OBTAINING MULTISTATE LIFE TABLE DISTRIBUTIONS FOR HIGHLY REFINED SUBPOPULATIONS FROM CROSS-SECTIONAL DATA: A BAYESIAN EXTENSION OF SULLIVAN'S METHOD*

SCOTT M. LYNCH AND J. SCOTT BROWN

Multistate life table methods are often used to estimate the proportion of remaining life that individuals can expect to spend in various states, such as healthy and unhealthy states. Sullivan's method is commonly used when panels containing data on transitions are unavailable and true multistate tables cannot be generated. Sullivan's method requires only cross-sectional mortality data and cross-sectional data indicating prevalence in states of interest. Such data often come from sample surveys, which are widely available. Although the data requirements for Sullivan's method are minimal, the method is limited in its ability to produce estimates for subpopulations because of limited disaggregation of data in cross-sectional mortality files and small cell sizes in aggregated survey data. In this article, we develop, test, and demonstrate a method that adapts Sullivan's approach to allow the inclusion of covariates in producing interval estimates of state expectancies for any desired subpopulation that can be specified in the cross-sectional prevalence data. The method involves a three-step process: (1) using Gibbs sampling to sample parameters from a bivariate regression model; (2) using ecological inference for producing transition probability matrices from the Gibbs samples; (3) using standard multistate calculations to convert the transition probability matrices into multistate life tables.

Multistate life table methods are used in demography to estimate the length of remaining life that individuals can expect to live in different states, such as healthy versus unhealthy states, married versus unmarried states, and so on. One of the most common applications has been the estimation of healthy life expectancy (HLE): the length or proportion of remaining life spent free from disability, chronic disease, or other health problem. A specific focus of multistate methods has been to estimate active life expectancy (ALE), which is the length of life spent free from physical limitations. To date, the method most often used to estimate ALE has been Sullivan's (1971) method, which is not a true multistate method but which provides good estimates of ALE with less stringent data requirements than true multistate approaches (Crimmins and Saito 2001; Crimmins, Saito, and Ingegneri 1997).

True multistate methods require panel data for computing transition probabilities reflecting the movement of individuals into and out of different states across time (see Land and Hough [1989] for an exception). Recent advances in multistate life table methodology include (1) the use of multivariate hazard models to produce smoothed transition probabilities for generating multistate life tables for specific subpopulations, and (2) the development of simulation-based methods for constructing interval estimates of state expectancies when sample data are used. As an example of the former, Land, Guralnik, and Blazer (1994) showed how predicted values from loglinear models with covariates could be used to generate expected age-specific transition probability matrices for input into multistate

*Scott M. Lynch, Department of Sociology and Office of Population Research, Princeton University, Princeton, NJ 08544; e-mail: slynch@princeton.edu. J. Scott Brown, Department of Sociology and Gerontology and Scripps Gerontology Center, Miami University, Oxford, OH 45056. This research was supported by NICHD Grant 1R03HD050374-01, and we thank the reviewers and members of Princeton's "Theorodology" group for helpful comments. Copies of the R programs used for hazard model and life table estimation can be obtained from the first author.

life table calculations, thereby leading to the construction of multistate life tables for any desired subpopulation, including those for which data disaggregation would have yielded sample sizes too small for stable estimation.

Multistate life tables derived under this approach, however, do not address the uncertainty inherent in using sample data. Thus, research over the past decade has concentrated on methods to capture uncertainty. Wolf and Laditka (1996) showed how a Markov process model can be estimated via multinomial logistic regression and used to simulate life histories and thereby to obtain individual-level variability in life table functions (see also Laditka and Wolf 1998). This method was extended by Lievre, Brouard, and Heathcoate (2003) via the derivation of standard errors using the delta method.

Hayward, Rendall, and Crimmins (1999) used a data bootstrap to obtain samples of hazard model parameters, generated transition probability matrices for each bootstrap sample, and computed multistate tables from these matrices, yielding bootstrapped confidence intervals (CI) for multistate expectancies. Lee and Rendall (2001) used a similar approach, but simulated the hazard model parameters directly from the sampling distribution implied by the parameters and standard error estimates from a single hazard model. More recently, Lynch and Brown (2005) used Gibbs sampling to sample hazard model parameters, generated transition probability matrices for each Gibbs sample, and computed multistate tables from them, yielding interval estimates. Most recently, Cai and Lubitz (2007) and Cai, Schenker, and Lubitz (2006) showed how to use bootstrapping procedures to compensate for complex sample design in generating standard errors for multistate life table quantities.[1]

Despite these recent advances, the availability of panel data remains a significant impediment to the use of multistate methods for up-to-date life table estimation and evaluation of trends over time in both developed and developing countries. Although panel data are more prevalent now than historically in developed countries, relatively few panels exist that cover significant portions of the age distribution for many birth cohorts. Consider the case of HLE in the United States, for example. The National Health and Nutrition Examination Survey (and its follow-ups—the National Health Epidemiologic Followup Surveys NHEFS)) is one of the longest running panel data sets investigating health and covering the entire adult age range. However, although the study spans 1971–1992 and covers persons ages 24–77 at baseline, it has limitations. First, there are only four (unevenly spaced) waves of data over the more than 20-year span. Second, the panel is not replenished at each wave, so estimates of HLE cannot be obtained for cohorts born after 1946. Third, the data are dated, given that the last follow-up occurred more than a decade ago. In developing countries, even this type of data is generally unavailable.

The limitations of panel data make Sullivan's method an attractive alternative to true multistate methods because Sullivan's method produces multistate-like estimates by using independent sets of cross-sectional data for mortality and for health (or other) states. Cross-sectional mortality data are produced annually in most developed countries, and annual cross-sectional prevalence estimates for health (or other states) can be obtained from various sources, including repeated cross-sectional surveys in particular. One such source in the United States is the National Health Interview Survey (NHIS), a repeated cross-sectional survey conducted every year since 1969. Thus, in theory, Sullivan's method can allow

1. At least two types of interval estimates can be constructed for state expectancies. One type of interval estimate concerns the extent to which a population-level measure—such as life expectancy—may vary, conditional on the observed data; the other concerns the extent to which individuals' experiences may vary around the population mean. The difference between these two types of variability can be best understood as the difference between a standard error and a standard deviation. In this article, we are concerned with the extent of possible variability in the population-level expectancy measure (as captured by a standard error of a mean). We do not attempt to estimate the extent of individual-level variation in the population (as captured by a standard deviation). This issue merits further discussion in the growing literature on multistate life table interval estimation.

estimation of HLE annually across roughly the past four decades, at a minimum making the evaluation of trends in HLE possible.

Sullivan's method, however, suffers from at least one significant limitation. Aggregate-level measurement of state expectancies, although important, ignores important and extensive heterogeneity in experience across subpopulations. Thus, Sullivan's method is often applied to disaggregated data to produce subpopulation-specific estimates. However, the subpopulations for which estimates can be produced are generally limited by two factors: (1) the level of *disaggregation* possible in mortality data, and (2) the subsample sizes for *aggregated* subpopulations in survey-based prevalence data. Annual U.S. vital statistics mortality data cover the entire population but are generally measured coarsely—usually only by age, sex, and race. On the other hand, survey data can be aggregated to a much more refined level, but survey sample sizes are often too small to produce stable transition or prevalence probabilities for highly refined subpopulations (see Land et al. 1994).

Even if cell sizes are adequate so that survey data can be aggregated to produce refined subpopulation-level prevalence estimates, it is unclear how to combine mortality and survey data at different levels of aggregation and compensate for the uncertainty in multistate estimates that such an approach would produce. Yet, it may be desirable to obtain state expectancy estimates for very specific subpopulations even if the mortality data cannot be disaggregated at that level. Answering key research questions concerning group differences in state expectancies requires such disaggregation. For example, in this article, we consider the extent to which black-white differences in HLE are attributable to socioeconomic status (SES) differences between blacks and whites even though annual mortality data cannot generally be disaggregated by SES.

In this article, we propose a regression-based approach to Sullivan's method that allows the inclusion of covariates measured at different levels in mortality and survey prevalence data and produces interval estimates of state expectancy that compensate for the uncertainty inherent in using such data. The method recasts Sullivan's approach as a true multistate approach and uses a two-stage Gibbs sampling strategy to produce distributions of multistate life table quantities. In the next section, we review Sullivan's method for a three-state model (the most common state space used). Then we develop a regression-based extension and present results of a brief simulation demonstrating the validity of the method. Finally, we provide an empirical example demonstrating the method's utility for addressing social science research questions. Although this method can be applied to *any* state space, we focus on the case of HLE, the original setting for the development of Sullivan's method.

ORIGINAL SULLIVAN'S (1971) METHOD

Sullivan's method for calculating HLE involves only four steps. First, a single-decrement life table is produced from age-specific, cross-sectional mortality rates/probabilities. Second, data from a cross-sectional survey are used to obtain age-specific prevalence proportions of persons in poor health. Third, these proportions are applied directly to the person-years ($L(x)$) column of the life table in order to apportion the years lived in each age interval, $[x,x + n)$ (where n is the interval width), into healthy and unhealthy years of life (say $L_h(x)$ and $L_u(x)$, respectively). Finally, the remaining life table calculations are carried out for these new person-years columns. That is, $T_h(x) = \sum_{a=x}^{\omega} L_h(a)$ is the total number of person-years to be lived healthy from age x to the oldest age (ω), and $e_h(x) = T_h(x) / l(x)$, where $l(x)$ is the total number of individuals alive at age x.

Standard errors of Sullivan estimates under the linear method for computing person-years lived can be obtained by using the binomial variance of the health proportions (see Molla, Wagener, and Madans 2001):

$$\sigma(e_h(x)) \approx \sqrt{\frac{1}{l^2(x)} \sum_{a=x}^{\omega} [L(a)^2 \times (\pi_h(a)(1 - \pi_h(a)))]/N(a)},$$

where $N(x)$ is the number of persons in the health survey sample aged x used to compute the health proportions. In this calculation, the standard errors are only *approximate*, in part because uncertainty in mortality is ignored: the mortality data are measured at the population level.

Table 1 presents an abbreviated empirical example of Sullivan's method applied to 2002 mortality data from the National Center for Health Statistics and 2002 data from the NHIS. The mortality data are disaggregated by age, sex, and race; and our table shows total and healthy life expectancy for black males. Being healthy is defined by a response of "excellent," "very good," or "good" to a self-rated health measure, and being unhealthy is defined by a response of "fair" or "poor." In the table, and throughout this article, we assume that the mortality data used are *probabilities* and not rates. If rates are all that are available, they should be converted to probabilities prior to their use in our method (see Preston, Heuveline, and Guillot 2000).

In constructing this table, we used the linear assumption for person-years lived in each age interval, and we estimated the table for ages 30–84. Although the vital statistics mortality data cover ages up to 100+, the measurement of age in the NHIS peaks at 85+. We therefore use data on the age interval 84–85 years to close our table because it is the last single-year interval for which both mortality probabilities and health proportions can be obtained.

The first seven columns of this table constitute a typical single-decrement life table (see Preston et al. 2000). The last four columns are Sullivan's additions. Column 8 contains age-specific proportions of healthy persons from the NHIS (n in parentheses). Column 9—the product of columns 5 and 8—contains the number of person-years lived healthy in the age interval. Column 10 is the analog to column 6: it is the sum of person-years lived healthy from age x forward. Finally, column 11 is the analog of column 7: it is the expectation of healthy years yet to be lived by a person at age x.

Table 1 is useful in helping describe, under a typical stationarity assumption, healthy life expectancy (HLE) and total life expectancy (TLE) for the black male population. As this table shows, TLE at age 30 for black males was just over 42 years in 2002, 33 of which could be expected to be spent healthy (78.7%).

This table also shows the sample sizes used to calculate the unhealthy prevalence proportions (column 8), highlighting a limitation of using Sullivan's method on disaggregated data, even at a coarse level as done here (by age, sex, and race). Although the sample consisted of more than 15,000 individuals, only 870 were black males. After the data were disaggregated by age, there were generally fewer than 30 observations used to calculate each age-specific health prevalence proportion. Thus, disaggregating the data further, even if doing so were possible in the mortality file, would lead to highly unstable prevalence proportions. A parametric method for smoothing these proportions, as discussed in Land et al. (1994) and Lynch and Brown (2005), is appropriate and is part of the method that we describe in this article.

Sullivan's method, although straightforward to implement, has not been without criticism. There has been long-standing debate regarding what Sullivan's method assumes about mortality and health prevalence, as well as regarding whether Sullivan's method produces reasonable estimates. Barendregt, Bonneux, and Van der Maas (1994), for example, argued that Sullivan's method works poorly when disability and mortality rates are changing rapidly. In contrast, Mathers and Robine (1997) argued that as long as disability and mortality changes are smooth, the method works quite well. Most recently, Imai and Soneji (2007) showed that Sullivan's method produces unbiased and consistent estimates under stationarity assumptions for both health and mortality, assumptions that are often required even if one is using panel data, given the relatively short-term nature of most panels. We would add (concurring with Mathers and Robine 1997) that although true multistate methods may be the best approach, when panel data are not available, Sullivan's method is the only option.

Table 1. Two-State Sullivan Life Table for Black Males From 2002 Vital Statistics Mortality Data and National Health Interview Survey Data

	Basic Life Table Calculations						Sullivan's Additions			
Age (1)	$l(x)$ (2)	$q(x)$ (3)	$d(x)$ (4)	$L(x)$ (5)	$T(x)$ (6)	$e(x)$ (7)	$p_h(x)(N)$ (8)	$L_h(x)$ (9)	$T_h(x)$ (10)	$e_h(x)$(SE) (11)
30	100,000	.002561	256.1	99,872	4,218,743	42.19	.95(20)	94,878	3,321,906	33.22(.43)
35	98,642	.003457	341.0	98,472	3,722,035	37.73	.91(23)	89,909	2,868,892	29.08(.42)
40	96,852	.004250	411.6	96,646	3,233,162	33.38	.91(33)	87,860	2,426,121	25.05(.42)
45	94,443	.006645	627.6	94,129	2,754,569	29.17	.80(20)	75,303	2,004,350	21.22(.41)
50	90,700	.010434	946.4	90,227	2,291,048	25.26	.78(27)	70,176	1,613,139	17.79(.40)
55	85,358	.015549	1,327.2	84,694	1,850,304	21.68	.84(25)	71,143	1,267,596	14.85(.40)
60	78,263	.021157	1,655.8	77,435	1,440,434	18.41	.73(15)	56,785	970,237	12.4(.39)
65	69,124	.029344	2,028.4	68,110	1,071,090	15.50	.67(12)	45,406	696,487	10.08(.39)
70	58,379	.040816	2,382.8	57,188	751,592	12.87	.73(11)	41,591	502,474	8.61(.41)
75	45,554	.059165	2,695.2	44,206	491,118	10.78	.83(6)	36,839	350,984	7.70(.46)
80	31,806	.087335	2,777.8	30,417	297,485	9.35	.33(3)	10,139	218,540	6.87(.55)
84+	21,571	.112654	21,571	191,483	191,483	8.88	.79(14)	150,450	150,450	6.97(.76)

Note: This is a complete table, although all years of age are not shown because of space constraints.

EXTENDING SULLIVAN'S METHOD

The method we develop reformulates Sullivan's method as a multistate method. The key difference between a true multistate approach and Sullivan's approach is that under a true multistate approach, age-specific transition probabilities are calculated directly from the data, or they are modeled, as in Land et al. (1994) and Lynch and Brown (2005). In other words, *incidence* is observed or modeled. In contrast, under Sullivan's method, no transitions are observed; instead, age-specific *prevalence* is observed. Nonetheless, under the typical stationarity assumption invoked to produce period life tables, we can estimate the transition probabilities that produced the prevalence probabilities by using ecological inference to ultimately produce true multistate life tables that compensate for both sampling uncertainty inherent in sample-based prevalence data and uncertainty inherent in the distinction between measuring prevalence versus incidence. Our method involves the following steps:

1. Construct a suitable data set.

2. Specify a bivariate regression model predicting prevalence proportions in healthy, unhealthy, and dead states, with desired covariates obtained from the survey prevalence file. Simulate G samples of the model parameters using Gibbs sampling.

3. Specify a fixed covariate profile and compute predicted age-specific prevalence proportions for each state for the covariate profile using each of the G Gibbs samples of parameters. Convert these prevalence proportions into transition probability matrices using ecological inference, again, for all G sets of prevalence proportions.

4. Convert each of the G sets of age-specific transition probability matrices into age-specific, continuous-time hazard matrices, and generate G multistate life tables by using standard demographic calculations.

Table 2 provides a graphic depiction of this multistep process. In the following sections, we discuss each step. In the penultimate section, we compare results obtained under this method with those obtained using two alternative methods by using panel data from the 1987 and 1992 waves of the NHEFS treated both as panel data and as cross-sectional data. In the final section, we illustrate the method empirically by using mortality probabilities from 2002 vital statistics life tables and health data from the 2002 NHIS to estimate the extent to which black-white differences in HLE are attributable to SES differences between blacks and whites in 2002.

Step 1: Constructing the Data

The data for our approach consist of a set of cross-sectional mortality probabilities and an n-individual cross-sectional survey data set measuring prevalence in two states (healthy/unhealthy). Repeated cross-sectional data may be used, with year (or cohort) included as a covariate. Age must be measured consistently across data sets and should be recoded as $x_i = x_i + (1/2)$—halfway through the observed single-year age interval—to compensate for the fact that individuals measured in a survey are not *exactly* age x at the time of interview. Let X_h be an $n \times k$ design matrix (including a column of ones for the intercept, age, and other covariates) constructed from the survey data, and let y_h be an $n \times 1$ vector of dichotomous health measures indicating whether respondents are healthy or unhealthy. Let X_m be a $T \times j$ matrix of covariates in the mortality file, and let y_m be a $T \times 1$ vector of mortality probabilities, where T is the product of the number of *combinations* of the

Table 2. Graphic Depiction of Process of Implementing New Method

	Step 3						Step 4			
Step 2	(a)		(b)		(c)		(a)		(b)	
β_1	→	$X_{a=x}\beta_1$	→	$\Pr(a = x)$	→	$P(x, x+1)$	→	$\mu(a = x)$		
	→	⋮	→	⋮	→	⋮	→	⋮	→	Life table 1 for profile X
	→	$X_{a=\Omega}\beta_1$	→	$\Pr(a = \Omega)$	→	$P(\Omega, \Omega+)$	→	$\mu(a = \Omega)$		
β_2	→	$X_{a=x}\beta_2$	→	$\Pr(a = x)$	→	$P(x, x+1)$	→	$\mu(a = x)$		
	→	⋮	→	⋮	→	⋮	→	⋮	→	Life table 2 for profile X
	→	$X_{a=\Omega}\beta_2$	→	$\Pr(a = \Omega)$	→	$P(\Omega, \Omega+)$	→	$\mu(a = \Omega)$		
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
β_m	→	$X_{a=x}\beta_m$	→	$\Pr(a = x)$	→	$P(x, x+1)$	→	$\mu(a = x)$		
	→	⋮	→	⋮	→	⋮	→	⋮	→	Life table *m* for profile X
	→	$X_{a=\Omega}\beta_m$	→	$\Pr(a = \Omega)$	→	$P(\Omega, \Omega+)$	→	$\mu(a = \Omega)$		

Notes: Step 1 (not shown): Construction of data and specification of a bivariate probit model predicting health and mortality.

Step 2: Samples of regression coefficients and error correlation obtained via *m* Gibbs sampling iterations (entire set of parameters represented by $\beta, \beta_1 \ldots \beta_m$).

Step 3a: Generation of two-dimensional age-specific predicted scores (in *z* units) from the Gibbs samples ($\hat{Z} = X_{age}\beta$).

Step 3b: Computation of age-specific prevalence proportions ($\Pr(age)$) from $\hat{Z}$ via bivariate normal integration.

Step 3c: Use of ecological inference to convert $\Pr(age)$ to age-specific transition probability matrices ($P(age, age + 1)$).

Step 4a: Conversion of $P(age, age + 1)$ into continuous-time hazard matrices ($\mu(age)$).

Step 4b: Generation of life tables from each set of age-specific hazard matrices.

j covariates' values for which the mortality data can be disaggregated. In our empirical example, the mortality data are disaggregated by age, sex, and race, with age having 55 values (30–84+), sex having two values, and race having two values (white, black). Thus, $T = 55 \times 2 \times 2 = 220$ and $\dim(X_m) = (220 \times 3)$.

Typically, $j < k$: that is, there are more covariates in the health file than in the mortality file. In our example, the health file contains education as an additional covariate, and so $k = 5$ (intercept, age, sex, race, education), while $j = 3$ (age, sex, race). These two data sources can be merged by X_m, a one-to-many merge from the mortality file into the health file. The mortality probabilities will not be unique for every individual, given that the level of covariate specificity in the mortality file will generally be lower than that contained in the health file. The resulting file will be $n \times (k + 2)$ containing X_h and Y, where $Y = [y_h\ y_m]$. Thus, for our example, the resulting data file is $n \times 7$ (a column of ones for the intercept, age, sex, race, education, a health indicator, and an age-sex-race–specific mortality probability).

Two comments are in order regarding the data setup. First, although we incremented age to halfway between the respondent's stated and next year of age—which is an appropriate strategy for the health outcome—mortality probabilities in vital statistics data are typically for persons at *exact* age *x* at the start of an interval, making our mortality probabilities slightly incorrect. However, two offsetting biases result. On the one hand, the mortality probabilities are too high, given that survey respondents at age $x + 0.5$ have (on average) survived half the age interval to which the mortality probabilities assigned to

them apply. On the other hand, the correct mortality probability that applies to individuals at age $x + 0.5$ should be an average of the mortality probabilities at age x and age $x + 1$, making the assigned mortality probability slightly too low. We have found that given how little mortality probabilities vary from one single-year age interval to the next, coupled with our parametric model specification (which itself contains some unavoidable error), our estimates of TLE are similar regardless of whether we adjust mortality. However, adjustment of mortality is relatively straightforward to make if one wishes to do so: average adjacent mortality probabilities. Second, mortality probabilities from vital statistics sources generally cover the entire population. In contrast, survey samples generally include only the noninstitutionalized population, which is a healthier population. Therefore, the mortality probabilities will tend to be slightly too high for survey respondents. Yet, the institutionalized constitute a very small proportion of the total population, and thus do not influence aggregate mortality probabilities greatly. Nonetheless, adjustment may be made to mortality probabilities if the bias is considered severe.

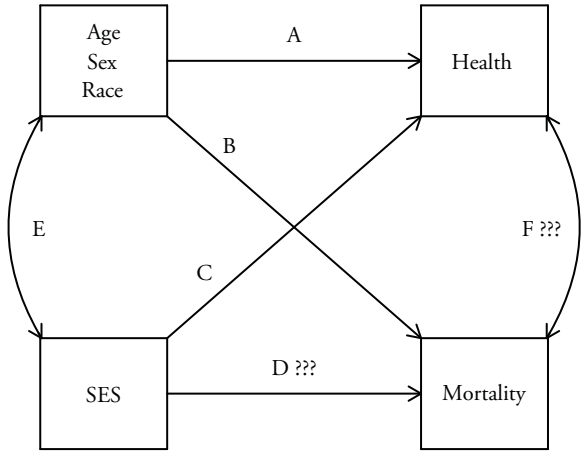
Step 2: Specifying and Estimating a Bivariate Hazard Model

Using the data constructed from Step 1, our goal is to specify a regression model predicting prevalence in healthy, unhealthy, and dead states by using age and whatever additional covariates are desired. The model parameters can then be used in subsequent steps to produce smoothed age- and covariate-specific expected transition probability matrices, similar to Land et al. (1994) and Lynch and Brown (2005). These matrices can then be used to generate multistate life tables.

Figure 1 shows a basic conceptual path model that contains all the relevant parameters for producing state prevalence estimates. For the sake of simplicity, we consolidated covariates based on their availability in the two merged data sets. Age, sex, and race are available in both data sets. Age, sex, race, and mortality are available in the mortality file; and age, sex, race, SES (education), and health are available in the health file.

In Figure 1, paths A and B contain the parameters representing the influence of age, sex, and race on poor health probabilities and mortality probabilities, respectively. Paths C

Figure 1. Conceptual Path Model for Bivariate Hazard Model Parameters to Be Estimated



and D contain the parameters representing the influence of SES on poor health probabilities and mortality probabilities, respectively. Finally, paths E and F represent the correlations between the covariates and outcomes, respectively.

Assuming that all these paths are estimable, the parameters can be used to generate age- and covariate-specific expected probabilities for use in producing transition probability matrices. However, as the figure suggests with question marks, paths D and F are only weakly identified because the mortality probabilities are not differentiated by SES. In a more general case, the mortality probabilities are not differentiated by values of any covariates not measured in the mortality file. These parameters are identified only to the extent that the age-sex-race–specific mortality probabilities are not perfectly predicted by linear combinations of age, sex, and race. Consider, for example, the case in which age is measured with dummy variables (as are sex and race), and all possible interactions between these variables are included. In such a case, age, sex, and race perfectly predict the mortality probabilities, leaving no residual variance to be explained by other covariates. More typically, however, a researcher would estimate the effect of age, sex, and race parametrically, leaving some residual variation that may be explained by the additional variables present in the health prevalence file.

With the data merged, we can develop a model for the observed outcome states, Y, based on the path diagram. The model we use is a bivariate dichotomous probit predicting the probability an individual is in a particular state (healthy, unhealthy, dead), conditional on age, sex, race, and other covariates. The two dimensions of the outcome are healthy versus unhealthy (0 vs. 1) and alive versus dead (0 vs. 1). Given that no transitions are observed, the model is not a hazard model, but we use the expected prevalence probabilities from this model in subsequent steps to produce transition probabilities.

One representation of this model is

$$Y^* = X\beta + e, \tag{1}$$

where Y^* is an $n \times 2$ matrix of latent propensities to be unhealthy and dead; X is an $n \times k$ matrix of covariates, one of which is age; β is a $k \times 2$ matrix of coefficients relating the covariates to the outcomes; and e is an $n \times 2$ matrix of errors that are assumed to be standard bivariate normally distributed with correlation ρ. The error correlation could be constrained to 0; however, because the level of detail on the covariates available in the health and mortality files differ, some efficiency in the estimation of the regression parameters may be gained by estimating ρ (Zellner 1962), and so we incorporate it as a parameter in the model. In this model, the latent propensities, Y^*, are linked to the observed data, Y, by

$$Y = \begin{cases} (0,0) \text{ if } y_1^* \le 0 \text{ and } y_2^* \le 0 \\ (0,1) \text{ if } y_1^* \le 0 \text{ and } y_2^* > 0 \\ (1,0) \text{ if } y_1^* > 0 \text{ and } y_2^* \le 0 \\ (1,1) \text{ if } y_1^* > 0 \text{ and } y_2^* > 0 \end{cases}. \tag{2}$$

In this specification, an individual is observed to be in poor health or dead if his/her latent propensity exceeds the threshold of 0 in the relevant dimension of the equation. We do not observe a dichotomous value for death; instead, we have mortality probabilities. We compensate for this deviation from a typical bivariate probit model in the process of model estimation (details available from the authors upon request).

An appropriate likelihood function for the observed data is based on the multinomial distribution:

$$L(data) \propto \prod_{i=1}^{n} \left(\prod_{r=0}^{1} \prod_{c=0}^{1} p(y_{i1} = r, y_{i2} = c)^{I(y_{i1}=r, y_{i2}=c)} \right), \tag{3}$$

where $p(y_{i1}=r, y_{i2}=c)$ is the probability that individual i is in cell (r,c) of the multinomial, and $I(y_{i1}=r, y_{i2}=c)$ is an indicator for whether the respondent is observed in cell (r,c).

The model parameters are incorporated into the likelihood function by combining Eqs. (1), (2), and (3) so that, for example,

$$\begin{aligned} p(Y_i = (0,0)) &= p(y_{i1}^* \leq 0, y_{i2}^* \leq 0) \\ &= p(X_i\beta(1) + e_{i1} \leq 0, X_i\beta(2) + e_{i2} \leq 0) \\ &= \Phi_2(X_i\beta(1), \infty; X_i\beta(2), \infty; \rho), \end{aligned}$$

where $\Phi_2(a,b;c,d;e)$ is the standard bivariate normal distribution function with correlation e evaluated from a to b in dimension 1 and from c to d in dimension 2. The last equality follows from the bivariate normality assumption for the error. The full likelihood function, then, is

$$L(\beta, \rho \mid data) \propto \prod_{i=1}^{n} \left(\prod_{r=0}^{1} \prod_{c=0}^{1} \Phi_2(\tau_{1,1-r}, \tau_{1,2-r}; \tau_{2,1-c}, \tau_{2,2-c}; \rho)^{I(y_{i1}=r, y_{i2}=c)} \right), \quad (4)$$

where $\tau_{e,f}$ is the fth threshold that divides the eth dimension of the bivariate normal distribution into two bins. There are three thresholds in each dimension, such that $\tau_{1,0}=\tau_{2,0}=-\infty$ and $\tau_{1,2}=\tau_{2,2}=\infty$, and $\tau_{1,1}=X_i\beta(1)$ and $\tau_{2,1}=X_i\beta(2)$ are the individual linear combinations of covariates and parameters for each dimension of the model, where $\beta(.)$ references the column of β.

Although the model parameters may be estimated by using maximum likelihood methods, we adopt a Bayesian approach because it facilitates interval estimation and interpretation as discussed in subsequent sections. (See Lynch [2007] for a review of Bayesian statistics, and see Chib and Greenberg [1998] for more discussion of the multivariate probit model.) To make the model fully Bayesian, the likelihood function is combined with prior distributions for all parameters. A common approach to specifying priors for parameters in multivariate regression models is to assume flat prior distributions for all regression parameters and a noninformative "reference prior" of $p(\sum) \propto |\sum|^{-(d+1)/2}$ for the error covariance matrix (Gelman et al. 1995). This approach to prior specification adds little information to the model, and so the data drive the results. We adopt this strategy of using a vague prior; under this specification, the *posterior* distribution for the parameters is proportional to the product of the likelihood function shown in Eq. (4) and the term $|\sum|^{-3/2}$.

With this prior specification, the influence on mortality of covariates that exist in the health file but not in the mortality file (SES, in this example) will be approximately 0. However, our Bayesian approach to estimation captures uncertainty in this estimate that is realized in the interval estimates for healthy and TLE. Furthermore, our approach produces valid estimates for the *proportion* of remaining life to be spent healthy, despite not having information on mortality disaggregated by all covariates.

In a Bayesian framework, the goal of model estimation is not to maximize the likelihood function and obtain a point estimate for the parameter and its standard error; rather, the goal is to summarize the entire posterior distribution. Many summary measures (such as the mean and variance) are integrals of the posterior distribution. Such integration can sometimes be performed analytically, but more commonly, summaries of parameters are obtained by simulating parameters from their posterior distributions and using basic sample statistics calculations applied to these samples. A benefit of having samples from the posterior distribution is that they can also be used to obtain posterior distributions for *functions* of the original parameters, such as state expectancies.

Gibbs sampling is a common method of producing samples from the posterior distribution for parameters (see Lynch 2007). Under Gibbs sampling, the posterior distribution for the parameters is decomposed into a set of conditional distributions for subsets of the

parameters. Then, samples are alternatingly drawn from each conditional distribution, treating the current values of other parameter subsets as known. In the limit, such sampling produces samples from the joint distribution of parameters. Gibbs sampling for our model involves decomposing the set of parameters into two components: the regression coefficients and the error correlation. The regression coefficients are simulated from their conditional distribution given the current value of the error correlation, and then the error correlation is simulated from its conditional distribution given the new values of the regression coefficients. This process is repeated $B + G$ times, and the first B samples are dropped. These early samples—the "burn-in"—reflect sampling prior to convergence of the algorithm on the true joint posterior distribution.

Step 3: Computing Transition Probability Matrices

After G samples from the posterior distribution of β and ρ are obtained (Step 2 of Table 2), G sets of age-specific expected transition probability matrices can be generated from them and used as input for multistate life table generation. This process involves three steps, repeated for *each* Gibbs sample:

1. Select a covariate profile, X, for which to generate the life table (e.g., black males with 12 years of education) and apply this profile to the parameters to generate expected, two-dimensional Z scores ($\hat{Z}$; propensities) for being healthy, unhealthy, or dead for each age, x, in the age range of the data.

2. Convert $\hat{Z}$ into age-specific expected prevalence probabilities, using bivariate normal integration.

3. Perform ecological inference to obtain transition probabilities from the prevalence probabilities obtained in the previous step. This step can be considered a second-stage model applied to the collection of the age-specific marginal probabilities obtained in the previous step.

For the first step, we choose a covariate profile (X), for which we would like life table estimates, and we apply the parameters obtained from each iteration of the Gibbs sampler, incrementally increasing age to obtain expected age-specific Z scores for being healthy, unhealthy, and dead ($\hat{Z}(x) = X_{age=x}\beta$). (See Step 3(a) in Table 2.) In the second step, we use bivariate normal integration to convert $\hat{Z}$ into age-specific prevalence proportions as follows:

$$\begin{aligned} p_d(x) &= \Phi_2(-\infty, \infty; -\infty, \hat{Z}(x,2); \rho) \\ p_u(x) &= \Phi_2(-\infty, \hat{Z}(x,1); -\infty, \infty; \rho) \\ p_h(x) &= 1 - (p_u(x) + p_d(x)) \end{aligned} \tag{5}$$

In these equations, $p_d(x)$ is the probability of being dead at exact age x, $p_u(x)$ is the probability of being unhealthy (but alive) at exact age x, $p_h(x)$ is the probability of being healthy (but alive) at exact age x, $\hat{Z}(a,b)$ is the value of $\hat{Z}$ for age a in dimension b, and $\Phi_2(.)$ is the standard bivariate normal distribution function with limits of integration $(a,b;c,d)$ and correlation ρ. These probabilities must be computed for all ages in the life table ($x = \alpha \ldots \omega$). (See Step 3(b) in Table 2.)

True multistate life table calculations require transition probabilities showing the movement between states in a state space and not simply prevalence proportions of persons in particular states at a given age/time. Indeed, one source of debate over the accuracy of Sullivan's method concerns the combination of period age-specific prevalence proportions of persons in healthy and unhealthy states with mortality probabilities (see Barendregt et al. 1994; and Mathers and Robine 1997). Here, we do not directly attempt to convert

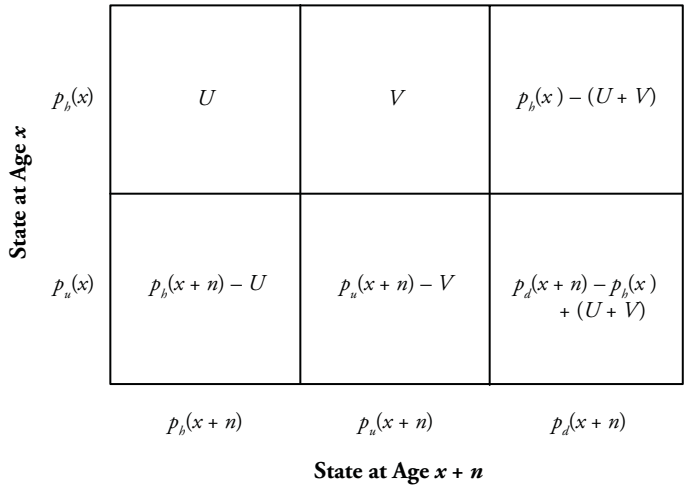
prevalence proportions to incidence probabilities mathematically because the conversion of estimated prevalence proportions to incidence probabilities is not one-to-one. Instead, state prevalence proportions at adjacent ages can imply multiple possible incidence probabilities between states. Under the same stationarity assumption that justifies the construction of period life tables, our estimated prevalence proportions can be converted into incidence probabilities by treating prevalence proportions at exact, adjacent ages as "starting" and "ending" probabilities for discrete age intervals, with one slight adjustment. When a set of these age-specific prevalence probabilities is considered as a set of ending probabilities for an age interval, the probabilities shown in Eq. (5) remain as computed. However, when they are considered "starting" probabilities for the next age interval, $p_h(x)$ and $p_u(x)$ must be conditioned on survival because decedents from a previous age interval are not eligible to "start" the next age interval. For example, for the age interval $[x - n,x)$, $p_u(x)$ remains as computed because it is the probability of ending the interval alive but unhealthy. However, for the age interval $[x, x + n)$, $p_u(x)$ is the proportion of survivors who start the interval unhealthy. Therefore, $p_u(x)$ is conditioned on survival (rescaled) as $p_u(x) = p_u(x) / (1 - p_d(x))$. Note that this approach—conditioning on survival—is equivalent to changing the limits of integration in the mortality equation for $p_u(x)$ in Eq. (5) from $(-\infty,\infty)$ to $(\hat{Z}(x,2),\infty)$.

These age-specific starting and ending *prevalence* probabilities constitute the marginals of 2×3 transition probability (i.e., *incidence*) matrices for each age interval. However, the marginals are not sufficient to complete the transition probability matrices for each age interval. Instead, we are left with an ecological inference problem. Figure 2 provides a graphic depiction of this ecological setting.

In Figure 2, $p_h(x)$ is the probability of being healthy at age x, $p_u(x)$ is the probability of being unhealthy at age x (both conditional on survival to age x), $p_h(x + n)$ is the probability of being healthy at age $x + n$, $p_u(x + n)$ is the probability of being unhealthy at age $x + n$, and $p_d(x + n)$ is the probability of being dead at age $x + n$. These marginals sum to 1 in both dimensions.

As Figure 2 shows, we must determine two transition probabilities in the cells of each age-specific table (labeled U and V) in order to completely determine the age-specific table.

Figure 2. Ecological Inference Setup for Transition Probability Matrices by Age Interval



Obviously, U and V cannot be determined as specific quantities; instead, they may take a range of allowable values. At first glance, given that U and V are probabilities, one may assume that they can take values anywhere on the unit square: that is, they can be drawn from a bivariate uniform distribution.

However, a number of constraints can—and, in some cases, must—be imposed that reduce the allowable values for U and V. Table 3 lists the constraints we impose. The first major set of constraints is that the values in all the cells must be nonnegative, and they cannot exceed the marginals. In the ecological inference literature, these constraints are collectively called the "methods of bounds" (see King 1997). These constraints can be reexpressed in terms of restrictions on the range of U, V, or their sum, as the third column of the table shows.

Ultimately, several such constraints, when expressed in terms of transition probabilities, are duplicated. For example, constraints 4 and 8 are identical, as are 5 and 10, and 6 and 11. If duplicate constraints are ignored, then the allowable ranges for U and V, in any particular age-specific table, based on the method of bounds are

$$\begin{aligned}
&\max\big(0, p_h(x+n) - p_u(x)\big) \le U \le \min\big(p_h(x), p_h(x+n)\big) \\
&\max\big(0, p_u(x+n) - p_u(x)\big) \le V \le \min\big(p_h(x), p_u(x+n)\big).
\end{aligned} \tag{6}$$

Additionally, the sum of U and V is constrained to fall in the range:

$$max(0, p_h(x) - p_d(x+n)) \le (U+V) \le min(1 - p_d(x+n), p_h(x)). \tag{7}$$

In addition to these constraints, two additional constraints can be imposed that are reasonable for most applications of the method. The first constraint is that the probability of remaining in any particular *living* state is greater than the probability of transitioning into another *living* state. This constraint imposes some "state dependence" into the model. The second constraint we impose is that the conditional probability of transitioning to death from the healthy state is less than the probability of transitioning to death from the unhealthy state. Although this constraint appears certainly to be specific to our focus on HLE, this constraint is easily transferred to many other state spaces that may be of interest. For example, it is well known that marriage is beneficial to longevity, and so an application concerned with marital expectancies could impose the constraint that transitioning from an unmarried state to the dead state is more probable than transitioning from the married state to the dead state. These constraints correspond to lines 15–17 in Table 3.

Finally, we impose one additional constraint that, in practice, is largely redundant with the state-dependence constraint. Recall that the underlying continuous-time model of a multistate life table is a piecewise (within age intervals), continuous-time (age) Markov process (see, e.g., Land and Schoen 1982). Singer and Spilerman (1976) showed that a sufficient condition for a discrete-time probability matrix to be embedded in—or have been generated by—such an underlying continuous-time process is that the eigenvalues, λ, of the transition probability matrix, P, be distinct, real, and positive; and that $det(P) > 0$. To ensure this, the diagonal elements of the matrix, conditioned on their respective rows, must sum to be greater than 1. This constraint can be found on line 18 of Table 3.

Figure 3 shows the theoretical space in the unit square to which U and V are limited based on these constraints. Given the marginal probabilities and these constraints, the posterior distribution for U and V, absent any additional prior information, is uniform on the intervals represented in Eqs. (6) and (7). Sampling values of U and V, therefore, simply requires the simulation from a rectangular bivariate uniform distribution implied by the lower and upper limits in each dimension and subjected to the state dependence, mortality, and embeddability constraints.

Table 3. Basic Constraints on *U* and *V* Cells in Ecological Table

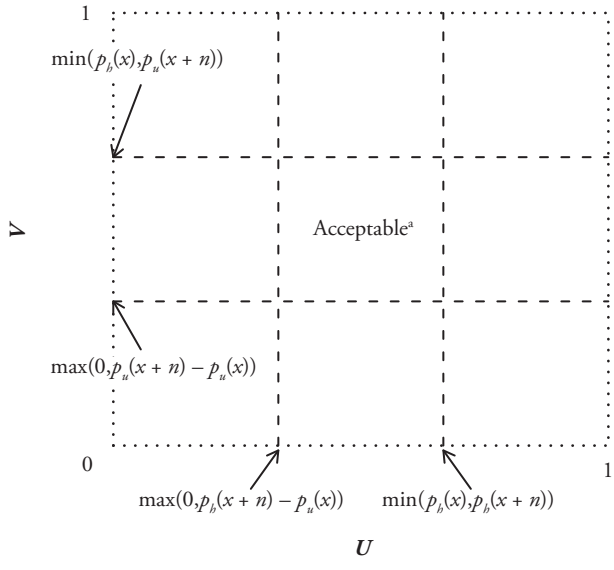
	Constraint	Simplified in Terms of *U* and *V*
All cells must be ≥ 0		
1.	$U \geq 0$	$U \geq 0$
2.	$V \geq 0$	$V \geq 0$
3.	$p_b(x) - (U+V) \geq 0$	$(U+V) \leq p_b(x)$
4.	$p_b(x+n) - U \geq 0$	$U \leq p_b(x+n)$
5.	$p_u(x+n) - V \geq 0$	$V \leq p_u(x+n)$
6.	$p_d(x+n) - p_b(x) + (U+V) \geq 0$	$(U+V) \geq p_b(x) - p_d(x+n)$
All cells must be ≤ marginals		
7.	$U \leq p_b(x)$	$U \leq p_b(x)$
8.	$U \leq p_b(x+n)$	$U \leq p_b(x+n)$
9.	$V \leq p_b(x)$	$V \leq p_b(x)$
10.	$V \leq p_u(x+n)$	$V \leq p_u(x+n)$
11.	$p_b(x) - (U+V) \leq p_d(x+n)$	$(U+V) \geq p_b(x) - p_d(x+n)$
12.	$p_b(x+n) - U \leq p_u(x)$	$U \geq p_b(x+n) - p_u(x)$
13.	$p_u(x+n) - V \leq p_u(x)$	$V \geq p_u(x+n) - p_u(x)$
14.	$p_d(x+n) - p_b(x) + (U+V) \leq p_u(x)$	$(U+V) \leq 1 - p_d(x+n)$
State Dependence or "Stickiness" of Health States		
15.	$U / p_b(x) > V / p_b(x)$	$U > V$
16.	$[p_u(x+n) - V] / p_u(x) > [p_b(x+n) - U] / p_u(x)$	$U - V > p_b(x+n) - P_u(x+n)$
17.	$[p_d(x+n) - p_b(x) + (U+V)] / p_u(x) > [p_b(x) - (U+V)] / p_b(x)$	$(U+V) > p_b(x)p_u(x) - p_b(x)p_d(x+n) + p_b(x)^2$
Embeddability		
18.	$[U / p_b(x)] + [[p_u(x+n) - V] / p_u(x)] > 1$	none

Note: See the text for a description of state dependence and "embeddability."

To obtain values for *U* and *V*, then, for each age-specific transition probability table (and for each *G* Gibbs-sampled parameter value), a bivariate uniform draw is simulated from the appropriate rectangular subregion delineated by the method-of-bounds constraints until it meets these latter constraints (see Step 3(c) in Table 2).

Several comments are in order pertaining to this ecological inference step. First, in some cases, the marginals are such that no values of *U* and *V* will meet the embeddability constraint. Indeed, this happens occasionally when the age intervals in the data exceed one year. However, the constraint we impose is a *necessary* and not *sufficient* condition for embeddability. In general, with the additional constraints imposed—namely, the state-dependence constraint—the transition probability matrices tend to be diagonally dominant and can therefore ultimately be embedded as part of a Markov process (see Singer and Spilerman 1976). From a methodological standpoint, if a particular set of *U* and *V* does not meet this stringent embeddability constraint after, say, 1,000 tries, our software drops this constraint. If a life table then cannot be computed, we argue that the

Figure 3. Ecological Inference Sample Space for U and V



[a]Subject to state dependence, mortality, and embedability constraints (see items 15–18 in Table 3).

marginal probabilities are not a legitimate draw from the posterior distribution for the desired life tables.

Second, ecological inference is commonly performed by developing a model applicable to a *set* of marginal values, and not a single instance of them as we use here. For example, a typical ecological modeling approach might involve using marginal probabilities for race (e.g., white vs. nonwhite) and voting outcomes (e.g., Democrat vs. Republican) for a large number of precincts in a state and attempt to estimate the cell probabilities (e.g., proportion black voting for Republicans) from all precincts simultaneously (see King 1997). In such an approach, this would correspond to using all G Gibbs sampled values simultaneously to obtain a (posterior) distribution for U and V. Instead, our method involves performing the ecological inference *individually*, one Gibbs sample at a time (and, within one age interval at a time). On the one hand, ours is therefore an inefficient approach because we do not use all data simultaneously. On the other hand, our approach *cannot* suffer from aggregation bias because our approach does not involve using pooled aggregate-level data. Instead, our approach can be viewed as using a set of ecological inference models applied to a set of independent aggregate data, eliminating aggregation bias at the expense of some efficiency.

These issues notwithstanding, given a value for U and V, in each age-specific ecological table, the entire transition probability matrix $P(x)$ for a specific age, x (and therefore for an entire age distribution) can be computed simply by conditioning U, V, and the other cell values on the row probabilities. For example, $p_{hh}(x) = U / p_h(x)$, $p_{hu}(x) = V / p_h(x)$, $p_{hd}(x) = (p_h(x) - (U + V)) / p_h(x)$, and so forth, where we have switched notation to double-subscript the transition probabilities between states over the age interval. To complete the construction of $P(x)$ ($x = \alpha\omega$), we must add a row vector, [001], to the bottom of $P(x)$ to make it a 3×3 matrix. This final row represents the probabilities of transitioning from the "dead" starting state.

Step 5: Generating Multistate Life Tables

Generation of life tables, given the discrete-time transition probability matrices $P(x)$ for each age, relies on basic demographic calculations. First, the radix for the life table can be computed by setting the diagonal of $l(\alpha)$ to $[p_h(\alpha)p_u(\alpha)]$, where α corresponds to the proportion of the radix population estimated to be in each state at exact age α (obtained from predicted values with age set to α and conditioned on survival to that age).

Next, the transition probability matrices are converted into continuous-time hazard matrices, $\mu(x)$. (See Step 4(a) in Table 2.) Under the assumption that the force of transition is constant over an age interval—the exponential assumption—$P(x) = \exp\{n\mu(x)\}$; thus, $\mu(x) = -(1/n)\ln(P(x))$. $\ln(P(x))$ can be obtained via Sylvester's formula (see Singer and Spilerman 1976). Under Sylvester's formula:

$$\ln P = \sum_{i=1}^{3} log(\lambda_i) \prod_{i \neq j} \frac{(P - \lambda_j I)}{(\lambda_i - \lambda_j)},$$

where λ_i is the ith eigenvalue of P.

Given $\mu(x)$, the remaining life table computations include the $l(x)$ matrix and the $L(x)$ matrix for each age. (See Step 4(b) in Table 2.) Note that $l(x)$ and $L(x)$ are diagonal matrices, but $l(x+n)$ and $L(x+n)$ are not; as we iterate the life table calculations across age, $l(x+n)$ and $L(x+n)$ must be converted to diagonal matrices (see Palloni 2000, and Schoen 1988). Under the exponential assumption, $l(x+n) = l(x)\exp\{-n\mu(x)\}$, where $\exp\{-n\mu(x)\} = I + \sum_{i=1}^{\infty}((-n)^i \mu(x)^i(x))/i!$ is the series expansion representing the exponential function. In practice, the summation generally requires fewer than 10 iterations to converge.

$L(x) = \int_x^{x+n} l(x)dx$ can then be computed as $L(x+n) = nl(x)\left[I + \sum_{i=1}^{\infty}((-n)^i \mu^i(x))/(i+1)!\right]$. Finally, we can compute state expectancies as $e(x) = L(x)l(x)^{-1}$; and for the oldest age group, to close out the table, we can compute $e(\omega) = l(\omega)\mu(\omega)^{-1}$. Given that we have specified a parametric pattern for age dependence of health and mortality (linear in the probit), we can carry out our life table calculations to any age.

The net result of this multistep process, after repeating it for each of the G samples of model parameters, is that we obtain *distributions* of multistate life tables for each covariate profile we select. We can summarize these distributions of life tables by using basic summary statistics. For example, we can compute the mean, median, mode, variance, range, and so on for multistate life table quantities. We can also compute empirical interval estimates by sorting the life table quantities and selecting the values that represent the quantile cutpoints of interest. We can also perform statistical tests to compare populations with different covariate profiles. We highlight this process in the next sections.

In sum, this process involves multiple steps, each of which produces high-dimensional matrices. For example, suppose that we are interested in constructing life tables for ages 65 and above (say 85+, yielding a total of 21 age groups), and we have age, sex, and race as covariates. After the data have been constructed appropriately, we would use Gibbs sampling to produce, say, $G = 1{,}000$ samples from the posterior distribution of the model parameters (Step 2 from Table 2). This would result in 1,000 sets of nine parameters (a regression coefficient for the intercept and each covariate for both health and mortality equations, plus an error correlation).

Following Step 3(a) from Table 2, we would select a covariate profile for which to produce the life tables (e.g., black males), and apply it to each of the 1,000 parameter samples, incrementing age—one of the covariates—from 65 to 85, to produce expected two-dimensional age-specific z scores. For *each* parameter sample, this step results in a 21×2 matrix of z scores (one row for each age). Following Step 3(b) from Table 2, we would then convert this 21×2 matrix of z scores into expected age-specific prevalence proportions for being healthy, unhealthy, and dead, using integration of the bivariate

normal distribution. The result would be a 21×3 matrix of expected prevalence proportions. This matrix would then be duplicated, and one set would condition out mortality, so that we would have a 21×2 matrix representing prevalence in healthy and unhealthy states at the *start* of each age interval. The other matrix would remain as is to represent prevalence in healthy, unhealthy, and dead states at the *end* of each age interval.

Following Step 3(c) from Table 2, ecological inference would be performed within each age interval to obtain a set of 21 age-specific 3×3 expected discrete-time transition probability matrices. In Step 4(a) from Table 2, we would convert each of these discrete-time matrices into hazard matrices, yielding a set of 21 age-specific 3×3 hazard matrices. Finally, in Step 4(b), we would apply standard multistate calculations to each age-specific set of hazard matrices to obtain a multistate life table. Repeating this process for each of the 1,000 parameter samples results in a distribution of 1,000 multistate life tables.

Although this process involves numerous steps and many tedious calculations, annotated software and a user's manual are available to make implementing the method straightforward.

COMPARISON OF APPROACHES TO MULTISTATE LIFE TABLE ESTIMATION

We demonstrate this new method, and compare results obtained with it to other approaches, using data from the NHEFS. The NHEFS is a panel study in which a subset of respondents from the original 1971 NHANES were followed up on at least three occasions: 1982, 1987, and 1992. Here, we use data from the 1987 and the 1992 waves. Our state space has three states: healthy, unhealthy, and dead, with "healthy" defined as a response of "excellent" or "good" to a self-rated health item, and age is measured in five-year intervals from 45 through 85+.

Given that the data are from a panel, age-specific transition probabilities between states between the two waves were observed, and so a true multistate life table can be produced directly from the data, following standard demographic calculations shown in the previous section. However, for purposes of comparison, approximate cross-sectional age-specific mortality and health prevalence proportions can also be constructed from the data. For mortality, we use the incidence of mortality between age intervals, and we replace individuals' known indicators for mortality with their age-specific sample-level mortality probability. For health, we use 1992 health values for individuals surviving to 1992; we use 1987 health values for those dying during the interval. Table 4 shows the age-specific health and mortality transition and prevalence probabilities. The prevalence values approximate, but are not exactly, those that would be obtained in a cross-section, but there is no absolutely correct approach to producing cross-sectional prevalence proportions from transition data that do not rely on assumptions about transitions. Furthermore, given that the NHEFS survey data are at the individual level, it is not clear how to assign health scores to individuals based on aggregate prevalence proportions, even if we were able to construct them perfectly. So, we note at the outset that our comparison between panel and cross-sectional data is not perfect.

We present results of three sets of analyses. In the first set of analyses, we include no covariates other than age, and we compare results from (1) true multistate calculations applied to the transition probabilities shown in Table 4, (2) the panel method described in Lynch and Brown (2005) applied to the individual-level NHEFS data shown in the table, and (3) the new method described in this article using the individual-level NHEFS data with the age-specific mortality probabilities shown in the table replacing the observed individual-level measure for death. This approach is akin to using a cross-sectional health survey with mortality probabilities from an external source merged into the individual-level file.

Given that we know the true transition probabilities, the multistate approach is the "gold standard": that is, it constitutes the "best" estimates for HLE and TLE that can be obtained from real data. We call estimates derived from this approach the "true" values although they are not true in an absolute sense. The panel method described in Lynch and

Table 4. Transition Probabilities and Prevalence Proportions From the 1987–1992 NHEFS Data

Age	Start ↓	Transition Probabilities: Healthy	End Unhealthy	Dead	Prevalence Proportions (*n*)
45–49	H	.892	.099	.009	.851 (423)
	U	.412	.510	.078	.149 (74)
	D	0	0	1	.016
50–54	H	.884	.100	.016	.811 (361)
	U	.303	.621	.076	.189 (84)
	D	0	0	1	.025
55–59	H	.898	.076	.026	.790 (347)
	U	.305	.568	.126	.210 (92)
	D	0	0	1	.048
60–64	H	.802	.134	.064	.737 (411)
	U	.307	.550	.143	.273 (154)
	D	0	0	1	.083
65–69	H	.761	.162	.076	.707 (384)
	U	.266	.500	.234	.293 (159)
	D	0	0	1	.112
70–74	H	.731	.141	.127	.665 (268)
	U	.208	.450	.342	.335 (135)
	D	0	0	1	.191
75–79	H	.589	.172	.239	.587 (165)
	U	.158	.406	.436	.413 (116)
	D	0	0	1	.310
80–84	H	.403	.151	.446	.548 (125)
	U	.079	.281	.640	.452 (103)
	D	0	0	1	.522
85+	H	.130	.043	.826	.468 (44)
	U	0	.146	.854	.532 (50)
	D	0	0	1	.840

Notes: Sample sizes are the number of healthy and unhealthy individuals at the start of each age interval. Death counts can be obtained by subtracting the total in one age group from the total in the previous one.

Brown (2005) involves true multistate calculations applied to estimated expected transition probabilities, where the estimated transition probabilities are obtained via bivariate probit regression. That approach is similar to that of the new method in its use of a parametric regression model to obtain smoothed estimates of transition probabilities. We refer to this method as the *LB method*.

In the second set of analyses, we examine the case in which a covariate—education—is measured in the health file but not in the mortality file. To accomplish this, we compute

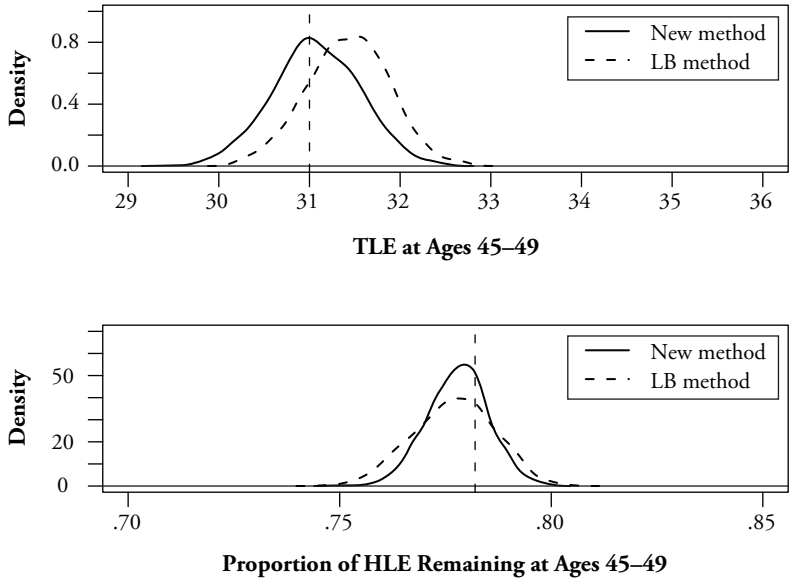
mortality probabilities by age, and we replace each individual's observed mortality measure with his/her age-specific mortality probability. Thus, mortality varies only by age, while health varies across both age and education. We restrict our covariates to age and education, where education is measured dichotomously (12+ years vs. <12 years) because of the limited sample size. That is, we wish to compare the results obtained via the true multistate approach with the new method, but the panel data cannot be disaggregated much beyond age and education ($9 \times 2 = 18$ cells) because of cell size limitations.

In the third set of analyses, we include sex and race in addition to age and education. Given sample size limitations, we are unable to compute true multistate life tables, and so we compare the LB method and the new method.

Figure 4 shows the results of the first set of analyses. The top panel shows TLE at ages 45–49 as computed by using typical multistate calculations (a point estimate), with the LB method and the method discussed here. The top panel shows that TLE at ages 45–49 is 31 years, and both parametric methods appear to estimate this quantity very well. There is a nonsignificant difference between all three approaches: that is, the interval estimates for the two parametric methods overlap substantially, and both capture the point estimate produced by traditional multistate calculations.

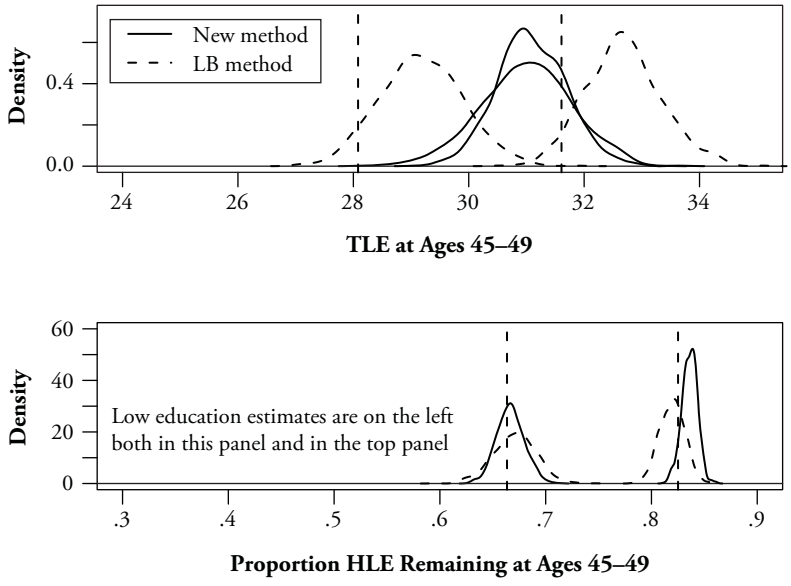
The bottom panel of Figure 4 shows the estimated proportion of remaining life to be spent healthy for this age group. The value estimated using standard multistate life table calculations was .782. The 95% probability interval estimates for the two parametric approaches were [.757,.794] for the LB method and [.763,.792] for the new method described in this article. As with TLE, the three sets of values are statistically indistinguishable.

Figure 4. Total Life Expectancy (top panel) and Proportion of Remaining Life to Be Lived Healthy (bottom panel) at Ages 45–49



Notes: Vertical lines in the figure are the true values as computed using (nonparametric) multistate life table calculations applied to observed transition probabilities. Histograms are the samples obtained using (1) the parametric panel regression method described in Lynch and Brown (2005) and (2) the new method described here.

Figure 5. Total Life Expectancy (top panel) and Proportion of Remaining Life to Be Lived Healthy (bottom panel) at Ages 45–49 for Low and High Education Groups



Notes: Vertical lines in the figure are the true values as computed using (nonparametric) multistate life table calculations applied to observed transition probabilities. Histograms are the samples obtained using (1) the parametric panel regression method described in Lynch and Brown (2005) and (2) the new method described here.

Figure 5 shows the results of the second set of analyses. The top panel of the figure shows TLE at ages 45–49. Under the true multistate approach and the LB approach, there are two sets of estimates (the vertical lines): one for persons with low education, and one for persons with high education. As computed under the true multistate methods, TLE was 28.08 years for persons with low education and 31.61 years for persons with high education, a difference of approximately 3.5 years. The LB estimates are approximately 1 year higher each, at 29.18 and 32.68 years, respectively, also reflecting a difference of about 3.5 years in total life. In contrast, as the figure shows, the life expectancy estimates for the new approach are roughly equal for both high- and low-education groups. The reason is that mortality is not differentiated by education level in the contrived "cross-sectional" data. As a consequence, the estimate for TLE for the new method is a compromise (weighted average) between the two education groups.

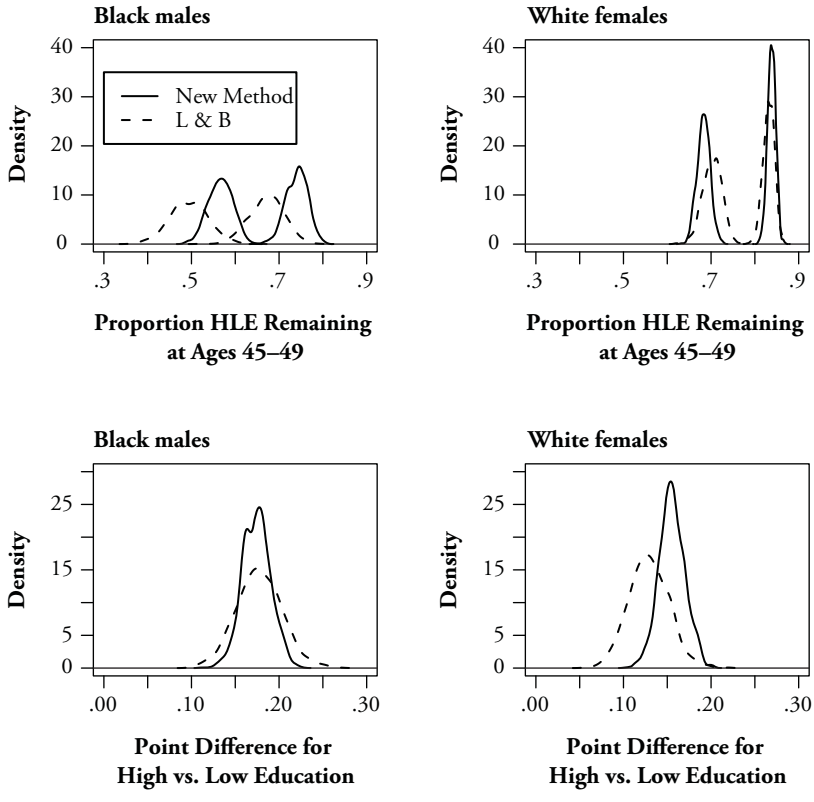
The bottom panel of Figure 5 shows the estimated proportion of remaining life to be spent healthy for high- and low-education groups at ages 45–49. According to the figure, the multistate results show that persons with low education can expect to spend about 66% of their remaining life healthy, and persons with high education can expect to spend about 82% of their remaining life healthy. The results show remarkable consistency between the results of the three approaches. Once again, the interval estimates are statistically indistinguishable from each other and the "true" values produced using traditional multistate calculations.

It may seem surprising that the proportion of remaining life to be spent healthy is accurately estimated using the new approach, given that education-specific mortality

probabilities were not "available" to the model, and thus total life remaining was estimated to be comparable across high- and low-education groups. However, the method works because the health equation predicts the proportion healthy and unhealthy by education level *among those who have survived to the age at which they were observed in the cross section*. That is, the model accurately predicts the probability a survivor is healthy (vs. unhealthy). The variance in this probability is increased, however, when it is used to produce transition probabilities because the mortality data were not measured at as refined of a level as the health data. Therefore, the proportion of remaining life to be spent healthy is accurate, but it is somewhat imprecise because of the poor measurement of mortality.

Figure 6 shows the results of the final set of analyses. The upper two panels show the distributions for the proportion of remaining life to be spent healthy for black males (upper left) and for white females (upper right) by level of education. The lower two panels show the distribution of the percentage point difference in the proportion of healthy life that is produced by education. The upper-left panel shows that the new method tends to produce slightly higher (although not statistically distinguishable) estimates than the LB method for the proportion of remaining life black males can expect to spend healthy. The lower-left

Figure 6. Proportion of Remaining Life to Be Lived Healthy (top panel) and Education-Based Difference in Proportions (bottom panel) for Black Males and White Females



Note: In upper plots, histograms on the left are the distributions for persons with low education; histograms on the right are for persons with high education.

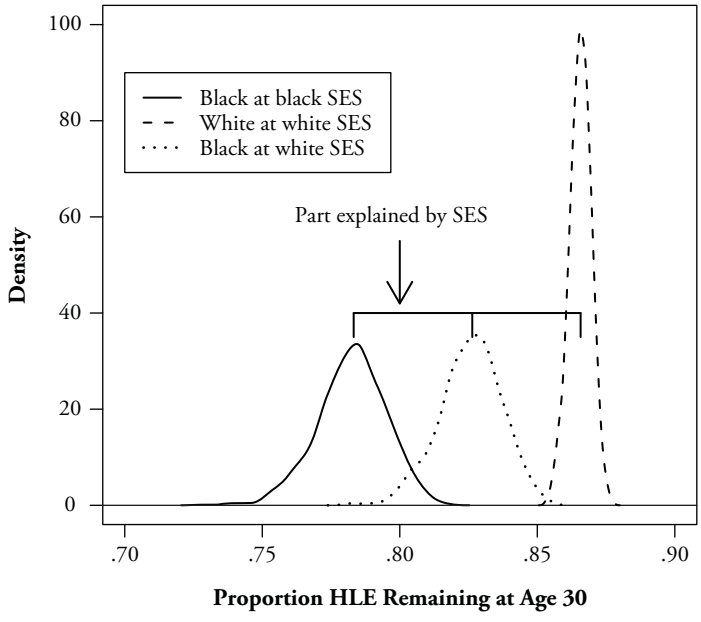
panel, however, shows that the influence of education, in terms of its percentage point effect on healthy life, is estimated virtually identically for both the new and LB methods. The results for white females yield a similar conclusion.

AN EMPIRICAL EXAMPLE

In this section, we provide a more realistic example involving the new method. The question we seek to address is, What proportion of the black-white difference in HLE is explained by SES differences between races? We use data from the 2002 NHIS and 2002 vital statistics mortality data. The mortality data are disaggregated by age, sex (male = 1), and race (black vs. white; others excluded), and merged into the NHIS data at that level. From the NHIS, we use age, sex, race, an interaction between age and race, region of the country (South = 1 vs. non-South), education (years of schooling), and income (log dollars). After obtaining samples of model parameters via Gibbs sampling as described earlier, we generate three sets of life tables. For all three sets, we fix sex and region at sample means (.4 and .3, respectively). With those variables held constant, we generate life tables (1) for blacks with education and income set to black means for these variables, (2) for whites with education and income set to white means, and (3) for blacks with education and income set to white means.

Figure 7 shows the three distributions of proportion of healthy life remaining. As the figure shows, at age 30, whites can expect to live about 86% of their remaining lives healthy (95% interval estimate of [.86,.87]), while blacks can expect to live about 78% of their remaining lives healthy (interval estimate of [.76,.81]). Thus, there is about a 10% disparity (8 percentage points) between races. If blacks had comparable levels of

Figure 7. Histograms of Proportion of Remaining Life to Be Lived Healthy at Age 30 for Blacks and Whites



education and income to whites, the black estimate would increase to approximately 83% (interval = [.80,.85]).

We can determine the extent to which the black-white difference in HLE is explained by SES differences between races by computing

$$\%\ \text{attributable to SES} = 1 - \frac{HLE(white \mid white\ SES) - HLE(black \mid white\ SES)}{HLE(white \mid white\ SES) - HLE(black \mid black\ SES)}.$$

Because this calculation can be performed for all 1,000 Gibbs samples for HLE, we can obtain a posterior mean and an empirical interval for this quantity. This calculation shows that 52.9% of the black-white difference is explained by SES differences, with an empirical interval for this difference of [41.2% , 68.4%].

CONCLUSIONS

Sullivan's method has been used extensively in place of multistate life tables when panel data are unavailable. However, Sullivan's method has been limited in its ability to incorporate covariates into the process of estimation because mortality data are often disaggregated only coarsely, and survey samples are often small enough that aggregation to highly refined subpopulation levels yields unstable prevalence estimates. In this article, we have proposed a regression-based extension to Sullivan's method that enables the construction of multistate life tables for highly refined subpopulations and produces estimates even when the mortality data cannot be disaggregated at the desired level. The method not only produces estimates of multistate life table quantities, but it also provides for interval estimation, thereby compensating for uncertainty inherent in using sample data and having data disaggregated/aggregated at different levels in mortality and survey prevalence files. The results of our comparisons to other methods indicate that the method works quite well, and our empirical example shows how the method can be used to address research questions that may be of interest to a broad social science audience.

Although our approach is an improvement over the use of traditional Sullivan's method, several issues should be discussed and addressed in future research. First, our approach inherently assumes that the sample data come from a simple random sample; and most, if not all, major data sets from which prevalence proportions would be obtained involve complex sampling. Therefore, one could argue that our interval estimates are incorrect. Our approach here could be adapted relatively easily to compensate for some types of complex sample designs by using a Bayesian bootstrap involving post-stratification weights provided in the data combined with our Gibbs sampler (see Rubin 1981). In fact, we have incorporated a similar weighted bootstrap into our program. We have also undertaken extensive analyses investigating the importance of sample design and found that compensating for design has little effect on substantive conclusions (detailed report available from the authors upon request).

Second, our method produces TLE estimates that are identical across values of covariates across which mortality probabilities cannot be disaggregated. The result is that years of healthy life (or years in some other state) will also be incorrect because they will reflect the number of years remaining in given states as a proportion of total life. Thus, one should use proportions as the metric for the results, not years. The method produces valid estimates of the proportion of remaining life to be lived in different states. This is not a limitation of the method; rather, it is a limitation of the data used in estimation. Although estimates of years of healthy life and total life can be obtained and used if the mortality data can be disaggregated across all covariates, if mortality cannot be disagregated, there simply is not enough information available to obtain unique estimates of total life for all subpopulations. If, however, one has strong prior information for the influence of a particular covariate on mortality, this information could be incorporated in the prior

distribution for the regression coefficients and would allow one to obtain unique estimates of total years of life for each subpopulation.

Third, although we limited our presentation here to healthy life expectancy and to only two non-absorbing states and one absorbing state, the model can be used with any state space, and the state space need not be limited to two non-absorbing (or one absorbing) states. Expanding the state space involves extending the bivariate dichotomous probit model described here to a full multivariate probit model or extending one dimension to more than two categories if the non-absorbing states can be ordered. In addition to changes in the regression modeling step, using alternative state spaces will require reconsidering some of the constraints imposed in the ecological inference step, and this process may be quite difficult.

Finally, it is important to note again that Sullivan's method technically requires that the populations to which the mortality probabilities and the state prevalence proportions apply are the same. In our example, they were not. The mortality probabilities were for the entire U.S. population, while the NHIS is a survey of noninstitutionalized individuals. Care should be taken when applying Sullivan's method, or this extension of it, to ensure that populations are equivalent or at least that the biases in estimates resulting from using data from noncomparable populations are known.

REFERENCES

Barendregt, J.J., L. Bonneux, and P.J. Van der Maas. 1994. "Health Expectancy: An Indicator for Change?" *Journal of Epidemiology and Community Health* 48:482–87.

Cai, L. and J. Lubitz. 2007. "Was There Compression of Disability for Older Americans From 1992 to 2003?" *Demography* 44:479–95.

Cai, L., N. Schenker, and J. Lubitz. 2006. "Analysis of Functional Status Transitions by Using a Semi-Markov Process Model in the Presence of Left-Censored Spells." *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 55:477–91.

Chib, S. and E. Greenberg. 1998. "Analysis of Multivariate Probit Models." *Biometrika* 85:347–61.

Crimmins, E.M. and Y. Saito. 2001. "Trends in Healthy Life Expectancy in the United States, 1970–1990: Gender, Racial, and Educational Differences." *Social Science and Medicine* 52:1629–41.

Crimmins, E.M., Y. Saito, and D. Ingegneri. 1997. "Trends in Disability-Free Life Expectancy in the United States, 1970–1990." *Population and Development Review* 23:555–72.

Gelman, A., J.B. Carlin, H.S. Stern, and D.B. Rubin. 1995. *Bayesian Data Analysis.* Boca Raton, FL: Chapman and Hall/CRC Press.

Hayward, M.D., M. Rendall, and E. Crimmins. 1999. "Evaluating Group Differences in Healthy Life Expectancy: The Estimation of Confidence Intervals for Multistate Life Table Expectancies." Presented at the annual meeting of the Gerontological Society of America, San Francisco.

Imai, K. and S. Soneji. 2007. "On the Estimation of Disability-Free Life Expectancy: Sullivan's Method and Its Extension." *Journal of the American Statistical Association* 102:1199–211.

King, G. 1997. *A Solution to the Ecological Inference Problem: Reconstructing Individual Behavior From Aggregate Data.* Princeton, NJ: Princeton University Press.

Laditka, S.B. and D.A. Wolf. 1998. "New Methods for Analyzing Active Life Expectancy." *Journal of Aging and Health* 10:214–41.

Land, K.C., J.M. Guralnik, and D.G. Blazer. 1994. "Estimating Increment-Decrement Life Tables With Multiple Covariates from Panel Data: The Case of Active Life Expectancy." *Demography* 31:297–319.

Land, K.C. and G.C. Hough, Jr. 1989. "New Methods for Tables of School Life, With Applications to US Data From Recent School Years." *Journal of the American Statistical Association* 84:63–75.

Land, K.C. and R. Schoen. 1982. "Statistical Methods for Markov-Generated Increment-Decrement Life Tables With Polynomial Gross Flow Functions." Pp. 265–346 *Multidimensional Mathematical Demography,* edited by K.C. Land and A. Rogers. New York: Academic Press.

Lee, M.A. and M.S. Rendall. 2001. "Self-employment Disadvantage in the Working Lives of Blacks and Females." *Population Research and Policy Review* 20:291–320.

Lievre, A., N. Brouard, and C. Heathcoate. 2003. "Estimating Health Expectancies From Cross-Longitudinal Surveys." *Mathematical Population Studies* 10:211–48.

Lynch, S.M. 2007. *Introduction to Applied Bayesian Statistics and Estimation for Social Scientists.* New York: Springer.

Lynch, S.M. and J.S. Brown. 2005. "A New Approach to Estimating Life Tables With Covariates and Constructing Interval Estimates of Life Table Quantities." *Sociological Methodology* 35:177–225.

Mathers, C.D. and J-M. Robine. 1997. "How Good is Sullivan's Method for Monitoring Changes in Population Health Expectancies?" *Journal of Epidemiology and Community Health* 51:80–86.

Molla, M.T., D.K. Wagener, and J.H. Madans. 2001. "Summary Measures of Population Health: Methods for Calculating Healthy Life Expectancy." *Healthy People Statistical Notes* No. 21. National Center for Health Statistics, Hyattsville, MD.

Palloni, A. 2000. "Increment-Decrement Life Tables." Pp. 256–72 in *Demography: Measuring and Modeling Social Processes,* edited by S.H. Preston, P. Heuveline, and M. Guillot. Oxford, England: Blackwell.

Preston, S.H., P. Heuveline, and M. Guillot. 2000. *Demography: Measuring and Modeling Social Processes.* Oxford, UK: Blackwell.

Rubin, D.B. 1981. "The Bayesian Bootstrap." *Annals of Statistics* 9:130–34.

Schoen, R. 1988. *Modeling Multigroup Populations.* New York: Plenum.

Singer, B. and S. Spilerman. 1976. "The Representation of Social Processes by Markov Models." *American Journal of Sociology* 82:1–54.

Sullivan, D.F. 1971. "A Single Index of Mortality and Morbidity." *HMSHA Health Reports* 86: 347–54.

Wolf, D.A. and S.B. Laditka. 1996. "Stochastic Modeling of Active Life and Its Expectancy." Paper presented at the annual meeting of REVES, Rome, December 11–13.

Zellner, A. 1962. "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias." *Journal of the American Statistical Association* 57:348–68.