Renshaw, A. E. & Haberman, S. (1998). Modelling the recent time trends in UK permanent health insurance recovery, mortality and claim inception transition intensities (Report No. Actuarial Research Paper No. 113). London, UK: Faculty of Actuarial Science & Insurance, City University London.

CITY UNIVERSITY LONDON

City Research Online

MODELLING THE RECENT TIME TRENDS IN UK PERMANENT HEALTH INSURANCE RECOVERY, MORTALITY AND CLAIM INCEPTION TRANSITION INTENSITIES

by

A E RENSHAW and S HABERMAN

Actuarial Research Paper No. 113

Department of Actuarial Science and Statistics
City University
London

June 1998

ISBN 1 901615 31 6

"Any opinions expressed in this paper are my/our own and not necessarily those of my/our employer or anyone else I/we have discussed them with. You must not copy this paper or quote it without my/our permission".

MODELLING THE RECENT TIME TRENDS IN UK PERMANENT HEALTH INSURANCE RECOVERY, MORTALITY AND CLAIM INCEPTION TRANSITION INTENSITIES.

By A. E. RENSHAW, Ph.D AND S. HABERMAN, Ph.D, F.I.A.

(of City University, London)

ABSTRACT

Models representing the underlying trends in UK Permanent Health Insurance (PHI) recovery, mortality and claim inception transition intensitives over the twenty year calendar period, form 1975 to 1994 inclusive, are proposed. The investigation of such trends is of special interest given that the three transition intensities, with stationary estimates based on the equivalent grouped data for the quadrennial observation window 1975-78, form an important part of the UK Continuous Mortality Investigation Bureau multiple-state model for PHI business and play an important role in the pricing and reserving for PHI sickness benefits.

KEYWORDS

Permanent Health Insurance; Multi-state transition intensities; Trends

1. INTRODUCTION

The sickness recovery and inception transition intensities, together with the force of mortality when sick, which form the basis of the UK Continuous Mortality Investigation (CMI) Bureau's multiple state model, derive from the pooled experience of leading UK insurance companies in the observation window 1975-78 (CMI, 1991). The introduction of this model represents a milestone in the ways in which permanent heath insurance (PHI) premiums and benefits could be valued, with implications for pricing and reserving, and in which the underlying transition experience could be measured and monitored. In this paper, our objective is to investigate whether any significant time trends can be established in these three fundamental transition intensities, subsequent to 1975-78, and which might have important implications for current practice in terms of the pricing and reserving for PHI sickness benefits. As discussed in Section 4, attempts to model the tends in *sickness* inception transition intensities are not successful but more progress is possible with the *claim* inception intensities. The investigation is made possible by the recent consolidation of the information provided by contributing offices to the PHI experience into a suitable data base. Separate sections (2 to 4) are devoted to each of the three intensities, which, for reasons of convenience, are referred to in places as 'rates'.

2. SICKNESS RECOVERY RATES

2.1 Preliminaries.

In this section we target the sickness recovery transition intensities and for convenience we refer to these as 'rates'. Numbers of recoveries with matching exposures, in the raw data, have been made available by individual weeks for sickness duration, individual years for age at sickness inception and by individual calendar years, 1975 to 1994. This applies for each of five deferred periods of 1, 4, 13, 26, 52 weeks (DP1, DP4, DP13, DP26, DP52), for both males and females. Since the search for time trend patterns in the recovery rates is a primary aim of the ensuing analysis, we have elected not to group the data by calendar year (typically presented in quadrennia), but have adopted instead the following grouping of the raw data prior to their analysis:

gender	male, female
deferred period	1, 4, 13, 26, 52 wks
duration (z)	1- ,2-, 3-, 4-, 8-, 13-, 17-, 26-, 30-, 39-52 wks, 1-, 2-, 5-11 yrs
age (x)	18-, 25-, 30-, 35-, 40-, 45-, 50-, 55-, 60-65 yrs
period (t)	1975, 76, 77, ..., 94

Note that there is a maximum of 13 duration levels subject to the specific deferred period, with 10 levels for DP4, 8 levels for DP13, and so on. This grouping with respect to both duration and age at sickness inception is consistent with that used in recent CMI Bureau commentaries on parts of the data set (see CMI, 1996). For each of the five deferred periods and each gender (10 separate cases), based on this cross-classification, let:

$$r_{txz} = \text{number of recoveries in cell } (t, x, z)$$
$$e_{txz} = \text{exposure (in years) to the possibility of recovery in cell } (t, x, z)$$

as the case may be. Further, for each period t, the data are assumed to be located at the centroids of their respective sub-cells (x, z), determined by weighted averages, with weights based on the relative exposures generated by the raw data. There is a relatively small number of cells with zero exposure on the fringes of the resulting data grids, and these are weighted out of the ensuing analysis. There is also a tendency for recoveries to be somewhat thinly spread across calendar years at the long sickness durations, under this cross-classification, a feature which is discussed later.

For a specific deferred period and specific gender, the recovery rate per year from sickness, or sick to health intensity ρ_{txz}, is targeted by declaring the numbers of reported recoveries in the various data cells to be Poisson response variables:

$$r_{txz} \sim \text{Poi}(e_{txz}\,\rho_{txz}) \text{ independently } \forall \text{ cells } (t, x, z)$$

for which:

$$m_{txz} = \mathrm{E}(r_{txz}) = e_{txz}\,\rho_{txz},\ \ \mathrm{Var}(r_{txz}) = \mathrm{V}(m_{txz}) = m_{txz}$$

where V is the variance function of the associated generalised linear model (GLM). (The scale parameter is one). This is implemented in combination with the log-link predictor relationship:

$$\eta_{txz} = \log m_{txz} = \log e_{txz} + \log \rho_{txz}$$

where η_{txz} denotes the linear predictor and the $\log e_{txz}$ terms are declared as offsets.

2.2 Male, DP1 experience.

Following exploratory analysis using data for individual calendar years, we take as our starting point the model structure:

$$\log \rho_{xz} = \mu + \alpha\sqrt{z} + \beta z + \theta x + \phi x\sqrt{z} + \psi xz$$

where z is in weeks and x is in years. This model has been fitted previously, (see Renshaw & Haberman, 1995), to an earlier version of the 1975-78 data set, which was grouped into a single calendar quadrennium (coupled with a less stringent grouping with respect to sickness duration). The structure is linear in the parameters and, as such, can be readily adjusted to allow for period effects. We begin by fitting this structure separately for each calendar year. This is achieved in stages by sequentially adding in the terms on the RHS of the expression:

$$\log \rho_{txz} = (\mu+\mu_t) + (\alpha+\alpha_t)\sqrt{z} + (\beta+\beta_t) z + (\theta+\theta_t) x + (\phi+\phi_t) x\sqrt{z} + (\psi+\psi_t) xz \qquad (2.1)$$

subject to the constraints:

$$\mu_1 = \alpha_1 = \beta_1 = \theta_1 = \phi_1 = \psi_1 = 0.$$

The full structure effectively involves a total of 120 unknown parameters, where $t = 1, 2, ..., 20$ is used to code the respective calendar years 1975, 1976, ..., 1994.

The resulting analysis of deviance, as the *nested* parameter count is increased sequentially, is reported in Table 2.1. For such Poisson GLMs, the model deviances reported in the second column of the table, are given by:

$$\sum_{t,z,x} \omega_{txz}\left\{r_{txz} \log \frac{r_{txz}}{\hat{m}_{txz}} - (r_{txz} - \hat{m}_{txz})\right\} = \sum_{t,z,x} \omega_{txz}\, d_{txz} \qquad (2.2)$$

where $\hat{m}_{txz}$ denote the predicted number of recoveries (fitted values) under the specific predictor structure and d_{txz} is the contribution to the model deviance from the data cell defined by (t, x, z), with

weights $\omega_{txz} = 0$ if a data cell has zero exposure, otherwise $\omega_{txz} = 1$. The number of degrees-of-freedom reported in the third column, is equal to the sum of these weights minus the number of effective parameters included in the linear predictor structure. The mean deviation based on the *differences* in the deviances of two *nested* model structures, reported in the final column of Table 2.1, is used to assess the statistical significance of the added predictor structure. As a working rule of thumb, values 'of the order 2 or less' imply that the added structure is not statistically significant. On this basis, scrutiny of Table 2.1 (together with a detailed examination of parameter estimates and their standard errors not reproduced here) leads us to set the parameters $\{\theta_t\}$ and $\{\phi_t\}$ equal to zero.

Further, a comparison of the matching components of the deviances which contribute to the 'test' for the parameters $\{\psi_t\}$ reveals that a single data cell contributes an astonishing 18.78 to the figure of 78.3 in the deviance difference reported in the penultimate row of Table 2.1. Additionally, the refitting of the two relevant model structures, with this one data cell omitted, results in a reduction of the resulting mean deviance (based on differences) from the reported value of 4.1 in the Table to 2.5. The relevant data cell, *which has a negligible effect on the other reported mean deviances* in the final column of Table 2.1, can be identified as an *outlier*, involving the reporting of 13 recoveries for duration 5-11 years, age of sickness inception 55-59, in calendar year 1986. There is also evidence, albeit less extreme, of a few other outliers in this region of the data grid (see Table 2.2). We understand that such outliers have been recognised by the CMI Bureau and their effects reported (see p62 CMI, 1996). On this evidence of marginal significance with a single outlier removed, the parameters $\{\psi_t\}$ are also set equal to zero.

We therefore refit the simplified model structure:

$$\log \rho_{txz} = (\mu+\mu_t) + (\alpha+\alpha_t)\sqrt{z} + (\beta+\beta_t)z + \theta x + \phi x\sqrt{z} + \psi xz \qquad (2.1a)$$

and search for possible time trends by plotting the resulting parameter estimates $\{\hat{\mu}_t\}$, $\{\hat{\alpha}_t\}$ and $\{\hat{\beta}_t\}$ against time t. These plots are reproduced in Figure 2.1a. Given the essentially linear nature of these plots, we proceed to simplify the model structure by representing each of the parameter sets as a linear function of time. This is done sequentially, starting by setting $\mu_t = \mathrm{a} + \gamma_1 t$ (with $\mu \mapsto \mu + \mathrm{a}$) and fitting the resulting model structure:

$$\log \rho_{txz} = \mu + (\alpha+\alpha_t)\sqrt{z} + (\beta+\beta_t)z + \gamma_1 t + \theta x + \phi x\sqrt{z} + \psi xz. \qquad (2.1b)$$

The new parameter estimates $\{\hat{\alpha}_t\}$ and $\{\hat{\beta}_t\}$, based on this structure, are plotted against t in Figure 2.1b. Then, on setting $\alpha_t = \mathrm{a} + \gamma_2 t$ (with $\alpha \mapsto \alpha + \mathrm{a}$) and fitting:

$$\log \rho_{txz} = \mu + \alpha\sqrt{z} + (\beta+\beta_t)z + \gamma_1 t + \gamma_2 t\sqrt{z} + \theta x + \phi x\sqrt{z} + \psi xz \qquad (2.1c)$$

the new parameter estimates $\{\hat{\beta}_t\}$ are plotted against time t in Figure 2.1c. Repeating the process, we set $\beta_t = a + \gamma_3 t$ (with $\beta \mapsto \beta + a$) and fit the model structure:

$$\log \rho_{txz} = \mu + \alpha\sqrt{z} + \beta z + \gamma_1 t + \gamma_2 t\sqrt{z} + \gamma_3 tz + \theta x + \phi x\sqrt{z} + \psi xz. \qquad (2.3)$$

The justification for this latter series of model simplifications has been presented in graphical form *viz* Figures 2.1(a,b,c). It would be possible to apply more sophisticated smoothing techniques to the sets of parameters, in recognition of some local variation in the parameter patterns lost under the suggested model simplifications; however, this graphical evidence in support of the overall linear trend patterns in the parameters is felt to be sufficiently compelling to justify the adoption of equation (2.3). A discussion of the deviance profile associated with these predictor simplifications is presented in Appendix I. We will return to these alternative smoothing techniques in a subsequent report.

Finally, noting that the RHS of equation (2.3) is a polynomial in the three variates t, x and $\sqrt{z}$, (an unexpected turn of events, as far as t is concerned), we experiment by introducing additional polynomial terms in the three *variates*, using as criteria a significant reduction in the model deviance *coupled* with the retention of a complete set of statistically significant parameter estimates. On this basis, equation (2.3) is modified to include an additional parameterised term in xt. (We note that this induces a reduction of 5.9 in the mean deviance, thereby more than matching the overall loss of 4.7 in the mean deviance discussed in Appendix I.) This leads finally to the adoption of the model structure:

$$\log \rho_{txz} = \mu + \alpha\sqrt{z} + \beta z + \gamma_1 t + \gamma_2 t\sqrt{z} + \gamma_3 tz + \theta x + \phi x\sqrt{z} + \psi xz + \kappa xt \qquad (2.4)$$

Before discussing the implications of this model, we note the relatively low numbers of reported recoveries at the higher durations, and draw attention to the acknowledged poor quality of certain data points involving sickness duration 5-11 years. These features are evident in Table 2.2, in which the recovery counts over all 20 calendar years, matched for age at sickness inception and sickness duration, are presented. Subject to this caveat we proceed to interpret the model. In so doing, we have chosen not to highlight the predictions for sickness duration 5-11 years.

The parameter estimates, together with their standard errors and t-statistics are presented in Table 2.3. Diagnostic checks of the model structure are conducted using deviance residuals, defined by:

$$\text{sign}(r_{txz} - \hat{m}_{txz})\, \omega_{txz} \sqrt{d_{txz}}$$

where d_{txz} are the components of the model deviance, defined in equation (2.2). Specifically, deviance residuals plotted against the index (or counter):

$$index(t, z^{l}) = z^{l} + z_{*} \times (t - 1)$$

based on duration *categories* $z^{l} = 1, 2, \ldots, z_{*}$, ($z_{*} = 13$ for DP1, $z_{*} = 10$ for DP4, etc), serialised by calandar year $t = 1, 2, \ldots, 20$, for each separate age category, are especially informative. Thus, the first z_{*} points on the index represent the full range of possible sickness durations, arranged in increasing order, for 1975, then for 1976, and so on. Such plots are reported in Appendix II.

The structured model (2.4) may be interpreted as a three dimensional surface in t, x, and $\sqrt{z}$. This may be viewed from a number perspectives, which include the following:

PERSPECTIVE 1:

$$\log \rho_{txz} = A_{xz} - B_{xz}\, t$$

where:

$$A_{xz} = \mu + \alpha\sqrt{z} + \beta z + \theta x + \phi x\sqrt{z} + \psi xz, \quad B_{xz} = -\gamma_1 - \gamma_2\sqrt{z} - \gamma_3 z - \kappa x$$

PERSPECTIVE 2:

$$\log \rho_{txz} = A_{tx} + B_{tx}\sqrt{z} + C_{tx}\, z$$

where:

$$A_{tx} = \mu + \gamma_1 t + \theta x + \kappa xt, \; B_{tx} = \alpha + \gamma_2 t + \phi x, \; C_{tx} = \beta + \gamma_3 t + \psi x$$

PERSPECTIVE 3:

$$\log \rho_{txz} = A_{tz} - B_{tz}\, x$$

where:

$$A_{tz} = \mu + \alpha\sqrt{z} + \beta z + \gamma_1 t + \gamma_2 t\sqrt{z} + \gamma_3 tz, \; B_{tz} = -\theta - \phi\sqrt{z} - \psi z - \kappa t$$

Perspective 1, which focuses on the model prediction that, for fixed (x, z), the log recovery rates have changed linearly in time over the 20 year calendar period, is of particular interest. The predicted values of (A_{xz}, B_{xz}) are presented in Table 2.4 Note that the behaviour of the signs of B_{xz} over the (x, z) grid is of special interest, since these dictate the cells for which the predicted recovery rates have increased or decreased over the period concerned. On writing the coefficient:

$$B_{xz} = -\gamma_3(\sqrt{z})^2 - \gamma_2\sqrt{z} - (\gamma_1 + \kappa x)$$

as a quadratic in $\sqrt{z}$ (*c.f.* $a(\sqrt{z})^2 + b\sqrt{z} + c$, with negative a and positive discriminant $b^2 - 4ac > 0$ for all x, see Table 2.3 for parameter values), it follows that B_{xz} is positive, and hence the predicted recovery rates decrease, for values of $\sqrt{z}$ between the roots of the quadratic. The corresponding limits for duration z are given in Table 2.5. In essence, these imply that recovery rates have increased over

time for durations of 4 weeks and under (3 weeks and under for ages in excess of 45-49 years) and for durations 300-315 weeks and over, but otherwise have been in decline during the period of investigation. Predicted log recovery rates, plotted against calendar period, for specific ages at sickness inception are reproduced in Figure 2.2.

Perspective 2 focuses on the log recovery rate, viewed as a function of sickness duration, for fixed (t, x). It is a quadratic in $\sqrt{z}$, with C_{tx}, the coefficient of $(\sqrt{z})^2$, positive for all realistic values of (t, x). Hence the quadratic is convex, with a turning point that is a minimum. A typical family of quadratics for each t, fixed x, is illustrated in Figure 2.3, in which the turning point lies 'off the chosen scale'.

Perspective 3, focusing on the prediction that log recovery rates have changed linearly with age at sickness inception, for fixed (t, z), is similar in detail to Perspective 1, with the roles of x and t reversed. Again the signs of:

$$B_{tz} = -\psi(\sqrt{z})^2 - \phi\sqrt{z} - (\theta + \kappa t)$$

determine whether the log recovery rates increase or decrease linearly (*c.f.* $a(\sqrt{z})^2 + b\sqrt{z} + c$ implies positive a and negative discriminant $b^2 - 4ac < 0 \ \forall \ t$, see Table 2.3 for parameter values). Hence $B_{tz} > 0$ for all (t, z) and the predicted log recovery rates decrease linearly with age at sickness inception, fixed (t, z).

2.3 Male, other deferred periods.

For each of the other deferred periods, analysis shows that a different model from the DP1 case is needed, with the $\sqrt{z}$ term no longer playing a significant role. We take as our starting point the model structure:

$$\log \rho_{txz} = (\mu+\mu_t) + (\beta+\beta_t)z + (\theta+\theta_t)x + (\zeta+\zeta_t)(z - z_0)_+ + (\psi+\psi_t)xz \qquad (2.5)$$

subject to constraints:

$$\mu_1 = \beta_1 = \theta_1 = \zeta_1 = \psi_1 = 0,$$

where $(z - z_0)_+ = z - z_0$ if $z > z_0$, and $(z - z_0)_+ = 0$ if $z \leq z_0$. It is a modified version, with period adjustments in the parameters, of the structure:

$$\log \rho_{txz} = \mu + \beta z + \theta x + \xi(z - z_1)_+ + \zeta(z - z_2)_+ + \psi x(z - z_1)_+$$

fitted to earlier versions of the DP4, DP13 and DP26 data sets, for the 1975-78 quadrennium, (see Renshaw & Haberman, 1995). Here we set the first of the knots $z_1 = 0$ ($\beta \mapsto \beta + \xi$), since its presence is effectively made redundant once the first 4 weekly durations for which sickness benefit becomes payable are grouped into a single category (for DP4, DP13, DP26), prior to the analysis of

these data. We note that the knot z_1 was designed (Renshaw & Haberman, 1995) to cater for the reported (CMI, 1991) sluggish recovery rates in the first few individual weeks immediately after sickness benefit becomes payable, for the quadrennial 1975-78 male experience in which the data were edited by weekly duration prior to analysis.

The separate analysis of each of the four cases follows a near identical pattern, subject to decreasing overall exposure with increasing (fixed) deferred period. Firstly, the optimum position of the knot z_0 is determined by the repeated fitting of model structure (2.5), with the knot positioned at the centre of a different duration category each time. The resulting deviance profiles are reported in Table 2.6. As a consequence of these results the ensuing analysis is based on setting the knot $z_0 = 78$, for each deferred period. (This compares with the setting of the second knot at 34.5 weeks duration by Renshaw & Haberman, 1995). Next, the deviance profiles, as the terms on the RHS of equation (2.5) are included sequentially, are reported in Tables 2.7(a to d). Here, we have arbitrarily elected to fit the terms in a different sequence to that reported in equation (2.5), introducing all five terms with period effects ignored in the first instance, and then adding in the period effects. The conclusions based on these deviance profiles, are remarkably similar in all four cases, leading to the adoption of the simplified model structure:

$$\log \rho_{txz} = \mu + \mu_t + \beta z + \theta x + \zeta(z - z_0)_+ + \psi xz \qquad (2.5a)$$

for DP4, DP13, DP26, and with additionally $\mu_t = 0$ for DP52. Finally since the ensuing $\{\mu_t\}$ time patterns, reproduced in Figure 2.4, are essentially linear, we set $\mu_t = a + \gamma t$ ($\mu \mapsto \mu + a$) and adopt the model structure:

$$\log \rho_{txz} = \mu + \beta z + \theta x + \gamma t + \zeta(z - z_0)_+ + \psi xz \qquad (2.6)$$

for DP4, DP13, DP26. The deviance profiles associated with this final modification are reported in Appendix I.

The parameter estimates, together with their standard errors and t-statistics, for all four cases, are presented in Table 2.8. Note that the structure (2.6) has also been fitted for the DP52 experience, although the period effects are statistically insignificant, a feature confirmed by the value of the t-statistic of the relevant parameter. Diagnostic checks are again based on deviance residual plots but are omitted for reasons of economy. They can be made available on request.

Equation (2.6) may be viewed from the same three perspectives as equation (2.4), namely:

PERSPECTIVE 1:

$$\log \rho_{txz} = A_{xz} - B_{xz} t$$

where:

$$A_{xz} = \mu + \beta z + \zeta(z - z_0)_+ + \theta x + \psi xz, \quad B_{xz} = -\gamma$$

PERSPECTIVE 2:

$$\log \rho_{txz} = A_{tx} + B_{tx}z + C_{tx}(z - z_0)_+$$

where:

$$A_{tx} = \mu + \theta x,\ B_{tx} = \beta + \psi x,\ C_{tx} = \zeta$$

PERSPECTIVE 3:

$$\log \rho_{txz} = A_{tz} - B_{tz}x$$

where:

$$A_{tz} = \mu + \beta z + \zeta(z - z_0)_+ + \gamma t,\ B_{tz} = -\theta - \psi z$$

Under perspective 1, B_{xz} is positive for all four deferred periods, so that the predicted log recovery rates decrease linearly, for fixed (x, z), over the period under investigation. This feature is illustrated in Figure 2.5.

Under perspective 2, $B_{tz} < 0$ and the predicted log recovery rates decrease linearly with increasing sickness duration, for fixed (t, x), but less rapidly so for durations in excess of 78 weeks. Figure 2.6 illustrates this.

Under perspective 3, $B_{tz} > 0$ and the predicted log recovery rates decrease linearly with age at sickness inception, for fixed (t, z).

2.4 Female, DP1 experience.

With no prior knowledge of any analysis of the female PHI experience, (for example, from the 1975-78 investigation) we follow the same approach as that adopted in Section 2.2 for the male experience, while noting that the female experience is much more sparse than the corresponding male experience. The deviance profile associated with the sequential fitting of the terms on the RHS of expression (2.1) is reported in Table 2.9. While the final column of this table reveals support for all five parameters α, β, θ, ϕ and ψ, the experience is too sparse to establish statistically significant time patterns in the sets of parameters with the possible exception of the $\{\mu_t\}$. However, scrutiny of the parameter estimates under the fitted model structure:

$$\log \rho_{txz} = (\mu+\mu_t) + \alpha\sqrt{z} + \beta z + \theta x + \phi x\sqrt{z} + \psi xz$$

reveals no decernible time pattern in the estimated μ_ts coupled with large standard errors for these parameters and mainly non-significant t-statistics. The details accordingly are not reproduced here. Although the conclusion to be drawn from this analysis is that no firm statistically significant time pattern can be established by this approach (but that a useful insight into female recovery rates 'averaged' over the whole 20 year period can be obtained), we present the detail for the fitted model

structure:

$$\log \rho_{txz} = \mu + \alpha\sqrt{z} + \beta z + \gamma t \ + \theta x + \phi x\sqrt{z} + \psi xz \qquad (2.7)$$

in which we have set $\mu_t = \gamma t$, and which can be rewritten as:

PERSPECTIVE 1:

$$\log \rho_{txz} = A_{xz} - B_{xz}\, t$$

with:

$$A_{xz} = \mu + \alpha\sqrt{z} + \beta z + \theta x + \phi x\sqrt{z} + \psi xz, \ \ B_{xz} = -\gamma.$$

The parameter estimates, standard errors, and t-statistics reported in Table 2.10. These indicate that the period effect is not strictly statistically significant, but there is some weak evidence to support the general statement that recovery rates have increased, if anything, over the 20 year period concerned. The quality of the fit has again been assessed by the customary residual plots associated with the fitted structure, and these can be made available on request.

2.5 Female, other deferred periods.

Structure (2.6), fitted to the equivalent male experiences in Section 2.3, has been fitted to the female DP4, DP13 and DP26 experience but with the product term in xz omitted, (since this term proved to be statistically non-significant in this case). The optimum position of the knot, again set at 78 weeks in all three cases, is verified by reference to the deviance profiles reported in Table 2.11, this time constructed by the repeated fitting of the adopted model structure:

$$\log \rho_{txz} = \mu + \beta z + \theta x + \gamma t + \zeta(z - z_0)_+$$

which again we write as:

PERSPECTIVE 1:

$$\log \rho_{txz} = A_{xz} - B_{xz} t$$

with:

$$A_{xz} = \mu + \beta z + \zeta(z - z_0)_+ + \theta x, \ \ B_{xz} = -\gamma$$

in order to highlight the dependence on t. Details of the parameter estimates are reported in Table 2.12. Since the estimated value of γ is negative for each deferred period, there is an implied deterioration in the recovery rates over the 20 year period concerned. Again the relevant residual plots are reproduced in Appendix I.

2.6 Discussion of results.

It is of interest to compare the magnitudes of B_{xz}, the rates by which the predicted log-recovery intensities change with time, across all deferred periods where feasible (see perspective 1 in each case). With the exception of the male DP1 experience, these rates of change with time are particularly simple, with B_{xz} = -γ, a constant $\forall$ (x, z). Consequently, it is a simple matter to compare rates in these cases by tabulating the values of -B_{xz} = γ, say, taken from Tables 2.8, 2.10 and 2.12. Thus:

	DP1	*DP4*	*DP13*	*DP26*	*DP52*
males	o	-0.033	-0.034	-0.037	-0.022†
females	0.0050†	-0.026	-0.036	-0.040	*

† not statistically significant * depleted experience o see comments below

with a *negative* sign implying decreasing recovery intensities over time. The similarity of the results for the male DP4, DP13 and DP26 experience, together with the female DP13, DP26 experience, and to a lesser extent the female DP4 experience, is noteworthy. The situation is somewhat more complex for the male DP1 experience, with B_{xz} dependent on (x, z). Essentially for this experience, recovery intensities decrease with time for durations in excess of 4 weeks but less that 6 years, subject to relative small changes in these limits with age at sickness inception, as given in Table 2.5; while values of B_{xz} are stated in Table 2.4. A possible explanation for the increase in recovery rates with time at these short sickness durations for DP1 policies could be the improved management of claims within insurers and a move to "active intervention in newly admitted claims". These changes are considered to have led to fewer "marginal cases" being accepted as new claims; to accepted claims being managed more actively at the short durations, leading to increased short duration recoveries; and to the residual claims, surviving this "initial, active management stage", being the more problematic and long term cases. These effects would be particulary noticeable in respect of shorter deferred policies, as these "offer greater opportunity for early intervention in newly notified claims". For longer deferred periods, claims would be notified somewhat later and be more "established" by the time that the insurer's claims management process can intervene. This could explan the difference in the nature of the results for DP1 compared to the longer deferred periods (Heeney, 1998).

3. Mortality From Sickness Transition Intensities

In this section we focus on the force of mortality when in the sick state. Because of the relatively low reported incidence of sick to death transitions, in keeping with previous work by the CMI Bureau on such transitions, the deaths and exposures in matching cells are combined by summation over all deferred periods (DP1, DP4, DP13, DP26, DP52), and the cross-classified data cells are defined by:

gender (g)	male, female	
duration (z)	1-, 4-, 8-, 13-, 17-, 26-, 30-, 39-52 wks, 1-, 2-, 5-11 yrs	(11 levels)
age (x)	18-, 35-, 40-, 45-, 50-, 55-, 60-65 yrs	(7 levels)
period (t)	1975-78, 79-82, 83-86, 87-90, 91-94	(5 quadrennia)

with calendar years grouped as quadrennia, unlike Section 2. Hence the raw data used in this analysis comprise:

$$d^g_{txz} = \text{number of deaths in sick cell } (g,\ t,\ x,\ z)$$
$$e^g_{txz} = \text{exposure (in years) to risk of death in sick cell } (g,\ t,\ x,\ z).$$

The force of mortality when sick (or mortality from sickness transition intensity) ν^g_{txz} is targeted by declaring the numbers of reported deaths in the various data cells to be independent Poisson response variables:

$$d^g_{txz} \sim \text{Poi}(e^g_{txz}\,\nu^g_{txz}) \text{ independently } \forall \text{ cells } (g,\ t,\ x,\ z)$$

for which:

$$m^g_{txz} = \text{E}(d^g_{txz}) = e^g_{txz}\,\nu^g_{txz},\ \ \text{Var}(d^g_{txz}) = \text{V}(m^g_{txz}) = m^g_{txz}$$

where V is the variance function of the associated GLM. (The scale parameter is one). This is implemented in combination with the log-link predictor relationship:

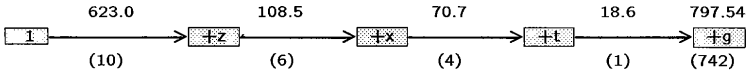
$$\eta^g_{txz} = \log m^g_{txz} = \log e^g_{txz} + \log \nu^g_{txz}$$

where η^g_{txz} denotes the linear predictor and the $\log e^g_{txz}$ terms are declared as offsets. Since the primary objective is to detect any patterns supported by the data, particularly in relation to gender and calendar period, as opposed to the detailed construction of graduated values for ν^g_{txz}, all four covariates (gender, period, age, duration) are initially modelled as categorical factors. In particular, the main effects structure takes the parametric form:

$$\log \nu^g_{txz} = \mu + \alpha_g + \beta_t + \gamma_x + \delta_z. \tag{3.1}$$

It is of interest to note that this is an extension, involving additional calendar period and gender effects, of the multiplicative structure (under the inverse log-link) investigated by Bayliss (1991) in relation to the 1975-78 male experience, and which was identified as a Poisson GLM with a log-link by Renshaw & Haberman (1995).

One of the possible ways in which the factors can be added sequentially to the predictor structure is recorded in the following diagram, reading from left to right:



where:

$$\boxed{1} \Rightarrow \log \nu^{g}_{txz} = \mu, \quad \boxed{+z} \Rightarrow \log \nu^{g}_{txz} = \mu + \delta_z, \text{ etc.}$$

The nodes therefore represent the various predictor structures, while the reduction in both the deviance and the number of degrees of freedom are recorded on the connecting branches of the lattice. As an approximation, an indication of the level of statistical support for all four main effects terms is obtained by referring the deviance differences to the appropriate chi-square distribution. The alternative is to focus on the mean deviance based on differences, as in Section 2. All four differences are highly significant on this basis.

The feasibility of incorporating additional paired interactions terms into the predictor structure was subsequently investigated and quickly abandoned, owing to the paucity of recorded deaths from sick in many of the cross-classified data cells, especially for females. Indeed the total number of recorded female deaths across all cells making up the five quadrennia are as follows:

1975-78	1979-82	1983-86	1987-90	1990-94
19	19	28	41	58

Deviance residual plots generated by the main effect model structure are presented in Appendix II. These comprise plots of the residuals against an age/duration index, separately for males in each of the five quadrennia, and for all females combined over all 20 calendar years. The index is defined as:

$$\text{index}(x, z) = z_i + 11 \times (x_i\text{-}1)$$

where z_i, with values 1 to 11, identify the eleven duration groups arranged in increasing order; and similarly x_i, with values 1 to 7, identify the seven age groups arranged in increasing order. Thus, the first eleven points on the index represent the full range of possible sickness durations, arranged in increasing order of magnitude, for ages 18-34, and so on.

The parameter estimates for the main effects structure, equation (3.1), make interesting reading. The details are presented in Table 3.1, The fitted model predicts that mortality from the sick state:

(a) deteriorates with increasing age,

(b) is lower in females than males,

(c) improves progressively over the five quadrennial calendar periods,

(d) is esentially bell shaped with duration, deteriorating up to duration 26-29 weeks,

followed by a relative improvement as duration further increases, with an apparent final upturn in mortality at very long durations.

These effects are graphically illustrated in Figure 3.1, in which the predicted log mortality rate $\log \nu^g_{txz}$, is plotted against duration (on the categorical scale z_i), for each age band x_i, for the 1975-78 male experience ($\beta_1 = 0$, $\alpha_g = 0$). We recall that the 1975-78 male experience formed the basis for the construction of the PHI standard model in current use in the UK. It is of interest to note that the general shape of the graphs in Figure 3.1 is consistent with the so-called duration factor profile, Figure B6 of Bayliss (1991), based on that author's analysis of the 1975-78 male experience. In addition, the age factor profile in Figure B7 of Bayliss (1991) is consistent with the pattern in the age effects parameter estimate $\hat{\gamma}_x$ of Table 3.1, with again statistically insignificant effects for ages under 40 years, as implied for the 1975-78 experience on p 36 of Bayliss (1991). The parallel profile representing each of the age bands in Figure 3.1 is a feature of the additive, non-interactive, nature of the linear predictor structure. Identical profiles, constructed by moving (*lowering* since both the estimated α_gs and the estimated β_ts are negative) the whole configuration vertically relative to the ordinate by an amount $\alpha_g + \beta_t$, apply for the nine remaining period/gender combinations under study. These profiles imply an improvement in mortality on an age/duration cell basis relative to the 1975-78 male experiance. It is also of interest to note, that while the CMI (1991) model is based on ν^g_{txz} depending on x (attained age) only, for sickness duration sickness duration z in excess of 5 years, it follows from Table 3.1 that the parameter estimates for durational main effects, $\hat{\delta}_z$, are of borderline statistical significance, for durations in excess of 1 year, that is $\hat{\delta}(j)$ for $j = 9, 10, 11$.

Given the patterns in the sets of parameter estimates (Table 3.1), it is further possible to reduce the number of parameters utilised by representing all three of the covariates- period, age and duration - as continuous variates and incorporating them into the formula:

$$\log \nu^g_{txz} = \mu + \alpha_g + \gamma t + \theta x + \beta z + \sum_{j=1}^{k} \beta_j (z - z_j)_+ \qquad (3.2)$$

where:

$$\alpha_1 = 0 \text{ (for males)}$$

and:

$$(z - z_j)_+ = \begin{cases} z - z_j, & z > z_j \\ 0, & z \leq z_j. \end{cases}$$

Here period t is coded from 1 to 5 (to match the respective quadrennia), while age x (in years) and z (in weeks) are allocated values at the centres of the relevant data cells defined at the outset of this section. In equation (3.2) duration effects are represented by hinged line segments, with k hinges or knots positioned at $z_j \in \{2.5, 6, 10.5, 15, 21.5, 28, 34.5, 45.5, 78, 182, 416\}$. A similar structure with $\alpha_2 = \gamma = 0$, $k = 2$, was fitted to an earlier version of the 1975-78 males experience, Renshaw &

Haberman (1995). The deviance profile for model (3.2), based on two knots, is reproduced in Table 3.2. This suggests the optimum knot settings (10.5, 182) weeks. The siting of knots at these two positions is otherwise implied from Figure 3.1. Further analysis based on the introduction of an additional knot, leads us finally to report the detail of the three knot model, with knots positioned at $k_1 = 10.5$, $k_2 = 78$ and $k_3 = 182$ weeks. This achieves a further reduction of 11.12 in the deviance for the loss of a single degree of freedom, while all parameters remain highly significant statistically. Details of the fit are reported in Table 3.3 and the resulting predicted log mortality rate is plotted against duration, this time measured in weeks, for each age, for the 1975-78 male experience (t =1, α_g = 0) in Figure 3.2. The overall patterns in the corresponding residuals are similar in nature to those associated with model (3.1), as reported in Appendix II, and are hence not reproduced. The conclusions to be drawn from model (3.2) are similar to those drawn from model (3.1), that is (a) to (d) as presented above.

4. Claim Inception Intensities

4.1 Preliminaries.

In this section we report on the claim inception intensities. The raw data comprise claim inception counts with matching exposures, cross-classified according to the format:

gender	male, female
deferred period	1, 4, 13, 26, 52 wks
age at inception (x)	18-, 25-, 30-, 35-, 40-, 45-, 50-, 55-, 60-65 yrs
calendar time at inception (t)	1975, 76, 77, ..., 94.

For each gender in combination with each of the five deferred periods, let:

$$i_{tx} = \text{claim inception count in cell } (t, x)$$
$$e_{tx} = \text{matching exposure in cell } (t, x),$$

as the case may be. Thus, i_{tx} represents the number of sicknesses which start in the observation cell (t, x) *and* which last beyond the deferred period of the policy. Exposures are measured in years and correspond to EH_x on page 3 of CMI (1996). We note that these quantities do not include any allowance for the exposure time spent as sick but not claiming - unlike CMI (1991) and Renshaw & Haberman (1995) who make adjustments for this factor. However, as reported by CMI (1996), the effect of ignoring this factor is only "to overstate the exposure by about 0.5%".

The claim inception transition intensity τ_{tx} is targeted for specific gender and deferred period

combinations. This is done by declaring the numbers of reported incidents of claims in the various data cells to be independent over-dispersed Poisson response variables, such that:

$$m_{tx} = \mathrm{E}(i_{tx}) = \mathrm{e}_{tx}\,\tau_{tx},\ \ \mathrm{Var}(i_{tx}) = \phi\,\mathrm{V}(m_{tx}) = \phi\,m_{tx}$$

with variance function V, scale parameter ϕ. The scale parameter is included in recognition of the presence of duplicates amongst the policies contributing to the data base. Such responses are implemented in combination with the log-link predictor relationship:

$$\eta_{tx} = \log m_{tx} = \log \mathrm{e}_{tx} + \log \tau_{tx}$$

where η_{txz} denotes the linear predictor and $\log \mathrm{e}_{tx}$ the offset.

It is informative to relate the targeting of claim inception transitions τ_{tx} to that of targeting sickness inception transitions σ_{tx}, since the latter play a more fundamental role in the PHI multiple state model. It is clear that $\tau_{tx} = \pi_{txd}\,\sigma_{tx}$, where π_{txd} denotes the probability that the time spent in the sick state exceeds the relevant deferred period of d (=1, 4, 13, 26, 52) weeks. Hence, when targeting σ_{tx} as opposed to τ_{tx}, it would first be necessary to determine values for π_{txd} so that $\log \pi_{txd}$ may be added to the offset term prior to model fitting. It is possible to do this by evaluating the integral:

$$\pi_{tx\omega} = \exp - \int_0^{\omega/52} (\rho_{t+u,x+u,u} + \nu_{t+u,x+u,u})\,\mathrm{d}u$$

However the resulting predicted sickness inceptions are sensitive to the values of π_{txd}, a feature noted in Renshaw & Haberman (1995), and we have elected to model and report on trends in the claim inception rates which we believe are also of great interest to practitioners, in particular for the measurement and monitoring of emerging claims experience. Others have followed this route of attempting to model directly the claim inception rates: see, for example, Dillner (1969) and Haberman and Walsh (1998), but see Appendix III for further comments on the evaluation of the integral expression for $\pi_{tx\omega}$.

Since our primary objective is to search for possible time trends in the claim inception intensities, we begin, as in Section 3, by representing both age at claim inception x and period t effects as factors, giving rise to the nested main effects model structures:

$$\log \tau_{tx} = \mu$$

$$\log \tau_{tx} = \mu + \alpha_x;\ \ (\alpha_1 = 0) \qquad (4.1)$$

$$\log \tau_{tx} = \mu + \alpha_x + \beta_t;\ \ (\alpha_1 = \beta_1 = 0) \qquad (4.2)$$

where the fully interactive model structure:

$$\log \tau_{tx} = \mu + \alpha_x + \beta_t + (\alpha\beta)_{xt};\ (\alpha_1 = \beta_1 = (\alpha\beta)_{1t} = (\alpha\beta)_{x1} = 0) \tag{4.3}$$

constitutes the saturated model. If adopted, this structure would imply that no modelling takes place and that we simply interpret the empirical inception rates calculated directly using the raw data.

The order in which the x and t effects are incorporated into the model structure is essentially arbitrary, but we have elected to include x effects before t effects since the sickness inception rates are known to vary with the former, see for example CMI (1991) and Renshaw and Haberman (1995). The resulting deviance profiles, for each gender/DP combination (with the exception of the female DP52 experience for which the data are too sparse), are presented in Table 4.1 for males and Table 4.2 for females. In these tables, the scale parameter ϕ is estimated by dividing the deviance by the degrees-of-freedom, for the model structure concerned. The so-called F-statistics reported in the final column of the tables are computed by dividing the mean deviance, based on differences, by the appropriate scale parameter. For normally distributed response models, such statistics have an exact F- distribution but not otherwise, and are interpreted merely on the basis of analogy with the interpretation of the well known ANOVA tables for normal response models. The 'F- statistics' need to be interpreted in conjunction with monitoring of the residual plots and scrutiny of the standard errors of the parameter estimates.

4.2 Male experiences.

For the male experiences, (Table 4.1) there are statistically significant main periods effects as well as main age at claims inception effects, for each DP. However, there is a marked difference in the pattern of 'F- statistics' for DP1 compared with all four DPs in excess of one week. In addition, the scale parameter is excessively high for DP1. This proves to be indicative of a poor fit, a feature confirmed by the residual plots for this model. Pronounced patterns in the residuals when plotted against t for certain ages hint at a degree of interaction between period and age effects.

For DP1 only, and as a consequence of experimenting with various partially interactive structures (recall that the fully interactive model, equation (4.3), is saturated), we focus on the partially interactive structure:

$$\log \sigma_{tx} = \alpha_x + \beta_{x't} \tag{4.3a}$$

where x' is a declared factor with three levels generated by clustering ages 18-24, 25-29 into one level, ages 30-34, 35-39, ..., 55-59 into a second level, and ages 60-65 into the third level. While acknowledging that such a structure is grossly over-parameterised, we note that our sole objective in this exploratory stage of analysis is to search for possible time trends in the claim inception rates, and we merely report the resulting predicted log claims inception rates in graphical form, Figure 4.1. The nature of the fitted structure is immediately obvious from Figure 4.1, comprising an additive non-

interactive structure in age and period effects within each age cluster, but incorporating interactions between the age clusters x'. Also note that since the third age cluster is composed of a single age grouping (60-65 years), the predicted values for these ages are equal to the crude inception rates, a feature reflected by the associated residuals, which are all zero. Figure 4.1 shows that, even with the interactive structure proposed, there is evidence of a minimum (perhaps only local) in the time trend of the claim inception rates in 1981 and of a maximum in 1988.

As already noted, the situation regarding each of the DP male experiences in excess of one week is relatively consistent and straight forward, leading to the adoption of the additive main effects structure under the log link, equation (4.2). Aspects of the predicted log claim inception rates are illustrated in Figure 4.2 and Figure 4.3. In Figure 4.2, predicted $\log \tau_{tx}$ values are plotted against t for DP4, DP13, DP26 and DP52, in each of two frames, with each frame representing a specific age at claim inception. In particular, note that the relative (vertical) displacements of the four curves between frames (ages) is indicative of the additive non-interactive nature of the structure on the log scale. Figure 4.2 indicates that the time variation for these modelled claim inception rates is smaller than for DP1. There is also evidence of a minimum for the trend for each deferred period in about 1982 and a subsequent (albeit less marked) maximum, the timing of which depends on the deferred period. In Figure 4.3, predicted $\log \tau_{tx}$ values are plotted against x for DP4, DP13, DP26 and DP52, in each of four frames, with each frame representing a specific calendar year. Again, the additive non-interactive nature of the structure is in evidence on comparing frames.

4.3 Female experiences.

For the female experiences, Table 4.2, there is scant supportive evidence of statistically significant period effects in the data by these methods, a conclusion supported by the parameter esimates and their standard errors under model (4.2). Consequently we report the predicted claim inception rates σ_x by x for each DP (DP52 excepted) based on model (4.1), in Table 4.3.

5. Further Work

We have identified a number of areas worthy of further investigation:

(a) consideration of the correlation between trends identified in Sections 2 - 4 (for example, depicted in Figures 2.1, 2.3, 4.1, 4.3) and the underlying economic variables (as attempted by Haberman and Walsh, (1998)), allowing for changes in the eligibility conditions and levels of DSS sickness and invalidity benefits and in taxation of DSS and PHI benefits;

(b) consideration of the recoveries split into two broad categories by type of disability

  i) all musculoskeletal, mental and nervous disorders

  ii) all other types

in order to investigate the hypothesis that type i) would be more susceptible to economic trends. This is dependent on the relevant data being provided by the CMI Bureau;

(c) comparison of mortality trends identified in Section 3 with those reported by the CMI Bureau for assured lives or annuitants;

(d) investigation of the sensitivity of the results to the exclusion of the data at the longest sickness durations;

(e) application of alternative smoothing techniques to the model parameters, including a time series approach to the modelling of parameter trends (as in McNown and Rogers (1989) for example).

ACKNOWLEDGEMENTS

The authors wish to thank the CMI Bureau for providing both financial support and the data, and in particular for generously giving of its time to respond to various queries concerning the data. The authors would also like to thank David Heeney, Mark Hills and Duncan Walsh (members of the PHI Working Party on Trends in Claims Experience) for helpful comments on earlier drafts.

REFERENCES

BAYLISS, P.H. (1991). The graduation of claim recovery and mortality intensities. *CMIR,* 12, 21-50.

CMI Committee (1991). The analysis of Permanent Health Insurance data. *CMIR,* 12, 1-263.

CMI Committee (1996). Recovery rates under PHI policies, individually 1975-90 and grouped 1975-86. *CMIR,* 15, 1-50.

DILLNER, C.G. (1969). New bases for non-cancellable sickness insurance in Sweden. *Skandinavisk Aktuarietidskrift,* 52, 113-124.

HABERMAN S. & WALSH, D.J.P. (1998). Analysis of trends in PHI claim inceptions data. *Trans 26th International Congress of Actuaries,* 6, 159-186.

HEENEY, D. (1998). *Personal communication.*

MCNOWN, R. & ROGERS, A. (1989). Forecasting mortality: A parameterised time series approach. *Demography,* 26, 645-660.

RENSHAW, A.E. & HABERMAN S. (1995). On the graduations associated with a multiple state model for permanent heath insurance. *Insurance: Mathematics and Economics,* 18, 1-17.

Model terms	*Deviance*	*D.F.*	*Difference*	*D.F*	*Mean deviance*
μ	87,386	2113			
			4,009	19	211.0
$+\mu_t$	83,377	2094			
			76,167	1	76,167
$+\alpha\sqrt{z}$	7,209.7	2093			
			281.1	19	14.8
$+\alpha_t\sqrt{z}$	6,928.6	2074			
			904.8	1	904.8
$+\beta z$	6,023.8	2073			
			149.5	19	7.9
$+\beta_t z$	5,874.3	2054			
			3,175	1	3,175
$+\theta x$	2,699.6	2053			
			32.6	19	1.7†
$+\theta_t x$	2,667.0	2034			
			12.3	1	12.3
$+\phi x\sqrt{z}$	2,654.7	2033			
			27.7	19	1.5†
$+\phi_t x\sqrt{z}$	2,627.0	2014			
			13.7	1	13.7
$+\psi xz$	2,613.3	2013			
			78.3	19	4.1
$+\psi_t xz$	2,535.0	1994			

† not statistically significant

Table 2.1 Deviance profile: male, DP1 experience

age	*18-24*	*25-*	*30-*	*35-*	*40-*	*45-*	*50-*	*55-*	*60-65*
d=1	228	1406	2116	2189	2025	1668	1340	1010	620
d=2	65	470	699	785	853	792	838	678	433
d=3	32	170	260	326	391	400	418	391	313
d=4-	44	253	351	467	576	680	683	731	605
d=8-	13	64	107	150	180	243	308	303	285
d=13-	1*	23	21	39	54	75	112	109	118
d=17-	2*	11	22	46	44	71	101	103	93
d=26-	0**	4	6	9	14	14	21	27	22
d=30-	1**	9	7	15	13	12	30	22	21
d=39-	1**	6	5	9	13	18	19	17	15
d=1yrs	0***	7	9	15	19	20	30	27	19
d=2-	-	4	4	7	8	14	12	9	22
d=5-11	-	2	1	4	1	9	7	26†	-

† comprises 0*14+1*2+2*2+7*1+13*1 (including two rank outliers, in periods 1985 & 1986)
* involving 14 (out of 20) zero weighted cells
** involving 16 (out of 20) zero weighted cells
*** involving 18 (out of 20) zero weighted cells
- indicates no exposure (ie involving 20 (out of 20) zero weighted cells)

Table 2.2 Total recoveries over all periods, age by duration: male, DP1 experience

parameter	*estimate*	*standard error*	*t- statistic*
μ	5.695	0.08739	65.2
α	-0.9355	0.05004	-18.70
β	0.02941	0.004484	6.56
θ	-0.03564	0.001840	-19.37
ϕ	0.004260	0.0009529	4.47
ψ	-0.0002717	0.00008406	-3.23
γ_1	0.05447	0.004856	11.22
γ_2	-0.02596	0.001681	-15.44
γ_3	0.001339	0.0001351	9.91
κ	-0.0002471	0.0001012	-2.44

Table 2.3 Parameter estimates: male, DP1 experience

	x=21.5	*x=27.5*	*x=32.5*	*x=37.5*	*x=42.5*	*x=47.5*	*x=52.5*	*x=57.5*	*x=63*
d=1.5	-1.970	-2.155	-2.309	-2.463	-2.618	-2.772	-2.926	-3.080	-3.249
	-.0194	-.0179	-.0167	-.0154	-.0142	-.0129	-.0117	-.0105	-.0091
d=2.5	-2.247	-2.425	-2.573	-2.721	-2.869	-3.016	-3.164	-3.312	-3.475
	-.0115	-.0100	-.0087	-.0075	-.0063	-.0050	-.0038	-.0026	-.0012
d=3.5	-2.468	-2.640	-2.783	-2.926	-3.069	-3.212	-3.355	-3.499	-3.656
	-.0053	-.0038	-.0026	-.0013	-.0001	.0011	.0024	.0036	.0050
d=6	-2.898	-3.059	-3.193	-3.327	-3.461	-3.595	-3.730	-3.864	-4.011
	.0064	.0079	.0091	.0103	.0116	.0128	.0141	.0153	.0166
d=10.5	-3.459	-3.607	-3.731	-3.854	-3.978	-4.101	-4.224	-4.348	-4.484
	.0209	.0224	.0236	.0248	.0261	.0273	.0286	.0300	.0311
d=15	-3.887	-4.026	-4.142	-4.258	-4.374	-4.491	-4.607	-4.723	-4.850
	.0313	.0328	.0340	.0352	.0365	.0377	.0390	.0402	.0415
d=21.5	-4.378	-4.509	-4.617	-4.726	-4.835	-4.943	-5.052	-5.161	-5.280
	.0424	.0439	.0451	.0464	.0476	.0488	.0501	.0513	.0527
d=28	-4.778	-4.902	-5.005	-5.109	-5.213	-5.316	-5.420	-5.523	-5.637
	.0507	.0522	.0534	.0547	.0559	.0571	.0584	.0596	.0610
d=34.5	-5.116	-5.236	-5.336	-5.436	-5.536	-5.636	-5.736	-5.836	-5.945
	.0571	.0586	.0598	.0611	.0623	.0635	.0648	.0660	.0674
d=45.5	-5.592	-5.708	-5.804	-5.901	-5.997	-6.093	-6.190	-6.286	-6.392
	.0650	.0665	.0677	.0690	.0702	.0714	.0727	.0739	.0753
d=1.5yr	-6.587	-6.702	-6.798	-6.895	-6.991	-7.087	-7.183	-7.279	-7.384
	.0757	.0771	.0784	.0796	.0808	.0821	.0833	.0845	.0859
d=3.5yr	-8.068	-8.234	-8.372	-8.510	-8.648	-8.786	-8.925	-9.063	-9.215
	.0573	.0588	.0601	.0613	.0625	.0638	.0650	.0662	.0676

Table 2.4 A_{xz} and B_{xz} coefficients: male, DP1 experience

age	21.5	27.5	32.5	37.5	42.5	47.5	52.5	57.5	63
z1 (wks)	4.52	4.22	3.98	3.74	3.52	3.30	3.09	2.89	2.68
z2 (wks)	298	300	303	305	307	309	311	313	315

Table 2.5 Predicted age specific durations (*z1, z2*) between which recovery rates decrease

z_0*(wks)*	*DP4*	*DP13*	*DP26*	*DP52*
28	2,240.0	2,032.8	1229.7	
34.5	2,102.3	1,885.1	1195.6	
45.5	1,970.0	1,709.0	1127.5	
78	1,925.9	1,592.4	1044.1	251.3
182	2,383.8	1,928.6	1113.3	260.2

Table 2.6 Male experience: deviance profiles for different knot settings

Model terms	*Deviance*	*D.F.*	*Difference*	*D.F*	*Mean deviance*
μ	16,186	1663			
			12,285	1	12,285
$+\beta z$	3,901.2	1662			
			503.0	1	503.0
$+\theta x$	3,398.3	1661			
			1,010	1	1,010
$+\gamma(z\text{-}z_0)_+$	2,388.8	1660			
			30.3	1	30.3
$+\psi xz$	2,358.4	1659			
			303.1	19	16.0
$+\mu_t$	2,055.4	1640			
			22.6	19	1.2 †
$+\beta_t z$	2,032.8	1621			
			28.6	19	1.5 †
$+\theta_t x$	2,004.2	1602			
			26.0	19	1.4 †
$+\gamma_t(z\text{-}z_0)_+$	1,978.2	1583			
			52.2	19	2.7 †
$+\psi_t xz$	1,925.9	1564			

† not statistically significant

Table 2.7a Male, DP4 experience

Model term	*Deviance*	*D.F.*	*Difference*	*D.F*	*Mean deviance*
μ	6,684.0	1355			
			3,440	1	3,440
$+\beta z$	3,243.6	1354			
			554.8	1	554.8
$+\theta x$	2,688.8	1353			
			852.2	1	852.2
$+\gamma(z\text{-}z_0)_+$	1,836.6	1352			
			9.8	1	9.8
$+\psi xz$	1,826.8	1351			
			132.7	19	7.0
$+\mu_t$	1,694.1	1332			
			36.9	19	1.9 †
$+\beta_t z$	1,657.2	1313			
			16.8	19	0.9 †
$+\theta_t x$	1,640.4	1294			
			19.5	19	1.0 †
$+\gamma_t(z\text{-}z_0)_+$	1,620.9	1275			
			28.5	19	1.5 †
$+\psi_t xz$	1,592.4	1256			

† not statistically significant

Table 2.7b Male, DP13 experience

Model term	*Deviance*	*D.F.*	*Difference*	*D.F*	*Mean deviance*
μ	2,416.2	994			
			624.6	1	624.6
$+\beta z$	1,791.6	993			
			336.7	1	336.7
$+\theta x$	1,454.9	992			
			206.4	1	206.4
$+\gamma(z\text{-}z_o)_+$	1,248.5	991			
			4.6	1	4.6
$+\psi xz$	1,243.9	990			
			65.2	19	3.4
$+\mu_t$	1,178.8	971			
			39.1	19	2.1 †
$+\beta_t z$	1,139.6	952			
			26.0	19	1.4 †
$+\theta_t x$	1,113.7	933			
			32.5	19	1.7 †
$+\gamma_t(z\text{-}z_o)_+$	1,081.2	914			
			37.0	19	1.9 †
$+\psi_t xz$	1,044.1	895			

† not statistically significant

Table 2.7c Males, DP26 experience

Model term	*Deviance*	*D.F.*	*Difference*	*D.F*	*Mean deviance*
μ	425.83	415			
			15.83	1	15.8
$+\beta z$	410.45	414			
			35.30	1	35.3
$+\theta x$	375.15	413			
			6.66	1	6.7
$+\gamma(z\text{-}z_o)_+$	368.48	412			
			3.48	1	3.5
$+\psi xz$	365.00	411			
			23.38	19	1.2 †
$+\mu_t$	341.62	392			
			16.17	19	0.9 †
$+\beta_t z$	325.46	373			
			15.26	19	0.8 †
$+\theta_t x$	310.20	354			
			37.96	19	2.0 †
$+\gamma_t(z\text{-}z_o)_+$	272.24	335			
			20.91	19	1.1 †
$+\psi_t xz$	251.33	316			

† not statistically significant

Table 2.7d Males, DP52 experience

	DP4	*DP13*	*DP26*	*DP52*
μ	3.196	3.197	4.040	3.227
	(.06603)	(.1160)	(.2412)	(.9214)
	48.4	27.6	16.7	3.5
β	-0.04800	-0.04228	-0.04764	-0.03506
	(.002378)	(.001971)	(.003247)	(.009828)
	-20.2	-21.5	-14.7	-3.6
θ	-0.01920	-0.03276	-0.06354	-0.07930
	(.001359)	(.002357)	(.004554)	(.01510)
	-14.1	-13.9	-14.0	-5.3
ζ	0.05616	0.04451	0.04158	0.02605
	(.001460)	(.001475)	(0.002877)	(.009367)
	38.5	30.2	14.4	2.8
ψ	-0.0002606	-0.0001133	0.00008436	0.0001638
	(.00004542)	(.00003395)	(.00004339)	(.00008995)
	-5.7	-3.3	1.9	1.8
γ	-0.03312	-0.03399	-0.03658	-0.02200
	(.002176)	(.003528)	(.005950)	(.01676)
	-15.2	-9.6	-6.1	-1.3

Table 2.8 Parameter estimates, (standard errors), t-statistics: male experience

Model terms	*Deviance*	*D.F.*	*Difference*	*D.F*	*Mean deviance*
μ	9,239.4	1684			
			138.6	19	7.3
$+\mu_t$	9100.9	1665			
			7,018	1	7,018
$+\alpha\sqrt{z}$	2,082.5	1664			
			14.9	19	0.8†
$+\alpha_t\sqrt{z}$	2,067.6	1645			
			108.7	1	108.7
$+\beta z$	1,958.9	1644			
			14.5	19	0.8†
$+\beta_t z$	1,944.4	1625			
			233.8	1	233.8
$+\theta x$	1,710.6	1624			
			19.3	19	1.0†
$+\theta_t x$	1,691.3	1605			
			18.6	1	18.6
$+\phi x\sqrt{z}$	1,672.7	1604			
			21.6	19	1.1†
$+\phi_t x\sqrt{z}$	1,651.1	1585			
			9.8	1	9.8
$+\psi xz$	1,641.2	1584			
			29.4	19	1.5†
$+\psi_t xz$	1,611.9	1565			

† not statistically significant

Table 2.9 Deviance profile: female, DP1 experience

parameter	*estimate*	*standard error*	*t- statistic*
μ	6.035	0.1717	35.15
α	-1.336	0.1013	-13.19
β	0.04960	0.008206	6.04
θ	-0.04107	0.004117	10.00
ϕ	0.01004	0.002390	4.20
ψ	-0.0005190	0.0002032	-2.55
γ	0.004979	0.002992	1.66

Table 2.10 Parameter estimates: females, DP1 experience

$z_0(wks)$	*DP4*	*DP13*	*DP26*
34.5	1,478.5	1,048.5	723.6
45.5	1,444.5	1,025.2	706.2
78	1,428.6	1,011.5	686.7
182	1,530.6	1,081.4	705.3

Table 2.11 Female experience: deviance profiles for different knot settings

	DP4	*DP13*	*DP26*
μ	2.711	3.140	3.331
	(.1225)	(.2423)	(.4068)
	22.1	13.0	8.2
β	-0.05294	-0.04153	-0.03890
	(.001938)	(.002688)	(.004969)
	-27.3	-15.5	-7.8
θ	-0.01194	-0.03506	-0.05333
	(.002598)	(.004626)	(.006911)
	-4.6	-7.6	-7.7
ζ	0.04864	0.03705	0.03705
	(.003033)	(.003538)	(0.005784)
	16.0	10.5	6.4
γ	-0.02589	-0.03624	-0.04033
	(.004825)	(.008301)	(.01251)
	-5.4	-4.4	-3.2

Table 2.12 Parameter estimates, (standard errors), t-statistics: female experience

parameter	*estimate*	*s.e.*	*t-statistic*
μ	-3.528	0.1985	-17.8
$\alpha(2)$	-0.3417	0.0825	-4.14
$\beta(2)$	-0.0719	0.0867	-0.83
$\beta(3)$	-0.3113	0.0809	-3.85
$\beta(4)$	-0.4178	0.0779	-5.36
$\beta(5)$	-0.5436	0.0801	-6.79
$\gamma(2)$	0.1334	0.1476	0.90†
$\gamma(3)$	0.3737	0.1313	2.85
$\gamma(4)$	0.5017	0.1246	4.03
$\gamma(5)$	0.6210	0.1200	5.18
$\gamma(6)$	0.6866	0.1179	5.82
$\gamma(7)$	0.8270	0.1251	6.61
$\delta(2)$	0.5554	0.2016	2.76
$\delta(3)$	0.9224	0.1983	4.65
$\delta(4)$	0.8955	0.1996	4.48
$\delta(5)$	1.1183	0.1811	6.17
$\delta(6)$	1.2046	0.1919	6.28
$\delta(7)$	0.9217	0.1827	5.04
$\delta(8)$	0.8483	0.1798	4.72
$\delta(9)$	0.3181	0.1723	1.85††
$\delta(10)$	-0.3470	0.1733	-2.00††
$\delta(11)$	-0.3073	0.1786	-1.72††

$\alpha(1) = \beta(1) = \gamma(1) = \delta(1) = 0$

† not satistically significant †† of marginal statistical significance

Table 3.1 Model (3.1): parameter estimates, standard errors, t- statistics

$k_2 \backslash k_1$	*2.5*	*6*	*10.5*	*15*	*21.5*	*28*	*34.5*
45.5	1038.8	986.93	952.11	941.79	928.98	923.48	948.31
78	987.53	922.68	883.81	875.24	868.45	873.36	892.56
182	907.22	853.42	831.45	834.10	839.71	855.78	872.70
416	1047.7	1019.7	1017.7	1026.8	1035.6	1045.4	1047.6

Table 3.2 Model (3.2): two knot deviance profile

parameter	*estimate*	*standard error*	*t- statistic*
μ	-4.358	0.2206	-19.76
α_2	-0.3368	0.08257	-4.08
γ	-0.1396	0.01750	-7.98
θ	0.02371	0.002663	8.90
β	0.1422	0.01787	7.96
β_1	-0.1542	0.01828	-8.44
β_2	0.005190	0.001570	3.30
β_3	0.006986	0.0008959	7.80

knots: $k_1 = 10.5$, $k_2 = 78$, $k_1 = 182$

deviance = 820.33 on 764 degrees-of-freedom

Table 3.3 Model (3.2): parameter estimates, standard errors, t-statistics

DP1

Model	*Deviance*	*D.F.*	$\hat{\phi}$	*Difference*	*D.F*	*'F-statistic'*
1	2277.8	179				
				468.5	8	5.5
$+x$	1809.3	171	10.58			
				1086.7	19	12.0
$+t$	722.6	152	4.75			

DP4

Model	*Deviance*	*D.F.*	$\hat{\phi}$	*Difference*	*D.F*	*'F-statistic'*
1	2681.8	179				
				2220.7	8	102.8
$+x$	461.1	171	2.70			
				188.7	19	5.4
$+t$	272.4	152	1.79			

DP13

Model	*Deviance*	*D.F.*	$\hat{\phi}$	*Difference*	*D.F*	*'F-statistic'*
1	2577.0	179				
				2302.4	8	178.8
$+x$	274.6	171	1.61			
				66.8	19	2.6
$+t$	207.8	152	1.37			

DP26

Model	*Deviance*	*D.F.*	$\hat{\phi}$	*Difference*	*D.F*	*'F-statistic'*
1	3386.4	179				
				2995.8	8	164.2
$+x$	390.6	171	2.28			
				201.3	19	8.5
$+t$	189.3	152	1.25			

DP52

Model	*Deviance*	*D.F.*	$\hat{\phi}$	*Difference*	*D.F*	*'F-statistic'*
1	1489.9	179				
				1119.3	8	64.5
$+x$	370.6	171	2.17			
				127.0	19	4.2
$+t$	243.6	152	1.60			

Table 4.1 Analysis of deviance, male experiences

DP1

Model	*Deviance*	*D.F.*	$\hat{\phi}$	*Difference*	*D.F*	*'F-statistic'*
1	828.4	179				
				318.4	8	13.4
$+x$	510.0	171	2.98			
				87.5	19	1.7
$+t$	422.5	152	2.78			

DP4

Model	*Deviance*	*D.F.*	$\hat{\phi}$	*Difference*	*D.F*	*'F-statistic'*
1	725.7	179				
				488.3	8	43.9
$+x$	237.4	171	1.39			
				30.9	19	1.2
$+t$	206.5	152	1.36			

DP13

Model	*Deviance*	*D.F.*	$\hat{\phi}$	*Difference*	*D.F*	*'F-statistic'*
1	395.4	179				
				182.1	8	18.2
$+x$	213.2	171	1.25			
				30.3	19	1.3
$+t$	182.9	152	1.20			

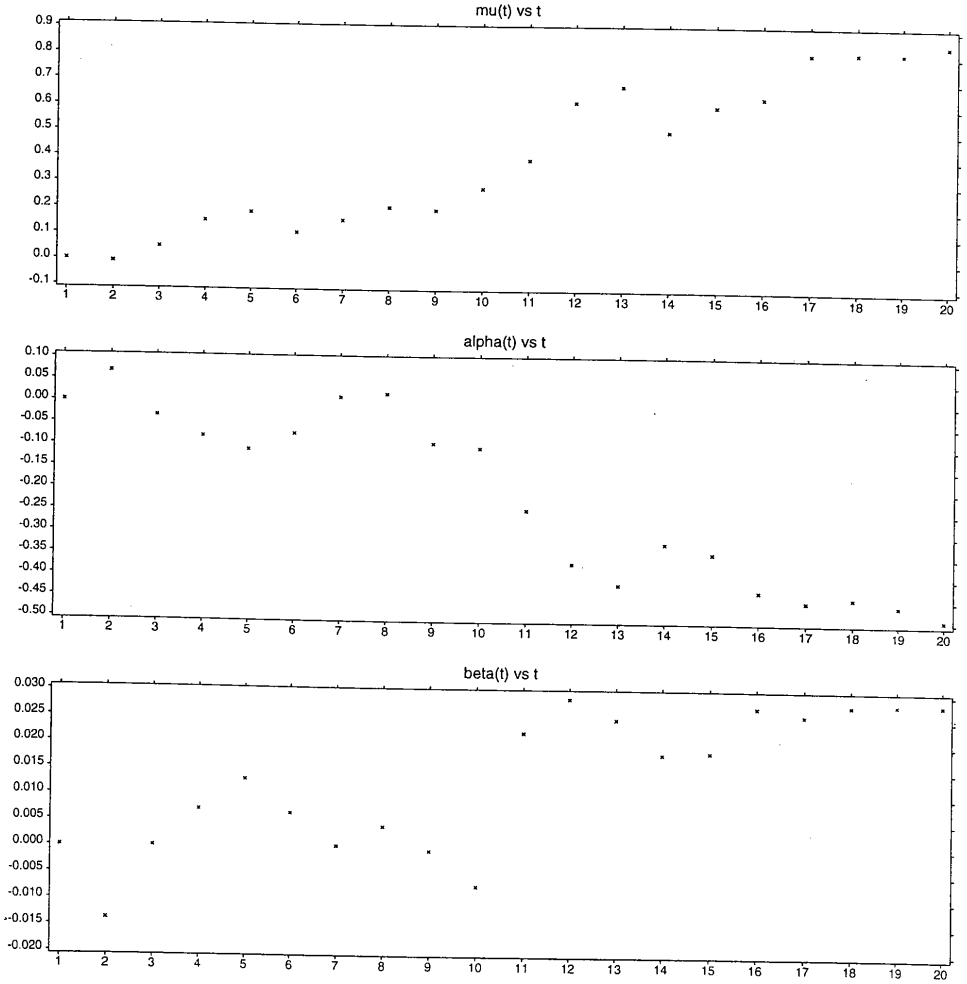
DP26

Model	*Deviance*	*D.F.*	$\hat{\phi}$	*Difference*	*D.F*	*'F-statistic'*
1	661.8	179				
				383.2	8	29.4
$+x$	278.6	171	1.63			
				65.1	19	2.4
$+t$	213.5	152	1.40			

Table 4.2 Analysis of deviance, female experiences

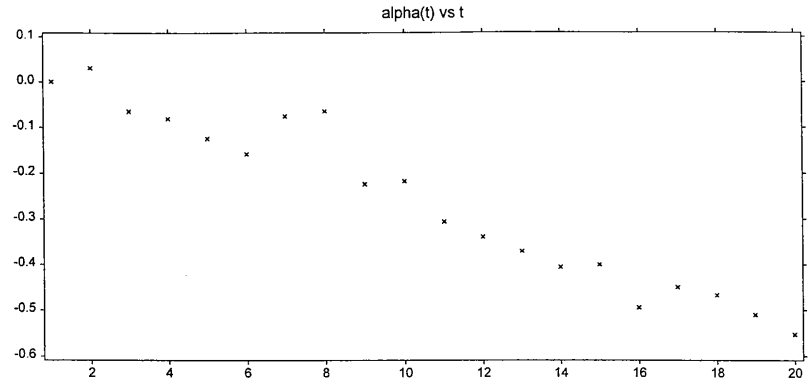
age	*DP1*	*DP4*	*DP13*	*DP26*
18-24	0.1001	0.0102	0.0036	0.0005
25-29	0.1192	0.0105	0.0033	0.0015
30-34	0.1477	0.0210	0.0043	0.0019
35-39	0.1840	0.0237	0.0050	0.0028
40-44	0.1985	0.0291	0.0070	0.0039
45-49	0.2104	0.0376	0.0084	0.0058
50-54	0.2151	0.0363	0.0120	0.0095
55-59	0.2020	0.0523	0.0170	0.0157
60-65	0.1229	0.0318	0.0105	0.0128

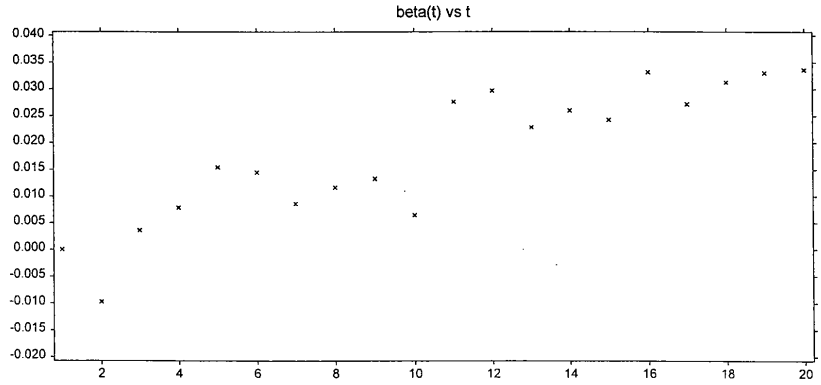
Table 4.3 Predicted claims inception rates by age, 1975-94 female experiences



Model (2.1a): parameter estimates $\hat{\mu}_t$, $\hat{\alpha}_t$, $\hat{\beta}_t$ *vs* period 1975-94 (coded $t = 1, ..., 20$)

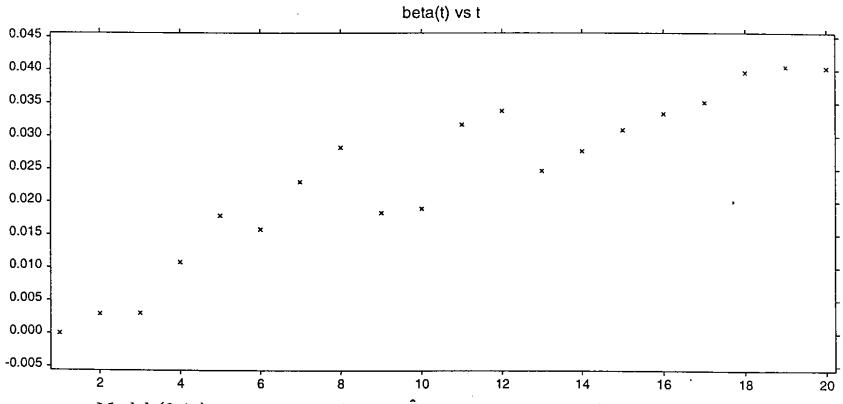
Figure 2.1(a)





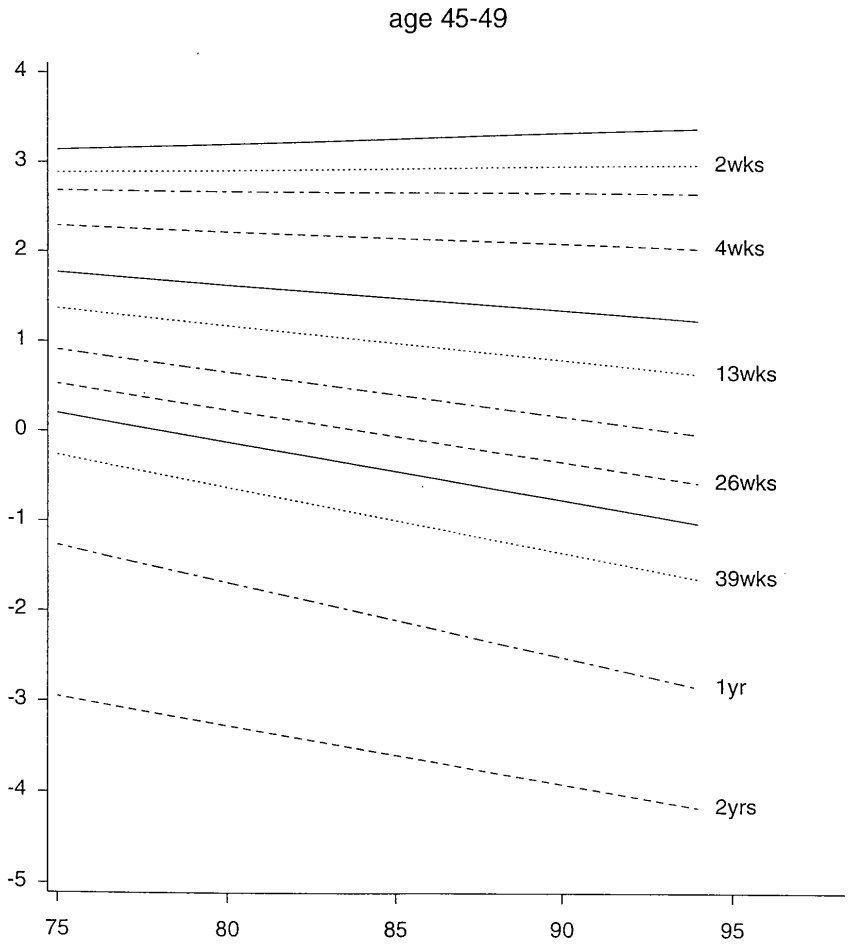
Model (2.1b): parameter estimates $\hat{\alpha}_t$, $\hat{\beta}_t$ *vs* period 1975-94 (coded $t = 1, ..., 20$)

Figure 2.1(b)



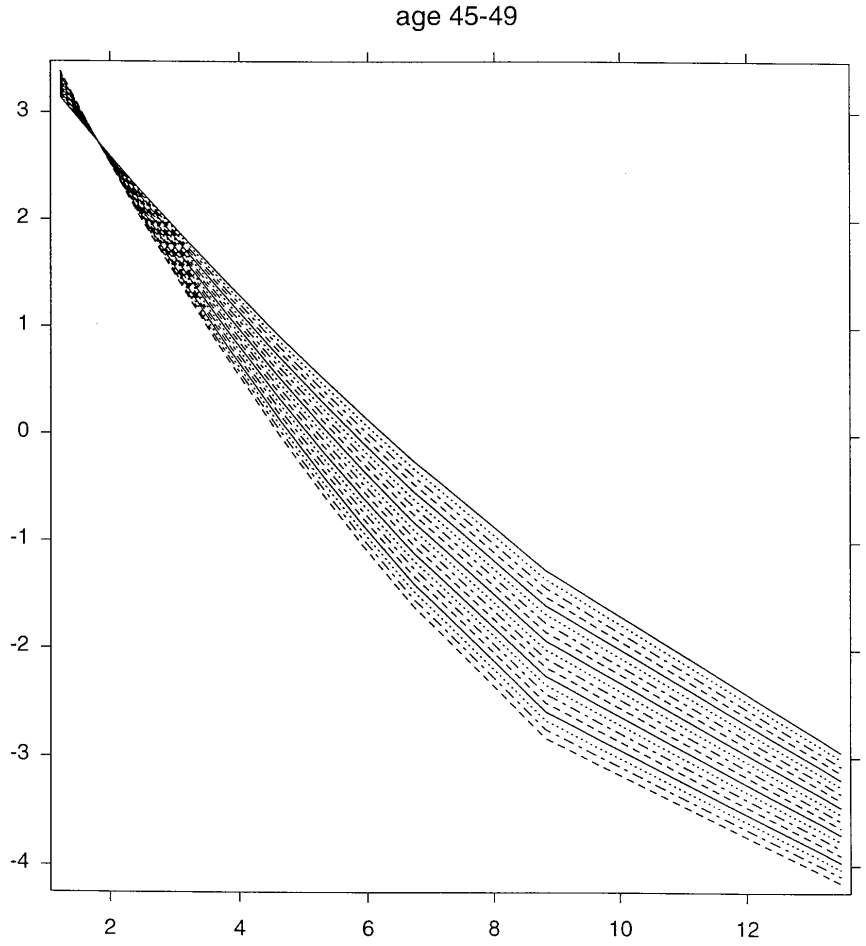
Model (2.1c): parameter estimates $\hat{\beta}_t$ *vs* period 1975-94 (coded $t = 1, ..., 20$)

Figure 2.1(c)



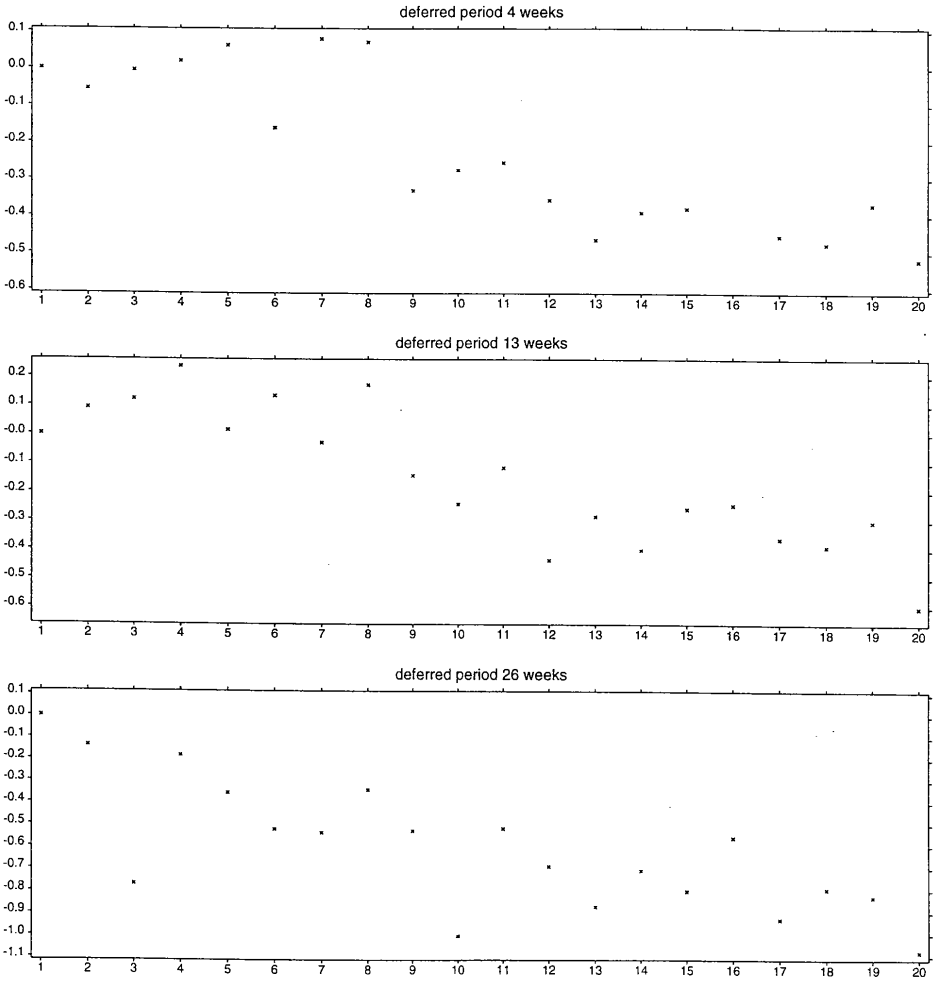
Model (2.4): predicted straight-line log recovery rates *vs* period 1975-94, by duration (separate contours), specific age, male DP1 experience

Figure 2.2



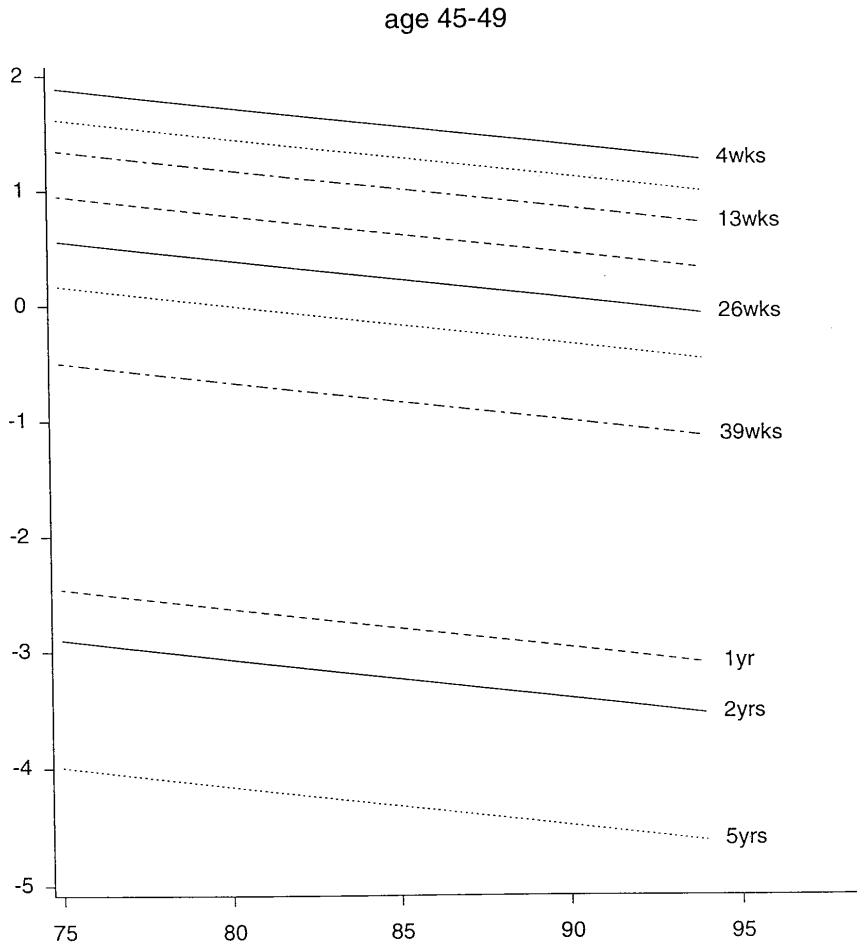
Model (2.4): predicted quadratic recovery rates *vs* square root duration, by period (separate contours), specific age, male DP1 experience

Figure 2.3



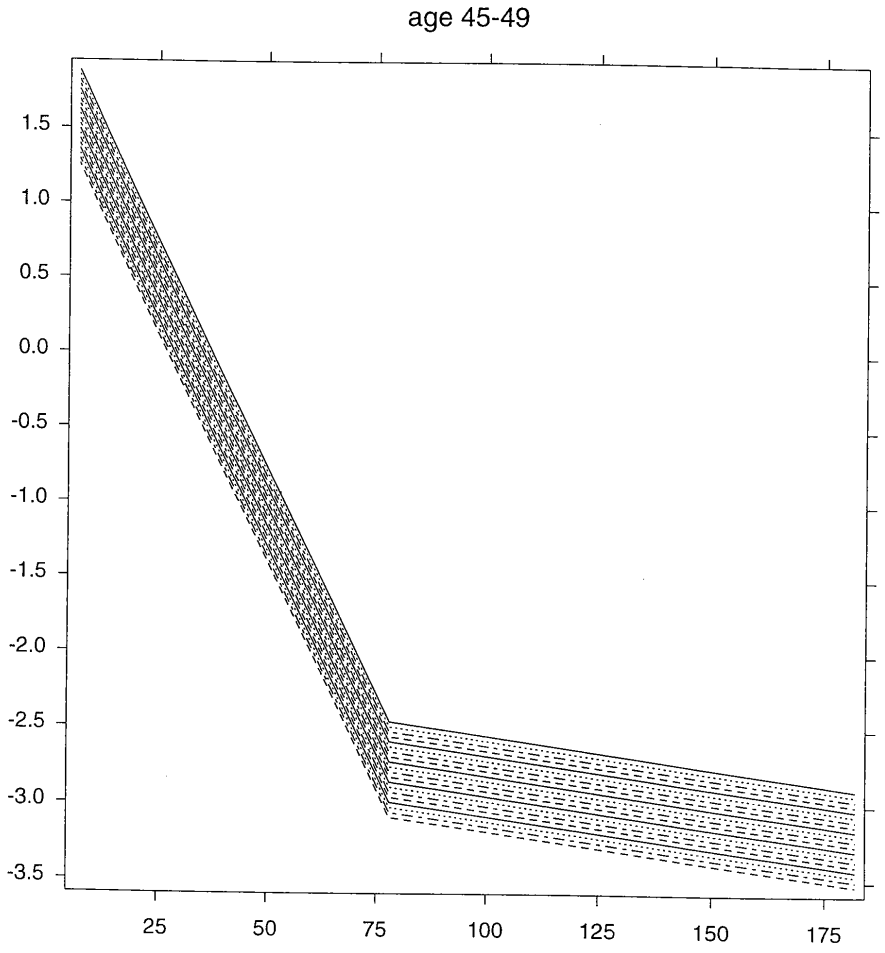
Model (2.5a): parameter estimates $\hat{\mu}_t$ *vs* period 1975-94 (coded $t = 1, ..., 20$), fixed DP

Figure 2.4



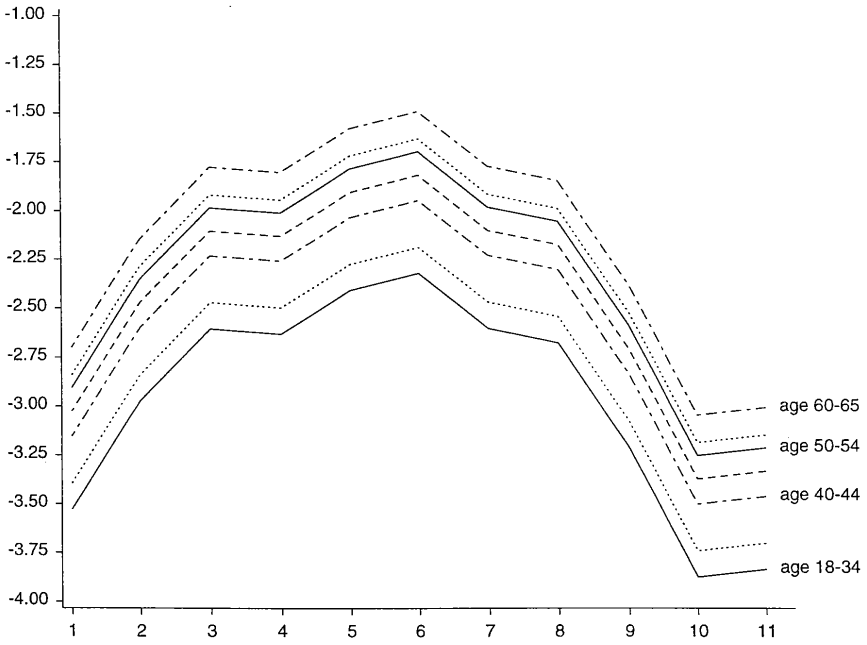
Model (2.6): predicted straight-line log recovery rates *vs* period 1975-94, by duration (separate contours), specific age, male DP4 experience

Figure 2.5



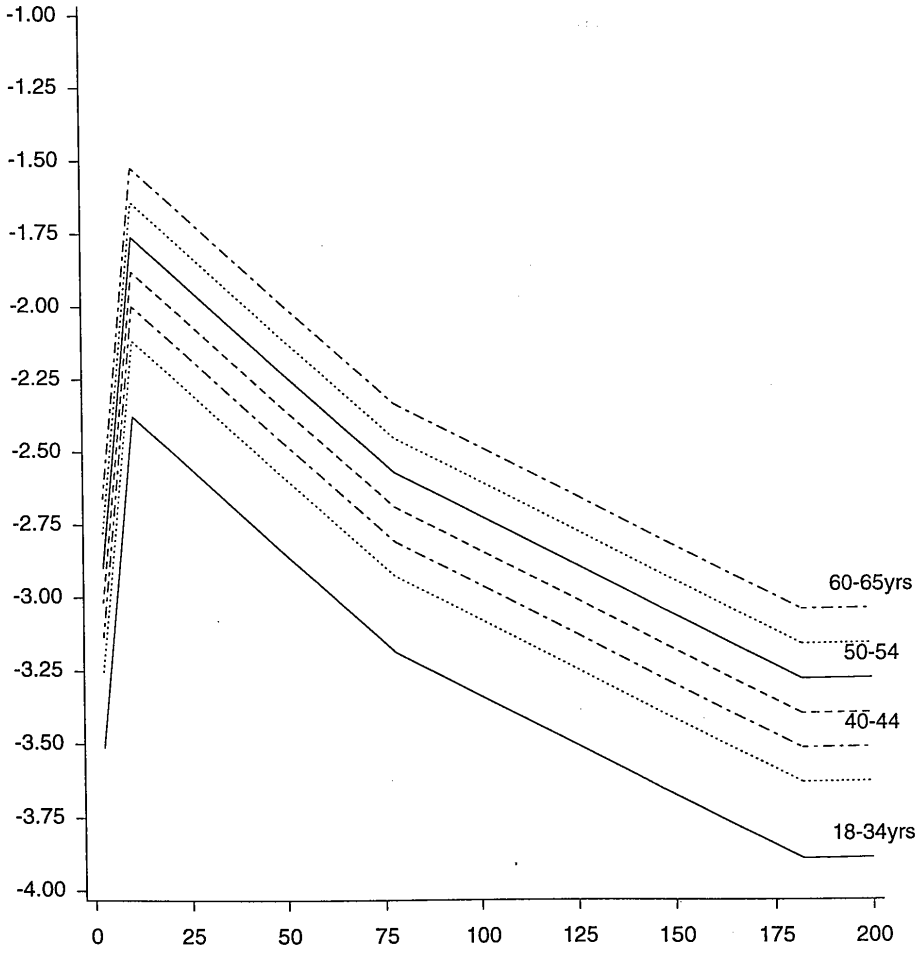
Model (2.6): predicted break-point recovery rates *vs* duration,
by period (separate contours), specific age, male DP4 experience

Figure 2.6



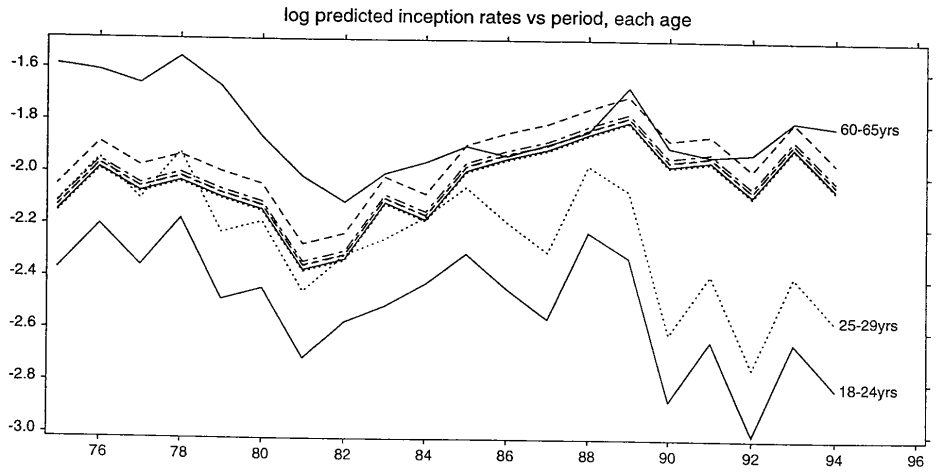
Model (3.1): predicted log mortality rates *vs* duration (categorical scale), specific ages, male 1975-78 experience

Figure 3.1



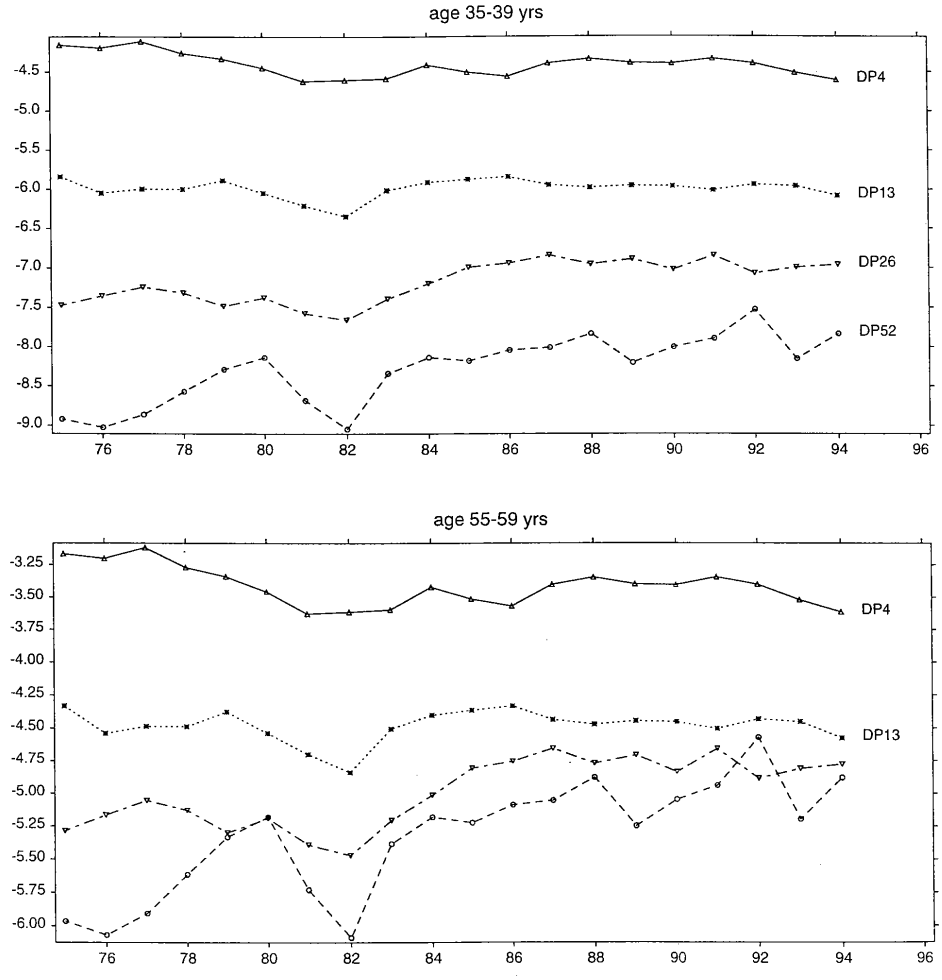
Model (3.2): predicted log mortality rates *vs* duration (weeks), specific ages, male 1975-78 experience

Figure 3.2



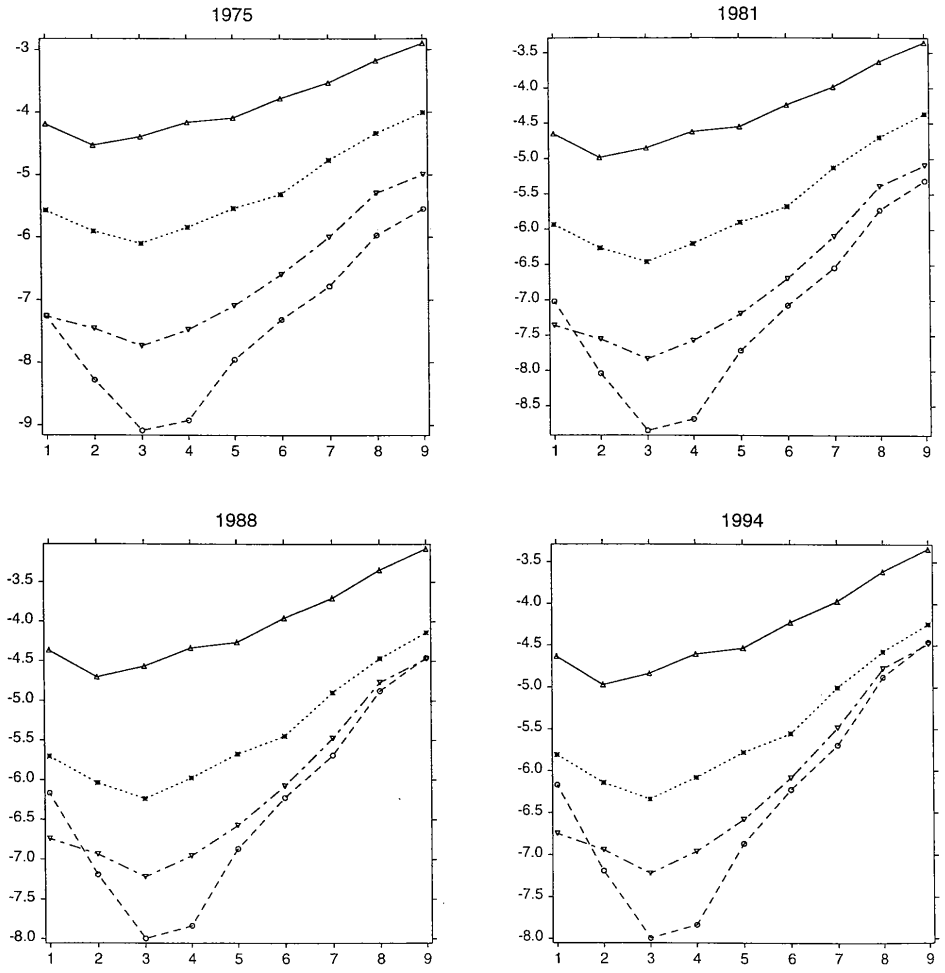
Model (4.3a): male DP1 experience

Figure 4.1



Model (4.2): predicted log claim inception rates *vs* period,
by DP (separate contours), different ages, male experience

Figure 4.2



Model (4.2): predicted log claim inception rates *vs* age,
by DP4 to DP52 (separate contours), different periods, male experience

Figure 4.3

Appendix I (deviance profiles).

Male, DP1 experience:

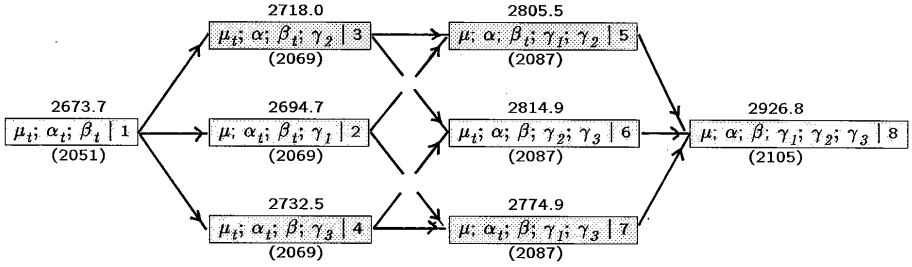
The graphical evidence supporting the progression from model structure:

$$\log \rho_{txz} = (\mu+\mu_t) + (\alpha+\alpha_t)\sqrt{z} + (\beta+\beta_t)z + \theta x + \phi x\sqrt{z} + \psi xz$$

to the model structure:

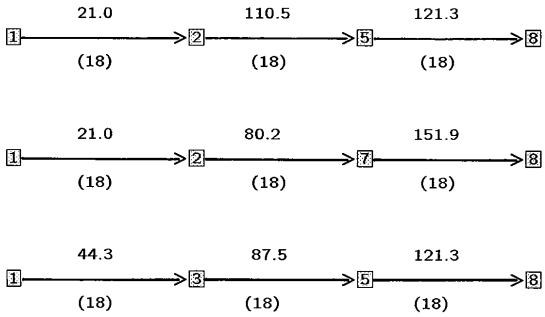
$$\log \rho_{txz} = \mu + \alpha\sqrt{z} + \beta z + \gamma_1 t + \gamma_2 t\sqrt{z} + \gamma_3 tz + \theta x + \phi x\sqrt{z} + \psi xz$$

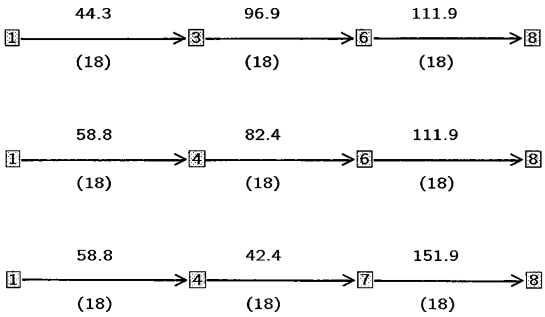
in Section 2.2, is based on just one of six possible sequences of model simplifications. The full set of possible simpifications is depicted by the various routes in the following lattice diagram:



Lattice of hypotheses

in which the model structures are represented by the various numbered frames and these, in turn, are identified by the listed parameters, (with θ, ϕ and ψ taken as read). The respective model deviances and degrees of freedom are displayed above and below each frame. The six possible routes through the lattice (the first of which is referred to in Section 2.2), together with the differences in deviance and degrees of freedom, are as follows:





Possible routes through the lattice diagram

Irrespective of the route, the net reduction in deviance is 253.1 on 54 degrees of freedom, with a mean deviance reduction of 4.7.

Male, other deferred periods:

The reduction in mean deviance in progressing from:

$$\log \rho_{txz} = \mu + \mu_t + \beta z + \theta x + \zeta (z - z_0)_+ + \psi xz$$

to:

$$\log \rho_{txz} = \mu + \beta z + \theta x + \gamma t + \zeta (z - z_0)_+ + \psi xz$$

for DP4, DP13, DP26, in which μ_t is replaced by a + γt ($\mu \mapsto \mu$ + a), is as follows:

DP4	*DP13*	*DP26*
73.6 on 18 d.f.	41.9 on 18 d.f.	28.2 on 18 d.f.
4.0	2.3	1.6

Mean deviance based on differences

Appendix II (residual plots).

By way of illustration, this appendix contains residual plots for the following cases:

RECOVERIES: Male DP1 experience, Section 2, Model (2.4).
Deviance residuals plotted against the 'end-on' duration by period index, for each age category. Pages II-1 to II-3.

DEATHS: All experience, 1975-94, Section 3, Model (3.1).
Deviance residuals plotted against the 'end-on' duration by age index, for males- each quadrennium, and for female. Pages II-4, II-5.

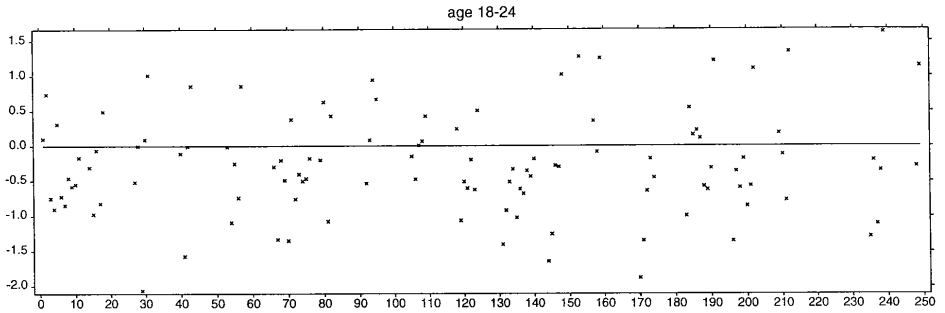
INCEPTIONS: Male DP1 experience, Section 4, Model (4.3a). Pages II-6 to II-8.
Male DP4 experience, Section 4, Model (4.2). Pages II-9 to II-11.
Deviance residuals plotted against period (coded 1 to 20), each age category.

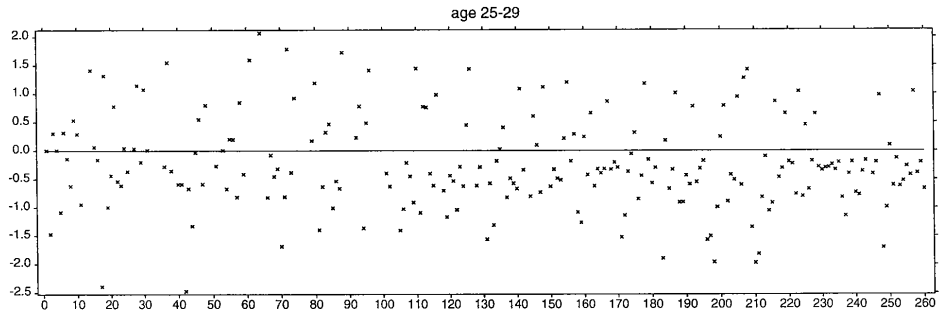
The following features are noteworthy:

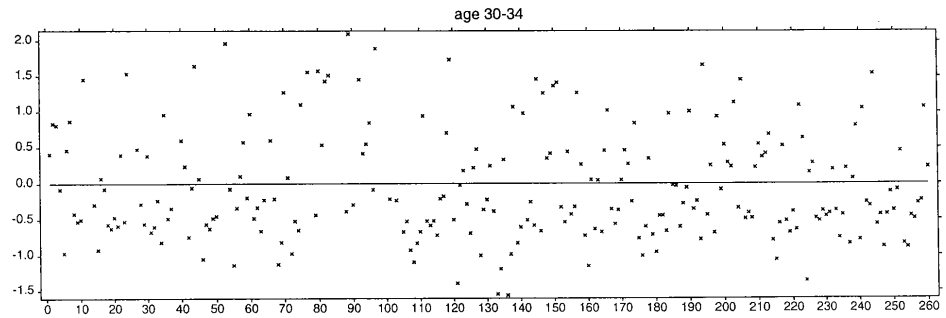
(i) Ideally the residuals should lie in horizontal bands about the x- axis, exhibiting no other obvious pattern.

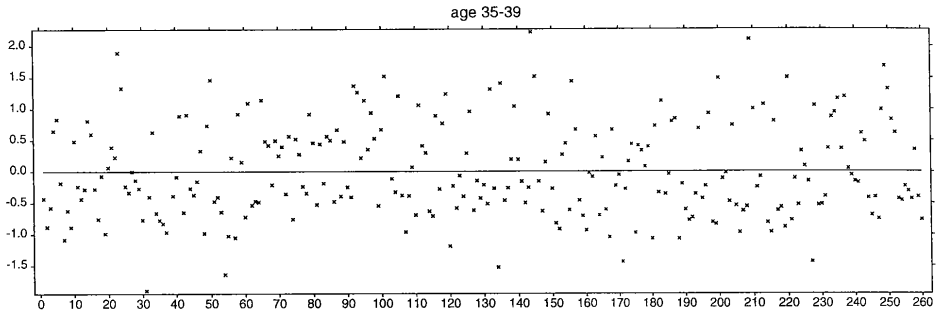
(ii) The marked bias towards negative residuals in the case of recoveries and deaths (Pages II-1 to II-5), which is more pronounced in certain plots than others, is induced by the relatively high proportion of zero response counts in cells, and which automatically trigger negative residuals.

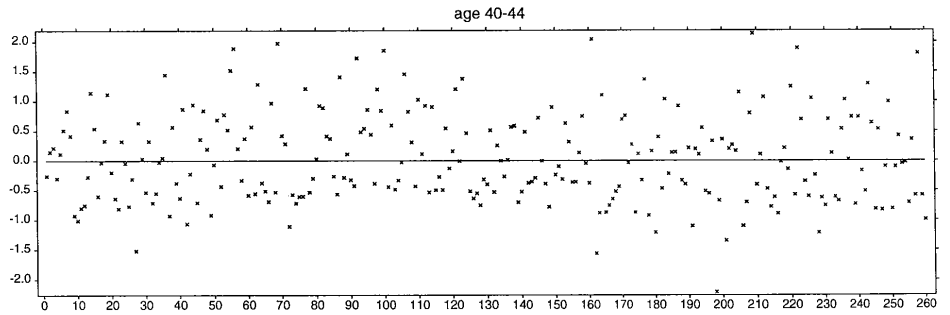
(iii) The outliers discussed in Section 2 are in evidence in the plots for recoveries (p II-3), for age categories 55-59 and 60-65.

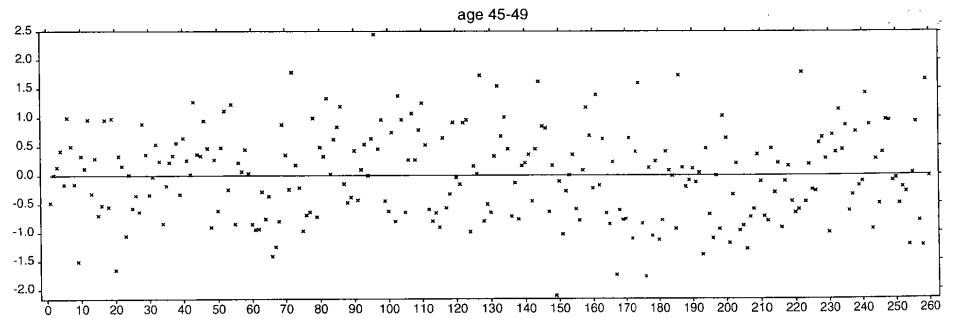

age 18-24
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250

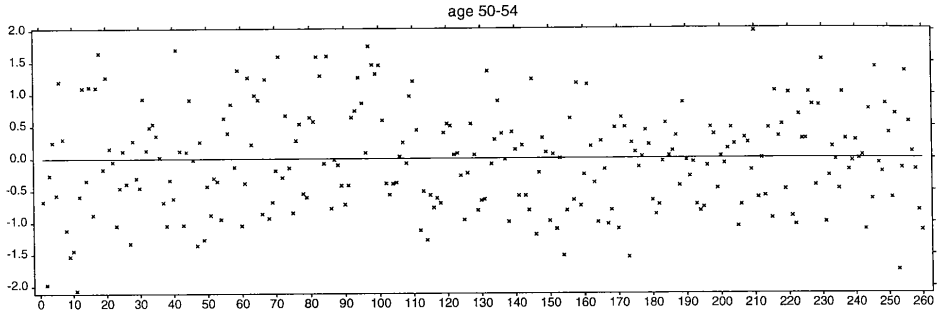

age 25-29
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
-2.5
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260

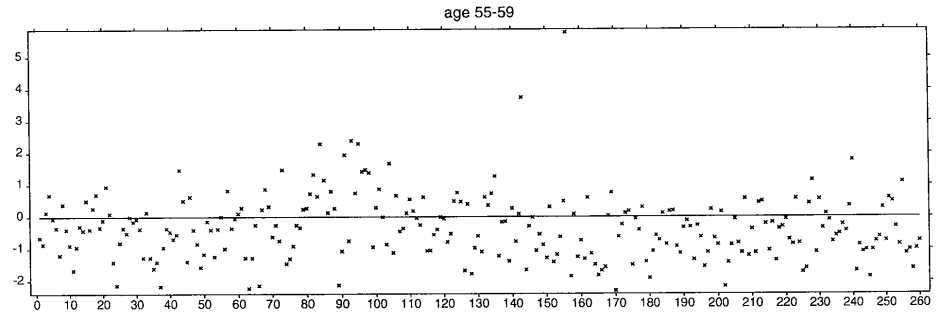

age 30-34
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260

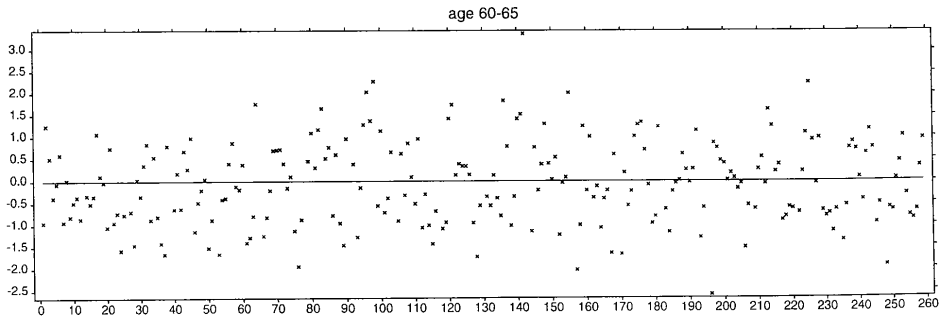

age 35-39
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260

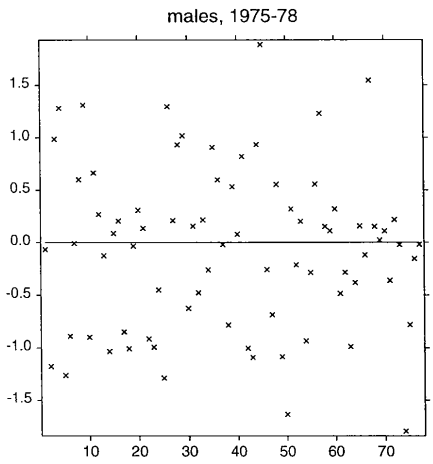

age 40-44
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260

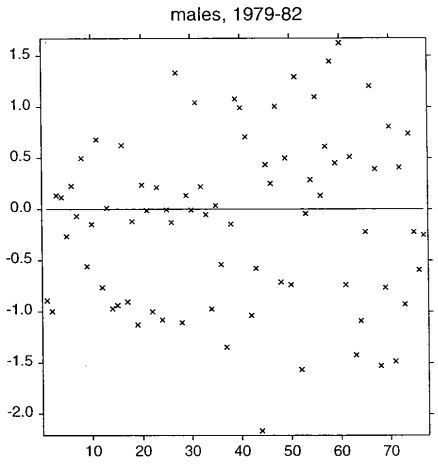

age 45-49
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260

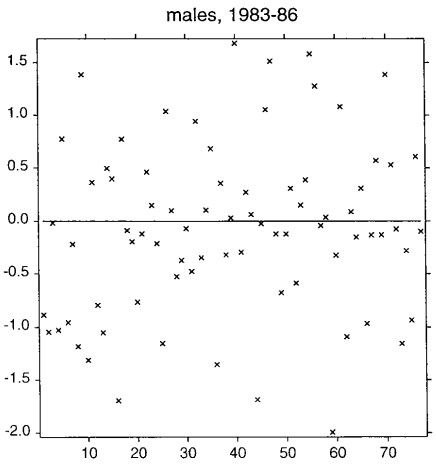

age 50-54
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260

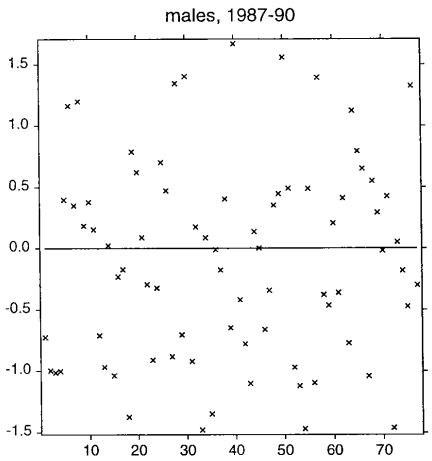

age 55-59
5
4
3
2
1
0
-1
-2
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260

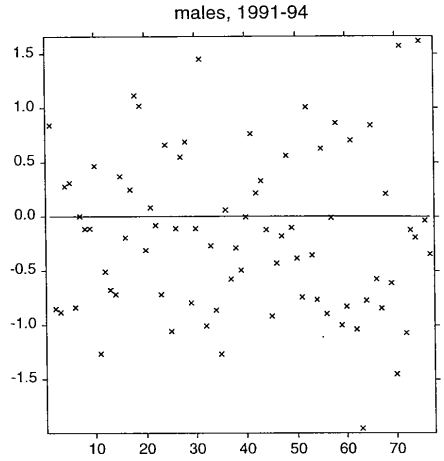

age 60-65
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
-2.5
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260

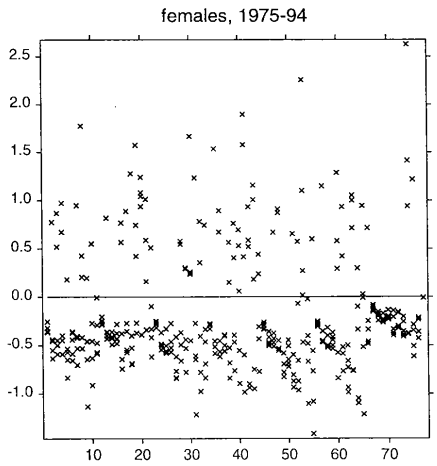

males, 1975-78
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
10
20
30
40
50
60
70

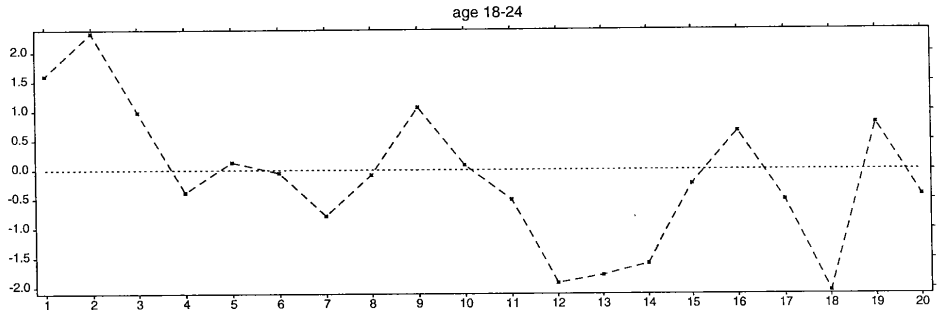

males, 1979-82
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
10
20
30
40
50
60
70

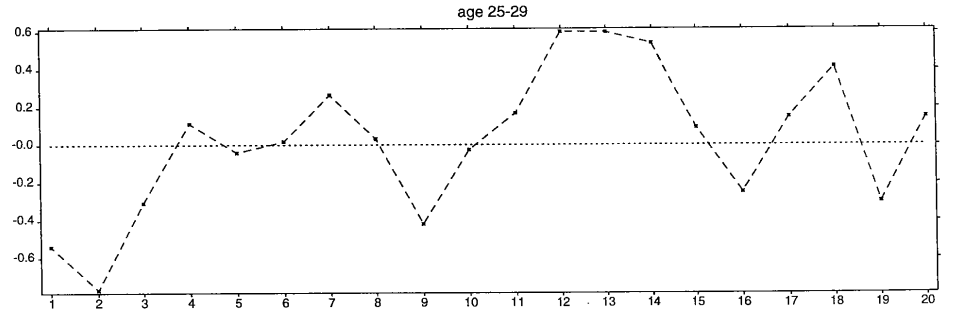

males, 1983-86
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
10
20
30
40
50
60
70

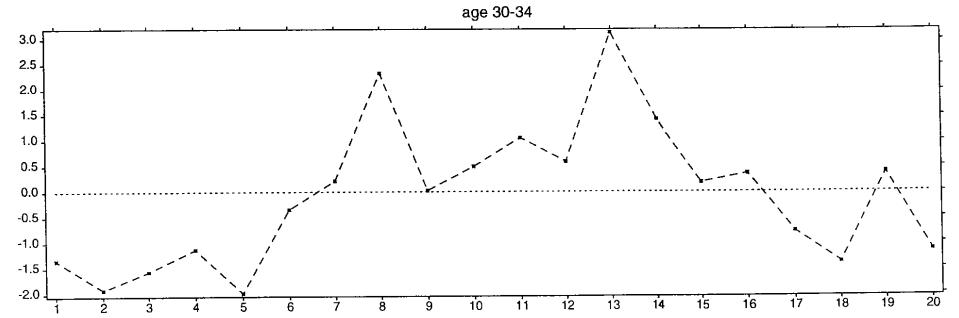

males, 1987-90
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
10
20
30
40
50
60
70

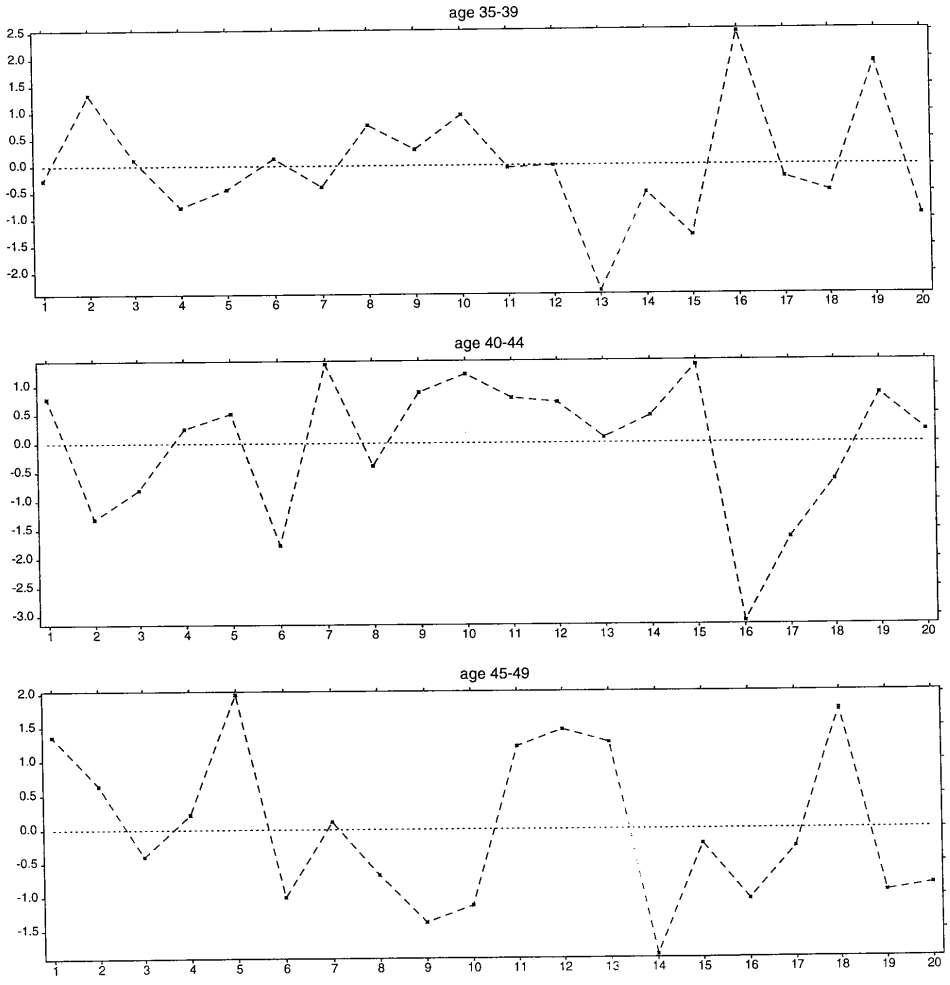

males, 1991-94
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
10
20
30
40
50
60
70

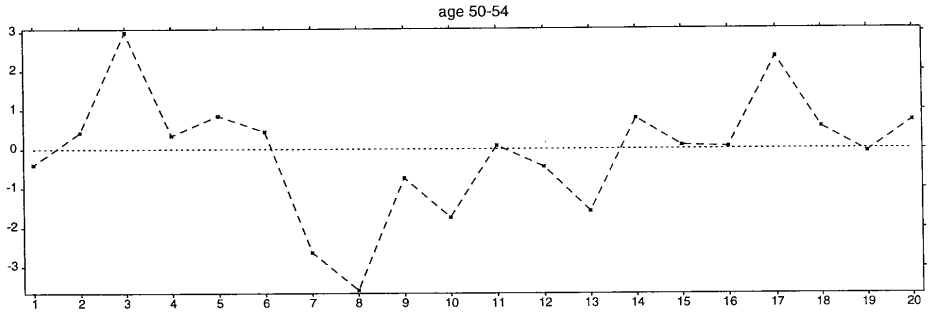

females, 1975-94
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
10
20
30
40
50
60
70

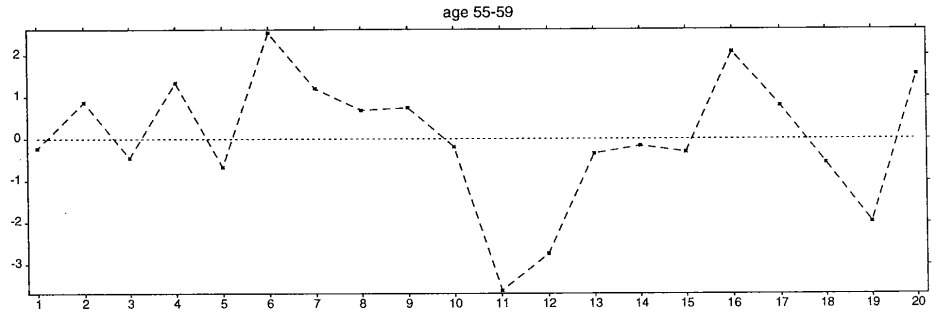

age 18-24
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

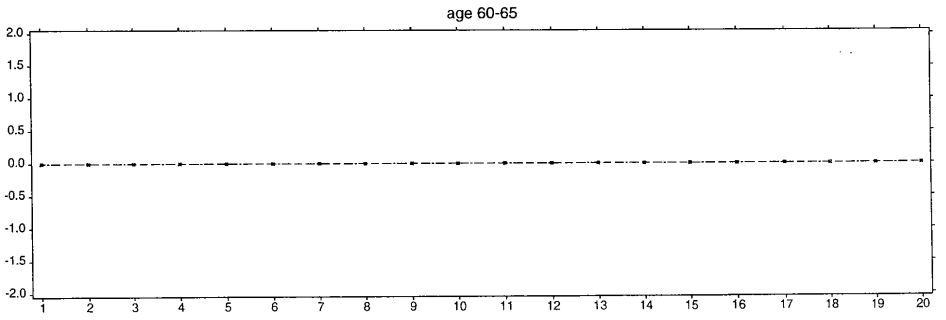

age 25-29
0.6
0.4
0.2
-0.0
-0.2
-0.4
-0.6
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

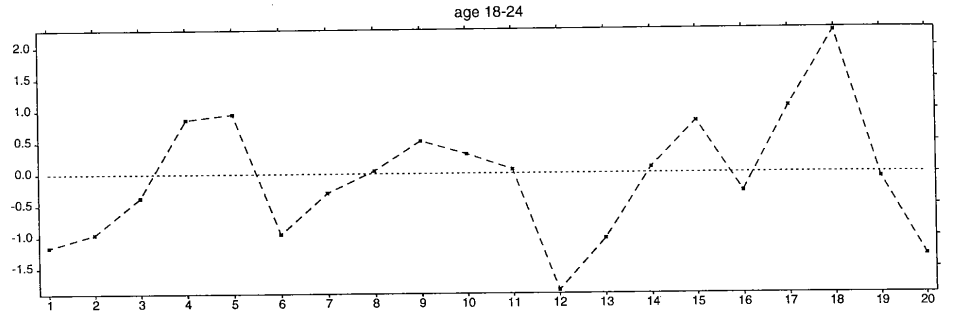

age 30-34
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

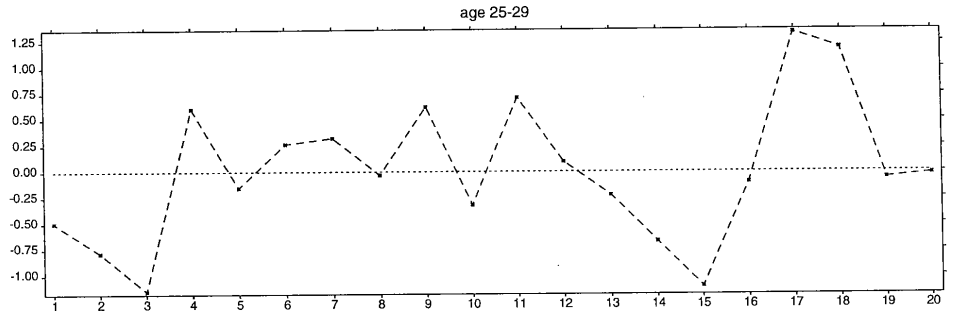

age 35-39
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
age 40-44
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
-2.5
-3.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
age 45-49
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

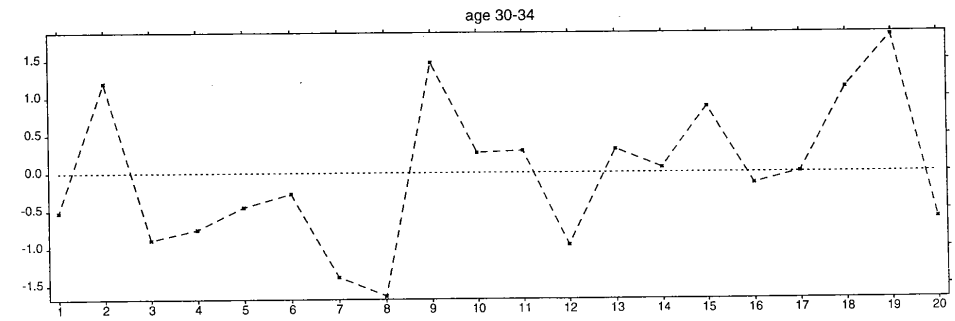

age 50-54
3
2
1
0
-1
-2
-3
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

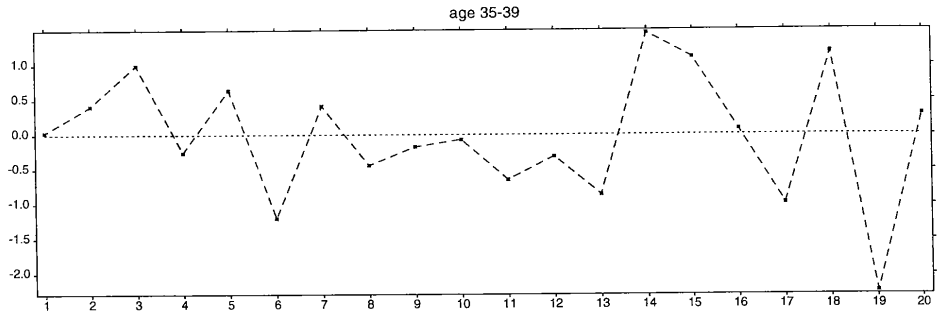

age 55-59
2
1
0
-1
-2
-3
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

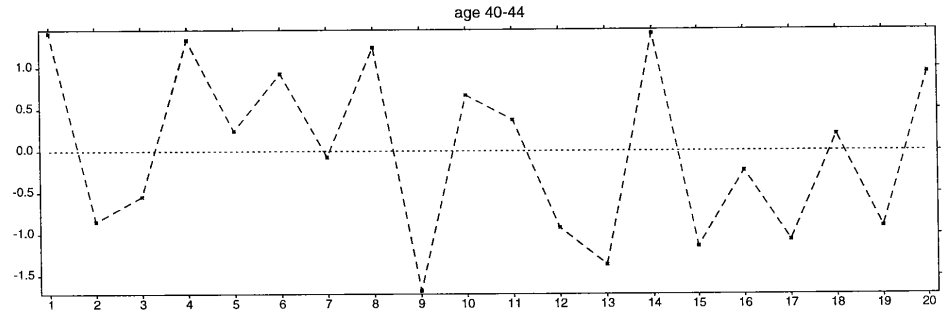

age 60-65
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

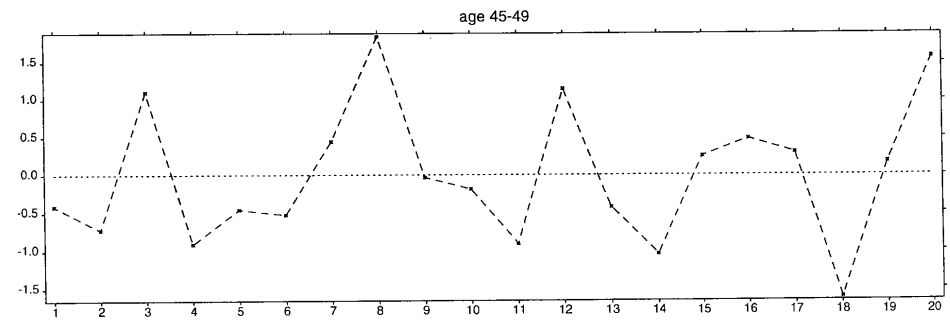

age 18-24
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

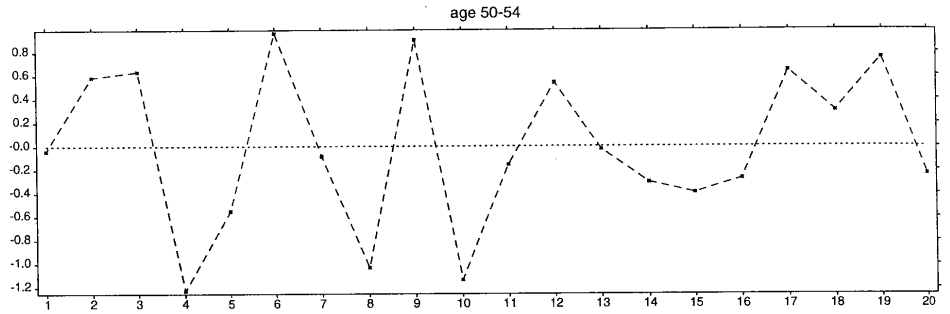

age 25-29
1.25
1.00
0.75
0.50
0.25
0.00
-0.25
-0.50
-0.75
-1.00
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

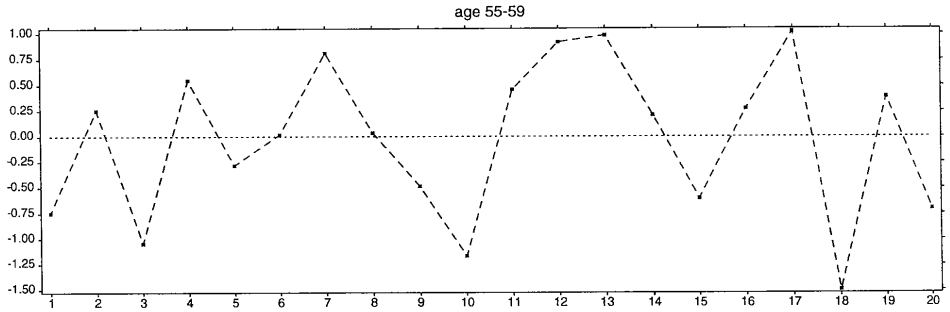

age 30-34
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

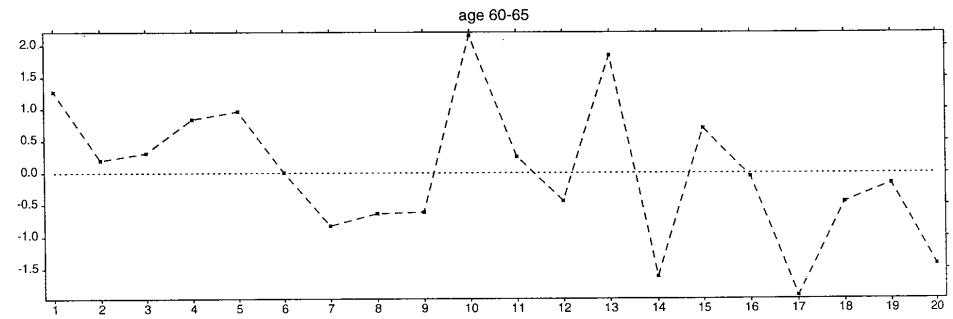

age 35-39
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20


age 40-44
1.0
0.5
0.0
-0.5
-1.0
-1.5
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20


age 45-49
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20


age 50-54
0.8
0.6
0.4
0.2
-0.0
-0.2
-0.4
-0.6
-0.8
-1.0
-1.2
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20


age 55-59
1.00
0.75
0.50
0.25
0.00
-0.25
-0.50
-0.75
-1.00
-1.25
-1.50
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20


age 60-65
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

Appendix III (computation of $\pi_{tx\omega}$).

For the record, we note that the integral:

$$\pi_{tx\omega} = \exp - \int_0^{\omega/52} (\rho_{t+u,x+u,u} + \nu_{t+u,x+u,u})\,\mathrm{d}u$$

(t and x in years, ω in weeks) may be evaluated numerically using the appoximation:

$$\pi_{tx\omega} \doteq \exp - \frac{1}{52h} \sum_{u=1}^{h\omega} \left(\rho_{t + \frac{(u\text{-}0.5)}{52h},\ x + \frac{(u\text{-}0.5)}{52h},\ \frac{(u\text{-}0.5)}{52h}} \right); \quad \omega = 1, 2, 3, \ldots$$

$$\pi_{tx0} = 1,$$

based on the 'trapezium' rule. The approximation recognises that the contribution to the integral from ρ far exceeds that from ν (recovery from sickness is far more likely than death from sickness) so that the latter contribution may be ignored for practical purposes. Note also that the week is partitioned into h (typically h= 3) intervals in order to implement the numerical approximation.

For the male experience, this formula might be used in combination with the following rearranged version of equation (2.4), relating to DP1:

$$\log \rho_{t+u,x+u,u} = A_{tx} + B_{tx}\sqrt{u} + C_{tx}u + D(\sqrt{u})^3 + Eu^2$$

where:

$$A_{tx} = \mu + \gamma_1 t + \theta x + \kappa xt, \quad B_{tx} = \alpha + \gamma_2 t + \phi x,$$
$$C_{tx} = \beta + \gamma_1 + \theta + (\gamma_3 + \kappa)t + (\psi + \kappa)x,$$
$$D = \gamma_2 + \phi, \quad E = \gamma_3 + \psi + \kappa;$$

and, for the female experience, it might be used in combination with the rearranged version of equation (2.7), relating to DP1:

$$\log \rho_{t+u,x+u,u} = A_{tx} + B_x\sqrt{u} + C_x u + D(\sqrt{u})^3 + Eu^2$$

where:

$$A_{tx} = \gamma t + \theta x, \quad B_x = \alpha + \phi x, \; C_x = \beta + \gamma + \theta + \psi x, \; D = \phi, \; E = \psi.$$

CITY University

DEPARTMENT OF ACTUARIAL SCIENCE AND STATISTICS

Actuarial Research Papers since 1992

32. England P.D. and Verrall R.J. Dynamic Estimation for Models of Excess Mortality. January 1992. 19 pages. ISBN 1 874 770 32 8

33. Verrall R.J. A State Space Formulation of Whittaker-Henderson Graduation, with Extensions. January 1992. 13 pages. ISBN 1 874 770 33 6

34. Verrall R.J. Graduation by Dynamic Regression Methods. January 1992. 37 pages. ISBN 1 874 770 34 4

35. Gerrard R.G. and Haberman S. Stability of Pension Systems when Gains/Losses are Amortized and Rates of Return are Autoregressive. March 1992. 12 pages. ISBN 1 874 770 35 2

36. Haberman S. HIV, AIDS and the Approximate Calculation of Life Insurance Functions, Annuities and Net Premiums. April 1992. 28 pages. ISBN 1 874 770 36 0

37. Cooper D.R. Savings and Loans: An Historical Perspective. May 1992. 29 pages. ISBN 1 874 770 37 9

38. Verrall R.J. Dynamic Bayes Estimation for AIDS Delay Tables. May 1992. 16 pages. ISBN 1 874 770 38 7

39. Kaye G.D. Data Sources Relating to Life Assurance Expenses. May 1992. 39 pages. Presented to the Staple Inn Actuarial Society and Royal Statistical Society - February 1992. ISBN 1 874 770 39 5

40. Renshaw A.E., and Haberman S. On the Graduation Associated with a Multiple State Model for Permanent Health Insurance. May 1992. ISBN 1 874 770 40 9

41. England P.D. Statistical Modelling of Excess Mortality Number 3. June 1992. 163 pages. ISBN 1 874 770 41 7

42. Bloomfield D.S.F. and Haberman S. Male Social Class Mortality Differences Around 1981: An Extension to Include Childhood Ages. 21 pages. June 1992. ISBN 1 874 770 42 5

43. Berg M.P and Haberman S. Trend Analysis and Prediction Procedures for Time Nonhomogeneous Claim Processes. 33 pages. June 1992. ISBN 1 874 770 43 3

44. Booth P.M. The Single Market for Insurance, Free Capital Movements and attempts to Fix Exchange Rates. October 1992. 28 pages. ISBN 1 874 770 44 1

45. Verrall R.J. Chain Ladder with Varying Run-off Evolutions. February 1993. 15 pages. ISBN 1 874 770 45 X

46. Gavin J., Haberman S. and Verrall R.J. Moving Weighted Average Graduation using Kernel Estimation. November 1992. 14 pages. ISBN 1 874 770 46 8

47. Gavin J., Haberman S. and Verrall R.J. On the Choice of Bandwidth for Kernel Graduation. November 1992. 21 pages. ISBN 1 874 770 47 6

48. S. Haberman. Pension Funding with Time Delays and the Optimal Spread Period. May 1993. 13 pages. ISBN 1 874 770 48 4

49. S. Haberman. Stochastic Investment Returns and the Present Value of Future Contributions in a Defined Benefit Pension Scheme. May 1993. 22 pages. ISBN 1 874 770 49 2

50. A. Zimbidis and S. Haberman. Delay, Feedback and Variability of Pension Contributions and Fund Levels. May 1993. 19 pages. ISBN 1 874 770 50 6

51. S. Haberman. Pension Funding: The Effect of Changing The Frequency Valuations. June 1993. 19 pages. ISBN 1 874 770 51 4

52. S Haberman. HIV, AIDS Markov Chains and PHI. June 1993. 15 pages. ISBN 1 874 770 52 2

53. S Haberman. A Consideration of Pension Credit and Termination Insurance. June 1993. 22 pages. ISBN 1 874 770 53 0

54. M Z Khorasanee. Survey of Actuarial Practice in the Funding of UK Defined Benefit Pension Schemes. July 1993. 19 pages. ISBN 1 874 770 54 9

55. P M Booth, R G Chadburn and A S K Ong. A Utility Maximisation Approach to Asset Allocation. September 1993. 40 pages. ISBN 1 874 770 55 7

56. R G Chadburn. Bias in Select Mortality Investigations. August 1993. 62 pages. ISBN 1 874 770 56 5

57. M Z Khorasanee. A Comparison of Pension Funding Strategies by means of Simulations for a Model Scheme. August 1993. 43 pages. ISBN 1 874 770 57 3

58. A E Renshaw, P Hatzopolous and S Haberman. Recent Mortality Trends in Male Assured Lives. June 1993. 23 pages. ISBN 1 874 770 58 1

59. E Pitacco. Disability Risk Models: Towards A Unifying Approach. September 1993. 33 pages. ISBN 1 874 770 59 X

60. M Boskov and R J Verrall. Premium Rating by Geographic Area Using Spatial Models. September 1993. 14 pages. ISBN 1 874 770 60 3

61. R G Chadburn. Managing Mutual Life Offices: The Use of an Expense Ratio in New Business Decision Making and Expense Control. October 1993. 21 pages. ISBN 1 874 770 61 1

62. Haberman S. Pension Funding Modelling and Stochastic Investment Returns. 56 pages. March 1994. ISBN 1 874 770 62 X

63. Renshaw A E and Verrall R J. The Stochastic Model Underlying the Chain-Ladder Technique. 25 pages. April 1994. ISBN 1 874 770 63 8

64. Haberman S and Sung J-H. Dynamic Approaches to Pension Funding. 22 pages. April 1994. ISBN 1 874 770 64 6

65. Renshaw A.E. On the Second Moment Properties and the Implementation of Certain GLIM Based Stochastic Claims Reserving Models. 36 pages. September 1994. ISBN 1 874 770 65 4

66. Booth P.M., J.N. Allan, and J.W. Jang. An Evaluation of the UK Life Insurance Mismatching Test. September 1994. ISBN 1 874 770 66 2

67. Booth P.M. and Stroinski K. Insurance and Investment Markets in Poland. September 1994. 35 pages. ISBN 1 874 770 67 0

68. Ong A. A Stochastic Model for Treasury-Bills: An Extension to Wilkie's Model. September 1994. 12 pages. ISBN 1 874 770 68 9

69. Bloomfield D.S.F. Moving on from Undergraduate Exams to Professional Exams: Actuaries. November 1994. 22 pages. ISBN 1 874 770 69 7

70. Huber P. A Review of Wilkie's Stochastic Investment Model. January 1995. 22 pages. ISBN 1 874 770 70 0

71. Renshaw A.E. On the Graduation of `Amounts'. January 1995. 24 pages. ISBN 1 874 770 71 9

72. Renshaw A.E. Claims Reserving by Joint Modelling. December 1994. 26 pages. ISBN 1 874 770 72 7

73. Renshaw A.E. Graduation and Generalised Linear Models: An Overview. February 1995. 40 pages. ISBN 1 874 770 73 5

74. Khorasanee M.Z. Simulation of Investment Returns for a Money Purchase Fund. June 1995. 20 pages. ISBN 1 874 770 74 3

75. Owadally M.I. and Haberman S. Finite-time Pension Fund Dynamics with Random Rates of Return. June 1995. 28 pages. ISBN 1 874 770 75 1

76. Owadally M.I. and Haberman S. Stochastic Investment Modelling and Optimal Funding Strategies. June 1995. 25 pages. ISBN 1 874 770 76 X

77. Khorasanee M.Z Applying the Defined Benefit Principle to a Defined Contribution Scheme. August 1995. 30 pages. ISBN 1 874 770 77 8

78. Sebastiani P. and Settimi R. Experimental Design for Non-Linear Problems. September 1995. 13 pages. ISBN 1 874 770 78 6

79. Verrall R.J. Whittaker Graduation and Parametric State Space Models. November 1995. 23 pages. ISBN 1 874 770 79 4

80. Verrall R.J. Claims Reserving and Generalised Additive Models. November 1995. 17 pages. ISBN 1 874 770 80 8

81. Nelder J.A. and Verrall R.J. Credibility Theory and Generalized Linear Models. November 1995. 15 pages. ISBN 1 874 770 81 6

82. Renshaw A.E., Haberman S. and Hatzopoulos P. On The Duality of Assumptions Underpinning The Construction of Life Tables. December 1995. 17 Pages. ISBN 1 874 770 82 4

83. Chadburn R.G. Use of a Parametric Risk Measure in Assessing Risk Based Capital and Insolvency Constraints for With Profits Life Insurance. March 1996. 17 Pages. ISBN 1 874 770 84 0

84. Haberman S. Landmarks in the History of Actuarial Science (up to 1919). March 1996. 62 Pages. ISBN 1 874 770 85 9

85. Renshaw A.E. and Haberman S. Dual Modelling and Select Mortality. March 1996. 30 Pages. ISBN 1 874 770 88 3

86. Booth P.M. Long-Term Care for the Elderly: A Review of Policy Options. April 1996. 45 Pages. ISBN 1 874 770 89 1

87. Huber P.P. A Note on the Jump-Equilibrium Model. April 1996. 17 Pages. ISBN 1 874 770 90 5

88. Haberman S and Wong L.Y.P. Moving Average Rates of Return and the Variability of Pension Contributions and Fund Levels for a Defined Benefit Pension Scheme. May 1996. 51 Pages. ISBN 1 874 770 91 3

89. Cooper D.R. Providing Pensions for Employees with Varied Working Lives. June 1996. 25 Pages. ISBN 1 874 770 93 X

90. Khorasanee M.Z. Annuity Choices for Pensioners. August 1996. 25 Pages. ISBN 1 874 770 94 8

91. Verrall R.J. A Unified Framework for Graduation. November 1996. 25 Pages. ISBN 1 874 770 99 9

92. Haberman S. and Renshaw A.E. A Different Perspective on UK Assured Lives Select Mortality. November 1996. 61 Pages. ISBN 1 874 770 00 X

93. Booth P.M. The Analysis of Actuarial Investment Risk. March 1997. 43 Pages. ISBN 1 901615 03 0

94. Booth P.M., Chadburn R.G. and Ong A.S.K. Utility-Maximisation and the Control of Solvency for Life Insurance Funds. April 1997. 39 Pages. ISBN 1 901615 04 9

95. Chadburn R.G. The Use of Capital, Bonus Policy and Investment Policy in the Control of Solvency for With-Profits Life Insurance Companies in the UK. April 1997. 29 Pages. ISBN 1 901615 05 7

96. Renshaw A.E. and Haberman S. A Simple Graphical Method for the Comparison of Two Mortality Experiences. April 1997. 32 Pages. ISBN 1 901615 06 5

97. Wong C.F.W. and Haberman S. A Short Note on Arma (1, 1) Investment Rates of Return and Pension Funding. April 1997. 14 Pages. ISBN 1 901615 07 3

98. Puzey A S. A General Theory of Mortality Rate Estimators. June 1997. 26 Pages. ISBN 1 901615 08 1

99. Puzey A S. On the Bias of the Conventional Actuarial Estimator of q_x. June 1997. 14 Pages. ISBN 1 901615 09 X

100. Walsh D. and Booth P.M. Actuarial Techniques in Pricing for Risk in Bank Lending. June 1997. 55 Pages. ISBN 1 901615 12 X

101. Haberman S. and Walsh D. Analysis of Trends in PHI Claim Inception Data. July 1997. 51 Pages. ISBN 1 901615 16 2

102. Haberman S. and Smith D. Stochastic Investment Modelling and Pension Funding: A Simulation Based Analysis. November 1997. 91 Pages. ISBN 1 901615 19 7

103. Rickayzen B.D. A Sensitivity Analysis of the Parameters used in a PHI Multiple State Model. December 1997. 18 Pages. ISBN 1 901615 20 0

104. Verrall R.J. and Yakoubov Y.H. A Fuzzy Approach to Grouping by Policyholder Age in General Insurance. January 1998. 18 Pages. ISBN 1 901615 22 7

105. Yakoubov Y.H. and Haberman S. Review of Actuarial Applications of Fuzzy Set Theory. February 1998. 88 Pages. ISBN 1 901615 23 5

106. Haberman S. Stochastic Modelling of Pension Scheme Dynamics. February 1998. 41 Pages. ISBN 1 901615 24 3

107. Cooper D.R. A Re-appraisal of the Revalued Career Average Benefit Design for Occupational Pension Schemes. February 1998. 12 Pages. ISBN 1 901615 25 1

108. Wright I.D. A Stochastic Asset Model using Vector Auto-regression. February 1998. 59 Pages. ISBN 1 901615 26 X

109. Huber P.P. and Verrall R.J. The Need for Theory in Actuarial Economic Models. March 1998. 15 Pages. ISBN 1 901615 27 8

110. Booth P.M. and Yakoubov Y. Investment Policy for Defined Contribution Pension Scheme Members Close to Retirement. May 1998. 32 Pages ISBN 1 901615 28 6

111. Chadburn R.G. A Genetic Approach to the Modelling of Sickness Rates, with Application to Life Insurance Risk Classification. May 1998. 17 Pages. ISBN 1 901615 29 4

112. Wright I.D. A Stochastic Approach to Pension Scheme Funding. June 1998. 24 Pages. ISBN 1 901615 30 8

113. Renshaw A.E. and Haberman S. Modelling the Recent Time Trends in UK Permanent Health Insurance Recovery, Mortality and Claim Inception Transition Intensities. June 1998. 57 Pages. ISBN 1 901615 31 6

Statistical Research Papers

1. Sebastiani P. Some Results on the Derivatives of Matrix Functions. December 1995. 17 Pages. ISBN 1 874 770 83 2

2. Dawid A.P. and Sebastiani P. Coherent Criteria for Optimal Experimental Design. March 1996. 35 Pages. ISBN 1 874 770 86 7

3. Sebastiani P. and Wynn H.P. Maximum Entropy Sampling and Optimal Bayesian Experimental Design. March 1996. 22 Pages. ISBN 1 874 770 87 5

4. Sebastiani P. and Settimi R. A Note on D-optimal Designs for a Logistic Regression Model. May 1996. 12 Pages. ISBN 1 874 770 92 1

5. Sebastiani P. and Settimi R. First-order Optimal Designs for Non Linear Models. August 1996. 28 Pages. ISBN 1 874 770 95 6

6. Newby M. A Business Process Approach to Maintenance: Measurement, Decision and Control. September 1996. 12 Pages. ISBN 1 874 770 96 4

7. Newby M. Moments and Generating Functions for the Absorption Distribution and its Negative Binomial Analogue. September 1996. 16 Pages. ISBN 1 874 770 97 2

8. Cowell R.G. Mixture Reduction via Predictive Scores. November 1996. 17 Pages. ISBN 1 874 770 98 0

9. Sebastiani P. and Ramoni M. Robust Parameter Learning in Bayesian Networks with Missing Data. March 1997. 9 Pages. ISBN 1 901615 00 6

10. Newby M.J. and Coolen F.P.A. Guidelines for Corrective Replacement Based on Low Stochastic Structure Assumptions. March 1997. 9 Pages. ISBN 1 901615 01 4.

11. Newby M.J. Approximations for the Absorption Distribution and its Negative Binomial Analogue. March 1997. 6 Pages. ISBN 1 901615 02 2

12. Ramoni M. and Sebastiani P. The Use of Exogenous Knowledge to Learn Bayesian Networks from Incomplete Databases. June 1997. 11 Pages. ISBN 1 901615 10 3

13. Ramoni M. and Sebastiani P. Learning Bayesian Networks from Incomplete Databases. June 1997. 14 Pages. ISBN 1 901615 11 1

14. Sebastiani P. and Wynn H.P. Risk Based Optimal Designs. June 1997. 10 Pages. ISBN 1 901615 13 8

15. Cowell R. Sampling without Replacement in Junction Trees. June 1997. 10 Pages. ISBN 1 901615 14 6

16. Dagg R.A. and Newby M.J. Optimal Overhaul Intervals with Imperfect Inspection and Repair. July 1997. 11 Pages. ISBN 1 901615 15 4

17. Sebastiani P. and Wynn H.P. Bayesian Experimental Design and Shannon Information. October 1997. 11 Pages. ISBN 1 901615 17 0

18. Wolstenholme L.C. A Characterisation of Phase Type Distributions. November 1997. 11 Pages. ISBN 1 901615 18 9

19. Wolstenholme L.C. A Comparison of Models for Probability of Detection (POD) Curves. December 1997. 23 Pages. ISBN 1 901615 21 9

Department of Actuarial Science and Statistics

Actuarial Research Club

The support of the corporate members

Commercial Union
Coopers & Lybrand
Government Actuary's Department
Guardian Insurance
Hymans Robertson
KPMG
Munich Reinsurance
Swiss Reinsurance

is gratefully acknowledged.

ISBN 1 901615 31 6