Household Projection Using Conventional Demographic Data

ZENG YI
JAMES W. VAUPEL
WANG ZHENGLIAN

IN MOST COUNTRIES, the size and structure of households are changing. Populations are aging 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 an older age. In some countries, marriage rates are declining and divorce rates and proportions of cohabiting couples are rising. People are living longer, so that an increasing number of middle-aged workers have living children, parents, and even grandparents. The gap between male and female life expectancy is widening, leaving more widows. Increased mobility is leading children to move to areas distant from their parents. Changes in attitudes about coresidence between adult children and parents in developing countries are resulting in reduction of extended family households. These factors, in various combinations and strengths for different populations, are yielding new patterns and distributions of household structures.

There is an important interaction between changes in household structure and the health status of the elderly. Living alone without nearby relatives can cause or worsen ill health and disability. Nearby family members often support the elderly who are ill or disabled. In the absence of such support, the need for nursing homes, social services, and health care services increases. Health care costs and social services provided to the elderly now account for over 10 percent of gross national product in many countries, including the United States. As the proportion elderly grows, these costs grow as well. Some scholars have proposed that household size and structure should be explicitly treated in the modeling of population, development, and environment. A strong argument in favor of this is that private consumption patterns are mostly defined in terms of household consumption, not individual consumption (Lutz and Prinz 1994: 225). For the preceding reasons and others, 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 households, including the single-parent household, the three-generation household, the household consisting of only one or two elderly people without children, the household of cohabiting couples? How many elderly persons will need assistance, but will not have children, spouse, or other close relatives to provide it? How many middle-aged persons will have to care both for elderly parents and for young children? This chapter develops a multidimensional model for projecting households using conventional demographic data that can be employed to address such questions as these.

A brief review of the three kinds of major models for family household projection is presented in the next section. We then discuss our modeling strategy and analytical framework. This is followed by an illustrative example of the application of the model with a sensitivity analysis on how demographic changes can affect households and population dynamics.

Models for household projection

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

The microsimulation model simulates the life course events and keeps detailed records of demographic status transitions for each individual. It has major advantages in studying the variability of individuals and households and their probability distributions (Hammel et al. 1976; Wachter 1987; Smith 1987; Nelissen 1991). It is particularly powerful in complex kinship simulation and projection, which macrosimulation and headship-rate methods cannot do. In a large population, in which households are classified by a large number of characteristics, however, the size of the representative sample to be used as the starting point of a projection should also be large. For instance, a sample of one percent of the populations of China and the United States consists of about 12.5 and 2.6 million persons, respectively. To simulate so many persons one by one would take substantial computing power and time. Another problem is that a census usually asks simple questions that cannot provide enough data for the microsimulation to model detailed characteristics of individuals. Hammel, Wachter, and their colleagues handled this problem by starting with a presimulation for a few decades before the beginning year of their projections. Using a manageable sample for this presimulation, they were able to approximate the family, household, and kinship distribution at the beginning year of the projection, and then simulate it forward. This approach is not as good as directly using the population, 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 accurately reflect 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.

The headship-rate method is a classic approach that has long been used by demographers to project households. In a census or a survey, a "head" is identified for each household. The age-sex-specific headship rates are computed by dividing the number of persons who are head of a household by the total number of persons of the same age and sex. The future households are projected by extrapolating the headship rates. Despite their widespread use, headship-rate methods suffer several serious shortcomings. The head is an arbitrary and vague choice 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: 510), although some recent work has shown that the regression approach to headship rates has some merit (Burch and Akaburskis 1993). 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 in such models (Mason and Racelis 1992: 510; Spicer, Diamond, and Ní Bhrolcháin 1992: 530).

The macrosimulation approach does not suffer the shortcomings inherent in headship-rate methods. Although not as flexible as microsimulation models in analyzing variability and probability distributions and in fully utilizing information at the individual level, macrosimulation models are not limited to the sample size at the beginning year of the projection and can effectively use the grouped data 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 there are some important advantages to macrosimulation, we do not wish to imply that microsimulation approaches are less useful. We believe that microsimulation is more powerful in some applications such as complex kinship simulations. It would be desirable to develop both kinds of approaches, since they have complementary strengths.

Keilman (1988), Van Imhoff and Keilman (1992), and Ledent (1992) reviewed dynamic household models based on the macrosimulation approach. Most of these models require data on transition probabilities among various household types or statuses, 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 data requirements, especially for data not commonly available on transition probabilities of household types or statuses in dynamic household models, are an important factor in the slow development and infrequent application of these models. Furthermore, the status-transition-based model cannot directly link changes in household structure with demographic rates. For example, changes in the probability of transition from households of husband and wife with more than two children to single-parent households with fewer than two children jointly depend on changes in divorce rates, death of spouse, fertility rates, and rates of leaving the parental home. It is extremely difficult to decompose the impacts of each demographic factor on changes in household status transition probabilities. Therefore, it is important to develop a dynamic household projection model that requires as input only conventional demographic rates that can be obtained from vital statistics, censuses, and ordinary surveys.

Benefiting from methodological advances in multidimensional demography (Rogers 1975; Land and Rogers 1982), and especially the multi-state marital-status life table model (Willekens et al. 1982; Willekens 1987), Bongaarts (1987) developed a nuclear-family-status life table model. Zeng (1986, 1988, 1991) extended Bongaarts's model and included both nuclear and three-generation households. The life table models by Bongaarts and Zeng are female-dominant one-sex models and assume that age-specific demographic rates are constant. Building on Zeng's family-status life table model, this chapter develops a two-sex dynamic projection model that permits demographic schedules to change over time. The model requires only data that are available from conventional demographic data sources. It can be used to identify effects of changes in demographic rates on household structure.

As pointed out by Lutz and Prinz (1994: 225), population models and household models cannot convert information based on individuals directly into information on households. Even if these two different aspects could be matched for the starting year, there is so far no way to guarantee consistent changes in both patterns when they are projected into the future. As is shown in this chapter, our new household model projects households and individuals simultaneously and consistently.

Demographic status identified

Following Brass's (1983) approach, we select the individual as the basic unit of the projection model. The major reason why we chose the individual is that demographic rates available from conventional population data sources can then be readily applied to individuals. The individuals of the base (or starting) population derived from a census or survey and the future projected population are classified according to the following dimensions of demographic statuses: age; sex; marital status (single, married, divorced, widowed, optionally cohabiting, and separated can be included); parity (optional); number of children living at home; coresidence with two parents or one parent or not living with parents;[1] rural or urban (optional); and whether living in a private household or a collective household.

Accounting system to link individuals' characteristics with households

We follow Brass's marker approach to identify households among individuals. Brass (1983) calls the reference person a household "marker." In Brass's original work and Zeng'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 chapter, both sexes are included, and a female adult, or a male adult when a female adult is not available, is identified as the reference person (or "marker").[2] The household type and size are derived from the characteristics of the reference person. For example, a person who is not married or cohabiting, not coresiding with parents, not living with a child, and not living in a collective household represents a one-person household. A married or cohabiting woman who is not coresiding with parents and not living with a child, represents a one-couple household. A married or cohabiting woman who is not coresiding with parents and living together with i children is a reference person representing a two-generation household of $2 + i$ persons. If this person is not married, the household size is $1 + i$.[3] A person who is married or cohabiting, coresiding with two parents or one parent, and living with i children is a reference person representing a three-generation household of $2 + 2 + i$ or $1 + 2 + i$ persons. If this reference person is not married, the household size is reduced by one. Readers interested in the mathematical formulas for computing the number of households with various types and sizes are referred to Zeng, Vaupel, and Wang (1997: 191–193).

Empirical test on the accuracy of the accounting system

We have tested the accounting system described above using real data sets. We identified each individual code by sex, marital status, and status of coresidence 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 above, we derived the distribution of households by types and sizes. The household distribution derived in this way may 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 household head. This kind of census tabulation may be called "direct-count." A comparison of distributions of household types derived by "model-counts" and "direct-counts" using the one percent data set of the China 1990 census shows that the relative differences are very small, all below one percent with one exception. The exception is that the error in the model-count of three-generation family households is 1.98 percent. The reason is that a joint family of two or more married brothers, with children living together, is counted as two or more stem-family households. This kind of 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 is likely in future years in China. The relative differences in frequency distributions of household sizes between model-count and direct-count using the Chinese data set are also reasonably small: 0.4 percent for one-person households, 2.5 percent for small households with 2–3 persons, 5.7 percent for middle-size households with 4–5 persons, and −10.7 percent for large households with 6 and more persons (Zeng, Vaupel, and Wang 1997).

To see how well our model works for populations in Western countries, we tested it on the 5 percent sample data file of the United States 1990 census and the German 1995 micro-census sample data file (sample size of the German mini-census is one percent of the total population). As shown in Tables 1 and 2, the distributions of the major household types derived by the "model-counts" and the "direct-counts" are almost identical.

TABLE 1 Comparison of number of family households by household type in the United States between the results derived from the model-count and the direct-count, using the 5 percent data tape of the 1990 census

	Number of family households				Frequency distribution			
Family type	**Mod-co.**	**Dir-co.**	**Dif. #**	**Dif. %**	**Mod-co.**	**Dir-co.**	**Dif. #**	**Dif. %**
One person	1229673	1228831	842	0.07	0.2696	0.2695	0.0001	0.03
One couple	1170281	1171011	−730	−0.06	0.2566	0.2568	−0.0002	−0.10
2 generations	2026992	2025525	1467	0.07	0.4444	0.4442	0.0002	0.04
3+ generations	134268	134319	−51	−0.04	0.0294	0.0295	0.0000	−0.07
Total	4561214	4559686	1528	0.03	1.0000	1.0000	NA	NA

NOTES: (1) 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. (2) The one-person household type refers to a person who does not live with spouse (or cohabiting partner) and children, but he or she may or may not live with other relative(s) or non-relative(s). NA = not applicable.

TABLE 2 Comparison of number of family households by household type in Germany between the results derived from the model-count and the direct-count, using the 1995 micro-census data

Family type	Number of family households				Frequency distribution			
	Mod-co.	Dir-co.	Dif. #	Dif. %	Mod-co.	Dir-co.	Dif. #	Dif. %
One person	90036	90036	0	0.00	0.3934	0.3934	0.00	0.00
One couple	55554	55554	0	0.00	0.2427	0.2428	0.00	0.00
2 generations	81054	81051	3	0.00	0.3542	0.3542	0.00	0.00
3+ generations	2209	2209	0	0.00	0.0097	0.0097	0.00	0.00
Total	228853	228850	3	0.00	1.0000	1.0000	NA	NA

Notes as in Table 1.
NA = not applicable.

As shown in Tables 3 and 4, the relative differences in frequency distributions of household sizes between model-counts and direct-counts are not small. Two factors contribute to the lower accuracy in the household size model-count. The first is that the model-count does not include those who are neither stem-family members nor spouse (or cohabiting partner) of the reference person. This problem is more serious in the United States and Germany than in China since some persons who live with their unmarried partners may not report as cohabiting, but rather report as "other relative" or "non-relative." The second factor is that we limit the highest parity to 5 in the illustrative examples, which underestimates the size of the large family households that have more than 5 children. As demonstrated by results not shown here, if we exclude those "other relatives" and "non-relatives" (i.e., who report as neither stem-family members nor spouse, or cohabiting partner, of the reference persons) in both the direct-

TABLE 3 Comparison of number of family households by household size in the United States between the results derived from the model-count and the direct-count, using the 5 percent 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 alone	1229673	1051897	177776	16.90	0.2696	0.2307	0.0389	16.86
2–3 persons	2225382	2289644	−64262	−2.81	0.4879	0.5021	−0.0143	−2.84
4–5 persons	982828	1023977	−41149	−4.02	0.2155	0.2246	−0.0091	−4.05
6+ persons	123331	194168	−70837	−36.48	0.0270	0.0426	−0.0155	−36.50
Total	4561214	4559686	1528	0.03	1.0000	1.0000	NA	NA
Average household size	2.66	2.52	−0.15	−5.68	NA	NA	NA	NA

NOTES: (1) Same as in Table 1. (2) "1 person alone" household here refers to one person who is living alone.
NA = not applicable.

TABLE 4 Comparison of number of family households by household size in Germany between the results derived from the model-count and the direct-count, using the 1995 micro-census data

	Number of family households				Frequency distribution			
Family type	Mod-co.	Dir-co.	Dif. #	Dif. %	Mod-co.	Dir-co.	Dif. #	Dif. %
1 person alone	90036	79128	10908	13.79	0.3934	0.3458	0.0477	13.78
2–3 persons	101375	110682	–9307	-8.41	0.4430	0.4836	–0.0407	–8.41
4–5 persons	35173	36512	–1339	-3.67	0.1537	0.1595	–0.0059	–3.67
6+ persons	2269	2528	–259	–10.25	0.0099	0.0110	–0.0011	-10.25
Total	228853	228850	3	0.00	1.0000	1.0000	NA	NA
Average household size	2.14	2.22	–0.08	–3.42	NA	NA	NA	NA

Notes as in Table 3.
NA = not applicable.

count and the model-count, the differences between the two become almost zero. This clearly demonstrates that the first factor is the major factor causing the lower accuracy in the household size model-count. Our model overestimates the number of one-person households by 15.9 percent using the US 1990 census data set and 13.8 percent using the German 1995 mini-census data set. In the US 1990 census, the living arrangement of cohabiting with an unmarried person is one of the codes of relationship to the reference person of the household, so that we could include "cohabiting" as one of the marital statuses distinguished. However, some cohabiting persons might be incorrectly reported as other relatives or non-relatives. In the case of the German 1995 mini-census, information about cohabitation was not required, so that a person who was non-married but living with a cohabiting partner was counted as a "one-person" household in our model, because of lack of information. The underestimation of the number of households with 6 or more persons is more serious in the case of the US 1990 census data set (36.5 percent), but less serious in the German case (10.3 percent). Large households with 6 or more persons are more likely to have member(s) who are other relative(s) or non-relative(s) of the reference person. But those large households made up a very small proportion of the total number of households, namely, 4.3 percent in the United States in 1990 and 1.1 percent in Germany in 1995. The inaccurate accounting of household-size distribution in the starting year and future projection years as a result of the lack of capacity to identify the reference person's living-together, other relative(s) or non-relative(s) in our model can be reasonably corrected by a procedure described in the appendix.

In sum, our model works better for developed countries than for developing countries in terms of accounting of household types because joint family households with two or more married siblings living together, which our model cannot identify, are negligible in the developed countries. However, our model counts household size less accurately in the United States and Germany than in China because of underreporting of cohabitation in the United States and Germany. The errors in household size accounting can be reasonably corrected by a simple procedure.

Computational strategy

It is theoretically possible for a woman or man to have any realistic combination of the statuses identified in the model. We may call the combination the composite state. Let $l_i(x,t)$ denote the number of persons of age x with composite state i (i=1, 2 ... 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,

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

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, as is the case in our model. In an illustrative numerical application, for instance, 2 statuses of residence, 4 marital statuses, 3 statuses of coresidence with parents, 6 parity statuses, and 6 statuses of coresidence with children are distinguished for females and males respectively, and the total number of composite statuses distinguished at each age for each sex in our model is:

$$T=2\times 4\times 3\times \sum_{p=0}^{5}(p+1)=504.$$

The total number of cells in the transition matrix is thus: 504x504=254,016. There would be one such large matrix for males and one for females at each single age. Although there are many zero cells in the matrices, the number of non-zero cells to be estimated is still much too large. Since so many categories have been distinguished, the number of observed events for some categories is too few to estimate the status-transition probabilities, even if the sample size is large. Therefore, the estimation of such large transition matrices is not practical.

Bongaarts's nuclear-family-status life table model also distinguishes numerous statuses. Bongaarts overcame the difficulty by assuming that particular events take place at particular points in time between age x and x+1 (Bongaarts 1987: 209–211). Here, we follow Bongaarts's 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. They depend on the status at the beginning and the middle of the year, respectively. 2) Deaths, migration, changes in status of coresidence with parents, marital status transition, and changes in number of surviving and coresiding children resulting from children's death or leaving or returning home occur in the middle of the year. These probabilities of occurrence of events refer to the whole year and depend on status at the beginning of the year. We calculate probabilities only for those transitions triggered by the specific demographic events included in the projection model—birth, death, marriage, divorce, remarriage, leaving and returning to parental home, migration, and so on. As shown mathematically and numerically by Zeng (1991: 61–63, 81–84), the strategy of assuming that births occur throughout the first and the second half of the year and other status transitions at the middle of the year leads to accurate numerical estimates.

Consistency in the two-sex and multi-generation model

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

1) Consistency between females and males. If the population under study is a closed marriage market with negligible inter-marriages with outsiders, such as a country with extremely few international marriages, we use the harmonic-mean procedure to ensure the two-sex consistency. 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; the number of newly widowed females (males) is equal to the number of new deaths of currently married men (women). When the cohabiting status is distinguished, the number of cohabiting males is equal to the number of cohabiting females; the number of males who exit from cohabiting status is equal to the number of female counterparts. Mathematical formulas for the harmonic mean used in our two-sex model to ensure consistency can be found elsewhere (e.g., Keilman 1985: 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). If the population under study is not a closed marriage market, such as a subregion within one country or a country with sizable international marriages, the two-sex consistency should not be computed. Our software ProFamy provides appropriate options for users to choose whether to compute the two-sex consistency.

2) Consistency between children and parents. We define three quantities from children's perspective[4]: C1: number of status transitions from living with two parents to not living with parents due to leaving home and due to death of parents; C2: number of status transitions from living with one parent to not living with parent due to leaving home and due to death of parent; C3: number of deaths of persons who lived with parents at time of death. Let S1 = C1 + C2 + C3. Define three quantities from the perspective of parents: P1: number of events of reductions in number of children living together[5] of wives (and cohabiting women if any) due to children leaving home and due to children's deaths; P2: number of events of reductions in number of children living together of single parent due to children leaving home and due to children's deaths; P3: number of deaths of a couple (both parties die in the same year) multiplied by number of their coresiding children plus number of deaths of single parent multiplied by number of their coresiding children. Let S2 = P1 + P2 + P3. In any year, S1 should be equal to S2. In the numerical calculation, however, S1 and S2 may not be exactly the same because of differences in estimation procedures. Therefore, an adjustment is needed to ensure that S1 and S2 are equal to each other. Readers interested in the mathematical formulas and their derivation, which follows the harmonic mean approach to ensure S1 = S2, are referred to Zeng, Vaupel, and Wang (1997: 195–196).

3) Consistency between births computed for the female and male populations. Changes in parity and status of coresidence 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 be equal to the total number of births computed based on the male population. Single-year age- and parity-specific fertility rates for male populations are rarely available. Therefore, we estimate male birth rates based on female birth rates and the average age difference between male and female partners. Clearly, computation of births for the female population is more accurate than for males. Therefore, the number of births produced by the male population is adjusted (by raising or lowering the male age-specific birth rates) to equal the number of births produced by the female population. In our model the assumption about how fertility depends on marital status for females is consistent with that for males.

4) Consistency between female and male status of coresidence with children before and after divorce (and dissolution of cohabitation). Children would stay with either the mother or the father after their parents' divorce or dissolution of cohabitation. Therefore, the number of children living together with the mother or father immediately after the parents' divorce should be equal to the number before divorce. The living arrangement of the children of divorced couples is a complicated social phenomenon, and data are sparse. Because young children in most societies are more likely to stay with their mother after divorce, 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 after divorce. If a couple has an even number of children living together before their divorce, each party would have half of their children after divorce. In societies where divorced couples do not wish their children to be separated from each other, our model and ProFamy software can provide an option for users to assume that all children stay with their mother after their parents' divorce.

5) Consistency between female and male status of coresidence with children before and after remarriage. Children living with a single mother and a single father would join the new household after a parent's remarriage. A newly remarried couple's number of children living together should equal the sum of children living with either of the parties before remarriage. For simplicity, we assume that the probability that a remarried woman or man will have additional children from their new partner's previous union depends only on the frequency distribution of the status of coresidence with children of newly married men or women in the year.[6]

The demographic accounting equations

Demographic accounting equations are used to compute the number of female and male persons and the changes in marital and cohabiting status, parity, the status of whether living with parents and children, residence status, deaths, and so on in each projection year. The basic structure of all 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 new entries into status i that occur in the year (t,t+1) among persons age x+1 at time t+1) –
(number of exits out of status i that occur in the year (t, t+1) among persons age x at time t).

The number of events including birth, death, migration, marriage, divorce, leaving and returning to parental home, and so on between age x and x+1 (and between time t and t+1) is 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 time t and t+1).

Based on the above-stated principles as well as the analytical framework presented in the previous sections, we derive the accounting equations and the procedures for estimating the status transition probabilities to calculate changes in coresidence with parents, marital status, parity, and number of children living at home between age x and x+1 and between time t and t+1 for all individuals of the population. Readers interested in these formulas are referred to Zeng, Vaupel, and Wang (1997: 197–199, 211–214).

The accounting equations discussed above include all individuals of the population at the starting year of projection and update their survival status and other demographic statuses as well as household status in future years. The distribution of household size and structure are derived from characteristics of the reference persons. The tabulations of population size, age/sex distributions, and other demographic indexes such as proportions of the elderly, school-age children, and youth, dependency ratios, and size of labor force are derived from all individuals including reference persons and non-reference persons. Our new household projection model has thus simultaneously projected households and population consisting of individuals, and the consistency between patterns of changes in individuals and households is guaranteed. This is not the case in other models for projecting households or population.

List of assumptions made

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

1) 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.

2) Homogeneity assumption: people with the same characteristics have the same status transition probabilities.[7]

3) 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.

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

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

6) Some of the 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 non-married parent, and leaving the parental home as well as returning home are independent.

7) If a couple has an odd number of children living together before divorce, the mother will have one more child than the father after divorce; if a couple has an even number of children living together, each party has

half of their children after divorce; or the user can assume that all children stay with their mother after their parents' divorce.

8) A remarried person's probability of having additional children from the new partner's previous union depends on the frequency distribution of the status of coresidence with children of the newly married persons with opposite sex in the same year.

Data needed

1) Base population derived from a census or a survey, and classified by age, sex, marital status, parity (optional), number of children living at home, and coresidence with parents as well as whether living in a collective household. Information on parity (number of children ever born) is optional. If no parity information is available, we assume that birth probabilities depend on age and number of children living at home.

2) Age-specific demographic schedules observed in the recent past:

a) Age-, sex- (and marital-status if possible) specific probabilities of surviving.

b) Age- and sex-specific occurrence/exposure rates or probabilities of marital status transitions.

c) Age- and parity-specific occurrence/exposure rates or probabilities of birth by married women. If births by non-married women are not negligible, the ratios of the overall fertility level of women with various non-married statuses to the overall fertility level of the married women will be needed.

d) Age- and sex-specific occurrence/exposure rates or probabilities of leaving and returning to the parental home. If data on returning home are not available, one can use the age- and sex-specific net rates of leaving the parental home.

e) Age-, sex- (and marital-status if possible) specific occurrence/exposure rates or probabilities of migration 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.

It is ideal to have the observed age-specific data described in a) through e) above from the country or region under study. When this is not available, however, one may use the standard schedules based on data from other countries or regions where the general age pattern of demographic processes is similar to that in the country or region under study.

3) Parameters to specify levels of parity-specific total fertility rates (TFR); propensity of eventual marriage, divorce, remarriage, and other possible transitions among marital statuses including cohabitation and separation if they are distinguished by the user; proportion eventually leaving

parental home; mean age at first marriage for males and females, mean age at births of all orders combined, mean age at leaving parental home, life expectancy at birth, and total number of male and female immigrants and out-migrants in the future projected years.

If the option of rural–urban regional classification is chosen by the user in the application, the data specified in 1) through 3) are region-specific. In this case, age-, sex- (and marital-status if available) specific net rates of rural-to-urban migration within the country or region under study will be needed. One will also need to specify the proportion of urban population in the future projected years.

The observed age-specific or age-parity-specific rates or probabilities define demographic schedules. The demographic schedules and the projected 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 demographic 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 future demographic schedules based on a standard schedule and projected changes in the median age and the inter-quartile ranges (Zeng et al. 1993).

At this stage of our research, only modest attention has been paid to incorporating the best methods for projecting the future demographic rates required by the projection model. Much more work on this is planned, including research on joint forecasting of economic and demographic variables.

Illustrative application: A sensitivity analysis on how demographic changes affect households and population dynamics

Our household projection model and its associated computer software ProFamy can produce a large number of output tables and graphics for each of the projection years, including cross-tabulations of distributions of age, sex, marital status, parity, and coresidence with one or two parents or not living with parents, 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 wife or husband; distribution of nuclear-family and three-generation households by household size and structure and by age, marital status, and other characteristics of the reference persons, as well as the age and sex distribution of the population; percent of school-age children; percent and age distribution of elderly; dependency ratios of children and elderly; labor force population; and so on. These distributions are presented in absolute numbers and percentages. Applications of our model and the software ProFamy can be used for purposes of policy analysis, social planning, or market analysis. They can also be used for academic purposes to answer such important questions as how demographic changes affect future households.

Elsewhere we have presented an illustrative application of two major scenarios with the same assumptions about medium fertility, first marriage, divorce, remarriage, and leaving parental home, but with medium and low mortality assumptions. Some straightforward analysis on how medium fertility, the assumed trends in future marriage, divorce, and leaving home, and medium versus low mortality may affect future Chinese household structure and population can be found in Zeng, Vaupel, and Wang (1997). Since then, we have been using Chinese and American data to conduct more in-depth and substantive applications; however, it is impossible to present them here because of space limitations. Such substantive applications together with careful interpretation and policy analysis will be presented in our subsequent publications on this subject. We therefore simply present our new sensitivity analysis on how demographic changes may affect households in future years, using the Chinese data to illustrate the application of the model mainly for academic purposes.

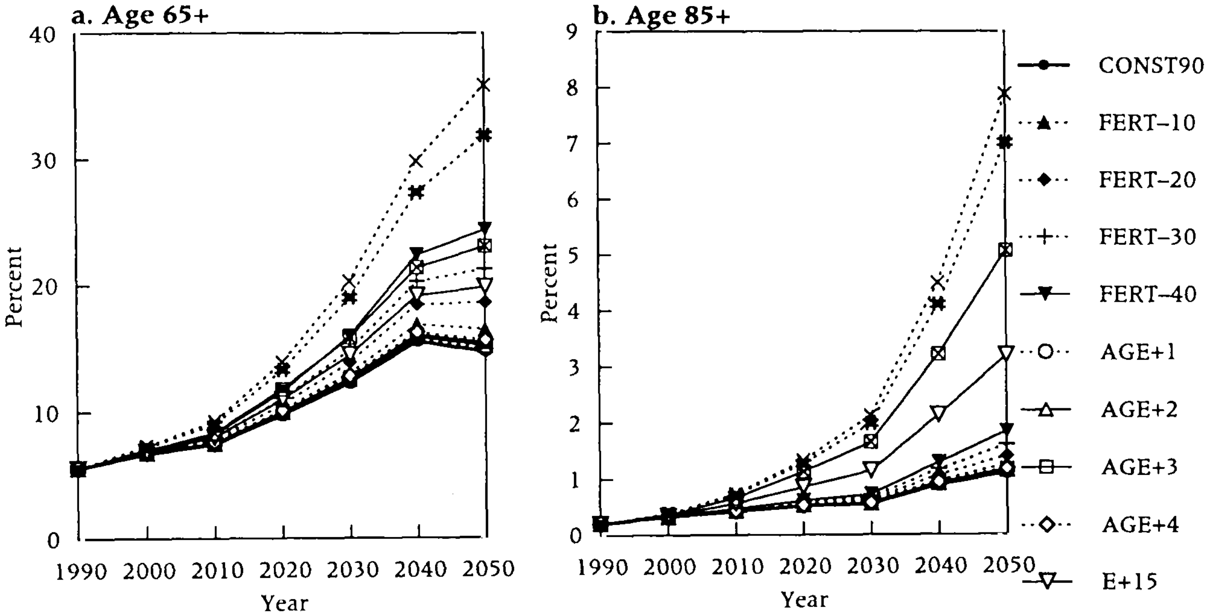
The base population and other required data of demographic schedules are derived from the one percent sample data tape of the 1990 census, the 1988 Two-Per-Thousand Fertility and Contraceptive Survey (State Family Planning Commission 1990), and the 1985 in-depth fertility survey (see Zeng, Vaupel, and Wang 1995 for more detailed resources, the estimates, and discussion).

The baseline scenario keeps everything constant at the 1990 level. In 1990, the TFR in China was 2.2, with parity-specific rates of 0.980, 0.767, 0.298, 0.096, and 0.056 for the first, second, third, fourth, and fifth and higher births, respectively. Life expectancy at birth was 70.7 and 67.5 years for females and males. The proportions eventually marrying were 0.99 and 0.97 for females and males. The probability of eventually divorcing was 0.071; the probabilities of eventually remarrying were 0.41 and 0.76 for the divorced and widowed.

The simulation results of all other scenarios with changing demographic parameters in one or two dimensions while everything else remains unchanged are compared with the baseline scenario. The numerical results are presented in Figures 1 and 2 and in Appendix Tables 1–3.

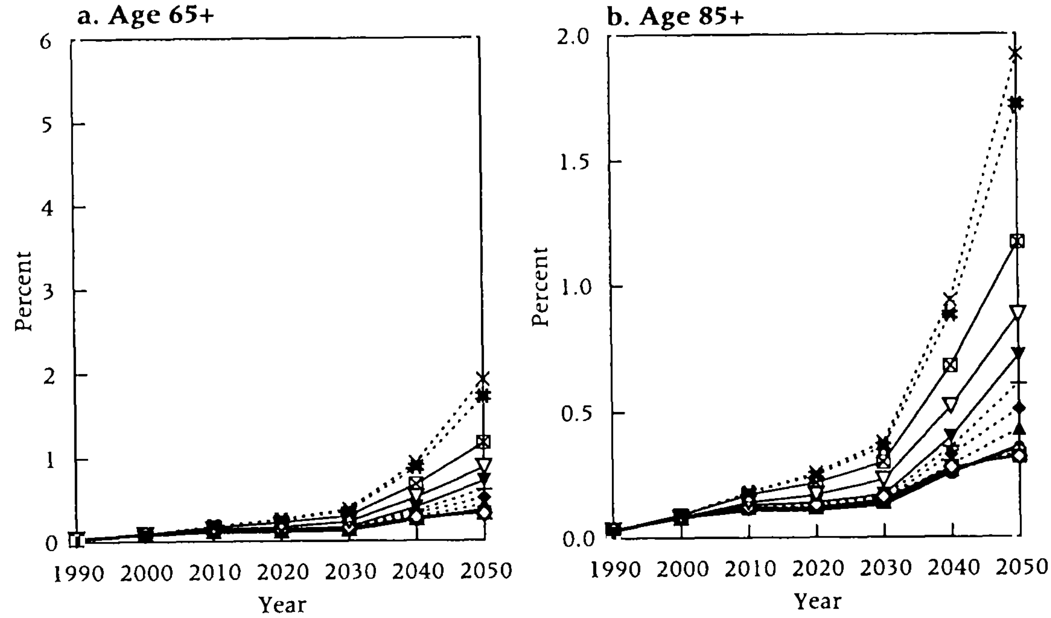
Comparing the baseline scenario with the four scenarios in which fertility levels are reduced by 10, 20, 30, and 40 percent respectively (marked as FERT–10, FERT–20, FERT–30, FERT–40 in the figures and tables),[8] we found that fertility decrease not only reduces total population and household size and increases the proportion of elderly, but also substantially increases the proportion of elderly living alone (see Figures 1 and 2) and the

FIGURE 1 Percent of total population that are elderly under various scenarios



NOTE: For explanations of the legends (i.e., assumptions in each scenario) see note to Appendix Table 1.

FIGURE 2 Percent of total population that are elderly living alone under various scenarios



NOTE: For explanations of the legends (i.e., assumptions in each scenario) see note to Appendix Table 1.

proportion of one-person and one-couple households. Further decline in fertility alone would slightly increase the proportion of three-generation households because there are fewer siblings to establish new nuclear households. When fertility decreases by 40 percent to an extremely low level, the proportion of three-generation households decreases by 6.6 percent in 2050 as compared with the baseline scenario—the extremely low fertility would cause some elderly to be unable to live with a married child even if they wish to do so (Zeng 1991: 136–137). It is interesting to note that as compared with the baseline scenario, fertility decline alone would have a very small effect on the proportion of single-parent households in the first quarter of the next century, but this effect would be much larger by the middle of the next century—the proportion would increase by 14 or 25 percent if fertility decreased by 30 or 40 percent (see Appendix Tables 2 and 3). This is perhaps because with the further substantial fertility reduction, other things constant, there will be proportionally more elderly because of population aging. The elderly are more likely to be in the status of widowhood or to remain non-married after divorce, as compared with younger persons.

Increase in median age at first marriage and at births of various parities by 1, 2, 3, and 4 years[9] as compared with the baseline scenario (marked as AGE+1, AGE+2, AGE+3, AGE+4 scenarios) would decrease total population size by 13, 26, 39, and 51 million in 2020; and by 22, 45, 68, and 91 million in 2050. The increase in age at first marriage and at births would modestly increase the dependency ratio of the elderly. It would increase the proportion of elderly living alone in 2020 because of the age structure change. By the middle of the next century, however, the increase in age at first marriage and at births would result in a modest decrease in the proportion of elderly who are living alone (see Figures 1 and 2), because the prolonged length between generations would enable more elderly to live with their young children who have not yet reached the age of leaving home. An increase in age at first marriage would lead to an increase in the proportion of one-person households; this impact will be larger in years 2000 and 2020 than in 2050. Delaying marriage may result in an increase in proportion of single-parent households (see Appendix Tables 1–3). The reason is that delayed marriage will reduce the total number of two- and three-generation households, whereas the rates of divorce and remarriage are assumed to be constant.

Two mortality reduction scenarios assume that life expectancy will increase linearly by 15 percent and 25 percent respectively from 1990 to 2050 (marked as E+15 and E+25), which are about the same medium and low mortality assumptions used in our previous population projection studies (Zeng and Vaupel 1989). The expected mortality decrease will substantially increase total population size and the proportion elderly. As compared with the baseline scenario, the increase in the proportion of the oldest-old in these two scenarios will be much more dramatic—a 187 or 355 percent increase of the oldest-old, age 85 and older, in 2050 depending on whether the mortality improvement is 15 or 25 percent, as compared with a 35 or 57 percent increase for all elderly age 65 and older. As compared with the baseline scenario, the percent of elderly age 65 and older living alone in 2050 would decrease by 8.7 or 4.5 percent if mortality improvement is 15 or 25 percent because of the general reduction of widowhood among all elderly. With marital status transition probabilities being constant, mortality improvement (15 percent and 25 percent respectively) alone would, however, increase the percent of the elderly age 85 and older who are living alone by 8 and 13 percent in 2000, 54 and 92 percent in 2020, and 145 and 227 percent in 2050, respectively (see Figures 1 and 2). This is because the mortality differentials between males and females tend to be larger at extremely high ages, which will result in more oldest-old widows when life expectancy increases substantially. Some recent studies (e.g., Kannisto et al. 1994) have shown that future improvement of mortality at advanced ages will be greater than at younger ages. Consequently, there will be more persons age 85 and older. It is clear that demographers should pay more attention to the study of these populations. Mortality improvement would result in a considerable decrease of single-parent households among two- and three-generation households because of the reduction of the likelihood of widowhood.

The combination of decrease in fertility and increase in life expectancy (marked as F–30E+25 and F–40E+25) would more sharply increase the proportion elderly and the elderly living alone, especially those age 85 and older (see Figures 1 and 2).

Increase in divorce rates by 100, 200, 300, and 400 percent (marked as DIV+100, DIV+200, DIV+300, and DIV+400) would slightly increase the proportion of elderly living alone. Its impact on reducing household size is not substantial, about 0.6–1.2 percent in 2020 and about 1.0–2.5 percent in 2050, but it will modestly increase the proportion of one-person households. The increase in divorce rates will, however, substantially increase the proportion of single-parent households by 8–31 percent in 2020 and 15–56 percent in 2050 (see Appendix Tables 1–3). The divorce propensity in the baseline scenario, which keeps the Chinese 1990 demographic rates unchanged, is very low (7 percent of all marriages would end in divorce in 1990). A 400 percent increase in the Chinese divorce level by the middle of the twenty-first century, which is probably the maximum possible change in the Chinese cultural and social context, would still be lower than the current divorce level in the United States. If a more dramatic increase in divorce rates in China were to occur in the next century, its impact on household dynamics would be more significant than the simulation results presented in Appendix Tables 1–3.

The direction and magnitude of decreasing remarriage rates (decreased by 25 percent and 50 percent in scenarios marked REM–25 and REM–50) on the proportion of elderly living alone and on average household size are minor. The impact on the increase of the proportion of single-parent households is about 4–8 percent in 2020 and 8–17 percent in 2050 (see Appendix Tables 1–3).

If the propensity for leaving the parental home after marriage increased by 20, 35, and 50 percent (marked as LH+20, LH+35, and LH+50), the proportion of elderly aged 65+ and 85+ living alone would increase by 17–69 percent and 22–77 percent in 2050. The average household size would decrease by 5–14 percent; the proportions of one-person and one-couple households would increase by 14–53 percent and 17–61 percent, respectively; the two- and three-generation households would decrease by 1–26 and 19–48 percent, respectively; and the single-parent household would decrease by 4–26 percent. It is clear that change in the propensity for coresidence between parents and married children is one of the most important factors affecting future household structure in a developing country such as China. The simulation also suggests that the rarity of married children coresiding with parents in Western countries, and the high proportion of married children doing so in developing countries, explain much of the large difference in household structure between Western and developing countries.

Although the results of the sensitivity analysis presented and discussed above are useful in understanding demographic impacts on household dynamics, they should be interpreted as a country-specific illustration and should not be generalized to other populations without caution. For example, our simulations have shown that the substantial decrease in the Chinese remarriage rates would have relatively minor impacts on household dynamics in the future. This is mainly because the divorce level and remarriage rates at ages over 50, when widowhood is more likely to occur, were very low in the Chinese 1990 baseline scenario. If, elsewhere, the baseline divorce rates are high and the remarriage and cohabiting rates of widows and widowers over age 50 are also relatively high, as is the case in the United States, the impact of changes in remarriage rates on household dynamics may be substantially larger.

Concluding remarks

As Keyfitz (1985) and Bongaarts (1983) observed, family demography is a difficult and underdeveloped field. Our model and its associated software ProFamy contribute to the development of methods for projecting family structures in the following ways.

First, the model permits projection of many characteristics of households and their members, using 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 demographic schedules for the population under study are not available (e.g., in a small area or a developing country with poor data resources), model standard schedules from another region or country with similar demographic conditions can be used. The user can then project summary measures, such as life expectancy, total fertility rate by birth order, proportion eventually leaving the parental home, proportion eventually married (and cohabiting, if any) and divorced, proportion eventually remarried after divorce and widowhood, and the like. We therefore expect it to be relatively easy to apply the model.

Second, unlike the traditional headship-rate method in which demographic factors are not systematically reflected in the rates themselves, our model can closely link the projected 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 demographic changes may affect households. As demonstrated by our illustrative numerical application and the sensitivity analysis, such an exercise offers useful insights into household and population dynamics and their policy relevance.

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

Fourth, our model includes all individuals in the population at the starting year of projection, and updates their survival status and other demographic statuses as well as household status in future years. The distribution of household size and structure is derived from characteristics of the reference persons. The tabulations of population size, age/sex distributions, marital status distribution, and other demographic indexes are derived from all individuals, including reference persons and non-reference persons. Consistency between patterns of changes in individuals and households is guaranteed in our model. This is not the case in other models for projecting households or population.

There is a need for long-term projections, 30 or 50 years or even a century into the future. Uncertainties about developments so far in the future are sufficiently great that such projections should not be interpreted as forecasts. Nonetheless, long-term projections may be useful to policy analysts as scenarios for comparing the relative effects of alternative policies. For instance, long-term projections provide insights into the relative impact on population aging of a one-child versus a two-child-plus-spacing policy in China (Vaupel and Zeng 1991). The model presented in this chapter will also be useful for such scenario analysis of the long-term consequences of alternative policy directions.

Most planners and analysts in governmental agencies and in business firms are primarily interested in short-term forecasts of trends over the next five or ten years. 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 a ten-year time frame because of the uncertainties of the demographic parameters and the associated socioeconomic variables. Furthermore, a long-term 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 deemed correct when the range straddles the subsequent outcome (Keyfitz 1985: 204). We believe that the model presented here has the potential of performing satisfactorily for shorter-term forecasts, even though it involves a substantially richer set of variables than most demographic projections.

Of course, the method of household projection presented in this chapter requires further development. For example, the formulation of more accurate assumptions about future demographic parameters deserves more attention. Ideally, the demographic assumptions should be formulated based on prediction of future changes in social and economic factors and their association with demographic parameters. Although the methodology for doing so is not mature and the prediction of future socioeconomic factors is itself inexact, we should continue to develop good models to forecast demographic parameters using socioeconomic factors as explanatory variables. In addition to formulating assumptions about the main parameters in the household projection—such as total fertility rate, life expectancy, and proportions eventually married, divorced, remarried, and leaving and returning to the parental home—we also need to formulate assumptions about model schedules of these demographic processes. Regional model life tables are well developed. Migration model schedules were developed at IIASA some years ago by Rogers and his colleagues (1975). However, very little work on regional model schedules of fertility, marriage, divorce, remarriage, and leaving and returning to the parental home has been done. Since model schedules are important for demographic projections, their development is another future research endeavor that needs more attention.

Appendix

Based on the census data set, we can derive $h(i,j,t)$, the proportion of households with i direct family members and j other relatives or non-relatives among the total number of households with i direct family members in year t. The term "direct family members" here refers to spouse (or cohabiting partner), children, and parents of the reference person.

$\sum_j h(i,j,t) = 1.0$, for all i. The maximum value of i in our model is 2+2+P, i.e., the largest three-generation household has two grandparents, two parents, and P (highest parity distinguished) children. j = 0, 1, 2, 3, ... O, where O is the largest number of other relatives or non-relatives living in a household. We chose O as 5 in our current version of software ProFamy since we think that the number of single households with more than five other relatives or non-relatives in modern societies is negligible.

Denote $H(i,t)$ as the number of households of size i accounted by our model before the adjustment. Denote $N(i,j,t)$ as number of households with i direct family members and j other relatives or non-relatives in year t. $N(i,j,t) = H(i,t) * h(i,j,t)$. The actual household size of $N(i,j,t)$ is $i+j$. Regrouping $N(i,j,t)$ by summing i and j as z, we obtain the adjusted number of households with size z in year t, which is denoted as $H(z,t)$, where $z=1,2,3,\ldots,2+2+P+O$ (i.e., the largest household size is $2+2+P+O$).

The average number of other relatives or non-relatives among all households with i direct family members is: $a(i,t) = h(i,j,t) * j$. We can allow $a(i,t)$ to change over time. We may assume that the relative changes in $h(i,j,t)$ for all $j>0$ in year t as compared with year $t-1$ is the same as the relative changes of $a(i,t)$ as compared with $a(i,t-1)$, and $h(i,0,t)$ will change accordingly to fulfill the constraint of the sum of $h(i,j,t)$ for all j is equal to one.

$h(i,j,t) = h(i,j,t-1) * a(i,t)/a(i,t-1)$ for all $j>0$. If the sum of $h(i,j,t)$ ($j>0$) is greater than one, which may usually not happen in the real world, we will have to standardize $h(i,j,t)$ ($j>0$) to make sure the sum is not greater than one. We then estimate $h(i,0,t)$ as:

$$h(i,0,t) = 1.0 - \sum_{j>0} h(i,j,t)$$

To help readers understand how this procedure works, we present a numerical example as follows. Based on the US 1990 census data set, we know that the proportions of American households with four direct family members and 0, 1, 2, 3, 4, 5 other relatives or non-relatives were 0.9320, 0.0516, 0.0102, 0.0040, 0.0012, and 0.0011 respectively, and the average number of other relatives or non-relatives among the households of four direct family members was 0.09 in 1990. If we assume that this average will become 0.11 in year 2000, we then estimate:
$h(4,1,2000) = h(4,1,1990) * 0.11/0.09 = 0.0516*1.222=0.0631$;
$h(4,2,2000) = h(4,2,1990) * 0.11/0.09 = 0.0102*1.222=0.0125$;
$h(4,3,2000) = h(4,3,1990) * 0.11/0.09 = 0.0040*1.222=0.0049$;
$h(4,4,2000) = h(4,4,1990) * 0.11/0.09 = 0.0012*1.222=0.0015$;
$h(4,5,2000) = h(4,5,1990) * 0.11/0.09 = 0.0011*1.222=0.0013$;
$h(4,0,2000) = 1.0 - (0.0631+0.0125+0.0049+0.0015+0.0013) =0.9167$.

APPENDIX TABLE 1 How changes in demographic rates may affect households in China in the year 2000

	Absolute values of households summary measures									Relative changes (%) as compared with baseline scenario (CONST90)								
SCENARIO	HSIZE	H1	H2–3	H4–5	H6+	COUP	GEN2	GEN3	SP2–3	HSIZE	H1	H2–3	H4–5	H6+	COUP	GEN2	GEN3	SP2–3
CONST90	4.1	8.1	26.3	48.3	17.3	6.7	62.7	22.5	4.6									
FERT–10	4.0	8.1	26.6	49.1	16.3	6.6	62.8	22.5	4.6	–1.0	–.4	1.2	1.5	–5.9	–1.0	.1	.1	–.2
FERT–20	4.0	8.0	27.4	49.1	15.5	6.7	62.7	22.6	4.6	–2.1	–.6	4.2	1.6	–10.5	–.1	–.1	.4	–.4
FERT–30	3.9	8.1	29.5	48.2	14.2	6.9	62.5	22.6	4.6	–3.3	–.2	12.3	–.2	–18.0	3.6	–.4	.2	–.3
FERT–40	3.9	8.1	32.7	46.0	13.1	7.3	62.1	22.4	4.6	–4.7	.6	24.2	–4.7	–23.9	9.5	–1.0	–.4	.2
AGE+1	4.0	9.0	26.3	47.1	17.6	7.0	62.0	22.0	4.8	–.6	11.1	.1	–2.5	1.8	4.9	–1.1	–2.2	4.0
AGE+2	4.0	10.0	26.3	45.9	17.8	7.3	61.2	21.4	5.0	–1.4	24.1	0	–5.1	2.9	9.8	–2.4	–4.9	8.1
AGE+3	4.0	11.2	26.3	44.6	17.9	7.6	60.5	20.6	5.2	–2.3	38.8	–.1	–7.6	3.4	13.9	–3.5	–8.3	12.6
AGE+4	3.9	12.6	26.2	43.4	17.8	7.8	59.9	19.7	5.4	–3.5	55.3	–.4	–10.2	3.1	16.9	–4.5	–12.4	17.2
E+15	4.1	8.0	26.2	48.3	17.4	6.7	62.6	22.6	4.5	.2	–.7	–.3	0	.8	.9	–.2	.6	–3.4
E+25	4.1	8.0	26.2	48.3	17.5	6.8	62.5	22.7	4.4	.3	–1.1	–.4	–.1	1.4	1.5	–.4	1.0	–5.6
F–30E+25	3.9	8.0	29.4	48.2	14.4	7.0	62.2	22.8	4.3	–3.1	–1.3	11.8	–.2	–16.8	5.0	–.8	1.2	–5.9
F–40E+25	3.9	8.0	32.5	46.1	13.3	7.4	61.9	22.7	4.4	–4.4	–.6	23.7	–4.6	–22.9	10.7	–1.3	.7	–5.6
DIV+100	4.1	8.1	26.3	48.3	17.3	6.7	62.7	22.5	4.8	0	.2	.1	–.1	–.1	–.3	0	.1	2.8
DIV+200	4.0	8.1	26.4	48.2	17.2	6.7	62.7	22.5	4.9	–.1	.5	.4	–.2	–.3	–.3	0	0	5.6
DIV+300	4.0	8.1	26.5	48.2	17.2	6.6	62.7	22.5	5.0	–.2	.8	.6	–.3	–.5	–.4	0	0	8.4
DIV+400	4.0	8.2	26.5	48.2	17.2	6.6	62.7	22.5	5.1	–.2	1.0	.8	–.4	–.6	–.6	–.1	0	11.2
REM–25	4.1	8.1	26.3	48.3	17.3	6.7	62.7	22.5	4.7	0	.2	.1	0	–.1	–.1	0	0	1.0
REM–50	4.0	8.1	26.3	48.3	17.2	6.7	62.7	22.5	4.7	–.1	.3	.1	0	–.3	–.1	0	–.1	2.1
LH+20	4.0	8.6	28.2	47.2	16.0	7.6	63.3	20.5	4.7	–2.2	6.6	7.4	–2.4	–7.7	13.2	.9	–8.9	2.1
LH+35	3.9	9.1	29.8	46.2	15.0	8.4	63.6	18.9	4.8	–4.0	12.4	13.2	–4.5	–13.4	25.1	1.5	–15.9	3.3
LH+50	3.8	9.6	31.4	45.0	13.9	9.3	63.8	17.2	4.8	–5.9	19.3	19.4	–6.9	–19.3	39.6	1.7	–23.4	3.9
LH+100	3.7	10.4	33.0	43.7	12.9	10.6	63.5	15.5	4.8	–8.0	28.6	25.3	–9.6	–25.1	58.3	1.2	–31.0	3.2

NOTES: (1) Codes of scenarios—CONST90: everything constant as in 1990; FERT–10, FERT–20, FERT–30, FERT–40: fertility reduces 10%,20%,30%,40%; AGE+1, AGE+2, AGE+3, AGE+4: age at 1st marriage and at parity-specific births increases 1,2,3,4 years; E+15, E+25: life expectation at birth increases 15%,25%; F–30E+25: fertility reduces 30%, life expectation increases 25%; F–40E+25: fertility reduces 40%, life expectation increases 25%; DIV+100, DIV+200, DIV+300, DIV+400: divorce propensity increases 100%,200%, 300%,400%; REM–25,REM–50: remarriage propensity reduces 25%, 50%; LH+20, LH+35, LH+50: leaving parental home propensity increases 20%,35%,50%; LH+100: everyone leaves parental home after marriage.
(2) Abbreviations of the households summary measures—HSIZE, average household size; H1, H2–3, H4–5, H+6: percent of households of size 1, 2–3, 4–5, 6+ persons; COUP: percent of one-couple households with no children; GEN2, GEN3: percent of 2-generation, 3-generation household; SP2–3: percent of single-parent households among two- and three-generation households.

APPENDIX TABLE 2 How changes in demographic rates may affect households in China in the year 2020

	Absolute values of households summary measures									Relative changes (%) as compared with baseline scenario (CONST90)								
SCENARIO	HSIZE	H1	H2–3	H4–5	H6+	COUP	GEN2	GEN3	SP2–3	HSIZE	H1	H2–3	H4–5	H6+	COUP	GEN2	GEN3	SP2–3
CONST90	3.7	8.3	35.7	43.8	12.2	10.6	62.9	18.2	6.8									
FERT–10	3.6	8.3	37.9	43.8	10.1	11.2	62.4	18.2	6.9	–2.9	–.6	6.2	–0.1	–17.3	5.1	–.7	–.2	.2
FERT–20	3.5	8.3	40.6	42.6	8.5	12.2	60.9	18.6	6.9	–5.7	.5	13.8	–2.9	–30.6	14.3	–3.1	2.3	.1
FERT–30	3.4	8.6	45.1	39.5	6.7	13.6	58.3	19.4	6.8	–8.7	3.6	26.5	–9.8	–44.7	27.9	–7.2	6.9	–.4
FERT–40	3.3	8.9	50.7	35.0	5.4	15.4	55.3	20.4	6.7	–11.6	7.6	42.0	–20.1	–55.9	44.2	–12.1	12.4	–1.4
AGE+1	3.7	9.4	34.9	43.1	12.6	10.7	61.3	18.6	7.0	–.4	12.9	–2.3	–1.7	3.7	.8	–2.5	2.3	2.6
AGE+2	3.7	10.7	33.8	42.3	13.1	10.9	59.3	19.2	7.2	–.9	28.8	–5.2	–3.4	7.7	2.1	–5.7	5.3	4.5
AGE+3	3.7	12.1	32.7	41.7	13.5	10.9	57.4	19.5	7.3	–1.5	46.2	–8.4	–4.9	10.7	2.6	–8.7	7.5	6.2
AGE+4	3.7	13.5	31.5	41.2	13.8	10.7	56.3	19.4	7.4	–2.0	62.7	–11.6	–6.1	13.2	.7	–10.4	6.8	8.6
E+15	3.8	7.9	35.5	43.8	12.8	11.0	62.5	18.6	5.9	.8	–5.2	–.4	–.1	4.8	3.6	–.6	2.5	–14.3
E+25	3.8	7.6	35.5	43.8	13.1	11.3	62.2	18.9	5.3	1.3	–8.6	–.4	–.2	7.8	6.0	–1.0	4.1	–22.5
F–30E+25	3.5	7.8	44.8	39.8	7.7	14.1	57.9	20.3	5.3	–7.2	–6.5	25.4	–9.1	–37.2	32.0	–7.9	11.5	–22.0
F–40E+25	3.4	8.1	50.3	35.6	6.1	15.6	54.9	21.4	5.3	–10.0	–3.0	40.9	–18.8	–50.0	46.6	–12.6	17.6	–22.6
DIV+100	3.7	8.5	35.8	43.7	12.1	10.5	62.8	18.2	7.4	–.3	1.9	.2	–.4	–.6	–1.1	–.1	.1	8.1
DIV+200	3.7	8.7	35.9	43.4	12.0	10.5	62.7	18.2	7.9	–.6	4.2	.7	–.9	–1.4	–1.6	–.2	–.1	16.0
DIV+300	3.7	8.8	36.0	43.2	11.9	10.4	62.6	18.2	8.5	–.9	6.3	1.0	–1.4	–2.1	–2.4	–.4	–.2	23.9
DIV+400	3.7	9.0	36.1	43.0	11.9	10.3	62.5	18.1	9.0	–1.2	8.4	1.3	–1.9	–2.8	–3.1	–.5	–.3	31.7
REM–25	3.7	8.4	35.7	43.8	12.1	10.6	62.8	18.2	7.1	–.2	1.4	.1	–.2	–.7	–.5	–.1	–.1	3.9
REM–50	3.7	8.5	35.8	43.7	12.0	10.5	62.8	18.1	7.4	–.4	3.0	.2	–.3	–1.5	–1.0	–.1	–.3	8.2
LH+20	3.6	10.0	38.4	41.0	10.6	13.2	61.0	15.8	6.5	–4.2	19.9	7.8	–6.5	–12.9	24.2	–3.0	–13.0	–4.4
LH+35	3.4	11.5	40.4	38.7	9.4	15.6	59.2	13.8	6.2	–7.6	37.9	13.2	–11.7	–22.7	46.0	–5.9	–24.0	–9.4
LH+50	3.3	13.3	42.1	36.4	8.2	18.4	56.6	11.6	5.7	–11.4	60.4	18.0	–17.0	–32.6	72.8	–9.9	–36.0	–16.7
LH+100	3.1	15.9	43.1	33.9	7.1	22.2	52.5	9.4	5.0	–15.6	91.4	20.8	–22.7	–41.4	108.8	–16.5	–48.3	–26.9

Notes as in Appendix Table 1.

APPENDIX TABLE 3 How changes in demographic rates may affect households in China in the year 2050

	Absolute values of households summary measures									Relative changes (%) as compared with baseline scenario (CONST90)								
SCENARIO	HSIZE	H1	H2–3	H4–5	H6+	COUP	GEN2	GEN3	SP2–3	HSIZE	H1	H2–3	H4–5	H6+	COUP	GEN2	GEN3	SP2–3
CONST90	3.4	13.2	39.3	36.9	10.6	14.6	52.9	19.2	8.1									
FERT–10	3.3	14.7	41.8	34.6	8.9	16.6	49.0	19.7	8.4	–4.7	11.2	6.4	–6.1	–16.5	13.1	–7.4	2.7	2.8
FERT–20	3.1	16.5	44.2	31.8	7.5	18.9	44.3	20.3	8.7	–9.4	24.7	12.6	–13.8	–29.3	29.2	–16.2	5.5	6.7
FERT–30	2.9	18.6	48.0	27.9	5.6	21.8	40.2	19.5	9.3	–14.7	40.2	22.1	–24.3	–47.5	48.7	–24.0	1.4	14.4
FERT–40	2.7	20.8	52.2	23.2	3.9	25.0	36.3	17.9	10.2	–20.1	57.0	32.7	–37.1	–63.5	70.4	–31.3	–6.6	25.2
AGE+1	3.4	13.3	39.9	36.8	10.0	13.9	54.8	17.9	9.1	–.7	.6	1.5	–.1	–6.0	–5.1	3.6	–6.5	12.1
AGE+2	3.4	13.8	39.9	36.7	9.6	13.3	55.7	17.2	10.0	–1.4	4.4	1.6	–.6	–9.5	–9.4	5.3	–10.5	22.3
AGE+3	3.4	14.7	39.6	36.4	9.3	12.8	55.9	16.7	10.6	–2.3	11.3	.7	–1.4	–11.8	–12.9	5.6	–13.2	30.0
AGE+4	3.3	15.7	38.9	36.2	9.2	12.1	56.2	16.0	11.4	–3.0	18.8	–1.1	–1.9	–13.1	–17.2	6.2	–16.8	39.3
E+15	3.5	12.7	39.3	35.7	12.3	16.6	49.7	21.1	6.0	1.4	–4.2	.1	–3.2	16.2	13.0	–6.2	9.9	–25.9
E+25	3.5	11.9	39.8	34.8	13.5	18.1	47.4	22.6	5.0	2.5	–9.8	1.2	–5.7	27.6	23.6	–10.5	17.6	–38.9
F–30E+25	3.0	15.1	49.5	27.8	7.6	25.0	37.9	22.0	6.4	–11.2	14.2	25.9	–24.7	–28.0	70.5	–28.3	14.6	–20.9
F–40E+25	2.9	16.3	54.4	24.1	5.2	27.9	35.4	20.4	7.4	–16.2	23.3	38.5	–34.7	–51.1	90.5	–33.1	6.2	–9.4
DIV+100	3.4	13.7	39.3	36.5	10.5	14.4	52.8	19.2	9.3	–.6	3.5	0	–.9	–1.3	–2.0	–.3	–.1	14.5
DIV+200	3.4	14.2	39.4	36.2	10.3	14.1	52.6	19.1	10.5	–1.3	7.0	.2	–1.9	–2.9	–3.6	–.5	–.6	28.7
DIV+300	3.4	14.6	39.4	35.8	10.1	13.9	52.5	19.0	11.6	–2.0	10.4	.3	–2.9	–4.3	–5.4	–.8	–.9	42.7
DIV+400	3.3	15.0	39.5	35.5	10.0	13.6	52.4	19.0	12.7	–2.6	13.7	.5	–3.8	–5.6	–7.1	–1.0	–1.3	56.3
REM–25	3.4	13.6	39.2	36.8	10.4	14.4	52.9	19.1	8.8	–.5	2.8	–.2	–.3	–1.8	–1.6	0	–.6	7.6
REM–50	3.4	14.1	39.1	36.6	10.2	14.1	52.9	19.0	9.5	–1.0	6.3	–.4	–.7	–3.9	–3.6	–.1	–1.3	17.1
LH+20	3.3	15.1	42.1	34.0	8.7	17.1	52.3	15.5	7.9	–5.0	14.2	7.3	–7.8	–17.7	17.0	–1.2	–19.3	–3.5
LH+35	3.1	17.2	43.7	31.6	7.4	19.8	50.1	12.8	7.2	–9.3	30.3	11.2	–14.3	–29.8	35.4	–5.3	–33.2	–11.0
LH+50	2.9	20.2	44.5	29.1	6.2	23.5	46.2	10.0	6.0	–14.2	52.8	13.3	–21.2	–41.5	60.6	–12.6	–47.8	–25.9
LH+100	2.7	24.8	44.2	26.0	5.1	28.8	39.3	7.1	4.3	–20.5	87.2	12.5	–29.6	–52.3	96.8	–25.7	–63.2	–47.3

Notes as in Appendix Table 1.

Notes

This chapter is based in part on Zeng, Vaupel, and Wang (1997). New material includes the accuracy tests on our model using the data sets from recent censuses of the United States and Germany, the sensitivity analysis, and discussions. We have sought to explain the model in an accessible manner.

The research reported in this chapter was supported by the Alfred P. Sloan Foundation (United States), the National Natural Science Foundation (China), 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 at Odense University (Denmark), the Center for Population Analysis and Policy of the Humphrey Institute at the University of Minnesota (United States), and the Netherlands Interdisciplinary Demographic Institute. The authors acknowledge comments and suggestions provided by Thomas Burch and Frans Willekens. Discussions with Eugene Hammel, Thomas Pullum, and Kenneth Wachter are also highly appreciated. The authors have developed the first version of user-friendly computer software with a manual associated with the model described in this chapter. Scholars interested in applying this model and software may write to: Computer and Information Service System, Max Planck Institute for Demographic Research, Doberaner Strasse 114, 18057, Rostock, Germany.

1 The status of "coresidence with parents" is here broadly defined. A child who is living with parent(s) or grandparent(s) or other senior family members who act as care providers when parents are not available is classified as "living with parent(s)."

2 When both an adult woman and an adult man are present in a household, we chose the woman as the reference person of the 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 are more likely to live with their mother.

3 Those elderly who live together with child(ren) and grandchild(ren) in a three-generation household are not reference persons of the household, since the child of the middle-generation adult with whom they live has already taken the position of a reference person. A household cannot have two reference persons. Therefore, the number of those nonmarried elderly and female married elderly that is equal to the number of three-generation households should be subtracted when we compute the number of reference persons who represent two-generation households. The subtraction involves some reasonable approximation since we do not track whether living with grandchild(ren) in the current version of our computer software. Those who are interested in the formulas are referred to Zeng, Vaupel, and Wang (1997: 191) or Zeng (1991: 195). Given the enhanced computer capacity of memory today, we may include an additional dimension in version 2.0 of our software ProFamy: living with grandchild(ren), so that the approximation will not be needed since those who are living with child(ren) and grandchild(ren) will be clearly distinguished as non-reference persons.

4 The definition of children here is relative to parents. For example, a person aged 60 and older is still a child if he or she lives with parent(s).

5 When status of number of children living together is reduced by i, i events are accounted.

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

7 The homogeneity assumption can be relaxed by introducing more characteristics. For instance, the assumption is less strong for a fertility model that considers age, parity, and maternal status, than for one 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.

8 Total fertility rates of cohorts who were age 10 years and younger in 1990 and those new cohorts born after 1990 are assumed to be reduced by 10, 20, 30, and 40 percent respectively as compared with the 1990 level.

9 The median ages at first marriage and at parity-specific births for cohorts who were age 10 years and younger in 1990 and those new cohorts born after 1990 are assumed to increase by 1, 2, 3, and 4 years respectively as compared with the 1990 level.

References

Bell, M. and J. Cooper. 1990. "Household forecasting: Replacing the headship rate model," paper presented at the Fifth National Conference, Australian Population Association, Melbourne, November.

Bongaarts, John. 1983. "The formal demography of family and households: An overview," *IUSSP Newsletter*, No. 17 (November).

———. 1987. "The projection of family composition over the life course with family status life tables," pp. 189–212 in John Bongaarts, Thomas K. Burch, and Kenneth W. Wachter (eds.), *Family Demography: Methods and Their Application*. Oxford: Clarendon Press.

Brass, W. 1983. "The formal demography of the family: An overview of the proximate determinants," in *The Family*, Occasional Paper 31. London: Office of Population Censuses and Surveys.

Burch, Thomas K. and Andrejs Akaburskis. 1993. "Projecting household headship: Exploration and comparison of formal and behavioural approaches," Project Summary for Research Division, Canada Mortgage and Housing Corporation.

Hammel, E. A. et al. 1976. "The SOCSIM demographic-sociological microsimulation program," Institute of International Studies, Research Series No. 27, University of California, Berkeley.

Kannisto, Väinö, Jens Lauritsen, A. Roger Thatcher, and James W. Vaupel. 1994. "Reduction in mortality at advanced ages: Several decades of evidence from 27 countries," *Population and Development Review* 20(4): 793–810.

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

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

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

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

Ledent, Jacques. 1992. "Vers des perspectives de familles/ménages sur la base d'un modèle de type multidimensional," unpublished manuscript.

Lutz, Wolfgang and Christopher Prinz. 1994. "The population module," pp. 221–231 in W. Lutz (ed.), *Population-Development-Environment: Understanding Their Interactions in Mauritius*. Berlin: Springer-Verlag.

Mason, A. and Rachel Racelis. 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 M. J. Murphy and J. Hobcraft (eds.), *Population Research in Britain*, Supplement to *Population Studies*.

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

Pollard, J. H. 1977. "The continuing attempt to incorporate both sexes into marriage analysis," pp. 291–310 in *International Population Conference: Mexico 1977*. Liège: IUSSP.

Rogers, Andrei. 1975. *Introduction to Multi-Regional Mathematical Demography*. New York: John Wiley.

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

———. 1988. *Modeling Multigroup Populations*. New York: Plenum Press.

Smith, James E. 1987. "The computer simulation of kin sets and kin counts," pp. 249–266 in John Bongaarts, Thomas K. Burch, and Kenneth W. Wachter (eds.), *Family Demography: Methods and Their Application*. Oxford: Clarendon Press.

Spicer K., Ian Diamond, and Máire Ní Bhrolcháin. 1992. "Into the twenty-first century with British households," *International Journal of Forecasting* 8: 529–539.

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

Van Imhoff, E. and N. Keilman 1992. *LIPRO 2.0: An Application of a Dynamic Demographic Projection Model to Household Structure in The Netherlands*. Netherlands: Swets and Zeitlinger.

Vaupel, J. and Yi Zeng. 1991. "Population tradeoffs in China," *Policy Science* 24: 389–406. Revised version forthcoming in *Scientific American*.

Wachter, Kenneth W. 1987. "Microsimulation of household cycles," pp. 215–227 in John Bongaarts, Thomas K. Burch, and Kenneth W. Wachter (eds.), *Family Demography: Methods and Their Application*. Oxford: Clarendon Press.

Willekens, Frans. 1987. "The marital status life table," pp. 125–149 in John Bongaarts, Thomas K. Burch, and Kenneth W. Wachter (eds.), *Family Demography: Methods and Their Application*. Oxford: Clarendon Press.

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

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

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

———. 1991. *Family Dynamics in China: A Life Table Analysis*. Madison, WI: University of Wisconsin Press.

Zeng, Yi and James W. Vaupel 1989. "The impact of urbanization and delayed childbearing on population growth and aging in China," *Population and Development Review* 15(3): 425–445.

Zeng, Yi, James W. Vaupel, and Wang Zhenglian. 1995. "A multidimensional model for projecting family households—with an illustrative numerical application," *Population Studies of Aging*, No. 17. Center for Health and Social Policy, Odense University, Denmark.

———. 1997. "A multidimensional model for projecting family households—with an illustrative numerical application," *Mathematical Population Studies* 6(3): 187–216.

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, Institute of Population Research, Peking University.

Library of Congress Cataloging-in-Publication Data
Frontiers of population forecasting / Wolfgang Lutz, James W. Vaupel,
Dennis A. Ahlburg, editors.
p. cm.
Includes bibliographical references (p.).
ISBN 0-87834-096-3 (pbk. : alk. paper)
1. Population forecasting. I. Lutz, Wolfgang. II. Vaupel, James
W. III. Ahlburg, Dennis A.
HB849.53.F76 1999 99-21220
304.6'01'12--dc21 CIP

ISSN 0098-7921

ISBN 0-87834-096-3

Printed in the United States of America.