DISCUSSIONS OF PAPERS ALREADY PUBLISHED

"A Quantitative Comparison of Stochastic Mortality Models Using Data from England and Wales and the United States," Andrew J. G. Cairns, David Blake, Kevin Dowd, Guy D. Coughlan, David Epstein, Alen Ong, and Igor Balevich, Vol. 13, No. 1, 2009

KAILIANG CHEN*, JIA LIAO†, XIAOYU SHANG‡, AND JOHNNY SIU-HANG LI§

It is our pleasure to congratulate Professor Andrew Cairns and his coauthors on this timely paper, which provides valuable information to users of stochastic mortality models. The findings from this paper indicate that the CBD class of models performs well, particularly when the robustness of parameter estimates is taken into account. The objective of this discussion is to expand the CBD class of models by considering two transformations, log and probit, in addition to the logit transformation that is used in the original specification. The new variants are compared against the original CBD model by some quantitative means.

D.1. EXTENDING MODEL M5

Let us consider the original CBD model (Model M5 in the paper):

$$\text{logit}(q(t, x)) = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x}),$$

where $q(t, x)$ is the probability that an individual aged exactly x at time t will die between t and $t + 1$, and $\bar{x}$ is the mean age in the sample range. This model, in the original work of Cairns et al. (2006), is written as

$$\text{logit}(q(t, x)) = A_1(t) + A_2(t)x,$$

where $A_1(t) = \kappa_t^{(1)} - \kappa_t^{(2)}\bar{x}$ and $A_2(t) = \kappa_t^{(2)}$. The success of the CBD model lies in the following two properties when the logit transformation is applied to $q(t, x)$:

1. The transformed death probability varies linearly with age
2. The patterns of $A_1(t)$ and $A_2(t)$ are highly linear.

It is interesting to find whether the properties above can be preserved when transformations other than logit are used. In this discussion we study two additional transformations:

1. log (natural-log):

$$\log(q(t, x)) = A_1(t) + A_2(t)x$$

---

* Kailiang (Kevin) Chen is an MMath candidate in the Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1, Kailiangc@hotmail.com.

† Jia Liao, PhD, is an MMath candidate in the Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1, j3liao@uwaterloo.ca.

‡ Xiaoyu (Mavis) Shang, ASA, MMath, is an Actuarial Associate in RBC Insurance, Mississauga, Ontario, Canada, L5A 3A1, mavis.shang@gmail.com.

§ Johnny Siu-Hang Li, Phd, FSA, is an Assistant Professor in the Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1, shli@unwaterloo.ca.

2. probit:

$$\Phi^{-1}(q(t, x)) = A_1(t) + A_2(t)x,$$

where Φ is the distribution function for a standard normal random variable.

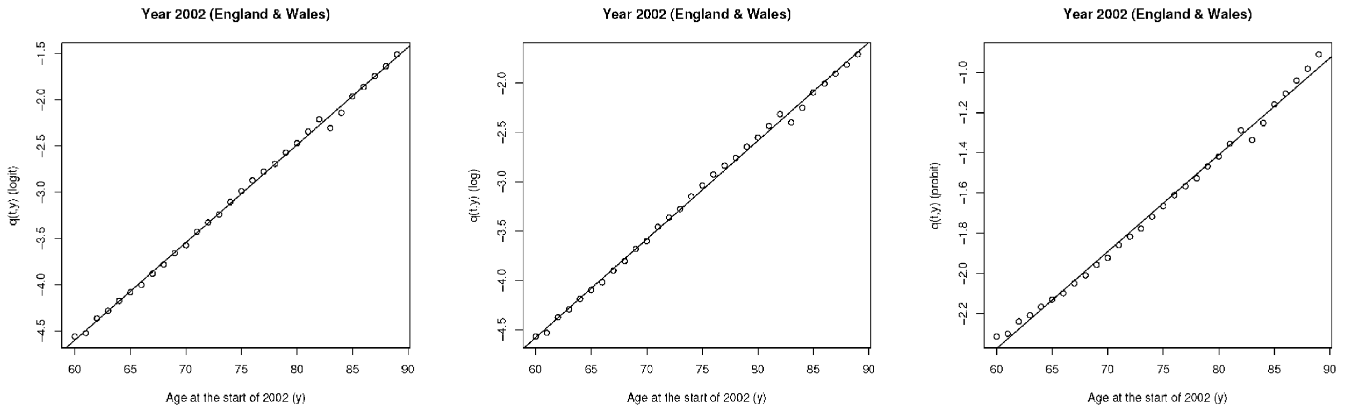
Note that under the log transformation, $q(t, x)$ is not bounded above by 1. However, given the trends of $A_1(t)$ and $A_2(t)$, it is not likely that the model will yield a projection of $q(t, x)$ that is greater than 1.

D.2. MODEL FITTING

We fit the original CBD model and our proposed variants to data from the United States and England and Wales. Following the paper that we are discussing, we consider the sample period of 1968–2004 for the United States and 1961–2004 for England and Wales. The data are obtained from the Human Mortality Database (2009).

Figure D.1

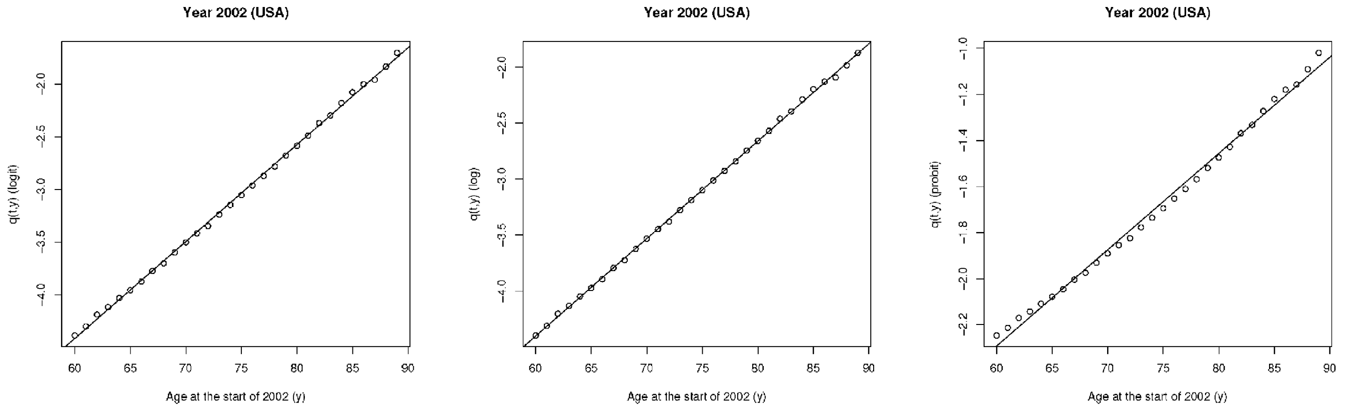
Ungraduated *q*(*t*, *x*) and Fitted Mortality Curves, England and Wales, 2002



(a) Logit (b) Log (c) Probit

Figure D.2

Ungraduated *q*(*t*, *x*) and Fitted Mortality Curves, United States, 2002



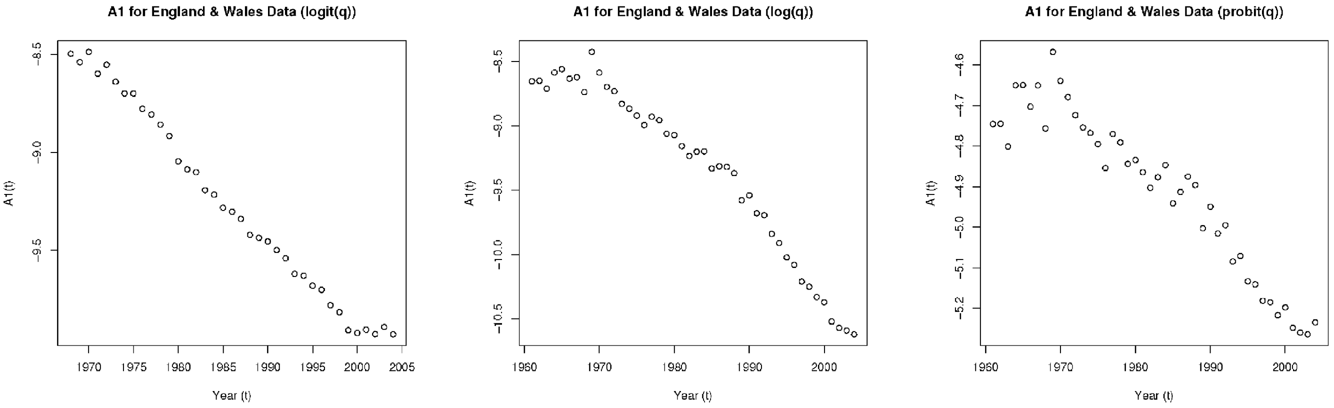
(a) Logit (b) Log (c) Probit

In Figures D.1 and D.2 we plot the transformed crude death probabilities for 2002 (the empty circles) against age. We observe that, under all three transformations, the relationships are highly linear. By comparing the fitted mortality curves (the solid lines) and the crude probabilities, we can tell all three transformations give a reasonable goodness-of-fit.

We then compare the stochastic factors $A_1(t)$ and $A_2(t)$ resulting from different transformations. From Figures D.3–D.6 we observe a clear linearity in the stochastic factors derived from all three transformations. Such a linearity facilitates the extrapolation of $A_1(t)$ and $A_2(t)$ using a linear bivariate time-series model.

Figure D.3

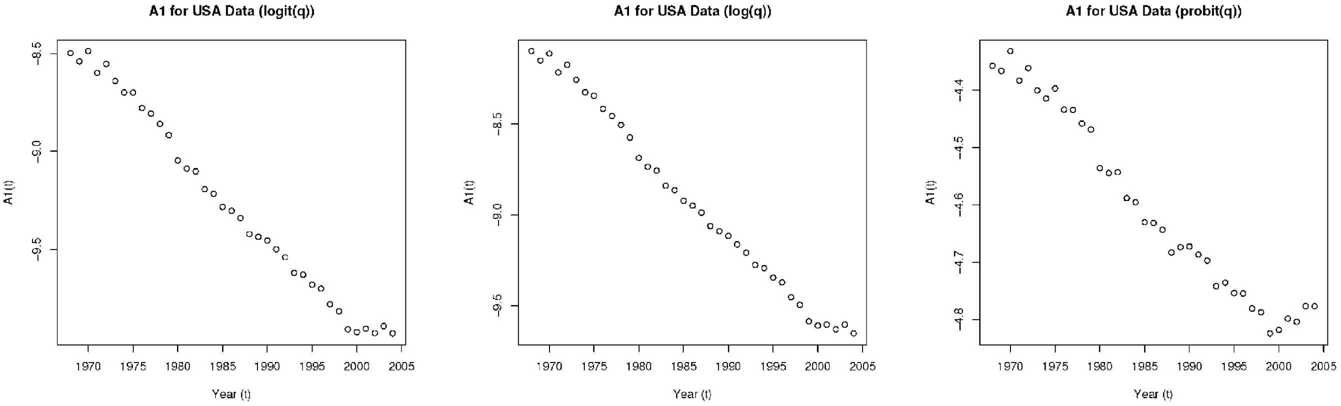
Estimated Values of $A_1(t)$, England and Wales



(a) Logit (b) Log (c) Probit

Figure D.4

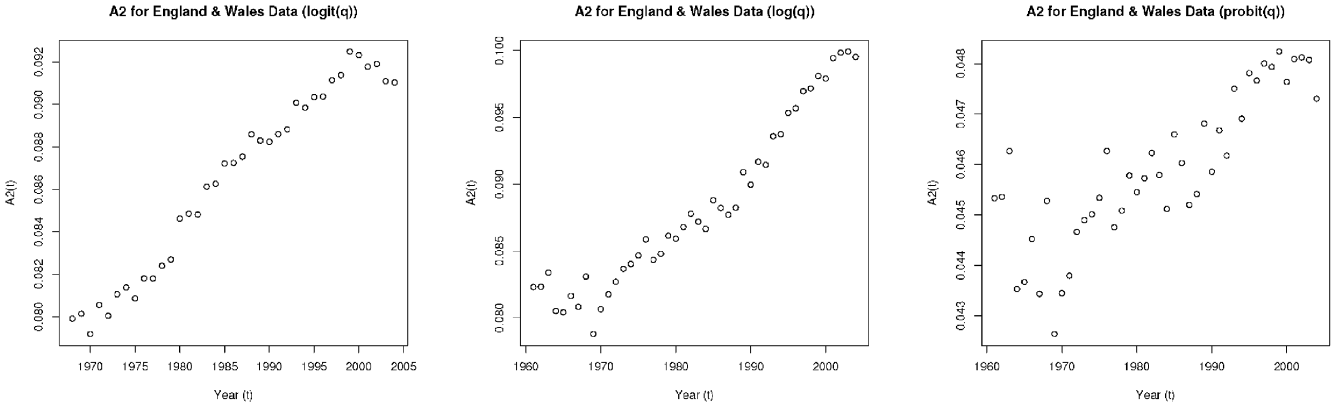
Estimated Values of $A_1(t)$, United States



(a) Logit (b) Log (c) Probit

Figure D.5

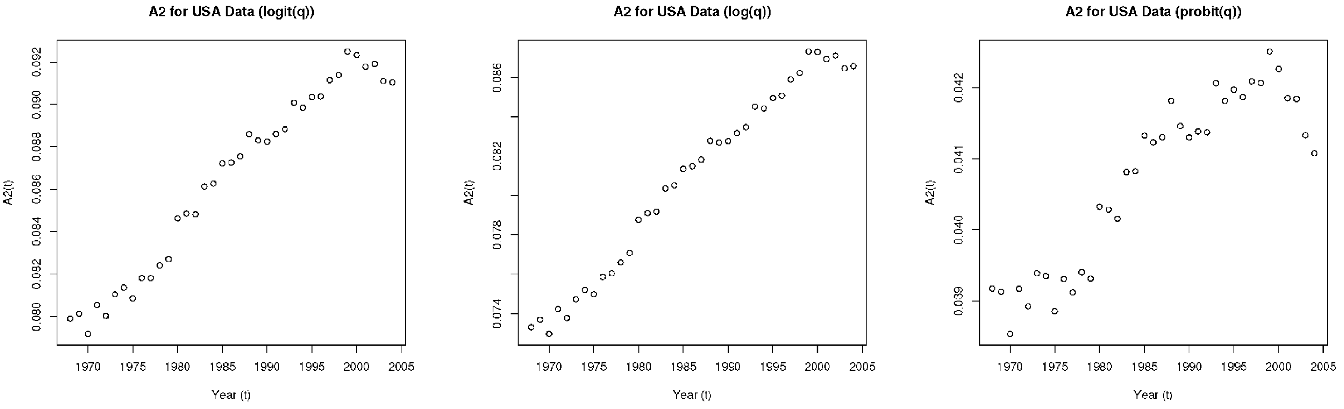
Estimated Values of $A_2(t)$, England and Wales



(a) Logit (b) Log (c) Probit

Figure D.6

Estimated Values of $A_2(t)$, United States



(a) Logit (b) Log (c) Probit

D.3. EVALUATING THE EXTENSIONS

We now evaluate the performance of our proposed extensions. First, we compare the sum of squared errors (SSE), which is defined by

$$\text{SSE} = \sum_{x,t} (q(x, t) - \hat{q}(x, t))^2,$$

where $\hat{q}(x, t)$ is the fitted value of the death probability at age x and time t, and the summation is taken over the entire sample period and sample age range. A smaller SSE value indicates a better fit to the data. In Table D.1 we display the SSE values derived from the three transformations. For England and Wales, the SSE from the logit model is significantly less than those from the other two models. However, the log model gives a slightly better fit to the data from the U.S. population than the logit model.

Table D.1

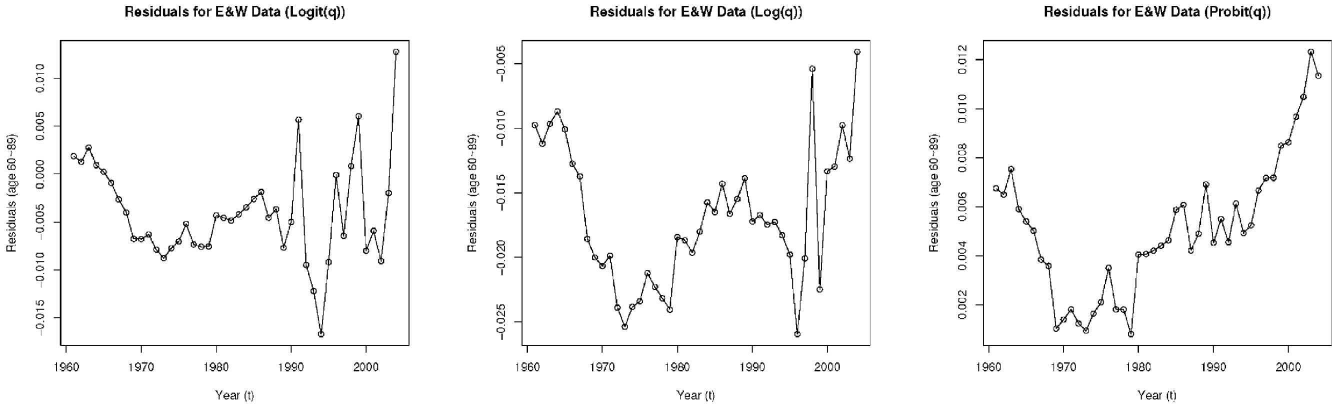
Sum of Squared Errors

	Logit	Log	Probit
England and Wales	0.008642	0.028757	0.017263
United States	0.004842	0.002941	0.019830

Next we examine the residuals. In particular, we consider the aggregate residuals, $\Sigma_x(q(x,\, t) - \hat{q}(x,\, t))$, in each year during the sample period.[1] A systematic departure of the residuals from zero indicates a bad fit to the data. We show in Figures D.7 and D.8 the patterns of the aggregate residuals. It is clear that the probit transformation underperforms, because it results in residuals that are sys-

Figure D.7

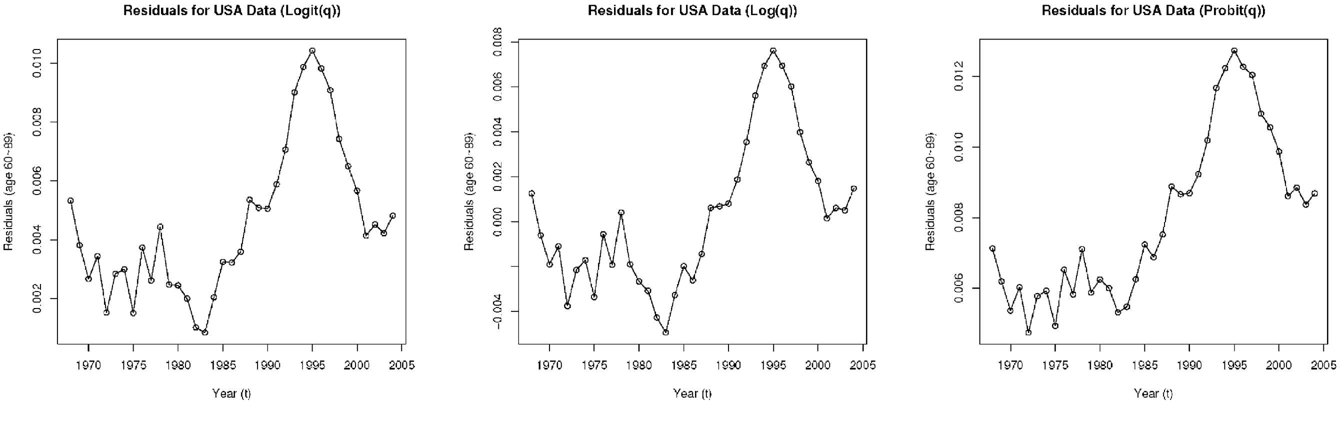
Patterns of Aggregate Residuals, England and Wales



(a) Logit (b) Log (c) Probit

Figure D.8

Patterns of Aggregate Residuals, United States



(a) Logit (b) Log (c) Probit

---

[1] There are other forms of residuals, such as standardized residuals.

tematically above zero for both populations. However, the conclusion for the other two transformations depends on the population. The logit transformation outperforms when it is applied to the data from England and Wales, but not when applied to the data from the United States.

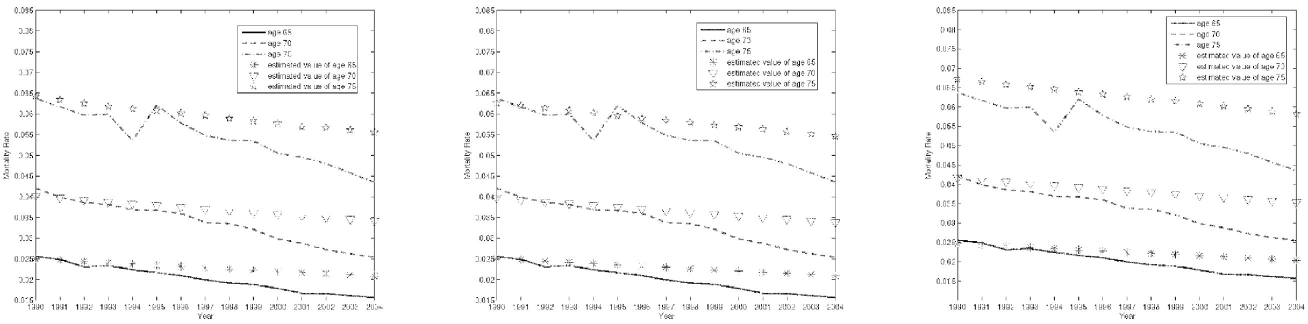
Besides goodness-of-fit, the ex post forecasting performance is also important. As such, we perform a simple backtest on the models.[2] The backtest is performed as follows:

1. Fit the models to a restricted sample period (1961–1989 for England and Wales; 1968–1989 for the United States);
2. Compare the forecast of $q(x, t)$, $t = 1990, \ldots, 2004$, with the actual death probabilities.

In Figures D.9 and D.10 we plot the actual and projected death probabilities at three representative ages (65, 70, and 75). For the U.S. population, the projections made by all three models are reasonably accurate. However, for the English and Welsh population, the predicted pace of mortality decline seems too slow. We also observe that the overestimation of $q(x, t)$ is the most significant when the probit version is used. The logit and log versions have a similar forecasting performance.

Figure D.9

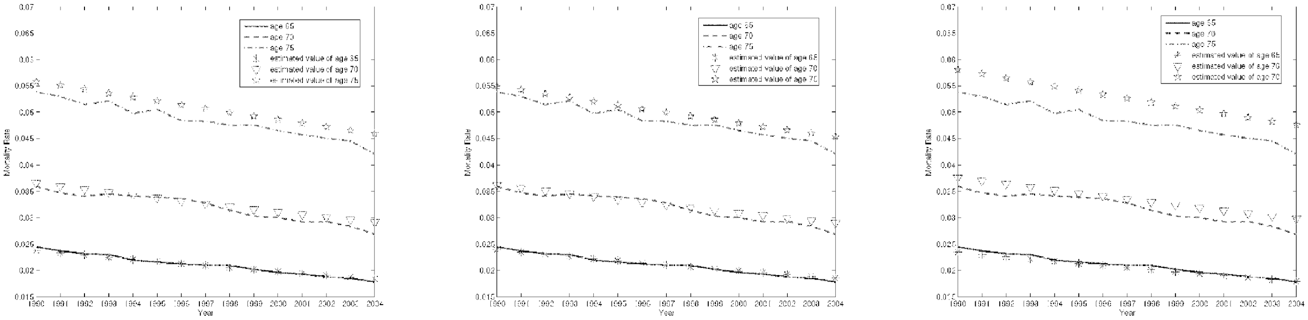
Actual and Projected Death Probabilities, England and Wales



(a) Logit (b) Log (c) Probit

Figure D.10

Actual and Projected Death Probabilities, United States



(a) Logit (b) Log (c) Probit

---

[2] Dowd et al. (2008) provide a comprehensive discussion on backtesting stochastic mortality models.

D.4. CONCLUSION

In Table D.2 we summarize the comparison between our proposed extensions and the original CBD model (the logit transformation).

The original CBD model outperforms the others in at least two of the three criteria we considered. However, the log version may be a good alternative when we make a forecast of the U.S. mortality. The choice depends partly on the data and partly, as Professor Cairns and his coauthors mentioned, on how the user weights the selection criteria.

Our findings suggest that, besides introducing extra stochastic factors (e.g., $\gamma_{t-x}^{(3)}$ in Model M6), the CBD model may be improved by considering a change in the model structure. Given the importance of the CBD class of models in the valuation of longevity securities, further research on this issue is warranted.

Table D.2

Comparison between Logit, Log, and Probit Transformations

Criterion	Best Transformation
SSE	Logit (England and Wales), logit/log (United States)
Aggregate residuals	Logit (England and Wales), log United States)
Backtesting	Logit, log

REFERENCES

CAIRNS, A. J. G., D. BLAKE, AND K. DOWD. 2006. A Two-Factor Model for Stochastic Mortality with Parameter Uncertainty: Theory and Calibration. *Journal of Risk and Insurance* 73: 687–718.

DOWD, K., A. J. G. CAIRNS, D. BLAKE, G. D. COUGHLAN, D. EPSTEIN, AND M. KHALAF-ALLAH. 2008. Backtesting Stochastic Mortality Models: An Ex-Post Evaluation of Multi-Period-Ahead Density Forecasts. Pensions Institute Discussion Paper PI-0803.

HUMAN MORTALITY DATABASE. 2009. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org or www.humanmortality.de (data downloaded on Feb. 1, 2009).

''Valuation of Discrete Dynamic Fund Protection under Lévy Processes,'' Hoi Ying Wong and Ka Wai Lam, April 2009

JUN YANG*

The authors must be congratulated for their paper on valuing discrete dynamic fund protection under Lévy processes. This discussion is motivated by Section 3.2, which is related to the method of *Esscher transforms,* a time-honored technique in actuarial science. I shall also present related results in McDonald (2006), which is the official textbook for Examination MFE/3F.

S. R. Srinivasa Varadhan of New York University was the recipient of the 2007 Abel Prize of the Norwegian Academy of Science and Letters. (The Abel Prize is often described as ''the mathematician's Nobel Prize.'') The prize was given ''for his fundamental contributions to probability theory and in particular for creating a unified theory of large deviation.'' Before the Abel Prize celebration in Oslo, Professor Varadhan was interviewed (Raussen and Skau 2008), and here are his remarks about Esscher transforms:

> The subject of large deviations goes back to the early 1930s. It in fact started in Scandinavia, with actuaries working for the insurance industry. The pioneer who started that subject was named Esscher. He was inter-

---

* Jun Yang is a PhD student in the Department of Statistics and Actuarial Science at the University of Iowa, Iowa City, IA 52242, junyangcc@gmail.com.

> ested in a situation where too many claims could be made against the insurance company, he was worried about the total claim amount exceeding the reserve fund set aside for paying these claims, and he wanted to calculate the probability of this. . . . And what you are really interested in is the probability that the sum of a large number of independent random variables exceeds a certain amount. You are interested in estimating the tail probabilities of sums of independent random variables. . . . Esscher came up with this idea that is called Esscher's tilt. . . . It is a way of changing the measure that you use in a very special manner. And from this point of view, what was originally a tail event, now becomes a central event. So you can estimate it much more accurately and then go from this estimate to what you want, usually by a factor which is much more manageable.

Gerber and Shiu (1994, 1996) extend the concept of Esscher transforms from single random variables to Lévy processes. Let $\{X(t)\}$ be a Lévy process and h be a real number. For a measurable function of the process up to time T,

$$g(X(t),\, 0 \le t \le T),$$

its expectation with respect to the changed probability measure, indexed by h, is

$$\begin{aligned} \mathrm{E}[g(X(t),\, 0 \le t \le T);\, h] &= \mathrm{E}\left[g(X(t),\, 0 \le t \le T)\, \frac{e^{hX(T)}}{\mathrm{E}[e^{hX(T)}]} \right] \\ &= \mathrm{E}[g(X(t),\, 0 \le t \le T) e^{hX(T)}]/\mathrm{E}[e^{hX(T)}] \\ &= \mathrm{E}[g(X(t),\, 0 \le t \le T) e^{hX(T)}]/[M_{X(1)}(h)]^T, \end{aligned} \tag{D.1}$$

where $M_{X(1)}(h)$ denotes the moment-generating function of the random variable $X(1)$ evaluated at h. It is assumed that the moment-generating function exists.

The elegant *factorization formula* is as follows. For real numbers h and k,

$$\mathrm{E}[e^{kX(T)} \times g(X(t),\, 0 \le t \le T);\, h] = \mathrm{E}[e^{kX(T)};\, h] \times \mathrm{E}[g(X(t),\, 0 \le t \le T);\, k + h]. \tag{D.2}$$

The expectation of a product of two factors that are not independent is factorized as a product of two expectations. Note that the last expectation in (D.2) is taken with respect to a different probability measure. Its proof is easy:

$$\begin{aligned} \text{L.H.S.} &= \mathrm{E}[e^{kX(T)} g(X(t),\, 0 \le t \le T) e^{hX(T)}]/\mathrm{E}[e^{hX(T)}] \\ &= \mathrm{E}[e^{(k+h)X(T)} g(X(t),\, 0 \le t \le T)]/\mathrm{E}[e^{hX(T)}] \\ &= \mathrm{E}[e^{(k+h)X(T)}] \cdot \mathrm{E}[g(X(t),\, 0 \le t \le T);\, k + h]/\mathrm{E}[e^{hX(T)}] \\ &= \mathrm{E}[e^{kX(T)};\, h] \cdot \mathrm{E}[g(X(t),\, 0 \le t \le T);\, k + h] \\ &= \text{R.H.S.} \end{aligned}$$

Section 3.2 of Wong and Lam (2009) can be viewed as a consequence of (D.2). In the paper,

$$F_t = F_0 e^{X(t)},$$

$$S_t = K/F_t,$$

and $M_N(S_t)$ denotes the maximum value of $\{S_t\}$ over N monitoring. By considering $\max(M_N(S_t) - 1, 0)$ as $g(X(t),\, 0 \le t \le T)$, we have

$$\begin{aligned} \mathrm{DFP}(0) &= e^{-rT} \mathrm{E}^Q[F_0 e^{X(T)} \max(M_N(S_t) - 1,\, 0)] \\ &= F_0 e^{-rT} \mathrm{E}^Q[e^{X(T)}] \cdot \mathrm{E}^Q[\max(M_N(S_t) - 1,\, 0);\, 1] \\ &= F_0 e^{-rT} e^{rT} \mathrm{E}^Q[\max(M_N(S_t) - 1,\, 0);\, 1] \\ &= F_0 \mathrm{E}^Q[\max(M_N(S_t) - 1,\, 0);\, 1], \end{aligned}$$

which is (3.15) on page 212.

The factorization formula (D.2) can be extended to multidimensional Lévy processes $\{\mathbf{X}(t)\}$ (Gerber and Shiu 1994, 1996):

$$\mathrm{E}[e^{<\mathbf{k},\mathbf{X}(T)>}g(\mathbf{X}(t),\ 0 \le t \le T);\ \mathbf{h}] = \mathrm{E}[e^{<\mathbf{k},\mathbf{X}(T)>};\ \mathbf{h}] \cdot \mathrm{E}[g(\mathbf{X}(t),\ 0 \le t \le T);\ \mathbf{k} + \mathbf{h}], \tag{D.3}$$

where $<\mathbf{k}, \mathbf{X}(T)>$ denotes the inner product of the vectors $\mathbf{k}$ and $\mathbf{X}(T)$. This formula has many elegant applications. For example, it prices *exchange options,* of which the ordinary call and put options are special cases.

Consider a European exchange option on two securities S_1 and S_2, with time-T payoff

$$[S_1(T) - S//I//_2(T)]_+ = S_1(T)I[S_1(T) > S_2(T)] - S_2(T)I[S_1(T) > S_2(T)], \tag{D.4}$$

where $I(\cdot)$ is the indicator function. The first term on the right-hand side of (D.4) is the payoff of an *asset-or-nothing call* option. Under appropriate assumptions so that the factorization formula (D.3) can be applied, its time-0 price is

$$\begin{aligned}\mathrm{E}^Q[e^{-rT}S_1(T)I[S_1(T) > S_2(T)]] &= \mathrm{E}^Q[e^{-rT}S_1(T)] \cdot \mathrm{E}^Q[I[S_1(T) > S_2(T)];\ (1, 0)'] \\ &= F^P_{0,T}(S_1) \cdot \mathrm{Pr}^Q[S_1(T) > S_2(T);\ (1, 0)'].\end{aligned} \tag{D.5}$$

The symbol $F^P_{0,T}(S_1)$ is due to McDonald (2006); it is the time-0 *prepaid forward price* for delivery of one unit of Security 1 at time T.

The second term on the right-hand side of (D.4) is the payoff of an *asset-or-nothing put* option. Similar to (D.5), its time-0 price is

$$F^P_{0,T}(S_2) \cdot \mathrm{Pr}^Q[S_1(T) > S_2(T);\ (0, 1)']. \tag{D.6}$$

Combining (D.5) and (D.6), the time-0 price of the exchange option with payoff (D.4) is

$$F^P_{0,T}(S_1) \cdot \mathrm{Pr}^Q[S_1(T) > S_2(T);\ (1, 0)'] - F^P_{0,T}(S_2) \cdot \mathrm{Pr}^Q[S_1(T) > S_2(T);\ (0, 1)']. \tag{D.7}$$

Now, consider

$$\mathbf{X}(t) = (\ln[F^P_{t,T}(S_1)/F^P_{0,T}(S_1)],\ \ln[F^P_{t,T}(S_2)/F^P_{0,T}(S_2)])',\ 0 \le t \le T, \tag{D.8}$$

where

$$F^P_{t,T}(S_j) = \mathrm{E}^Q_t[e^{-r(T-t)}S_j(T)]. \tag{D.9}$$

Assume that $\{\mathbf{X}(t)\}$ is a two-dimensional Brownian motion with

$$\mathrm{Var}[\mathbf{X}(t)] = t\begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix} = t\Sigma. \tag{D.10}$$

The mean vector of $\mathbf{X}(t)$ under the risk-neutral measure Q is

$$\mathrm{E}^Q[\mathbf{X}(t)] = t(r - \tfrac{1}{2}\sigma_1^2,\ r - \tfrac{1}{2}\sigma_2^2)'. \tag{D.11}$$

It follows from Gerber and Shiu (1994, 1996) that

$$\begin{aligned}\mathrm{E}^Q[\mathbf{X}(t);\ (1, 0)'] &= \mathrm{E}^Q[\mathbf{X}(t)] + t\Sigma(1, 0)' \\ &= \mathrm{E}^Q[\mathbf{X}(t)] + t(\sigma_1^2,\ \rho\sigma_1\sigma_2)' \\ &= t(r + \tfrac{1}{2}\sigma_1^2,\ r - \tfrac{1}{2}\sigma_2^2 + \rho\sigma_1\sigma_2)'.\end{aligned} \tag{D.12}$$

Note that

$$\begin{aligned}\Pr[S_1(T) > S_2(T)] &= \Pr[F^P_{T,T}(S_1) > F^P_{T,T}(S_2)]\\ &= \Pr[\ln F^P_{T,T}(S_1) > \ln F^P_{T,T}(S_2)]\\ &= \Pr[X_1(T) + \ln F^P_{0,T}(S_1) > X_2(T) + \ln F^P_{0,T}(S_2)]\\ &= \Pr[X_2(T) - X_1(T) < \ln[F^P_{0,\mathrm{T}}(S_1)/\ln F^P_{0,T}(S_2)]]. \qquad \text{(D.13)}\end{aligned}$$

From (D.12), we have

$$\begin{aligned}\mathrm{E}^Q[X_2(T) - X_1(T);\ (1,0)'] &= T(-\tfrac{1}{2}\sigma_2^2 + \rho\sigma_1\sigma_2 - \tfrac{1}{2}\sigma_1^2)\\ &= -\tfrac{1}{2}\mathrm{Var}[X_2(T) - X_1(T)]\\ &= -\tfrac{1}{2}\sigma^2 T, \qquad \text{(D.14)}\end{aligned}$$

by defining

$$\sigma^2 = \mathrm{Var}[X_2(1) - X_1(1)]. \qquad \text{(D.15)}$$

By (D.13) to (D.15) and the fact that the variance of $\mathbf{X}(T)$ is unchanged under an Esscher transform, we obtain

$$\Pr{}^Q[S_1(T) > S_2(T);\ (1,\ 0)'] = N\left(\frac{\ln[F^P_{0,T}(S_1)/F^P_{0,T}(S_2)]}{\sigma\sqrt{T}} + \frac{1}{2}\,\sigma\sqrt{T}\right), \qquad \text{(D.16)}$$

where $N(.)$ is the cumulative distribution function of a standard normal random variable. Similarly,

$$\Pr{}^Q[S_1(T) > S_2(T);\ (0,\ 1)'] = N\left(\frac{\ln[F^P_{0,T}(S_1)/F^P_{0,T}(S_2)]}{\sigma\sqrt{T}} - \frac{1}{2}\,\sigma\sqrt{T}\right). \qquad \text{(D.17)}$$

It follows from (D.16) and (D.17) that the exchange option pricing formula (D.7) becomes

$$F^P_{0,T}(S_1)N\left(\frac{\ln[F^P_{0,T}(S_1)/F^P_{0,T}(S_2)]}{\sigma\sqrt{T}} + \frac{1}{2}\,\sigma\sqrt{T}\right) - F^P_{0,T}(S_2)N\left(\frac{\ln[F^P_{0,T}(S_1)/F^P_{0,T}(S_2)]}{\sigma\sqrt{T}} - \frac{1}{2}\,\sigma\sqrt{T}\right), \qquad \text{(D.18)}$$

which generalizes a formula in the Examination MFE/3F syllabus (McDonald 2006, eq. 14.16).

We end this discussion with a derivation of Proposition 21.1 in McDonald (2006) by using the factorization formula (D.3). Instead of (D.8), we now consider

$$\mathbf{X}(t) = (\ln[S_1(t)/S_1(0)],\ \ln[S_2(t)/S_2(0)])'. \qquad \text{(D.19)}$$

We also assume that $\{\mathbf{X}(t)\}$ is a two-dimensional Brownian motion satisfying (D.10). Furthermore, we assume that for $j = 1, 2$, Security j pays dividends of amount $S_j(t)\delta_j dt$ between time t and time $t + dt$. Instead of (D.11), we now have

$$\mathrm{E}^Q[\mathbf{X}(t)] = t(r - \delta_1 - \tfrac{1}{2}\sigma_1^2,\ r - \delta_2 - \tfrac{1}{2}\sigma_2^2)'. \qquad \text{(D.20)}$$

Let $V(S_1, \delta_1)$ denote the time-0 price of a European derivative claim on Security 1 maturing at time T, that is,

$$V(S_1, \delta_1) = \mathrm{E}^Q[e^{-rT}\pi(S_1, T)], \qquad \text{(D.21)}$$

where $\pi//I//(S_1, T)$ denote the time-T payoff of the derivative claim. Proposition 21.1 gives a formula for pricing a claim that pays at time T

$$\pi(S_1, T)[S_2(T)]^b. \qquad \text{(D.22)}$$

It follows from the *fundamental theorem of asset pricing* that the time-0 price of the claim is

$$\mathrm{E}^Q[e^{-rT}\pi(S_1, T)[S_2(T)]^b] = \mathrm{E}^Q[[S_2(T)]^b]\mathrm{E}^Q[e^{-rT}\pi(S_1, T);\ (0, b)'] \tag{D.23}$$

by (D.3).

The first expectation on the right-hand side of (D.23) is

$$[S_2(0)]^b\mathrm{E}^Q[e^{bX_2(T)}] = [S_2(0)]^b M^Q_{X_2(T)}(b), \tag{D.24}$$

where $M^Q_{X_2(T)}(\cdot)$ is the moment-generating function of $X_2(T)$ calculated under the risk-neutral measure Q. It follows from (D.20) that

$$\begin{aligned} M^Q_{X_2(T)}(b) &= \exp[(r - \delta_2 - \tfrac{1}{2}\sigma_2^2)Tb + \tfrac{1}{2}\sigma_2^2Tb^2] \\ &= e^{(r-\delta*)T}, \end{aligned} \tag{D.25}$$

where, as defined by McDonald (2006, p. 694),

$$\delta^* = r - b(r - \delta_2) - \tfrac{1}{2}b(b - 1)\sigma_2^2. \tag{D.26}$$

For the second expectation on the right-hand side of (D.23), note that, because

$$\begin{aligned} \mathrm{E}^Q[\mathbf{X}(t);\ (0, b)'] &= \mathrm{E}^Q[\mathbf{X}(t)] + t\boldsymbol{\Sigma}(0, b)' \\ &= \mathrm{E}^Q[\mathbf{X}(t)] + tb(\rho\sigma_1\sigma_2, \sigma_2^2)', \end{aligned}$$

we have

$$\begin{aligned} \mathrm{E}^Q[X_1(t);\ (0, b)'] &= t(r - \delta_1 - \tfrac{1}{2}\sigma_1^2 + b\rho\sigma_1\sigma_2) \\ &= t[r - (\delta_1 - b\rho\sigma_1\sigma_2) - \tfrac{1}{2}\sigma_1^2]. \end{aligned} \tag{D.27}$$

It follows from (D.21) and (D.27) that

$$\mathrm{E}^Q[e^{-rT}\pi(S_1, T);\ (0, b)'] = V(S_1, \delta_1 - b\rho\sigma_1\sigma_2). \tag{D.28}$$

Thus, the time-0 price of the derivative claim with payoff (D.22) is

$$V(S_1, \delta_1 - b\rho\sigma_1\sigma_2)[S_2(0)]^b e^{(r-\delta^*)T}, \tag{D.29}$$

which is equivalent to expression (21.37) on page 694 of McDonald (2006).

REFERENCES

GERBER, H. U., AND E. S. W. SHIU. 1994. Option Pricing by Esscher Transforms. *Transactions of Society of Actuaries* 46: 99–140.

GERBER, H. U., AND E. S. W. SHIU. 1996. Actuarial Bridges to Dynamic Hedging and Option Pricing. *Insurance: Mathematics and Economics* 18: 183–218.

MCDONALD, R. L. 2006. *Derivatives Markets*. 2nd ed. Boston: Pearson.

RAUSSEN, M., AND C. SKAU. 2008. Interview with Srinivasa Varadhan. *Notices of the American Mathematical Society* 55: 238–46.

WONG, H. Y., AND K. W. LAM. 2009. Valuation of Discrete Dynamic Fund Protection under Lévy Processes. *North American Actuarial Journal* 13: 202–16.