic assumptions about future changes in fertility, marriage, divorce and mortality (Mason and Racelis, 1992, p. 510; Spicer *et al.*, 1992, p. 530).

The macrosimulation approach does not suffer from the shortcomings inherent in headship-rate methods. Although not as flexible as microsimulation models in analysing variability and probability distributions, macrosimulation models are not limited to the sample size of the beginning year of the projection and can use fully the information from a census or a large survey as a starting point. Furthermore, planners and policy analysts can conduct macrosimulation projections relatively easily on a personal computer if user-friendly software and a lucid manual are provided.

Although we believe that there are some important advantages to macrosimulation, as outlined above, we do not wish to imply that microsimulation approaches are less useful. It would be desirable to develop both kinds of approaches, since they have complimentary strengths, and since the issues they can be applied to are so important.

Keilman (1988), Van Imhoff and Keilman (1992) and Ledent (1992) reviewed dynamic family-household models based on the macrosimulation approach. Most of these models require data on transitions among various household types, data that have to be collected in a special survey because they are not available in the conventional demographic data sources of vital statistics, censuses and ordinary surveys. As stated by Van Imhoff and Keilman (1992), the high demands of data, especially the not-commonly available data on transitions between household types, in most dynamic household models are an important factor in the slow development and infrequent application of these models. Therefore, it is important to develop a dynamic family-household model that requires only conventional demographic data that can be obtained from vital statistics, census and ordinary surveys.

Benefiting from methodological advances in multi-dimensional demography (Rogers, 1975; Land and Rogers, 1982), and especially the multistate marital-status life-table model (Willekens *et al.*, 1982; Willekens, 1987), Bongaarts (1987) developed a nuclear family-status life-table model. Zeng Yi extended Bongaarts' model and included both nuclear and three-generation families (Zeng, 1986; 1988; 1991). The life-table models by Bongaarts and Zeng Yi are female-dominant models and they assume that age-specific demographic rates are constant. Building on Zeng Yi's family-status life-table model, this article develops a two-sex dynamic projection model. As will be shown later, the model presented in this article permits demographic schedules to change over time and does not require data that are not available from conventional demographic data sources.

DEMOGRAPHIC STATUS IDENTIFIED

Following Brass' approach (Brass, 1983), the individual is chosen as the basic unit of the projection model. The major reason why we chose the individual as the unit is that demographic rates that are available from the conventional population data sources can then be readily linked to individuals. The individuals of the base (or starting) population derived from a census or survey and the future projected population are all classified according to eight dimensions of demographic statuses listed in Table 1.

ACCOUNTING SYSTEM TO LINK INDIVIDUALS' CHARACTERISTICS WITH FAMILY HOUSEHOLD

It is theoretically possible for a woman or man to have any realistic combination of the statuses identified in Table 1. We may call the combination the composite state. In a numerical application, for instance, 2 residence statuses, 5 marital statuses, 3

TABLE 1
Individual's demographic statuses identified in the model

Demographic status	Code and definition
Age x	0 to the highest age. The highest age (denoted as W) is to be determined by the user.
Sex s	1. Female; 2. Male.
Marital status m	1. Single and not cohabiting; 2. Married; 3. Widowed; 4. Divorced; 5. Cohabiting. Cohabit status refers to those who live with a partner but are not married. The five marital statuses are assumed to be mutually exclusive.
Co-residence with parent(s) k	1. Living with two parents; 2. Living with one parent only; 3. Not living with parents; 4. Living in a collective household in which members have no any family or intimate relationship to each other.
Parity p	$p=0$ to the highest parity. The highest parity (denoted as P) is chosen by the user depending on the fertility level in the country or area under study.
Co-residence with children c	$c=0$ to the highest parity.
Residence r	1. Rural; 2. Urban.
Projection year t	t_1(starting year) and t_2 (ending year) of the projection are determined by the user.

Notes: (1) $m=5$, and $k=4$ are optional statuses, and users can omit them when such living arrangements are rare or the data are not available. (2) The status of "co-residence with parents" here is broadly defined. A child who is living with parent(s) or grandparent(s) or other senior family members who act as care provider when parents are not available is classified as "living with parent(s) ($k=1$ or 2)". Consequently, this nonindependent child is not a reference person (or marker) of the household. (3) If the demographic differences between rural and urban sectors are small, as is the case in many developed countries, or if the data classified by rural/urban sectors are not available, the rural–urban dimension can be omitted. (4) In terms of time reference t, stock variables and flow variables are distinguished. The stock variables are demographic status such as age, marital status, parity, maternal status, residence, etc., and they all refer to a particular point of time. The flow variables are demographic events, rates and probabilities and they all refer to a period (e.g. one year).

TABLE 2
Family household types and sizes identified in the model

Family household types	Reference person's status: s	k	m	p	c	Household size	Number of households in the year t
One-generation							
One-woman	1	3	1, 3, 4	≥ 0	0	1	$G_a^1(t) = \sum_{x=0,\dots,W} \sum_{m=1,3,4} \sum_{p=0,\dots,P} N_{3,m,p,0}(x, t, 1)$
One man	2	3	1, 3, 4	≥ 0	0	1	$G_b^1(t) = \sum_{x=0,\dots,W} \sum_{m=1,3,4} \sum_{p=0,\dots,P} N_{3,m,p,0}(x, t, 2)$
One couple	1	3	2, 5	≥ 0	0	2	$G_c^1(t) = \sum_{x=\alpha,\dots,W} \sum_{m=2,5} \sum_{p=0,\dots,P} N_{3,m,p,0}(x, t, 1)$
Two-generation							
a couple + children	1	3	2, 5	>0	>0	$2+c$	$G_a^2(t) = \sum_{x=\alpha,\dots,W} \sum_{m=2,5} \sum_{p=1,\dots,P} N_{3,m,p,c}(x, t, 1) - (G_a^3 + G_c^3 + G_e^3)$
Nonmarried mother	1	3	1, 3, 4	>0	>0	$1+c$	$G_b^2(t) = \sum_{x=\alpha,\dots,W} \sum_{m=1,3,4} \sum_{p=1,\dots,P} N_{3,m,p,c}(x, t, 1) - (G_b^3 + G_d^3 + G_f^3) \times R$
Nonmarried father	2	3	1, 3, 4	>0	>0	$1+c$	$G_c^2(t) = \sum_{x=\alpha,\dots,W} \sum_{m=1,3,4} \sum_{p=1,\dots,P} N_{3,m,p,c}(x, t, 2) - (G_b^3 + G_d^3 + G_f^3) \times (1 - R)$
Three-generation							
two grand parents + a couple at 2nd generation + children	1	1	2, 5	>0	>0	$2+2+c$	$G_a^3(t) = \sum_{x=\alpha,\dots,W} \sum_{m=2,5} \sum_{p=1,\dots,P} N_{1,m,p,c}(x, t, 1)$
One grandparent + a couple at 2nd generation + children	1	2	2, 5	>0	>0	$1+2+c$	$G_b^3(t) = \sum_{x=\alpha,\dots,W} \sum_{m=2,5} \sum_{p=1,\dots,P} N_{2,m,p,c}(x, t, 1)$
Two grandparents + lone-mother + children	1	1	1, 3, 4	>0	>0	$2+1+c$	$G_c^3(t) = \sum_{x=\alpha,\dots,W} \sum_{m=1,3,4} \sum_{p=1,\dots,P} N_{1,m,p,c}(x, t, 1)$
One grandparent + lone-mother + children	1	2	1, 3, 4	>0	>0	$1+1+c$	$G_d^3(t) = \sum_{x=\alpha,\dots,W} \sum_{m=1,3,4} \sum_{p=1,\dots,P} N_{2,m,p,c}(x, t, 1)$
Two-grandparents + lone-father + children	2	1	1, 3, 4	>0	>0	$2+1+c$	$G_e^3(t) = \sum_{x=\alpha,\dots,W} \sum_{m=1,3,4} \sum_{p=1,\dots,P} N_{1,m,p,c}(x, t, 2)$
One grandparent + lone-father + children	2	2	1, 3, 4	>0	>0	$1+1+c$	$G_f^3(t) = \sum_{x=\alpha,\dots,W} \sum_{m=1,3,4} \sum_{p=1,\dots,P} N_{2,m,p,c}(x, t, 2)$

Notes: (1) $x, s, m, k, p, c, t, W, P$, are defined in Table 1. $N_{k,m,p,c}(x, t, s)$ is the population of age x, sex s and k, m, p, c status at year t; (2) Categories of three-generation family households can be omitted when there are a negligible number of three-generation family households; (3) Those elderly couples, who live together with an ever-married child (and the child's spouse if the child is currently married) and grandchildren, are not reference persons of a family household, since the ever-married child with whom they live has already taken the position of a reference person. A family household cannot have two reference persons. Therefore, the number of those elderly couples that is equal to $(G_a^3(t) + G_c^3(t) + G_e^3(t))$ should be subtracted when we compute $G_a^2(t)$; (4) Similarly, $(G_b^3(t) + G_d^3(t) + G_f^3(t)) \times R$ and $(G_b^3(t) + G_d^3(t) + G_f^3(t)) \times (1 - R)$ should be subtracted when we compute $G_b^2(t)$ and $G_c^2(t)$. R is equal to the number of nonmarried women over age 49, not living with parents and with at least one child or grandchild living at home, divided by the total number of nonmarried women over age 49 and men over age 51, not living with parents and with at least one child living at home.

TABLE 3

Comparison of number of family households by household types in China between the results derived from the model-count and the direct-count, using the 1% data tape of 1990 census

Family type	Number of family households				Frequency distribution			
	Mod-co.	Dir-co.	Dif. #	Dif. %	Mod-co.	Dir-co.	Dif. #	Dif. %
One-person	202547	202493	54	0.03	0.0694	0.0697	−0.0003	−0.43
One-couple	193416	193376	40	0.02	0.0662	0.0666	−0.0004	−0.60
2 generation	1981961	1977370	4591	0.23	0.6787	0.6806	−0.0019	−0.28
3+ generation	542496	531983	10513	1.98	0.1858	0.1831	0.0027	1.47
Total	2920420	2905222	15198	0.52	1.00	1.00	/	/

Notes: Mod-co. = Model-count; Dir-co. = Direct-count; Dif. # (Absolute difference) = (Mod-co. − Dir-co.); Dif. % (Relative difference) = 100∗(Mod-co. − Dir-co.)/Dir-co.

statuses of co-residence with parents, 6 parity statuses and 6 statuses of co-residence with children are distinguished. The total number of composite statuses at each age is thus

$$T = 2 \times 5 \times 3 \times \sum_{p=0}^{5}(p+1) = 630.$$

Brass (1983) calls the reference person of a family household a "marker", and we follow Brass' marker approach to identify family households among individuals with possible composite statuses. In Brass' original work and Zeng Yi's family-status life-table model, only a senior female is chosen as a marker, which implies a female-dominant one-sex model. In the model discussed in this article, both sexes are included, and a female adult, or a male adult when female adult is not available, is identified as the reference person. The family household type and size are derived from the characteristics of the reference person, as listed in Table 2.[1]

We have tested the accuracy of this accounting system. Using the one per cent sample data tape of China's 1990 census, we identified each individual's code of sex, marital and status of co-residence with parents and children. According to these codes, we identified the reference persons. Based on the characteristics of the reference persons and following the accounting system described in Table 2, we derived the distribution of households by types and sizes. The household distribution derived in this way might be called a "model-count". Second, we followed the standard census tabulation approach and derived the household distribution directly using the codes that record household membership and relationship to the head of the household. This kind of census tabulation might be called "direct-count". As shown in Table 3, a comparison of distributions of household types derived by the "model-counts" and the "direct-counts" shows that the relative differences are very small, all

[1] When both an adult woman and an adult man are present in a family household, we chose the adult women as the reference person of the family household, because women marry earlier and live longer, reliable age-parity specific fertility data for women are much easier to obtain than for men, and following divorce, young children more likely live with their mother.

TABLE 4

Comparison of number of family households by household sizes in China between the results derived from the model-count and the direct-count, using the 1% data tape of 1990 census

Family Type	Number of family households				Frequency distribution			
	Mod-co.	Dir-co.	Dif. #	Dif. %	Mod-co.	Dir-co.	Dif. #	Dif. %
1 person	202547	202493	54	0.03	0.0694	0.0697	−0.0003	−0.43
2–3 persons	1018034	1038631	−20597	−1.98	0.3486	0.3575	0.0089	2.49
4–5 persons	1326508	1248036	78472	6.29	0.4542	0.4296	0.0246	5.74
6+ persons	373331	416062	−42731	−10.27	0.1278	0.1432	−0.0154	−10.74
Total	2920420	2005222	15198	0.52	1.00	1.00	/	/

Notes: The same as in Table 3.

below one per cent with one exception.[2] The relative differences in frequency distributions of household sizes between model-counts and direct-counts are also reasonably small: 0.4 per cent for one-person households, 2.5 per cent for small households with 2–3 persons, 5.7 per cent for the middle size households with 4–5 persons, and −10.7 per cent for the large households with 6 and more persons.[3] The overall average household sizes derived from model-counts and direct-counts are 3.83 and 3.88 (Table 4).

COMPUTATIONAL STRATEGY

Let $l_i(x,t)$ denote the number of persons of age x with composite state i ($i=1, 2,\ldots,T$) in year t. Let $P_{ij}(x,t)$ denote the probability that a person of age x with composite state i in year t will survive and be in composite state j at age $x+1$ in year $t+1$. Thus

$$l_j(x+1,t+1) = \sum_{i=1}^{T} l_i(x,t)P_{ij}(x,t).$$

[2]The reason why the error in the model-counts of three-generation family households is relatively large is that a joint family of two or more married brothers, with children living together, is counted as two or more stem family households. The error will be essentially eliminated in societies where the proportion of joint families is negligible. This is the case at present in Western countries and likely in near future years in China and most other developing countries.

[3]Two factors contribute to the lower accuracy in the household size model-count. The first one is that the model-count does not include those who are neither stem-family members nor spouse (or cohabiting partner) of the reference person. The second factor is that we limit the highest parity to 5 in the illustrative example, which underestimates the size of the large family households that have more than five children. The underestimation due to the first factor can be corrected by multiplying an adjustment factor: $z'(i,t) = a(i,t) \times z(i,t)$, where $z(i,t)$ and $z'(i,t)$ are the number of households with size i in year t before and after the adjustment respectively; $a(i,t)$ is the ratio of the numbers of households accounting and without accounting for those who are neither stem-family members nor spouse (or cohabiting partner). The value of $a(i,t)$ can be estimated based on the census base population at the starting year of the projection t_1, and assumptions about the future trend. The underestimation due to the second factor can also be reduced to a minimum degree by increasing the highest parity considered.

If $P_{ij}(x, t)$, which are elements of a $T \times T$ matrix, were properly estimated, the calculation of $l_j(x+1, t+1)$ would be straightforward. Unfortunately, the estimation of $P_{ij}(x, t)$ is usually not practical when the total number of states T is large, which is the case in our model. As stated earlier, for example, if a total number of 630 statuses is distinguished, the total number of cells in the transition matrix is thus: $630 \times 630 = 396{,}900$. There would be one such large matrix for males and females respectively at each single age. Although there are many zero cells in the matrices, the number of nonzero cells to be estimated is still much too large. Since so many categories have been distinguished, the number of observed events for some categories are too few to estimate the transition probabilities, even if the sample size is large. Therefore, the estimation of such large transition matrices is not practical.

In Bongaarts's nuclear family-status life-table model, a large number of statuses are also distinguished. Bongaarts overcame the difficulty by assuming that particular events take place at particular points in time between age x and $x+1$ (Bongaarts, 1987). Here, we follow Bongaarts' useful approach, and assume and compute the status transitions at different points of time in the single-year age interval. In particular: (1) births occur throughout the first half and the second half of the year. The birth probabilities used refer to the corresponding half year and they depend on status at the beginning and middle of the year; (2) deaths, migration, changes in status of co-residence with parents and marital status transitions, and changes in number of co-residing children due to children's death or leaving or returning home all occur at the middle of the year. These transition probabilities refer to the whole year and they depend on the corresponding status at beginning of the year. As shown mathematically and numerically by Zeng (1991, pp. 61–63; pp. 81–84), the strategy of assuming that births occur throughout the first and the second halves of the year and other status transitions occur at the middle of the year leads to accurate numerical estimates.

CONSISTENCY IN THE TWO-SEX MODEL

Because our model deals with two sexes and both children and parents, the following procedures are adopted to ensure the consistency.

(1) *Consistency between females and males.* We use the harmonic-mean procedure to ensure that in any year, the number of male marriages is equal to the number of female marriages; the number of male divorces is equal to the number of female divorces; and the number of newly widowed females or males is equal to the number of new deaths of currently married men or women. When cohabitation is modeled, the number of newly cohabiting males is equal to the number of newly cohabiting females; the number of males who exit from cohabiting status is equal to the number of female counterparts; the number of females or males whose cohabiting partners died in the year is equal to the number of new deaths of currently cohabiting men or women. Mathematical formulas for the harmonic mean can be found elsewhere (see, for example, Keilman, 1985, pp. 216–221). It has been shown that the harmonic mean satisfies most of the theoretical requirements and practical considerations for

handling consistency problems in a two-sex model (Pollard, 1977; Schoen, 1981; Keilman, 1985; Van Imhoff and Keilman, 1992).

(2) *Consistency between children and parents.* Define three quantities from the perspective of children:[4] C_1 – number of k status transitions from $k = 1$ or 2 (living with two parents or one parent) to $k = 3$ (not living with parents) due to leaving home; C_2 – number of k status transitions from $k = 1$ or 2 to $k = 3$ due to death of parents; C_3 – number of deaths of persons who lived with parents at time of death. Let $S_1 = C_1 + C_2 + C_3$. Note that S_1, C_1, C_2, C_3 and all other variables defined in this sub-section are year-specific, but we omit the time dimension for simplicity.

Define three quantities from the perspective of parents: P_1 – number of events of c status reductions[5] of wives (and cohabiting women if any) and single parents due to children's leaving home; P_2 – number of events of c status reductions of wives (and cohabiting women if any) and single parents due to children's deaths; P_3 – number of deaths of a couple (both parties die in the same year) multiplied by number of their co-residing children plus number of deaths of single parents multiplied by their number of co-residing children. Let $S_2 = P_1 + P_2 + P_3$.

In any year, S_1 should equal to S_2. In the numerical computation, however, S_1 and S_2 may not be exactly the same due to differences in estimation procedures. Therefore, an adjustment is needed to ensure that S_1 and S_2 are equal to each other. Following the harmonic mean approach, two equations must be satisfied:

$$C_1 \times a_1 + C_2 \times a_2 + C_3 \times a_3 = (2 \times S_1 \times S_2/(S_1 + S_2)), \tag{1}$$

$$P_1 \times b_1 + P_2 \times b_2 + P_3 \times b_3 = (2 \times S_1 \times S_2/(S_1 + S_2)), \tag{2}$$

where a_1, a_2, a_3, b_1, b_2, b_3 are adjustment factors to be estimated. Note that an adult's status of co-residence with children (i.e. status c) only includes her or his children, but not other nonindependent family members under the adult's care. This simplification causes a small underestimation of household size, as discussed in footnote 3. The number of events of leaving home and death computed from the perspective of children is slightly larger than the corresponding number computed from the perspective of parents. From the census base population, we can compute the total number of persons living with parents or other care provider (i.e. those with $k = 1$ or 2) from the perspective of children. We can also compute the total number of children living at home from the perspective of parents, based on the data of number of parents and their status of co-residence with children (i.e. c status). The ratio of these two total numbers from children's and parents' perspectives can be reasonably assumed to be stable and close to one.[6] Denote this ratio as R. The right side of Eqs. (1) and (2) should be multiplied by factors r_1 and r_2 respectively, where r_1 and r_2 satisfy the constraints that $r_1/r_2 = R$ and $r_1 + r_2 = 2$. Solving the equations of the constraints, we obtain formulas for estimating r_1 and r_2:

$$r_1 = 2 \times R/(R + 1), \tag{3}$$

[4] The definition of children here is relative to parents. For example, a person aged 60 and over is still a child if he or she lives with parents.

[5] When c status is reduced by i, i events are accounted for.

[6] This ratio is 1.025 in 1990 in China.

$$r_2 = 2/(R+1). \tag{4}$$

Taking into account the r_1 and r_2 adjustment factors, Eqs. (1) and (2) become:

$$C_1 \times a_1 + C_2 \times a_2 + C_3 \times a_3 = (2 \times S_1 \times S_2/(S_1 + S_2)) \times r_1, \tag{5}$$

$$P_1 \times b_1 + P_2 \times b_2 + P_3 \times b_3 = (2 \times S_1 \times S_2/(S_1 + S_2)) \times r_2. \tag{6}$$

Note that the following relations should be satisfied: $C_2 \times a_2 = P_3 \times b_3$ and $C_3 \times a_3 = P_2 \times b_2$. We have used the exact single-year age-specific number of cohort members at risk and the death probabilities to compute C_3 and P_3. Only the age of each cohort members is tracked; no detailed age information for the cohort members' parents and children are distinguished in the model. Consequently, we have to estimate age of parents when we compute C_2, and to estimate children's age when we compute P_2. Obviously, C_3 and P_3 are much more accurate than C_2 and P_2. Therefore, a_1, a_2, a_3, b_1, b_2, and b_3 are estimated as follows:

$$a_3 = 1; \quad b_3 = 1; \quad a_2 = P_3/C_2; \quad b_2 = C_3/P_2,$$

$$a_1 = ((2 \times S_1 \times S_2/(S_1 + S_2)) \times r_1 - C_2 \times a_2 - C_3 \times a_3)/C_1,$$

$$b_1 = ((2 \times S_1 \times S_2/(S_1 + S_2)) \times r_2 - P_2 \times b_2 - P_3 \times b_3)/P_1.$$

(3) *Consistency between births computed for the female and male populations.* Changes in parity and status of co-residence with children are computed for both female and male populations in our two-sex model. The total number of births computed based on the female population should equal the total number of births computed based on the male population; the number of married females with i children living together should equal the number of married males with i children living together, and the number of cohabiting females with i children living together should equal the number of cohabiting males with i children living together. Data on single-year age and parity specific fertility rates for male populations are rarely available. Therefore, we estimate males' birth rates based on females' birth rates and the average age difference between the male and female partners. Clearly, computation of births for the female population is more accurate than that for the males. Therefore, the number of births produced by the male population is adjusted to equal to numbers of births produced by the female population.

(4) *Consistency between the females' and males' status of co-residence with children before and after divorces (and dissolution of cohobation).* Children would stay either with mother or father after their parents' divorce or dissolution of cohabitation. Therefore, the number of children living together with mother or father immediately after the parents' divorce should equal the number before divorce. The living arrangement of the children of the divorced couples is a complicated social phenomenon, and data are sparse on this issue. We assume that if a couple has an odd number of children living together before divorce, the mother will have one more child than the father; if a couple has an even number of children living together before their divorce, both parties would have half of the children after divorce. In societies where divorced couples do not wish their children to be separated from each

other, our model has an option for the users to choose to assume all children stay with their mother after their parents' divorce.

(5) *Consistency between females' and males' status of co-residence with children before and after remarriage*. Children living with a single mother or single father all join the new family household after a parent's remarriage. A newly remarried couple's number of children should equal the sum of children living with either of the parties before remarriage. For simplicity, we assume that a remarried woman's or man's probability of having additional children from a new partner's previous union depends only on the frequency distribution of the status of co-residence with children of newly married men or women in the year.[7]

THE DEMOGRAPHIC ACCOUNTING EQUATIONS

Demographic accounting equations are used to compute the number of female and male persons and the changes in marital status, parity, and status of co-residence with parents and children in each projection year. The basic structure of all the accounting equations is:

> Number of persons age $x+1$ with status i at time $t+1$ = (number of persons age x with status i at time t) + (number of entries into state i which occur in the year $(t,t+1)$ among persons age $x+1$ at time $t+1$)–(number of exits out of state i which occur in the year $(t,t+1)$ among persons age $x+1$ at time $t+1$).

The number of events between age x and $x+1$ (and between years t and $t+1$) are calculated as the number of persons age x at risk of the occurrence of the events in the year multiplied by the probability of occurrence of the events between age x and $x+1$ (and between years t and $t+1$).

Calculation of the changes in k, m, p, and c statuses between age x and $x+1$ and between years t and $t+1$ consists of three steps as described below. Note that in the following discussion, the dimension "t" refers to the year, and dimension "s" refers to sex ($s=1$ for females and $s=2$ for males). If rural–urban is modeled, all stock variables and transition probabilities in this article should add one additional dimension r ($r=1$ for rural areas and $r=2$ for urban areas).

Step 1. Updating p and c statuses due to births occurring in the first half of the year:

$$\begin{aligned} N_{k,m,0,0}(x+0.5,t+0.5,s) &= N_{k,m,0,0}(x,t,s)(1-{}_{1/2}b_{0,m}(x,t,s)),\\ N_{k,m,p,c}(x+0.5,t+0.5,s) &= N_{k,m,p,c}(x,t,s)(1-{}_{1/2}b_{p,m}(x,t,s))\\ &\quad + N_{k,m,p-1,c-1}(x,t,s)\,{}_{1/2}b_{p-1,m}(x,t,s), \quad p>0, \end{aligned}$$

where $N_{k,m,p,c}(x,t,s)$ and $N_{k,m,p,c}(x+0.5,t+0.5,s)$ are the populations of age x, sex s and k, m, p, c statuses at beginning of year t and at middle of year t after

[7] We exclude the persons who are newly married for the first time with no pre-marital births from computing the frequency distribution of the maternal status of the newly married persons, since those young people are much less likely to choose a partner whose previous marriage was dissolved.

updating p and c statuses, respectively. ${}_{1/2}b_{p,m}(x, t, s)$ is the probability of having a birth of order $p+1$ in the first half of year t by persons of age x, sex s, parity p and marital status m. The formulas for estimating half-year age and parity specific birth probabilities based on whole-year age and parity specific data are derived and justified in Zeng (1991, pp. 61–63).

Step 2. Compute deaths, in- and out-migration, rural–urban migration (if it is included in the application), marital status transitions, and changes of co-residence with parents and children at middle of the year.

External migration (i.e., moving in and moving out of the country or region under study) is handled by multiplying the projected number of female and male in- and out-migrants by the standard age, sex, and marital-status specific frequency distribution, and then adding the immigrants to and subtracting the out-migrants from the population. If a distinction is made between rural and urban areas, the number of rural–urban migrants is determined by the projected proportion of the population that lives in urban areas in each year. We assume that the out-migrants' k, p, and c status distribution before migration are the same as nonmigrants with the same age, sex and marital status. The in-migrants are assumed to have the same p and c status distribution as nonmigrants with the same age, sex and marital status, but they are assumed to be not living with parents immediately after migration.

Procedures to compute marital status changes and deaths are well documented in the literature, see, for example, Willekens *et al.* (1982); Bongaarts (1987); Schoen (1988), and therefore not being detailed here.

Let $w_{ij}(x, t, s, m)$ denote the probability of transition from co-residence status with parents i at age x in year t to co-residence status with parents j at age $x+1$ in year $t+1$, for persons of sex s and marital status m. The events which cause transitions in the co-residence status are death of one or two parents, divorce of parents, remarriage of the nonmarried parent, leaving parental home, returning to parental home, and switching to and departing from the collective household status. We assume that these events are independent, so that the transition probabilities $w_{ij}(x, t, s, m)$ can be estimated based on the age-specific probabilities of death, divorce, remarriage, leaving and returning home, and switching to and departing from the collective household status (see Appendix). Hence,

$$N''_{i,m,p,c}(x+0.5, t+0.5, s) = \sum_{j=1}^{4} N'_{j,m,p,c}(x+0.5, t+0.5, s)\, w_{ji}(x, t, s, m),$$

where $N'_{k,m,p,c}(x+0.5, t+0.5, s)$ is the population at middle of year t after computing deaths, in- and out-migration, rural–urban migration (if it is included in the application) and marital status transitions. $N''_{k,m,p,c}(x+0.5, t+0.5, s)$ is the population at middle of year t after updating the k status.

To illustrate the accounting equations for computing the status of co-residence with children due to children's leaving or returning home or death and to simplify the presentation, we assume that the highest parity is 3 here (the calculation method will be the same when the highest parity is larger than 3). In year t, let $s_1(t)$, $s_2(t)$, and $s_3(t)$ be the probabilities that the one child, two children, or three children, who were living at home at the beginning of the year, will survive and live at home at the end of the year. Similarly, let $d_1(t)$, $d_2(t)$, and $d_3(t)$ denote the probabilities that the one child, both of the two children, and all of the three children will die or leave home during the year. Let $d_{12}(t)$, $d_{13}(t)$, and $d_{23}(t)$ be the probabilities that one of the two children, one of the three children, and two of the three children will die or leave home at the end of the year. Assuming that the events of leaving home and death are independent, we can estimate $s_1(t)$, $s_2(t)$, $s_3(t)$, $d_1(t)$, $d_2(t)$, $d_3(t)$, $d_{12}(t)$, $d_{23}(t)$, and $d_{13}(t)$ easily:

$$N'''_{k,m,0,0}(x+0.5,t+0.5,s) = N''_{k,m,0,0}(x+0.5,t+0.5,s),$$

when $p > 0$ and $p \geq c$,

$$\begin{aligned} N'''_{k,m,p,0}(x+0.5,t+0.5,s) &= N''_{k,m,p,0}(x+0.5,t+0.5,s) \\ &+ N''_{k,m,p,1}(x+0.5,t+0.5,s)d_1(t) + N''_{k,m,p,2}(x+0.5,t+0.5,s)d_2(t) \\ &+ N''_{k,m,p,3}(x+0.5,t+0.5,s)d_3(t), \end{aligned}$$

$$\begin{aligned} N'''_{k,m,p,1}(x+0.5,t+0.5,s) &= N''_{k,m,p,1}(x+0.5,t+0.5,s)s_1(t) \\ &+ N''_{k,m,p,2}(x+0.5,t+0.5,s)d_{12}(t) + N''_{k,m,p,3}(x+0.5,t+0.5,s)d_{23}t, \end{aligned}$$

$$\begin{aligned} N'''_{k,m,p,2}(x+0.5,t+0.5,s) &= N''_{k,m,p,2}(x+0.5,t+0.5,s)s_2(t) \\ &+ N''_{k,m,p,3}(x+0.5,t+0.5,s)d_{13}t, \end{aligned}$$

$$N'''_{k,m,p,3}(x+0.5,t+0.5,s) = N''_{k,m,p,3}(x+0.5,t+0.5,s)s_3t,$$

where $N'''_{k,m,p,c}(x+0.5,t+0.5,s)$ is the population at middle of year t after updating the c status due to children's deaths and leaving/returning home.

For the computation of family-household status changes at the highest open-ended age interval (at which we assume that no births, marriages, divorce, and migration and co-residence status changes would occur), see Zeng (1990, p. 89, p. 176).

Step 3. Updating p and c status due to births occurring in the second half of the year:

$$N_{k,m,0,0}(x+1,t+1,s) = N'''_{k,m,0,0}(x+0.5,t+0.5,s)(1-{}_{1/2}b_{0,m}(x+0.5,t+0.5,s)),$$

$$\begin{aligned} N_{k,m,p,c}(x+1,t+1,s) = {} & N'''_{k,m,p,c}(x+0.5,t+0.5,s)(1-{}_{1/2}b_{p,m}(x+0.5,t+0.5,s)) \\ & + N'''_{k,m,p-1,c-1}(x+0.5,t+0.5,s)_{1/2}b_{p-1,m}(x+0.5,t+0.5,s), \quad (p>0). \end{aligned}$$

A LIST OF ASSUMPTIONS WE HAVE MADE

To present a clearer picture of the nature of our model and to aid interpretation of the output of the model, we list the major assumptions.

Markovian assumption: Status transitions depend on age and the status occupied at the beginning of the single year interval, but are independent of duration in the status. More specifically, we assume that fertility depends on age, parity, and marital status. Mortality, first marriage, widowhood, divorce and remarriage depend on age, sex, and marital status, as does the probability that a child will leave his or her parents' home.

Homogeneity assumption: People with the same characteristics have the same status transition probabilities.[8]

Births occur throughout the *first half and the second half of the year*, and other status transitions and deaths occur at the *middle of the year*.

Parents may or may not live with one married child and his (or her) spouse and their unmarried children. *No married brothers or sisters live together.*

Multiple births in a single age interval for one woman are counted as independent single births.

Events are assumed to be *locally independent*. The events of a child's death and leaving home are independent. Events of deaths and births are independent. Deaths and marital status changes are independent of parity and number of children living at home. Events of death of one or two parents, divorce of parents, remarriage of the nonmarried parent, and leaving the parental home as well as returning home are independent.

If a couple has an odd number of children living together before divorce, the mother will have one more child than the father; if a couple has an even number of children living together, both parties will have half of their *children after divorce*; or the user can choose the option to assume that all children will live with their mother after their parents' divorce.

A remarried person's probability of having additional children from the new partner's previous union depend only on the frequency distribution of the status of co-residence with children of newly married persons in the same year.

DATA NEEDED

The following data are needed for the model:

(1) Base population derived from a census or a survey, and classified by age, sex, marital status, parity, number of children living at home, and co-residence with parents.

(2) Age, sex, (and marital-status, if possible) specific death rates or probabilities.

(3) Age and sex-specific rates or probabilities of first marriage, widowhood, divorce and remarriage;

[8]The homogeneity assumption can be relaxed by introducing more characteristics. For instance, the homogeneity assumption is less strong for a fertility model that considers age, parity and maternal status, than for a fertility model that takes account of age only. Since our family household projection model accounts for more characteristics of the population under study than most other demographic projection models, the Markovian and homogeneity assumptions in our model are less restrictive than in most other models of demographic projection.

(4) Age and parity-specific rates or probabilities of birth. If there is no parity related birth control in the study population, age-specific rates or probabilities of births can be used.

(5) Age and sex-specific rates or probabilities of leaving home.

(6) Age, sex (and marital status if possible) specific frequency distribution of out-migration to the rest of the world outside the country or region under study; age, sex (and marital status if possible) specific frequency distribution of immigration from the rest of the world to the country or region under study.

(7) Age, sex and marital status specific frequency distribution of rural–urban migration within the country or region under study if the option of rural–urban classification is chosen by the user.

Note that the age-specific data in (2)–(7) define standard schedules which shape the general age pattern of the corresponding demographic process. It is ideal to have the observed age-specific data from the country or region under study. When this is not available, however, one may borrow the standard schedules from another country or region where the general age pattern of demographic processes is similar to that in the country or region under study.

(8) Parameters to specify levels of parity-specific Total Fertility Rates, proportion eventually married, proportion not live with parents after becoming an adult, proportion eventually divorced for those married, proportion eventually remarried for those divorced or widowed, proportion of urban population (if rural–urban classification is identified), median age and inter-quartile range of parity-specific fertility, first marriage, and leaving home, life expectation at birth in the future projected years.

Note that if the option of rural–urban regional classification is chosen by the user in the application, the data specified in (1)–(6) and (8) are region-specific.

If the data are rates rather than the probabilities, our model and computer program convert the rates into probabilities based on the assumption of constant intensity in the single-year age interval.[9] Probabilities are estimated based on a general formula:

$$p = \exp(-hm),$$

where p is a probability, m is the rate given by the number of events divided by the person-years at risk in the age interval, and h is the length of the age interval.

The observed age-specific or age-parity-specific rates or probabilities define standard schedules. The standard schedules and the parameters that specify future levels of fertility, first marriage, divorce, remarriage, mortality, leaving home, and migration are used to project the corresponding age-specific probabilities in future years. This can be done by assuming that future age patterns are some function of the standard schedule, as in the Brass model life tables, or by choosing an appropriate schedule among a set of model schedules, as in the Coale–Demeny regional model life tables. Another possibility is to estimate the future demographic schedules based on a standard schedule and projected changes in the median age and the interquartile ranges (Zeng *et al.*, 1993).

[9] One may also employ the uniform distribution assumption, but the difference in outcome of the estimates between the constant intensity and uniform distribution assumptions is extremely small when the age interval is one year (Keilman and Gill, 1990).

It is important to note that, at this stage of our research only modest attention has been paid to incorporating the best available methods and data for justifying and projecting the demographic rates required by the model. It would be useful to do some sensitivity analysis that would indicate which rates are most important for which outcome. How would outcomes change if demographic rates changed? What would policy implications be? How important is heterogeneity? Would the overall outcomes of a projection with rural and urban sectors pooled be close to the weighted outcomes from a projection distinguishing the dimension and the dynamics of rural and urban sectors, or would they be biased in some way? Are there other sources of heterogeneity – e.g. regional or ethnic – that might be important? How could one jointly project or forecast economic and demographic variables? How could we introduce a probabilistic approach and confidence intervals in our household projection model? Much more work on these issues are planned for the next stage of our research.

AN ILLUSTRATIVE NUMERICAL APPLICATION: PROJECTING CHINESE FAMILY HOUSEHOLDS INTO THE NEXT CENTURY

(A) *Data resources.* Base population and other required data are derived from the 1% sample data tape of the 1990 census, the 1988 Two-Per-Thousand Fertility and Contraceptive Survey (SFPC, 1990), and the 1985 In-Depth Fertility Survey (see Zeng *et al.*, 1995, for more detailed resources, the estimates and discussion).

(B) *Assumptions.* To illustrate how such a multi-dimensional family household projection model can be used for policy analysis and academic studies, we present two major scenarios with the same medium fertility assumption, but with medium and low mortality assumptions. In these scenarios, the state space includes six statuses (0, 1, 2, 3, 4, 5+) of parity and number of children living at home, four marital statuses (single, married, widowed and divorced), four k statuses (live with two parents, live with one parent, not live with parents, live in a collective household), and rural–urban residence. The assumptions about parity-specific total fertility rates, median age at births at different parities, first marriage, divorce, remarriage, and leaving parental home for the two scenarios are the same. The medium fertility assumption, which is our educated guess, is adopted from Zeng and Vaupel (1989). It assumes that rural cohort fertility will gradually decline from its current level to 2.1 children per women, and that urban fertility will increase gradually from about 1.5 children per woman for the cohort aged 30 in 1990 to 1.8 children per woman in 2050. First marriage propensity[10] is assumed to remain at the current level. The medium mortality assumption, which allows for slow progress in reducing mortality rates over the next 60 years, is essentially the same as that used in conventional population projections for China, such as the projection by the United Nations (United Nations, 1989). The low mortality scenario assumes a more optimistic increase in life expectancy: it has been adopted from the analysis of Ogawa (1988). No marital status differentials in

[10]The term "propensity" in this article is defined as the cohort proportion eventually experiencing the event.

mortality are assumed since the necessary data are not available. The rural and urban propensities of divorce are assumed to increase from 6 and 10 percent in 1990 to 20 and 40 per cent in 2050 respectively. By the middle of the next century, the propensities of remarriage after divorce and widowhood are assumed to decline by 20 to 50 per cent in rural and urban areas. The rural and urban proportions of parents who have at least one married child but do not live with any of them are assumed to increase by 30 to 45 per cent. The proportion of the population residing in urban areas is assumed to increase from 26.5 per cent in 1990 to 70 per cent in 2050. Because this article is a methodological study rather than a substantive analysis, we only have the above mentioned simple and limited designation of the scenarios. Because the nature of this paper and the space limit, we do not present detailed tables which list assumptions of all parameters in different years for females and males in rural and urban areas. Readers who are interested in looking at those tables are referred to Zeng *et al.* (1995).

(C) *Output*. Our model and its associated computer software can produce a large number of output tables for rural and urban rectors in each of the projection years, including the age and sex distribution of population; per cent of children under age 18 or 15; per cent of elderly age 65–74, 75–84, and 85+; ratio of working age population to children and elderly population; cross-tabulations of distributions of age, sex, marital status, co-residence with two, one, or no parents, parity, and number of children living together; distribution of one-person households by age, sex, and marital status; distribution of one-couple households by age of wives or husbands; distribution of nuclear and three-generation households by household size, type, age, sex, marital status and the other characteristics of the reference persons, etc. These distributions can be presented in either absolute numbers or percentages.

Because space is limited and the main purpose of this numerical application is for illustration, only a very small portion of the output is presented, together with some straightforward analysis on how medium fertility, the assumed future marriage, divorce and leaving home trends, and medium *vs.* low mortality may affect future Chinese family-structure and population. We focus on the elderly.

Tables 5 and 6 present the projected per cent of the population that is elderly in different years up to 2050, by types of living arrangement. Some 29.9 and 24.0 per cent of the population in rural and urban areas in 2050 would be above age 65, if low mortality is achieved. Moreover, 30.8 million persons (6.5 per cent of the population) would be elderly and living alone in rural areas, and 65.2 million elderly persons (5.9 per cent of the population) would be living alone in urban areas, for a total of 96.0 million elderly living alone. Some 33.7 million of this group would be age 85 and older. Clearly, the middle of the next century will be a difficult time for the country due to serious problems of population ageing.

Table 7 presents the simulated percentage distribution of household size in future years under the two scenarios. It appears that Chinese household size will steadily decline, from an average of 4.1 persons in 1990, to 4.0 in 2000, 3.3 in 2020, and 2.7 in 2050 in the rural areas. In the urban areas, the average household size was 3.6 in 1990, and it would be 3.4 in 2020, and 2.8 in 2050. The urban household size is slightly larger than the rural one after 2020 because the expected massive rural–urban migration would bring many young people to the urban areas and leave their middle-aged or elderly parents living alone in rural areas. The per cent of one-person and two-person households would more than triple by the middle of the next century, and such small households would become the most common variety. Such a

TABLE 5

Total population and per cent of elderly among total population, by types of living arrangement, under the same assumption of fertility, marriage, divorce, and propensity of co-residence between parents and married children, but different mortality, Rural Areas

	1990	2000		2010		2020		2030		2040		2050	
		M	L	M	L	M	L	M	L	M	L	M	L
Total population (million)	837	820	823	770	778	699	713	631	652	539	566	443	474
65–74 years old													
living alone	0.34	0.39	0.39	0.47	0.45	0.97	0.88	1.36	1.14	1.86	1.44	1.73	1.23
w. spouse only	0.86	0.69	0.71	0.82	0.87	2.42	2.61	4.12	4.27	5.72	5.71	5.45	5.14
w. children +	2.82	3.80	3.85	4.29	4.42	5.52	5.71	5.09	5.53	5.55	6.35	4.23	5.16
Institution	0.02	0.03	0.03	0.04	0.03	0.08	0.07	0.12	0.10	0.18	0.14	0.16	0.12
Sub total	4.04	4.90	4.97	5.63	5.78	8.99	9.28	10.69	11.02	13.31	13.64	11.58	11.65
75–84 years old													
living alone	0.20	0.35	0.35	0.39	0.40	0.51	0.53	1.26	1.28	2.02	1.95	2.93	2.68
w. spouse only	0.21	0.28	0.30	0.27	0.31	0.40	0.51	1.33	1.76	2.59	3.47	3.76	5.07
w. children +	1.05	1.46	1.51	2.21	2.41	2.55	2.87	3.31	3.71	2.86	3.19	2.93	3.42
Institution	0.02	0.04	0.04	0.04	0.04	0.06	0.06	0.14	0.14	0.25	0.24	0.37	0.33
Sub total	1.48	2.13	2.20	2.90	3.16	3.52	3.96	6.05	6.88	7.73	8.85	9.99	11.50
85+ years old													
living alone	0.03	0.09	0.09	0.17	0.21	0.21	0.28	0.33	0.46	0.93	1.31	1.79	2.55
w. spouse only	0.01	0.01	0.02	0.03	0.04	0.03	0.07	0.08	0.14	0.29	0.61	0.63	1.50
w. children +	0.15	0.26	0.29	0.45	0.57	0.75	1.06	0.97	1.48	1.40	2.17	1.40	2.25
Institution	0.00	0.02	0.02	0.02	0.03	0.03	0.05	0.05	0.06	0.14	0.19	0.29	0.42
Sub total	0.20	0.38	0.42	0.68	0.85	1.04	1.45	1.42	2.15	2.76	4.29	4.10	6.72
65+ years old													
living alone	0.58	0.84	0.85	1.04	1.06	1.70	1.70	2.94	2.87	4.82	4.71	6.44	6.45
w. spouse only	1.08	0.98	1.02	1.12	1.24	2.86	3.19	5.52	6.16	8.60	9.79	9.85	11.72
w. children +	4.03	5.52	5.65	6.95	7.39	8.83	9.64	9.38	10.71	9.81	11.71	8.56	10.82
Institution	0.04	0.08	0.08	0.10	0.10	0.16	0.17	0.31	0.31	0.56	0.57	0.83	0.87
Sub total	5.73	7.41	7.59	9.20	9.80	13.55	14.69	18.16	20.05	23.80	26.78	25.67	29.87

Notes: M: Medium mortality scenario; L: Low mortality scenario; w. children +: living with children and others; w. spouse only: living with spouse only; institution: living in collective household or nursing house, etc., which are estimated based on the age-marital status-specific proportions of elderly who were

TABLE 6

Population size and per cent of elderly among total population, by types of living arrangement, under the same assumption of fertility, marriage, divorce, and propensity of co-residence between parents and married children, but different mortality, Urban Areas

	1990	2000		2010		2020		2030		2040		2050	
		M	L	M	L	M	L	M	L	M	L	M	L
Total population (million)	294	461	463	605	611	757	772	872	900	959	1007	1033	1105
65–74 years old													
living alone	0.29	0.33	0.32	0.48	0.45	1.04	0.93	1.57	1.37	2.22	1.91	1.88	1.57
w. spouse only	0.86	0.80	0.82	1.22	1.30	3.02	3.23	4.51	4.80	6.01	6.45	4.94	5.23
w. children +	2.41	3.12	3.16	2.71	2.80	2.31	2.43	2.24	2.44	3.47	3.67	2.70	2.89
Institution	0.03	0.03	0.03	0.05	0.05	0.12	0.11	0.19	0.17	0.27	0.24	0.23	0.19
Sub total	3.59	4.28	4.34	4.46	4.59	6.49	6.70	8.51	8.78	11.98	12.28	9.75	9.88
75–84 years old													
living alone	0.18	0.26	0.26	0.31	0.31	0.48	0.49	1.15	1.13	1.78	1.72	2.51	2.35
w. spouse only	0.21	0.25	0.27	0.30	0.35	0.53	0.67	1.49	1.93	2.43	3.20	3.39	4.54
w. children +	0.91	1.08	1.12	1.60	1.73	1.41	1.57	1.19	1.33	1.12	1.22	1.82	2.05
Institution	0.03	0.05	0.05	0.06	0.06	0.11	0.11	0.25	0.24	0.40	0.37	0.55	0.50
Sub total	1.32	1.64	1.70	2.26	2.46	2.52	2.84	4.09	4.63	5.72	6.51	8.27	9.44
85 + years old													
living alone	0.03	0.07	0.07	0.12	0.14	0.16	0.22	0.30	0.41	0.80	1.10	1.37	1.96
w. spouse only	0.01	0.01	0.02	0.03	0.04	0.05	0.08	0.09	0.19	0.33	0.68	0.60	1.36
w. children +	0.16	0.21	0.23	0.33	0.41	0.53	0.74	0.52	0.81	0.51	0.78	0.52	0.80
Institution	0.00	0.02	0.02	0.03	0.04	0.05	0.05	0.08	0.11	0.23	0.31	0.41	0.55
Sub total	0.21	0.32	0.35	0.50	0.64	0.78	1.09	1.00	1.52	1.86	2.86	2.91	4.67
65 + years old													
living alone	0.50	0.65	0.65	0.90	0.91	1.69	1.64	3.02	2.92	4.80	4.73	5.76	5.88
w. spouse only	1.08	1.07	1.12	1.54	1.69	3.59	3.97	6.10	6.92	8.76	10.34	8.93	11.14
w. children +	3.48	4.41	4.52	4.63	4.95	4.24	4.74	3.96	4.58	5.09	5.67	5.04	5.73
Institution	0.07	0.10	0.10	0.15	0.15	0.27	0.27	0.53	0.52	0.90	0.91	1.21	1.25
Sub total	5.13	6.23	6.38	7.22	7.69	9.79	10.63	13.60	14.93	19.56	21.65	20.93	24.00

Notes: The same as in Table 5.

TABLE 7

Percentage distribution of family household sizes, under the same assumption of fertility, marriage, divorce, and propensity of co-residence between parents and married children, but different mortality

	1990	2000		2010		2020		2030		2040		2050	
		M	L	M	L	M	L	M	L	M	L	M	L
Rural areas													
1 person	6.30	8.53	8.46	9.33	9.13	11.74	11.26	16.10	15.24	21.85	20.80	26.72	25.66
2 person	8.37	9.45	9.48	14.84	14.93	21.30	21.58	26.42	26.92	29.22	30.10	30.32	31.73
3 person	19.31	16.42	16.33	20.98	20.92	22.96	22.93	20.72	20.69	17.29	17.10	14.29	13.87
4 person	27.20	28.47	28.42	26.39	26.26	23.15	22.97	19.81	19.62	16.79	16.46	14.95	14.48
5 person	20.34	21.94	21.98	17.83	17.86	13.68	13.72	10.96	11.05	9.59	9.69	8.71	8.74
6 person	10.28	9.72	9.79	7.53	7.69	5.73	5.99	4.85	5.22	4.28	4.69	4.04	4.38
7+ person	8.20	5.47	5.54	3.09	3.22	1.44	1.55	1.13	1.27	0.99	1.16	0.98	1.13
Average size	4.14	3.98	3.99	3.64	3.65	3.29	3.31	3.02	3.05	2.80	2.83	2.65	2.67
Urban areas													
1 person	6.30	10.37	10.28	10.33	10.04	13.71	13.14	17.56	16.77	22.05	21.27	25.40	24.75
2 person	10.90	12.02	12.07	18.79	19.05	23.16	23.67	25.95	26.81	27.25	28.93	27.79	29.88
3 person	35.22	23.31	23.19	14.39	14.20	11.48	11.21	10.91	10.51	10.81	10.05	10.17	9.28
4 person	24.16	33.02	33.07	31.63	31.70	27.93	27.95	24.82	24.64	22.28	21.82	20.96	20.34
5 person	13.79	9.96	9.94	9.31	9.11	9.83	9.60	9.01	8.88	8.19	7.90	7.51	7.16
6 person	5.68	9.58	9.68	13.94	14.24	12.39	12.83	10.26	10.76	8.13	8.60	6.98	7.28
7+ person	3.94	1.74	1.76	1.61	1.65	1.50	1.59	1.50	1.63	1.28	1.43	1.20	1.32
Average size	3.63	3.57	3.57	3.60	3.61	3.41	3.43	3.19	3.22	2.97	2.98	2.83	2.83

large increase in one-person and two-person households can be attributed to more young people living away from their parents, rapidly increasing divorce rates, and more elderly living alone or living with spouse only. The differences between the medium and low mortality scenarios are small, because lower mortality has various effects that balance out.

Projected percentage distributions of family-household types are presented in Tables 8 and 9. One-couple households in the middle of next century would be about three times as frequent as in 1990. This substantial rise is mainly due to an increase in the number of elderly couples living with no children. The frequency of three-generation family households in the year 2000 in rural areas would be more or less the same as that in 1990, and then steadily decline to about 14 per cent in 2050. In urban areas, the frequency of three-generation family households would fall from 16.8 per cent in 1990 to about 10 per cent in years 2040 and 2050.

The results for the year 2000 may seem puzzling: the frequency of three-generation family households in rural areas does not decrease, even though we have assumed a substantial decrease in propensity of co-residence between parents and married children. A plausible explanation is that the children who were born after the late 1970s (when fertility levels fell to half the levels in the 1960s and earlier) will reach family formation stage around the end of this century: these children have fewer siblings than in earlier generations. Given the fact that a certain (albeit declining) proportion of Chinese parents and one of their married children may still wish to live together, the reduction of the number of siblings will result in a smaller proportion of young couples who can move out of their parental home to form nuclear households. (For a more detailed discussion, see Zeng, 1986.)

Single-parent households would increase from 5.7 per cent in 1990 to 6.3 per cent in 2000, 9.8 per cent in 2020, and 11.3 per cent in 2050 in rural areas, under the medium mortality scenario. Single-parent family households in urban areas would also substantially increase: 4.5 per cent in 1990, 5.9 per cent in 2000, 10.4 per cent in 2020, and 14.9 per cent in 2050. Under the low mortality assumption, the per cent of single-parent family households would be slightly lower than that under the medium mortality scenario. The percent of single parent households observed in the 1990 census was higher in rural areas than in urban areas, and this differential would be reversed by the middle of next century, because a substantially higher divorce level and a lower remarriage level in the urban areas are assumed in the future.

This illustrative application shows that the rapid fall in births in China in recent decades, together with the age difference at first marriage, and the increase in sex ratio at birth (Zeng *et al.*, 1993) in recent years will create tensions in the marriage market, namely, more males who are eligible for marriage than their female counterparts, in the next few decades. For example, given the consistency requirement in our two-sex projection model, marriage propensities for men will have to be adjusted downwards to 0.96 in 2000, 0.92 in 2030, and 0.92 in 2050.[11] This implies that there

[11] The age-specific marriage probabilities for women would have to be adjusted upwards by about 16 per cent, which implies that the Chinese women would have to marry earlier due to the pressure of marriage squeeze. After the adjustment, the life table proportion of eventually marry for women would remain being very close to one, that is almost the same as before adjustment, because the universal marriage for women in China has been prevalent for a very long time (see for example, Coale, 1984).

TABLE 8

Percentage distribution of family household types, under the same assumption of fertility, marriage, divorce, and propensity of co-residence between parents and married children, but different mortality, rural areas

	1990	2000		2010		2020		2030		2040		2050	
		M	L	M	L	M	L	M	L	M	L	M	L
One person	6.30	8.53	8.46	9.33	9.13	11.74	11.26	16.10	15.24	21.85	20.80	26.72	25.66
One couple	6.91	7.23	7.31	11.29	11.54	16.99	17.47	21.68	22.43	24.91	25.87	26.79	28.21
2-generation													
Husband-wife	63.01	60.29	60.29	57.08	57.22	49.26	49.42	41.40	41.53	33.91	33.65	28.61	27.90
Lone mother	2.41	2.87	2.76	3.97	3.70	4.12	3.92	3.78	3.68	2.93	3.06	2.16	2.37
Lone father	1.89	1.61	1.56	1.99	1.84	2.06	1.83	2.36	1.97	2.34	1.97	2.05	1.66
Sub. total	67.31	64.77	64.61	63.04	62.76	55.44	55.17	47.54	47.19	39.18	38.67	32.81	31.93
3-generation													
MG H-W and 1P	9.96	6.81	6.79	4.80	4.64	4.48	4.20	4.52	4.24	4.77	4.47	4.90	4.59
MG H-W and 2P	8.88	11.83	12.02	10.79	11.19	10.57	11.14	9.31	10.09	8.37	9.27	7.75	8.59
MG L. M and 1P	0.17	0.25	0.24	0.20	0.19	0.18	0.17	0.20	0.18	0.24	0.23	0.27	0.26
MG L. M and 2P	0.04	0.23	0.23	0.25	0.25	0.26	0.26	0.26	0.27	0.26	0.28	0.25	0.29
MG L. F and 1P	0.33	0.18	0.18	0.13	0.13	0.13	0.11	0.15	0.13	0.19	0.16	0.24	0.20
MG L. F and 2P	0.09	0.16	0.16	0.18	0.18	0.20	0.20	0.23	0.23	0.24	0.25	0.26	0.27
Sub total	19.48	19.47	19.61	16.34	16.57	15.82	16.09	14.67	15.15	14.07	14.66	13.68	14.19
SP amo. 2-gen	6.39	6.91	6.69	9.45	8.83	11.15	10.42	12.91	11.98	13.47	12.98	12.82	12.63
MSP amo. 3-gen	3.26	4.23	4.12	4.65	4.47	4.89	4.68	5.70	5.43	6.59	6.26	7.52	7.15
SP amo. 2/3-gen	5.69	6.29	6.09	8.46	7.92	9.76	9.12	11.21	10.39	11.65	11.14	11.26	10.94

Notes: M: Medium mortality scenario; L: Low mortality scenario; H-W: husband and wife; L. M: Lone mother; L. F: Lone father; MG: middle generation; 1P: one grandparent; 2P: two grand parents.

SP amo. 2-gen: single parent households among two-generation households; SP amo. 3-gen: single parent households among three-generation households; SP amo. 2/3-gen: single parent households among all two-generation and three-generation households.

TABLE 9

Percentage distribution of family household types, under the same assumption of fertility, marriage, divorce, and propensity of co-residence between parents and married children, but different mortality, Urban Areas

	1990	2000		2010		2020		2030		2040		2050	
		M	L	M	L	M	L	M	L	M	L	M	L
One person	6.30	10.37	10.28	10.33	10.04	13.71	13.14	17.56	16.77	22.05	21.27	25.40	24.75
One Couple	9.49	8.86	8.96	14.78	15.12	18.82	19.40	21.17	22.20	22.26	24.24	22.61	25.07
2-generation													
Husband–wife	64.06	63.19	63.19	58.09	58.12	50.70	50.74	44.86	44.60	39.43	38.48	35.26	33.87
Lone Mother	2.39	2.73	2.65	3.43	3.30	3.61	3.49	3.61	3.45	3.48	3.28	3.35	3.15
Lone Father	0.91	1.37	1.34	2.10	2.03	2.55	2.45	2.91	2.73	3.04	2.75	3.10	2.73
Sub. Total	67.37	67.28	67.19	63.62	63.44	56.85	56.68	51.38	50.77	45.95	44.52	41.71	39.74
3-generation													
MG H–W and 1P	8.31	4.14	4.11	2.93	2.81	2.78	2.62	2.89	2.73	3.12	2.91	3.43	3.17
MG H–W and 2P	8.05	8.71	8.83	7.62	7.86	6.98	7.32	6.01	6.52	5.50	5.93	5.52	5.94
MG L. M and 1P	0.20	0.16	0.16	0.14	0.13	0.15	0.14	0.18	0.17	0.23	0.22	0.29	0.27
MG L. M and 2P	0.05	0.23	0.23	0.27	0.27	0.29	0.30	0.31	0.33	0.32	0.35	0.36	0.40
MG L. F and 1P	0.17	0.09	0.09	0.10	0.09	0.13	0.12	0.18	0.16	0.23	0.21	0.29	0.26
MG L. F and 2P	0.05	0.15	0.15	0.22	0.23	0.28	0.29	0.32	0.34	0.33	0.36	0.38	0.41
Sub. Total	16.84	13.49	13.57	11.28	11.40	10.62	10.79	9.89	10.26	9.73	9.97	10.27	10.44
SP amo. 2-gen	4.91	6.09	5.94	8.69	8.39	10.82	10.47	12.69	12.16	14.20	13.56	15.46	14.79
MSP amo. 3-gen	2.83	4.72	4.65	6.49	6.39	8.01	7.91	10.05	9.84	11.43	11.39	12.81	12.84
SP amo. 2/3-gen	4.49	5.86	5.73	8.36	8.08	10.38	10.06	12.26	11.77	13.72	13.16	14.94	14.38

Notes: The same as in Table 8.

will be about 8 per cent of men who wish to marry but are not able to do so due to shortage of eligible females in the next century.[12] In our illustrative application, we assumed that the sex ratio at birth in 1990[13] remained unchanged until 2050. If the sex ratio at birth in China continues to increase in the future, tensions in the marriage market will be more serious. This policy-relevant problem certainly deserves more research.

CONCLUDING REMARKS

As Keyfitz (1985) and Bongaarts (1983) observed, family demography had been a difficult and underdeveloped field. Our multi-dimensional model contributes toward the development of methods for projecting family structures in the following ways.

First, the model permits projection of many characteristics of family households and their members, using ordinary demographic data that are usually available from conventional data sources in most developed countries and some developing countries. When the necessary data to establish the standard demographic schedules for the population under study (e.g., in a small area or a developing country with poor data resources) are not available, a standard from another country or region with similar demographic conditions can be used. The user can then project summary measures, such as life expectancy, cohort total rates and median age of first marriage, total fertility rates by birth order, proportion of co-residence between parents and married children, proportion eventually divorced, proportion eventually remarried after divorce and widowhood, etc. We therefore expect that it will be relatively easy to apply the model.

Second, unlike the traditional headship-rate method, our model can closely link the projected family household and its members' characteristics with future demographic rates, so that the model can be used for policy analysis and academic studies in exploring how future demographic changes may affect family households.

Third, the model includes both nuclear and three-generation family households, so that it can be used to project family households in western countries where only nuclear family households are dominant, and in Asian countries as well as some other developing countries where nuclear and three-generation family households are both important.

There is a growing need for longer-term projections, thirty or fifty years or even a century into the future. Uncertainties about developments so far in the future are so immense that such projections should not be interpreted as forecasts. Nonetheless, long-term projections may be useful to policy analysts as scenarios that can be used to compare the relative effects of alternative policies. For instance, long-term projections provided insights into the relative impact on population ageing of a one-child *vs.* a two-child plus spacing policy in China (Vaupel and Zeng, 1991). The model

[12]This finding is consistent with what is discovered by Tuljapurkar *et al.* (1995, p. 874) by using projected sex ratio of potential partners in the first half of the next century in China.

[13]The sex ratio at birth under normal conditions is about 106 (106 male babies per 100 female babies). In 1990, the sex ratio at birth reported by a large sample survey on births delivered in the hospitals where possibility of the under-reporting of females could be ruled out are 111.9 in rural areas and 108.9 in urban areas (Zeng *et al.*, 1993).

presented in this article is useful, as demonstrated by the illustrative application, for such scenario analysis of the long-term consequences of alternative policy-directions.

Most planners and analysts in governmental agencies and business firms are primarily interested in short-term forecasting of trends over the next five to ten years or so. They need forecasts that are as accurate as possible, forecasts that capture the actual details of future events. Unfortunately, however, even a careful demographic forecast may not necessarily produce precise results over ten-year time frames because the uncertainties of the demographic parameters and the associated socio-economic variables. Furthermore, a forecast is nearly certain to be wrong if it consists of a single number; if it consists of a range with probability attached, it can be correct in the sense that the range straddles the subsequent outcome (Keyfitz, 1985, p. 204). We believe that the model presented here has the potential of performing satisfactorily for such shorter-term forecasts, even though it involves a substantially richer set of variables than most demographic projections. Of course the method requires further development, including how to estimate probabilistic confidence intervals.

ACKNOWLEDGEMENTS

The research reported in this paper was supported by the National Natural Science Foundation of China, the Sloan Foundation and the Dutch Foundation for Promotion of Cultural Exchanges between The Netherlands and China. Part of the work was carried out while the first author was a visiting professor at the Center for Health and Social Policy (CHS) of Odense University, the Center for Population Analysis and Policy (CPOP) of the Humphrey Institute at the University of Minnesota, and Netherlands Interdisciplinary Demographic Institute (NIDI). The authors wish to acknowledge the comments, suggestions, and encouragement provided by Professor F. Willekens. Discussions with Professors Nico Keilman, E.A. Hammel, K.W. Wachter, and Thomas Pullum, and very helpful comments by the three anonymous referees, are also highly appreciated. The authors are solely responsible for any errors that may remain.

APPENDIX A

PROCEDURES TO ESTIMATE TRANSITION PROBABILITIES OF STATUS OF CO-RESIDENCE WITH PARENTS

Let $w_{ij}(x, t, s, m)$ denote the probability of transition from co-residence status i at age x in year t to j at age $x+1$ in year $t+1$ for persons of sex s and marital status m, where i (= 1,2,3,4) and j (= 1,2,3,4). Denote the probabilities of death of an x-year old person's mother and father as $q_m(x, t)$ and $q_f(x, t)$; and probabilities of divorce of an x-year old person's mother and father in year t as $d_m(x, t)$ and

$d_f(x,t)$:

$$q_m(x,t) = \sum_{i=\alpha}^{49} q_1(x+i,t)f_1(i); \quad q_f(x,t) = \sum_{i=\alpha}^{49} q_2(x+z+i,t)f_2(i);$$

$$d_m(x,t) = \sum_{i=\alpha}^{49} d_1(x+i,t)f_1(i); \quad d_f(x,t) = \sum_{i=\alpha}^{49} d_2(x+z+i,t)f_2(i);$$

where $q_1(x,t)$ and $q_2(x,t)$ are female and male death probabilities in year t; $d_1(x,t)$ and $d_2(x,t)$ are female and male divorce probabilities in year t; z is the average age difference between the male and female partners; and $f_1(i)$ and $f_2(i)$ are the frequency distributions of the product of age-specific fertility rate and conditional survival probabilities:

$$f_1(i) = (b(i)l_1(x+i)/l_1(i)) / \sum_{i=\alpha}^{49} (b(i)l_1(x+i)/l_1(i)),$$

$$f_2(i) = (b(i)l_2(x+i)/l_2(i)) / \sum_{i=\alpha}^{49} (b(i)l_2(x+i)/l_2(i)),$$

where $b(i)$ are age-specific fertility rates, and $l_1(x)$ and $l_2(x)$ are female and male survival probabilities of surviving from age 0 to x. Ideally $b(i)$, $l_1(x)$ and $l_2(x)$ should be based on cohort data, but it will not introduce significant bias if one employs period data since the frequency distribution rather than the fertility and mortality level is used.

The events which cause transitions of the co-residence status from 1 to 2 are the death of one of the parents or divorce of the parents. If the death of one parent occurs first, divorce cannot occur. However, divorce may precede death. Therefore,

$$\begin{aligned} w_{12}(x,t,s,m) = {} & q_m(x,t) + q_f(x,t) + d(x,t) - q_m(x,t)q_f(x,t) \\ & - q_m(x,t)d(x,t)/2 - q_f(x,t)d(x,t)/2, \end{aligned} \tag{A.1}$$

where $d(x,t) = (d_m(x,t) + d_f(x,t))/2$.

The events which cause change in k status from 1 to 3 are an x-year old person's leaving parental home or death of both parents. If the deaths of both parents occur first, the event of leaving parental home cannot occur. However, a person can leave home before either parent dies or after one of them dies. Therefore,

$$w_{13}(x,t,s,m) = l(x,t,s,m) + q_m(x,t)q_f(x,t)(1 - (2/3)l(x,t,s,m)), \tag{A.2}$$

where $l(x,t,s,m)$ is probability of leaving parental home at age x in year t for persons of sex s and marital status m.

The events which cause transition of k status from 2 to 3 are the death of the nonmarried parent or an x-year old person's leaving parental home. If the death of the lone parent occurs first, the event of leaving parental home cannot occur.

Therefore,

$$w_{23}(x,t,s,m) = l(x,t,s,m) + q(x,t) - (l(x,t,s,m)q(x,t))/2, \tag{A.3}$$

where $q(x,t) = (q_m(x,t) + q_f(x,t))/2$.

The event that causes transition of k status from 2 to 1 is the remarriage of the nonmarried parent. Denote $r_{d1}(x,t)$ and $r_{d2}(x,t)$ as divorced female and male remarriage probabilities in year t; and $r_{w1}(x,t)$ and $r_{w2}(x,t)$ as widowed female and male remarriage probabilities in year t. Then,

$$w_{21}(x,t,s,m) = \left(\sum_{i=\alpha}^{49} r_{d1}(x+i,t)f_1(i)\right) g_{d1}(x) + \left(\sum_{i=\alpha}^{49} r_{d2}(x+z+i,t)f_2(i)\right) g_{d2}(x+z)$$
$$+ \left(\sum_{i=\alpha}^{49} r_{w1}(x+i,t)f_1(i)\right) g_{w1}(x) + \left(\sum_{i=\alpha}^{49} r_{w2}(x+z+i,t)f_2(i)\right) g_{w2}(x+z), \tag{A.4}$$

where,

$$g_{d1}(x) = \sum_{i=\alpha}^{49} N_{d1}(x+i) \Big/ \sum_{i=\alpha}^{49} [N_{d1}(x+i) + N_{d2}(x+z+i) + N_{w1}(x+i) + N_{w2}(x+z+i)],$$

$$g_{d2}(x+z) = \sum_{i=\alpha}^{49} N_{d2}(x+z+i) \Big/ \sum_{i=\alpha}^{49} [N_{d1}(x+i) + N_{d2}(x+z+i) + N_{w1}(x+i) + N_{w2}(x+z+i)],$$

$$g_{w1}(x) = \sum_{i=\alpha}^{49} N_{w1}(x+i) \Big/ \sum_{i=\alpha}^{49} [N_{d1}(x+i) + N_{d2}(x+z+i) + N_{w1}(x+i) + N_{w2}(x+z+i)],$$

$$g_{w2}(x+z) = \sum_{i=\alpha}^{49} N_{w2}(x+z+i) \Big/ \sum_{i=\alpha}^{49} [N_{d1}(x+i) + N_{d2}(x+z+i) + N_{w1}(x+i) + N_{w2}(x+z+i)],$$

where $N_{d1}(x+i)$, $N_{d2}(x+z+i)$, $N_{w1}(x+i)$, and $N_{w2}(x+z+i)$ are the number of divorced females, divorced males, widowed females, widowed males, aged $x+i$ or $x+z+i$, all living with at least one child in year t.

The event that causes transition of k status from 3 to 1, from 3 to 2, from 4 to 1, and from 4 to 2 is an x-year old person's returning home to join her or his two parents or one nonmarried parent, so that

$$w_{31}(x,t,s,m) = h(x,t,s,m)N_{k1}(x,t,s,m)/[N_{k1}(x,t,s,m) + N_{k2}(x,t,s,m)], \tag{A.5}$$
$$w_{32}(x,t,s,m) = h(x,t,s,m)N_{k2}(x,t,s,m)/[N_{k1}(x,t,s,m) + N_{k2}(x,t,s,m)], \tag{A.6}$$
$$w_{41}(x,t,s,m) = h(x,t,s,m)N_{k1}(x,t,s,m)/[N_{k1}(x,t,s,m) + N_{k2}(x,t,s,m)], \tag{A.7}$$
$$w_{42}(x,t,s,m) = h(x,t,s,m)N_{k2}(x,t,s,m)/[N_{k1}(x,t,s,m) + N_{k2}(x,t,s,m)], \tag{A.8}$$

where $h(x,t,s,m)$ is the probability of returning home between age x and $x+1$ to join parental home in year t, for persons of sex s and marital status m. $N_{k1}(x,t,s,m)$ and $N_{k2}(x,t,s,m)$ are number of x-year old persons of sex s and marital status m, who are living with two parents and one parent, respectively.

The event that causes transition of k status from 4 to 3 is an x-year old person's switching from collective household status to a private household status and not living with parents:

$$w_{43}(x,t,s,m) = u(x,t,s,m), \tag{A.9}$$

where $u(x,t,s,m)$ is the probability of switching from collective household status to private household status and not living with parents between age x and $x+1$ in year t, for persons of sex s and marital status m.

The event that causes transition of k status from 1 to 4, from 2 to 4, and from 3 to 4 is an x-year old person's switching to collective household status from private household status. Hence

$$w_{14}(x,t,s,m) = v(x,t,s,m), \tag{A.10}$$

$$w_{24}(x,t,s,m) = v(x,t,s,m), \tag{A.11}$$

$$w_{34}(x,t,s,m) = v(x,t,s,m), \tag{A.12}$$

where $v(x,t,s,m)$ is probability of switching to collective household status from private household status between age x and $x+1$ in year t, for persons of sex s and marital status m.

Note that $l(x,t,s,m)$, $h(x,t,s,m)$, $u(x,t,s,m)$, and $v(x,t,s,m)$ can be parity-status specific if the data are available.

In addition:

$$w_{11}(x,t,s,m) = 1 - w_{12}(x,t,s,m) - w_{13}(x,t,s,m) - w_{14}(x,t,s,m), \tag{A.13}$$

$$w_{22}(x,t,s,m) = 1 - w_{21}(x,t,s,m) - w_{23}(x,t,s,m) - w_{24}(x,t,s,m), \tag{A.14}$$

$$w_{33}(x,t,s,m) = 1 - w_{31}(x,t,s,m) - w_{32}(x,t,s,m) - w_{34}(x,t,s,m), \tag{A.15}$$

$$w_{44}(x,t,s,m) = 1 - w_{41}(x,t,s,m) - w_{42}(x,t,s,m) - w_{43}(x,t,s,m). \tag{A.16}$$

REFERENCES

Bell, M. and Cooper, J. (1990) Household forecasting: replacing the headship rate model. Paper presented at the Fifth National Conference, Australian Population Association, Melbourne, November.

Bongaarts, J. (1983) The formal demography of family and households: an overview. *IUSSP Newsletter*, no. 17 (November).

Bongaarts, J. (1987) The projection of family composition over the life course with family status life tables. In: Bongaarts, J., Burch, T. and Wachter, K.W. (eds.) *Family Demography: Methods and Applications*. Oxford: Clarendon Press. pp. 189–212.

Brass, W. (1983) The formal demography of the family: an overview of the proximate determinants. In: *The Family*, Occasional paper 31. Office of Population Censuses and Surveys, London.

Coale, A.J. (1984) *Rapid Population Change in China, 1952–1982*. Committee on Population and Demography, National Research Council, report no. 27, Washington D.C.: National Academy Press.

Hammel, E.A., Hutchinson, D.W., Wachter, K.W., Lundy, R.T. and Deuel, R.Z. (1976) The SOCSIM demographic-sociological microsimulation program, Institute of International Studies, Research Series No. 27, University of California, Berkeley.

Van Imhoff, E. and Keilman, N. (1992) *LIPRO 2.0: An Application of A Dynamic Demographic Projection Model To Household Structure in The Netherlands.* Amsterdam: Swets and Zeitlinger Publisher.

Keilman, N. (1985) Nuptiality models and two-sex problem in national population forecasts. *European Journal of Population* **1**: 207–235.

Keilman, N. (1988) Dynamic household models. In: Keilman, N., Kuijsten, A. and Vossen, A. (eds.) *Modelling Household Formation and Dissolution.* Oxford: Clarendon Press.

Keilman, N. and Gill, R. (1990) On the estimation of multidimensional demographic models with population registration data. *Mathematical Population Studies* **2**, 2: 119–143.

Keyfitz, N. (1985) *Applied Mathematical Demography (Second Edition)*, New York: Springer-Verlag.

Land, K.C. and Rogers, A. (eds.) (1982) *Multidimensional Mathematical Demography.* New York: Academic Press.

Ledent, J. (1992) Vers des perspectives de famies/ménages sur la base d'un modèle de type multidimensional, unpublished manuscript.

Mason, A. and Racelis, R. (1992) A comparison of four methods for projecting households. *International Journal of Forecasting* **8**: 509–527.

Murphy, M.J. (1991) Modelling households: a synthesis. In: Murphy, M.J. and Hobcraft, J. (eds.) *Population Research in Britain,* A supplement to population studies, Vol. 45.

Nelissen, J.H. (1991) Household and education projections by means of a microsimulation model. *Economic Modelling* **8**: 480–511.

Pollak, R.A. (1990) Two-sex demographic models. *Journal of Political Economy* **98**, 2: 399–420.

Pollard, J.H. (1977) The continuing attempt to incorporate both sexes into marriage analysis. Proceedings of International Population Conference: Mexico 1977 (IUSSP, Liege), pp. 291–310.

Rogers, A. (1975) *Introduction to Multi-regional Mathematical Demography,* New York: John Wiley and Sons.

Schoen, R. (1981) The harmonic mean as the basis of a realistic two-sex marriage model. *Demography* **18**, 2: 201–216.

Schoen, R. (1988) *Modelling Multi-Group Population.* New York: Plenum Press.

Smith, J. (1987) Computer simulation of kin sets and kin counts. In: Bongaarts, J., Burch, T. and Wachter, K.W. (eds.) *Family Demography: Methods and Applications.* Oxford: Clarendon Press, pp. 249–266.

Spicer, K., Diamond, I. and Bhrolchain, M.N. (1992) Into the twenty-first century with British households. *International Journal of Forecasting* **8**: 529–539.

SFPC (State Family Planning Commission) (1990) *Data Collections of the Two-Per-Thousand Fertility and Contraceptive Survey.* Beijing.

Tuljapurkar, S., Nan Li, Feldman, M.W. (1995) High sex ratios in China's future. *Science* **267**: 874–866, 10 February.

United Nations (1989) *World Population Prospects, 1988.* Population Studies No. 106, Department of International Economic and Social Affairs. New York.

Vaupel, J. and Zeng Yi (1991) Population tradeoffs in China. *Policy Science* **24**: 389–406.

Wachter, K.W. (1987) Microsimulation of household cycles. In: Bongaarts, J., Burch, T. and Wachter, K.W. (eds.) *Family Demography: Methods and Applications.* Oxford: Clarendon Press, pp. 215–227.

Willekens, F.J., Shah, I., Shah, J.M. and Ramachandran, P. (1982) Multistate analysis of marital status life tables: theory and application. *Population Studies* **36**, 1: 129–144.

Willekens, F.J. (1987) The marital status life table. In: Bongaarts, J., Burch, T. and Wachter, K.W. (eds.) *Family Demography: Methods and Applications.* Oxford: Clarendon Press, pp. 125–149.

Zeng Yi (1986) Changes in family structure in China: a simulation study. *Population and Development Review* **12**, 4: 675–703.

Zeng Yi (1988) Changing demographic characteristics and the family status of Chinese women. *Population Studies* **42**, 2: 183–203.

Zeng Yi (1991) *Family Dynamics in China: A Life Table Analysis.* Wisconsin: The University of Wisconsin Press.

Zeng Yi (1990) *User's Manual and Software of "FAMY" – A PC Computer Program for Family Status Life Table Analysis.* Groningen: iec ProGAMMA (Interuniversity Expert Center for Computer Software Development and Distribution).

Zeng Yi and Vaupel, J. (1989) Impact of urbanization and delayed childbearing on population growth and ageing in China. *Population and Development Review* **15**, 3: 425–445.

Zeng Yi, Wang Zhenglian, Ma Zhongdong and Chen Chunjun (1993) A simple method for projecting fertility, first marriage, and leaving home rates. Working Paper, no. 6, The Institute of Population Research, Peking University.

Zeng Yi, Coale, A.J., Minja Kim Choe, Liang Zhiwu and Liu Li, (1994) Leaving parental home: census based estimates for China, Japan, South Korea, The United States, France, and Sweden. *Population Studies* **48**, 1: 65–80.

Zeng Yi, Vaupel, J. and Wang Zhenglian (1995) A multidimensional model for projecting family households – with an illustrative numerical application. Working Paper no. 17, Aging Research Unit, Center for Health and Social Policy, Odense University.