Testing the Validity and Consistency of
Hierarchical Age-Period-Cohort Cross-Classified Models

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

ABSTRACT

Hierarchical Age-Period-Cohort (HAPC) Cross-Classified Fixed Effects Models (CCFEM) and Random Effects Models (CCREM) have been increasingly used by social scientists to investigate temporal variation in numerous outcomes across ages, time periods, and birth cohorts. The models have received recent scrutiny and testing, with some researchers cautioning that HAPC models estimate (1) biased and (2) inconsistent age, period, and cohort effects. These previous findings, however, were based on a misrepresentation of the HAPC modeling framework and were derived from exercises that applied HAPC models to unrealistic simulated data. In this article we discuss the scope and application of HAPC models and test the validity and consistency of HAPC estimates of age-, period-, and cohort-based variation in outcomes using simulated data. We replicate previous findings and show that that existing criticisms of HAPCs apply to rare, select circumstances, with the previous poor performance of HAPC models stemming from (1) misapplication of the models on (2) unrealistic simulated data. Findings from simulated data in this paper show that fitted HAPC-CCREMs and HAPC-CCFEMs estimate the "true" age, period, and cohort effects in simulated data when applied to (1) Age-Period-Cohort data structures in which cohort membership is not a function of one's age and period, or (2) data in which the functional forms of age effects, period effects, and cohort effects on the outcome are not assumed to be linear. Further, we show that distributions of the age, period, and cohort effects estimated from Markov chain Monte Carlo simulations using Gibbs sampling for the HAPC-CCREM are also consistent with "true" effects in both individual-level data and aggregate rate data.

We thank the Robert Wood Johnson Foundation's Health and Society Scholars Program for its financial support.

Our paper will proceed across several analytical steps.

Step 1: We replicate the data, methods, and findings from a paper “The Cross-Classified Age-Period-Cohort Model as a Constrained Estimator” presented at 2013 PAA Annual Meeting Session 24, Innovative Theory and Methods for Demographic Research, by Liying Luo and James Hodges. The authors concluded that the HAPC fitted to three simulated datasets produced biased and inconsistent results. We show that these authors made two critical errors in their analytic design: (1) the “true” age, period, and cohort effects selected by the authors create data that bury temporal variation by assuming a linear functional form for all three temporal dimensions. As such, these authors applied HAPC models to data that did not exhibit any temporal-based variation, thereby assuring the models’ failures to detect any period or cohort effects. This is shown in Table1 (the “true” effects from three simulated datasets created by the authors), Table 2 (the observed age-specific outcomes across time periods as functions of the “true” effects), and Figure 1 (graphical plots of the observed outcomes) below.

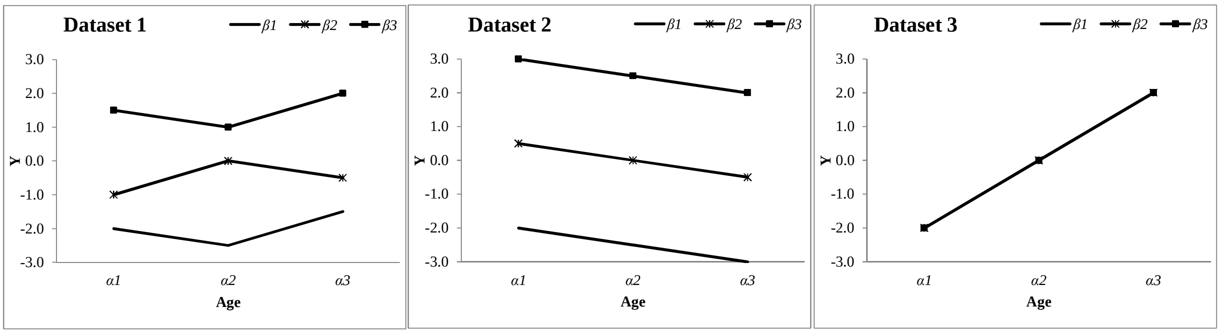
Table 1. Replication of Luo's "True" Age, Period, and Cohort Effects.

Dataset	Age			Period			Cohort				
	α1	*α2*	*α3*	*β1*	*β2*	*β3*	*γ1*	*γ2*	*γ3*	*γ4*	*γ5*
NO. 1	-1	0	1	-1	0	1	-1.5	-1.5	0	0	1.5
NO. 2	-1	0	1	-1	0	1	-3	-1.5	0	1.5	3
NO. 3	-1	0	1	-1	0	1	2	1	0	-1	-2

Table 2. "True" Observed Outcomes from the Combined "True" Age, Period, and Cohort Effects.

Dataset	NO. 1			NO. 2			NO. 3		
	β1	*β2*	*β3*	*β1*	*β2*	*β3*	*β1*	*β2*	*β3*
α1	-2	-1	1.5	-2	-1	1.5	-2	-2	-2
α2	-2.5	0	1	-2.5	0	1	0	0	0
α3	-1.5	-0.5	2	-1.5	-0.5	2	2	2	2

Figure 1. “True” Observed Outcomes from the Combined “True” Age, Period, and Cohort Effects.



Dataset 1 Observations exhibit variation in Age and Period, and Cohort (i.e., Age-specific outcomes vary across Period).
Dataset 2 Observations exhibit variation in Age and Period only (i.e., parallel Age effects).
Dataset 3 Observations exhibit variation in Age only (i.e., Age effects are uniform across time).

As seen in Figure 1, only in Dataset NO. 1 do all three temporal dimensions exhibit some variation in the outcome. The age-specific outcomes in Dataset NO. 2 vary across period in a uniform/parallel manner, and thus only exhibit period-based temporal variation. No cohort variation is detectable. Finally, the age-specific outcomes in Dataset NO. 3 are exactly the same across time periods and birth cohorts. Thus, in this case there is neither cohort variation nor period variation in the outcome. Had Luo and Hodges estimated model fit statistics for Datasets NO. 2 & 3 they would have seen that an age-period model (in the case of Dataset No. 2) or an age-only model (in the case of Dataset NO. 3) would have been preferred to an APC model. Indeed, the need to perform model fit tests before applying APC models is a point stressed by multiple APC researchers (Yang and Land 2013).

The second critical error committed by Luo and Hodges was to frame the application of HAPC-CCREMs as only occurring to data from tabular rates of age-specific data across time periods, wherein cohort is produced as a direct relationship from period-age = cohort. Data of this structure suffer the identification problem in which the values of age, period, and cohort are absolutely dependent on each other. However, the authors did not highlight the fact that multiple real-life applications of the HAPC-CCREM modeling framework have been applied to individual-level data wherein respondents self-report their year of birth (such as the National Health Interview Survey) or other data structures in which the year of respondents' births are known (such as knowing the birth year of mothers in the National Vital Statistics Birth Data). In these cases, cohort groupings can be created from the self-reported year of birth rather than being a direct linear outcome from period-age in tabular data.

In short, multiple instances show that the HAPC-CCREM have been applied to data structures that do not reflect the C=P-A identification problem, yet Luo and Hodges present the HAPC-CCREM as performing well only under the circumstances in which a researcher must constrain the effects of neighboring cohorts to be equal. This is not the case.

Taken together, we show that Luo and Hodges misrepresent the application of HAPC-CCREMs and misapply the models themselves to data that should not be analyzed with an APC framework.

Step 2: We next discuss the assumptions behind the linear dependency of C = P-A, and the appropriate application of HAPC models.

Here, for example, is the classic C = P-A identification problem stemming from tabulated Age-specific outcome across Periods. In these data we do not know Cohort, so we must assign it as a linear function of Period-Age:

Linear Assumption / Tabular Data

	$\beta 1$	$\beta 2$	$\beta 3$
$\alpha 3$	$\gamma 1$	$\gamma 2$	$\gamma 3$
$\alpha 2$	$\gamma 2$	$\gamma 3$	$\gamma 4$
$\alpha 1$	$\gamma 3$	$\gamma 4$	$\gamma 5$

Assumes five C = P-A

Example of Tabular Data

	1990	1991	1992
52	1938	1939	1940
51	1939	1940	1941
50	1940	1941	1942

Age, Period, and Cohort obtained from individual-level data in which survey respondents self-report their birth year, however, do not suffer from the linear dependency behind the identification problem.

Table. 3. Age, Period, and Cohort from Individual-level Data

	$\beta 1$		$\beta 2$		$\beta 3$	
$\alpha 3$	$\gamma 0$	$\gamma 1$	$\gamma 1$	$\gamma 2$	$\gamma 2$	$\gamma 3$
	$\gamma 0$	$\gamma 1$	$\gamma 1$	$\gamma 2$	$\gamma 2$	$\gamma 3$
$\alpha 2$	$\gamma 1$	$\gamma 2$	$\gamma 2$	$\gamma 3$	$\gamma 3$	$\gamma 4$
	$\gamma 1$	$\gamma 2$	$\gamma 2$	$\gamma 3$	$\gamma 3$	$\gamma 4$
$\alpha 1$	$\gamma 2$	$\gamma 3$	$\gamma 3$	$\gamma 4$	$\gamma 4$	$\gamma 5$
	$\gamma 2$	$\gamma 3$	$\gamma 3$	$\gamma 4$	$\gamma 4$	$\gamma 5$

Uses Individual-level Self-reported Birth Year to Create C ≠ P-A

Just as in a classic Lexis diagram, every person experiences a calendar year at two different ages. Thus, when birth cohort is self-reported in individual-level data we can observe respondents from a given cohort at one age across two periods. In the design above, for instance, persons in birth cohort $\gamma 2$ experience age $\alpha 2$ during the time periods $\beta 1$ and $\beta 2$. When we have individual-level information on birth year the APC data can be scaled up, such that, for example, we can set α_i, β_j, and γ_k to all be five years wide, 10 years wide, or to be various widths. To illustrate, here are actual data from the National Health Interview Survey, waves 1986-2004, linked to mortality records at the National Death Index through December 31, 2006. Because we have survey respondents' self-reports of age, birth year, and enumerator-reported survey year, we can freely create Age, Period, and Cohort groupings that need not be direct, linear functions of one another. In the case below we observe five year cohorts age across five year periods in terms of five year age intervals.

Age-Period-Cohort Design from Individual-level Data in the National-Health Interview Survey, 1990-2005.

5yr X 5yr X 5yr Age-Period-Cohort Design for 50-65 year Old Age Groups

	Period 2, [1990-1995)					Period 3, [1995-2000)					Period 4, [2000-2005)				
	1926	1927	1928	1929	1930	1931	1932	1933	1934	1935	1936	1937	1938	1939	1940
	Cohort 1, [1925-1930)			1930	1931	1932	1933	1934	1935	1936	1937	1938	1939	1940	1941
Age 3, [60-65)	1928	1929	1930	1931	1932	1933	1934	1935	1936	1937	1938	1939	1940	1941	1942
	1929	Cohort 2, [1930-1935)					1935	1936	1937	1938	1939	1940	1941	1942	1943
	1930	1931	1932	1933	1934	1935	1936	1937	1938	1939	1940	1941	1942	1943	1944
	1931	1932	1933	1934		Cohort 3, [1935-1940)				1940	1941	1942	1943	1944	1945
	1932	1933	1934	1935	1936	1937	1938	1939	1940	1941	1942	1943	1944	1945	1946
Age 2, [55-60)	1933	1934	1935	1936	1937	1938	1939	Cohort 4, [1940-1945)					1945	1946	1947
	1934	1935	1936	1937	1938	1939	1940	1941	1942	1943	1944	1945	1946	1947	1948
	1935	1936	1937	1938	1939	1940	1941	1942	1943	1944	1945	1946	1947	1948	1949
	1936	1937	1938	1939	1940	1941	1942	1943	1944		Cohort 5, [1945-1950)				1950
	1937	1938	1939	1940	1941	1942	1943	1944	1945	1946	1947	1948	1949	1950	1951
Age 1, [50-55)	1938	1939	1940	1941	1942	1943	1944	1945	1946	1947	1948	1949	1950	1951	1952
	1939	1940	1941	1942	1943	1944	1945	1946	1947	1948	1949	Cohort 6, [1950-1955)			
	1940	1941	1942	1943	1944	1945	1946	1947	1948	1949	1950	1951	1952	1953	1954

The lesson as it applies to the current exercises is this: the design of the APC data depends on our assumptions about the source. Are the data individual-level (or contain some other way by which we know birth year such that cohorts do not need to be linearly produced from period-age), or are they tabulated rate data that suffer from the linear dependency of period-age=cohort? Luo's and Hodges's paper and presentation assumed only the latter.

Example of Individual-level Data

	1990		1991		1992	
52	1937	1938	1938	1939	1939	1940
	1937	1938	1938	1939	1939	1940
51	1938	1939	1939	1940	1940	1941
	1938	1939	1939	1940	1940	1941
50	1939	1940	1940	1941	1941	1942
	1939	1940	1940	1941	1941	1942

Uses Individual-level Self-reported Birth Year to Create C ≠ P-A

Example of Individual-level Data Assuming Linear Dependence

	1990		1991		1992	
52	1938	1938	1939	1939	1940	1940
	1938	1938	1939	1939	1940	1940
51	1939	1939	1940	1940	1941	1941
	1939	1939	1940	1940	1941	1941
50	1940	1940	1941	1941	1942	1942
	1940	1940	1941	1941	1942	1942

Uses Age at time Period to Assign Cohort

Step 3: We next introduce two different data designs that each addresses the two separate errors in Luo's and Hodges exercises: (1) introduce data as though it were obtained at the individual-level, and thus has a sixth cohort, with two cohorts for every age x period cell; and (2) assume non-linear functional form of cohort's effect on Y.

Here, we introduce a sixth cohort to the original three datasets, assuming the data to be individual-level and, thus, following the structure depicted in Table 3.

Table 4. "True" Age, Period, and Cohort Effects from Individual-level Data

Dataset	**Age**			**Period**			**Cohort**					
	α1	*α2*	*α3*	*β1*	*β2*	*β3*	*γ1*	*γ2*	*γ3*	*γ4*	*γ5*	***γ6***
NO. 1	-1	0	1	-1	0	1	-1.5	-1.5	0	0	1.5	**2**
NO. 2	-1	0	1	-1	0	1	-3	-1.5	0	1.5	3	***4.5**
NO. 3	-1	0	1	-1	0	1	2	1	0	-1	-2	***-3**

* Continue to assume linear functional form

Next, we then create three alternative datasets from Luo's and Hodges's original data sets that assume non-linear effects of cohort on Y:

Table 5. "True" Age, Period, and Cohort Effects in Data with Non-linear Functional Form for Cohort

Dataset	**Age**			**Period**			**Cohort**				
	α1	*α2*	*α3*	*β1*	*β2*	*β3*	*γ1*	*γ2*	*γ3*	*γ4*	*γ5*
NO. 1'	-1	0	1	-1	0	1	-1.5	-1.5	0	0	*1.5
NO. 2'	-1	0	1	-1	0	1	-3	**-0.5**	0	**0.5**	3
NO. 3'	-1	0	1	-1	0	1	**1**	**0.9**	**0.25**	**-0.25**	**-0.1**

* Dataset NO. 1 already assumed a non-linear functional form of cohort, so no changes were made

Figure 2. "True" Cohort Effects, Original Data from Table 2 and Non-linear Data from Table 5.

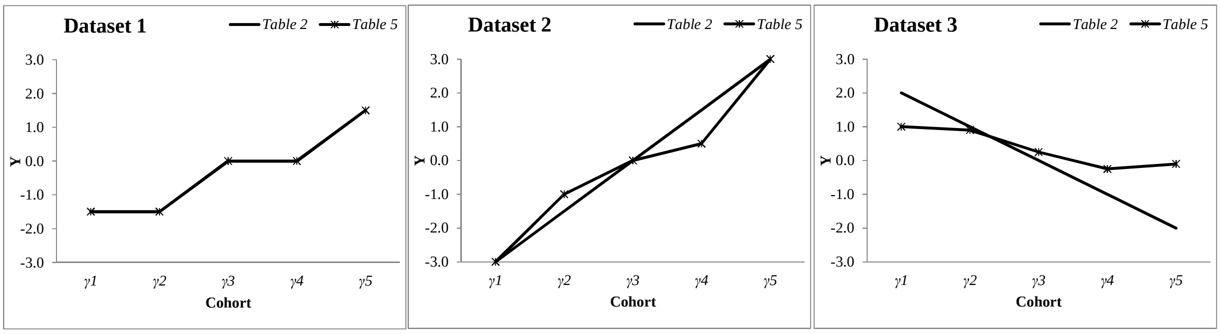
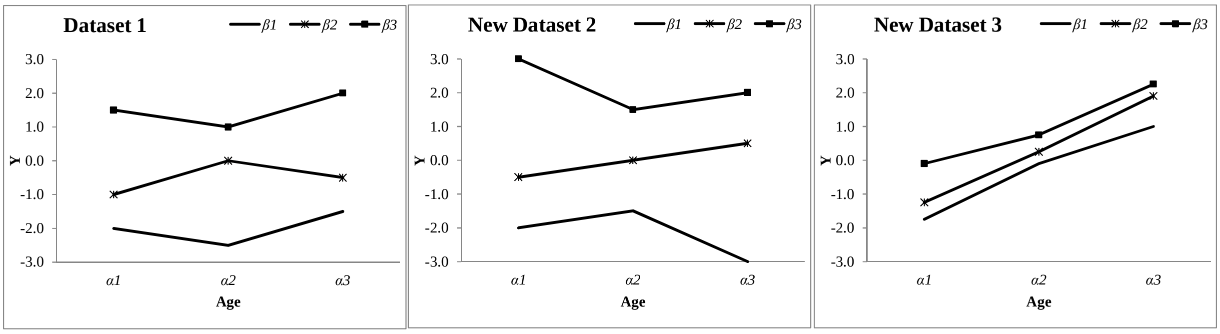


Figure 3. "True" Observed Outcomes from the Combined "True" Age, Period, and Cohort Effects in Table 5.



Dataset 1 Observations exhibit variation in Age and Period, and Cohort (i.e., Age varies across Period).
Dataset 2 Observations exhibit variation in Age and Period, and Cohort (i.e., Age varies across Period).
Dataset 3 Observations exhibit variation in Age and Period, and Cohort (i.e., Age varies across Period).

Step 4 – Fit HAPC-CCFEM and HAPC-CCREM on data from Table 2 to Replicate Luo's & Hodges's findings:

Table 6. Simulation Results: CCFEM and CCREM estimates for the three datasets in Table 1.

		Dataset NO. 1			Dataset NO. 2			Dataset NO. 3		
		Assigned	CCFEM	CCREM	Assigned	CCFEM	CCREM	Assigned	CCFEM	CCREM
Age	1	-1	-1.00	-1.00	-1	0.47	0.48	-1	-2.03	-2.02
	2	0	-0.01	-0.01	0	-0.01	0.01	0	-0.01	0.01
	3	1	1.01	1.01	1	-0.45	-0.48	1	2.05	2.02
Period	1	-1	-0.97	-0.99	-1	-2.44	-2.48	-1	0.06	0.00
	2	0	0.00	-0.02	0	0.00	-0.01	0	0.00	0.00
	3	1	1.03	1.01	1	2.50	2.50	1	0.00	0.00
Cohort	1	-1.5	-1.49	-1.47	-3	-0.06	0.00	2	-0.06	0.00
	2	-1.5	-1.49	-1.47	-1.5	-0.02	0.00	1	-0.02	0.00
	3	0	0.00	0.02	0	0.00	0.00	0	0.00	0.00
	4	0	0.00	0.02	1.5	0.03	0.00	-1	0.03	0.00
	5	1.5	1.44	1.46	3	0.00	0.00	-2	0.00	0.00

Note: CCFEM and CCREM models in Dataset NO.1 constrained Cohort 1=2 and Cohort 3=4.

These results are the same as those presented by Luo and Hodges. They show the inability of the HAPC-CCFEM and HAPC-CCREM to estimate the "true" APC effects from Table 1.

Step 5: Refit HAPC-CCFEM and HAPC-CCREM on data from Table 4: individual-level data containing 3 Age, 3 Period, and 6 Cohort Effects.

Table 7. Simulation Results: CCFEM and CCREM estimates for the three datasets in Table 4.

		Dataset NO. 1*			Dataset NO. 2			Dataset NO. 3		
		Assigned	CCFEM	CCREM	Assigned	CCFEM	CCREM	Assigned	CCFEM	CCREM
Age	1	-1	-1.00	-1.00	-1	-1.00	-0.97	-1	-0.97	-0.97
	2	0	**-0.04	**0.06	0	-0.02	0.75	0	0.04	-0.51
	3	1	1.00	1.00	1	1.00	1.00	1	1.01	1.01
Period	1	-1	-0.99	-0.99	-1	-0.98	-1.04	-1	-1.07	-1.03
	2	0	0.00	0.00	0	0.00	0.01	0	0.00	0.03
	3	1	1.00	1.00	1	0.98	1.03	1	0.97	1.00
Cohort	1	-1.5	-1.53	-1.63	-3	-2.98	-3.61	2	2.09	2.60
	2	-1.5	-1.50	-1.60	-1.5	-1.46	-2.24	1	0.97	1.48
	3	0	0.00	-0.10	0	0.00	-0.73	0	0.00	0.51
	4	0	0.02	-0.07	1.5	1.53	0.75	-1	-1.03	-0.51
	5	1.5	1.57	1.47	3	3.07	2.17	-2	-2.04	-1.52
	6	2	2.02	1.92	4.5	4.53	3.66	-3	-3.08	-2.56

*Cohorts 1 and 2, and Cohorts 3 and 4 were not constrained to be equal in Dataset NO. 1

** Age 2 "Effect" in the CCFEM and CCREM Colums are model intercepts

In all datasets, the HAPC-CCFEMs and HAPC-CCREMs accurately and consistently estimate the "true" effects in all three datasets. All estimated coefficients are nonsignificantly different from the true effects. Thus, when applied to individual-level Age-Period-Cohort data, the HAPC method is both valid and reliable, and need not apply any constraints on the Age, Period, or Cohort terms. The primary mistake Luo and Hodges made was in trying to inappropriately apply the HAPC method to tabulated data in which C=P-A and in which no temporal variation was detectable.

Next, we fit HAPC-CCREM using Markov Chain Monte Carlo simulations using Gibbs sampling on tabulated rate data in which C=P-A, but in which the effect of C is not assumed to be linear. That is, we fit MCMC HAPC-CCREMs on data in Table 5.

Step 6: Fit MCMC HAPC-CCREM on data from Table 5:group-level data containing 3 Age, 3 Period, and 5 Cohort Effects in which C=P-A. Also, we test if we can replicate estimates from CCFEMs and CCREMs on data from Table 4 using MCMC HAPC-CCREM.

Table 8. Simulation Results: MCMC estimates for the three datasets in Table 4 and Table 5.

		Dataset NO. 1*				Dataset NO. 2				Dataset NO. 3			
		Table 4	MCMC	Table 5	MCMC	Table 4	MCMC	Table 5	MCMC	Table 4	MCMC	Table 5	MCMC
	1	-1	-0.99	-1	-0.95	-1	-0.99	-1	-0.64	-1	-1.00	-1	-1.08
Age	2	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00
	3	1	1.00	1	0.97	1	1.00	1	0.66	1	1.01	1	1.03
	1	-1	-0.99	-1	**-1.02	-1	-0.99	-1	-1.33	-1	-0.98	-1	-0.95
Period	2	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00
	3	1	0.98	1	1.08	1	0.98	1	1.39	1	0.98	1	0.92
	1	-1.5	-1.47	-1.5	-1.40	-3	-2.98	-3	-2.28	2	2.02	1	0.60
	2	-1.5	-1.45	-1.5	-1.44	-1.5	-1.45	-1	-0.63	1	1.04	.9	0.59
Cohort	3	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	.25	0.00
	4	0	0.03	0	-0.05	1.5	1.53	.5	0.14	-1	-0.97	-.25	-0.43
	5	1.5	1.57	1.5	1.34	3	3.07	3	2.22	-2	-1.92	-.1	-0.18
	6	2	2.03			4	4.53			-3	-2.96		

*Cohorts 1 and 2, and Cohorts 3 and 4 were not constrained to be equal in Dataset NO. 1

** Period and Cohort Coefficients Centered on P2 and C3

First, the MCMC HAPC-CCREMs accurately and consistently estimate the "true" A, P, and C effects in the individual-level data. Furthermore, the point estimates are more accurate than those estimated by the HAPC-CCREM in Table 7.

Second, even for tabular data in which C=P-A identification problem hinders model convergence for the HAPC-CCFEM and HAPC-CCREM, the MCMC HAPC-CCREM estimates A, P, and C effects that are *consistent* with the "true" effects.

Step 7: Finally, we fit MCMC HAPC-CCREMs to data simulated from empirically estimated APC effects from both individual-level data and from tabulated rate data. That is, we use the HAPC-CCREM modeling framework to determine if these models can retrieve the estimated APC effects in data simulated from (1) results presented by Powers (2013) looking at variation in US infant mortality rates by mother's age, mother's birth cohort, and year of birth; and (2) results estimated from APC models using the intrinsic estimator (IE) to estimate age-, period-, and cohort-based variation in non-Hispanic black men's mortality from infectious diseases between 1960 and 2009. In the first case, cohort is self-reported by the mother and thus the APC data structure is a real-life example of the data structure depicted in Table 4. In the second case, five year age-specific death rates among US black men are tabulated across five year time periods. Thus, in this case the data structure is a real-life example of the data structure depicted in Table 5, where P-A=C.

Preliminary results (not presented) show that the MCMC HAPC-CCREMs are able to retrieve the original estimates of age, period, and cohort effects in both sets of simulated fake data.

References

Luo, L and J Hodges. 2013. "The Cross-Classified Age-Period-Cohort Model as a Constrained Estimator." Paper Presented at *Session 24, Innovative Theory and Methods for Demographic Research*, Population Association of America 2013 Annual Meeting. New Orleans, LA.

Powers, DA. 2013. "Black-White Differences in Maternal Age, Maternal Birth Cohort, and Period Effects on Infant Mortality in the US (1983-2002)." *Social Science Research* 42(4): 1033-1045.

Yang, YC and KC Land. 2013. Age-Period-Cohort Analysis: New Models, Methods, and Empirical Applications. Chapman & Hall/CRC Interdisciplinary Statistics Series. New York, NY: CRC Press, Taylor & Francis Group.