Different Shades of Risk: Mortality Trends Implied by Term Insurance Prices*

Qiheng Guo†
Department of Risk Management and Insurance. Georgia State University
35 Broad Street. Atlanta, GA 30303. USA

Daniel Bauer‡
Department of Risk Management and Insurance. Georgia State University
35 Broad Street. Atlanta, GA 30303. USA

December 2016

Abstract

To infer forward-looking, market-based mortality trends, we estimate a flexible affine stochastic mortality model based on a set of US term life insurance prices using a generalized method of moments approach. We find that neither mortality shocks nor stochasticity in the aggregate trend seem to affect the prices. In contrast, allowing for heterogeneity in the mortality rates across carriers is crucial. We conclude that for life insurance, rather than aggregate mortality risk, the key risks emanate from the composition of the portfolio of policyholders. These findings have consequences for mortality risk management and emphasize important directions of mortality-related actuarial research.

Keywords: Mortality Catastrophe. Affine Stochastic Mortality Model. GMM Estimation. Basis Risk. Mortality Heterogeneity.

---

*A previous version was presented at the Twelfth International Longevity Risk and Capital Markets Solutions Conference (Longevity 12) under the title “Mortality Trends Implied by Term Insurance Prices”, and parts are taken from the earlier working paper “The Risk in Catastrophe Mortality Securitization Transactions” by the second author. The authors are grateful for helpful comments from the Longevity 12 conference participants and for financial support under a Society of Actuaries CAE grant.

†Corresponding author. Phone: +1-(678)-999-6385. Fax: +1-(404)-413-7499. E-mail: qguo5@gsu.edu.

‡E-mail address: dbauer@gsu.edu.
1 Introduction

Term life insurance policies are typically considered to be fairly homogeneous products. Aside from conversion options and financial strength ratings—the relevance of which is mitigated to some extent by guaranty funds protection—the ensuing cash flows are typically relative congenerous across issuers. The key risk factors are investment/interest and mortality risks, where in view of the former investment opportunities and strategies again do not vary much across companies. Thus, given competitive forces in this large and undifferentiated market, it seems proximate to infer forward-looking, market-based estimates of future mortality dynamics from insurance prices (Mullin and Philipson, 1997).

Building on this logic, we estimate the stochastic mortality model from Bauer and Kramer (2016) to a set of US (term) life insurance prices using a generalized method of moments (GMM) approach, where we allow for selection/underwriting effects, surrenders, pertinent expenses, etc. The model includes a catastrophe component to pick up mortality shocks associated with pandemics or natural disasters, a fairly flexible mortality trend component, and a diffusion term to capture stochastic variations in the mortality trend. Our results are striking: Neither the catastrophe component nor the diffusion term are significant, and a model comparison favors a simple deterministic mortality model. In contrast, allowing for heterogeneity among the carriers is of utmost importance: A model that does not incorporate differences in mortality trends between the different companies is strongly rejected. The company effects are large in magnitude and point towards three possible sub-markets.

Our interpretation is that for pricing and managing term life insurance products—potentially in contrast to annuity and pension products—the key risks emanate from the composition of the pool of policyholders, rather than the uncertainty in aggregate mortality trends. Differences in underwriting criteria, primary distribution channels and distribution area, and other factors jointly determine the composition of the insurer’s portfolio of policies in each rating class. The relevant mortality rates for a given company will be based on the demographic
evolution of this subgroup, and their behavior with regards to lapsing their contracts. Given heterogeneity in mortality trends for different population subgroups and different causes of death, the cross-sectional risk dimension dwarfs uncertainties in the aggregate trend. In other words, *basis risk*, i.e. the deviation of the experience in the particular pool relative to the aggregate population, seems to dominate *systematic mortality risk*, i.e. uncertainties in the aggregate trend component. And, indeed, an active management of the composition of the pool, by accepting certain risks and rejecting others, may be a way to compete in the marketplace.

Our findings have consequences for mortality risk management and for the corresponding actuarial literature. On the one side, they do not support assertions that aggregate mortality risks are important for life insurance products. For instance, papers presenting “natural hedging” between life insurance and annuity lines as an internal way to manage a company’s mortality exposure implicitly assume that the corresponding populations of policyholders are subject to the same variations. Thus, our results cast doubt on the effectiveness of such strategies, pointing to the necessity of market-based solutions for managing longevity. On the other side, our conclusion that mortality trends for different partitions of the population and different conditions determine the mortality profile of the relevant portfolio of policyholders, and that it may be possible to actively manage this profile, emphasizes the importance of studying mortality at a more granular level.

**Related Literature and Organization of the Paper**

A closely related paper to ours is Mullin and Philipson (1997), who argue that under the assumption that insurance companies are close to risk-neutral with respect to their mortality exposure, it is possible to derive market-based estimates from zero expected profit (moment) conditions. In contrast to Mullin and Philipson, however, we include additional institutional factors affecting life insurance prices such as expenses, selection/underwriting effects, and policy surrenders/lapses. Furthermore, we consider variations in mortality in the time and
the cross-sectional (across companies) direction. The latter aspect, in particular, allows us to draw our primary conclusions. Several papers in the actuarial literature also rely on prices of insurance products to obtain parameters in mortality models. In particular, Lin and Cox (2005) and Bauer et al. (2010) use annuity quotes to estimate risk premiums for longevity risk.

In the estimation process, we use the stochastic mortality model from Bauer and Kramer (2016). As discussed in their paper, the model is flexible enough to fit a relatively long time series of US mortality data, it includes a mortality catastrophe component, and it is tractable. The latter property originates from it falling in the class of affine mortality models (Biffis, 2005; Dahl and Møller, 2006). In particular, this allows for an efficient computation of survival probabilities and, thus, basic life insurance prices, which is essential for our estimation approach.

Our results relate to recent results in the economic and medical literature on the heterogeneity of mortality trends across different subpopulations that show that there are large disparities in life expectancy between different racial, regional, and socio-economic groups (Chetty et al., 2016; Case and Deaton, 2015). Furthermore, they relate and endorse analyses of cause-specific mortality rates, and how these in aggregate affect the mortality of a certain population (Arnold et al., 2016; Arnold and Sherris, 2016, and references therein). An insurer’s underwriting process, together with its regional presence, its advertising strategy, etc., determines the composition of the portfolio of policyholders, and it is the mortality dynamics of that group that is relevant for the insurer’s future cash flows.

As indicated, our findings are relevant for so-called “natural hedging” approaches to managing longevity risk by trading off annuity and life insurance exposures (Cox and Lin, 2007; Li and Haberman, 2015, and references therein). In particular, our results suggest that beyond difficulties with natural hedging within the same population (Zhu and Bauer, 2015), basis risk may inhibit the effectiveness of such strategies.

The remainder of the paper is structured as follows: Section 2 introduces the affine mor-
tality model from Bauer and Kramer (2016) with some extensions to suit our setting. Section 3 introduces our GMM estimation method, presents results, and their discussion. Finally, Section 4 concludes.

2 Model

In what follows, we introduce the mortality model from Bauer and Kramer (2016), which we rely on in the remainder of the text. As discussed in their paper, the model consists of three parts: 1) a catastrophe component, 2) an age-dependent affine mortality component, and 3) a temporary component. In contrast to their analysis, we use the third part to include selection/underwriting effects, rather than a temporary deterministic trend. Furthermore, we subsequently augment the model by company, risk class, and calendar year effects that account for heterogeneity in mortality trends.

As in Bauer and Kramer (2016), we introduce the stochastic force of mortality by relying on conventional concepts from credit risk modeling (Lando, 1998). More precisely, it given a stochastic process \( X = (X_t)_{0 \leq t \leq T} \) and positive, continuous function \( \mu(s, \cdot) \), we define an individual’s time of death as the first jump time of a Cox-process with intensity \( \mu(x_0 + t, X_t) \):

\[
\tau_{x_0} = \inf \left\{ t : \int_0^t \mu(x_0 + s, X_s) \, ds \geq E \right\},
\]

(1)

where \( E \) is a \( \text{Exp}(1) \)-distributed random variable and independent among individuals.

Considering only one single insured for now, let the filtrations \( G = \{ \mathcal{G}_t \}_{0 \leq t \leq T} \) and \( H = \{ \mathcal{H}_t \}_{0 \leq t \leq T} \) be given as the augmentations of the filtrations generated by \( (X_t)_{0 \leq t \leq T} \), and \( \{1_{\{\tau_{x_0} \leq t\}}\}_{0 \leq t \leq T} \), respectively, and set \( \mathcal{F}_t = \mathcal{G}_t \vee \mathcal{H}_t \). From Equation (1), we can then derive the \( (T - t) \)-year survival probability at time \( t \) for an \( x_t = x_0 + t \) year old individual as

\[
P_{x_t}(t) := \mathbb{E} \left[ 1_{\{\tau_{x_0} > T\}} \, \mathcal{G}_t, \tau_{x_0} > t \right] = \mathbb{E} \left[ \exp \left\{ - \int_t^T \mu(x_0 + s, X_s) \, ds \right\} \, \mathcal{G}_t, \tau_{x_0} > t \right].
\]
and, from results of Lando (1998),

\[ \mathbb{E} \left[ 1_{\{\tau_{x_0} > T\}} \mid \mathcal{F}_T \right] = 1_{\{\tau_{x_0} > t\}} T - t p_{x_0+t}(t) \]

Following Bauer and Kramer (2016), we use the following model for the baseline stochastic force of mortality:

\[ \mu_t(x_0) = \mu(x_0 + t, Y_t, \Gamma_t) = e^{b(x_0 + t)} Y_t + \Gamma_t + D_t(x_0), \quad \Gamma_0, Y_0 > 0. \]  

(3)

The catastrophe component follows the dynamics:

\[ d\Gamma_t = -\kappa \Gamma_t dt + dJ_t, \quad \Gamma_0 \geq 0, \]

where \((J_t)\) is a compound Poisson process with intensity \(\lambda\) and \(Exp(\xi)\)-distributed jumps.

And the baseline component follows the dynamics:

\[ dY_t = \alpha \left( \left( Y_0 - \beta^{(2)} \right) \frac{e^{-\beta^{(1)} t}}{\beta^{(3)}} + \beta^{(2)} - Y_t \right) dt + \sigma \sqrt{Y_t} dW_t, \quad Y_0 > 0, \]

where \((W_t)_{0 \leq t \leq T^*}\) is a one-dimensional Brownian motion and \(\alpha, \beta^{(1)}, \beta^{(2)},\) and \(\sigma\) are positive constants with \(\alpha \neq \beta^{(1)}, \beta^{(2)} + \beta^{(3)}\) describes the trend level at time 0. We refer to their paper for the motivation of the model in the context of demographic research.

We use the temporary component \(D_t(x_0)\) to introduce selection effects acting to temporary reduce mortality. Here the *selection effect* does not refer to the potential impact of adverse se-
lection on life insurance prices but the impact of underwriting during the early policy years.\footnote{In insurance practice, actuaries rely on so-called “select-and-ultimate” tables to account for this type of selection, where the “select” tables used in early policy years display lower mortality due to underwriting examinations.} In particular, this is important for the evaluation of life insurance prices as here mandatory health examinations lead to significant selection effects. Within our model, this can be captured via the temporary component $D$ by a roughly proportional structure akin to a proportional hazards model:

$$D_t(x_0) = c - (c + e^{b(x_0+t)}) \times \frac{\gamma}{T} \times (\tau - t)^+.$$  

Note that the component is still deterministic and it is only roughly proportional since we use an approximation of the relevant force of mortality based on the initial values. Moreover, this specification of $D$ only relies on the four additional parameters ($c, \tau, \gamma, \bar{T}$) minding the complexity of the estimation process. We obtain:

$$\bar{D}_{t,T}(x_0) = \int_t^T D_s(x_0) \, ds$$

$$= c (T - t) - \frac{c \gamma}{T} \left[ ((\bar{T} - t)^+)^2 - ((\bar{T} - T)^+)^2 \right]$$

$$- \frac{\tau \gamma}{T \bar{T}} \left[ (T - T)^+ \exp \{ b(x_0 + \min\{T, \bar{T}\}) \} - (T - t)^+ \exp \{ b(x_0 + \min\{t, T\}) \} \right]$$

$$- \frac{\tau \gamma}{T \bar{T}^2} \left[ \exp \{ b(x_0 + \min\{T, \bar{T}\}) \} - \exp \{ b(x_0 + \min\{t, \bar{T}\}) \} \right].$$

The (exponential-)affine structure enables us to write down an analytical representation of the survival function (cf. Prop. 1 in Duffie et al. (2000)):

$$T_{t} p_{x_0+t}(t) = \exp \left\{ u(T - t) + v(T - t) Y_t - \frac{\Gamma_t}{\kappa} \left( 1 - e^{-\kappa(T-t)} \right) - \frac{\lambda(T - t)}{\bar{T} \kappa + 1} \right\}$$

$$\times \exp \left\{ \frac{\lambda \kappa}{\bar{T} \kappa + 1} \log \left[ 1 + \frac{1}{\bar{T} \kappa} \left( 1 - e^{-\kappa(T-t)} \right) \right] - \bar{D}_{t,T}(x_0) \right\},$$

(4)
where \( u \) and \( v \) satisfy the following Riccatti ODEs:

\[
\begin{align*}
  v'(s) &= -e^{b(x_0 + T - s)} - \alpha v(s) + \frac{1}{2}\sigma^2 v^2(s), \quad v(0) = 0, \\
  u'(s) &= v(s)\alpha \left( e^{-\beta(1)(T - s)Y_0} + (1 - e^{-\beta(1)(T - s)})\beta(2) \right), \quad u(0) = 0.
\end{align*}
\] (5)

Here, \( \mu_t(x_0) \) and \( T - t p_{x_0 + t} \) represent the baseline force of mortality and survival function, respectively.

We introduce heterogeneity in mortality trends by assuming that the aggregate force of mortality in each company’s portfolio is proportional to the baseline. Thus, the force of mortality for company \( i \) can be written as:

\[
\mu_i^t(x_0) = \mu_t(x_0)(1 + E_{co}^i), \quad \sum_i E_{co}^i = 0, \quad i = 1, 2, \ldots, I
\]

where \( E_{co}^i \) is the company effect. We can further add the risk class effect \( E_{rc}^j \) and calendar year effect \( E_{year}^h \) to the model to account for variation within a company and across different calendar year:

\[
\mu_i^{(i,j,h)}(x_0) = \mu_t(x_0)(1 + E_{co}^i + E_{rc}^j + E_{year}^h),
\]

\[
\sum_i E_{co}^i = \sum_j E_{rc}^j = \sum_h E_{year}^h = 0, \quad E_{co}^i + E_{rc}^j + E_{year}^h > -1 \forall i, j, h
\]

with risk class \( j \) ranges from 1 (highest mortality risk, e.g. regular class) to \( J \) (lowest mortality risk, e.g. preferred plus class) and year \( h \) spans all calendar years of insurance prices. The three effect components do not relate to the industry baseline force of mortality. Therefore, the survival function for company \( i \), risk class \( j \) in calendar year \( h \) is:

\[
T - t p_{x_0 + t}^{(i,j,h)}(t) = T - t p_{x_0 + t}(t)^{1 + E_{co}^i + E_{rc}^j + E_{year}^h}
\]

Given risk class, calendar year, and company effect parameters, it remains to calculate the sur-
survival function (4) and actuarial present values by simply solving the ODEs from Equation (5). The full analytical representation of the survival function facilitates the estimation process.

3 Estimation to Insurance Price Data

3.1 Insurance Price GMM Estimator

We are given annual term-insurance premiums $P_{x,n}^{(i,j,h)}$ for age $x$, term $n$, risk class $j \in 1, 2, ..., J$, calendar year $h \in 1, 2, ..., H$ and a fixed benefit $B$ payable upon death at the end of the year of from $I$ companies, i.e. $i \in \{1, 2, \ldots, I\}$, which we take to be i.i.d. Denote by $C$ the collection of all available age-term combinations $(x, n)$ and we have $N_{x,n}$ such combination. As is common in actuarial modeling and in contrast to Mullin and Philipson (1997), we consider various types of expenses to be reflected in our insurance price data. In particular, we include initial expenses both as a percentage of the first premium $c_{IP}^{(1)}$ and as a fixed amount depending on the death benefit $c_{IP}^{(2)}$, as well as a fixed maintenance expense $c_{M}$ (Society of Actuaries, 2004).

We assume policyholders surrender at a fixed proportion $q_u^{(l)}$ in policy year $u$, $u \geq 1$, immediately before premiums become due. Thus, policies remain in force for $k$ years with a probability of

$$kP_{x,n}^{(r)}(0) = kP_{x,n}^{(i,j,h)}(0) \prod_{1 \leq u \leq k} (1 - q_u^{(l)}), \ k \geq 1.$$  

Upon surrendering at time $k$, policyholders may be entitled to so-called cash surrender values (CSVs) $kC_{x,n}^{(i,j,h)}$, which are to be calculated according to the National Association of Insurance Commissioners’ (NAIC) Standard Nonforfeiture Law for Life Insurance. This regulation essentially entails the calculation of guaranteed reserve levels according to given interest rates, a given mortality table, and given expense levels. For the simplicity of model, we assume the
cost and surrender probabilities are the same across all companies.\(^2\)

We start by writing down the an insurance company’s loss function, or the negative profit function:

\[
\begin{align*}
    f(\theta, P^{(i,j,h)}_{x,n}) &= \sum_{k=0}^{n-1} p(0, k + 1) \left[ kp_x^{(l)} \left( p_x^{(i,j,h)}(0) - k_{+1} p_x^{(i,j,h)}(0) \right) B + k_{+1} p_x^{(i,j,h)} q_k^{(l)} q_{k+1} C_{x,n} \right] \\
    &+ c^{(2)}_{IP} + c_M \sum_{k=0}^{n-1} p(0, k) k_{\tau x0}^{(r)} (0) - P_{x,n}^{(i,j,h)} \left( \sum_{k=0}^{n-1} p(0, k) k_{\tau x0}^{(r)} (0) - c_{(1)} \right),
\end{align*}
\]

where \(p(t, \tau)\) denotes the time \(t\) price of a zero coupon bond with maturity \(t + \tau\). Of course, equation \((6)\) depends on the (risk-neutral) parameters of the mortality model as well as on the parameters governing expenses, selection, and surrenders, which we stack in the parameter vector \(\theta\). The key assumption is now that the so-called equivalence principle holds, i.e. that the expected present value of future benefits (including expenses) equals the expected present value of future premiums for all of the company’s policies for one mortality profile, that is, for one risk class in one calendar year. Thus, the equivalence principle must hold for each \(i, j, h\) combination. That is to say, for each company every risk class in one calendar year, the expected profit (loss) from selling insurance contracts is zero. Therefore, we can write down our moment conditions for GMM estimation as follows:

\[
\mathbb{E}_{x,n} \left[ f(\theta, P^{(i,j,h)}_{x,n}) \right] = 0, \ (x, n) \in \mathcal{C}.
\]

The efficient GMM estimator can be obtained using so-called “two-step feasible GMM” method, which yields efficient and consistent estimators. First, choose a weighting matrix \(W\), where \(W\) is a \(I \times J \times H\) dimensional square matrix. \(W\) can be any such matrix in the first

\(^2\)This assumption, again, is motivated by competitive pressures in the market. Large differences in costs should be competed away unless they are linked to aspects relating to underlying heterogeneity.
step. We choose $W$ such that

$$W^{-1} = \text{diag} \left\{ \sigma^2_{x,n}, (x, n) \in C \right\},$$

where $\sigma^2_{x,n} = \text{Var} \left[ P_{x,n} \right]$ with corresponding sample version $\hat{\sigma}^2_{x,n}$. We obtain a GMM estimate $\hat{\theta}_1$ by minimizing the following function of $\theta$:

$$\hat{\theta}_1 = \arg \min_\theta \left[ \frac{1}{N_{x,n}} \sum_{x,n} f(\theta, P_{x,n}^{(i,j,h)}) \right]' W \left[ \frac{1}{N_{x,n}} \sum_{x,n} f(\theta, P_{x,n}^{(i,j,h)}) \right],$$

Then, we update the weighting matrix using $\hat{\theta}_1$:

$$\hat{W}(\hat{\theta}_1)^{-1} = \text{diag} \left\{ \frac{1}{N_{x,n}} \sum_{j=1}^{N_{x,n}} \left( f(\hat{\theta}_1, P_{x,n}^{(i,j,h)}) \right)^2, (i, j, h) \right\}$$

and minimize function (7) using $W = \hat{W}(\hat{\theta}_1)$ to obtain our estimate $\hat{\theta}$. We carry out both minimization numerically.

### 3.2 Data and Estimation

We use quotes for annual life insurance premiums from the CompuLife price quotation system (historical data) for April 2012, 2013, 2014 and 2015. More precisely, we focus on contracts with a face amount of $500,000 for male non-smokers under two underwriting categories: regular (Rg, residual standard), and preferred plus (Pf+, super preferred). From 31 companies, we retrieved data for 33 age-term combinations, with terms of 10, 15, 20, and 30 years, and ages from 25 to 75, where for numerical convenience we use ages of multiple of five only. We only allow for a single quote per company per age/term-combination. Since the records include several instances of multiple quotes from the same company for different states—though prices usually coincide—we averaging over these quotes. All in all, we rely on 7,688
different quotes in our estimation process.

Figure (1a) and (1b) plots the quotes of insurance by term in April 2012, for age 25 and 40 respectively. For each risk class, the variance of insurance prices in each term is about the same and the coefficient of variation decreases as term increases. This goes against the intuition of stochastic mortality diffusion because it implies the increasing variation in mortality (and likely insurance prices) as the term increases. Similar characteristics on the variance of the insurance price are observed in the other calendar years and ages.

![Figure 1: Illustration of the insurance price data](image)

In addition to the parameters of the mortality model, the GMM estimator depends on various business-related parameters. The CSVs \( kC_{x,n}, k \geq 0 \), are calculated according to the NAIC Standard Nonforfeiture Law for Life Insurance. The interest rates derive from Moody’s Corporate Bond Yield Averages Index (we refer to Towers Watson (2015) for details and an illustration of the relevant rates), and the relevant mortality rates are taken from the Commissioners Standard Ordinary (CSO) 2017 mortality table (mandatory in 48 U.S. states and optional in all 50; see American Academy of Actuaries (2008)). We do not include expense, surrender and other business-related parameters in the minimization, but rather fix them according to Table 1, whose column 2 and 3 list their values and the source, respectively.\(^3\)

\(^3\)Including them in the estimation procedure leads to problems when including a complex mortality model,
DIFFERENT SHADES OF RISK: MORTALITY TRENDS IMPLIED BY TERM INSURANCE PRICES

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_{IP}^{(1)}$</td>
<td>60%</td>
<td>Avg. values from the 2005</td>
</tr>
<tr>
<td>$c_{IP}^{(2)}$</td>
<td>$882.5$</td>
<td>“Generally Recognized Expense Table”</td>
</tr>
<tr>
<td>$c_{M}$</td>
<td>$45$</td>
<td>(see e.g. Society of Actuaries (2004))</td>
</tr>
<tr>
<td>$\bar{T}$</td>
<td>18 years</td>
<td>Avg. value from Society of Actuaries (2007)</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>60%</td>
<td>Value matched to CSO 2017 mort. table</td>
</tr>
<tr>
<td>$q_i$</td>
<td>$15%, i \leq 3, 5%, i &gt; 3$</td>
<td>Roughly matches pattern (non-renewable) according to LIMRA and SOA (2005)</td>
</tr>
</tbody>
</table>

Table 1: Parameters relevant to the insurance contracts.

For the numerical optimization in (7), we rely on Julia language and apply the COBYLA (constrained optimization by linear approximation) algorithm, suitable for a minimization task with a large number of equality and inequality constraints of the parameters. For the numerical solution of the ordinary differential equations arising in each time step, we rely on an implementation of a Runge-Kutta method with a variable time step as available within Julia (ode45).

We estimate five versions of the model. The first model version (Full model) includes catastrophe component parameters ($\kappa$, $\lambda$ and $\zeta$). In the second version (w/o CAT), we exclude the catastrophe component. In the third model (w/o Sigma), we assume that the stochastic diffusion parameter $\sigma$ is zero. In the fourth model (w/o Trend), we set the trend parameters $\beta_1$ and $\beta_3$ to zero so that $Y_t$ is fixed at $\beta_2$. The fifth model (w/o Co. effect) takes away company effect from the w/o CAT model. We record the estimates, standard errors and also a likelihood-like J-statistics, which is the basis for testing the overall specification and parametric restrictions since the impact on insurance prices is similar to those originating from certain components. We obtain reasonable magnitudes not too different from the set values in the context of simple mortality models.
as in Sargan (1958) and Hansen (1982). It is calculated as follows:

\[
J = N_{x,n} \left[ \frac{1}{N_{x,n}} \sum_{x,n} f(\hat{\theta}, P^{(i,j,h)}_{x,n}) \right] \left( \frac{1}{N_{x,n}} \sum_{x,n} f(\hat{\theta}, P^{(i,j,h)}_{x,n}) \right)^2 \]

\[
J, a Wald statistic, converges in distribution to \chi^2(L - K) under the null hypothesis that the model is valid, or well specified, where \( L \) is the number of moment equations and \( K \) is the number of parameters. In terms of testing parametric restrictions, according to Newey and West (1987), the difference of two J-statistics, the GMM counterpart to the likelihood ratio test statistic, converges in distribution to \( \chi^2(K_1 - K_2) \), where \( K_1 - K_2 \) is the number of restricted parameters in model 2 compared to model 1. This “Likelihood-ratio-like” statistic can be used to test the parameter restrictions and offer guidance to model comparison, which is highlighted in the next part.

### 3.3 Results and Discussion

Table 2 provides the parameter estimates for all five model versions, with the first four having the company effect parameters. In the Full model, the catastrophe parameter estimates \((\kappa, \lambda, \zeta)\) have large standard errors and are not significant, so do the trend parameter \(\beta_1\) and diffusion parameter \(\sigma\). In contrast, the trend parameter \(\beta_2\) and Gompertz parameter \(b\) are significant, pointing to a simple deterministic mortality model. Parameters related to selection/underwriting effect \((c, \tau, \gamma)\) are significant in the first four model estimations, highlighting such effect in life insurance underwriting.

The estimation results for company effects are shown in Figure 2. Figure 3 summarizes the significance of the estimates compared to zero and we found around one third of companies having mortality rate significantly above the industry average, and around one third of companies having mortality rate significantly below average. In the fifth model, without the company effects capturing the heterogeneity of mortality, parameter estimates turn out to have
large standard errors and the overall specification statistic $J$ is much larger compared to the other four, indicating that the w/o Co. effect model is unlikely to be favorable. The estimation results strongly favors the model with company effects, which overshadow the mortality catastrophe and diffusion parameters.

Figure 2: Estimated company effect

Figure 3: Company effect estimates and their significance to zero at 95% confidence level (red lines serve as visual aid only and are not classification boundaries)
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Full</th>
<th>w/o CAT</th>
<th>w/o Sigma</th>
<th>w/o Trend</th>
<th>w/o Co. effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>1.8793</td>
<td>1.8176</td>
<td>1.8178</td>
<td>1.8270</td>
<td>1.3609</td>
</tr>
<tr>
<td></td>
<td>(6.1387)</td>
<td>(1.2603)</td>
<td>(1.2613)</td>
<td>(5.12E10)</td>
<td>(13.7273)</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>2.4470</td>
<td>2.4440</td>
<td>2.4446</td>
<td>-</td>
<td>2.6631</td>
</tr>
<tr>
<td></td>
<td>(10.5424)</td>
<td>(2.2557)</td>
<td>(2.2593)</td>
<td>-</td>
<td>(50.4369)</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>1.20E-06</td>
<td>1.20E-06</td>
<td>1.20E-06</td>
<td>1.09E-06</td>
<td>1.01E-06</td>
</tr>
<tr>
<td></td>
<td>(5.58E-08)</td>
<td>(5.99E-08)</td>
<td>(6.00E-08)</td>
<td>(5.90E-08)</td>
<td>(2.64E-06)</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.0002</td>
<td>0.0004</td>
<td>-</td>
<td>-</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>(1.4153)</td>
<td>(0.4272)</td>
<td>-</td>
<td>-</td>
<td>(98.2278)</td>
</tr>
<tr>
<td>$b$</td>
<td>0.1015</td>
<td>0.1018</td>
<td>0.1017</td>
<td>0.1027</td>
<td>0.0439</td>
</tr>
<tr>
<td></td>
<td>(0.0008)</td>
<td>(0.0008)</td>
<td>(0.0008)</td>
<td>(0.0010)</td>
<td>(0.0383)</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>2.2554</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(76.0239)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.0007</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.0043)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$\zeta$</td>
<td>19.4661</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(367.7243)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>2.15E-06</td>
<td>2.15E-06</td>
<td>2.15E-06</td>
<td>-</td>
<td>3.87E-06</td>
</tr>
<tr>
<td></td>
<td>(1.60E-06)</td>
<td>(8.24E-07)</td>
<td>(8.24E-07)</td>
<td>-</td>
<td>(2.33E-05)</td>
</tr>
<tr>
<td>$c$</td>
<td>0.0047</td>
<td>0.0047</td>
<td>0.0046</td>
<td>0.0047</td>
<td>0.0057</td>
</tr>
<tr>
<td></td>
<td>(0.0001)</td>
<td>(0.0001)</td>
<td>(0.0001)</td>
<td>(6.00E-05)</td>
<td>(0.0001)</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.9628</td>
<td>0.9606</td>
<td>0.9603</td>
<td>0.9372</td>
<td>0.9603</td>
</tr>
<tr>
<td></td>
<td>(0.0071)</td>
<td>(0.0070)</td>
<td>(0.0070)</td>
<td>(0.0073)</td>
<td>(0.0073)</td>
</tr>
<tr>
<td>$\tau$</td>
<td>3.00E-07</td>
<td>3.00E-07</td>
<td>3.00E-07</td>
<td>3.10E-07</td>
<td>3.53E-07</td>
</tr>
<tr>
<td></td>
<td>(1.06E-07)</td>
<td>(1.13E-07)</td>
<td>(1.13E-07)</td>
<td>(1.14E-07)</td>
<td>(4.52E-06)</td>
</tr>
<tr>
<td>$Y_0$</td>
<td>3.35E-06</td>
<td>3.35E-06</td>
<td>3.35E-06</td>
<td>1.09E-06</td>
<td>4.88E-06</td>
</tr>
<tr>
<td></td>
<td>(1.60E-06)</td>
<td>(8.26E-07)</td>
<td>(8.26E-07)</td>
<td>(5.90E-08)</td>
<td>(2.34E-05)</td>
</tr>
<tr>
<td>$E_{i}^{co}$</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>$E_{i}^{cc}$</td>
<td>0.3652</td>
<td>0.3651</td>
<td>0.3651</td>
<td>0.3652</td>
<td>0.3747</td>
</tr>
<tr>
<td></td>
<td>(0.0087)</td>
<td>(0.0086)</td>
<td>(0.0086)</td>
<td>(0.0090)</td>
<td>(0.0110)</td>
</tr>
<tr>
<td>$E_{gear}^{2012}$</td>
<td>0.0131</td>
<td>0.0130</td>
<td>0.0130</td>
<td>0.0129</td>
<td>0.0148</td>
</tr>
<tr>
<td></td>
<td>(0.0122)</td>
<td>(0.0121)</td>
<td>(0.0121)</td>
<td>(0.0127)</td>
<td>(0.0155)</td>
</tr>
<tr>
<td>$E_{gear}^{2013}$</td>
<td>-0.0043</td>
<td>-0.0044</td>
<td>-0.0044</td>
<td>-0.0043</td>
<td>-0.0037</td>
</tr>
<tr>
<td></td>
<td>(0.0121)</td>
<td>(0.0120)</td>
<td>(0.0120)</td>
<td>(0.0125)</td>
<td>(0.0153)</td>
</tr>
<tr>
<td>$E_{gear}^{2014}$</td>
<td>0.0172</td>
<td>0.0173</td>
<td>0.0173</td>
<td>0.0173</td>
<td>0.0183</td>
</tr>
<tr>
<td></td>
<td>(0.0122)</td>
<td>(0.0121)</td>
<td>(0.0121)</td>
<td>(0.0126)</td>
<td>(0.0155)</td>
</tr>
<tr>
<td>J-stat</td>
<td>33.0087</td>
<td>38.5897</td>
<td>38.5755</td>
<td>35.3664</td>
<td>112.6926</td>
</tr>
<tr>
<td>df</td>
<td>202</td>
<td>205</td>
<td>206</td>
<td>208</td>
<td>235</td>
</tr>
</tbody>
</table>

Table 2: Estimated parameters (company effect estimates not shown) for four models based on GMM. Standard errors for each parameter are shown in the parentheses.
We conduct model comparison using the difference of $J$ statistics as discussed in the previous subsection. Figure 5 shows the value of the statistics (referred to as “LR”) between models, with arrow pointing toward the more favorable model. Table 3 shows the hypothesis testing results at significance level of 95%. The results suggest strong rejection of w/o Co. effect model, again highlighting the importance of including the company heterogeneity in the model. Among four models with company effects, we always fail to reject simpler model with less parameters. As a result, the preferred specification turns out to be “w/o Trend” model, which has a simple deterministic Gompertz mortality component and selection effect. These results suggest that company effects matters while the catastrophe components and the stochastic diffusion are not important in explaining the prices.
Both our estimation results and the model comparisons point towards the relevance of company effects in the life insurance market, whereas other aspects (mortality catastrophes, stochastic trends) do not seem to be of key relevance. The distribution of the company effects
takes a roughly tri-modal shape as shown in Figure 4, with several company bunching at positive effects (worse mortality experience) around 0.1, some companies bunching around 0 (average mortality experience), and some companies bunching with negative effect (better mortality experience) around -0.1, also seen in Figure 3. The results suggests that there are roughly three segments in the market.

We retrieved information from A.M. Best to check whether these groups can be explained by company characteristics. More precisely, from A.M. Best Rating/Information Services, we obtain company ratings and company size (two companies’ information are not available). Here, financial strength rating means A.M. Best’s independent opinion of “an insurer’s financial strength and ability to meet its ongoing insurance policy and contract obligations” and issuer credit rating refers to A.M. Best’s independent opinion of “an entity’s ability to meet its ongoing financial obligations”. Table 4 provides the information together with the company effects from our estimation. In addition to including the effect, we code the companies according to the three segments, -, o, +, belonging to the group with negative, insignificant, and positive company effects, respectively.

There are no obvious relationships to company size or company ratings. There are very highly rated companies and very large companies in each segment, and there are smaller and relatively low-rated companies in each segment. In other words, company characteristics do not seem to explain the market segmentation—or, more generally, the significant heterogeneity in prices. Thus, it appears that differences in the pools of policyholders between companies must lead to the segmentation.

These differences in the mortality profile of each company’s risk profile may arise from different channels. First, underwriting criteria, and particularly the categorization of individuals with a certain health history to rating classes, differ among carriers. Since in the end the mortality experience is driven by morbidity rates and cause-specific mortality rates, and since these in aggregate shape the mortality profile of the relevant pool, heterogeneity may arise
(Arnold et al., 2016; Arnold and Sherris, 2016, and references therein). Moreover, different companies penetrate different regions dissimilarly and they use different distribution channels, which in turn affect the demographic and socio-economic distribution of policyholders. Since there are large differences in mortality across different subpopulations (Chetty et al., 2016; Case and Deaton, 2015), again these aspects may generate heterogeneity in company mortality profiles.

<table>
<thead>
<tr>
<th>Company Effect</th>
<th>Financial Strength Rating</th>
<th>Long-Term Issuer Credit Rating</th>
<th>Financial Size Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>A- (Excellent)</td>
<td>a- (Excellent)</td>
<td>X ($250 mil ~ $500 mil)</td>
</tr>
<tr>
<td>-</td>
<td>A (Excellent)</td>
<td>a+ (Excellent)</td>
<td>XV (&gt;$2 billion)</td>
</tr>
<tr>
<td>o</td>
<td>A (Excellent)</td>
<td>a+ (Excellent)</td>
<td>XIV ($1.5 bil ~ $2 bil)</td>
</tr>
<tr>
<td>+</td>
<td>A- (Excellent)</td>
<td>a- (Excellent)</td>
<td>IX ($250 mil ~ $500 mil)</td>
</tr>
<tr>
<td>-</td>
<td>A+ (Superior)</td>
<td>aa- (Superior)</td>
<td>IX ($250 mil ~ $500 mil)</td>
</tr>
<tr>
<td>-</td>
<td>A (Excellent)</td>
<td>a+ (Excellent)</td>
<td>VIII ($100 mil ~ $250 mil)</td>
</tr>
<tr>
<td>+</td>
<td>A+ (Superior)</td>
<td>aa- (Superior)</td>
<td>XV (&gt;$2 billion)</td>
</tr>
<tr>
<td>o</td>
<td>A (Excellent)</td>
<td>a+ (Excellent)</td>
<td>XV (&gt;$2 billion)</td>
</tr>
<tr>
<td>o</td>
<td>A+ (Superior)</td>
<td>aa- (Superior)</td>
<td>XV (&gt;$2 billion)</td>
</tr>
<tr>
<td>+</td>
<td>A- (Excellent)</td>
<td>a- (Excellent)</td>
<td>VIII ($100 mil ~ $250 mil)</td>
</tr>
<tr>
<td>-</td>
<td>A+ (Superior)</td>
<td>aa- (Superior)</td>
<td>XV (&gt;$2 billion)</td>
</tr>
<tr>
<td>o</td>
<td>A (Excellent)</td>
<td>a+ (Excellent)</td>
<td>XV (&gt;$2 billion)</td>
</tr>
<tr>
<td>o</td>
<td>A+ (Superior)</td>
<td>aa- (Superior)</td>
<td>XV (&gt;$2 billion)</td>
</tr>
<tr>
<td>o</td>
<td>A (Excellent)</td>
<td>a+ (Excellent)</td>
<td>XV (&gt;$2 billion)</td>
</tr>
<tr>
<td>o</td>
<td>A+ (Superior)</td>
<td>aa- (Superior)</td>
<td>XV (&gt;$2 billion)</td>
</tr>
<tr>
<td>o</td>
<td>A+ (Superior)</td>
<td>aa+ (Superior)</td>
<td>XV (&gt;$2 billion)</td>
</tr>
<tr>
<td>o</td>
<td>A (Excellent)</td>
<td>a+ (Excellent)</td>
<td>VIII ($100 mil ~ $250 mil)</td>
</tr>
<tr>
<td>o</td>
<td>A+ (Superior)</td>
<td>aa- (Superior)</td>
<td>XV (&gt;$2 billion)</td>
</tr>
<tr>
<td>+</td>
<td>A- (Excellent)</td>
<td>a- (Excellent)</td>
<td>VIII ($100 mil ~ $250 mil)</td>
</tr>
<tr>
<td>+</td>
<td>A+ (Superior)</td>
<td>aa+ (Superior)</td>
<td>XV (&gt;$2 billion)</td>
</tr>
<tr>
<td>+</td>
<td>A+ (Superior)</td>
<td>aa (Superior)</td>
<td>XV (&gt;$2 billion)</td>
</tr>
<tr>
<td>+</td>
<td>A+ (Superior)</td>
<td>aa- (Superior)</td>
<td>IX ($250 mil ~ $500 mil)</td>
</tr>
</tbody>
</table>

Table 4: Company effects v.s. A.M. Best ratings and company size (-,o,+, belonging to the group with negative, insignificant, and positive company effects, respectively)
4 Conclusion

Our study of term life insurance prices goes against the common perception that life insurance is fairly homogeneous and that aggregate mortality trends are highly relevant in this marketplace. Our estimates do not reflect significance of mortality catastrophe or stochastic mortality trends. The model comparison strongly rejects the model without company specific mortality effects, however, pointing to a significant heterogeneity in the mortality profiles among different life insurance companies. Thus, “basis risk” dominates systematic mortality risk in life insurance. This questions the efficiency of so-called “natural hedging” of longevity risk using life insurance exposure. Furthermore, it points to the relevance of understanding granular mortality trends.

References


