

Reference

- Le Bras, Hervé and Kenneth W. Wachter (1978), 'Living Forbears in Stable Populations.' In K. W. Wachter, E. A. Hammel, P. Laslett (*et al.*), *Statistical Studies of Historical Social Structure*. Academic Press, New York, 163–88.

14 Some Mathematical Models of Kinship and the Family

THOMAS W. PULLUM

Population Research Center, University of Texas, Austin, Texas, USA

1 Introduction

Mathematical techniques and models have been used in a great variety of ways within the context of the study of kinship and the family. All the chapters in the present section of this volume, and many of those elsewhere in the volume, use mathematics to formalize a set of assumptions and to generate the implications of these assumptions. The value of mathematics in this context, as elsewhere in the social sciences, is essentially to provide a language with which abstract structures and processes can be described relatively unambiguously, as well as to provide techniques for manipulating those abstractions. This chapter will attempt to describe specific instances in which formal modelling of kinship, the family or the household has been useful and holds the possibility of further utility.

An excellent brief summary of the objective of mathematical modelling in the social sciences has been given by Herbert Simon (1957). A comprehensive inventory of the techniques in general use, with a number of social science applications, has been provided by Fararo (1973). More specialized models in population study have been described by Keyfitz (1968, 1977).

The minimal requirements for modelling are simply stated. First, there must be a set of units. These can be individual human beings, or abstractions such as household heads, families, households, relationships, lines of descent, and so on. Second, there must be a set of states, either discrete categories or continuous qualities, which differentiate those units. Sometimes the articulation between the units and the states they occupy can be problematic. For example, marital status is a classificatory property of individuals, but if the unit of analysis becomes the married couple, then the married state is a defining characteristic and not a variable. There is a trade-off between the use of characteristics to define units and their use to differentiate units. Generally, there is some gain in flexibility if units are defined as simply as possible, and characteristics are seen as variables rather than defining qualities.

A third required feature of almost all models is some process, mechanism,

or dynamic by which the units can change from one state to another. There may be an explicit time dimension, as in the familiar Lotka model of population growth, or time may be implicit and unspecified, as in structural equations models, but there are generally some formalized statements about how change occurs or is caused. What are the likely elements of models in the area of family demography, as distinct from the demography of individuals? It might at first seem surprising that each of the models used as illustrations in this chapter takes the individual as the unit of analysis. It will be argued that compound units, such as the couple or the family, tend to be more difficult to manipulate because they can dissolve or recombine over time. The distinctive character of the family enters through the kinds of states or variables which are attributed to individuals. For example, the states can consist of different roles within the household, such as household head; marital or parental statuses, such as number of children ever born; length of time in a role or status; and so on. The individual can also be assigned properties which derive from an aggregate of which he or she is a member, such as the size of the household, occupation of the spouse, the influence of other family members, and so on. The third element of the models, the forces which describe or cause a change in states, are rooted in household dynamics and demand the greatest ingenuity. This can be implemented with a variety of techniques, which will depend for example upon the kinds of states, whether or not a time dimension is included, and the number of variables, but conceptual clarity is always the principal challenge.

This chapter will review a set of research traditions which formally represent kinship and the family, arranged to highlight structural features first, and then passing on to increasingly dynamic models. This sequence will also lead from the rather diffuse notion of kinship as a part of the general social environment toward the more immediate and integrated unit of the co-resident family. The starting point is the modelling of kinship by mathematical anthropologists.

2 Representation of kinship structures

Within the tradition of anthropological research, one of the most important tasks of the ethnographer is to record information on kinship relationships. Thus, the ethnographer lists the labels which are used for the various biological categories of kin and anticipates or calls attention to the social implications of these labels. In English, for example, one's (first) cousin is any person with whom one shares a grandparent, regardless of the gender not only of that person but also of the grandparents and of the intervening parents. It would be possible to distinguish four or even more different kinds of first cousins. That these distinctions are not made in English is correlated with the cultural fact that English-speaking countries, and Western cultures more generally, even in the historical past, do not distinguish their obliga-

tions to or their expectations from these potentially differentiable cousins. In fact, without exception, English kin terms are bilateral, distinguishing at most the gender of the kin category (e.g. son or daughter, father or mother) but never the gender of the intermediate kin.

Recognized or socially meaningful categories of kin arise from normative interpretations of biological relationships. During the past three decades, considerable effort has been invested by anthropologists in the development of formal models to describe structures within which there is a social equivalence of kin who are quite different in their biological relationships. These structures are built up from the following kinds of principles. First, individuals at their birth are assigned membership in one of a set of moieties, clans, etc. Second, individuals are assigned these memberships on the basis of memberships of one or both parents. Third, one's membership at birth is a determinant of whom one is eligible to marry.

These algebraic models are based on the individual as the unit of analysis, clan membership or identity as the attribute, and descent and marriage as the processes which determine these identities. The limitation on potential marriage partners is a consequence of one's identity but is also a system dynamic which leads to the birth and identity of subsequent individuals.

Overlaid on this structure is a set of terms which link a selected individual, male or female, to his or her kin. In a traditional culture with a clan structure, these terms involve an elaboration of the kinds of terms used in the West (more precisely, their linguistic equivalents), incorporating some of the distinctions which are suppressed in the West plus some distinctions related to clan identity. In cultures which employ or formerly employed these elaborate distinctions, they are the fundamental basis of social organization and role relationships.

We shall not describe these models in detail. They require specialized terminology and some familiarity with group theory. They also tend to be highly idealized. Relatively accessible presentations are given by White (1963), the highlights of which are given also by Fararo (1973, 525-45). Instead, we shall summarize succinctly some recent developments in this line of research which begin to link it with family demography. The contribution of this recent research rests upon its empirical basis — using data to compare actual terminology with the normatively expected terminology — and in the inclusion of preferred age at marriage as an influence on the differences between these two terminologies.

Denham, McDaniel, and Atkins (1979) represent an effort to determine whether an elaborate (by Western standards) normative structure of kin relations in the Alyawara tribe of Australian aborigines is actually followed in practice. This structure, a classic in the literature on kinship, has 22 distinct kinship terms which were specified by Radcliffe-Brown. Denham, having elicited biological genealogies for 264 informants, asked these

informants to apply the correct Alyawara kinship term when presented with photographs of their kin. The expected term predicted by Radcliffe-Brown's classification failed to match the reported term for many of the intermediate and distant kin. For example, in a high proportion of cases, the term which should have been applied to the mother's brother's daughter (MBD), actually quite a close and important relation in this structure, was not used, but was replaced by another one.

The Alyawara were also found by early ethnographers to have two strong preferences regarding marriage (within the stipulation of exogamy). The first was for bilateral second cross-cousin marriage. That is, a male would prefer to marry his mother's mother's brother's daughter's daughter (MMBDD). Denham's data showed clearly that first cousins, especially matrilineal cross-cousins, are preferred to second cousins. Second, the Alyawara were claimed to prefer sibling exchange marriage, whereby a brother and sister would marry a sister and brother from another family. Among the 114 marriages in the data base, there were no instances of sibling marriage. In short, Denham *et al.* found that although 'the model is internally consistent, and it is compatible with Alyawara and Aranda ideology . . . it does not correlate well with actual Alyawara kinship term applications and marriage practices.' (1979, 16).

Denham *et al.* propose a modification to the original model which takes into account the preferred age difference between spouses. Hammel (1976) and a few other anthropologists since Rose (1960) have noted that if the husband is to be substantially older than the wife, then certain types of marriage rules are difficult to implement, if not impossible. In the case of the Alyawara, Denham's fieldwork found that the average age difference between spouses was 14 years. Moreover, since a mother is on the average 28 years older than her children, it followed that a father will on the average be 42 years older than his children. It is then easy to see that a male will on the average be 14 years older than his MBD first cousin, i.e. of the correct age difference for marriage, but he will on the average be 14 years younger than his FZD first cousin and 28 years older than his MMBDD second cousin, sharply reducing the acceptability of such marriages. Sibling exchanges will also be difficult to accomplish because they must be limited to pairs of siblings in which a brother is either much older or much younger than a sister.

Denham's co-author, Atkins, proposed a scheme by which these mean age differences can be incorporated into the lattice of a genealogy, taking it into the three-dimensional form of a double helix. In the case of the Alyawara, where the mean age difference between spouses is half the mean maternal age at childbearing, the diagram is relatively simple (although not simple enough to permit a detailed description here). The diagram shows that the age difference is responsible for the slippage in kin terminologies noted above, as well as for the rarity of MMBDD marriages. The MBD first cousin

tends to be in the same age range as the individual's wife or his daughter, and it is their category labels which tend to be substituted for the proper one. Unfortunately, the Denham *et al.* study does not indicate whether the actual ages, of those individuals for whom the terms are misapplied, conform to the pattern of age deviations in the model, but this could be regarded as a hypothesis for future work. One might even seek evidence that in Western cultures and terminology, there will be a tendency to refer to a late-born aunt or uncle as a cousin, or to refer to a late-born cousin as a niece or nephew, or at least to behave toward them as if that were the role they occupied, even though fully aware of the actual biological relationship.

The impression should not be given that the specific research reviewed here in some detail is unique. In addition to Hammel (1976), a few other anthropologists such as Fredlund and Dyke (1976) have used analytic methods as well as computer simulation for such purposes. This line of research has also been extended by Tjon Sie Fat (1981). He presents a powerful algebraic model which subsumes a broad range of marriage and descent rules and has the possibility for even greater generalization.

At this point there exists no model which adequately incorporates age preferences with normative marriage preferences in these settings — at least, no analytical model, as distinguished from simulation. And, of course, the use of the mean age difference between husbands and wives is an extremely crude demographic consideration. The numbers of kin of the relevant types and their age distribution are important as well, and the distribution or at least the acceptable range of husband-wife age differences would be preferable to the mean.

This discussion has considered a type of modelling which may appear to have limited application to modern industrialized societies. Nevertheless, some of the fundamental features of these ethnographically-derived models remain central today. We are not now encouraged to marry first or second cousins, for example, but we may, for intervals of time, include kin in our households or in a larger network of regular interaction, exchange, and support. A general model which allows for the availability of kin for such relationships, through demographic variables, and also specifies a structure of preferred or socially mandated relationships, with some room for substitutions, is a potential consequence of this work by mathematical anthropologists. Family demographers have found ways to deal with the issues of whether kin are actually available, so to speak, for incorporation into a household or into regular interaction. The next section will describe some of this work. They have tended to ignore, at least in formal modelling, a normative structure corresponding in any way to the algebraic models to which we have alluded. Such a structure could prioritize the biological roles in terms of their importance for support and interaction after ego has reached maturity.

3 Numbers of kin

Analytical models for frequencies of kin are of two types. The first type is exemplified by the work of Le Bras (1973) and of Goodman, Keyfitz, and Pullum (1974, 1975). It concerns the expected number of kin, ever-born or still alive, at specific ages of an individual, in a population. This research has only taken a deterministic format, and has used fixed age-specific schedules of fertility and mortality.

The deterministic models began with Lotka (1931), who wished to calculate the probability that a randomly selected woman of specified age a , would have her mother still living, under specified schedules of mortality and fertility. He reasoned that this probability, $M_1(a)$, would be given by:

$$M_1(a) = \int_{\alpha}^{\beta} l(a+x)m(x)e^{-rx}dx,$$

where $l(x)$ = the probability of surviving from birth to age x ,
 $m(x)dx$ = the probability of a daughter being born to a woman in the age interval $(x, x + dx)$,
 r = the intrinsic rate of natural increase, and
 (α, β) = the lower and upper age limits of childbearing.

The logic behind this formula is not difficult. The age of the mother at the time of the woman's birth is taken to be x , so that her age (if she still lives) when the respondent is age a , will be $a + x$. Let $M_1(a|x)$ be the probability that a woman now aged a , whose mother was aged x at the birth, will have a currently living mother; this will simply be $l(a+x)/l(x)$, the chance of a woman surviving from age x to $a+x$. The unconditional probability, $M_1(a)$, can then be obtained by multiplying $M_1(a|x)$ by the probability that a woman aged a , will have been born to a mother age x , and integrating over x . In a stable population this latter probability will not depend at all upon the value of a , and will simply be the product of the maternity function, $m(x)$, and the stable age distribution, $l(x)e^{-rx}$. The formula for $M_1(a)$ then follows. Lotka's orphanhood formula was used in reverse by Brass (1953). Vital rates can be estimated from orphanhood levels if the latter are available at several ages a and if the survivorship and maternity schedules can be parameterized.

Goodman *et al.* (1974) generalized Lotka's orphanhood formula in two major directions. First, they extended it to the probabilities of survivorship of more distant progenitors (Le Bras, 1973, dealt independently with grandparents), and the expected numbers of other close kin, ever-born and still alive, by age a , of the respondent. They showed, in principle, how a formula such as that for $M_1(a)$ can be developed to relate the individual to any category of kin whatsoever. Second, they moved away from the assumption of a stable population. Thus, the more general form of $M_1(a)$ is

$$M_1(a) = \int_{\alpha}^{\beta} l(a+x)/l(x)W(x|t-a)dx,$$

where $W(x|t-a)$ is the age distribution (at time $t-a$) of the women who gave birth to a daughter at time $t-a$. Goodman *et al.* also calculated the probabilities of surviving progenitors and the expected frequencies of other kin under illustrative stable population regimes.

The strength of such work lies in its analytical coherence, supplemented by the computational convenience of the computer program to implement the formulas. It is possible to estimate the sensitivity of the various expected numbers to changes in the schedules of survivorship and fertility. Also, emulating the applications by Brass (1953) of the orphanhood formula, it is possible to obtain indirect estimates of vital rates from the empirically observed numbers of kin. Goldman (1978), Wachter (1980), and McDaniel and Hammel (1984) have developed estimates of the intrinsic growth rate using the formulas for numbers of older sisters and younger sisters of the individual.

The weaknesses of the deterministic formulas arise from their simplifying assumptions, some of which were explored in Goodman *et al.* (1975). The most serious assumptions, in terms of explaining why the calculated numbers can deviate substantially from the true numbers, merit discussion here because they apply to most models of family building. Of the four most serious assumptions, the first is demographic stability. This requirement was relaxed in the Goodman *et al.* generalization, in that the age distribution is not required to be fixed, but even so the $l(x)$ and $m(x)$ functions are not indexed by time and are therefore not subject to change during the lifespans of pertinent kin. Even this limited kind of stability will have severe implications for relationships which are traced across more than one generation.

The second weakness of this approach is that it is not based on parity progression. It is assumed that a fertility rate — more precisely, the chance of having a child during an interval of time and age — does not depend upon the number of children a woman has already had. This is the basic assumption of a Poisson process. The more fertility is controlled by delayed marriage and contraception, the more serious this assumption becomes. In a controlled setting, fertility can be broadly dispersed across ages, yet with relatively little variation in completed family sizes. The formulas will imply a broader dispersion in numbers of children and all other kin than is empirically observed, and this will affect the estimated mean frequencies because of the complex linkages between kin categories.

Thirdly, the deterministic model assumes that the population is homogeneous. In reality, of course, it is stratified by fecundity and by fertility targets. The overall observed schedules are implicitly averaged across different strata, but the expected values produced by the formulas need not be the averages across these strata. The fourth major weakness is the limitation to a single sex: the individual, the kin of interest, and all intermediate kin must be females. However, most formal demographic modelling is limited to a single sex, and even if the so-called two-sex problem were solved con-

ceptually, the necessary input data for a two-sex model would almost never be available.

Formally, it would be possible to extend the formulas of Goodman *et al.* to make the rates time-dependent, order-specific, and stratified. (A two-sex formulation would be difficult.) But if this were done, the data requirements would become immense, and in addition the heuristic or conceptual value of the model might be eroded. If more complex data are available, then one may wish to turn to simulation, as described by Reeves in chapter 12 of this volume.

The limitations listed above for the deterministic models, and other limitations for other formal models, must also be kept in perspective. A model is not intended to serve just as a computing algorithm or as an accounting scheme, although algorithms and schemes can be very useful. Rather, a model is an effort to explicate a process. Much of its value comes from simplifying a process to its most important features, and then using standard mathematical techniques to simulate behaviour. Almost every useful model then goes through subsequent revisions, adding complexities which utilize available relevant data, in order to replicate better the mechanisms by which inputs are transformed into outputs in the real world. The deterministic models of kin frequencies have not yet seen such revisions, largely because the output quantities which the model seeks to estimate, as well as more elaborate or detailed input data, are all unavailable empirically.

The second type of model here is stochastic. The defining feature of a stochastic model is that basic events, such as having a birth or dying at a specific age, occur randomly with specified probabilities. These probabilities imply full distributions of output quantities, such as numbers of children ever born, and not just the mean values of these distributions. In the present context, a stochastic model gives the distribution of kin of each type under stable population assumptions. Time is marked by the length of a generation rather than by years of time or age. The stochastic models do not extend to the distribution of the number of survivors in a specific kin category at each age of the individual, which would correspond to the expected values developed by Goodman *et al.* These models have been developed out of the theory of branching processes. We shall briefly review their basic structure and suggest how they may be extended and used.

The adaptation of branching processes appears to have been made only recently by Waugh (1981) and Joffe and Waugh (1982) and independently by Pullum (1982), even though the theory has been available for many years. The starting point is the 'reproduction' probability generating function (p.g.f.) of a randomly sampled woman at her birth. If p_i is the probability, defined at birth, that a woman will have exactly i female births during her lifetime, then this probability generating function is given by $f(x) = p_0 + p_1x + p_2x^2 + \dots$. The moment of this distribution can be obtained

from successive derivatives of $f(x)$ at $x = 0$. The first moment, the mean, $p_1 + 2p_2 + 3p_3 + \dots$, is the Net Reproduction Rate.

The papers by Waugh (1981) and by Joffe and Waugh (1982) use the probability-generating function to yield the distributions of all kin in generation k for a woman sampled in generation n . For example, if the woman is sampled in generation 1, then there will be one person in generation 0 (her mother); she and her sister will comprise the set in generation 1; and her daughters and nieces will comprise the set in generation 2. If the woman is sampled in generation 2, then there will be one person in generation 0 (her mother's mother); her mother and aunts in generation 1; she, her sisters, and her cousins in generation 2; and so on. (This is stated as a one-sex model.) They give the formulas for the moments of these distributions and also the probabilities of each specific combination of numbers of kin in each generation. With the latter formulas, for example, it is possible to calculate the probability distribution of the number of daughters plus nieces, conditional on a specified number of sisters. Waugh (1981) considers the case of a large population with time-homogeneity and independence of individuals. Joffe and Waugh (1982) relax these assumptions, although the probability generating function $f(x)$ remains the same in each generation.

Pullum (1982) approaches essentially the same problem with reference to specific categories of kin, rather than entire generations, giving the probability-generating functions for sisters, aunts, nieces, cousins, etc., distinct from other kin. He also gives illustrations of these distributions in a hypothetical setting of replacement fertility. The distributions are of 'eventual' frequencies of kin, a generalization of the notion of reproductivity in the first generation. One obtains the probability of exactly n kin of a specified type, say, consisting of all those persons who will ever be born into this type, regardless of how long they may live or how long the individual may live. When mortality is high, there is a high probability that a woman will produce no births at all (because she will not herself survive to childbearing age) and a corresponding high probability that she will have no descendants of other more remote types. Of course, kin types which do not depend upon ego's survival will be less affected by mortality. When mortality is low, as in developed countries, the eventual frequencies of kin will not be a great deal higher than the number who are actually available for interaction, making reasonable allowances for generational differences.

Other related uses of branching processes deal with the extinction of kin lines, the probability of being a first born child (Jagers, 1982), and so on.

What is the value of these models? First, they have a complementary relationship to the structural models of the preceding section. With some information about reproductivity, one can actually attach frequencies to the nodes in the lattices of kin categories. These can be summed in various ways, corresponding to culturally recognized combinations of relationships, such as 'first cousin.' Numbers like these are the demographic component of

kinship referred to at the end of the previous section. Second, they permit estimation of the numbers of kin who are available for participation in kin-based households or communication networks. An individual has a chance of living with or depending upon or communicating with a specific type of kin, which is approximately proportional to the number of living kin of that type. Third, the models allow estimation of the proportion of the population which have no or few kin of a specific type. This sub-population will have limited opportunities for certain kinds of socialization as children or limited opportunities for social support when they are older. When fertility is low, and as childlessness becomes more acceptable again in America and Europe, increasing proportions of people will have no descendants for support in their old age. Small families mean that fewer children will have siblings of the same sex and fewer will have siblings of the opposite sex. These models permit one to quantify such obvious implications of low fertility. Fourth, models permit one to estimate the sensitivity of these numbers to changes in the vital rates or reproduction p.g.f.

The available models actually fall short of these potential uses. As noted, the branching process models do not include age or mortality, and the models which include age and mortality are deterministic and do not produce probability distributions. A reconciliation of these approaches seems likely in the near future, however, at the very least through a hybridization.

A different kind of limitation is that these models characterize an aggregate and do not include any more substantive components. In the investigation of theoretical propositions about kin contact, for example — say, the null hypothesis that after marriage a couple will have equal contact with the wife's kin and the husband's kin, taking account of availability — both contact and availability must be measured empirically for a sample of individuals. A model could suggest how the distribution would look, and might permit some generalizations across broad sweeps of time or social groupings for which empirical data are incomplete, but it would be insufficient for the testing of such a hypothesis. To repeat, these models are primarily of descriptive value at an aggregate level and enable one to see the consequences of changes in vital rates, as mentioned above, but their present formulation does not allow inclusion of non-demographic variables.

4 Family size distributions

The preceding section referred to the reproduction probability-generating function as the basic source of the eventual frequencies of kin in a stable population. Apart from such a role, this function has interest in its own right. Particularly in developed countries, in which the nuclear family predominates and child mortality is low, the number of children per woman will be the foremost determinant of the size of the co-resident family.

The reproduction probability-generating function refers to the daughters ever born to a female, defined at the point of her own birth. We shall also consider the distribution of children ever born (sons and daughters) defined at the point when the woman reaches age 45, i.e. has completed and survived the childbearing period. The linkage between these two distributions is described briefly in Pullum (1982). If a woman has had *k* children, then the maximum size of her nuclear family at any point will be *k* + 2, including herself and her spouse. It will be smaller during those years before all the children were born, or after they begin to leave home, or as a result of parental or child mortality. Divorce and separation and other kinds of family decomposition and recomposition may also be important, but the distribution of children ever born will be the major source of variation in co-resident family size during the part of the family life cycle when most individuals are being socialized.

The completed parity distribution has undergone truly dramatic changes in recent generations in Western countries, and in the past decade or so in several developing ones. It is remarkable, in view of the extent of this change, that the distribution is so rarely described by anything other than its mean.

Table 14.1 gives the percentages of women who had 0, 1, 2, 3, or 4 or more children, out of those cohorts of American women who were age 45–59 in 1900, 1925, 1950, and 1975. Among the earlier cohorts, survivorship is certainly an important factor for any translation of these figures to household membership, and we shall not attempt to make such a translation here. However, there can be no doubt that some of the major features of these differences have had in impact on the co-resident family size.

Table 14.1 Children ever born to women cohorts (USA)

Age 45–59 in:	Number				
	0	1	2	3	4 or more
1900	14	11	13	12	50 percentage
1925	22	12	14	12	40
1950	20	22	22	14	22
1975	9	11	24	22	34

Particular attention is drawn to the changing frequency of no children or one child. These two parities accounted for about 44 per cent of the American women who completed their childbearing at the middle of this century, a figure which cannot be attributed to the Great Depression of the 1930s because it was part of a long-term trend. A major factor in the post-war baby boom was the sharp reduction of these two lowest parities, with two children becoming the effective minimum. There are indications, of course, that the lowest parities are again becoming relatively common,

whether because of non-marriage or for voluntary reasons within marriage.

Mathematical models can be used to describe this distribution, how it varies from one setting to another, and how it changes over time. For example, these models show that the increasing concentration or lack of dispersion in the family size distribution in the United States is not simply a consequence of the decline in the mean; it is a manifestation of greater consensus — whether normative or circumstantial — that the number of children should occupy a narrow range: not too small and not too large. This concentration sharply contrasts with the great dispersion in parities which characterizes high fertility settings, both historically and at present in developing countries. When fertility is high, the number of children ever born as well as the number surviving to any specific age is always highly dispersed. The World Fertility Survey, in its survey of ever-married women in the Philippines in 1978, found a mean of 5.5 children ever born to women aged 40–44, but 12 per cent had 0–2 and 18 per cent had 8 or more. The 1976 Colombian survey of all women showed an even more dispersed distribution for women aged 40–44. Their mean was 5.1 children, but 18 per cent had 0–2 and 22 per cent had 8 or more. These are only illustrations from a larger set of distributions (Hodgson and Gibbs, 1980).

Mathematical models to describe these patterns can take different forms. Lotka (1939, 121–2) was one of the first demographers to model the reproduction distribution. He proposed a two-parameter function which could be written as follows: $p_0 = a$ and $p_i = br^i$. (Since these proportions must sum to unity, there are in effect only two free parameters.) This function implies that all successive pairs of family sizes have the same ratio 1:r, with the exception of p_0 and p_1 which have the ratio $a:b$. Waugh (1981) has noted that this function has some desirable properties for mathematical manipulation. However, it has a very weak behavioural interpretation and does not conform well to a broad range of available data. This distribution can only have a mode of $i=0$ or $i=1$ (daughters) and it is unable to describe real settings where fertility is much above replacement level.

Shifting again from reproductivity to completed childbearing, several demographers, notably Henry and Ryder, have advocated the transformation or representation of the distribution with parity progression ratios, the conditional probabilities of attaining successively higher parities. Ryder (1982) and Feeney (1983) have recently proposed refinements of the PPRs which incorporate lengths of birth intervals.

The present author has reviewed a large number of family size distributions under a variety of transformations. A particularly promising representation is based on an analogy between childbearing and survivorship in the usual life table. In this analogy, 'age' is marked by successive births, and the termination of childbearing, i.e. the failure to progress to the next higher parity, is the decrement. 'Survivorship' at parity i , say, is then measured by the proportion of women in a cohort who ever achieved parity

i , beginning with a radix of 1 for parity 0. This survivorship function declines monotonically until the highest parity achieved by the cohort, just as the life table function $l(x)$ declines monotonically up to the highest age at death. This representation has another desirable property, in that the 'survivorship' function can be calculated at any point in a cohort's experience, and it converges toward the final distribution at a rate which describes the tempo of childbearing.

Chiang and van den Berg (1982) have developed many of the implications of this analogy. There is evidence that this representation of childbearing can be reduced to very few parameters, capitalizing upon the research of Brass and others on mortality, e.g. through the relational logit transformation of the survivorship function. These parameters can be interpreted as statistical measures of the mean, the dispersion, and the skew, say. Preferably, they can be re-stated to have more substantive interpretations, regarding the normative acceptability of very low parities, the consensus about the most desirable parities, and so on.

Models which can summarize or reduce the information in the childbearing pattern, especially with the inclusion of timing components, have the potential for integration with models of the family as a whole. Brass, building upon earlier work of Burch (1970), has developed a model for family size which uses the age-specific fertility rates of an indexed woman in the household. It seems that a re-casting of such models, using parity progression as captured by a few interpretable parameters, is necessary if they are to be extended to include full distributions, not just the mean. This is because age-specific fertility rates can be used to calculate mean parity, completed or cumulative to a specific age, but they are not sufficient to estimate the distribution of parities. Krishnamoorthy (1979) has developed Poisson-type approximations to the parity distribution based on age-specific fertility rates, and although these approximations may be satisfactory in a setting of natural fertility, they are much too broad when fertility is controlled by contraception. The same point was made in the discussion of models of kin frequencies.

5 Family transitions

We have discussed some topics in the demography of kin which have narrowed the focus from the structure of kinship categories to the frequencies of any kind of kin and then to the frequency distributions of children. The final stage of this sequence is the co-resident family or domestic group, as contrasted with the larger set of blood relatives. One can, of course, conceptualize intermediate groupings which depend upon levels of dependency or communication, say, but they will lack the objective clarity of the two extremes, the kin structure and the co-resident family.

Several chapters in this volume, including those by Willekens and

Feichtinger present new models of family statuses and transitions from one status to another. With these authors, we prefer to treat the family in which an individual resides as a characteristic of that individual. An entity which is subject to the decomposition and recomposition that characterize the family can be an awkward choice, in itself, for the unit of analysis. There is no loss of information if one attaches to each individual, at each year (or some other interval) of age, the characteristics of his or her family, in as much detail as desired, and including the role of the individual (e.g. relation to head) within the larger unit. Such an approach admittedly duplicates much information. One variation is to identify a single individual within the family as the 'marker' to achieve a one-to-one correspondence between families and a subset of individuals. Brass (1983) has proposed the use of a marker, specifically an adult woman, and has ingeniously linked the life course of the family to this woman.

Without attempting to summarize these new approaches to the life course of the co-resident family, so to speak, one can indicate in broad terms the potential of such modelling. First, the specification of statuses or dimensions and their possible inter-connections will allow one to conceptualize transitions more clearly. Currently, Grady and McLaughlin (1983) are using the US Current Population Surveys of the 1970s to identify the sources of increases in the population living alone. Single-person families or households typically arise when, say, one member of an elderly couple dies; or a younger couple are divorced, and their children remain with one parent, so that one parent lives alone; or an adult child leaves home to live alone. There are other sources as well. One can apply the notion of risk: individuals have a chance of moving into the state of 'living alone' which depends upon their current state. The increase in living alone in the 1970s may have been due to a change in the size of the prior groups at risk or due to an increase in the rates of transition out of these risk groups. The distinction between these two sources of change involves a careful specification of the pathways between different family statuses.

Second, as in stable population theory and in much of multi-dimensional mathematical demography, one can generate synthetic or stable distributions which serve to summarize complex schedules of rates. This is precisely the character of the mean family sizes which Brass computes under alternative schedules of mortality and fertility. Such distributions should not be interpreted as projections, but by comparison with current observed means and distributions, inferences can be made about trends in transition rates.

Thirdly, such models allow one to investigate the sensitivity of the overall distribution to changes or disturbances, particularly hypothetical ones, in the underlying rates. Some of these rates are of the traditional demographic type, such as mortality rates; others pertain more specifically to household formation and change, such as the rates at which adult children leave the parental home and elderly parents move in with their children. By simulation

or possibly by more analytic methods, a particular type of rate can be changed, by a plausible amount, with everything else held constant, and the impact on the synthetic or stable distribution can be evaluated.

At present, household models do not explicitly include a linkage to the larger network of kin. Heuristically, and with reference to the concluding comments in section 2, it would ideally be possible to take account of the number (and possibly age and accessibility) of kin who have some risk of joining the family. In a model for the Alyawara, say, it would be particularly important to incorporate kin categories which are allegedly favoured for marriage. Cultural preferences would be expressed by the actual rate at which persons were transferred to the 'spouse' category. Some kin, such as elderly parents, siblings, etc., are also at risk, not of marriage, but of co-residence. Here, as in the work of Grady and McLaughlin, one wishes to distinguish between the fact of availability, the 'denominator at risk', and the actual rate of transition, of joining the chosen individual's family. At a more practical level, such information is difficult to obtain empirically, particularly in developing countries where the extended family occurs with non-trivial frequency.

6 Conclusion

This chapter has not presented new models or empirical analyses, in contrast with most of the chapters of this volume. It has had the rather more modest objective of conveying, in a brief space, the character of formal modelling of kin networks and family formation. We have tried to indicate, with almost no explicit use of mathematics, a broad range of issues. These issues touch upon both the nature of modelling itself and, more importantly, the conceptual linkages between kinship structures, frequencies of kin, the child-bearing of individual women, and the formation of the co-resident household.

References

- Bongaarts, J. (1983), 'The Formal Demography of Families and Households: an Overview', *Newsletter*, 17. International Union for the Scientific Study of Population, Liège.
- Brass, W. (1953), 'The Derivation of Fertility and Reproduction Rates from Restricted Data on Reproductive Histories', *Population Studies* 7, 137-66.
- (1983), 'The Formal Demography of the Family: An Overview of the Proximate Determinants', in *The Family*, *British Society for Population Studies, OPCS Occasional Paper*, No. 31, Office of Population Censuses and Surveys (OPCS), London, 37-49.
- Burch, T. K. (1970), 'Some Demographic Determinants of Average Household Size: An Analytic Approach', *Demography* 7, 61-9.
- Chiang, C. L., and B. J. van den Berg (1982), 'A Fertility Table for the Analysis of

- Human Reproduction', *J. of Mathematical Biosciences* 62, 237-51.
- Denham, W. W., C. K. McDaniel, and J. R. Atkins (1979), 'Aranda and Alyawara Kinship: A Quantitative Argument for a Double Helix Mode', *American Ethnologist* 6, 1-24.
- Fararo, T. J. (1973), *Mathematical Sociology*, Wiley, New York.
- Feeney, G. (1983), 'Population Dynamics Based on Birth Intervals and Parity Progression', *Population Studies* 37, 75-89.
- Fredlund, E. V., and B. Dyke (1976), 'Measuring Marriage Preference', *Ethnology* 15, 35-45.
- Goldman, N. (1978), 'Estimating the Intrinsic Rate of Increase of a Population from the Average Number of Older and Younger Sisters', *Demography* 15, 499-508.
- Goodman, L. A., N. Keyfitz, and T. W. Pullum (1974), 'Family Formation and the Frequency of Various Kinship Relationships', *Theoretical Population Biology* 5, 1-27.
- , N. Keyfitz, and T. W. Pullum (1975), 'Addendum to Family Formation and the Frequency of Various Kinship Relationship', *Theoretical Population Biology* 8, 376-81.
- Grady, W. R., and S. D. McLaughlin (1983), 'Living Arrangements Among the Never Married: Changes in the Propensity to Live Alone in the United States', presented at annual meetings of the Population Association of America.
- Hammel, E. (1976), 'The Matrilateral Implications of Structural Cross-Cousin Marriage', in E. Zubrow (ed.), *Demographic Anthropology: Quantitative Approaches*, University of New Mexico Press, Albuquerque.
- Hodgson, M., and J. Gibbs (1980), 'Children Ever Born', *Comparative Studies No. 12*, World Fertility Survey.
- Jagers, P. (1982), 'How Probable is it to be First Born? And Other Branching-Process Applications to Kinship Problems', *J. of Mathematical Biosciences* 59, 1-15.
- Joffe, A., and W. A. O'N. Waugh (1982), 'Exact Distributions of Kin Numbers in a Galton-Watson Process', *J. of Applied Probability* 19, 767-75.
- Keyfitz, N. (1968), *Introduction to the Mathematics of Population*, Addison-Wesley, Reading, Mass.
- (1977), *Applied Mathematical Demography*, Wiley, New York.
- Krishnamoorthy, S. (1979), 'Family Formation and the Life Cycle', *Demography* 16, 121-9.
- Le Bras, H. (1973), 'Parents, Grandparents, Bisaïeux', *Population* 1, 9-38.
- Lotka, A. J. (1931), 'Orphanhood in Relation to Demographic Factors: A Study in Population Analysis', *Metron* 9, 37-109.
- (1938), *Théorie Analytique des Associations Biologiques*, Vol. 2. Hermann, Paris.
- McDaniel, C. K., and E. A. Hammel (1984), 'A Kin-based Measure of r and an Evaluation of its Effectiveness', *Demography* 21, 41-51.
- Pullum, T. W. (1982), 'The External Frequencies of Kin in a Stable Population', *Demography* 19, 549-65.
- Rose, F. (1960), *Classification of Kin, Age Structure and Marriage Amongst the Groote Eylandt Aborigines*. Akademie Verlag, Berlin.
- Ryder, N. B. (1977), 'Models of Family Demography', *Population Bulletin of the United Nations*, 9, United Nations, New York.

- (1982), 'Progressive Fertility Analysis', *Technical Bulletin No. 8*, World Fertility Survey.
- Simon, H. (1957), *Models of Man*, Wiley, New York.
- Tjon Sie Fat, F. E. (1981), 'More Complex Formulae of Generalized Exchange', *Current Anthropology* 22, 377-90.
- Wachter, K. W. (1980), 'The Sisters' Riddle and the Importance of Variance when Guessing Demographic Rates from Kin Counts', *Demography* 17, 103-14.
- Waugh, W. A. O'N. (1981), 'Application of the Galton-Watson Process to the Kin Number Problem', *Advanced Applied Probability* 13, 631-49.
- White, H. (1963), *An Anatomy of Kinship*, Prentice-Hall, Englewood Cliffs.