The formal demography of kinship: Descendants, generations, and the persistence of lineages

EPC 2024 Extended Abstract

Hal Caswell
University of Amsterdam

Draft: revised October 27, 2023

Contents

1 Introduction

The matrix kinship model Caswell (2019) projects the expected age structures of the kinship network of a Focal individual. The initial one-sex, time-invariant, age-classified model has been expanded to include stage structure (Caswell, 2020), time-varying rates (Caswell and Song, 2021), both sexes (Caswell, 2022), and bereavement and causes of death (Caswell, Verdery, and Margolis, 2023). This paper focuses in on one important part of the kinship network: the lineage of descendants, in subsequent generations, of Focal (children, grandchildren, . . .). I will derive a model to project the age or stage structure of successive generations. The model admits the possibility of generation-dependent demographic rates. It makes it possible to compute a variety of measures of multigenerational overlap, and it permits the distinction between one-sex and two-sex definitions of lineages.

Lineages are delicate; the stochastic nature of survival and reproduction imply the possibility of extinction of a lineage. The model also leads to a new analytical calculation, based on a multistate, multitype branching process, for the probability of persistence or extinction of lineages, including the possibility of generation-dependent rates, and accounting for the properties of lineage founders and for one-sex and two-sex definitions of lineages.

The word 'generation' has multiple meanings in current discourse (e.g., Duffy, 2021). My focus here is on biological generations; if Focal is in one generation, her children form the second generation, their children form the third, and so on. Mapping this well-defined concept onto generations defined by social events (boomers, gen X, millenials, and so on) is an open research question.

In Section 2 I present the deterministic kinship model and in Section 3 extend it to project the dynamics of generations. Section 4 develops the stochastic version of this model and uses it to calculate the dynamics of extinction probability.

2 The matrix kinship model

A brief overview of the matrix kinship model follows. The network of Focal is shown in Figure 1. Each kin type is denoted by a unique letter. The network can be extended if desired. In the time-invariant version of the model the dynamics of the age distribution of some type **k** of kin are

$$\begin{aligned} \mathbf{k}(x+1) &= \mathbf{U}\mathbf{k}(x) + \boldsymbol{\beta}(x) \qquad x = 0, \ldots, \omega & (1) \\ \mathbf{k}(0) &= \mathbf{k}_0 & (2) \end{aligned}$$

where $\mathbf{k}(x)$ is the age structure vector for this type of kin at age x of Focal and $\boldsymbol{\beta}(x)$ is the the recruitment vector of new kin. New kin might be created by reproduction of some other type of kin, in which case

$$\boldsymbol{\beta}(x) = \mathbf{F}\mathbf{k}^*(x) \qquad (3)$$

where $\mathbf{k}^*(x)$ is the age structure vector of the other type of kin (e.g., granddaughters of Focal are produced by daughters of Focal). For some kin, the recruitment term is zero (e.g., Focal cannot accumulate any older sisters after she is born).

An initial condition $\mathbf{k}_0$ must be specified, giving the age structure of the kin at the birth of Focal. For some types of kin, the initial condition is zero (e.g., we know that Focal has no daughters at the time of her birth). For other types of kin, the initial condition is calculated as a mixture over the age distribution $\boldsymbol{\pi}$ of mothers at the birth of Focal (e.g., the older sisters of Focal at her birth are the daughters of Focal's mother at her age at the birth of Focal). For calculation of this distribution, see Caswell (2019).

2.1 The model for lineages and generations

The lineage of descendants of Focal comprise a population jointly classified by age and generation. We project the lineage of these descendants with a multistate age×generation-classified matrix

$$\tilde{\mathbf{A}} = \tilde{\mathbf{U}} + \tilde{\mathbf{F}} \qquad (4)$$

with age classes $(1, \ldots, \omega)$ within generations $(1, \ldots, \gamma)$. As an example, consider three generations (Focal, daughters, granddaughters); then

$$\tilde{\mathbf{U}} = \begin{pmatrix} \mathbf{U} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{U} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{U} \end{pmatrix} \qquad \tilde{\mathbf{F}} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{F} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{F} & \mathbf{0} \end{pmatrix} \qquad (5)$$

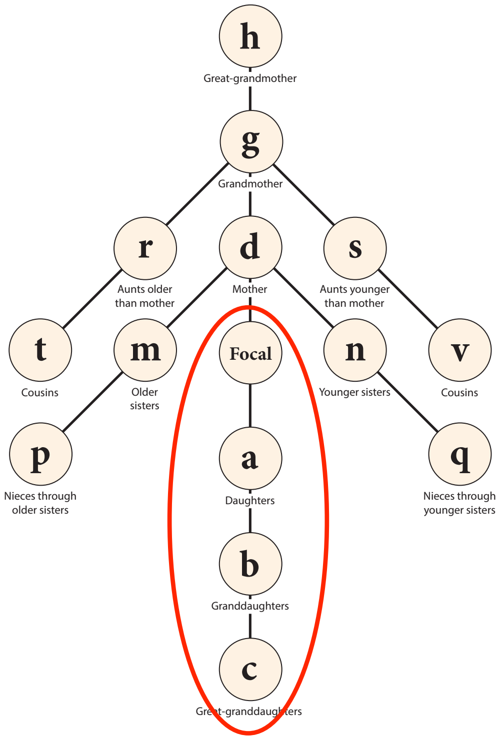


Figure 1: The kinship network of Focal, with the lineage of descendant generations highlighted.

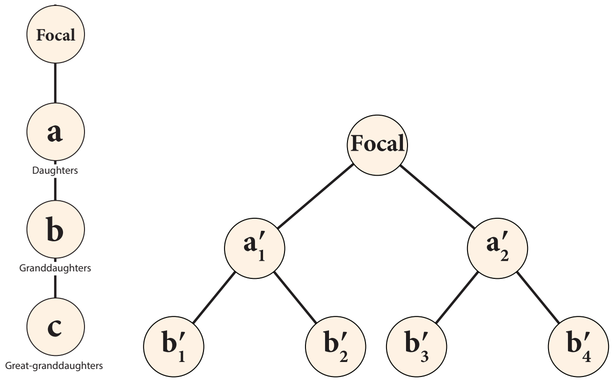


Figure 2: One-sex and two-sex lineages

Let $\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_N$ denote a sequence of N generations. For the three-generation case, the projection is

$$\left(\begin{array}{c}\mathbf{k}_1\\ \hline \mathbf{k}_2\\ \hline \mathbf{k}_3\end{array}\right)(t+1) = \left[\begin{pmatrix}\mathbf{U} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{U} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{U}\end{pmatrix} + \begin{pmatrix}\mathbf{0} & \mathbf{0} & \mathbf{0}\\ \mathbf{F} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{F} & \mathbf{0}\end{pmatrix}\right]\left(\begin{array}{c}\mathbf{k}_1\\ \hline \mathbf{k}_2\\ \hline \mathbf{k}_3\end{array}\right)(t). \tag{6}$$

The survival and transition matrix $\mathbf{U}$ appears on the diagonal of $\tilde{\mathbf{U}}$ because individuals within a generation age, but cannot move to another generation. The fertility matrix $\mathbf{F}$ appears on the subdiagonal of $\tilde{\mathbf{F}}$ because the fertility of each generation provides new offspring to the next generation.[1]

2.2 One-sex and two-sex lineages

The lineage of descendants of Focal may be defined through one sex only (e.g., daughters, granddaughters as daughters of daughters, etc.) or through both sexes (sons and daughters, grandchildren as the children of sons and daughters, etc.). Many times in history an obsession with one-sex (usually male) lineages has characterized royal or noble families. This model accommodates both of these possibilities.

Caswell (2022) has presented a flexible two-sex version of the matrix kinship model. It classifies individuals by both age and sex, and operates with matrices containing female and male survival, and female and male fertility. Various approximations are possible in this context, depending on whether separate female and male rates are available. The extreme case uses the same rates for both sexes; since females and males are indistinguishable in this model, it is called the androgynous approximation. It corresponds to the factors used by Goodman, Keyfitz, and Pullum (1974) to translate numbers of female offspring to combined numbers of males and females under the same assumption.

It can be shown that equation (5 projects both sexes of kin when $\mathbf{F}$ gives children of both sexes per mother, and projects single lines of descent when $\mathbf{F}$ gives daughters (or sons) per mother. The two-sex fertility is used in computing total fertility rates (TFR); the one-sex fertility is used in calculating net reproductive rates (R_0).

Thus if $\mathbf{F}$ is in terms of children per mother, the two-sex lineage projection under the androgynous approximation is

$$\tilde{\mathbf{k}}(t+1) = \left(\tilde{\mathbf{U}} + \tilde{\mathbf{F}}\right)\tilde{\mathbf{k}}(t) \tag{7}$$

and the single-sex lineage projection is

$$\tilde{\mathbf{k}}(t+1) = \left(\tilde{\mathbf{U}} + 0.5\tilde{\mathbf{F}}\right)\tilde{\mathbf{k}}(t) \tag{8}$$

Of course, if separate male and female rates were available, they could be incorporated, doing away with the necessity of the androgynous approximation.

---

[1] Note that this formulation makes it easy to incorporate survival or fertility rates specific to each generation. This is not equivalent to time-specific demographic rates. Generation-specific rates could be useful for studying maternal effects (e.g., Hernández et al., 2020; van Daalen et al., 2022, for rotifers), epigenetic effects, and intergenerational transmission (e.g., Mare, 2011; Lee, 2003). The common expectation that one's children will have better lives than one's self refers not to a change in time (i.e., not that things will be better next year) but that things will be better for members of the next generation, no matter the year in which they appear.

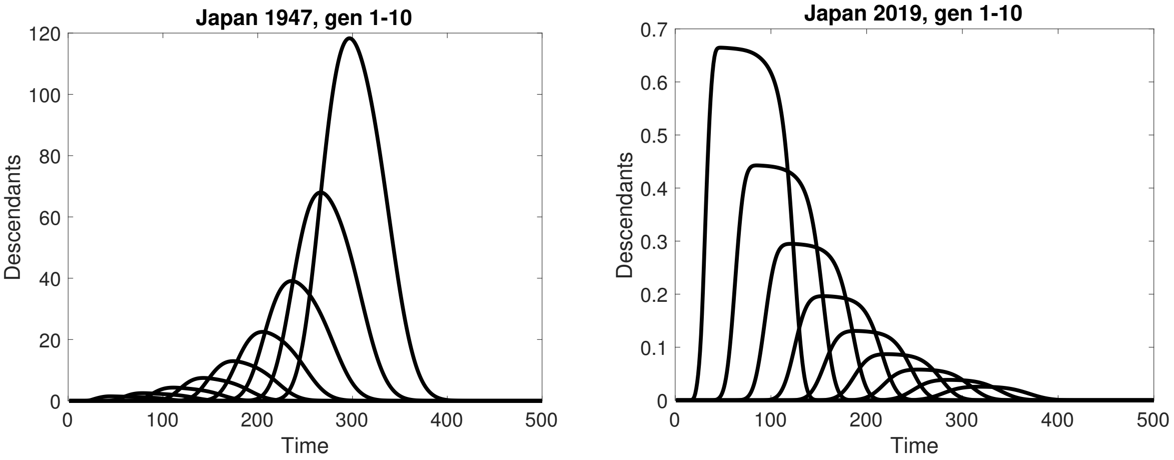


Figure 3: Projections of total descendant numbers by generation for Japan with 1947 and 2019 rates. Each curve represents one generation.

3 Generations: projection and overlap

As an example[2], I analyze the generation projection using rates for Japan in 1947 and in 2019. This time difference saw a dramatic change in the demography, from a low survival, high fertility, increasing population to a high survival, low fertility, decreasing population.

	1947	2019	%change
life exp	54	87	+61%
TFR	4.61	1.34	−70%
R_0	1.7	0.7	−59%

Figure 3 shows the total numbers of descendants in generations 1–10 under these two sets of rates.

The number of descendants and their overlap are clearly visible in the example. Note that these results show the *total number* of descendants, but are obtained by integrating over the age distribution in each generation at each time. Thus it is easy to calculate the overlap of particular parts of the age distribution in a series of generations.

4 Generations: persistence and extinction

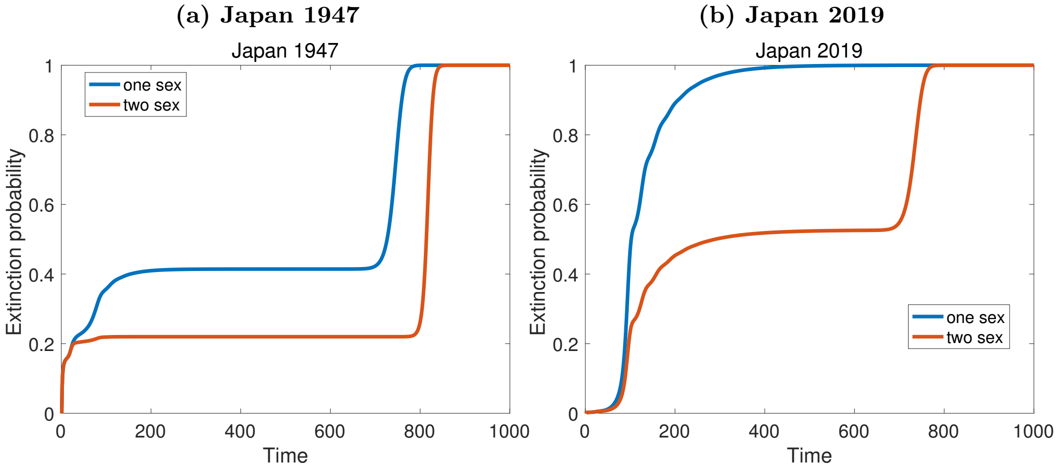
> *Then this is a day of independence*
> *For all the Munchkins and their descendants . . . if any.*
> The Wizard of Oz (1939)

Because the generation model is formulated as an age×generation matrix model, it is, like any matrix model, the expected value operator of a multitype branching process (Harris, 1963; Caswell, 2001). A general stage-classified theory was presented in Caswell (2001, Chap. 15). This theory includes that of Pollard (1966) as a special case. It is tailored to the model structure of a population projection matrix, and my approach here follows it as closely as possible. From this model, the probability of extinction can be calculated.

This is not the place for a detailed derivation (which will appear in a final paper). Suffice it to say that the extinction probability is calculated using the probability generating functions

---

[2] This example will probably be replaced by the time of the conference.


(a) Japan 1947

(b) Japan 2019

Figure 4: Extinction probability as a function of time for one-sex and two-sex lineages, using rates for Japan in 1947 and 2019. Maximum of 20 generations.

(pgfs) for transitions and for reproduction. The vector $\mathbf{X}$ is the age-generation distribution of new copies produced by an individual in stage j. These individuals are the sum of the copies by transition and by reproduction. Given independence, the pgf is the product of the multinomial distribution from $\mathbf{U}$ and the binomial distribution from $\mathbf{F}$:

$$G^{(j)}_{\mathbf{X}}(\mathbf{s}) = G^{(j)}_{\tilde{\mathbf{U}}}(\mathbf{s})\; G^{(j)}_{\tilde{\mathbf{F}}}(\mathbf{s}). \tag{9}$$

We are now in a position to calculate extinction probability as a function of time. Let $\tilde{\mathbf{q}}(t)$ contain the probabilities of extinction, by time t, for lineages started by a single individual in each age-generation combination. Then

$$\tilde{\mathbf{q}}(t+1) = \tilde{\mathbf{G}}_{\mathbf{X}}\left(\tilde{\mathbf{q}}(t)\right) \tag{10}$$

As is well known, extinction probability goes to 1 if $\lambda \leq 1$, and converges to a value less than 1 if $\lambda > 1$, where λ is the growth rate from the matrix (4).

4.1 The dynamics of extinction probability

Figure 4 shows some results that give an idea of the kind of dynamics that I will explore more. It compares extinction probability, starting with a single newborn individual in generation 1, based on the demographic rates of Japan in 1947 and 2019.

This figure deserves careful attention. Under 1947 rates, $\lambda > 1$ for one-sex and two-sex lineages. For both, the extinction probability reaches an asymptote, with the two-sex lineage having a lower probability, as makes sense. However, both probabilities abruptly jump to an extinction probability of 1 after about 700 years.This sudden transition occurs because there is an extinction barrier built into the model; in $\tilde{\mathbf{F}}$. In equation (5), note that after the three generations shown, there is no subsequent generation to continue. Whenever the last members of the third generation die, the lineage is extinct.

The asymptote reflects the well known theorem (Harris, 1963, Theorem 8.1) that, with probability 1, every lineage either goes extinct or escapes towards infinity. Those that do escape become large enough that extinction becomes vanishingly rare. Thus if the model contains enough generations to permit an asymptote, that represents the lineage extinction probability.

The right side of Figure 4 shows the results for Japan under 2019 rates. Under these rates, $\lambda < 1$ for the one-sex lineage, and $\lambda > 1$ for the two-sex lineage. The extinction probability for a one-sex lineage under these conditions converges smoothly to 1. If the lineage is defined by both sexes, the probability of eventual extinction is less than one. Including both sexes thus rescues the lineage.

From results such as this it is possible to examine in detail the impact of the demographic rates on the probability of extinction, or persistence, of a lineage.

5 Beyond this extended abstract

This is probably enough for an extended abstract. Here are some points about the rest of the study.

Asymptotic extinction probability. The asymptote apparent in Figure 4 can be calculated directly by appending a persistent state after the last generation.

Founders of lineages. Extinction probability can be calculated as a function of the age of the founder (trivially, a lineage founded by a single post-reproductive female has no chance of persistence), or of a group of individuals of any specified ages.

Survival analysis of lineages. The quantity $1-\tilde{\mathbf{q}}(t)$ is a survivorship function for lineages. Rescaling and conditioning on extinction provides the distribution of 'age (in years) at death' for extinct lineages. Other survival analyses are possible.

But there's more. Comparing $\tilde{\mathbf{q}}(t)$ calculated with different limits on the number of generations provides a distribution of extinction by generations, complementing the distribution by time. This has never been possible before, as far as I know.

Variation, by generation and by time. The analyses shown here are for fixed demographic rates. The same calculations would apply to rates that vary from generation to generation. The theory for time variation in branching processes is much more difficult. However, the population projection results of Section 3 apply directly to time varying rates.

References

Caswell, H. (2001). *Matrix Population Models: Construction, Analysis, and Interpretation.* Sunderland, MA: Sinauer Associates, 2nd ed.

Caswell, H. (2019). The formal demography of kinship: A matrix formulation. *Demographic Research* 41: 679–712.

Caswell, H. (2020). The formal demography of kinship. II. Multistate models, parity, and sibship. *Demographic Research* 42: 1097–1144.

Caswell, H. (2022). The formal demography of kinship IV: Two-sex models and their approximations. *Demographic Research* 47: 359–396.

Caswell, H. and Song, X. (2021). The formal demography of kinship. III. kinship dynamics with time-varying demographic rates. *Demographic Research* 45: 517–546.

Caswell, H., Verdery, A., and Margolis, R. (2023). The formal demography of kinship V: Kin loss, bereavement, and causes of death. *Demographic Research* (accepted).

Duffy, B. (2021). *Generations: Does When You're Born Shape Who You Are?* London, United Kingdom: Atlantic Books.

Goodman, L.A., Keyfitz, N., and Pullum, T.W. (1974). Family formation and the frequency of various kinship relationships. *Theoretical Population Biology* 5(1): 1–27.

Harris, T.E. (1963). *The theory of branching processes.* Berlin: Springer.

Hernández, C.M., van Daalen, S.F., Caswell, H., Neubert, M.G., and Gribble, K.E. (2020). A demographic and evolutionary analysis of maternal effect senescence. *Proceedings of the National Academy of Sciences* .

Lee, R.D. (2003). Rethinking the evolutionary theory of aging: transfers, not births, shape senescence in social species. *Proceedings of the National Academy of Sciences* 100(16): 9637–9642.

Mare, R.D. (2011). A multigenerational view of inequality. *Demography* 48(1): 1–23.

Pollard, J.H. (1966). On the use of the direct matrix product in analysing certain stochastic population models. *Biometrika* 53: 397–415.

van Daalen, S.F., Hernández, C.M., Caswell, H., Neubert, M.G., and Gribble, K.E. (2022). The contributions of maternal age heterogeneity to variance in lifetime reproductive output. *The American Naturalist* 199(5): 603–616.