Annuity Valuation with Dependent Mortality

Author(s): Edward W. Frees, Jacques Carriere and Emiliano Valdez
Source: *The Journal of Risk and Insurance*, Vol. 63, No. 2 (Jun., 1996), pp. 229-261
Published by: American Risk and Insurance Association
Stable URL: http://www.jstor.org/stable/253744
Accessed: 13-11-2017 06:04 UTC

© The Journal of Risk and Insurance, 1996, Vol. 63, No. 2, 229–261.

Annuity Valuation with Dependent Mortality

Edward W. Frees
Jacques Carriere
Emiliano Valdez

ABSTRACT

Annuities are contractual guarantees that promise to provide periodic income over the lifetime(s) of individuals. Standard insurance industry practice assumes independence of lives when valuing annuities where the promise is based on more than one life. This article investigates the use of dependent mortality models to value this type of annuity. We discuss a broad class of parametric models using a bivariate survivorship function called a copula. Using data from a large insurance company, we calculate maximum likelihood estimates to calibrate the model. The estimation results show strong positive dependence between joint lives with real economic significance. Annuity values are reduced by approximately 5 percent when dependent mortality models are used compared to the standard models that assume independence.

INTRODUCTION

Financial service organizations offer contractual promises to provide periodic level incomes over the lifetime of individuals. These contracts, called annuities, typically provide a level monthly amount payable until the death of a named individual, called an annuitant. Annuity obligations are offered by insurance companies, pension and other employee benefit funds, and state and federal retirement systems. To illustrate the importance of these obligations, U.S. insurance companies alone made $40.3 billion in annuity payments in 1993 (American Council of Life Insurance, 1994).

An important variation of the standard life annuity is the joint and last-survivor annuity. Under this contract, periodic level payments are made until the last of a group of individuals dies. To illustrate, a prime example of a group is a married couple, where the last-survivor annuity pays as long as either spouse survives.

Edward W. Frees is Professor of Business and Statistics at the University of Wisconsin–Madison. Jacques Carriere is Associate Professor of Actuarial and Management Sciences at the University of Manitoba. Emiliano Valdez is a doctoral student at the University of Wisconsin–Madison.

The research of Jacques Carriere and Emiliano Valdez was supported by the Society of Actuaries. The research of Edward Frees was partially supported by a professorship funded by the Time Insurance Company and the Social Security Administration, Grant No. 10-P-98353-5-01.

Many variations are offered in the marketplace, including a joint and 50 percent annuity that pays a level amount while both annuitants survive with a 50 percent reduction of that amount upon the death of one annuitant.

Valuation of life annuities depends upon the time value of money and the probability of the annuitants' survivorship. The time value of money is important because annuity payments are made in the future with respect to the valuation of the annuity obligation. However, conditional on survivorship of the annuitants, these payments can be valued using standard theory from financial economics, such as the term structure of interest rates. We do not pursue this aspect of annuity valuation here. Instead, we apply the traditional approach of assuming a constant discount rate.

In this article, we focus on estimating the probability of joint survivorship of two annuitants. We focus on two annuitants because this type of contract is issued far more frequently and involves far larger amounts than contracts with more than two annuitants. Estimation of survival probabilities for more than two annuitants can be done by direct extensions of the methods of this article.

Traditionally, estimation of joint survival probabilities of a pair has been done by assuming independence of lives. With this assumption, the probability of joint survival is the product of the probability of survival of each life. This assumption reduces the joint estimation to a single life estimation problem. Estimation of the probability of survivorship of a single annuitant is a well developed area (see, e.g., Elandt-Johnson and Johnson, 1980, or Cox and Oakes, 1990).

However, several empirical studies of joint lives in noncommercial contexts have established that survival of pairs are not independent events. To illustrate, Hougaard, Harvald, and Holm (1992) analyze the joint survival of Danish twins born between 1881 and 1930. Another type of empirical study involves measuring the impact on mortality induced by the death of one's spouse. Parkes, Benjamin, and Fitzgerald (1969) and Ward (1976) provide early examples of this impact, often called the "broken heart" syndrome. A more recent study, with references to many other works, is by Jagger and Sutton (1991).

There are several ways to model the impact of survivorship of one life upon another. For example, as in Jagger and Sutton (1991), the question of increased mortality after an event such as the death of a spouse is well-suited to a survival model called proportional hazards that allows for time-varying explanatory variables. This type of model was also used by Hougaard, Harvald, and Holm (1992) to assess the generations effect by studying twins over a long period of time.

Classical models of dependent lives are called "common shock" models (see, for example, Marshall and Olkin, 1967). These models assume that the dependence of lives arises from an exogenous event that is common to each life. For example, in lifetime analysis this shock may be an accident or the onslaught of a contagious disease. Although there are many other types of dependencies in human lifetimes that are not captured by shock models, their particularly simple form turns out to be convenient for annuity valuation purposes. These models are discussed further below.

Other parametric bivariate survival models include the "frailty" models described by Oakes (1989), the mixture models of Marshall and Olkin (1988) and the "copula" models, as described in Genest and McKay (1986). Nonparametric estimation of bivariate survival estimation has been summarized by Pruitt (1993). Because of the several complications that appear in annuity data, we focus on the copula models. These models provide tractable parametric models of the bivariate distribution and are described in more detail below.

To calibrate our models, we consider data from a large Canadian insurer. We study the mortality experience by observing approximately 15,000 policies over a five-year period, 1989 through 1993. The next section further describes the sources and characteristics of the data.

Because of the nature of our data and our interest in annuity valuation, this article differs from other empirical studies of bivariate distributions in several aspects. First, our data sampling period, five years, is much shorter than other lifetime studies (for example, the Danish twins study used an observation period of 110 years). Thus, we need not discuss cohort effects of mortality as in that study. However, because of the short time frame of our observation period, our data are (right) censored in that most policyholders survived through the end of the observation period. Further, our data are (left) truncated in that policyholders who had died prior to the beginning of the study were not available for analysis. This complication is called "left-truncation with right-censoring" in survival analysis (see, for example, Cox and Oakes, 1990, p. 177).

Second, in this article, the scientific interest is different. Works that study the "broken heart" syndrome often wish to establish predictive models, that is, identifying an event such as the death of a spouse to improve the predictions of the probability of death. Epidemiological studies often wish to isolate explanatory variables that induce the onslaught of a certain disease or infection. Our interest lies in the valuation of annuity contracts. As such, we are interested in assessing the strength of dependence and the effects of the dependence on contract values. In particular, the choice of the model of bivariate dependence is influenced by the desire for simplicity in our annuity valuation procedures.

Third, reporting mechanisms for industry data tend to be different than for population data that might be gathered by the U.S. Census Bureau or data from a carefully designed clinical trial. For industry, two important issues are the mortality patterns within a contractual guarantee period and the reporting of the first death for some contracts. Some joint-life contracts offer a guarantee of annuity payments, typically over a period of five or ten years from contract initiation. Thus, there is no economic incentive for reporting a death within the guarantee period. Alternatively, those that elect the guarantee option may exhibit higher mortality than those who do not. Further, for joint and last-survivor policies, payments are made until the second, or last, of the pair dies. Thus, although policyholders should notify the insurance company of a change in the mortality status of annuitants, there is no economic purpose for reporting the death of the first annuitant. These two issues are addressed in greater detail below.

The next two sections introduce the data and discuss models of dependence. Then, we summarize the effects of dependence on annuity values and address the problem of underreporting. Finally, we examine the robustness of choice of parametric families by considering an alternative marginal distribution and an alternative bivariate distribution function, the common shock model.

DATA CHARACTERISTICS

This article analyzes mortality patterns based on information from 14,947 contracts in force with a large Canadian insurer over the period December 29, 1988, through December 31, 1993. These contracts are joint and last-survivor annuities that were in the payout status over the observation period.

For each contract, we have the date of birth, date of death (if applicable), date of contract initiation, and sex of each annuitant. Table 1 presents the frequency distribution of annuitants by sex, entry age, and mortality status group. Entry age is defined to be the age at which the annuitant entered the study and was computed from the date of birth and contract initiation date. For mortality status, we classify annuitants according to whether they survived until the end of the observation period. In addition to the dates discussed above, we also have the date that the annuity guarantee expired (if applicable). This will be discussed further below.

Table 1
Number of Policies by Sex, Entry Age, and Mortality Status

	Mortality Status		
Entry Age	Survive	Death	Total
Males			
Less than 60	1,170	42	1,212
60 – 70	7,620	534	8,154
70 – 80	4,355	806	5,161
Greater than 80	229	177	406
Total	13,374	1,559	14,933
Females			
Less than 60	2,962	30	2,992
60 – 70	8,222	239	8,461
70 – 80	3,014	245	3,259
Greater than 80	186	63	249
Total	14,384	577	14,961

There was roughly an equal number of males and females in our study, 14,933 and 14,961, respectively. Roughly three times more males as females died during the study period, due in part to the higher average entry age for males (68) than for females (65). It also suggests higher mortality rates for males than for females.

Figure 1 displays a graphical summary of the distribution of lifetimes for the 14,933 male annuitants. Because our data primarily concern policyholders who are at least middle-aged, we graph distribution functions that are conditional on survival to age 40. The jagged line in Figure 1 is the Kaplan-Meier product limit estimator of the distribution function. We use this as our baseline estimator of the distribution function because it is the nonparametric maximum likelihood estimator. See, for example, Elandt-Johnson and Johnson (1980) or Cox and Oakes (1990) for an introduction and further discussion of the properties of this estimator. From the Kaplan-Meier estimates, the median age at death is approximately 82 years. Further, the 25th and 75th percentiles are approximately 68 and 90 years, respectively.

Figure 1
Gompertz and Kaplan-Meier Fitted Male Distribution Functions

Note: The Gompertz curve is smooth, the Kaplan-Meier is jagged. The distribution is conditional on survival to age 40.

Superimposed in Figure 1 is a smooth curve that was fit using the Gompertz distribution. The Gompertz distribution function can be expressed as:

$$F(x) = 1 - \exp(e^{-m/\sigma}(1 - e^{x/\sigma})), \tag{1}$$

where the mode, m, and the scale measure, σ, are parameters of the distribution. To actuaries, the familiar Gompertz force of mortality, or hazard rate, is $\mu_x = F'(x)/(1-F(x)) = Bc^x$, that yields

$$F(x) = 1 - \exp((B/\ln c)\,(1 - c^x)).$$

However, with the transformations $B/\ln c = e^{-m/\sigma}$ and $c = e^{1/\sigma}$, we see that equation (1) is simply a reparameterized version of the usual expression for the Gompertz distribution. As pointed out by Carriere (1994), equation (1) is convenient for estimation purposes.

As shown in Figure 1, the Gompertz fit closely replicates the nonparametric Kaplan-Meier fit. The main advantage of the Gompertz fit is that only two parameter estimates are necessary to reproduce this curve. For male annuitants, the estimates turn out to be $\hat{m} \approx 86.4$ and $\hat{\sigma} \approx 9.8$ years. With these estimates and equation (1), the 25th, 50th, and 75th percentiles turn out to be 74.2, 82.8, and

89.6, respectively. However, to reproduce the Kaplan-Meier estimate, all 14,933 male lifetimes would be needed because of the continuous nature of our data. The parsimonious representation provided by the Gompertz curve is particularly important for the annuity calculations below.

Of course, to achieve a parsimonious representation of a lifetime distribution, many other families of distribution functions could be used. To illustrate, Figure 2 shows a fitted Weibull distribution function for the male annuitants with the Kaplan-Meier curve superimposed for reference.

Figure 2
Weibull and Kaplan-Meier Fitted Male Distribution Functions

Note: The Weibull curve is smooth, the Kaplan-Meier is jagged. The distribution is conditional on survival to age 40.

The fit is close, although the Gompertz distribution may be a better approximation of the Kaplan-Meier estimator. For purposes of annuity valuation, it turns out that any parametric representation of the lifetime distribution suffices. For our data set of older policyholders, the Gompertz distribution seems to provide an adequate fit. There is certainly a long history of fitting Gompertz distributions to the mortality of human populations, as described by Carriere (1994).

To provide background on the estimation procedures on which Figures 1 and 2 are based, we now give details on the limitations of our data, including truncation and censoring. Consider the bivariate ages-at-death random vector (X, Y), where X and Y represent the ages at death of the primary and secondary annuitant, respectively. In joint annuity contracts, one annuitant is usually designated as "primary" and the other "secondary" because some contracts provide for a reduced payout upon the death of the primary annuitant. An example of this is an annuity provided by a firm to an employee. This distinction turns out to be unimportant in our analysis of the data.

Industry data are truncated in the sense that data are observed only after a contract has been entered into by policyholders. Thus, we use standard notation and let x and y be the contract initiation ages of the primary and secondary annuitants, respectively. Further, for our data set, we observe the joint annuity contract if both annuitants are alive at the beginning of the observation period or if the annuitants enter the study during the observation period. With t_0 as the time of contract initiation, we define

$$a = \max(12/29/88 - t_0, 0)$$

to be the time from contract initiation to the beginning of the observation period. Thus, $x + a$ and $y + a$ are the entry ages of the primary and secondary annuitants, respectively. Under our left-truncation, we observe the contract only if $X > x + a$ and $Y > y + a$.

Our data are also censored from the right. Let $b = 1/1/94 - \max(12/29/88, t_0)$ denote the length of time that the policy was under observation. Denote $T_1 = X - x - a$ and $T_2 = Y - y - a$ to be the future annuitant lifetimes. Then, for $j = 1, 2$, we observe $T_j^* = \min(T_j, b)$, the censored future lifetime, and δ_j, a variable to indicate whether censoring has occurred. That is, δ_j is defined to be one if $T_j > b$ and zero otherwise.

Using this notation, our full data set consists of $\{T^*_{ij}, \delta_{ij}\}$, $j = 1, 2$, and $i = 1,\ldots, 14{,}947$. That is, there are a total of 29,894 ($2 \times 14{,}947$) univariate observations. Univariate distributions were fit for each sex using maximum likelihood techniques to produce the fitted parametric curves. We do not present the details here because the more complex bivariate situation is discussed in the next section. As mentioned above, the nonparametric curves were fit using the standard Kaplan-Meier product limit estimator.

Beginning in the next section, we focus our estimation procedures on bivariate observations, that is, observations of the joint mortality of both annuitants. For our data, we have 22 contracts where both annuitants are male, 36 contracts where both annuitants are female, and 14,889 contracts where one annuitant is male and the other female. Because of the preponderance of data in the third category, we focus our attention on male-female joint annuity mortality. Henceforth, we refer to x as the male life and y as the female life. Of course, the estimation techniques that we introduce also could be applied to the other two categories.

A classical nonparametric measure of dependence is Spearman's rank correlation. Appendix E shows how we can use this measure when the data are left-truncated, right-censored, and not identically distributed. The test of independence that we present assumes that the law of mortality is known for individual lives. Obviously, our knowledge of this law for annuity products is substantial but not perfect and so this method must be used with caution. Using this technique, we find that the correlation would be 0.41 and that a 95 percent confidence interval is (0.28, 0.55). If the lives were independent, then the correlation would be zero. Therefore, this crude preliminary analysis suggests that the lifetimes are dependent. In subsequent sections, we corroborate this analysis using maximum likelihood techniques.

MODELS OF DEPENDENCE

Bivariate Distributions

In this article, we express bivariate distributions using a function called a copula. Consider a bivariate age-at-death random vector (X, Y) with distribution function H, that is, $H(x, y) = \text{Prob}(X \le x, Y \le y)$. Let F_1 and F_2 denote the respective marginal distribution functions so that $F_1(x) = H(x, \infty)$ and $F_2(y) = H(\infty, y)$. We consider bivariate distribution functions of the form

$$H(x, y) = C(F_1(x), F_2(y)). \qquad (2)$$

Here, C is a real-valued function called a copula. Copulas are bivariate distributions with uniform marginal distributions.

Copulas are useful because they provide a link between the marginal distributions and the bivariate distribution. From equation (2), it is clear that, if F_1, F_2, and C are known, then H can be determined. Sklar (1959) proved a converse: if H is known and if F_1 and F_2 are known and continuous, then C is uniquely determined. In this sense, C "couples" the marginal distributions to the bivariate distribution.

There are many possible choices of the copula function. This article focuses on a one-parameter family due to Frank (1979) that can be expressed as

$$C(u,v) = \ln(1 + \frac{(e^{\alpha u}-1)(e^{\alpha v}-1)}{e^{\alpha}-1}) / \alpha . \qquad (3)$$

Advantages of this family have been presented by Nelsen (1986) and Genest (1987). The Frank, Genest, and Nelsen articles present the copula in terms of the parameter $\gamma = e^{\alpha}$. Similar to the case of the univariate Gompertz distribution, we work with the reparameterized version in equation (3). This transformation turns out to be more convenient for estimation purposes.

The parameter α captures the dependence between X and Y. The case of independence corresponds to $\alpha = 0$. This is because it can easily be shown, from equation (3), that $\lim_{\alpha \to 0} C(u, v) = uv$. Thus, the bivariate distribution function is the product of marginal uniform distributions. In addition to the dependence parameter α, we also present Spearman's correlation coefficient $\rho(\alpha)$. Spearman's correlation coefficient is a nonparametric measure, defined to be the ordinary Pearson correlation coefficient after taking a (marginal) uniform transformation of each random variable. Here, $\rho(\alpha)$ is a straightforward function of α that Nelsen (1986) showed to be

$$\rho(\alpha) = 1 - 12\,(D_2(-\alpha) - D_1(-\alpha))/\alpha, \qquad (4)$$

where $D_k(x) = kx^{-k} \int_0^x t^k(e^t - 1)^{-1}dt$, k = 1, 2 is called the Debye function.

To complete our specification of the bivariate distribution, we assume that each marginal distribution is Gompertz. Thus, using equation (1), we assume that

$$F_j(x) = 1 - \exp(\exp(-m_j/\sigma_j)(1 - \exp(x/\sigma_j))), j = 1, 2. \qquad (5)$$

Our model is then specified by equations (2), (3), and (5). This model has five parameters which can be represented by the vector

$$\theta = (m_1, \sigma_1, m_2, \sigma_2, \alpha)'. \tag{6}$$

Estimation Results

Using the left-truncated, right-censored data summarized in Table 1, we estimate the model described above. The method of estimation is maximum likelihood; the details are presented below. The results of the estimation are summarized in Table 2.

Table 2
Bivariate Data Parameter Estimates

	Bivariate Distribution		Univariate Distribution	
Parameter	Estimate	Standard Error	Estimate	Standard Error
m_1	85.82	0.26	86.38	0.26
σ_1	9.98	0.40	9.83	0.37
m_2	89.40	0.48	92.17	0.59
σ_2	8.12	0.34	8.11	0.38
α	–3.367	0.346	Not Applicable	Not Applicable

Table 2 shows that the "average," or modal, age at death is approximately fours years later for females than males. The estimates of variability are roughly the same. Using equation (4), the estimate of the dependence parameter can be converted to a correlation estimate. This turns out to be $\rho(\hat{\alpha}) = \rho(-3.367) = 0.49$. Recall that Spearman's correlation, like Pearson's correlation, is bounded by –1 and 1 with a correlation of zero implying no relationship. A value of $\rho(\hat{\alpha}) = 0.49$ indicates a strong statistical dependence. This is because a rough 95 percent confidence interval for α is $\hat{\alpha} \pm 1.96\ se(\hat{\alpha}) = -3.367 \pm 1.96\ (0.346) = (-4.045, -2.689)$. Translated into the correlation scale, a 95 percent confidence interval of Spearman's correlation is (0.41, 0.56). The parameter estimates presented in Table 2 can be directly used to value annuities.

Maximum Likelihood Estimation

We now develop the likelihood function to be maximized. Having developed the likelihood function, standard function maximization routines will yield the maximum likelihood estimates. Because our sampling satisfies standard regularity conditions (see, for example, Serfling, 1980), we can easily obtain asymptotic normality and subsequent standard errors for the estimates.

In our development, we need the following partial derivatives: $H_1(x, y) = \partial H(x, y)/\partial x$, $H_2(x, y) = \partial H(x, y)/\partial y$, and $h(x, y) = \partial^2 H(x, y)/\partial x \partial y$. Our assumption of Gompertz marginals and Frank's copula assures that these derivatives exist. Their explicit representation in terms of the vector of parameters is given in Appendix A.

To develop the likelihood function, we first consider truncated observations. Recall that the future lifetime random variables $T_1 = X - x - a$ and $T_2 = Y - y - a$ are observed only if $T_1 > 0$ and $T_2 > 0$. Therefore, define the conditional distribution function of T_1 and T_2 as

$$H_T(t_1, t_2) = \text{Prob}\,(T_1 \le t_1, T_2 \le t_2 | T_1, T_2 \text{ are observed}) = \frac{\text{Prob}(0<T_1\le t_1,\ 0<T_2\le t_2)}{\text{Prob}(T_1>0,\ T_2>0)}$$

$$= \frac{H(x+a+t_1, y+a+t_2) - H(x+a, y+a+t_2) - H(x+a+t_1, y+a) + H(x+a, y+a)}{1 - H(x+a, \infty) - H(\infty, y+a) + H(x+a, y+a)}. \tag{7}$$

Turning now to the case of right-censoring, recall that $T_j^* = \min(T_j, b)$. Four types of censoring may occur. The lifetimes may be both uncensored, the first uncensored and the second censored, the first censored and the second uncensored, and both censored. We handle each type in turn.

If both lifetimes are uncensored, then we may assume $t_1 < b$ and $t_2 < b$. In this case, we have $\delta_1 = 0$ and $\delta_2 = 0$ and

$$\text{Prob}(T_1^* < t_1, T_2^* < t_2 \mid T_1^*, T_2^* \text{ are observed})$$
$$= \text{Prob}(\min(T_1, b) < t_1, \min(T_2, b) < t_2 | T_1 > 0, T_2 > 0) = H_T(t_1, t_2).$$

Thus, using equation (7), the contribution to the likelihood function is

$$\frac{\partial^2 H_T(t_1, t_2)}{\partial t_1\, \partial t_2} = \frac{h(x+a+t_1, y+a+t_2)}{1 - H(x+a, \infty) - H(\infty, y+a) + H(x+a, y+a)}. \tag{8}$$

If the first lifetime is uncensored and the second is censored, then we may assume $t_1 < b$ and $t_2 \ge b$. In this case, we have $\delta_1 = 0$ and $\delta_2 = 1$ and

$$\text{Prob}(T_1^* < t_1, T_2^* = b \mid T_1^*, T_2^* \text{ are observed})$$
$$= \text{Prob}(T_1 < t_1, T_2 \ge b \mid T_1 > 0, T_2 > 0) = H_T(t_1, \infty) - H_T(t_1, b).$$

Thus, the contribution to the likelihood function is

$$\frac{\partial (H_T(t_1, \infty) - H_T(t_1, b))}{\partial t_1} = \frac{H_1(x+a+t_1, \infty) - H_1(x+a+t_1, y+a+b)}{1 - H(x+a, \infty) - H(\infty, y+a) + H(x+a, y+a)}. \tag{9}$$

If the first lifetime is censored and the second is uncensored, then we may assume $t_1 \ge b$ and $t_2 < b$. This case is similar to the previous case; thus, $\delta_1 = 1$ and $\delta_2 = 0$ and the contribution to the likelihood function is

$$\frac{H_2(\infty, y+a+t_2) - H_2(x+a+b, y+a+t_2)}{1 - H(x+a, \infty) - H(\infty, y+a) + H(x+a, y+a)}. \tag{10}$$

If both lifetimes are censored, then we may assume $t_1 \ge b$ and $t_2 \ge b$. In this case, we have $\delta_1 = 1$, $\delta_2 = 1$ and contribution to the likelihood function is

$$\text{Prob}(T_1^* = b, T_2^* = b \mid T_1^*, T_2^* \text{ are observed})$$

$$= \frac{1 - H(x+a+b, \infty) - H(\infty, y+a+b) + H(x+a+b, y+a+b)}{1 - H(x+a, \infty) - H(\infty, y+a) + H(x+a, y+a)}. \tag{11}$$

Combining equations (8) through (11), we can express the logarithm of the likelihood function for a single observation as

$$\begin{aligned}\ln L(x,y,t_1,t_2,\delta_1,\delta_2,a,b) &= (1-\delta_1)(1-\delta_2)\ \ln\ h(x+a+t_1,y+a+t_2)\\ &+ (1-\delta_1)\delta_2\ \ln(H_1(x+a+t_1,\infty)-H_1(x+a+t_1,y+a+b))\\ &+\delta_1(1-\delta_2)\ \ln(H_2(\infty,y+a+t_2)-H_2(x+a+b,y+a+t_2))\\ &+ \delta_1\delta_2\ \ln(1-H(x+a+b,\infty)-H(\infty,y+a+b)+H(x+a+b,y+a+b))\\ &- \ln(1-H(x+a,\infty)-H(\infty,y+a)+H(x+a,y+a))\,. \end{aligned} \tag{12}$$

Using equation (12), the log-likelihood for the data set can be calculated as

$$\ln \mathcal{L} = \sum_{i=1}^{n} \ln\ L(x_i,\ y_i,\ t_{1i},\ t_{2i},\ \delta_{1i},\ \delta_{2i},\ a_i,\ b_i)\,. \tag{13}$$

The maximum likelihood estimator of θ is the value $\hat{\theta}$ that maximizes $\ln \mathcal{L}$. Standard maximum likelihood estimation theory provides that $n^{1/2}(\hat{\theta} - \theta)$ is asymptotically normally distributed with mean zero and variance-covariance $V_n(\theta) = -(n^{-1}\partial^2(\ln \mathcal{L})/(\partial\theta\partial\theta'))^{-1}$. The variance-covariance matrix $V_n(\theta)$ can be consistently estimated using $V_n(\hat{\theta})$, which is an output from standard function maximization routines. The standard error of each parameter estimate in $\hat{\theta}$ may be determined by the square root of the corresponding diagonal element of $V_n(\hat{\theta})/n$.

Effects of Dependence on Annuity Values

As described in Section 1, the main purpose of this article is to assess the effects of our models of dependent mortality on annuity values. We use the basic model of annuity valuation described in Chapter 8 of Bowers et al. (1986). To this end, the illustrations below assume a constant effective interest rate i with associated discount rate $v = 1/(1 + i)$. The net single premium for a joint and last-survivor annuity issued to lives aged x and y is

$$\ddot{a}_{\overline{xy}} = \sum_{k=0}^{\infty} v^k\ {}_k p_{\overline{xy}}, \tag{14}$$

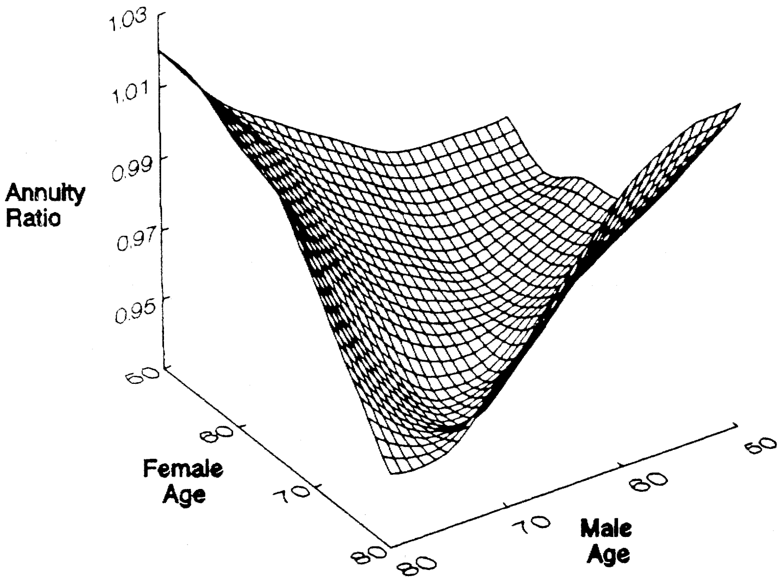
where ${}_k p_{\overline{xy}} = 1 - H_T(k, k)$ is the conditional probability that at least one life survives an additional k years. Here, the conditional distribution H_T is as defined in equation (7) with $a = 0$. Because H_T is a function of the vector of parameters θ, so are ${}_k p_{\overline{xy}}$ and $\ddot{a}_{\overline{xy}}$. We occasionally use the notation $\ddot{a}_{\overline{xy}}(\theta)$ to emphasize this dependence.

Effects of Age and Interest

We identify contract initiation ages x and y and interest environment i where models of dependent mortality really matter. The approach is straightforward: we estimate the annuity value $\ddot{a}_{\overline{xy}}$ in equation (14) with and without assuming independence. The parameter estimates with and without independence are discussed above, with Table 2. We compare the annuity values by calculating the ratio of annuity values estimated without an independence assumption to those estimated with the independence assumption.

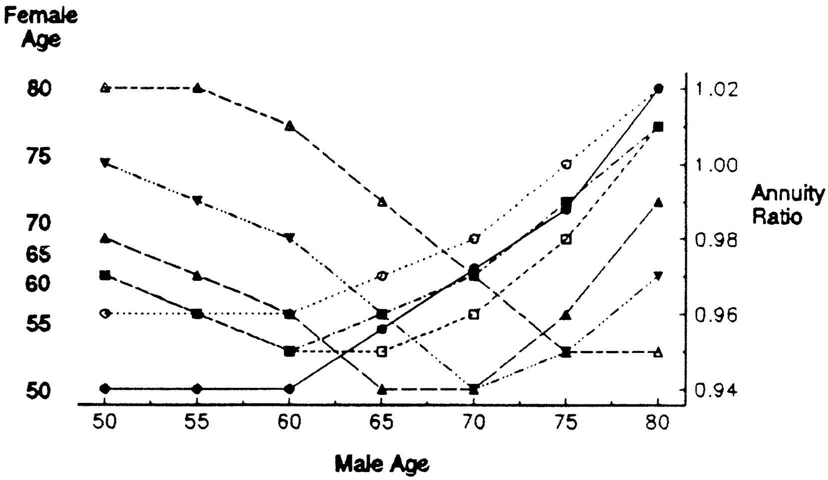
To assess the effects of contract initiation ages, Figure 3 presents a three-dimensional plot of the ratio of annuity values, by male (x) and female (y) ages. The curve is roughly symmetric in x and y, indicating that, although the two marginal distributions are different, they have approximately the same effect on the ratio. Further, there appears to be an interaction effect of x and y on the ratios. That is, the ratio is much smaller for large values of both x and y when compared to large values for either x or for y.

Figure 3
Three-Dimensional Plot of the Ratio of Dependent to Independent Annuity Values for Males and Females Aged 50 to 80
(Five Percent Interest Is Assumed)



To gain further understanding of this interaction effect, Figure 4 presents a multiple scatter plot of the ratio versus male ages, over several female ages. Here, for young female ages, we see that the ratio increases as male age increases. However, for older female ages, the ratio decreases as male age increases.

Figure 4
Multiple Scatter Plot of the Ratio of Dependent to Independent Annuity Values to Male Age, Over Several Female Ages (Five Percent Interest Is Assumed)



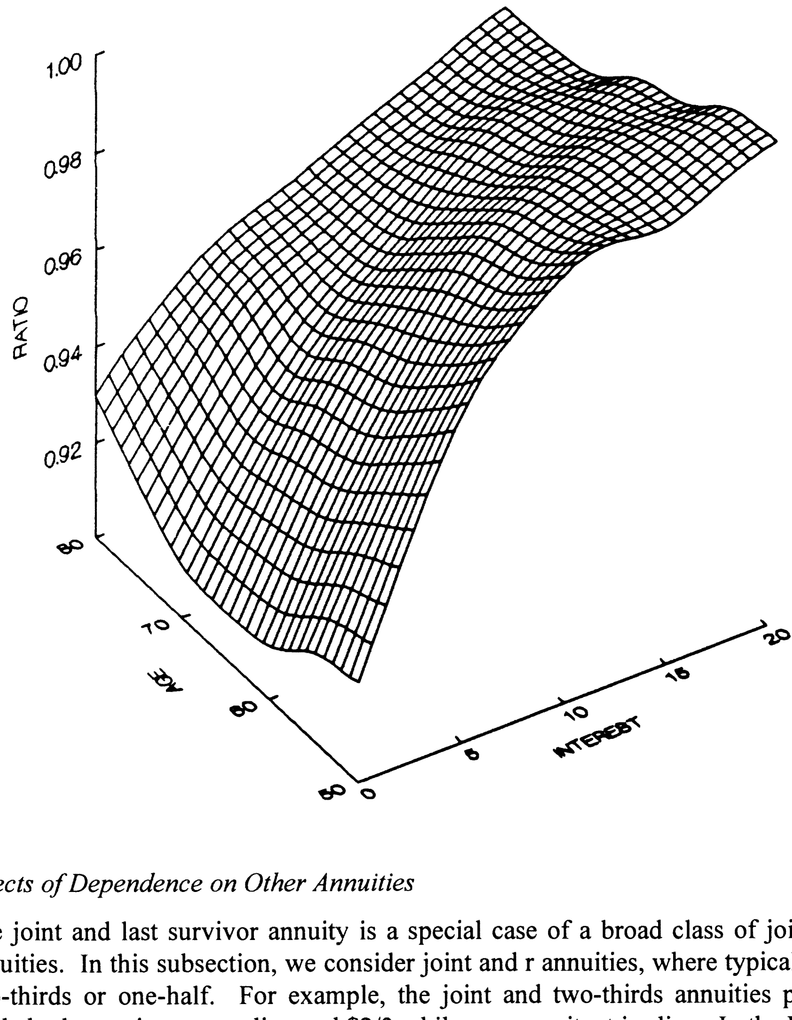
Ratios of less than one indicate that annuity values calculated assuming independence of lives are larger than those calculated without assuming independence. For this data set, it turns out that the average contract initiation age is approximately 65 for males and 63 for females. The average age at December 31, 1993, is 72.5 for male lives and 69.6 for female lives. We interpret the higher ratios for younger ages to mean that the effect of assuming independence is smaller for premium determination compared to annuity reserve setting. This suggests that reserves for annuities already paid-up are larger than necessary.

Our data set displays a strong relationship between x and y. An examination of the data shows that the median age difference is 2.4 years, the middle 50 percent of the data is between 0.1 and 5.2 years, and the middle 90 percent is between –3.7 and 11.0 years. Because of this concentration, for brevity in the subsequent analyses, we present only the special case of $x = y$.

To assess the effects of interest, Figure 5 presents a three-dimensional plot of the ratio of annuity values, over several interest rates and ages. Here, the male age is assumed equal to the female age. This figure shows a quadratic effect of joint age that can also be observed in Figure 3. The effects of the interest rate i seem to be linear.

To investigate these effects further, Figure 6 presents a multiple scatter plot of the ratio of annuity values to age, over several interest rates. This plot also demonstrates the quadratic effect of joint age and the linear effect of interest rates. From this plot, we see that the assumption of independence will matter more in times of low interest rates than in times of high interest rates.

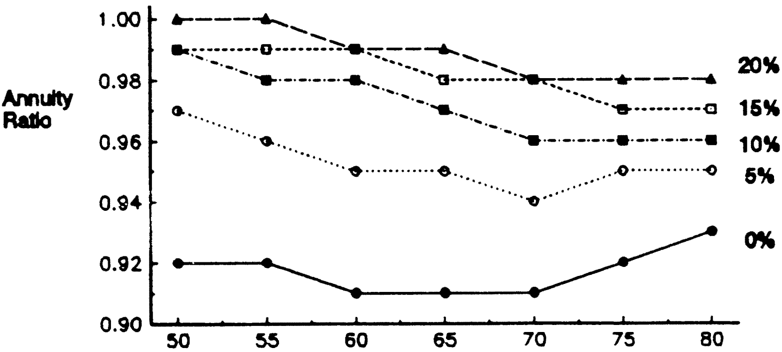
Figure 5
Three-Dimensional Plot of the Ratio of Dependent to Independent Annuity Values, Over Several Interest Rates and Ages (Equal Annuitant Ages Are Assumed)



Effects of Dependence on Other Annuities

The joint and last survivor annuity is a special case of a broad class of joint-life annuities. In this subsection, we consider joint and r annuities, where typically r is two-thirds or one-half. For example, the joint and two-thirds annuities pay \$1 while both annuitants are alive and \$2/3 while one annuitant is alive. In the United States, there may be a larger market for these annuities than the joint and last-survivor annuities that corresponds to r = 1 because the Employee Retirement Income Security Act (ERISA) mandates that all qualified pension plans offer to qualified beneficiaries a joint and survivor annuity with r at least 50 percent. For this purpose, in addition to the usual requirements, beneficiaries must be married to their current spouse for at least one year. Joint-life annuities, corresponding to r = 0, are not widely marketed.

Figure 6
Multiple Scatter Plot of the Ratio of Dependent to Independent Annuity Values to Age, Over Several Interest Rates (Equal Annuitant Ages Are Assumed)



Similar to equation (14), the net single premium for a joint and r annuity can be expressed as

$$\ddot{a}_{xy}(r) = \sum_{k=0}^{\infty} v^k(r\,{}_kp_x + r\,{}_kp_y - (2r-1)\,{}_kp_{xy}), \tag{15}$$

where ${}_kp_x = 1 - H_T(k, \infty)$ is the conditional probability that a life age x survives an additional k years, ${}_kp_y = 1 - H_T(\infty, k)$ is the conditional probability that a life age y survives an additional k years, and ${}_kp_{xy} = {}_kp_x + {}_kp_y - {}_kp_{\overline{xy}} = 1 - H_T(k, \infty) - H_T(\infty, k) + H_T(k, k)$ is the conditional probability that both lives ages x and y survive an additional k years.

Table 3 summarizes the effects of dependence on the reduced annuities. The reduction factor r has little effect on the annuity ratios. As discussed above, we are concerned primarily with reduction factors r = 1/2, 2/3, and 1 because these are the most widely marketed types of annuities.

Annuity Standard Errors

The annuity values calculated above are based on the point estimate $\hat{\theta}$ and thus depend on the sample. To measure the reliability of $\hat{\theta}$, in the maximum likelihood estimation section above, we discussed how the estimated variance-covariance matrix $V_n(\hat{\theta})$ could be used to derive parameter estimate standard errors. This subsection develops standard errors for $\ddot{a}_{\overline{xy}}(\hat{\theta})$. These standard errors, together with the asymptotic normality, allow us to provide confidence intervals for our joint and last-survivor annuity values.

Table 3
Ratios of Dependent to Independent Joint and r Annuity Values (Five Percent Interest and Equal Annuitant Ages Are Assumed)

	r					
Age	0	1/4	1/3	1/2	2/3	1.0
50	1.00	0.99	0.99	0.98	0.98	0.97
55	1.00	0.98	0.98	0.98	0.97	0.96
60	0.99	0.98	0.98	0.97	0.96	0.95
65	0.98	0.97	0.97	0.96	0.96	0.95
70	0.97	0.96	0.96	0.95	0.95	0.94
75	0.94	0.94	0.94	0.94	0.94	0.94
80	0.89	0.91	0.92	0.93	0.94	0.95

The asymptotic normality of $\ddot{a}_{\overline{xy}}(\hat{\theta})$ is based on the asymptotic normality of $\hat{\theta}$ and the so-called "delta-method" (see, for example, Serfling, 1980). Recall from the discussion of maximum likelihood estimation that

$$n^{1/2}(\hat{\theta} - \theta) \text{ is } AN(0, V_n(\theta)),$$

where AN(0, A) means asymptotically normal with mean vector 0 and variance-covariance matrix A. Define the gradient vector $G(\theta) = \partial\ddot{a}_{\overline{xy}}(\theta)/\partial\theta$. (See Appendix C for details of the calculation of $G(\theta)$.) From the delta method, we have

$$n^{1/2}(\ddot{a}_{\overline{xy}}(\hat{\theta}) - \ddot{a}_{\overline{xy}}(\theta)) \text{ is } AN(0, G(\theta)' V_n(\theta)\, G(\theta)). \tag{16}$$

Thus, we may define the standard error of $\ddot{a}_{\overline{xy}}(\hat{\theta})$ as

$$se(\ddot{a}_{\overline{xy}}(\hat{\theta})) = (G(\hat{\theta})'V_n(\hat{\theta})G(\hat{\theta})/n)^{1/2}. \tag{17}$$

From equation (16), we have that $\ddot{a}_{\overline{xy}}(\hat{\theta}) \pm 1.96\ se(\ddot{a}_{\overline{xy}}(\hat{\theta}))$ provides an approximate 95 percent confidence interval for our annuity value $\ddot{a}_{\overline{xy}}(\theta)$.

When computing the standard error, the most difficult component is the gradient vector. This is because, as noted above, the matrix $V_n(\hat{\theta})$ is an automatic output from standard function maximization routines. To compute the gradient vector, from equation (14), we have

$$G(\theta) = \frac{\partial}{\partial\theta}\ \ddot{a}_{\overline{xy}}(\theta) = \sum_{k=0}^{\infty} v^k \frac{\partial}{\partial\theta}\ {}_k p_{\overline{xy}}(\theta) = -\sum_{k=0}^{\infty} v^k \frac{\partial}{\partial\theta}\ H_T(k, k)\,.$$

From equation (17) and the chain rule, we have

$$-\frac{\partial}{\partial\theta} H_T(k, k) = (1 - H(x,\infty) - H(\infty,y) + H(x,y))^{-2}$$

$$[(H(x+k,y+k) - H(x,y+k) - H(x+k,y) + H(x,y))\frac{\partial}{\partial\theta}(1 - H(x,\infty) - H(\infty,y) + H(x,y)) \quad (18)$$

$$- (1 - H(x,\infty) - H(\infty,y) + H(x,y))\frac{\partial}{\partial\theta}(H(x+k,y+k) - H(x,y+k) - H(x+k,y) + H(x,y))].$$

Using equation (18), joint and last-survivor annuity standard errors are computed over several ages and interest rates. Table 4 presents the results for a five percent interest rate. The most important aspect of Table 4 is the magnitude of $se(\ddot{a}_{\overline{xy}}(\hat{\theta}))$. To illustrate, consider our largest and smallest estimated annuity values, which are $\ddot{a}_{\overline{50:50}}(\hat{\theta})$ = 17.45 and $\ddot{a}_{\overline{80:80}}(\hat{\theta})$ =9.65. For the largest annuity values, the standard error represents a typical error that is 0.002/17.45 = 0.011 percent of the annuity. For the smallest annuity value, the standard error represents a typical error that is 0.025/9.65 = 0.26 percent. Thus, the standard errors indicate that the estimated annuity values are very accurate, assuming that the model is correct.

Table 4
Annuity Standard Errors by Male and Female Age
(Five Percent Interest Is Assumed)

	Female Age						
Male Age	50	55	60	65	70	76	80
50	0.002	0.002	0.002	0.002	0.002	0.003	0.005
55	0.002	0.002	0.003	0.002	0.002	0.003	0.004
60	0.002	0.002	0.003	0.005	0.004	0.004	0.004
65	0.002	0.002	0.003	0.006	0.008	0.008	0.006
70	0.002	0.003	0.003	0.006	0.010	0.014	0.013
75	0.002	0.003	0.004	0.006	0.009	0.016	0.023
80	0.003	0.004	0.007	0.009	0.011	0.015	0.025

THE PROBLEM OF UNDERREPORTING

As noted above, the data analyzed in this article come from internal records of a large insurance company. Thus, in most cases, the accuracy of the observed lifetimes depends on the reporting behavior of policyholders. This section investigates two instances where substantial measurement errors may exist. The first instance involves underreporting of deaths within the guarantee period and the second involves underreporting of the first death. In each case, the approach is to reformulate the likelihood equation so that we are essentially re-estimating the models of dependence using only subsets of our data. The subsets are chosen to circumvent the potential bias due to underreporting. Unfortunately, by using only subsets of the data, we are unable to estimate the parameter values accurately when assessing potential underreporting of the first death. However, we do present the theoretical development of the likelihood equation to handle this type of underreporting.

Underreporting Within the Guarantee Period

Many policyholders elect a standard option that guarantees annuity payments will be made within a contractually specified period regardless of the mortality status of the annuitants. Because of the lack of financial incentives, there is concern that policyholders may not accurately report deaths that occur during the guarantee period. Further, it may be that mortality patterns for those electing a guarantee option may differ from those who do not. Of the 14,889 joint life contracts that we used for estimation, 10,011 contracts were at least partially guaranteed during the observation period.

To handle this potential bias, we re-estimate the model disregarding mortality events within the guarantee period by censoring our lifetime data from the left at the expiration of the guarantee period. A consequence of this left censoring is that, for policies whose contract guarantee exceeded the observation period, there is no variability in the observed lifetimes and hence these policies were completely excluded from the likelihood calculations. Of the 14,889 contracts, 9,172 had guarantees that exceed the observation period.

To define the variables needed for the new likelihood function, let c be the time since contract initiation of the guarantee period. Define $g = c - a$ to be the time from the beginning of the observation period to the end of the guarantee period (which may be negative). For contracts without a guarantee, we define $g = 0$.

Our new likelihood is based on the right- and left-censored times at death

$$T_j^{**} = \min(\max(T_j, g), b)$$

and indicators of the type of censoring

$$\delta_{jg} = \begin{cases} 1 \text{ if } T_j^{**} = g \text{ (left-censored)} \\ 0 \text{ otherwise} \end{cases}$$

$$\delta_{jb} = \begin{cases} 1 \text{ if } T_j^{**} = b \text{ (right-censored)} \\ 0 \text{ otherwise} \end{cases}$$

for $j = 1, 2$. Because contracts with guaranteed period exceeding the observation period are excluded from the likelihood function, we may assume $g < b$ without loss of generality.

The development of the guarantee period likelihood function is similar to that described in the maximum likelihood estimation section above. Because it is more complex, the details are included in Appendix B. The guarantee likelihood is analogous to the classic "select-and-ultimate" tables in life analysis, where the experience of policyholders during the select period is not used for calculating ultimate mortality rates. However, our analysis assumes that there is one mortality law. We use the information in the guarantee period in a specific way to reduce the potential bias of the parameter estimates.

Table 5 presents the estimation results for the new likelihood that accounts for the presence of the guarantee period. The parameter estimates from the guarantee period likelihood do not differ significantly from those of the full likelihood. As anticipated, the standard errors are larger for the guarantee period than for those of the full likelihood, due to the fact that we are using less information under the guarantee period censoring.

The data presented in Table 5 do not suggest the presence of adverse selection by policyholders who elect the guarantee option. On one hand, if there was an underreporting of deaths within the guarantee period, then we would expect the modal ages to decrease. On the other hand, if policyholders with poorer health elect guarantee options, then we would expect the modal ages to increase. Our guarantee likelihood estimates do not significantly differ from the full likelihood estimates, thus providing no conclusive evidence of the presence of adverse selection.

Table 5
Guarantee Period Parameter Estimates

	Full Likelihood		Guarantee Period Likelihood	
Parameter	Estimate	Standard Error	Estimate	Standard Error
m_1	85.82	0.26	84.78	0.40
σ_1	9.98	0.40	9.58	0.49
m_2	89.40	0.48	89.53	0.78
σ_2	8.12	0.34	7.82	0.40
α	–3.367	0.346	–2.92	0.623

The data presented in Table 5 also show that our estimate of dependence has decreased to $\rho(\hat{\alpha}) = \rho(-2.92) = 0.44$. Although not statistically different from the full likelihood estimates, there may be some economic significance. The data presented in Table 6 show that this is not the case. The ratios in Table 6 are approximately equal to the ratios presented in Table 3. This again illustrates the highly nonlinear nature of the dependence parameter; large changes in α are needed to induce even small changes in the ratios of annuity values.

Table 6
Ratios of Dependent to Independent Joint and Last Annuity Values Based on Guarantee Likelihood Estimates (Five Percent Interest Is Assumed)

	Female Age						
Male Age	50	55	60	65	70	75	80
50	0.97	0.96	0.96	0.96	0.97	0.99	1.01
55	0.97	0.96	0.95	0.95	0.96	0.97	1.00
60	0.97	0.96	0.95	0.94	0.94	0.96	0.99
65	0.98	0.97	0.96	0.94	0.93	0.94	0.97
70	0.99	0.98	0.97	0.95	0.94	0.93	0.94
75	1.01	1.00	0.99	0.98	0.96	0.93	0.92
80	1.01	1.01	1.01	1.01	0.99	0.96	0.93

Underreporting of First Death

For joint and last-survivor annuity policies, payments are made until the second, or last, annuitant dies. There is concern that policyholders might not report the death of only one annuitant, especially if there is no effect on the level of payment. For our data, of the 2,126 deaths (1,554 males + 572 females), 1,668 deaths left the other annuitant surviving. The other 458 deaths resulted in the cessation of payments on 229 contracts by the end of the observation period.

To handle this potential bias, we re-estimated the model by redefining "failure" to be the time of second death, that is, cessation of the policy. This approach treats the 1,668 single deaths as policies that "survive" the observation period.

Our new likelihood is based on the time of second death

$$T^* = \max(T_1^*, T_2^*) = \max(\min(T_1, b), \min(T_2, b)) = \min(\max(T_1, T_2), b))$$

and the indicator of censoring, δ^*, which is one if $T^* = b$ and zero otherwise. The likelihood function is based on two cases.

If the second death is uncensored, then we may assume $t < b$. In this case, we have $\delta^* = 0$ and

$$\text{Prob}(T^* < t|T_1>0,T_2 > 0) = \text{Prob}(T_1 < t,\ T_2 < t|T_1 > 0,T_2 > 0) = H_T(t, t).$$

Using the chain rule, we have $\partial H(t, t)/\partial t = H_1(t, t) + H_2(t, t)$. Thus, the contribution to the likelihood function is

$$\frac{\partial H_T(t,t)}{\partial t} = \frac{H_1(x+a+t,y+a+t)+H_2(x+a+t,y+a+t)-H_2(x+a,y+a+t)-H_1(x+a+t,y+a)}{1-H(x+a,\infty)-H(\infty,y+a)+H(x+a,y+a)}. \quad (19)$$

If the second death is censored, then we may assume $t \geq b$. In this case, we have $\delta^*=1$ and

$$\text{Prob}(T^* = b|T_1> 0,T_2 > 0) = 1 - H_T(b, b)\,. \quad (20)$$

Combining equations (19) and (20), we can express the logarithm of the likelihood function for a single observation as

$$\begin{aligned} &\ln L(x,y,t,\delta^*,a,b) \\ &= (1-\delta^*) \ln (H_1(x+a+t,y+a+t) + H_2(x+a+t,y+a+t) - H_2(x+a,y+a+t) - H_1(x+a+t,y+a)) \\ &\quad + \delta^* \ln(1-H(x+a,\infty) - H(\infty,y+a) + H(x+a,y+a+b) \\ &\quad + H(x+a+b,y+a) - H(x+a+b,y+a+b)) \\ &\quad - \ln(1-H(x+a,\infty) - H(\infty,y+a) + H(x+a,y+a)). \quad (21) \end{aligned}$$

The log-likelihood for the data can be calculated using the equation (21) expression in equation (13). Maximizing this log-likelihood function yields parameter estimates and standard errors.

Although this approach is technically sound, the maximum likelihood method yields unreliable parameter estimates using our data. Despite having nearly 15,000 contracts (and almost 30,000 lives) available for estimation, with this reduced data set we had only 229 deaths. Intuitively, most of the parameter information comes from the deaths, and we are attempting to estimate five parameters (in addition to the variance-covariance matrix). Thus, larger data sets, or longer observation periods, may be required to implement this method. This is interesting because data analysts generally do not consider 15,000 observations in a data set to be too small to employ likelihood methods.

ALTERNATIVE MODELS OF DEPENDENCE

This section investigates the robustness of the choice of Gompertz marginals and Frank's family of copulas by presenting some alternative choices.

Weibull Marginal Distribution

The Weibull distribution function can be expressed as

$$F(x) = 1 - \exp(-(x/m)^{m/\sigma}), \tag{22}$$

where m and σ are location and scale parameters. The mode of this distribution is $m(1-\sigma/m)^{\sigma/m}$, which is approximately 0.98m for m = 80 and $\sigma = 10$. Thus, because our estimated values of m and σ turn out to be close to 80 and 10, respectively, we may interpret m to be an approximate mode for this distribution, similar to the Gompertz. A more traditional expression for the Weibull is

$$F(x) = 1 - \exp(-Bx^c),$$

which is equivalent with the transformations $B = m^{-m/\sigma}$ and $c = m/\sigma$. Similar to the case of the Gompertz distribution, we find the parameterization in equation (22) to be more convenient for computational purposes.

Appendix D presents the Weibull parameter estimates and annuity ratios. We use maximum likelihood to estimate the parameter values. The annuity ratios are computed following the same format used above.

As suggested by Figures 1 and 2, the annuity ratios from the Weibull and Gompertz marginals are very similar. In most cases, the ratios differ by 0.01 or less, suggesting that our ratio values are not sensitive to the choice of marginal distributions. It is an interesting area of future research to measure the extent of the dependence of the ratios on the underlying marginal distributions.

Shock Models of Dependence

This section investigates the effects of the choice of the copula by considering an alternative family, the "common shock" models. As pointed out by Panjer (1994), the primary advantages of the common shock models are that they are easy to interpret and are computationally convenient.

To define this bivariate distribution, we begin with independent age-at-death random variables X and Y. We denote their marginal distribution functions by F_j so that $F_1(x) = \text{Prob}(X \le x)$ and $F_2(y) = \text{Prob}(Y \le y)$. We assume there exists an independent exponential random variable Z with parameter λ, that is, $\text{Prob}(Z \le t) = 1 - e^{-\lambda t}$. The bivariate time-until-death random vector is (T(x), T(y)), where T(x) =

$\min(X - x - a, Z)$ and $T(y) = \min(Y - y - a, Z)$. With this, we interpret Z to be a "shock" that is common to both lives. Our new underlying lifetime random variables are $X_c = T(x) + x + a = \min(X, Z + x + a)$ and $Y_c = T(y) + y + a = \min(Y, Z + y + a)$.

Under these assumptions, it is straightforward to compute the bivariate distribution. The survival distribution can be expressed as, for $t_1, t_2 \geq 0$,

$$\begin{aligned}
&\mathrm{Prob}(X_c > x + a + t_1, Y_c > y + a + t_2) \\
&= \mathrm{Prob}(\min(X, Z + x + a) > x + a + t_1, \min(Y, Z + y + a) > y + a + t_2) \\
&= \mathrm{Prob}(Z > \max(t_1, t_2))\ \mathrm{Prob}(X > x + a + t_1)\ \mathrm{Prob}(Y > y + a + t_2) \\
&= \exp(-\lambda \max(t_1, t_2))\ (1 - F_1(x + a + t_1))\ (1 - F_2(y + a + t_2)). \qquad (23)
\end{aligned}$$

Thus, the bivariate distribution function, for $t_1, t_2 \geq 0$, is

$$\begin{aligned}
&H(x + a + t_1, y + a + t_2) = \mathrm{Prob}(X_c \leq x + a + t_1, Y_c \leq y + a + t_2) \\
&= 1 - \exp(-\lambda t_1)(1 - F_1(x + a + t_1)) \\
&- \exp(-\lambda t_2)(1 - F_2(y + a + t_2)) \\
&+ \exp(-\lambda \max(t_1, t_2))(1 - F_1(x + a + t_1))(1 - F_2(y + a + t_2)). \qquad (24)
\end{aligned}$$

From equation (24), note that $H(x + a + t, \infty) = 1 - e^{-\lambda t} (1 - F_1(x + a + t)) \neq F_1(x + a + t)$. Thus, unlike the case of the copula bivariate function, the marginal common shock distributions are a function of the dependence parameter λ.

Parameters are estimated using the bivariate distribution function in equation (24) and the method of maximum likelihood. The only difference between the analysis conducted here and that based on Frank's copula is the likelihood of a common shock, which is given by the instantaneous probability

$$\frac{\partial}{\partial t} \mathrm{Prob}\ (T(x) = T(y) \leq t) = \lambda\, e^{-\lambda t} (1 - F_1(x + a + t))(1 - F_2(y + a + t)).$$

Table 7 presents the parameter estimates with the associated standard errors. The common shock location and scale parameters are close to the corresponding univariate estimates. All are within one standard error except m_1, which is only 1.3 ((86.66 – 86.38)/0.27) standard errors away. The measure of dependence, $\hat{\lambda}$, is more than five standard errors from zero, indicating strong statistical dependence. Although not significant, the location estimates are higher under the bivariate distribution than the univariate. Recall from Table 2 that the location estimates were significantly lower under the Gompertz/Frank model than under the univariate models.

The common shock model is intuitively appealing because bivariate conditional probabilities can be related easily to the marginals. To illustrate, recall from the section on effects of dependence on annuity values that ${}_k p_{\overline{xy}} = 1 - H_T(k, k)$ is the conditional probability that at least one life survives an additional k years. Using equations (23) and (24), straightforward calculations show that

$${}_k p_{\overline{xy}} = {}_k p_x + {}_k p_y - e^{\lambda k}\, {}_k p_x\, {}_k p_y. \qquad (25)$$

Here, ${}_kp_x = (1 - H(x + k, \infty))/(1 - H(x, \infty)) = e^{-\lambda k}(1 - F_1(x + k))/(1 - F_1(x))$ is the conditional probability that a life aged x survives an additional k years and similarly for ${}_kp_y$. Thus, for example, we may express our joint and last-survivor annuity as

$$\ddot{a}_{\overline{xy}} = \sum_{k=0}^{\infty} v^k({}_kp_x + {}_kp_y - e^{\lambda k}\, {}_kp_x\, {}_kp_y). \tag{26}$$

Table 7
Common Shock Parameter Estimates

Parameter	Bivariate Distribution Estimate	Bivariate Distribution Standard Error	Univariate Distribution Estimate	Univariate Distribution Standard Error
m_1	86.66	0.27	86.38	0.26
σ_1	9.89	0.37	9.83	0.37
m_2	92.69	0.64	92.17	0.59
σ_2	8.09	0.40	8.11	0.38
λ	0.00054	0.00010	Not Applicable	Not Applicable

An intuitively appealing feature of the common shock model is that the dependence parameter, λ, can be absorbed into the interest parameter, as follows. Define the pseudo conditional probabilities ${}_kp_x^* = (1-F_1(x + k))/(1 - F_1(x)) = e^{\lambda k}\, {}_kp_x$ and similarly for ${}_kp_y^*$. Using equations (25) and (26), we have

$$\ddot{a}_{\overline{xy}} = \sum_{k=0}^{\infty} e^{-(\delta+\lambda)k}({}_kp_x^* + {}_kp_y^* - {}_kp_x^*\, {}_kp_y^*) = {}^{I}\ddot{a}_{\overline{xy}} \ @\ (\delta+\lambda)\,. \tag{27}$$

Here, $\delta = \ln(1+i)$ is the so-called "force of interest," the symbol ${}^{I}\ddot{a}_{\overline{xy}}$ means calculate the annuity assuming independence using ${}_kp_x^*$ and ${}_kp_y^*$, and the notation @ $(\delta + \lambda)$ means at force of interest $\delta + \lambda$. Because ${}_kp_x^*$ and ${}_kp_y^*$ do not depend on λ, equation (27) shows that the joint and last-survivor annuity is a decreasing function of λ. In other words, the greater is the dependency, the smaller is the joint and last-survivor annuity.

To assess the real impact of dependency, Table 8 compares annuity values calculated under the common shock model to those calculated under independence. Unlike our copula models, annuity values are higher under the common shock model for most age combinations. This is interesting because, from equations (26) and (27), we expect annuity ratios less than one. On one hand, the increase in λ (from 0 to 0.00054) produces only a small decrease in annuity values. On the other hand, the larger location parameters mean that the individual forces of mortality are lower under the case of dependence. Lower forces of mortality result in larger annuity values.

Thus, despite the computational simplicity, the common shock model does not seem to provide the same pleasing intuitive results as the copula model. Further, the common shock model does not seem to fit the data as well as the Gompertz/Frank model. When estimating the models, the log-likelihood associated with the Gompertz/Frank model, –9,977, was larger than the log-likelihood associated with the common shock model, –10,078. The two models are not hierarchical and thus traditional likelihood ratio tests are not applicable. However, this does provide additional evidence that the Frank/Gompertz model provides a better fit to the data.

Table 8
Ratios of Dependent to Independent Joint and Last Annuity Values Based on the Common Shock Model (Five Percent Interest Is Assumed)

	Female Age						
Male Age	50	55	60	65	70	75	80
50	1.00	1.00	1.00	1.00	1.00	1.00	1.00
55	1.00	1.00	1.00	1.00	1.00	1.00	1.00
60	1.00	1.00	1.00	1.00	1.00	1.00	1.00
65	1.00	1.00	1.00	1.00	1.00	1.00	1.00
70	1.00	1.00	1.00	1.00	1.01	1.01	1.01
75	1.00	1.00	1.00	1.00	1.01	1.01	1.01
80	1.00	1.00	1.00	1.01	1.01	1.01	1.01

Finally, we note that the common shock does not seem to take into account all the dependencies that we observe in the data. Of the 229 pairs of deaths within our observation period, 29 occurred with one day and hence were "simultaneous." The data also revealed proximity of other deaths:

Number of Pairs of Deaths Within:
1 Day = 29
5 Days = 63
10 Days = 70
20 Days = 85
30 Days = 86.

Thus, there appears to be some dependency of lives that the common shock model does not detect. Of course, one can always alter the definition of the time scale to redefine what "simultaneous" means. An advantage of the copula models is that this is not necessary because the dependency is assessed in a smooth fashion.

CONCLUSION

This article discusses methods for estimating the probability of joint survival using insurance data. Although our focus is on annuity valuation, our methods can be applied easily to other types of insurance products. For example, Bragg (1994) discusses the growing importance of last-survivor, or "second-to-die," life insurance.

Throughout the article, our illustrations focus on valuing level annuities using fixed interest rates. However, with additional complexity, the methods also can be applied to variable products and can be used for valuing level annuities using a model from financial economics. This is because economic models assume that probabilities are exogenous inputs into an economic system. Thus, our dependent mortality models could be used to determine probabilities that are inputs to an economic model.

Because of the heavy truncation and censoring of our data, our models of the bivariate distribution are completely parametric. Maguluri (1993) provides some theoretical results on the efficiency of using a parametric family for the copula, such as Frank's family, with nonparametric distributions for the marginals, such as Kaplan-Meier. It would be interesting to fit data using a parametric copula and standard insurance industry tables, such as the 1983 Individual Annuity Table. We leave this for future research.

APPENDIX A
DERIVATION OF RESULTS NEEDED TO EVALUATE THE LIKELIHOOD USING FRANK'S COPULA AND GOMPERTZ MARGINALS

This appendix derives the results needed to evaluate the log-likelihood in equation (13) or, in particular, equation (12), in the case where we assume the Frank's copula function as given in equation (3) and Gompertz marginals as given in equation (5).

Consider the bivariate age-at-death random vector (X,Y) whose distribution function is given in equation (2) and where the copula is given in equation (3). Deriving the first partial derivatives of C, we have

$$C_1(u,v)=\frac{\partial}{\partial u}C(u,v)=\frac{e^{\alpha u}(e^{\alpha v}-1)}{e^{\alpha}-1+(e^{\alpha u}-1)(e^{\alpha v}-1)} \tag{A1}$$

and

$$C_2(u,v)=\frac{\partial}{\partial v}C(u,v)=\frac{e^{\alpha v}(e^{\alpha u}-1)}{e^{\alpha}-1+(e^{\alpha u}-1)(e^{\alpha v}-1)}. \tag{A2}$$

The second partial derivative of C is given as

$$C_{12}(u,v)=\frac{\partial^2}{\partial u\partial v}C(u,v)=\frac{\alpha(e^{\alpha}-1)e^{\alpha(u+v)}}{[(e^{\alpha}-1)+(e^{\alpha u}-1)(e^{\alpha v}-1)]^2}. \tag{A3}$$

We denote the density functions of X and Y as f_1 and f_2, respectively. In other words, we have

$$f_1(x)=\frac{\partial}{\partial x}F_1(x) \text{ and } f_2(y)=\frac{\partial}{\partial y}F_2(y).$$

Using the chain rule of differentiation, we then have the first and second partial derivatives of the distribution function H(x,y):

$$H_1(x,y)=\frac{\partial}{\partial x}H(x,y)=f_1(x)C_1(F_1(x),F_2(y)) \tag{A4}$$

and

$$H_2(x,y)=\frac{\partial}{\partial y}H(x,y)=f_2(y)C_2(F_1(x),F_2(y)) \tag{A5}$$

and

$$h(x,y)=\frac{\partial^2}{\partial x\partial y}H(x,y)=f_1(x)f_2(y)C_{12}(F_1(x),F_2(y)). \tag{A6}$$

If we suppose that the marginals follow a Gompertz distribution as in equation (1), then we have the following density functions:

$$f_j(x)=\frac{1}{\sigma_j}e^{(x-m_j)/\sigma_j}\exp[e^{-m_j/\sigma_j}(1-e^{x/\sigma_j})],\ j=1,2. \tag{A7}$$

Equations (A1) to (A3) and (A7) are then used to evaluate equations (A4) to (A6). Equations (A4), (A5), and (A6) are used in maximizing the log-likelihood as expressed in equation (12).

Note that the parameter α is not necessarily a standard measure of association. However, we can express the more familiar Spearman's correlation coefficient as a function of α as follows:

$$\rho(\alpha)=12\int_0^1\int_0^1 C(u,v)\,du\,dv-3.$$

If C is the Frank's copula, then we have $\rho(\alpha)$ as expressed in equation (4).

APPENDIX B
DEVELOPMENT OF THE LIKELIHOOD EQUATION FOR UNDERREPORTING OF DEATH WITHIN THE GUARANTEE PERIOD

To determine the conditional distribution of (T_1^{**},T_2^{**}), we consider nine cases of (t_1, t_2). The general expression for the conditional distribution function for the guaranteed case is

$$H_g(t_1,t_2) = \text{Prob}\ (T_1^{**} \le t_1,\ T_2^{**} \le t_2|T_1 > 0,\ T_2 > 0)$$

$$= \text{Prob}\ (\min(\max(T_1,g),b) \le t_1,\ \min(\max(T_2,g),b) \le t_2|T_1 > 0,\ T_2 > 0). \tag{B1}$$

Case 1

If both lifetimes are right-censored, then we have $t_1 \geq b$, $t_2 \geq b$. Hence, we have $\delta_{1g} = \delta_{2g} = 0$, $\delta_{1b} = \delta_{2b} = 1$ and

$$\begin{aligned} H_g(t_1,t_2) &= \text{Prob}\,(T_1^{**} = b, T_2^{**} = b \mid T_1 > 0, T_2 > 0) \\ &= \text{Prob}\,(T_1 \geq b, T_2 \geq b \mid T_1 > 0, T_2 > 0) \\ &= 1 - H_T(\infty,b) - H_T(b,\infty) + H_T(b,b). \end{aligned} \tag{B2}$$

Case 2

If the first lifetime is uncensored and the second lifetime is right-censored, then we have $g < t_1 < b$, $t_2 \geq b$. Hence, we have $\delta_{1g} = \delta_{2g} = \delta_{1b} = 0$, $\delta_{2b} = 1$ and

$$H_g(t_1,t_2) = \text{Prob}\,(T_1 \leq t_1, T_2 \geq b \mid T_1 > 0, T_2 > 0) = H_T(t_1,\infty) - H_T(t_1,b). \tag{B3}$$

Case 3

If the first lifetime is left-censored and the second lifetime is right-censored, then we have $t_1 \leq g$, $t_2 \geq b$. Hence, we have $\delta_{1g} = \delta_{2b} = 1$, $\delta_{1b} = \delta_{2g} = 0$ and

$$H_g(t_1,t_2) = H_T(g,\infty) - H_T(g,b). \tag{B4}$$

Case 4

If the first lifetime is right-censored and the second lifetime is uncensored, then we have $t_1 \geq b$, $g < t_2 < b$. Hence, we have $\delta_{1b} = 1$, $\delta_{1g} = \delta_{2g} = \delta_{2b} = 0$ and

$$H_g(t_1,t_2) = H_T(\infty,t_2) - H_T(b,t_2). \tag{B5}$$

Case 5

If both lifetimes are uncensored, then we have $g < t_1 < b$, $g < t_2 < b$. Hence, we have $\delta_{1g} = \delta_{1b} = 0$, $\delta_{2g} = \delta_{2b} = 0$ and

$$H_g(t_1,t_2) = H_T(t_1,t_2). \tag{B6}$$

Case 6

If the first lifetime is left-censored and the second lifetime is uncensored, then we have $t_1 \leq g$, $g < t_2 < b$. Hence, we have $\delta_{1g} = 1$, $\delta_{2g} = \delta_{1b} = \delta_{2b} = 0$ and

$$H_g(t_1,t_2) = \text{Prob}(T_1 \leq g, T_2 \leq t_2) = H_T(g,t_2). \tag{B7}$$

Case 7

If the first lifetime is right-censored and the second lifetime is left-censored, then we have $t_1 \geq b$, $t_2 \leq g$. Hence, we have $\delta_{1b} = \delta_{2g} = 1$, $\delta_{1g} = \delta_{2b} = 0$ and

$$H_g(t_1,t_2) = \text{Prob}(T_1 \geq b, T_2 \leq g) = H_T(\infty,g) - H_T(b,g). \tag{B8}$$

Case 8

If the first lifetime is uncensored and the second lifetime is left-censored, then we have $g < t_1 < b$, $t_2 \le g$. Hence, we have $\delta_{1g} = \delta_{1b} = \delta_{2b} = 0$, $\delta_{2g} = 1$ and

$$H_g(t_1,t_2) = H_T(t_1,g). \tag{B9}$$

Case 9

Finally, if both lifetimes are left-censored, we have $t_1 \le g$ and $t_2 \le g$. Hence, we have $\delta_{1g} = \delta_{2g} = 1$, $\delta_{1b} = \delta_{2b} = 0$ and

$$H_g(t_1,t_2) = H_T(g,g). \tag{B10}$$

Recall from equations (8), (9), and (10) that

$$\frac{\partial H_T(t_1,t_2)}{\partial t_1} = \frac{H_1(x+a+t_1, y+a+t_2) - H_1(x+a+t_1, y+a)}{1 - H(x+a,\infty) - H(\infty, y+a) + H(x+a, y+a)},$$

$$\frac{\partial H_T(t_1,t_2)}{\partial t_2} = \frac{H_2(x+a+t_1, y+a+t_2) - H_2(x+a, y+a+t_2)}{1 - H(x+a,\infty) - H(\infty, y+a) + H(x+a, y+a)},$$

and

$$\frac{\partial^2 H_T(t_1,t_2)}{\partial t_1 \partial t_2} = \frac{h(x+a+t_1, y+a+t_2)}{1 - H(x+a,\infty) - H(\infty, y+a) + H(x+a, y+a)}.$$

Combining equations (B2) to (B10), we then have the contribution of a single observation to the log-likelihood as follows:

$$\begin{aligned}
&\log L_g(x,y,a,b,t_1,t_2,\delta_{1g},\delta_{2g},\delta_{1b},\delta_{2b},g)\\
&= \delta_{1b}\delta_{2b}\log[1-H(\infty,y+a+b) - H(x+a+b,\infty) + H(x+a+b,y+a+b)]\\
&+ (1-\delta_{1g}-\delta_{1b})\delta_{2b}\log[H_1(x+a+t_1,\infty) - H_1(x+a+t_1,y+a+b)]\\
&+ \delta_{1g}\delta_{2b}\log[H(x+a+g,\infty) - H(x+a,\infty) - H(x+a+g,y+a+b) + H(x+a,y+a+b)]\\
&+ \delta_{1b}(1-\delta_{2g}-\delta_{2b})\log[H_2(\infty,y+a+t_2) - H_2(x+a+b,y+a+t_2)]\\
&+ (1-\delta_{1b}-\delta_{1g})(1-\delta_{2g}-\delta_{2b})\log[h(x+a+t_1,y+a+t_2)]\\
&+ \delta_{1g}(1-\delta_{2g}-\delta_{2b})\log[H_2(x+a+g,y+a+t_2) - H_2(x+a,y+a+t_2)]\\
&+ \delta_{1b}\delta_{2g}\log[H(\infty,y+a+g) - H(\infty,y+a) - H(x+a+b,y+a+g) + H(x+a+b,y+a)]\\
&+ (1-\delta_{1b}-\delta_{1g})\,\delta_{2g}\log[H_1(x+a+t_1,y+a+g) - H_1(x+a+t_1,y+a)]\\
&+ \delta_{1g}\delta_{2g}\log[H(x+a+g,y+a+g) - H(x+a+g,y+a) - H(x+a,y+a+g) + H(x+a,y+a)]\\
&-\log[1-H(x+a,\infty) - H(\infty,y+a) + H(x+a,y+a)].
\end{aligned}$$

APPENDIX C
CALCULATION OF THE GRADIENT OF $\ddot{a}_{\overline{xy}}(\theta)$

Let $\theta = (m_1, m_2, \sigma_1, \sigma_2, \alpha)' = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5)'$. Thus, from equation (18) for i = 1,2,... 5, we have

$$\frac{\partial}{\partial\theta_i}\ddot{a}_{\overline{xy}} = \sum_{k=0}^{\infty} v^k \frac{\partial}{\partial\theta_i}\,{}_k p_{\overline{xy}}$$

$$= \sum_{k=0}^{\infty} \frac{v^k}{[1-H(x,\infty)-H(\infty,y)+H(x,y)]^2}\left\{\begin{array}{l}(H(x+k,y+k)-H(x+k,y)-H(x,y+k)+H(x,y))\times \\ \left(\frac{\partial}{\partial q_i}(H(x,y)-F_1(x)-F_2(y))\right)- \\ (1-H(x,\infty)-H(\infty,y)+H(x,y))\times \\ \left(\frac{\partial}{\partial q_i}(H(x+k,y+k)-H(x,y+k)-H(x+k,y)+H(x,y))\right)\end{array}\right\}. \quad (C1)$$

Using equations (1) and (A7), for i = 1, 2, we have

$$\begin{aligned}\frac{\partial}{\partial m_i}F_i(x) &= -\exp[e^{-m_i/\sigma_i}(1-e^{x/\sigma_i})]\cdot\frac{\partial}{\partial m_i}(e^{-m_i/\sigma_i}(1-e^{x/\sigma_i})) \\ &= -(1-F_i(x))(1-e^{x/\sigma_i})\frac{\partial}{\partial m_i}e^{-m_i/\sigma_i} \\ &= \frac{1}{\sigma_i}(1-F_i(x))(1-e^{x/\sigma_i})e^{-m_i/\sigma_i} \\ &= f_i(0)-f_i(x) \qquad (C2)\end{aligned}$$

and

$$\begin{aligned}\frac{\partial}{\partial\sigma_i}F_i(x) &= -(1-F_i(x))\frac{\partial}{\partial\sigma_i}(e^{-m_i/\sigma_i}-e^{(x-m_i)/\sigma_i}) \\ &= -(1-F_i(x))(m_i e^{-m_i/\sigma_i}+(x-m_i)e^{(x-m_i)/\sigma_i})/\sigma_i^2 \\ &= \frac{-1}{\sigma_i}(m_i f_i(0)+(x-m_i)f_i(x)) \\ &= \frac{1}{\sigma_i}((m_i-x)f_i(x)-m_i f_i(0)). \qquad (C3)\end{aligned}$$

Equations (A4), (A5), (C2), and (C3) and the chain rule yield

$$\begin{aligned}\frac{\partial}{\partial m_1}H(x,y) = \frac{\partial}{\partial m_1}C(F_1(x),F_2(y)) &= C_1(F_1(x),F_2(y))\frac{\partial}{\partial m_1}F_1(x) \\ &= H_1(x,y)(f_1(0)-f_1(x))/f_1(x) \qquad (C4)\end{aligned}$$

$$\frac{\partial}{\partial m_2}H(x,y) = H_2(x,y)(f_2(0)-f_2(y))/f_2(y) \qquad (C5)$$

and

$$\frac{\partial}{\partial \sigma_1}H(x,y)=C_1(F_1(x),F_2(y))\frac{\partial}{\partial \sigma_1}F_1(x)$$
$$= H_1(x,y)((m_1-x)f_1(x)-m_1f_1(0))/(\sigma_1 f_1(x)) \qquad (C6)$$
$$\frac{\partial}{\partial \sigma_2}H(x,y)=H_2(x,y)((m_2-y)f_2(y)-m_2f_2(0))/(\sigma_2 f_2(y)). \qquad (C7)$$

To calculate $(\partial/\partial\alpha)H(x,y)$, we use equation (3) and let $K(\alpha) = K = \exp(\alpha C(u,v)) = 1 + \frac{(e^{\alpha u}-1)(e^{\alpha v}-1)}{(e^{\alpha}-1)}$. Thus,

$$\frac{\partial}{\partial\alpha}K(\alpha)=\frac{(e^{\alpha}-1)[ue^{\alpha u}(e^{\alpha v}-1)+ve^{\alpha v}(e^{\alpha u}-1)]-(e^{\alpha u}-1)(e^{\alpha v}-1)e^{\alpha}}{(e^{\alpha}-1)^2}.$$

Making the substitutions

$$e^{\alpha u}(e^{\alpha v}-1) = C_1(u,v)(e^{\alpha}-1)K$$

and

$$e^{\alpha v}(e^{\alpha u}-1) = C_2(u,v)(e^{\alpha}-1)K,$$

we have $\frac{\partial}{\partial\alpha} K(\alpha) = uKC_1(u,v) + vKC_2(u,v) - \frac{e^{\alpha}}{(e^{\alpha}-1)}(K-1)$.

Define $C_\alpha(u,v) = \frac{\partial}{\partial\alpha}C(u,v)$. Thus, we have

$$C_\alpha(u,v)=\frac{\partial}{\partial\alpha}\left(\frac{\log(K)}{\alpha}\right)=\frac{\alpha\frac{1}{K}\frac{\partial K}{\partial\alpha}-\log(K)}{\alpha^2}$$
$$=\frac{1}{\alpha}\frac{1}{K}\left(uKC_1(u,v)+vKC_2(u,v)+\frac{e^{\alpha}}{(e^{\alpha}-1)}(1-K)\right)-\frac{1}{\alpha^2}\alpha C(u,v).$$

Rearranging terms, we then have

$$C_\alpha(u,v)=\frac{1}{\alpha}\left\{\frac{e^{\alpha}(e^{-\alpha C(u,v)}-1)}{(e^{\alpha}-1)}+(uC_1(u,v)+vC_2(u,v)-C(u,v))\right\}.$$

Thus,

$$\frac{\partial}{\partial\alpha}H(x,y)=\frac{\partial}{\partial\alpha}C(u,v)\Big|_{u=F_1(x),v=F_2(y)}=C_\alpha(F_1(x),F_2(y)). \qquad (C8)$$

To get the gradient of $\ddot{a}_{\overline{xy}}(\theta)$, plug equations (C4) through (C8) into equation (C1).

APPENDIX D
WEIBULL ESTIMATION RESULTS

Table D1
Weibull Parameter Estimates

	Bivariate Distribution		Univariate Distribution	
Parameter	Estimate	Standard Error	Estimate	Standard Error
m_1	86.22	0.27	86.73	0.28
σ_1	10.16	0.39	10.12	0.37
m_2	89.91	0.55	93.00	0.69
σ_2	8.75	0.40	9.26	0.47
α	–3.354	0.338	Not Applicable	Not Applicable

Table D2
Ratios of Dependent to Independent Joint and Last Annuity Values Based on Weibull Marginal Distributions (Five Percent Interest Is Assumed)

	Female Age						
Male Age	50	55	60	65	70	75	80
50	0.97	0.96	0.96	0.97	0.98	1.00	1.02
55	0.97	0.96	0.95	0.95	0.97	0.99	1.02
60	0.97	0.96	0.95	0.94	0.95	0.97	1.01
65	0.98	0.97	0.95	0.94	0.94	0.95	0.99
70	0.99	0.98	0.96	0.95	0.93	0.93	0.96
75	1.00	0.99	0.99	0.97	0.95	0.94	0.94
80	1.01	1.01	1.01	1.00	0.99	0.96	0.94

APPENDIX E
SPEARMAN'S TEST OF INDEPENDENCE

This appendix shows how to apply Spearman's test of independence on our data, assuming that the marginal distributions are known. This last assumption is reasonable because extensive information is available on the law of mortality for individual lives. The following technique is a quick nonparametric way of identifying and measuring the dependence. The results from this method must be used with caution because we are assuming that the marginals are known.

Suppose we observe the time of deaths $(T_{1,k}, T_{2,k})$ for $k = 1, 2, \ldots, n$, where $T_{i,k}$ has a known continuous distribution function $G_{i,k}(t)$, $t > 0$, $i = 1, 2$. In our case, these distributions will not be identical. Consider the uniform random variables $U_{i,k} \equiv G_{i,k}(T_{i,k})$. Assume that the pairs $(U_{1,k}, U_{2,k})$ for $k = 1, 2, \ldots, n$ are independent and identically distributed with a common copula $C(u,v)$. The assumption of a common copula allows us to calculate Spearman's sample correlation coefficient and use it to test independence, that is, $C(u,v) = uv$. Let $R_{i,k}$ denote the rank of $U_{i,k}$; then Spearman's correlation is

$$\hat{\rho} = \frac{\sum_{k=1}^{n}[R_{1,k} - (n+1)/2][R_{2,k} - (n+1)/2]}{n(n^2 - 1)/12}.$$

An estimate of the asymptotic variance of this statistic is $(n-1)^{-1}$. So we reject the null hypothesis of independence at a 5 percent level if $|\hat{\rho}| > 1.96\,(n-1)^{-1/2}$.

Let's apply this technique to our data where $n = 229$ policies had both annuitants die during the observation period. Let $F_i(x)$ denote a Gompertz distribution as defined in equation (1) with parameters m_i, σ_i. If $i = 1$, then these parameters refer to a male life, and, if $i = 2$, then they refer to a female life. Consulting Table 2, we let $m_1 = 86.38$, $\sigma_1 = 9.83$ and $m_2 = 92.17$, $\sigma_2 = 8.11$. We find that the estimate of Spearman's correlation coefficient does not change very much when other reasonable parameter values are used. In our case, $G_{i,k}$ is the distribution of the time of death, given that the death occurs during the observation period. Let $x_{i,k} + a_k$ denote the age at the start of the observation period, and let $x_{i,k} + a_k + b_k$ denote the age at the end of the observation period; then

$$G_{i,k}(t) = \frac{F_i(x_{i,k} + t) - F_i(x_{i,k} + a_k)}{F_i(x_{i,k} + a_k + b_k) - (F_i(x_{i,k} + a_k)}.$$

The sample correlation is $\hat{\rho} = 0.414$ and a 95 percent confidence interval is (0.282, 0.547). Our estimate of ρ, given in the section on estimation results, is equal to 0.49, which lies within this confidence interval.

REFERENCES

American Council of Life Insurance, 1994, *1994 Life Insurance Fact Book (*Washington, D.C.: ACLI).

Bowers, N. L., H. U. Gerber, J. C. Hickman, D. A. Jones, and C. J. Nesbitt, 1986, *Actuarial Mathematics* (Schaumburg, Ill.: Society of Actuaries).

Bragg, J., 1994, The Last-Survivor Phenomenon, *The Actuary*, 28(1): 5.

Carriere, J., 1994, An Investigation of the Gompertz Law of Mortality, *Actuarial Research Clearing House,* 2.

Cox, D. R. and D. Oakes, 1990, *Analysis of Survival Data* (New York: Chapman and Hall).

Elandt-Johnson, R. C. and N. L. Johnson, 1980, *Survival Models and Data Analysis* (New York: John Wiley).

Frank, M. J., 1979, On the Simultaneous Associativity of $F(x,y)$ and $x+y-F(x,y)$, *Aequationes Math,* 19: 194–226.

Genest, C., 1987, Frank's Family of Bivariate Distributions, *Biometrika,* 74: 549–555.

Genest, C. and J. McKay, 1986, The Joy of Copulas: Bivariate Distributions with Uniform Marginals, *American Statistician,* 40: 280–283.

Hougaard, P., B. Harvald, and N. V. Holm, 1992, Measuring the Similarities Between the Lifetimes of Adult Danish Twins Born Between 1881–1930, *Journal of the American Statistical Association,* 87: 17–24.

Jagger, C. and C. J. Sutton, 1991, Death after Marital Bereavement—Is the Risk Increased? *Statistics in Medicine,* 10: 395–404.

Maguluri, G., 1993, Semiparametric Estimation of Association in a Bivariate Survival Function, *Annals of Statistics,* 21: 1648–1662.

Marshall, A. W. and I. Olkin, 1967, A Multivariate Exponential Distribution, *Journal of the American Statistical Association,* 62: 30–44.

Marshall, A. W. and I. Olkin, 1988, Families of Multivariate Distributions, *Journal of the American Statistical Association,* 83: 834–841.

Nelsen, R. B., 1986, Properties of a One-Parameter Family of Distributions with Specified Marginals, *Communications in Statistics—Theory and Methods,* A15: 3277–3285.

Oakes, D., 1989, Bivariate Survival Models Induced by Frailties, *Journal of the American Statistical Association,* 84: 487–493.

Panjer, H., 1994, Second-to-Die with Possibility of Simultaneous Death, *Product Development News*, 36 (Schaumburg, Ill.: Society of Actuaries).

Parkes, C. M., B. Benjamin, and R. G. Fitzgerald, 1969, Broken Heart: A Statistical Study of Increased Mortality among Widowers, *British Medical Journal,* i: 740–743.

Pruitt, R. C., 1993, Identifiability of Bivariate Survival Curves from Censored Data, *Journal of the American Statistical Association,* 88: 573–579.

Serfling, R., 1980, *Approximation Theorems in Mathematical Statistics* (New York: John Wiley).

Sklar, A., 1959, Fonctions de répartitions à *n* dimensions et leurs marges, *Inst. Statist. Univ. Paris Publ.,* 8: 229–231.

Ward, A., 1976, Mortality of Bereavement, *British Medical Journal,*, i: 700–702.