Biometrika (2008), **95**, 4, *pp*. 979–986
doi: 10.1093/biomet/asn026

Printed in Great Britain
Advance Access publication 15 September 2008

Miscellanea

Identification of the age-period-cohort model and the extended chain-ladder model

BY D. KUANG

Department of Statistics, University of Oxford, Oxford OX1 3TG, U.K.
di.kuang@some.ox.ac.uk

B. NIELSEN

Nuffield College, Oxford OX1 1NF, U.K.
bent.nielsen@nuffield.ox.ac.uk

AND J. P. NIELSEN

Cass Business School, City University London, 106 Bunhill Row, London EC1Y 8TZ, U.K.
Jens.Nielsen.1@city.ac.uk

SUMMARY

We consider the identification problem that arises in the age-period-cohort models as well as in the extended chain-ladder model. We propose a canonical parameterization based on the accelerations of the trends in the three factors. This parameterization is exactly identified and eases interpretation, estimation and forecasting. The canonical parameterization is applied to a class of index sets which have trapezoidal shapes, including various Lexis diagrams and the insurance-reserving triangles.

Some key words: Age-period-cohort model; Chain-ladder model; Identification.

1. INTRODUCTION

The age-period-cohort model used in epidemiology and demography describes the logarithm of mortality in an additive form, involving three interlinked time scales,

$$\mu_{ij} = \alpha_i + \beta_j + \gamma_{i+j-1} + \delta, \tag{1}$$

where i is the cohort, j is the age and $i + j - 1$ is the period. The indices i and j vary bivariately in an index set $I \in N^2$. The parameters of the model, $\alpha_i, \beta_j, \gamma_{i+j-1}$ and δ, describe the trends of the three factors in the model. It has long been appreciated that this parameterization is not identified. Holford (1983) therefore used generalized inverses when solving maximum-likelihood equations, remarking that the choice of generalized inverse can have a large effect on the parameter estimates. A similar solution has implicitly been used in the insurance literature; see Zehnwirth (1994). Clayton & Schifflers (1987) suggested that the ratios of the relative risks are identifiable. On a logarithmic scale, they are the second differences, which will be the key element in this paper. Carstensen (2007) represented the variation of the parameterization of (1) by adding and subtracting linear trends from $\alpha_i, \beta_j, \gamma_{i+j-1}$ and δ, which relates to a group-theoretic description of the identification suggested here. He also pointed out that an ideal parameterization should be simple in both estimation and computation.

In this paper, we revisit the identification problem. We propose a canonical parameterization which includes the identifiable second differences suggested by Clayton & Schifflers (1987), and prove that it has a one-to-one correspondence with μ_{ij}, for all $(i, j) \in I$. It will be shown that the interpretation of such a parameterization is straightforward, and that its design matrix can easily be deduced. Initially, we consider the simple case in which the indices i and j vary in a square. The canonical parameter ξ is then given by the second differences of α_i, β_j and γ_{i+j-1} and the three corner points μ_{11}, μ_{21} and μ_{12}. We then proceed to show that the three corner points can be replaced by three other points. Finally, this is extended to more general index sets.

As examples, we shall consider the three leading cases of the age-period-cohort model related to the Lexis diagram as discussed by Keiding (1990). In the terminology of Keiding, the first principal set of dead is data from certain cohorts that die within a given age range. This is where the indices vary in an age–cohort rectangle. The other two cases are where the indices vary in an age–cohort trapezoid. The second principal set of dead studies the deaths of certain cohorts in a given period, as in a longitudinal study. The third principal set of dead studies the death within a certain period and within a given age range, as in a repeated cross-sectional study.

We shall also consider the extended chain-ladder model used for reserving in non-life insurance. The issue in reserving is that claims relating to a given accident year may be reported many years after the accident. Thus, the available data in any given calendar year k are a simplex of size k with claims indexed by their accident year and by their reporting or development year. The accident year and the development year add up to the calendar year plus 1. This simplex is referred to as a run-off triangle. The classical chain-ladder model, discussed for instance by England & Verrall (2002), involves only two time scales relating to the accident and the development year. An extended chain-ladder model parameterized using three time scales as in (1) has been introduced by Zehnwirth (1994) and Barnett & Zehnwirth (2000).

2. Identification for square index sets

Consider a simple square index set given by Definition 1. In this section, we propose a canonical parameterization for model (1) for this situation.

Definition 1. *The set I is a square index set if, for some $k \in N$, $I = \{(i, j): i, j = 1, \ldots, k\}$.*

For a square index set, the parameters of (1) are

$$\theta = (\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_k, \gamma_1, \ldots, \gamma_{2k-1}, \delta) \in \mathbb{R}^{4k}.$$

Now let $\mu = \{\mu_{ij} : (i, j) \in I\}$, as given by (1). The map from θ to μ is surjective, but not injective. As pointed out by Carstensen (2007), linear trends in α_i, β_j and γ_{i+j-1} can be added without changing the value of μ_{ij}. This can be phrased as θ being overparameterized.

Clayton & Schifflers (1987) worked with a multiplicative formulation of (1) and suggested that ratios of ratios of the parameters would be invariant. In the linear set-up (1), the linear trends can correspondingly be removed from α_i, β_j and γ_{i+j-1} by taking second differences, such as $\Delta^2\alpha_i = \alpha_i - 2\alpha_{i-1} + \alpha_{i-2}$. To generate a canonical parameterization ξ, we rewrite (1) in terms of these second differences, and three initial points. This can be done by introducing the telescopic sums $\alpha_i = \alpha_1 + \sum_{t=2}^{i} \Delta\alpha_t$ and $\Delta\alpha_t = \Delta\alpha_2 + \sum_{s=3}^{t} \Delta^2\alpha_s$, so that

$$\alpha_i = \alpha_1 + (i-1)\Delta\alpha_2 + \sum_{t=3}^{i}\sum_{s=3}^{t} \Delta^2\alpha_s.$$

Substitute this expression for α_i and similar expressions for β_j and γ_{i+j-1} into (1). Writing $\triangle\alpha_2 + \triangle\gamma_2 = \mu_{21} - \mu_{11}$ and $\triangle\beta_2 + \triangle\gamma_2 = \mu_{12} - \mu_{11}$, we obtain

$$\mu_{ij} = \mu_{11} + (i-1)(\mu_{21} - \mu_{11}) + (j-1)(\mu_{12} - \mu_{11}) + a_{ij}, \tag{2}$$

for all $i, j \in I$, where

$$a_{ij} = \sum_{t=3}^{i}\sum_{s=3}^{t} \Delta^2\alpha_s + \sum_{t=3}^{j}\sum_{s=3}^{t} \Delta^2\beta_s + \sum_{t=3}^{i+j-1}\sum_{s=3}^{t} \Delta^2\gamma_s.$$

The expression for a_{ij} can be simplified further by exchanging the double sums; for instance, $\sum_{t=3}^{i}\sum_{s=3}^{t} \Delta^2\alpha_s$ equals $\sum_{s=3}^{i}(i-s+1)\Delta^2\alpha_s$, and thus

$$a_{ij} = \sum_{s=3}^{i}(i-s+1)\Delta^2\alpha_s + \sum_{s=3}^{j}(j-s+1)\Delta^2\beta_s + \sum_{s=3}^{i+j-1}(i+j-s)\Delta^2\gamma_s. \tag{3}$$

Based on formula (2), we define a parameter vector $\xi \in \mathbb{R}^{4k-4}$ as

$$\xi = \left(\mu_{11}, \mu_{21}, \mu_{12}, \Delta^2\alpha_3, \ldots, \Delta^2\alpha_k, \Delta^2\beta_3, \ldots, \Delta^2\beta_k, \Delta^2\gamma_3, \ldots, \Delta^2\gamma_{2k-1}\right). \tag{4}$$

Theorem 1 below shows that ξ gives a unique parameterization of μ. We therefore call it a canonical parameter.

For estimation, a design matrix for the canonical parameter ξ can be deduced from (2). In the case of a square index set, I, of dimension $k = 3$, the design matrix is given by

$$\begin{pmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{21} \\ \mu_{22} \\ \mu_{13} \\ \mu_{31} \\ \mu_{23} \\ \mu_{32} \\ \mu_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 2 & 0 & 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 2 & 0 & 1 & 2 & 1 & 0 \\ 1 & 2 & 1 & 1 & 0 & 2 & 1 & 0 \\ 1 & 2 & 2 & 1 & 1 & 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} \mu_{11} \\ \mu_{21} - \mu_{11} \\ \mu_{12} - \mu_{11} \\ \Delta^2\alpha_3 \\ \Delta^2\beta_3 \\ \Delta^2\gamma_3 \\ \Delta^2\gamma_4 \\ \Delta^2\gamma_5 \end{pmatrix}.$$

Theorem 1 shows that ξ is unique in general. The uniqueness of ξ implies that the design matrix has full column rank; this can be checked by inspection when $k = 3$. The proof is provided in the Appendix.

THEOREM 1. *Let $\mu = \{\mu_{ij} : (i, j) \in I\}$, where I is a square index set, and μ_{ij} satisfies (1). The parameterization ξ given by (4) satisfies that*

(i) *ξ is a function of θ, and*
(ii) *μ is a function of ξ, because of (2).*

The parameterization of μ by ξ is exactly identified in that $\xi^\dagger \neq \xi^\ddagger$ implies $\mu(\xi^\dagger) \neq \mu(\xi^\ddagger)$.

The result could also be cast in terms of group-theoretic arguments. As in Carstensen (2007), we define the group g by

$$g : \begin{pmatrix} \alpha_i \\ \beta_j \\ \gamma_{i+j-1} \\ \delta \end{pmatrix} \mapsto \begin{Bmatrix} \alpha_i + a + (i-1)d \\ \beta_j + b + (j-1)d \\ \gamma_{i+j-1} + c - (i+j-2)d \\ \delta - a - b - c \end{Bmatrix}, \tag{5}$$

where a, b, c and d are arbitrary constants. The parameter μ is a function of θ, which is invariant to g; that is, $\mu(\theta) = \mu\{g(\theta)\}$. Based on invariance arguments such as those in Cox & Hinkley (1974, § 5·3), Theorem 1 shows that ξ is a maximal invariant function of θ under g.

The assigned parameter θ can be constructed from ξ using (5). For instance, if we choose $\alpha_1 = \beta_1 = \gamma_1 = \gamma_2 = 0$, then θ can be computed from ξ as

$$\alpha_i = (i-1)(\mu_{21} - \mu_{11}) + \sum_{t=3}^{i}\sum_{s=3}^{t} \Delta^2\alpha_s,$$

$$\beta_j = (j-1)(\mu_{12}-\mu_{11}) + \sum_{t=3}^{j}\sum_{s=3}^{t}\Delta^2\beta_s,$$

$$\gamma_{i+j-1} = \sum_{t=3}^{i+j-1}\sum_{s=3}^{t}\Delta^2\gamma_s,$$

$$\delta = \mu_{11}.$$

Formula (2) shows that these components add up to μ_{ij}. If other values for α_1, β_1, γ_1 and γ_2 are desired, corresponding linear trends can be added as set out in (5) by choosing appropriate values of levels a, b and c and slope d.

Since we can choose a, b, c and d arbitrarily, interpretation of the original parameters α_i, β_j and γ_{i+j-1} is difficult. The visual impression of the parameters α_i, β_j and γ_{i+j-1} will depend on the choice of a, b, c and d. For instance, if we vary d, a plot of the α_i-parameters can appear to be increasing or decreasing. Correspondingly the level, and hence the sign, of the first differences $\Delta\alpha_i$ is arbitrary. However, the second differences $\Delta^2\alpha_i$, $\Delta^2\beta_j$ and $\Delta^2\gamma_{i+j-1}$ do have a unique interpretation. The interpretation of such second differences, or accelerations, is standard in time-series analysis. Likewise, any forecasting can be done more safely on the second differences rather than the levels. In applications, it would therefore be helpful to make graphs of the second differences.

In some applications the components α_i, β_j, γ_{i+j-1} and δ themselves are unimportant, whereas the original parameters μ_{ij} are of main interest. Plots of the parameters μ_{ij} will be meaningful as μ_{ij} is a function of ξ via (2) and (3), and it is therefore identified. An example is when the objective of interest is to forecast how many children there will be in different grades in the school system in the year 2010. In that case, let i be cohort and j be age, and plot μ_{ij} as a function of either age or cohort such that the period is $i+j-1 = 2010$. Other examples could be how mortality of people of age 80 varies with either the period or the cohort, or how mortality of people born in 1930 varies with either period or age. Similarly, in insurance the intrinsic issue is to predict outstanding claims relating to a given accident year rather than to forecast the calendar parameters γ_k, say.

3. The role of initial points

The choice of canonical parameterization is not unique. Any bijective mapping of ξ would also identify μ exactly. In particular, the three initial points in ξ given by (4) can be replaced by another set of three points without changing the content of Theorem 1.

The argument for changing the initial points is based on a manipulation of equation (2). It is convenient to introduce the matrix notation

$$Y = \begin{pmatrix}\mu_{i_1j_1}\\ \mu_{i_2j_2}\\ \mu_{i_3j_3}\end{pmatrix}, \quad X = \begin{pmatrix}\mu_{11}\\ \mu_{21}-\mu_{11}\\ \mu_{12}-\mu_{11}\end{pmatrix}, \quad A = \begin{pmatrix}a_{i_1j_1}\\ a_{i_2j_2}\\ a_{i_3j_3}\end{pmatrix}, \quad B = \begin{pmatrix}b_{i_1j_1}\\ b_{i_2j_2}\\ b_{i_3j_3}\end{pmatrix},$$

with a_{ij} as in formula (3) and $b_{ij} = (1, i-1, j-1)$. With this notation, it holds from (2) that

$$Y = BX + A, \tag{6}$$

which can be solved for X when B is invertible. We find that B is invertible when $\det(B) = i_2j_3 - i_3j_2 + i_3j_1 - i_1j_3 + i_1j_2 - i_2j_1$ is different from zero. As a consequence, X can be replaced by Y, which gives a new canonical parameter,

$$\xi = \left(\mu_{i_1j_1}, \mu_{i_2j_2}, \mu_{i_3j_3}, \Delta^2\alpha_3, \ldots, \Delta^2\alpha_k, \Delta^2\beta_3, \ldots, \Delta^2\beta_k, \Delta^2\gamma_3, \ldots, \Delta^2\gamma_{2k-1}\right). \tag{7}$$

The following corollary to Theorem 1 holds.

COROLLARY 1. *Suppose μ_{ij} satisfies (1) on a square index set I and consider the parameter ξ given by (7). If the matrix B is invertible, then the conclusions of Theorem 1 remain true.*

A design matrix can be constructed from ξ as given by (7). This is done by combining (6) and (2). This shows that, for all $(i, j) \in I$, it holds that

$$\mu_{ij} = b_{ij}X + a_{ij} = b_{ij}B^{-1}Y + \left(a_{ij} - b_{ij}B^{-1}A\right), \tag{8}$$

which is a linear function of ξ as defined in (7). The inverse of the matrix B is given by

$$B^{-1} = \frac{1}{\det(B)} \begin{pmatrix} \{(i_2-1)(j_3-j_2) & \{(i_1-1)(j_1-j_3) & \{(i_1-1)(j_2-j_1) \\ -(i_3-i_2)(j_2-1)\} & -(i_1-i_3)(j_1-1)\} & -(i_2-i_1)(j_1-1)\} \\ j_2-j_3 & j_3-j_1 & j_1-j_2 \\ i_3-i_2 & i_1-i_3 & i_2-i_1 \end{pmatrix}.$$

4. Identification for general index sets

In many situations, the parameterization (1) has an index set which is not a square as considered in § 2. For instance, I can be a parallelogram in a Lexis diagram, or a simplex in an insurance run-off triangle. More generally, I could be any irregular shape, with one or more missing points. It is therefore useful to construct a canonical parameterization for (1) with a non-square index set.

A convenient generalization of the square index sets is index sets of rectangular shapes, where the period $i + j - 1$ can be constrained to a certain interval. We will call such index sets generalized trapezoids and give a precise definition below. With such index sets it is immediately clear how to define a canonical parameterization from knowing the dimensions of the generalized trapezoid. The generalized trapezoid covers the most important situations encountered in practice, namely the three types of Lexis diagram and the insurance run-off triangle.

Definition 2. *The index set I is a generalized trapezoid if, for some $l, k, m \in N$, $h \in N_0$, and $h + m \leqslant l + k - 1$, then*

$$I = \{(i, j) \colon i = 1, \ldots, k,\ j = 1, \ldots, l,\ i + j - 1 = h + 1, \ldots, h + m\}.$$

In the following, we illustrate with diagrams some applications of the general trapezoid. Figure 1(a)–(c) shows examples of the three types of Lexis diagram discussed by Keiding (1990). The first principal set of dead gives a rectangular index set, whereas the second and third principal sets of dead are trapezoids. Figure 1(d) gives an example of an insurance run-off triangle as discussed by Zehnwirth (1994) and Barnett & Zehnwirth (2000).

For every generalized trapezoid I, we define the canonical parameter ξ from the dimensions l, k, m and h by restricting the three time scales; that is,

$$\xi = \left(\mu_{i_1,j_1}, \mu_{i_2,j_2}, \mu_{i_3,j_3}, \Delta^2\alpha_3, \ldots, \Delta^2\alpha_k, \Delta^2\beta_3, \ldots, \Delta^2\beta_l, \Delta^2\gamma_{h+3}, \ldots, \Delta^2\gamma_{h+m}\right), \tag{9}$$

where $(i_1, j_1), (i_2, j_2), (i_3, j_3) \in I$ and satisfy $\det(B) \neq 0$. The following corollary to Theorem 1 then holds.

Corollary 2. *Suppose μ_{ij} satisfies* (1) *on a generalized trapezoid index set I and consider the parameter ξ given by* (9)*. If the matrix B is invertible, then Theorem 1 remains true.*

Corollary 2 is proved by analyzing formula (8) and showing that the terms $\Delta^2\gamma_3, \ldots, \Delta^2\gamma_{h+2}$ are not needed. To see this, we isolate the $\Delta^2\gamma$ terms in a_{ij} of (8), that is $\sum_{s=3}^{i+j-1}(i + j - s)\Delta^2\gamma_s = \sum_{s=3}^{i+j}(i + j - s)\Delta^2\gamma_s$. It is then seen that $\Delta^2\gamma_s$ with index $s < h + 3$ is weighted by

$$w_s = (i + j - s) - b_{ij}B^{-1}(i_1 + j_1 - s, i_2 + j_2 - s, i_3 + j_3 - s)'.$$

The last vector can easily be written in terms of the matrix B, giving

$$w_s = (i + j - s) - (1, i - 1, j - 1)B^{-1}B(2 - s, 1, 1)' = 0.$$

A design matrix can be constructed from ξ as given by (9). As in § 3, this is done directly from formula (8). The design matrix has a number of zero elements; for instance, we can

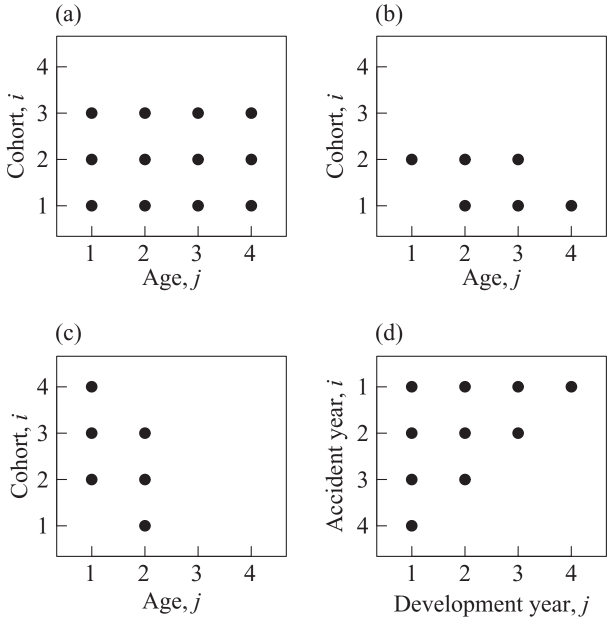

Fig. 1. Panels (a)–(c) show Lexis diagrams for first, second and third, respectively, principal sets of dead. Panel (d) shows an insurance run-off triangle.

show that $\Delta^2\gamma_{h+3}, \ldots, \Delta^2\gamma_{h+p}$, with $p \geqslant 3$, has weight zero if $i_1 + j_1 - 1, i_2 + j_2 - 1, i_3 + j_3 - 1$ and also $i + j - 1$ are all larger than $h + p$, following a procedure similar to the proof of Corollary 2.

Corollary 2 gives a sufficient condition only for the type of index set in which ξ in (9) is a canonical parameterization. Figure 2(a) shows an example of an index set which is not a generalized trapezoid. Figure 2(b) shows an extended index set which is a generalized trapezoid. Corollary 2 gives a canonical parameter ξ for the latter set. This parameter ξ is also a canonical parameter for the original set. To see this, we decompose ξ into elements ξ_6 say, related to the 3×3 simplex and the second differences $\Delta^2\beta_4$, $\Delta^2\gamma_4$ and $\Delta^2\gamma_5$. It turns out that there is a bijective mapping from those three elements to μ_{14}, μ_{24} and μ_{33}, which can be formulated as

$$\begin{pmatrix} \mu_{14} \\ \mu_{24} \\ \mu_{33} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} \Delta^2\beta_4 \\ \Delta^2\gamma_4 \\ \Delta^2\gamma_5 \end{pmatrix} + f(\xi_6),$$

where $f(\cdot)$ contains some functions of ξ_6. The design matrix here has rank three. However, if any one of the three points μ_{14}, μ_{24} and μ_{33} is missing, ξ would be overparameterized.

In Fig. 2(c), a canonical parameter for a general index set was found by extending the set to a generalized trapezoid. However, this strategy will not work in general, as shown by Figs. 2(c) and (d). In Fig. 2(c), the index set has four points. By adding the point μ_{22}, we obtain a generalized trapezoid with canonical parameter ξ of dimension five, which overparameterizes the original set. In the second, the same occurs when μ_{21} is added.

5. Discussion

While the canonical parameter is indeed unique, its interpretations are in terms of accelerations, which can be somewhat complicated to communicate. The level parameters α_i, β_j and γ_{i+j-1} are not unique, and their plots can be visually misleading as they evolve around arbitrarily chosen linear trends. In some

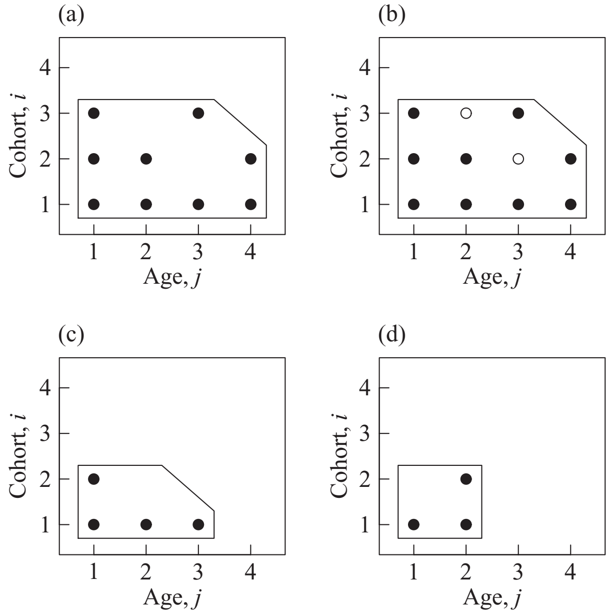

Fig. 2. Panel (a) shows an index set which is not a generalized trapezoid. This set is extended to a generalized trapezoid in panel (b). The sets in panels (a) and (b) have the same canonical parameter. Panels (c) and (d) show examples of sets which are not generalized trapezoids and for which the associated generalized trapezoid has a canonical parameter of larger dimension than the set itself.

applications, one could instead communicate plots of the original parameter μ_{ij} for a fixed value of i, j or $i + j - 1$.

Another sufficient condition for the permissible index sets can be based on a recursive argument. First, find a set $I_k \subset I$, which is a generalized trapezoid with canonical parameter ξ_k. Then add a point $(i, j) \in I \setminus I_k$ to I_k if it introduces at most one double difference that is not in ξ_k. Thus, $I_{k+1} = \{I_k \bigcup (i, j)\} \subseteq I$ is exactly identified by ξ_{k+1}. Figure 2(a) shows a set that cannot be obtained by this one-step recursive scheme, but by adding three points, $(1, 4)$, $(2, 4)$ and $(3, 3)$, to the identifiable 3×3 simplex.

Acknowledgement

The graphs were drawn in R (R Development Core Team, 2006). The second author received a grant from the U.K. Economic and Social Research Council.

Appendix

Proof of Theorem 1

Theorem 1 is proven by induction using formula (2).

As a preliminary step, consider the first two diagonals, i.e. the initial three elements. If one of $\mu^{\dagger}_{11} \neq \mu^{\ddagger}_{11}$, $\mu^{\dagger}_{12} \neq \mu^{\ddagger}_{12}$ or $\mu^{\dagger}_{21} \neq \mu^{\ddagger}_{21}$ holds, then the statement is true.

Then, to initialize the argument, consider the third diagonal. If $\mu^{\dagger}_{11} = \mu^{\ddagger}_{11}$, $\mu^{\dagger}_{12} = \mu^{\ddagger}_{12}$ and $\mu^{\dagger}_{21} = \mu^{\ddagger}_{21}$, but $\Delta^2\gamma^{\dagger}_3 \neq \Delta^2\gamma^{\ddagger}_3$, then $\mu^{\dagger}_{22} \neq \mu^{\ddagger}_{22}$ by formula (2).

If $\mu^{\dagger}_{11} = \mu^{\ddagger}_{11}$, $\mu^{\dagger}_{21} = \mu^{\ddagger}_{21}$, $\mu^{\dagger}_{12} = \mu^{\ddagger}_{12}$ and $\Delta^2\gamma^{\dagger}_3 = \Delta^2\gamma^{\ddagger}_3$, but $\Delta^2\beta^{\dagger}_3 \neq \Delta^2\beta^{\ddagger}_3$ or $\Delta^2\alpha^{\dagger}_3 \neq \Delta^2\alpha^{\ddagger}_3$, then $\mu^{\dagger}_{1,3} \neq \mu^{\ddagger}_{1,3}$ or $\mu^{\dagger}_{3,1} \neq \mu^{\ddagger}_{3,1}$ by formula (2).

For the induction step, consider the diagonal $(r+1)$, where $r+1 = 4, \ldots, 2k-1$. Assume that $\mu^{\dagger}_{11} = \mu^{\ddagger}_{11}$, $\mu^{\dagger}_{12} = \mu^{\ddagger}_{12}$ and $\mu^{\dagger}_{21} = \mu^{\ddagger}_{21}$, and, for $s = 3, \ldots, r$, that $\Delta^2\gamma^{\dagger}_s = \Delta^2\gamma^{\ddagger}_s$, $\Delta^2\beta^{\dagger}_s = \Delta^2\beta^{\ddagger}_s$ and $\Delta^2\alpha^{\dagger}_s = \Delta^2\alpha^{\ddagger}_s$ but $\Delta^2\gamma^{\dagger}_{r+1} \neq \Delta^2\gamma^{\ddagger}_{r+1}$. Then $\mu^{\dagger}_{2,r} \neq \mu^{\ddagger}_{2,r}$ by formula (2).

We then can show that $\mu^{\dagger}_{11} = \mu^{\ddagger}_{11}$, $\mu^{\dagger}_{12} = \mu^{\ddagger}_{12}$ and $\mu^{\dagger}_{21} = \mu^{\ddagger}_{21}$, for $s = 3, \ldots, r$, $\Delta^2\gamma^{\dagger}_s = \Delta^2\gamma^{\ddagger}_s$, $\Delta^2\beta^{\dagger}_s = \Delta^2\beta^{\ddagger}_s$, $\Delta^2\alpha^{\dagger}_s = \Delta^2\alpha^{\ddagger}_s$, $\Delta^2\gamma^{\dagger}_{r+1} = \Delta^2\gamma^{\ddagger}_{r+1}$, but $\Delta^2\beta^{\dagger}_{r+1} \neq \Delta^2\beta^{\ddagger}_{r+1}$ or $\Delta^2\alpha^{\dagger}_{r+1} \neq \Delta^2\alpha^{\ddagger}_{r+1}$. Then we have $\mu^{\dagger}_{1,r+1} \neq \mu^{\ddagger}_{1,r+1}$ or $\mu^{\dagger}_{r+1,1} \neq \mu^{\ddagger}_{r+1,1}$ by formula (2).

References

Barnett, G. & Zehnwirth, B. (2000). Best estimates for reserves. *Proc. Casualty Actuar. Soc.* **87**, 245–321.
Carstensen, B. (2007). Age-period-cohort models for the Lexis diagram. *Statist. Med.* **26**, 3018–45.
Clayton, D. & Schifflers, E. (1987). Models for temporal variation in cancer rates. II: Age-period-cohort models. *Statist. Med.* **6**, 469–81.
Cox, D. R. & Hinkley, D. V. (1974). *Theoretical Statistics*. London: Chapman and Hall.
England, P. D. & Verrall, R. J. (2002). Stochastic claims reserving in general insurance. *Br. Actuar. J.* **8**, 519–44.
Holford, T. R. (1983). The estimation of age, period and cohort effects for vital rates. *Biometrics* **39**, 311–24.
Keiding, N. (1990). Statistical inference in the Lexis diagram. *Phil. Trans. R. Soc.* A **332**, 487–509.
R Development Core Team. (2006). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
Zehnwirth, B. (1994). Probabilistic development factor models with applications to loss reserve variability, prediction intervals, and risk based capital. *Casualty Actuar. Soc. Forum: 1994 Spring Forum*, 447–605.

[*Received June* 2007. *Revised January* 2008]