Demography (2013) 50:1977–1979
DOI 10.1007/s13524-013-0255-8

COMMENTARY

Comment on "Assessing Validity and Application Scope of the Intrinsic Estimator Approach to the Age-Period-Cohort (APC) Problem"

Leonhard Held · Andrea Riebler

Published online: 17 October 2013

This article looks at a problem that has been studied over decades: the unidentifiability issue in the age-period-cohort (APC) model. The impossibility to identify the full set of age, period, and cohort effects in the analysis of mortality or other vital rates has created much confusion in the epidemiological and sociological literature. The misconception that the problem is a shortcoming of current statistical methods still seems to be prevalent, and apparently a new "solution" has been proposed recently: the intrinsic estimator (IE).

However, the difficulty is inescapable and arises from the fact that subjects cannot move in one time scale without an identical move in others (Clayton and Hills 1993:315). In order to develop some intuition for this phenomenon, it helps to visualize the simultaneous development of age, period, and cohort in a Lexis diagram. Similar graphical devices date back to sociological literature from the 1870s published in German; see Keiding (2011) for an interesting review. Alternatively, the identifiability problem in APC models can be described as follows. Suppose that we have estimates of age, period, and cohort effects, all of which can be centered to zero with appropriate adjustment of the global mean μ. Then the three sets of estimates can be rotated in a certain way, such that the linear predictor (Eq. (1) in the Luo article) remains the same for all observations (Knorr-Held and Rainer 2001:111, Eq. (3)). This phenomenon is illustrated in Fig. 1. Any pattern of age, period, and cohort effects can be transformed to an equivalent pattern, which will give exactly the same fit to the data. This applies also to estimates obtained via the IE. Therefore, Clayton and Schifflers (1987: figure 3) recommended displaying only the identifiable second differences parameter estimates, but those are difficult to interpret. A unified approach to validate and derive further estimable functions is given in O'Brien (2012).

The online version of the original article can be found at http://dx.doi.org/10.1007/s13524-013-0243-z.

L. Held (✉)
Institute of Social and Preventative Medicine, Division of Biostatistics, University of Zurich, Hirschengraben 84, 8001 Zurich, Switzerland
e-mail: leonhard.held@ifspm.uzh.ch

A. Riebler
Norwegian University of Science and Technology, Trondheim, Norway

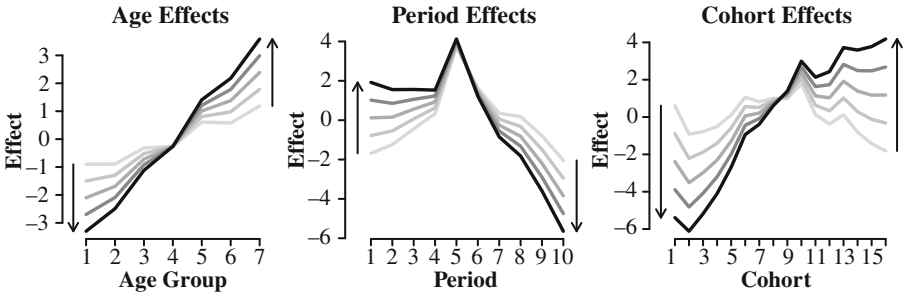

Fig. 1 Equivalent patterns of age, period, and cohort effects. Rotating the period effects in a certain direction with a corresponding rotation of age and cohort effects in the opposite direction does not affect the model fit

From Fig. 1, it is clear that the parameters in the APC model can be identified only if an additional constraint is imposed. The specific choice of this constraint is arbitrary, if it is not motivated from subject-specific considerations, and will obviously influence the resulting estimates. Luo identifies the IE as an approach that uses a specific linear constraint, as arbitrary as any other constraint. We are somewhat puzzled to see that the IE has received so much attention in the sociological literature even though, by its very nature, the identifiability problem cannot be solved. In addition, as Luo points out, the IE dates back to a suggestion by Kupper et al. (1983:2797), so it is not even new.

The preceding discussion makes clear that it does not make sense to see whether the "true" but unidentifiable APC effects are estimated "unbiasedly," "consistently," "efficiently," or similarly, and Luo rightly criticizes Yang et al. (2008) for doing that (see Luo's footnote 6). However, then Luo walks into the same statistical quicksand, investigating through simulation whether certain age, period, and cohort effects can be revealed under different constraints. She considers scenarios in which the IE performs worse than other constraints, but it does not make sense to perform such simulation studies in the first place, either in Yang et al. (2008) or in the Luo article. We are not convinced that such an approach is "particularly helpful for nontechnical researchers." It will just increase confusion.

We were surprised to see that the basic age-period-cohort model is still the main focus of the current discussion. The statistical and biostatistical literature has looked in recent years at several important extensions of this formulation, which seem to have relevance also in sociological and demographic applications. For example, Holford (2006) discussed extensions to the common case in which the interval widths are not the same for age and period, and additional identifiability problems arise. Riebler and Held (2010) developed APC methodology for the analysis of multiple outcomes. In contrast to the standard APC model, this extension allows to identify time trends under certain assumptions. Indeed, if at least one set of parameters—for example, the age effects—are assumed to be the same across outcomes, differences of the other time effects between outcomes are identifiable and interpretable as log relative risks. The estimated effects can be smoothed using either a Bayesian (Riebler and Held 2010) or a frequentist (Held and Riebler 2012) approach. Thus, the analysis of multiple outcomes circumvents the identifiability problem of the standard APC model, provides interpretable parameter estimates, and seems to deserve further attention in epidemiological and sociological research. For a recent application with particular focus on prediction, see Riebler et al. (2012).

References

Clayton, D., & Hills, M. (1993). *Statistical models in epidemiology.* Oxford, UK: Oxford University Press.

Clayton, D., & Schifflers, E. (1987). Models for temporal variation in cancer rates. II: Age-period-cohort models. *Statistics in Medicine, 6,* 469–481.

Held, L., & Riebler, A. (2012). A conditional approach for inference in multivariate age-period-cohort models. *Statistical Methods in Medical Research, 21,* 311–329.

Holford, T. R. (2006). Approaches to fitting age-period-cohort models with unequal intervals. *Statistics in Medicine, 25,* 977–993.

Keiding, N. (2011). Age-period-cohort analysis in the 1870s: Diagrams, stereograms, and the basic differential equation. *Canadian Journal of Statistics, 39,* 405–420.

Knorr-Held, L., & Rainer, E. (2001). Projections of lung cancer mortality in West Germany: A case study in Bayesian prediction. *Biostatistics, 2,* 109–129.

Kupper, L. L., Janis, J., Salama, I. A., Yoshizawa, C. N., Greenberg, B. G., & Winsborough, H. H. (1983). Age-period-cohort analysis: An illustration of the problems in assessing interaction in one observation per cell data. *Communications in Statistics—Theory and Methods, 12,* 2779–2807.

O'Brien, R. M. (2012). Estimable functions in age-period-cohort models: A unified approach. *Quality and Quantity.* Advance online publication. doi:10.1007/s11135-012-9780-6

Riebler, A., & Held, L. (2010). The analysis of heterogeneous time trends in multivariate age-period-cohort models. *Biostatistics, 11,* 57–69.

Riebler, A., Held, L., & Rue, H. (2012). Estimation and extrapolation of time trends in registry data—Borrowing strength from related populations. *Annals of Applied Statistics, 6,* 304–333.

Yang, Y., Schulhofer-Wohl, S., Fu, W. J., & Land, K. C. (2008). The intrinsic estimator for age-period-cohort analysis: What it is and how to use it. *American Journal of Sociology, 113,* 1697–1736.