

I welcome the opportunity to address the comments made by Murphy (2009) on Richards (2008). As Murphy points out, “the term cohort effect [...] is widely used” — Willets (1999) and Willets (2004) are seminal papers in the actuarial literature here. I am sympathetic to his semantic point about “patterns” instead of “effects”. So sympathetic, in fact, that I have even used his preferred term “pattern” in the title. However, I have also used the term “cohort effect” because it is widely used and understood, not only in the actuarial community but also by medics, academics and demographers working outside financial services, e.g. Doll *et al* (1997), Lawlor *et al* (2001) and Doblhammer (2004).

It is not quite clear if Murphy is purely arguing a point of semantics, or whether he has a particular problem with the idea of there being a causal mechanism operating on a cohort basis, although his comment about “the absence of evidence for clear-cut causal mechanisms” does strongly suggest the latter. However, evidence does indeed point towards there being plausible causal mechanisms: Figure 8 in my paper suggests there is a link for lung cancer, for example. Lawlor *et al* (2001) observe that “sex ratios for mortality [...] for lung cancer show a cohort effect”, while according to Doll *et al* (1997) “the trends [in lung cancer mortality] provide an almost perfect example of a cohort effect”. This latter article appeared in a book co-edited by Murphy himself — Charlton and Murphy (1997) — so I am left unsure if he really is disputing that there are causal mechanisms which operate on a cohort basis.

Note also that cohort effects are not restricted to the narrow example of lung cancer. For example, Davenport *et al* (1969) wrote of the influenza virus that “the major antigens of the strains of first infection of childhood permanently orient the antibody-forming mechanisms so that [...] the cohort [...] would respond with marked reinforcement of the primary antibody”. One reason for continuing interest in cohort effects is that there is a pressing need for society to know more, particularly in relation to the ultimate costs of pension provision. To illustrate the speed of change, an analysis of 1,138 deficit-recovery plans submitted to the Pensions Regulator (2008) revealed that 67% were reserving using projections of cohort effects. Just five years earlier this proportion would have been zero.

Murphy questions the use of the modal age at death and asks why I did not use either the mean or the median. In the first case the answer is simple: the mean age at death cannot be comparable across cohorts with differing degrees of closeness to extinction. An example is given in Figure 1 of my paper, which contrasts the cohorts born in 1901 and 1914. As the later-born cohort, the 1914 generation naturally has thirteen fewer ages at which death counts are available. Since the missing ages are obviously all at older ages, this would automatically give it a *lower* mean age at death than the 1901 cohort. Since Figure 1 clearly shows that the 1914 generation is dying later, using the mean age at death would be wholly misleading as a means of comparing these two cohorts. Similar comments apply to the median age at death. However, if the exposure figures are combined with the death counts to calculate rates, then the option of using the median survival age would also be sensible as the 50% point on the survival curve would be similarly robust. This latter point may have been what Murphy meant, but the point of this part of the paper was to explore what could be done with deaths data alone in the face of uncertain reliability of the population estimates.

Murphy asks why I focused on the modal age in later life, thus ignoring the fact that there is a mode at age zero which has historically often been higher than the mode in later life. The answer lies in the first line of the abstract: “*Late-life mortality patterns* are of crucial interest to actuaries assessing [...] annuities and defined-benefit pension schemes” (my emphasis). The introduction makes it plain that the focus of interest is the longevity risk in pension schemes for UK companies. Immediately post-natal mortality patterns are not strongly enough relevant for this, notwithstanding the insights of Doblhammer (2004) and many others. As an aside, it is interesting to note that for *period* mortality in England & Wales, deaths at age zero ceased to be the mode across the entire age range for males in 1968, and before 1961 for females. It is clearly too early yet to know, but it seems entirely plausible that age zero will not be the modal age at death for the generations born in the last few decades.

On a purely practical note, mortality rates at age zero must typically be excluded from smoothing models intended to separate noise from underlying patterns. In the P-spline models used in Richards, Kirkby and Currie (2006), if mortality at age zero is not excluded then it leads to under-smoothing.

Murphy mentions that drawing inferences from death counts is vulnerable to the effects of migration. This is correct, and was stated in the paper. It is worth particular note that migration varies by age group, with international migrants over age 45 comprising between 6.4% and 9.1% of the total international migrants between 1991 and 2007 (ONS data). However, over the same period the proportion of the England & Wales population over age 45 has varied between 36.5% and 40.2% (own calculations from ONS population exposures), a reflection of the fact that older people are less likely to be migrants than younger people. This was another reason for choosing to only use death counts at age 60 and over, as this is the adult age group least likely to be affected by migration.

Murphy asked for more detail in Table 3, in particular the AIC (Akaike, 1987). This information is given below, which shows that model 4 is a materially better fit than the others, regardless of whether one judges this by the X^2 statistic or the AIC.

Extended version of Table 3. Results of various models with time-based and cohort-based parameters.

Model	Description	X^2	X^2 relative to model 1	AIC	AIC relative to model 1
1	α, β constant, $\delta = 0$	22,112.7	n/a	69,209,428	n/a
2	α, β constant, $\delta = -0.01468$	4,456.0	-17,656.7	68,907,414	-302,014
3	α and β varying by cohort, $\delta = 0$	2,157.6	-19,955.1	68,878,129	-331,299
4	α and β varying by cohort, $\delta = -0.00791$	1,453.8	-20,658.9	68,872,379	-337,049

To illustrate the age-period-cohort identity, Murphy rewrites the equation:

$$\mu_{x,c,t} = \exp(\alpha + \beta x + \delta t) \quad (1)$$

as follows:

$$\mu_{x,c,t} = \exp(\alpha + \beta c + (\beta + \delta)x) \quad (1*)$$

but I think he actually means:

$$\mu_{x,c,t} = \exp(\alpha + \delta c + (\beta + \delta)x) \quad (1\dagger)$$

Murphy writes that “it is not logically possible to isolate such a linear effect since there is an identity between age, period and cohort”. This is correct, but it is irrelevant here for two reasons. First, none of the models in Appendix C are structured as per Equation (1). Specifically, where birth cohort appears in Models 3 and 4 it is as a discrete-level factor, not as a linear covariate as per Equation (1). Secondly, the models in Appendix C are primarily concerned with the overall fit, not the specific parameter values. Even if the identity problems in Equations (1) and (1†) did apply, the fitted values — and thus the X^2 statistic and the AIC — are invariant, as Murphy himself notes. Thus, conclusions based on comparing the overall model fits are valid. The subsequent remark “both points (i) and (ii) are incorrect” refers to a straw man of Murphy’s own creation beginning with the line “this appears to suggest ...”, so further comment from me is unnecessary.

Murphy takes issue with my claim to have “formally demonstrated the dominance of cohort effects over period effects among males in England and Wales over the past 40 years” in an earlier paper (Richards et al., 2006) *using similar modelling approaches as those discussed above*. The emphasis is mine because Murphy’s additional comment is incorrect: the models in Richards, Kirkby and Currie (2006) have a completely different structure to Equations (1) and (1†). As to the question of dominance itself, consider Figure A, which shows the mortality improvements for males in England and Wales by year of birth and year of observation. The definition of mortality improvement is that of Willets (1999), namely $1 - m_{x,t}/m_{x,t-1}$, where $m_{x,t}$ is the smoothed mortality rate at age x in year t . We have used both an age-period smoothing penalty (left panel) and an age-cohort penalty (right panel) to ensure the improvement patterns are not artefacts of the smoothing process. Period effects appear as horizontal patterns, whereas cohort effects

manifest themselves as vertical ones. Visually, it is clear that cohort patterns dominate under both smoothing approaches. Statistically we have the same result: the BIC for the age-cohort model is 100 units smaller than the age-period model, suggesting that cohort patterns do indeed dominate. Of course, this is not to say that there is no time-trend component at all, merely that cohort effects have contributed more strongly to mortality improvements in England and Wales.

Figure A about here. The title should be as follows:

“Smoothed male mortality improvements in England and Wales using penalised splines. The left panel shows the results of a model based on age and year of observation (age-period). The right panel shows the results of a better-fitting alternative model based on age and year of birth (age-cohort). A cohort-period model was also considered, but gave a much poorer fit due to the strength of the age signal. Source: Richards, Kirkby and Currie (2006).”

Are all attempts to separate age, period and cohort effects a “futile quest”? They may have seemed so when Glenn (1976) wrote this, but things do not seem so categorical now. It is certainly tricky, and Clayton and Schifflers (1987a and 1987b) provide a careful discussion of the perils of interpreting estimated cohorts effects. The Age-Period-Cohort (APC) model requires (arbitrary) identifiability constraints to fit, but Currie *et al* (2009) show how to smooth data with respect to the APC model (or any other non-identifiable model) by using penalties that are invariant with respect to the constraints used to define the model. As shown by Equations (1) and (1†), the age (or period or cohort effects) within the APC model are not themselves uniquely estimable. However, what *is* estimable are the fitted values for the whole table and smoothing can be done with respect to these invariants. The two-dimensional models used in Richards, Kirkby and Currie (2006) are in a different category entirely, since these are identifiable models. The two-dimensional improvements shown in Figure A are invariants.

On a technical point, Murphy questions whether the Poisson assumption is correct. Although the data naturally lend themselves to this, there is nevertheless considerable evidence of over-dispersion, as detailed in the latter part of Section 5 of the paper. This over-dispersion violates the Poisson assumption of the variance of death counts being equal to the mean value. Using a similar model to Biatat and Currie (2009), we can approximately control for this over-dispersion, but this leaves the mean values little changed and the non-random residual pattern in Figure 7 in any case remains, albeit with less-extreme absolute values for the residuals.

Murphy appears confused about my close look at the 1919/1920 birth cohort. As he himself suggests, I would indeed expect higher mortality for this cohort due to the combination of influenza exposure and post-war privation, so a large positive residual could have been expected in Figure 12. The fact that there is a huge *negative* residual for this generation was surprising and therefore potentially suspect. My investigations of the data suggest that this counter-intuitive result may well be due to incorrect population estimates. Table 1 of my paper suggests this as plausible, while Figures 13 and 14 show just how challenging it is for official statisticians to make population estimates for the generations born in 1919 and 1920. This is particularly the case when they only have decennial census data to work with and are forced to infer exposures for intervening years. Other authors have raised questions over these exposure estimates, including Cairns *et al* (2009). Faced with this evidence of mundane data-quality issues, I prefer this likelier explanation of the residual to one for which I know of no biological basis. Murphy says that “the alternative would be to accept the effect is real”, so I would be most interested to hear any theory as to how a cohort with “two apparently considerable disadvantages” could have beneficial mortality.

AKAIKE, H. (1987). Factor analysis and AIC, *Psychometrika*, **52**, 317–333.

BIATAT, V. A. D. AND CURRIE, I. D. (2009) On the problem of duplicates, *Working paper available on request*.

CAIRNS, A. J. G., BLAKE, D., DOWD, K., COUGHLAN, G. D., EPSTEIN, D., ONG, A. BALEVICH, I. (2009) A quantitative comparison of stochastic mortality models using data from England & Wales and the United States, *North American Actuarial Journal*, **13** No. 1.

- CHARLTON, J. AND MURPHY, M. J. (EDITORS) (1997) *The Health of Adult Britain 1841-1994 Decennial Supplement, Volume 1*, ISBN 9780116916952.
- CLAYTON, D. AND SCHIFFLERS, E. (1987a) Models for temporal variation in cancer rates I: Age-period and age-cohort models. *Statistics in Medicine* **6**, 449–467.
- CLAYTON, D. AND SCHIFFLERS, E. (1987b) Models for temporal variation in cancer rates II: Age-period-cohort models. *Statistics in Medicine* **6**, 469–481.
- COCEVAR, P. (2007) An analysis of recent mortality trends in the Italian population using penalised B-spline regression, *Giornale dell'Istituto Italiano degli Attuari* **70** pp. 21–43.
- CURRIE, I. D., KIRKBY, J. G, DURBAN, M. AND EILERS, P. H (2009) Smoothing Age-Period-Cohort models with P-splines: a mixed-model approach, *Working paper available on request*.
- DAVENPORT, F. M., MINUSE, E., HENNESSY, A. V. AND FRANCIS JR, T. (1969) Interpretations of influenza antibody patterns of Man, *Bulletin of the World Health Organisation* 1969 **41** pp. 453–460.
- DOBLHAMMER, G. (2004) The late life legacy of very early life, *Demographic Research Monographs*, Heidelberg: Springer, Chapter 6 “Cohort and Age Effects”.
- DOLL, R., DARBY, S. AND WHITLEY, E. (1997) *The Health of Adult Britain 1841-1994 Decennial Supplement, Volume 1*, ISBN 9780116916952. Edited by Charlton, J. and Murphy, M. J.
- LAWLOR, D. A., EBRAHIM, S. AND DAVEY SMITH, G. (2001) Sex matters: secular and geographical trends in sex differences in coronary heart disease mortality, *BMJ* 2001 **323** pp. 541–545.
- MURPHY, M. (2009) Some comments on a paper by S. J. Richards, Detecting year-of-birth mortality patterns with limited data, *Journal of the Royal Statistical Society A*, **171**, Part 1, pp. 1–20.
- PENSIONS REGULATOR (2008) Good practice when choosing assumptions for defined benefit pension schemes with a special focus on mortality, Consultation document, February 2008.
- RICHARDS, S. J., KIRKBY, J. G. AND CURRIE, I. D. (2006) The Importance of Year of Birth in Two-Dimensional Mortality Data, *British Actuarial Journal*, **12** (I) pp. 5–61.
- RICHARDS, S. J. (2008) Detecting year-of-birth mortality patterns with limited data, *Journal of the Royal Statistical Society A*, **171**, Part 1, pp. 1–20.
- WILLETS, R. C. (1999) Mortality in the next Millennium, *Staple Inn Actuarial Society*, London.
- WILLETS, R. C. (2004) The Cohort Effect: Insights and Explanations, *British Actuarial Journal*, **10**, Part IV, No. 48.

Stephen Richards,
Edinburgh,
3rd February 2010.