Published in final edited form as:
Demography. 2013 December ; 50(6): 1969–1971. doi:10.1007/s13524-013-0254-9.

Misunderstandings, Mischaracterizations, and the Problematic Choice of a Specific Instance in Which the IE Should Never Be Applied

Yang Claire Yang and
Department of Sociology, Lineberger Cancer Center, and Carolina Population Center, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3210 yangy@unc.edu

Kenneth C. Land
Department of Sociology and Center for Population Health and Aging, Population Research Institute, Duke University, Durham, NC 27708-0088 kland@soc.duke.edu

It is our pleasure to respond to the Editor's request for a commentary on Luo's article (this issue), which critically evaluates the utility of the intrinsic estimator (IE) that we first introduced to demography and sociology in Yang et al. (2004).

We first respond to the Editor's request for "an assessment of the author's argument, evidence, and conclusions." It is truly unfortunate and fundamentally incorrect to interpret our stance on the IE method as seeing it a "holy grail" or "magic bullet" for the identification problem of the age-period-cohort (APC) accounting model/multiple classification model. Nowhere in our previous publications did we make such a claim. We were crystal clear about the circumstances in which a full-blown APC models should be used. And such circumstances equally apply to any estimator of full three-factor APC models, not just the IE. Luo completely lost sight of this starting point.

Works as early as Yang (2008) and as recent as Yang and Land (2013: chapter 5) have laid out a three-step procedure that should be thoroughly applied to APC analysis using the accounting model. It is so important that we believe it is worth repeating here. Step 1 is to conduct descriptive data analyses using graphics, with the objective being to provide qualitative understanding of patterns of temporal variations. Step 2 is model fitting and calculation of model fit statistics, such as the Bayesian information criterion (BIC). The objective is to ascertain whether the data are sufficiently well described by any single factor or two-factor model of age (A), time period (P), and cohort (C) effects for which there is no identification problem. Only when these analyses suggest that all three dimensions are operative should one proceed with Step 3: a three-factor APC model to which a constrained estimator can be applied to identify the A, P, and C effects. By revisiting Glenn's (2005) numerical example, Yang and Land (2013:109) emphasized that "imposition of a full APC model on data when a reduced model fits the data equally well or better constitutes a model misspecification and should be avoided." Empirical examples of chronic disease mortality in Yang (2008) and cancer mortality in Yang and Land (2013) showed the necessity of all three steps, whereas those of cancer incidence for certain sites in the latter show that the first two steps suffice. A blind application of the IE, or any other constrained estimator, of the full three-factor APC model was never recommended.

It follows that any of Luo's exercises based on scenarios in which no full APC models should be applied are invalid. Because space limitations imposed on this comment do not

permit a more detailed analysis of Luo's article , we posted a full response elsewhere (http://www.unc.edu/~yangy819/apc/index.html) that explains the fundamental flaws. Briefly, Luo claims that (1) there is nothing new about the IE for the age, period, and cohort effect coefficients of the classical APC accounting model; (2) the IE is not an unbiased estimator of the unidentified coefficient vector of this model; (3) the constraint imposed by the IE on the unidentified coefficient is "implicit"; and (4) the IE performs poorly as a statistical estimator of the unidentified coefficient vector when that vector has very large effects of the design matrix. In our more detailed analysis, we respond that point 1 represents a misunderstanding of the IE, point 2 is a claim that we never made, point 3 is a mischaracterization of our work, and point 4 disregards the asymptotic properties of the IE. In short, there is little merit in Luo's article other than an algebraic demonstration of a situation—identical linear or nonlinear algebraic trends in the effect coefficients for all three temporal dimensions—in which the IE should never be applied.[1]

We next respond to the Editor's suggestion to comment on the implications of the article for previous research that has used the IE approach. Given the aforementioned fallacies in Luo's article, it should have no bearing on previous research that has used the IE approach. More importantly, since the publications of our earlier works on the IE, research has found increasing support for the utility of the IE method when the full APC accounting model can be properly applied. Additional methodological studies of its properties have continued. For example, Powers (2013:1039) couched the IE in a longer statistical tradition of estimable functions, showing "a number of equivalent numerical methods (i.e., principal components, singular value decomposition, generalized inverse, and so on) may be employed in various ways to provide solutions." And Powers' (2012) work provides a flexible extension (*ie_rate*) of Schulhofer-Wohl and Yang's (2006) APC IE Stata program (*apc_ie*). In addition, model validation studies have further demonstrated the robustness of the statistical properties of the IE through comparisons of results from an IE analysis of empirical data with results from an analysis of the same data by application of a different family of models that do not use the same identifying constraint. There is strong evidence that validates the IE method across multiple data sets and outcomes. For example, Masters et al. (2013) found estimates of temporal trends in official U.S. adult mortality rates from models using the IE to be entirely consistent with coefficient estimates from Bayesian hierarchical models using Gibbs sampling (Yang 2006). Other examples of consistent A, P, and C patterns estimated by the IE as well as alternative models such as mixed-effects models include verbal test scores from the General Social Survey (Yang et al. 2008), adult mortality from vital statistics (Yang 2008) and National Health and Interview Surveys (Masters 2012; Masters et al. 2012), and cancer incidence and mortality from national tumor registries (Yang and Land 2013).

Finally, the Editor asked that we share advice about the options for future researchers who seek to identify separate age, period, and cohort effects. Our advice is that researchers should never naively apply any estimator to APC data and expect to obtain meaningful results. In all cases, APC analysis should be approached with great caution and awareness of its many pitfalls. In this context, the IE has been shown to be a useful approach to the identification and estimation of the three-factor APC accounting model. The IE, however, is not a "final" or "universal" solution to the identification problem of linear or APC

[1]This statement is based on mathematical analysis. In addition, our statistical analyses of the three models used in the simulations reported in Luo's (2013) Table 3 show that the data are well described by two-factor models with period and cohort effects (data sets 1 and 2) and age and cohort effects (data set 3), respectively. In other words, as expected on the basis of mathematical analysis, the two-factor models show superior model fits to the data generated by the simulations to full, three-factor APC models, and therefore estimation of a three-factor model is completely erroneous and specious.

accounting models. There will never be such a solution within the confines of conventional linear models that necessarily beget the identification problem.

References

Glenn, ND. Cohort analysis. 2nd ed.. Sage Publications; Thousand Oaks, CA: 2005.

Masters RK. Uncrossing the U.S. black-white mortality crossover: The role of cohort forces in life course mortality risk. Demography. 2012; 49:773–796. [PubMed: 22729715]

Masters, RK.; Hummer, RA.; Lin, S.; Beck, AN.; Finch, BK. Long-term trends in adult mortality for U.S. blacks and whites: An examination of period- and cohort-based changes.. 2013. Manuscript under review

Masters RK, Hummer RA, Powers DA. Educational differences in U.S. adult mortality risk: A cohort perspective. American Sociological Review. 2012; 77:548–572.

Powers, DA. IE_RATE: Stata module to conduct age, period, and cohort (APC) analysis of tabular rate data using the intrinsic estimator. 2012. Retrieved from http://ideas.repec.org/c/boc/bocode/s457445.html

Powers DA. Black-white differences in maternal age, maternal birth cohort, and period effects on infant mortality in the U.S. (1983–2002). Social Science Research. 2013; 42:1033–1045. [PubMed: 23721672]

Schulhofer-Wohl, S.; Yang, Y. APC: Stata module for estimating age-period-cohort effects. 2006. Retrieved from http://econpapers.repec.org/software/bocbocode/S456754.htm

Yang Y. Bayesian inference for hierarchical age-period-cohort models of repeated cross-section survey data. Sociological Methodology. 2006; 36:39–74.

Yang Y. Trends in U.S. adult chronic disease mortality: Age, period, and cohort variations. Demography. 2008; 45:387–416. [PubMed: 18613487]

Yang Y, Fu WJ, Land KC. A methodological comparison of age-period-cohort models: The intrinsic estimator and conventional generalized linear models. Sociological Methodology. 2004; 34:75–110.

Yang, Y.; Land, KC. Age-period-cohort analysis: New models, methods, and empirical applications. Chapmen & Hall/CRC Press; New York: 2013.

Yang Y, Schulhofer-Wohl S, Fu WJ, Land KC. The intrinsic estimator for age-period-cohort analysis: What it is and how to use it? American Journal of Sociology. 2008; 113:1697–1736.