# Age-Period-Cohort Analyses in Epidemiological Studies: Clarifying Assumptions and Validating Results

Ryan K. Masters, University of Colorado Boulder
Daniel A. Powers, University of Texas at Austin

| | |
|---|---|
| Abstract: | 220 |
| Text: | 3240 |
| Figures: | 3 |
| Tables: | 0 |

We thank the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD)-funded University of Colorado Population Center (Award Number P2C HD066613) and University of Texas Population Research Center (Award Number P2CHD042849) for development, administrative, and computing support. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NICHD or the National Institutes.

**Abstract**

Background: Age-Period-Cohort (APC) models are often used to decompose health trends into period- and cohort-based sources, but their use in epidemiology remains contentious. Central to the contention are researchers' failures to 1) clearly state their analytic assumptions and/or 2) thoroughly evaluate model results. These failures generate confusion about the merits of APC methods, and different APC approaches are often treated and discussed as one and the same. Consequently, scholarly exchanges about APC methods usually result in strong disagreements that rarely offer practical advice to users or readers of APC methods.

Methods: We propose three simple guidelines to help practitioners of APC methods articulate their assumptions and validate their results. To demonstrate the usefulness of the guidelines, we apply them to results recently published in *American Journal of Epidemiology* about black-white differences in U.S. heart disease mortality.

Results: The application of the guidelines reveals two important findings. On the one hand, some APC methods produce inconsistent results that are highly sensitive to researcher manipulation. On the other hand, other APC methods estimate results that are robust to researcher manipulation and consistent across APC models.

Conclusions: The exercise shows the simplicity and effectiveness of the guidelines in resolving disagreements over APC results. The cautious use of APC models can generate results that are consistent across methods and robust to researcher manipulation.

Keywords: Age-Period-Cohort, trends, heart disease, mortality

A recent article in the *American Journal of Epidemiology* (AJE) (1) featured cohort analyses to examine trends in black-white differences in U.S. heart disease mortality. Accompanying the article was a comment (2) and a reply by two of the article's authors (3) that discussed age-period-cohort (APC) methods in epidemiologic studies. The back-and-forth was unsatisfactory, as the exchange provided little in the way of practical advice for users of APC methods. On one side was a group of APC practitioners who enthusiastically supported the use of a readily available "APC toolbox" (3, p. 1) and on the other side was a skeptic warning of severe "potential pitfalls" in APC analyses (2, p. 1). Harper (2) rightly emphasized that the utility of APC models rests on the plausibility of one's assumptions, but casted doubt on the use of all APC methods. Concluding that APC models have estimated "wildly differing conclusions regarding the influence of period and cohort effects" on U.S. heart disease mortality, Harper (2, p. 2) warned readers, "it seems likely that researchers could end up choosing APC models that are most consistent with their favorite hypotheses." In response to Harper's critical assessment of APC methods and his dire warnings about their use, Kramer and Casper (3, p. 1) believed the exchange produced "more areas of agreement than disagreement," and left with their results largely intact. They concluded, "the APC toolbox accomplished the task for which it is suited."

Readers of the exchange were left at an impasse, with neither a way to assess the exchange nor a clear way forward to use APC models or interpret their results. Should readers distrust all results from APC analyses or does an "APC toolbox" exist for their easy use? We suggest it's neither of these alternatives and recommend three guidelines for the cautious use of APC methods in social scientific research. We revisit Kramer et al.'s (1) analyses and apply these guidelines to provide evidence against the notion that an "APC toolbox" exists for easy

analysis of APC trends in health outcomes. Some APC methods are preferred to others and researchers need to adjudicate between them when fitting APC models. Specifically, some APC methods estimate results that are *internally consistent* (i.e., results are invariable to the choice of model constraint) as well as *consistent with other methods* (i.e., results estimated from different APC methods are statistically indistinguishable from each other). Conversely, other APC methods estimate results that are neither internally consistent nor consistent with other methods' estimates. However, the evidence also rebukes Harper's contention that APC models produce "wildly differing conclusions" (2: p. 1) and that researchers can manipulate APC models to be "most consistent with their favorite hypotheses" (2, p. 2). Conclusions differ across APC studies largely because researchers misapply APC methods and fail to validate results, not from the methods themselves. What follows is a simple test and application of APC guidelines to a) help clarify applications of APC methods in epidemiologic studies and b) use sensitivity tests to validate results.

**Guidelines for APC Analyses**

Cohort analyses have a long history in epidemiology (4), but doubts remain about using APC statistical methods (5-19). The first cohort analyses plotted age-specific rates by period to descriptively show variation in trends. These practices continue to serve as first steps for possibly identifying "non-parallelism" in period-based trends in outcomes (12, p. 815; 20, 21), which can indicate whether or not cohort-based variation might exist in the outcome. Thus, they help researchers decide if they ought to employ a full "three-dimensional" (i.e., APC)

statistical model over a simpler age-period (AP) or age-cohort model (AC) (1, 18, 20). These descriptive plots do not help determine what statistical approach is preferred for fitting an APC model (12, 20). This is an important point of clarification that challenges Kramer et al.'s (3, p. 1) notion of an "APC toolbox." Although graphical plots might suggest an APC model would be useful for detecting age, period, and cohort variation in an outcome, researchers must choose from a number of statistically-based APC methods to help investigate their research questions (20). The following guidelines apply to this choice.

*Simplify models and explicitly state assumptions.*

A central concern with fitting statistical APC models is the constraint imposed by a given method (8-12, 20). A constraint is necessary in order to identify *a* solution from the *infinite* number of solutions that exist as a result of the linear dependency between age, period, and cohort (e.g., cohort=period-age). Clearly stating the assumptions behind a statistical method's constraint is necessary as some constraints are more appropriate than others for a given data structure. Further, model estimates from some APC methods can vary considerably depending on the choice of constraint (9, 13-15, 20). Finally, some constraints are imposed directly by researchers whereas other methods impose constraints that are largely uninfluenced by researchers' decisions.

Kramer et al. (1) specified an APC model using dummy coding with equality constraints, which requires strong theory to assume equality between APC referent categories and the constrained period values. In models with highly collinear predictors, dummy-variable designs

like this privilege a particular solution in the solution space (12). It is for this reason that early developers of some APC methods used centered effects coding (i.e., sum to zero constraints over each set of APC parameters ) and constrained parameter invariance to small regions in the parameter space (22-24). The further use of an equality constraint on parameters requires even stronger theory or *a priori* "side information" to justify specifying parameter values to be equal to each other (9, p. 22; 11, 12, 16, 20). Kramer et al. (1), for instance, constrained the first two periods in their model to be equal to each other, and used the seventh age and sixth cohort as referent categories: $\beta_7 = \gamma_1 = \gamma_2 = \delta_6 = 0$ (8, p. 242). Kramer et al. (1) applied the constraint in a "somewhat unorthodox" way by fitting an APC model that constrained the period parameters associated with the rate *ratio* between black and white mortality, rather than fitting APC models separately to the mortality rates (3, p. 1). The approach unnecessarily complicates the constraint assumptions and is incredibly difficult to theoretically justify or to empirically evaluate. The constraint is both difficult for readers to follow and also increases the likelihood that model estimates of APC parameters are biased because the imposed $\gamma_1 = \gamma_2$ constraint must hold simultaneously in separate populations. This assumption receives little support in Kramer et al.'s (1) own Figure 1 and Figure 2, which show, respectively, differences between black and white men's and women's period-based trends in age-standardized mortality rates and rate ratios. The heart disease mortality trends during the 1970s – the time period constrained to be equal in Kramer et al.'s (1) model, 1973-1977=1978-1982 – appear to differ considerably for black and white men and women, with faster declines observed in the white population than in the black population.

To make constraint assumptions explicit and simple for readers to follow, we recommend:

a. Researchers fit APC models on rates themselves, not rate ratios.
b. Researchers fit APC models separately by subpopulations, not on pooled data using interaction terms.
c. Researchers use APC methods that minimize researcher involvement in the application of the model constraint.

Regarding point c., Kramer et al.'s (1) use of the "Mason et al. method" requires a researcher to set equal the variation between two ages, two periods, or two cohorts in order to identify the APC model (e.g., $\gamma_1 = \gamma_2$). This method therefore requires strong theory to inform the researcher about which two parameters should be constrained equal to each other in order to identify the APC model. Several alternative APC identification strategies exist that minimize a researcher's involvement in the constraint. For example, the intrinsic estimator (IE) applies a Moore-Penrose (MP) generalized inverse to the singular design matrix in tabular APC data. This approach builds off early work (22) that parameterizes the APC model using centered effects coding and yields a solution with smaller variance than other constrained approaches (20, 23-27). Alternative APC methods employ a maximum entropy estimator (MEE) to a bounded range of the response variable to estimate APC parameters that are set-identified (28). Fu (2016) shows the consistency of MP estimators such as the IE as the number of cells in the age x period table increase (24). Other work on MP estimators for APC models have discussed desirable shared properties of MP estimators as well as how APC estimates can diverge depending on design matrices (29). Additionally, Hierarchical-Age-Period-Cohort Cross-Classified Radom Effects Models (HAPC-CCREM) treat age as fixed effects and nest individuals in

periods and cohorts to estimate coefficients on a multiplicative scale (20, 30, 31). Although researchers can influence model estimates by changing the number and size of the APC groupings (e.g., three-year vs. five-year groupings), the constraints used in hierarchical APC methods and those in MP estimators remain largely outside the researcher's control. Constraints are not made by researchers arbitrarily setting values equal to one another. Instead, for example, the constraints in MP estimators depend on the structure of the data and the minimum-norm constraint applied in the solution space. Finally, although it is known that all APC models yield biased estimates of APC parameters unless the model's constraint exists in the actual population, it has been shown that some methods are better than others at minimizing this bias (12, 20, 22, 24, 32, 33). Overall, researchers using APC methods should make clear the constraint they use, apply it in the simplest way possible, and favor the use of methods that minimize researcher influence on the choice and application of the constraint.

*Test within-method consistency of APC estimates.*

Researchers should evaluate the sensitivity of an APC model's estimates by changing the model's referent categories, constraining different model parameters, and altering model specifics (9, 12, 20). To demonstrate the utility of validating one's results, we revisit Kramer et al.'s (1) use of constrained generalized linear models (CGLIM) and their Table 1 data. We refit APC variation in black and white men's heart disease mortality rates by first using Kramer et al.'s (1) equality constraint on the first two period parameters. We then fit the model with an alternative constraint that set equal the last two period parameters. The differences between

APC coefficients from Kramer et al.'s model and the alternative model are shown in the right-hand panels of Figure 1 and Figure 2. In contrast, also included in Figure 1 and Figure 2 (left-hand panels) are APC coefficients from two models fitted using the IE, with one model fitted using the first APC categories as referents and the other model fitted using the last APC categories as referents (25).

[Figure 1 about here]

[Figure 2 about here]

APC coefficients estimated from the CGLIM model fitted by using Kramer et al.'s (1) constraint are substantively different from the APC coefficients estimated by the CGLIM model that constrained the last two period parameters. The variation in APC estimates is seen among both black and white men's heart disease mortality. Consistent with others' warnings, APC coefficients estimated from CGLIM models are highly sensitive to a researcher's choice of constraint (9, 12, 20). Conversely, the sets of APC coefficients estimated from the two IE models are very much consistent with each other in both the black and white men's populations. Thus, contrary to Harper's (2) concern that researchers can tinker with APC methods to find results that support their favorite hypotheses, APC estimates from the IE models are largely invariant to the choice of referent categories (25).

*Test between-method consistency of APC estimates.*

A central worry raised by Harper (2, p. 1) is the "flexibility of APC Models" and the ease with which researchers can use computational programs to fit the models. Taken together, these concerns imply that researchers might be able to modify APC models to produce results that are "most consistent with their favorite hypotheses" (2, p. 2). This claim can be directly tested, however, and to some extent was tested in point 2 above. In addition to testing the within-method consistency of APC estimates by varying referents (in the case of the IE method) or changing the parameter constraints (in the case of CGLIMs), researchers should validate their APC estimates by comparing them to estimates from alternative APC methods. That is, directly test the variation in APC model estimates across methods.

[Figure 3 about here]

In their online supplement, Kramer et al. (1) tested between-method consistency of APC estimates by plotting rate ratios from three different methods: the CGLIM method used in their paper, the IE, and the Median Polish model (34, 35). Kramer et al.'s between-method comparison makes it difficult to assess the consistency of the models' results because they compare between-population estimates within a method (e.g., they plot the IE method's estimated APC rate ratios for white men, black men, white women, and black women together in one graph). A more useful comparison of between-method results would be to contrast within-population estimates from different methods (e.g., plot APC rate ratios of black women's heart disease mortality estimated from both the IE method and the CGLIM method on the same graph). Because our concern is with between-method variation in the APC estimates,

researchers should make the comparison of alternative methods' estimates as easy to gauge as possible.

Figure 3 plots APC coefficients for white men's heart disease mortality rates estimated from four different APC models: (a) Kramer et al.'s use of CGLIM models that constrain the first two periods, $\gamma_1 = \gamma_2$ (1), (b) the average of two models using the intrinsic estimator (IE), one using the first categories of APC as referents and the second using the last categories of APC as referents (25), (c) the maximum entropy estimator (MEE) (28), and (d) Hierarchical Age-Period-Cohort Cross-Classified Random Effects Models (HAPC-CCREM) (20, 30, 31) (an online supplement briefly describes the analytic approaches of these models and also provides the Stata and R codes used to fit the models). Results show that the age, period, and cohort coefficients estimated from the IE, MEE, and HAPC-CCREM models are consistent with one another. In all comparisons, we see age-based increases in heart disease mortality and strong declines in heart disease mortality risk across both periods and cohorts, albeit with slowing declines among recent cohorts. These findings are consistent with existing evidence suggesting cohort-based stalling in U.S. heart disease mortality declines, and strong period-based declines across this time (32, 36, 37). In Figure 3 we also see that the CGLIM model yields APC coefficients that are inconsistent with these patterns. The age patterns are more attenuated than the age-based patterns suggested by the three other models, period-based variation is estimated to be virtually non-existent, and cohort-based declines are estimated to be much more pronounced. Overall, the results show that three different APC modeling strategies (IE, MEE, HAPC-CCREM) estimate very similar APC variation in white men's heart disease mortality. Conversely, the constrained CGLIM model estimates APC patterns that seriously conflict with

the others. This evidence counters Harper's (2, p. 1) worry that different APC methods can produce "wildly different conclusions," and further undermines his notion that researchers can strong-arm APC models into generating results that are "most consistent with their favorite hypotheses" (2, p. 2). The evidence also supports our earlier suggestion that researchers favor APC methods that minimize their involvement in the model constraint. We see that the three APC methods that minimize researcher involvement in the application of the constraint – the IE, the MEE, and HAPC-CCREM – estimated APC coefficients that are very consistent with one another. Conversely, Kramer et al.'s (1) identification strategy of forcing equal variation in heart disease mortality during the 1973-1978 and 1978-1983 periods estimated very different results that are sensitive to choice of the equality constraint.

**Discussion**

Researchers aiming to document period and/or cohort trends in health outcomes will likely continue to use APC models. Some researchers will strongly caution against the use of APC models, while others will perhaps defend and even advocate their use. We urge readers to consider that the best practice for cohort analyses is to cautiously adjudicate between different APC approaches, as some techniques are more advantaged than others for identifying cohort-based variation in health outcomes. Although Kramer and colleagues (1, 3) and Harper (2) both acknowledged that APC models are at best descriptive tools, neither recognized the important differences that exist across APC approaches for estimating APC variation in outcomes. On the one hand, Kramer and colleagues (3, p. 1) advocated the use of a general "APC toolbox." On the

other hand, Harper (2, p.2) suggested "that APC models should be used with caution," but made little distinction between the vast APC approaches in his warning. Given the difficulties and "potential pitfalls" in APC analyses, researchers need to recognize that not all APC approaches are the same – they constitute neither a "toolbox" for easy application nor a black box susceptible to manipulation. In this paper, we offered three easy-to-implement steps for the cautious use of APC models, which build upon others' suggestions (e.g., 12, 20, 21). To recap, we encourage APC researchers to first use graphs as exploratory exercises. Plot age-specific rates across time periods in order to identify possible "non-parallelism" in the trends (12, 20, 21). If the trends appear to be parallel, an APC statistical model fitted to the data will likely estimate biased results (18, 20, 37). Also, use GOF statistics to help assess if APC models likely account for more variation in the outcome than do simpler AP or AC models. An APC model fitted to data where a simpler model is preferred will estimate biased results (18, 20, 21, 32, 33, 38). Next,

1. Simplify models and clearly state your assumptions.
    a. Fit APC models on rates themselves, not rate ratios.
    b. Fit APC models separately by subpopulations, not on pooled data using interaction terms.
    c. Favor APC methods that minimize researcher involvement in the application of the model constraint.
2. Test within-method consistency (i.e., specify the model differently and compare the new results to those from the original model)

3. Test between-method consistency (i.e., compare model results to those estimated by models identified from different constraints)

Using these guidelines to assess the validity of Kramer and colleagues' (3, p. 1) "APC toolbox," we find (a) it took a "somewhat unorthodox approach" to constraining the model, making the identification strategy difficult to follow, (b) it required the researchers to impose an arbitrary equality constraint on the model parameters ($\gamma_1 = \gamma_2$), (c) when the model was fitted with a different constraint, the estimates differed substantially from those reported in the paper, and (d) the model estimated results that were inconsistent with several alternative APC methods. These shortcomings are strong evidence against the existence of an "APC toolbox." However, by applying our guidelines to Kramer et al.'s (1) data, we also found little support for Harper's (2, p.2) contentions that APC models arrive at "wildly differing conclusions regarding the influence of period and cohort effects" or that "researchers could end up choosing APC models that are most consistent with their favorite hypotheses." On the contrary, when we used APC constraints that are beyond researcher manipulation, results were consistent across the models and estimates from the IE were insensitive to alternative referent categories. With these guidelines, we can likely move conversations about APC methods beyond sweeping criticisms and/or support for generic "APC toolboxes" toward more practical discussions about how researchers using APC methods should articulate assumptions, specify multiple models, and scrutinize their results.

Figure 1. Age, Period, and Cohort Variation in U.S. Black Men's Heart Disease Related Mortality Rates, Ages 35-85 between Years 1973-2010. Variations in Left Column Estimated from IE, Variations in Right Column Estimated from CGLIM.

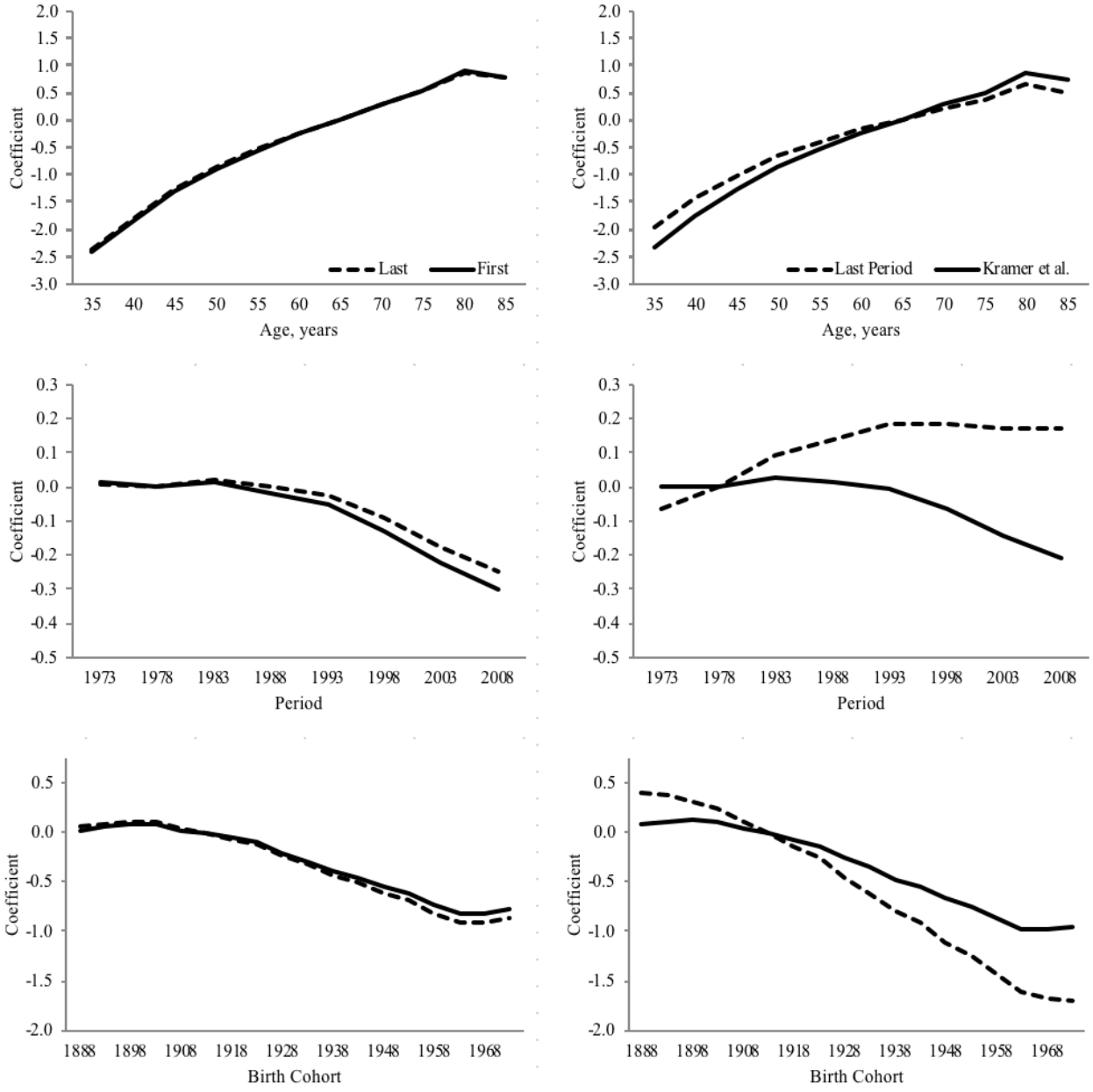


Figure 2. Age, Period, and Cohort Variation in U.S. White Men's Heart Disease Related Mortality Rates, Ages 35-85 between Years 1973-2010. Variations in Left Column Estimated from IE, Variations in Right Column Estimated from CGLIM.

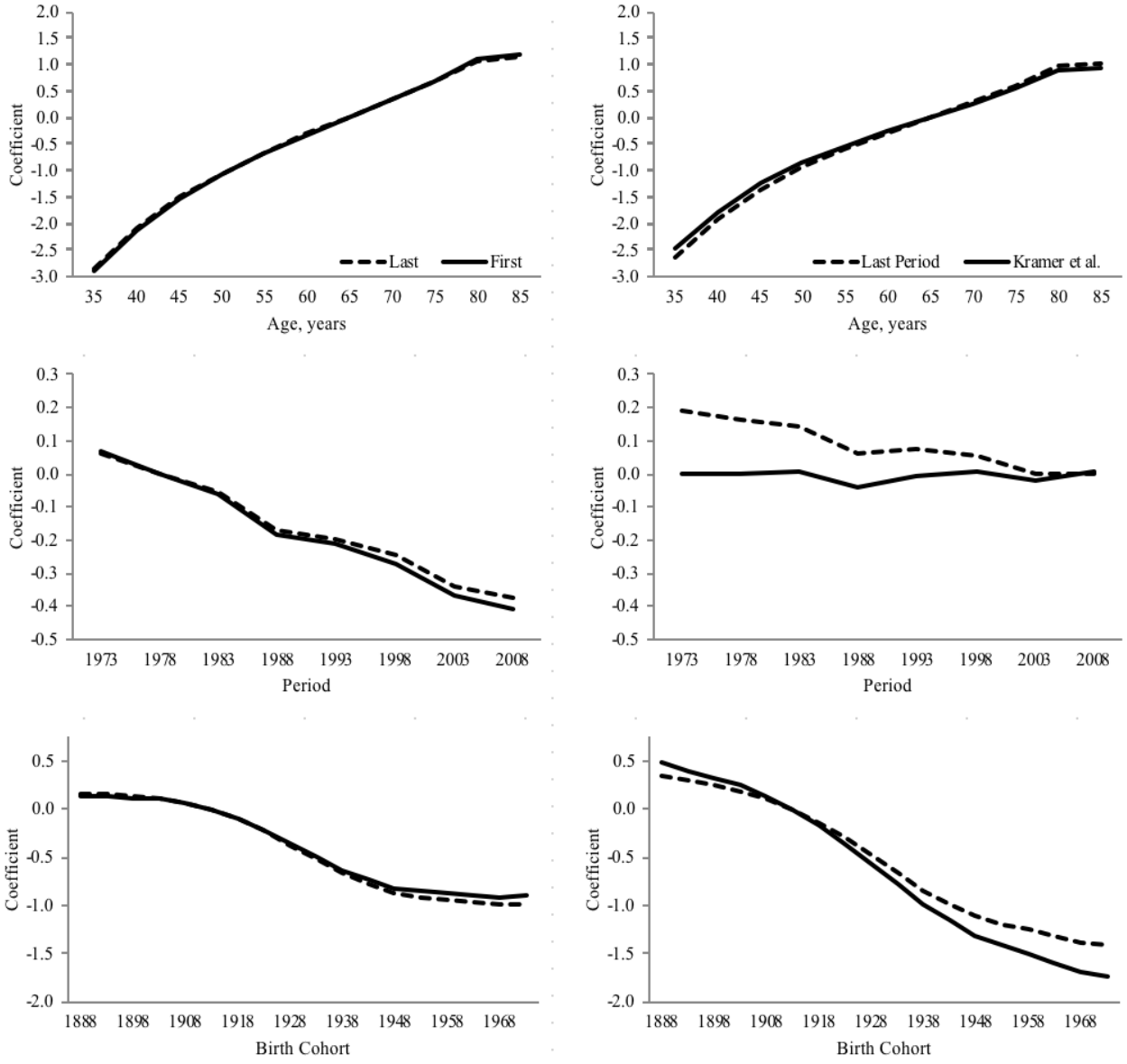
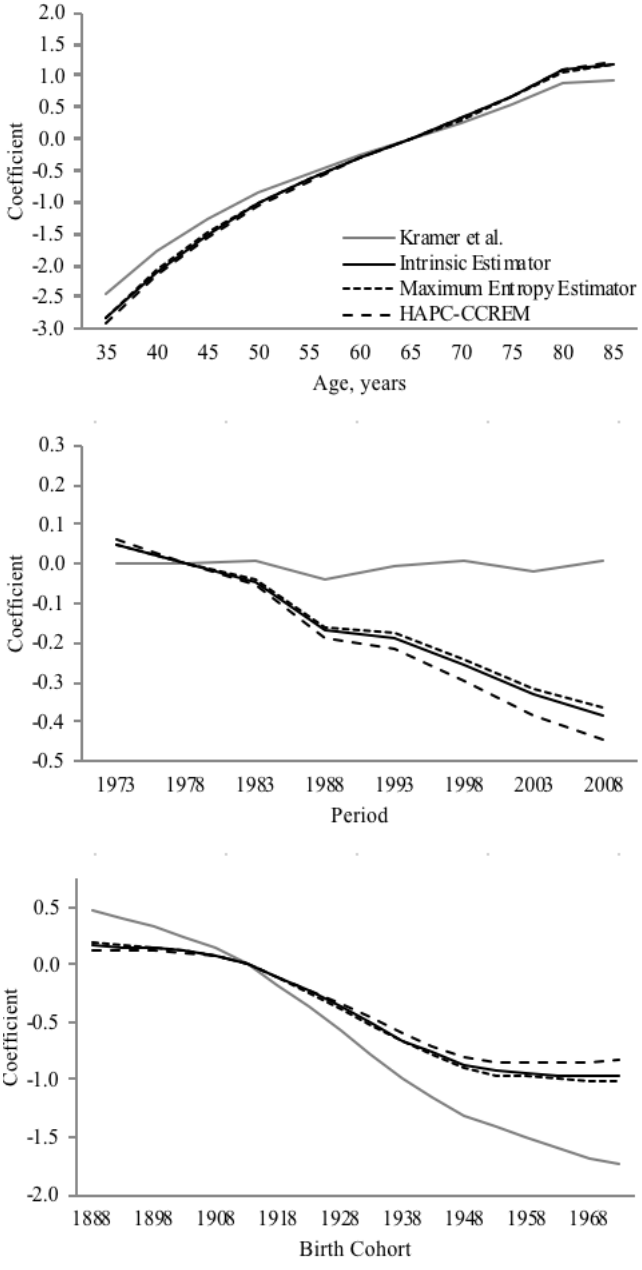


Figure 3. Age, Period, and Cohort Variation in U.S. White Men's Heart Disease Related Mortality Rates, Ages 35-85 between Years 1973-2010, Estimated from IE, MEE, HAPC-CCREM, and CGLIM.



**References**


1. Kramer, MR, Valderrama AL, and Casper ML. Decomposing black-white disparities in heart disease mortality in the United States, 1973-2010: an age-period-cohort analysis. *American Journal of Epidemiology.* 2015; 182(4):302-312.

2. Harper, S. Invited commentary: APC… it's as easy as 1-2-3! *American Journal of Epidemiology.* 2015; kwvo52.

3. Kramer, MR, and Casper ML. Kramer and Casper respond to "APC…it's as easy as 1-2-3!" *American Journal of Epidemiology.* 2015; kwv049.

4. Frost, WH. The age selection of mortality from tuberculosis in successive decades. *The Milbank Memorial Fund Quarterly*. 1940; 18(1):61-66.

5. Bell, A and Jones K. The impossibility of separating age, period, and cohort effects. *Social Science & Medicine*. 2013; 93:163-165.

6. Bell, A and Jones K. Another "futile quest"? A simulation study of Yang and Land's hierarchical age-period-cohort model. *Demographic Research.*2014; 30:333.

7. Bell, A and Jones K. Don't birth cohorts matter? A commentary and simulation exercise on Reither, Hauser, and Yang's (2009) age-period-cohort study of obesity. *Social Science & Medicine*. 2014; 101:176-180.

8. Glenn, ND. Cohort analysts' futile quest: statistical attempts to separate age, period and cohort effects. *American Sociological Review.* 1976; 41(5):900-904.

9. Glenn, ND. *Cohort Analysis*. Vol. 5. Sage; 2005.

10. Fienberg, SE and Mason WM. Identification and estimation of age-period-cohort effects in the analysis of discrete archival data. *Sociological Methodology.* 1979; 10(1):67.

11. Holford, TR. The estimation of age, period and cohort effects for vital rates. *Biometrics.* 1983; 39:311-324.

12. Kupper, LL, Janis JM, Karmous A and Greenberg BG. Statistical age-period-cohort analysis: a review and critique. *Journal of Chronic Disease*. 1985; 38(10):811-830.

13. Luo, L. Assessing validity and application scope of the intrinsic estimator approach to the age-period-cohort problem. *Demography.* 2013; 50(6):1945-1967.

14. Luo, L and Hodges JS. Block constraints in age-period-cohort models with unequal-width intervals. *Sociological Methods & Research.* 2015; 0049124115585359.

15. Luo, L, Hodges JS, Winship C, and Powers D. The sensitivity of the intrinsic estimator to coding schemes: a comment on Yang, Schulhofer-Wohl, Fu, and Land. *American Journal of Sociology*. 2016; 122: 930-961.

16. Mason, KO, Mason WM, Winsborough HH, and Poole WK. Some methodological issues in cohort analysis of archival data. *American Sociological Review.* 1973; 38(2):242-258.

17. Hobcraft, J, Menken J, and Preston S. Age, period, and cohort effects in demography: a review. Springer, New York. 1985; 89-135.

18. Reither, EN, Masters RK, Yang YC, Powers DA, Zheng H, and Land KC. Should age-period-cohort studies return to the methodologies of the 1970s? *Social Science & Medicine.* 2015; 128:356-365.

19. Smith, HL. Advances in age-period-cohort analysis. *Sociological Methods & Research.* 2008; 36(3):287-296.

20. Yang, YC and Land KC. Age-period-cohort analysis: new models, methods, and empirical applications. CRC Press; 2013.

21. Frenk, SM, Yang YC, and Land KC. Assessing the significance of cohort and period effects in hierarchical age-period-cohort models: applications to verbal test scores and voter turnout in U.S. presidential elections. *Social Forces.* 2013; 92(1):221-248.

22. Kupper, LL and Janis JM. The multiple classification model in age, period, and cohort analysis: theoretical considerations. *Institute of Statistics Mimeo*. 1980; 1311. Chapel Hill: Department of Biostatistics, University of North Carolina.

23. Fu, WJ. Ridge estimator in singular design with applications to age-period-cohort analysis of disease rates. *Communication in Statistics Theory and Method*. 2000; 29:263-278.

24. Fu, WJ. Constrained estimators and consistency of a regression model on a Lexis diagram. *Journal of the American Statistical Association*. 111(513): 180-199.

25. O'Brien Age period cohort models and the perpendicular solution. Epidemiol. Methods. 2015; 4(1): 87-99.

26. Yang, YC, Schulhofer-Wohl S, Fu WJ, and Land KC. The intrinsic estimator for age-period-cohort analysis: what it is and how to use it. *American Journal of Sociology.* 2008; 113(6):1697-1736.

27. Yang, YC, Fu WJ, and Land KC. A methodological comparison of age-period-cohort models: the intrinsic estimator and conventional generalized linear models. *Sociological Methodology.* 2004; 34(1):75-110.

28. Browning, M, Crawford I, and Knoef M. An entropy-based approach to the APC problem with applications to female mortality and labour-force participation. Department of Economics, Oxford University; 2012

29. Fosse, E, and Winship E. Moore-Penrose Estimators of Age-Period-Cohort Effects. Their Interrelationships and Properties. *Sociological Science*. June 7, 2018. 10.15195/v5.a14

30. Yang, YC, and Land K. A mixed model approach to the age-period-cohort analysis of repeated cross-section surveys with an application to data on trends in verbal test scores. *Sociological Methodology.* 2006; 36(1):75-97.

31. Yang, YC, and Land KC. Age-period-cohort analysis of repeated cross-section surveys: fixed or random effects? *Sociological Methods & Research*. 2008; 36(3):297-326.

32. Masters, RK, Hummer RA, Powers DA, Beck A, Lin SF, and Finch BK. Long-term trends in adult mortality for US blacks and whites: an examination of period- and cohort-based changes. *Demography*. 2014; 51(6):2047-2073.

33. Masters, RK, Powers DA, Hummer RA, Beck A, Lin SF, and Finch BK. Fitting age-period-cohort models using the intrinsic estimator: assumptions and misapplications. *Demography.* 2016; forthcoming.

34. Keyes, KM, Utz RL, Robinson W, and Li G. What is a cohort effect? Comparison of three statistical methods for modeling cohort effects in obesity prevalence in the United States, 1971-2006. *Social Science & Medicine.* 2010; 70(7):1100-1108.

35. Tukey, JW. Exploratory data analysis. Addison-Wesley, Reading, MA. 1977.

36. Reither, EN, Yang YC, and Olshansky SJ. New forecasting methodology indicates more disease and earlier mortality ahead for today's younger Americans. *Health Affairs.* 2011; 30(8):1-7.

37. Masters, RM, Tilstra AM, and Simon DH. Explaining recent mortality trends among younger and middle-aged white Americans. *International Journal of Epidemiology.* 2017; 47(1): 81-88.

38. Yang, YC, and Land KC. Misunderstandings, mischaracterizations, and the problematic choice of a specific instance in which the IE should never be applied. *Demography.* 2013; 50(6):1969.