Book Section:

Bell, A. orcid.org/0000-0002-8268-5853 and Jones, K. (2020) Multilevel models for age–period–cohort analysis. In: Bell, A., (ed.) Age, Period and Cohort Effects: Statistical Analysis and the Identification Problem. Routledge , Abingdon , pp. 23-40. ISBN 9780367174422

https://doi.org/10.4324/9780429056819

Chapter 3: Multilevel models for age-period-cohort analysis

Andrew Bell and Kelvyn Jones

To be published in Bell, A (ed, 2021) *Age, Period and Cohort Effects: Statistical Analysis and the Identification Problem.* Routledge

Abstract

Multilevel models (aka mixed models, random effects models, hierarchical linear models) have been used widely when considering age, period and cohort effects. In some cases, this is presented as a solution to the identification problem, for example the 'Hierarchical APC model'. This chapter will show why this is not the case, using both simulations and real-data examples, illustrating why mixed models produce the results that they do, and showing that those results are often not in line with the underling reality. Whilst mixed models may allow APC models to be identified, they do so by making strong assumptions that are implicit in the model design and data structure. The chapter will go on to show ways in which mixed models can be used, not as a solution to the identification problem, but as a way of framing APC models that makes the assumptions of those models explicit. This is illustrated using examples of mental health in the UK at the turn of the 20th/21st century, and changes in mortality across the 20th Century.

Last updated 10.12.18

4.1 Introduction

Over the last 30 years, multilevel models (also called mixed models, random effects models, and hierarchical linear models, depending on the discipline) have become one of the most-used statistical methods in the social science. The models are able to separate variation in a dependent variable into a number of different 'levels' with different units, and have allowed a nuanced understanding of how much different sources of variance matter. They have allowed, for instance, understanding of how much schools (in comparison to individual attributes) matter for educational attainment (O'Connell and McCoach, 2008); how much hospitals matter to patient outcomes (Leyland and Goldstein, 2001), and how much neighbourhoods or countries affect individuals that live within them (Jones, 1991).

The age-period-cohort identification problem can also be thought of as a problem of partitioning variance into different parts – that is, understanding to what extent change in a given outcome's variance is dependent on age, period and cohort. Not only that, but longitudinal data that is often used to attempt APC identification is inherently multilevel (Bell, 2019; Fitzmaurice et al., 2011). Panel data consists of individuals measured on multiple occasions, creating a multilevel structure of occasions nested in individuals. Repeated cross-sectional data consists of individuals nested within surveys. As such, it is perhaps unsurprising that multilevel models have become a key focus of age-period-cohort analysis, both as a framework through which to specify identifying constraints (see chapter 2), and as a potential solution to the identification problem itself.

This chapter discusses those models, and the extent to which they provide a useful framework for the analysis of age-period-cohort effects. We will also show why these models do not work as a solution to the age-period-cohort identification problem, and why multilevel age-period-cohort models produce the potentially biased results that they do. If you have read the preceding chapters, it will not surprise you to learn that much like the Intrinsic Estimator (chapter 3), these models are not panaceas that solve the identification problem – indeed we show that these models make implicit assumptions that are as strong as any made by other models. The purpose of this chapter is to make the models implicit assumptions explicit, so that researchers can fully understand the strengths and limitations of multilevel models, and make decisions about when these models will and will not be useful.

The chapter proceeds as follows. First, we give a brief introduction to multilevel models and how they work, before proceeding to thinking about how such models could work in an age-period-cohort framework. We then discuss the different combinations of fixed and random parameters that we can use to estimate as age, period and cohort effects, and discuss what those different parameterisations mean in terms of the explicit assumptions that they make. We then focus in particular on the Hierarchical Age Period Cohort model (HAPC) (Yang and Land, 2006), which uses a multilevel model to attempt to disentangle APC and solve the identification problem – we show that it does not work as an all-purpose solution, and explain why it finds the results that it finds. Finally, we finish with a discussion of what the models discussed in this chapter can achieve, with examples focussing on mental health and mortality.

4.2 What are multilevel models?

Multilevel models are an extension of regression models, used when data spans multiple 'levels' – that is there are multiple units of analysis at which an outcome variable varies. These models are used extensively across the social science. In education research, there is often interest in how attainment is affected by different attributes of pupils, classes, and schools. In this instance, multilevel models can be used to find both how pupil, class and school attributes affect attainment, but also consider how individual classes and schools achieve higher and lower attainment, controlling for their measured attributes.

A multilevel model might be specified as follows:

$$Y_{ij} = \beta_0 + \beta_1 X_{1ij} + \beta_2 X_{2j} + u_j + e_{ij} \tag{4.1}$$

In this 2-level model, Y_{ij} measures the attainment of a pupil i in a school j. X_{1ij} is a pupil-level variable (for example, past performance in an exam). X_{2j} is a school-level variable (for example, the size of the school. There are then two residual terms: u_j the school-level residual, and e_{ij} the pupil-level residual. Both of these are assumed to be Normally distributed, with a mean of zero and a variance that is estimated.

$$u_j \sim N(0, \sigma_u^2), \quad e_{ij} \sim N(0, \sigma_e^2)$$

(4.2)

We can tell, from such a model, the effects of the measured variables (through β_1 and β_2), just like with single-level regression – we call this the 'fixed part' of the model as its effect is unchanging across school and pupil. Additionally, we are estimating the model's 'random part', which includes the effect of unmeasured school effects (through the variance of the school-level residuals u_j) where 'random' simply means 'allowed to vary'. Thus, we could answer questions around how much an attribute of schools (such as X_{2j} above) is related to attainment, as well as how much schools seem to matter generally in comparison to unmeasured student attributes (through the estimates of the school variance σ_u^2 in comparison to the student-level variance σ_e^2. We can also use these models to consider attainment differences between specific schools once the variables in the fixed part of the model have been accounted for (on the basis of different estimated values of u_j).

This is a two-level model (where the two levels are students and schools). However, the models are extendable to include additional levels (e.g. extending the above model to include three levels: pupis, classes, and schools). These levels do not need to be exactly nested in one another, such that one could include for example both a school and neighbourhood level in a model, to understand how both a pupil's school and their home neighbourhood are related to their attainment. So, students at a given school might live in multiple neighbourhoods, and students that live in a particular neighbourhood might go to different schools. These models, where the levels are not exactly nested, are called crossclassified models. Clearly, multilevel models are highly flexible at capturing the complex structures present in many social situations.

Such models are frequently used with longitudinal data. When using panel data – data that follows individuals over time – we can use a multilevel structure of occasions (or, repeated measures) nested in individuals (that could be nested in further levels such as schools or neighbourhoods). When using repeated cross-sectional data – repeated surveys of the same population, but different samples each time – again, there is an inherent multilevel structure of individuals nested within surveys/years. Because of this, multilevel models are a standard way of modelling data over time.

4.3 Why use multilevel models for age-period-cohort analysis?

Given the use of multilevel models in longitudinal analysis, it makes sense that such models would be used, in some form or another, for analysing age, period and cohort analysis. Indeed, conceptualising at least period and cohort as 'levels' in a multilevel model makes a lot of sense. Periods and cohorts can be understood as contexts in which people exist – individuals are situated within the occasion of measurement, and they are situated within the generation (birth cohort) that they were born into. Just like neighbourhoods and schools influence individuals that reside within them, generations and occasions also have a conceptual top down effect on individuals. And whilst it is less conceptually clear that age can be thought of as a context (rather than an attribute of an individual), there is no technical reason why it should not be treated in that way, given individuals of the same age share common experiences at a given time of their lives.

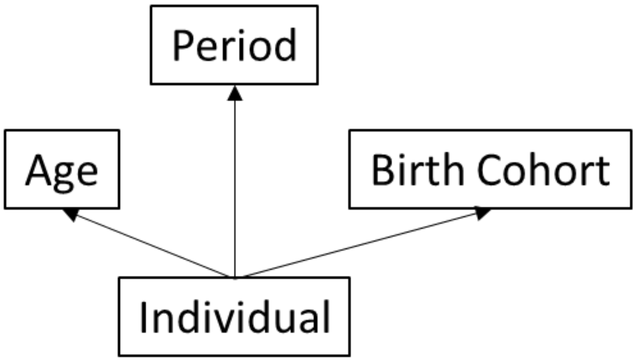
One could, therefore, at least in theory, estimate a model that treats all of age, period, and cohort as levels in a multilevel model, with each discrete value of the three terms treated as additive random effects:

$$Y_{ij} = \beta_0 + u_p + u_c + u_a + e_i$$

$$u_p \sim N(0, \sigma_{up}^2),\ \ u_c \sim N(0, \sigma_{uc}^2),\ \ u_a \sim N(0, \sigma_{ua}^2) \quad e_{ij} \sim N(0, \sigma_e^2) \tag{4.3}$$

Here, p c and a represent discrete period, cohort, and age groups respectively, with u_p giving the effect of being in year p, u_c giving the effect of being born in birth cohort c, and u_a giving the effect of being in age group a. This would imply a multilevel structure as shown in figure 4.1

Figure 4.1: The multilevel structure specified by equation 4.3



There are, however, a number of issues with models such as these. First, as mentioned in previous chapters, the identification problem is likely to be a problem in models like these if there are any linear effects present in the processes that generated the data. Interestingly, because age, period and cohort are treated as random effects, and so are subject to shrinkage (that is being pulled back towards a zero effect when unreliably estimated), these models will be identifiable, even when there is no grouping across APC years. This would not be the case in a fixed-classification model in which dummies are used to represent each and every age group, year, and birth year. However, this doesn't mean that the estimates that are produced by the multilevel model will be correct. Second, the assumption that the random effects u_a, u_p and u_c are independent and identically distributed is likely to be incorrect. Consecutive years are likely to be more related to each other than years that are far apart in time; people born in 1950 are likely to be more similar to those born in 1951 than to those born in 1980; and so on.

Given this, it might make more sense to model a mixture of fixed and random classifications, with the fixed part modelling continuous, long-run changes, and the random part of the model providing estimates of discrete changes net of any long-run changes. However, this needs to be done on the basis of theory and an understanding of the APC processes that are being modelled. For instance, age is a parameter that is likely to have only a continuous effect – that is, it is unlikely there is a specific effect of being, for example, age 24, but rather an underlying smooth effect of getting older across a longer age range. It will usually make more sense, therefore, to model this as a linear (or polynomial) trend in the fixed part of the model, showing a smooth change in an outcome variable through the life course. This is not necessarily the case for periods and cohorts, where a specific event (like a war or an economic recession) might lead to an effect on an outcome related to a very specific year or birth cohort. There might also be a combination of these two sorts of effects – that is, there might be a continuous effect of successive birth cohorts, as well as more discrete effects associated with specific birth cohorts born in very specific moments in history. For example, in general people born longer ago have higher rates of mortality, whilst there are also specific events (being born during the Spanish flu epidemic of 1919, or the Dutch famine of 1944) that also additionally impact individuals' mortality.

As such, we want to develop a model that can model both of the above – smooth changes over time, and isolated events with discrete effects. Not only that, but if we model smooth effects in the fixed part of a multilevel model, it will account for much of the dependency across APC units, meaning the random effects assumptions are more likely to be met in modelling discrete effects. For example, we might believe that there is a smooth trend associated with age (with no additional discrete effects associated with particular ages). We might also think that there are only discrete period effects, and both discrete and smooth cohort effects. This would result in a model along the lines of:

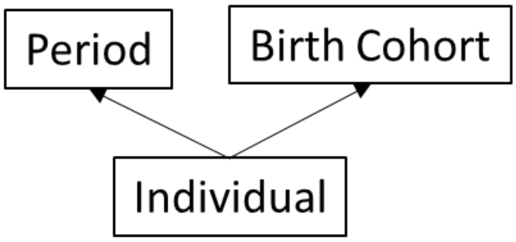
$$Y_{ij} = \beta_0 + \beta_1 Age_i + \beta_2 Age_i^2 + \beta_3 Cohort_i + \beta_4 Cohort_i^2 + u_p + u_c + e_i$$

$$u_p \sim N(0, \sigma_{up}^2), \; u_c \sim N(0, \sigma_{uc}^2), \quad e_{ij} \sim N(0, \sigma_e^2)$$

(4.4)

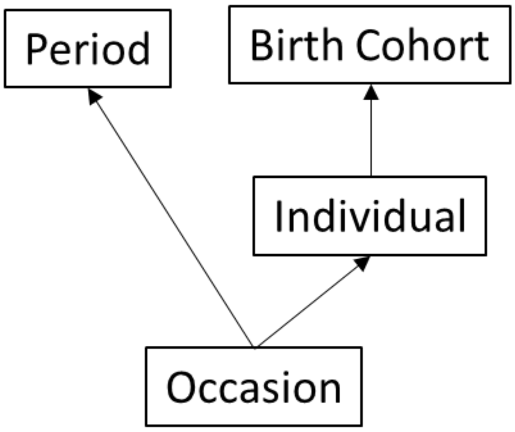
Here, we have included a polynomial effect for age (estimated by β_1 and β_2) and cohort (β_3 and β_4), and discrete random effects for period (u_p) and cohort (u_c). It would also simplify the multilevel structure implied by the model to that in figure 4.2 where Age is no longer a 'structure' but treated as a measured variable.

Figure 4.2: Multilevel structure implied by equation 4.4



One of the advantages of multilevel models is that they are highly extendable, and that applies to models like the above as well. We could add additional levels into a model like this – if we have panel data, we would usually include an additional individual-person level (see figure 4.3), and potentially further spatial levels (like neighbourhoods) as well (Bell, 2014). We could also add additional explanatory variables to the fixed part of the model which may represent measured attributes of individuals, birth cohorts and periods. The random effects are then the unexplained residual differences at each level net of the fixed effects of measured variables in the model.

Figure 4.3: Multilevel structure implied by a model with panel data, extending figure 4.2



We are, however, constrained in how many fixed classifications we can include in our model – that is, we can only include two of APC as linear effects in the model because of the identification problem (the other is constrained by the estimates of the other two to zero). This means that, in the model above, we are assuming that there is no continuous period trend – only discrete variation with no trend. As shown by O'Brien (2017), the choice of what variables we model as fixed effects will change the results that we find. Modelling two of APC as fixed effects will effectively set the trend of the third to zero. However, as we see later in this chapter, failing to model two of APC as fixed effects can lead to apparent arbitrary apportioning of effects to random effects which (a) will mean those effects are not independently and identically distributed, but more seriously (b) will produce solution to the identification problem that is based on the data's structure and groupings, rather than the true data generating process. In other words, including fewer than two of APC in the fixed part of the model will not make the assumptions being made less strong – it will just make them less visible.

4.4 The Hierarchical Age Period Cohort model

One of a number of methods to appear in the literature in the early noughties is the Hierarchical Age-Period-Cohort (HAPC) model (Yang and Land, 2006, 2013). The HAPC model is a version of the models described in the above section. It uses a specific combination of fixed continuous predictors and discrete random effects: age as a polynomial in the fixed part of the model, and period and cohort as discrete random effects (as in figure 4.2), meaning the model can be specified as in the equation above but without the cohort fixed parameters:

$$Y_{ij} = \beta_0 + \beta_1 Age_i + \beta_2 Age_i^2 \ + \ u_p + u_c + e_i$$

$$u_p \sim N(0, \sigma_{up}^2), \ u_c \sim N(0, \sigma_{uc}^2), \quad e_{ij} \sim N(0, \sigma_e^2)$$

(4.5)

The logic of the model may be apparent to the reader given the above discussion. It makes sense to think of age as a continuously changing random effect, because discrete effects of specific ages are rarely plausible, and because it is conceptually an attribute of the individual (i in the equation above). Similarly, period and cohort are indeed contexts in which an individual resides, much like neighbourhoods or schools, and other spatial contexts that are frequently modelled in this way. However, this does not mean that it solves the problems identified above, most notably the identification problem. And yet, the model has been used in a range of different social science and health disciplines as if it were a solution to the age-period-cohort identification problem (for examples, see: Dassonneville, 2013; Reither et al., 2009; Schwadel, 2010), and the authors of the method have claimed that the model does indeed solve the identification problem:

> "An HAPC framework does not incur the identification problem because the three effects are not assumed to be linear and additive at the same level of analysis" (Yang & Land, 2013:191).

> "The underidentification problem of the classical APC accounting model has been resolved by the specification of the quadratic function for the age effects" (Yang & Land, 2006:84).

If you have read the previous chapters, it is likely that you will already be somewhat sceptical of these claim, and that, with only one of APC specified in the fixed part of the model, it is likely that near-linear APC effects will be mis-apportioned. And indeed, simulation studies have shown that linear or near-linear APC trends can be incorrectly apportioned using the HAPC model (Bell and Jones, 2014a; Luo and Hodges, 2016). For instance, we were able to replicate Reither et al's (2009) study of obesity, using data generated in a quite different way from the results found by both them and us (Bell and Jones, 2014b). Follow-up commentaries have shown that, even for data that is not linearly generated, and even when there are all of age, period and cohort effects present in the processes that generated the data, the model can radically mis-apportion APC effects (Bell and Jones, 2015c, 2018).

4.5 Why do multilevel APC models produce the results that they do?

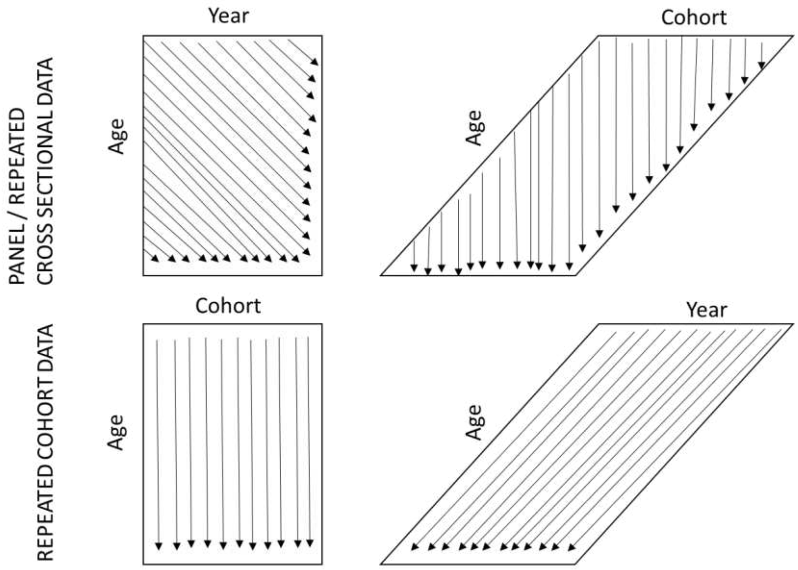
We have already discussed that the choice of fixed-part APC parameters affect the results that are likely to be produced. If we include two of APC in the fixed part of the model, the third one will automatically be set to a trend of zero, and the other two trends will adjust to accommodate this constraint. For example, if the true data generating process consists of a period trend with a gradient of 1 unit, but only age and cohort are included in the fixed part of the model, that period trend will be estimated as zero, and the cohort and age trends will be overestimated by 1 unit. That is, APC trends are tied to each other – if we constrain one to be wrong, the other two will also be wrong to the same extent to adjust the predictions of the model to be accurate.

Given the inflexibility of a model with two of APC in the fixed part of the model, it might be tempting to think that including fewer than two of APC in the fixed part of the model might be more effective. It is this that inspires the use of the HAPC model (equation 5), and could also be used to justify the use of a model along the lines of equation 3, which treats all of APC as random variables. Unfortunately, this doesn't solve the problem – all it does is make it less clear precisely what assumptions the model is making. However, the model is still making strong, but implicit, assumptions, meaning that misleading inference is still likely to occur.

So what does drive a model like the HAPC model, or the fully-random model in equation 4.3, to produce the results that it does? The answer lies in two things: the data collection process and resultant structure of the data being used, and the ways in which multilevel model estimators aim to maximise model fit.

The vast majority of APC analysis uses data that is collected in waves. That is, data is collected for an approximately representative sample, across all age groups, on a number of occasions. The result of this is that we could plot an age-by-period diagram of our data, and that diagram would be rectangular. It might be easy to think that this is the only way data can be collected, but in reality a number of other structures are, or could be used. For instance, cohort studies collect a sample of people born in a given year, and follow them through the rest of their lives. It isn't possible to study APC effects with a single cohort (because age and period are exactly collinear, and cohort is non-varying), but we can have variation in all three by combining multiple cohort studies together. This produces a dataset that, when arranged in an age-by-cohort array is rectangular but will be in the form of a parallelogram when arranged age-by-period (see figure 4.1).

Figure 4.4: data structures associated with panel / repeated cross-sectional data and repeated cohorts, when arranged age-by-year and arranged age-by-cohort. The arrows represent cohorts progressing through the life course. It should be noted the large number of cohort groups that exist with panel / repeated cross-sectional data compared to the number of years.



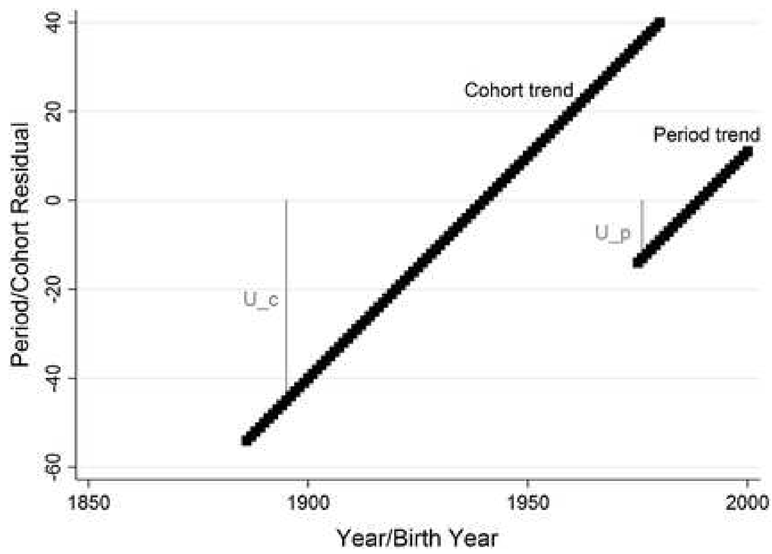
These differences in data structures may seem like they should be unimportant for the inference that models would produce, but it turns out that they have important characteristics that can influence the results produced by the models. Most notably, with repeated cross-sectional and panel data there will always be a wider range of cohorts than there will be years of measurement. For example, consider a panel dataset that runs every year between 1991 and 2008 (such as the British Household Panel Survey) with individuals aged between 18 and 70. This will have year groups spanning 18 years, but cohorts spanning from 1921 (those 70 in 1991) to 1990 (those who are 18 in 2008) – a 69 year range for birth cohorts.

Now, imagine there is a linear effect of period, of strength 1 (such that for every year that passes, we would expect an increase in the dependent variable of 1). If that were to be modelled in a multilevel model with period residuals, the size of those residuals would depend on the range of the periods – a wider range would lead to bigger residuals to account for bigger differences. Similarly, if there were a linear effect of cohort, of strength 1, that were modelled by birth cohort residuals in a multilevel model, the size of these residuals would also depend on the residuals for that data frame. This is visualised in figure 4.5 (from Bell and Jones, 2018, p787).

Given the identification problem, we know that statistical models will reapportion APC linear effects in such a way to maximise model fit. In this case, high values of residuals would imply a worse model fit, since these residuals count as unexplained variance in the random part of the model. If we consider the HAPC model (equation 4.5), we would expect the model to apportion any trend to period rather than cohort, because the 'cost' of the residuals in terms of model fit is lower. If there were a linear cohort trend of magnitude 1, it would make sense for the model to reapportion this; given Cohort = Period - Age, it could reapportion this as a Period trend of magnitude 1 and an Age trend of magnitude -1. The Age trend has no additional cost to the model's fit, since it is absorbed in the fixed part of the model, whilst the period trend is low cost in comparison to the equivalent cohort trend of the same magnitude. As such, in that situation the model is likely to find the wrong answer – it is in effect assuming there is no cohort effect, and modelling any change over time as a period effect.

Note that, if using repeated cohort data, we would expect the results to be reversed, because the range of years would be much greater than the range of cohorts. For other multilevel models, for example that in equation 4.3, the answer is more complex, since an age trend would also need to be modelled in the random part of the model, making the trend costly in terms of model fit. The results are likely to depend on the range of the age variable, but O'Brien, (2017) finds that, at least in some cases, such a model sets the cohort trend to close to zero, so would be likely to produce results similar to that produced by the HAPC model. The result is also likely to vary as a result of grouping periods and/or cohorts unevenly (Bell and Jones, 2018). The key point, however, is that these models are not apportioning affects based on actual APC processes. They are being apportioned based on the structure of the data being used.

Figure 4.5: Estimated cohort and period residuals associated with a cohort and a period linear effect of magnitude 1. Reproduced from Bell and Jones (2018, p787).



4.6 What Multilevel APC models should researchers use?

This should make for sober reading for anyone considering using multilevel models as an automatic way of getting around the APC identification problem. Whilst these models do not always make any obvious explicit assumptions with regard to APC trends, they do always make implicit assumptions that are as strong as those made in other APC models, such as those outlined in the previous chapters.

However, this is not to say that models such as these do not have value. The ability to estimate linear (and other polynomial) long-run effects in the fixed part of the model, as well as discrete random changes in the random part, is really powerful and allows for quite nuanced analysis of how APC effects operate. However, this needs to be done with an awareness that the identification problem cannot be solved, and that we need to make explicit assumptions and justify them with theory.

Given this, a model such as that in equation 4.4 might be a sensible one, *if* we are willing to assume that period trends have no continuous trends (that is, whilst there is random fluctuations from one year to the next, perhaps because of economic shocks, there are no long-run changes that are a result of period trends). This is often a reasonable assumption to make, where there are theoretical reasons why we would expect change over time to be a result of successive cohort replacement rather than years passing. This is the approach taken by Bell (2014) in his analysis of APC effects on mental health, and discussed further below (see also Delaruelle et al., 2015 for a similar approach). Alternatively, if we were able to assume the opposite (that periods drive change over time), we would want to include period in the fixed part of the model, and not cohort. Theory, plausibility and research questions, accompanied by a sceptical openness, are needed to derive an appropriate model specifications and analysis.

We do not have to assume a particular parameter has a linear slope of zero. We could, instead, constrain one of the fixed part parameters in the multilevel model to a particular value, potentially in a Bayesian framework by applying a strongly informative prior to that parameter (Bell and Jones, 2015b). This could be useful if, for instance, you have a strong idea of what the age trend of a variable should look like; for medical outcomes, we might have medical reasons for being able to assume an age trend, constrain the age parameter to that value, and then estimate period and cohort trends assuming the constraint on the age parameter is reasonable (for an example of this, see Van der Bracht and Van de Putte, 2014). It might be that we cannot know an exact constraint on a parameter, but might be able to impose some bounds on a parameter: for example, that any age effect will be positive (greater than zero). A combination of such constraints might lead to boundaries within which the true linear APC effects must lie (see chapter 7, also Fosse and Winship, 2016). Finally, one could use a range of constraints, compare the different combinations of APC trends that those different constraints produce, and come up with an argument for which combination is the most plausible. The important point is that, in each case, the constraint that is made and the process by which it is assumed is made explicitly so that a reader can judge the validity, or otherwise, of that assumption.

Finally, in some situations we are not interested in long run change, and only interested in discrete shocks. In this situation we can use a multilevel model like that in equation 4.4, but ignore the fixed part estimates entirely, only interpreting the random part estimates. We can use a method like this to find, for instance, period and cohort effects related to specific years of measurement or years of birth. However, we need to be careful not to mis-interpret these results by assuming that the linear effects do not exist – for instance, a non-linear trend in random effects might mean something very substantively different when a linear trend is included. However, sharp discrete changes in those random effects can often be interpreted, as we will see in the second example below.

We now present two examples that take some of the approaches outlined above: first a study of mental health that assumes a zero period effect, and second a study of mortality that ignores the estimated fixed part estimates entirely, and considers only discrete changes around those trends, estimated in the random part of the model. In each case, R code to guide readers in how to implement such multilevel models is provided online.

4.7 Example 1: Mental Health in the UK, 1991-2008

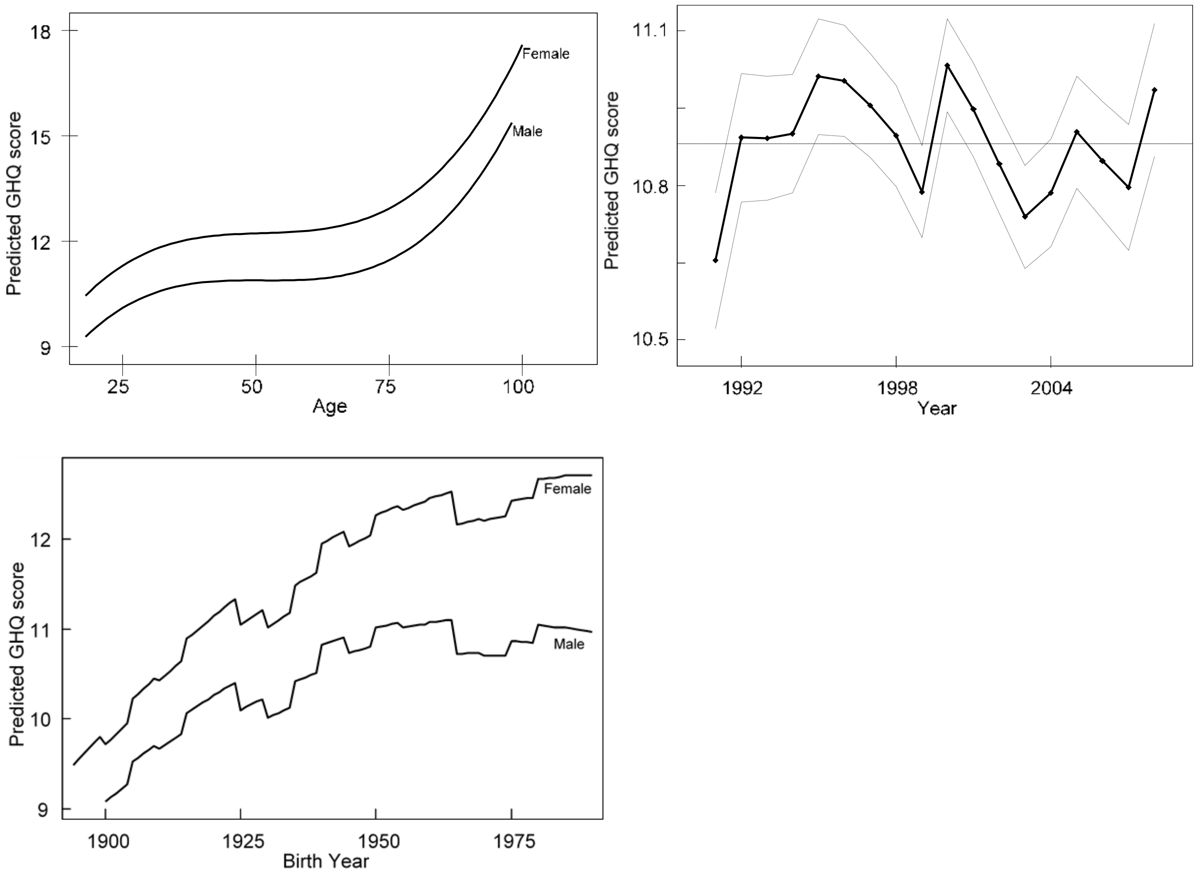
How does an individual's mental health change over their life course? There is some literature (e.g. Blanchflower and Oswald, 2008) which suggests that it takes a U-shaped pattern, with mental health worsening to midlife, and then improving into old age. However, this is often based on cross-sectional data (where the life course effect is confounded with cohort, and even when longer-term data is used, the problem of APC identification problem.

Bell (2014) and Bell and Jones (2015a) used a multilevel model in order to attempt to find whether this U-shape really exists when cohorts are controlled, using data from the British Household Panel Survey (BHPS), which runs from 1991-2008 in the UK. They used a model like that in equation 4.4, where Y is the score in the General Health Questionnaire (GHQ - Goldberg and Williams, 1988), a measure of mental health measured from 0 to 36, where high scores indicate worse mental health. The model used included age and cohort polynomials in the fixed part of the model, and so assumed that there were no period trends, on the basis that “there is no reason to expect a continuous period trend across periods affecting all ages. Cohorts, through the nature of individuals’ upbringings, more plausibly explains how changes in mental health could occur over time” (Bell, 2014, p23). Because the data is panel, the structure was that depicted in figure 4.3, with additional household and local authority levels.

The results showed that, in fact, when the assumption of no period effects is made, there is no U-shaped relationship between age and mental health – in contrast, mental health worsens throughout the life course (see figure 4.6a). Whilst the assumption of no period effects is arguably contentious, it is no more contentious than any other assumption that would need to be made to meaningfully identify an APC model.

The model was further able to identify discrete period effects (figure 4.6b), and both discrete and continuous cohort effects (combined in figure 4.6c) that suggested a general trend of worsening mental health with successive cohorts.

Figure 4.6: Predicted GHQ scores for different values of age, period and cohort. (a) fixed part continuous age effect, (b) fixed part continuous cohort effect combined with discrete random-part cohort effect, (c) discrete random-part period effect. Taken from Bell and Jones (2015a, p208-210).



4.8 Example 2: Mortality in the UK through the twentieth century

There are many things that have led to changes in mortality over the last 100+ years. First, mortality has reduced, as a result of medical and public health advances, implying the presence of continuous period effects, or cohort effects, or both. There is also, of course, changes in mortality as individuals age – that is, the likelihood of death increases as an individual ages. However, these trends are likely to be subject to the identification problem, and, in the absence of good theoretical reasons to constrain one of age, period and cohort, it would be impossible to find such APC effects robustly.

However, as well as this more long-run changes, we also expect to find discrete event-based period and cohort effects related to events in particular years. In particular, we would expect wars and disease epidemics to have effects both on those who lived in those times (a period effect), and also those who were born or brought up in those times and carried on through their lives (a cohort effect). Because these are not linear or near-linear trends, they can be identified, potentially through a multilevel model.

This was the approach undertaken by Jones et al. (2018) alongside some more graphical techniques (see chapter 5). They used a Poisson multilevel model to model data from the Human Mortality Database (1922-2016), that has data on mortality across all years and all age groups (University of California and Max Plank Institute for Demographic Research). The model is thus specified as follows:

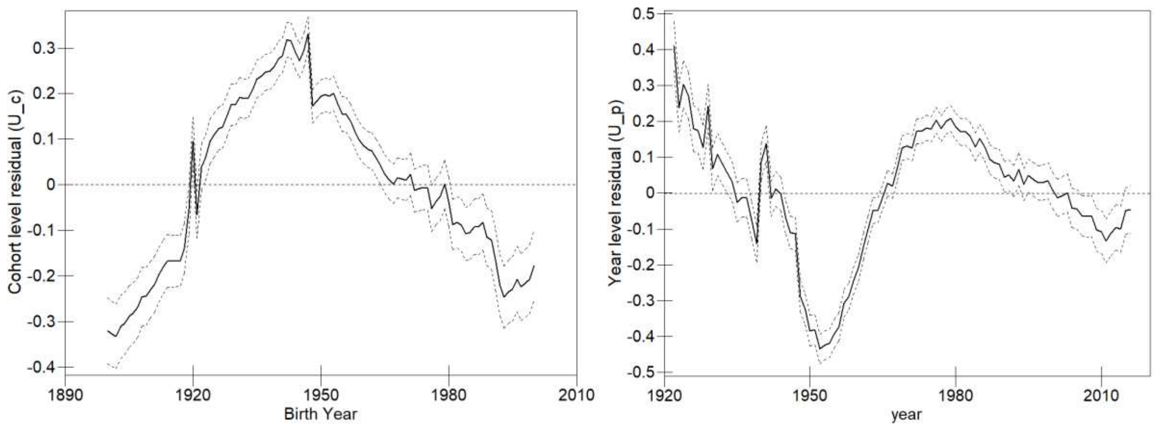
$$Log_e(Deaths_i) = Log_e(Exp_i) + \beta_0 + \beta_1 Age_i + \beta_2 Age_i^2 + \beta_3 Period_i + u_p + u_c \tag{4.6}$$

The outcome is the natural log of the number of deaths, with an offset being the natural log of the expected number of deaths (that is, the number of deaths, given the population size, we would expect if deaths were distributed evenly. The inclusion of the offset means that, instead of modelling the outcome of (log) number of deaths, the outcome effectively becomes the (log) mortality rate for the given age-year cell. Age and period are specified in the fixed part of the model, but any two of APC could be used, given the aim is simply to soak up the long run APC effects. The result is that the period and cohort residuals should be free of linear trends, allowing us to model non-linear discrete changes more appropriately. The level 1 residuals are assumed to follow a Poisson distribution, whereby the variance is equal to the mean, and the period and cohort residual differentials are assumed to be Normally distributed.

These period and cohort residuals are shown below in figure 4.7 for men. It can be seen that there is some non-linearity in these figures – however this should not be over-interpreted – this is a long-run but non-linear change not fully captured by the fixed-part age and period parameters, and it's meaning may well be different when combined with any linear trends that have been controlled out in the fixed part of the model. What we can consider, however, are any non-linear discrete changes that occur in this data. In this regard we can see some very clear cohort effects that appear to have produced an increased mortality for people born in 1919, and decreased mortality for people born in 1948. In terms of period effects, again there are potentially misleading non-linearities, but there also appear to be a significant increase in mortality associated with the early 1940s.

There are some clear reasons why such effects might have occurred. 1919 corresponds with the outbreak of the Spanish flu, which as well as causing high levels of mortality at the time, is also known to have had more long-lasting damage to young people at the time, including children born with pre-natal exposure to the disease (Almond, 2006). This corroborates the idea that a higher mortality risk associated with the Spanish flu followed individuals in that birth cohort through the rest of their lives. The opposite effect, associated with the year 1948, corresponds to the formation of the NHS and other post-war improvements in public health. It seems that those measures had a positive effect on those born at the time, suggesting that improvements in prenatal and antenatal healthcare were particularly valuable, and stayed with those individuals through their continuing lives. For period effects, there appears to be an effect of the second world war – unsurprising given the large number of lives lost at that time.

Figure 4.7: Cohort and period level residuals conditional on continuous trends estimated in the fixed part of the model (u_c and u_p in equation 4.6), on male mortality.



4.9 Conclusion

Multilevel models present a useful tool for considering age period and cohort effects. This is because data that can be used for APC analysis is inherently multilevel, because longitudinal data always has some kind of structure. However, multilevel models do not provide a solution to the identification problem – rather a structure around which the identification problem can be considered, and appropriate and strong assumptions made in order to make the models produce robust results if those assumptions are justified.

Whilst others have suggested multilevel models present a potential automatic solution to the identification problem, this is not the case: such multilevel models may be identifiable, but they tend to apportion APC near-linear effects on the basis of the structure the data being analysed, rather than the true effects present in the processes that generated the data. As such, multilevel models need to make assumptions that are justified by theory and made explicit, or ensure that only discrete non-linearities are interpreted.

In sum, multilevel models present opportunities for APC analysis, but are not a magic bullet – they are not a solution to the identification problem because no such solution can exist.

4.10 References

Almond D (2006) Is the 1918 Influenza Pandemic Over? Long-Term Effects of *In Utero* Influenza Exposure in the Post-1940 U.S. Population. *Journal of Political Economy* 114(4). The University of Chicago Press : 672–712. DOI: 10.1086/507154.

Bell A (2014) Life course and cohort trajectories of mental health in the UK, 1991-2008: a multilevel age-period-cohort analysis. *Social Science & Medicine* 120: 21–30.

Bell A (2019) Cross-sectional and longitudinal studies. In: Morin J, Olson C, and Atikcan E (eds) *Key Concepts in Research Methods*. Routledge.

Bell A and Jones K (2014a) Another 'futile quest'? A simulation study of Yang and Land's Hierarchical Age-Period-Cohort model. *Demographic Research* 30: 333–360.

Bell A and Jones K (2014b) 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 and Medicine* 101: 176–180.

Bell A and Jones K (2015a) Age, period and cohort processes in longitudinal and lifecourse analysis: a multilevel perspective. In: Burton-Jeangros C, Cullati S, Sacker A, et al. (eds) *A life course perspective on health trajectories and transitions.* Springer, pp. 197–213. DOI: 10.1007/978-3-319-20484-0_10.

Bell A and Jones K (2015b) Bayesian Informative Priors with Yang and Land's Hierarchical Age-Period-Cohort model. *Quality and Quantity* 49(1): 255–266.

Bell A and Jones K (2015c) Should age-period-cohort analysts accept innovation without scrutiny? A response to Reither, Masters, Yang, Powers, Zheng, and Land. *Social Science and Medicine* 128: 331–333.

Bell A and Jones K (2018) The hierarchical age–period–cohort model: Why does it find the results that it finds? *Quality and Quantity* 52(2): 783–799.

Blanchflower DG and Oswald AJ (2008) Is well-being U-shaped over the life cycle? *Social Science & Medicine* 66(8): 1733–1749. DOI: 10.1016/j.socscimed.2008.01.030.

Dassonneville R (2013) Questioning Generational Replacement. An Age, Period and Cohort Analysis of Electoral Volatility in the Netherlands, 1971-2010. *Electoral Studies* 32(1): 37–47.

Delaruelle K, Buffel V and Bracke P (2015) Educational expansion and the education gradient in health: A hierarchical age-period-cohort analysis. *Social Science & Medicine* 145: 79–88. DOI: 10.1016/j.socscimed.2015.09.040.

Fitzmaurice GM, Laird NM and Ware JM (2011) *Applied Longitudinal Analysis*. 2nd ed. Hoboken, NJ: Wiley.

Fosse E and Winship C (2017) Bounding Analyses of Age-Period-Cohort Models. Working Paper. Available at: https://q-aps.princeton.edu/sites/default/files/q-aps/files/apcanalysis.pdf (accessed 31 October 2018).

Goldberg D and Williams P (1988) *A User's Guide to the General Health Questionnaire*. Windsor: NFER-NELSON.

Jones K (1991) Specifying and estimating multi-level models for geographical research. *Transactions of the Institute of British Geographers* NS 16(2): 148–159.

Jones PM, Minton J and Bell A (2018) Period and cohort changes in mortality risk over the twentieth century in the UK: an exploratory analysis. *OSF Preprints*. DOI: 10.31219/OSF.IO/4F7JR.

Leyland AH and Goldstein H (2001) *Multilevel modelling of health statistics*. Wiley.

Luo L and Hodges JS (2016) Block Constraints in Age-Period-Cohort Models with Unequal-width Intervals. *Sociological Methods & Research* 45(4): 700–726.

O'Brien RM (2017) Mixed models, linear dependency, and identification in age-period-cohort models. *Statistics in Medicine* 36(16): 2590–2600.

O'Connell AA and McCoach DB (2008) *Multilevel modelling of educational data*. Charlotte NC: Information Age.

Reither EN, Hauser RM and Yang Y (2009) Do birth cohorts matter? Age-period-cohort analyses of the obesity epidemic in the United States. *Social Science & Medicine* 69(10): 1439–1448.

Schwadel P (2010) Age, Period, and Cohort Effects on US Religious Service Attendance: The Declining Impact of Sex, Southern Residence, and Catholic Affiliation. *Sociology of Religion* 71(1): 2–24.

University of California B and Max Plank Institute for Demographic Research (n.d.) Human Mortality Database. Available at: www.mortality.org (accessed 2 November 2018).

Van der Bracht K and Van de Putte B (2014) Homonegativity among first and second generation migrants in Europe: The interplay of time trends, origin, destination and religion. *Social Science Research* 48. Academic Press: 108–120. DOI: 10.1016/J.SSRESEARCH.2014.05.011.

Yang Y and Land KC (2006) A mixed models 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* 36: 75–97.

Yang Y and Land KC (2013) *Age-Period-Cohort Analysis: New models, methods, and empirical applications*. Boca Raton, FL: CRC Press.