Age-Period-Cohort with *hysteresis* APC-H model / A Method

Louis Chauvel chauvel@louischauvel.org

Abstract

Despite the improvements of the APC methodology over the last decade, the problem of stability of cohort effects remains. Even if it at been expressed by the main specialist of the cohort issue, explicitly mentioned in the early Mannheim (1928) or Ryder (1965) discussions, and also in the "cohort inversion" versus "accumulating cohort effects" models of Hobcraft and colleagues (1982) and in Yang and colleagues APC- intrinsic estimator (2008), it has not received sufficient methodological answer. Here, a better APC method is proposed: first, the APC-D (detrended) model delivers a DCE (detrended cohort effect, having a zero-sum zero-slope shape) set of parameters, and second, the APC-H (*hysteresis*) model assesses a H index that tests the stability of the DCE over age span. An appropriate standardization of H offers a test between three configurations: increase or "Mathew effect" (H>0), linear stability or *hysteresis* (H=0), decrease or resilience (H<0) of the DCE over the age span. Moreover, a complete absorption of the cohort effect over life span is obtained when H=-1. Conversely, in case of development over life span of a DCE from zero at minimum age, we have H=+1. The methodological difficulty of the APC-H model is that we can't estimate both the DCE and the H coefficient inside the same model since H is not a linear interaction.

The strategy here is to develop an iterative process that converges toward a unique solution including a final DCE(n) and a final H(n) hysteresis index. The initial step is based on the APCD model, which delivers an initial detrended cohort effect DCE(0) defining a mean cohort effect over life course. On the following iterations n, we alternate the estimation of H(n) when DEC(n-1) is given, and thus of DCE(n) on the base of the previously estimated H(n). This process converges to a final DCE and H index.

Expression of the model

An OLS specification[1] of the APCH model is presented here: we consider a dependant variable y, for instance the logged annual income, in a microdata set of independent cross-sectional surveys. Here, y_i^{apc} pertains to individual i of age a in period p, and thus belonging to the birth cohort c=p-a. The intervals $[a_{min},a_{max}]$ and $[p_{min},p_{max}]$ denotes respectively the age-span and the periods of observation defined by a series of cross-sectional independent sample surveys carried with a regular pace p (yearly, or each 5 years, for example); the age groups a are based on the same pace of time than periods; we will suppose no hole in the a p rectangle. We exclude the first and the last cohorts of the estimations of the models, in order to improve the confidence intervals of the parameters. Then, c is in the interval $[p_{min}-a_{max}+1,\ p_{max}-a_{min}-1]$. Control variables (continuous or dichotomic ones) X_j, such as gender, race, education, state, etc. are to be included as covariates.

The initial DCE(0) step is the research of non zero cohort effects that denote a specific behaviour of the members of given birth cohorts, after control by period and age effects, and by control variables. The initial APCD performs that estimation where two constraints on the DCE estimates (zero-sum and zero-slope) that are imputed so that a unique stable solution is obtained (Wilmoth, 2001, Chauvel, 2001) and the traditional identification problem of the APCis solved[2].

---

[1] The GLM procedure of STATA offers a wide range of specification from OLS to logit or Poisson models of different kinds. Unfortunately ordinal or polytomic logit models, and also quantile regressions, can not be easily handled.

[2] Several solutions have been proposed, exist such as the Yang Yang and colleagues APC-IE model which performs accurate results. The only problem of APC-IE is the production of unstable cohort estimates since their slope is unstable when one changes the metrics of the dependant variable. For instance, the estimated APC-IE cohort effects are different whether one considers nominal wages or real ones when the APCD provides a unique DCE since the difference between real and nominal wages is absorbed by the period coefficients. See examples at : http://www.louischauvel.org/apchex.htm

$$
(0)\begin{cases}
y^{apc} = \alpha_a + \pi_p + \gamma_c + \alpha_0 rescale(a) + \gamma_0 rescale(c) + \beta_0 + \sum_j \beta_j X_j + \varepsilon_i \\
\begin{cases}
p = c + a \\
\sum_a \alpha_a = \sum_p \pi_p = \sum_c \gamma_c = 0 \\
Slope_a(\alpha_a) = Slope_p(\pi_p) = Slope_c(\gamma_c) = 0 \\
\min(c) < c < \max(c)
\end{cases}
\end{cases}
\qquad \text{(APCD)}
$$

- The set of constraints –on the zero-sum, zero-slopes and on the domain of estimation of the cohort effects that excludes the first and the last cohort– produce a unique estimate of DCE and solve the old APC identification problem.
- Rescale(a) is the linear function that rescales the index a (pertaining to age) from -1 to 1. This rescaling allows a better interpretation of the H index of the APC-H.
- $Slope_a(\alpha_a)$ is the linear slope of the α_a estimates. $Slope_a(\alpha_a)=0$ if and only if $\Sigma_a [(2a - a_{min} - a_{max})\ \alpha_a] = 0$
- α_a, π_p and γ_c are respectively the detrended age, period and cohort effects. The π_p effects fit the categorical period changes, and are able to absorb the period-specific changes in measurements of the dependant variable, effects of inflation, etc. The α_a represent the non-linear age changes. For our purpose, γ_c (also named "detrended cohort effects" DCE) are the most important estimates of this model since significantly non-zero γ_c coefficients will detect cohort effects.
- β_0 is the general intercept, and $\beta_j X_j$ pertain to the controls by selected covariates such as gender, race, education, etc. that can be introduced in the model.
- α_0 is the interperiod, intercohort linear slope of age. γ_0 is the interperiod interage linear slope of cohort. Since we have a linear relation p=a+c, these two coefficients are not to be naively analysed in terms of age and cohort effects but as linear time intercepts.

In this APCD model, if no cohort effect $\hat{\gamma}_c$ is significantly different to 0, the simple age and period AP model is sufficient representation of data. Raftery's (1986) BIC could help to decide between AP and APCD. Conversely, when at least one of the DCE(0) estimates is significantly different to 0, the pertaining cohort effect can be defined as an average specific behavior for the cohort, averaged on the available age-span. However, this cohort effect can be either stable or not over age-span, and we have to test this by a specific interaction between DCE(0) and age. Different configuration can be observed:

- The cohort effect is stable over age span (this is a real cohort effect) and we are in a case of *hysteresis*.
- The cohort effect diminishes or even vanishes in a configuration called "resilience".
- The cohort effect increases in a configuration of cumulative (dis-) advantages over age-span that denotes a cohort "Mathew effect".

In order to test the type of configuration, a solution is to introduce in the APCD model a specific interaction, H-hysteresis, between the estimated cohort effect DCE and the rescaled age (from -1 for the younger age group to +1 for the elder one). The H index will be 0 in case of linear stability of the cohort, +1 in case of increase from nil of the cohort effect over age span, and -1 in case of complete vanishing of the cohort effect for the elder age group.

It is impossible to simultaneously assess, in the same model, the DCE coefficient and the H-hysteresis index that is an interaction between rescaled age and DCE. The strategy is to build an iterative process of estimation where at step (n) we alternate (n,a) the estimation of the H(n) as an interaction between rescaled age and the DCE(n-1) that had been estimated in the previous step; and then (n,b) where the DCE(n) is estimated on the base of the estimate of H(n) and the interaction between rescaled age and DCE(n-1). At step (1), H(1) is estimated with the results DCE(0) of the initial APCD model.

$$
(\text{n,a})\begin{cases} y^{apc} = \alpha_a + \pi_p + \hat{\gamma}_c^{(n-1)} + H(rescale(a) * \hat{\gamma}_c^{(n-1)}) + \alpha_0 rescale(a) + \gamma_0 rescale(c) + \beta_0 + \sum_j \beta_j X_j + \varepsilon_i \\ \begin{cases} p = c + a \\ \sum_a \alpha_a = \sum_p \pi_p = 0 \\ Slope_a(\alpha_a) = Slope_p(\pi_p) = 0 \\ \min(c) < c < \max(c) \end{cases} \end{cases}
$$

$$
(\text{n, b})\begin{cases} y^{apc} = \alpha_a + \pi_p + \gamma_c + \hat{H}^{(n)}(rescale(a) * \hat{\gamma}_c^{(n-1)}) + \alpha_0 rescale(a) + \gamma_0 rescale(c) + \beta_0 + \sum_j \beta_j X_j + \varepsilon_i \\ \begin{cases} p = c + a \\ \sum_a \alpha_a = \sum_p \pi_p = \sum_c \gamma_c = 0 \\ Slope_a(\alpha_a) = Slope_p(\pi_p) = Slope_c(\gamma_c) = 0 \\ \min(c) < c < \max(c) \end{cases} \end{cases}
$$

This iterative process of estimation converges towards a unique result of an appropriate final DCE and of a final hysteresis index H. The standard error of H and BIC differences of the models are able to assess if APCD is better than AP, and if APCH is to be preferred to APCD.

References

Chauvel L., 2001, "Education and Class Membership Fluctuations by Cohorts in France and United-States (1960-2000)", presented at the Mannheim workshop of the International sociological association Research committee 28 (Stratification and social mobility), mimeo, http://www.louischauvel.org/MANNHEIM.pdf

Hobcraft J., J. Menken, and S. Preston. 1982, "Age, Period, and Cohort Effects in Demography: A Review." *Population Index* 48:4–43.

Mannheim K.,1928, "Das Problem der Generationen", *Kölner Vierteljahres Hefte für Soziologie* 7: 157-85, 309-30.

Raftery A.E., 1986, "Choosing models for cross-classifications", *American Sociological Review*, 51, 145-146.

Ryder N.B., 1965, "The Cohort as a Concept in the Study of Social Change", *American Sociological Review*, 30: 843-861.

Wilmoth J.R., 2001, "Les modèles âge-période-cohorte en démographie." In: G. Caselli, J. Vallin and G. Wunsch (eds.), *Démographie: Analyse et Synthèse. I. La Dynamique des Populations*, Paris: INED, pp. 379-397.

Yang Y., S. Schulhofer-Wohl, W. Fu, and K. Land. 2008. "The Intrinsic Estimator for Age-Period-Cohort Analysis: What It is and How to Use it?" *American Journal of Sociology*, 113: 1697-1736.