Introduction: identifying age-period-cohort accelerations in a regression framework with individual-level data

Zoë Fannon

August 2020

1 Introduction

This chapter provides the background to the contributions made in the three substantive chapters of the thesis. I explain the relevance of age, period, and cohort effects to economics and outline the analytical framework which will be used throughout the thesis.

I motivate the thesis in §2 with a discussion of existing papers in economics which use age, period, or cohort as explanatory variables. Examples are particularly prevalent in labour economics, health economics, and industrial organisation. I suggest ways in which accelerations in age, period, and cohort could be used to address the research questions of these applications.

The bulk of this chapter is devoted to an outline of the approach used in this thesis to identify age, period, and cohort (APC) accelerations from regressions, given in §3. I discuss the general APC identification problem and the fact that accelerations are robust to this. I then introduce the framework used to identify APC accelerations from regression. This framework is based on a reparametrization of what is known as the classical APC model. This approach was developed in a series of papers working with aggregate data, starting with Kuang et al. (2008).

In the final sections of this chapter I describe the advantages of the acceleration-based regression framework as an approach to the analysis of APC effects, as compared to a number of other common approaches. Many of these approaches have the undesirable property that their estimates are sensitive to untestable identifying constraints, which is avoided by the APC acceleration framework.

2 Age, period, and cohort (APC) effects in economics

In this section, I provide a brief tour of the applied literature that uses individual-level data in which age, period, or cohort are explanatory variables. This highlights the range of potential use cases for the framework developed in this thesis.

Many of the papers cited here use what I call "constraint methods" to estimate age, period, and cohort effects. These methods may be favoured by researchers because they give estimates of slopes in age, period, and cohort, and researchers may be more accustomed to working with slopes than the accelerations used in this thesis. The problem with constraint methods is that they only obtain estimates of slopes by imposing untestable constraints on the age, period, and cohort effects. The choice of constraints impacts the resulting estimates of the slopes in age, period, and cohort. This is explored further in §5.1. The framework I develop, in terms of accelerations, does not require these constraints.

I argue that the use of these constraint methods to estimate age, period, and cohort effects is unnecessary and can lead to incorrect inference. Many of the underlying questions of interest in the applications described below could be answered using the accelerations, which can be identified without imposing such constraints. Where the question of interest can truly only be answered using slopes, it seems unwise to rely on an answer to that question which depends on an untestable constraint imposed by the researcher. Even in the best examples of the application of constraint methods, such as Lagakos et al. (2018), where the constraints are carefully selected and argued for based on the context, it is impossible to formally test the constraints. It therefore is preferable to use accelerations, which are identified without the need for constraints. In the following, I describe how accelerations could be used to address research questions in economics.

In life-cycle models, age effects are naturally the primary object of interest. Life-cycle models may be estimated from individual-level data to better understand the shape of the relationship between age and an outcome of interest, controlling for other factors. For example, Hanoch & Honig (1985) estimate age profiles of employment and earnings for white Americans using social security data, and Browning et al. (2016) use British survey data to estimate an age profile of purchases of durable goods in later life. Accelerations are very suitable for exploring age (or period or cohort) profiles, as I demonstrate in my exploration of the age and cohort profiles of obesity in §**??**. Life-cycle models may be estimated in order to evaluate structural models, as is the case in Low et al. (2010) where a constraint method is used to estimate the age profile of employment rates from the American SIPP dataset. It would be straightforward to compare accelerations in the data with accelerations implied by the structural model, rather than comparing estimates of levels which rely on an untestable assumption. Other applications involve comparing life-cycle profiles between groups, which could easily be done in terms of accelerations. For example, Fitzenberger et al. (2004) use German micro-census data to compare the male and female age profiles of labour force participation and employment, and Lagakos et al. (2018) harmonized repeated cross section wage surveys to study variation in the experience-wage profile across countries.

Several studies have investigated whether a policy change or exogenous shock creates a discontinuity in one of age, period, or cohort. Krueger & Pischke (1992) use data from the repeated cross section Current Population Survey to investigate cohort discontinuities in American labor supply due to social security reforms. Almond (2006) looks for cohort discontinuities in various outcomes which could be linked to the 1918 influenza pandemic, using individual-level US census data. Discontinuities can be readily identified using APC accelerations; an example with aggregate data is McKenzie (2006), in which discontinuities in consumption by period are detected and attributed to the Mexican peso crisis.

Age, period, and cohort often appear as controls, rather than direct objects of interest, in studies which use individual-level data. An example is the literature on the economics of health and well-being, where it is essential to control for the impact of age (see for example Künn-Nelen, 2016; Roberts et al., 2011; Van Landeghem, 2012; Dickerson et al., 2014). Functional form restrictions are often imposed on age, period, and cohort controls. The APC acceleration framework can be used to determine whether functional form restrictions on age, period, and cohort controls are appropriate. An example is given in §**??**.

Exploratory analyses of the importance of age, period, or cohort effects for some outcome of interest are seen in many sub-disciplines of economics. They are common in labour economics (in addition to the studies cited in the disussion of life-cycle models, see Méndez & Sepúlveda, 2012; Meghir & Whitehouse, 1996) and in the study of consumption and savings (in addition to the studies cited in the disussion of life-cycle models, see Kapteyn et al., 2005; Bíró, 2017; Attanasio, 1998). There is a literature on the relative importance of firm vintage, age, and year of observation for outcomes such as firm productivity (Jensen et al., 2001; Fukuda, 2013). In energy economics, Bardazzi & Pazienza (2018) seek to separate the contributions of age and cohort to energy demand.

A particularly concerning phenomenon is the existence of studies where one of age, period, or cohort is the main explanatory variable of interest, but the author seems unaware of the risk of confounding with the effects of the other two. An example is Rosenthal (2014), published in the *American Economic Review*. This paper examines the relationship between the age of a house and the income of the occupier, to examine whether the "filtering" process - where houses are passed down the income distribution as they age - provides a sufficient supply of housing for low-income families. There is no provision for cohort effects, although the year in which a house was built may also be related to the income of its occupiers.

The APC identification problem is not unique to economics; it appears in in applied work across the social sciences. Yang & Land (2013) and O'Brien (2015) describe examples in criminology, epidemiology, and sociology. The use of constraint methods is also common in these literatures. I discuss the constraint methods used in economics and other social sciences in more detail in §5.

The above discussion shows that age, period, and cohort are ubiquitous as explanatory variables. The use of regression-based constraint methods in applications with individual-level data is widespread, despite the general knowledge that accelerations can be identified and the ease with which accelerations could be used to answer the questions of interest. One of the factors preventing the wider use of accelerations may be the lack of a straightforward, regression-based framework from which to estimate them. It is also possible that there is uncertainty around the use and interpretation of accelerations. This thesis attempts to fill both of these gaps, by developing a framework and software for estimating APC accelerations from individual-data regressions, and by providing applications and examples which show how accelerations may be used.

3 The reparametrized classical APC model: A framework to estimate APC accelerations from aggregate data regressions

The framework I develop to estimate APC accelerations from individual-level data regressions is based on an existing framework for estimation of APC accelerations from aggregate data regressions (Kuang et al., 2008; Nielsen, 2015). In this section, I explain that aggregate data framework. At the heart of the framework is a reparametrization of a model known as the classical APC model. The classical APC model is not identified; the reparametrization is identified, and permits isolation of the APC accelerations as regression parameters.

The remainder of this section is as follows. In §3.1, I provide a very brief overview of the aggregate data framework, which is explained in detail in the rest of the section. In §3.2, I explain the structure of the aggregate data, introduce the classical APC model, and explain why the classical APC model is not identified. In §3.3 I first outline the key ideas of the reparametrization approach to the classical APC model, and then provide a detailed derivation of the reparametrization. In §3.4 I explain the four use cases of the APC acceleration parameters as estimated from the reparametrized model.

3.1 Overview

The data for which the framework is developed is an array of aggregate outcome outcomes indexed by two of the three of age, period, and cohort. These are denoted by a, p, c respectively, and related by $p = a + c - 1$; this is explained in §3.2.

The analysis of this data begins with a model called the classical APC model, which is defined in terms of age, period, and cohort fixed effects:

$$\begin{aligned} \mu_{ac} =& \delta + \alpha_a + \beta_p + \gamma_c \\ =& d'_{ac}\theta \end{aligned} \tag{1}$$

Here μ_{ac} is the linear predictor for the outcome indexed by a, c; δ is a general intercept; and α_a, β_p, and γ_c are age, period, and cohort fixed effects. These are collected in the parameter vector θ which is associated with a design vector d_{ac}. The setting up of this classical APC model is detailed in §3.2.1.

The parameter vector of the classical APC model,

$$\theta = \{\delta, \alpha_1, \ldots, \alpha_A, \beta_1, \ldots, \beta_P, \gamma_1, \ldots, \gamma_C\} \tag{2}$$

is not identified. It has dimension $A + P + C + 1$, which is greater than can be identified from the data. This identification problem is well-known, see §3.2.2.

To address the identification problem, the model in (1) is reparametrized in terms of a parameter vector ξ, which is of lower dimension than θ and is therefore just-identified. The reparametrized model is

$$\mu_{ac} = x'_{ac}\xi, \tag{3}$$

for x_{ac} the design vector associated with ξ. The parameter vector ξ is unique with respect to the linear predictors μ_{ac}, and is invariant to the set of θ consistent with μ_{ac}. The details of the reparametrization are given in §3.3.

The APC accelerations that are the primary object of interest appear as elements of the new parameter vector ξ. The parameter vector is defined

$$\xi = \{v_o, v_a, v_c, \Delta^2\alpha_3, \ldots, \Delta^2\alpha_A, \Delta^2\beta_3, \ldots, \Delta^2\beta_P, \Delta^2\gamma_3, \ldots, \Delta^2\gamma_C\}. \tag{4}$$

Here age acceleration parameters are denoted $\Delta^2\alpha_a$, while period and cohort acceleration parameters are denoted $\Delta_2\beta_p$ and $\Delta^2\gamma_c$ respectively. There is also a linear plane defined by v_o, v_a, v_c. These parameters are explained in detail in §3.3.4.

There are at least four use cases of the accelerations, which are discussed in §3.4. First, the accelerations can be used independently to identify discontinuities due to policy changes or exogenous shocks; see §3.4.1. Second, accelerations can be cumulated and detrended to construct a visual representation of the non-linear part of the relationship between an outcome of interest and age, period, or cohort;

see §3.4.2. Third, restrictions on the model in terms of ξ can be imposed and tested using standard hypothesis testing frameworks; such restrictions may be implied by economic theories or simply desired for the sake of parsimony, see §3.4.3. Finally, the accelerations and other elements of ξ can be used in forecasting, see §3.4.4.

3.2 The classical APC model

The APC acceleration framework has its roots in the classical APC model. The classical APC model includes an indicator for each age, period, and cohort appearing in the dataset. This model is agnostic about functional form. It is well-known that the classical APC model is not identified due to the relationship $period = cohort + age - 1$[1], and there is an extensive literature addressing this lack of identification. The APC acceleration framework for aggregate data developed in Kuang et al. (2008) and Nielsen (2015) inherits features from this literature, as will be seen in this section. Other branches of the literature which do not make use of accelerations will be examined in §5.

3.2.1 Setting up the classical APC model

The classical APC model includes an indicator for each age, period, and cohort appearing in the dataset. Therefore, the first step in setting up the model is to determine the set of required indicators from the dataset.

Any aggregate dataset used for age-period-cohort analysis should be aggregated at the level of age-cohort combinations. An example is a dataset recording the unemployment rate at a given age for each of several cohorts, as seen in Table 1. This dataset is shown as an age-cohort array, where each row is an age and each column is a cohort. Each cell is referred to as an age-cohort cell.

For an aggregate dataset represented as an age-cohort array, the classical model will include an indicator for each cohort (column), $c = 1 \ldots C$, and an indicator for each age (row), $a = 1 \ldots A$. In the example dataset in Table 1, $A = 5$ and $C = 5$. Indicators for period also need to be included. In this age-cohort array, the periods are given by the diagonals. In period 1 (1960, in the example), the first cohort is at the first age. In period 2, the first cohort is at the second age and the second cohort is at the first age. This continues for all periods, $p = 1 \ldots P$. The relationship between the age, period, and cohort indices is thus given by the formula $p = a + c - 1$.

[1]Why -1? This relation holds if age, period, and cohort are indexed such that cohort 1 is aged 1 in period 1. This indexing, referred to as time stamp accounting, is preferred because it simplifies later calculations. However, human ages are counted differently: babies are "zero years old" for all of their first year, only turning one in the beginning of their second year, yielding a relation of $period = age + cohort$. This is referred to as calendar accounting.

Table 1: US unemployment rate in an age-cohort array

age, cohort	1936-1940	1941-1945	1946-1950	1951-1955	1956-1960
20-24	0.076	0.060	0.082	0.136	0.115
25-29	0.038	0.046	0.086	0.080	0.075
30-34	0.037	0.067	0.057	0.064	0.051
35-39	0.060	0.049	0.054	0.044	0.046
40-44	0.042	0.049	0.039	0.040	0.029

Data from the OECD online database, copied from Fannon & Nielsen (2019)
Example: Top-left cell records unemployment in 1960 for the cohort 1936-1940.
In 1960 those born in 1936 were 24 and those born in 1940 were 20.

The linear predictor of the classical APC model for a dataset of this form is

$$\begin{aligned}\mu_{ac} = \delta &+ \alpha_1 \mathbb{1}(a=1) + \cdots + \alpha_A \mathbb{1}(a=A)\\ &+ \beta_1 \mathbb{1}(p=1) + \cdots + \beta_P \mathbb{1}(p=P)\\ &+ \gamma_1 \mathbb{1}(c=1) + \cdots + \gamma_C \mathbb{1}(c=C).\end{aligned} \tag{5}$$

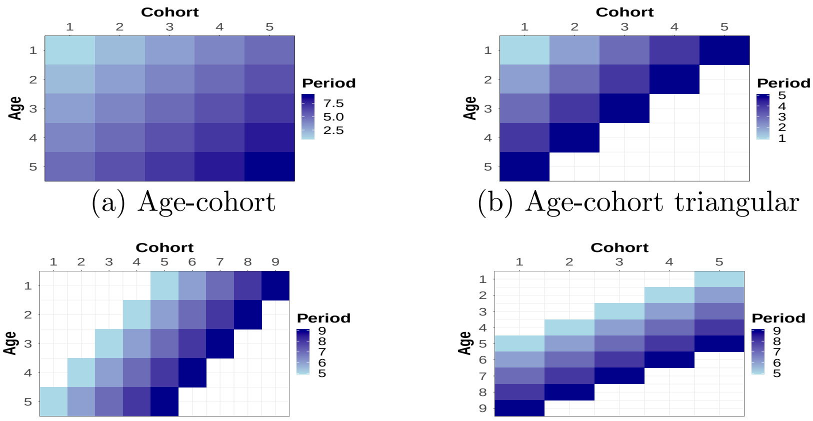
The left-hand side term, μ_{ac} is the value of the linear predictor implied by the classical APC model at the age-cohort cell $\{a, c\}$. This linear predictor is related to the outcome value y_{ac} at cell $\{a, c\}$ via a statistical model. For example, a linear model that can be estimated by OLS is

$$y_{ac} = \mu_{ac} + \varepsilon_{ac}, \tag{6}$$

for ε_{ac} a well-behaved noise term. On the right-hand side of equation (5) there are indicators for each age, period, and cohort appearing in the data. For example, $\mathbb{1}(a = 1)$ is an indicator which takes the value 1 if the condition in brackets is satisfied for cell $\{a, c\}$ and 0 otherwise. Greek letters represent parameters: δ is a generalised intercept, α_a is the coefficient on the indicator for being at age a, β_p is the coefficient on the indicator for being in period p, and γ_c is the coefficient on the indicator for being a member of cohort c. This model is not identified, as will be explained in §3.2.2.

The dataset described above has an "age-cohort" structure, because it records values at given ages for a set of cohorts, but other structures exist. Three other structures are of interest. The first structure is age-cohort triangular, which is half of an age-cohort dataset (later periods have not been observed yet). The second structure is period-cohort, in which a set of cohorts are recorded for a fixed number of periods. The third structure is age-period, in which observations are recorded over a fixed range of ages for a fixed number of periods (a labour force survey is a good example). Each of these datasets can be represented as an portion of an age-cohort array, as seen in Figure 1.

Figure 1: Data structures as age-cohort arrays



(a) Age-cohort (b) Age-cohort triangular

(c) Age-period (d) period-cohort

For these three alternative structures, the classical APC model is set up in a very similar way as it was for data with an age-cohort structure. The only difference is in the period indexation. I now present a unified indexation which can account for all four data structures. Consider the age-period data in Figure 1c. The first periods in the age-cohort array - those periods in which the first cohorts passed through the first ages - are not recorded in the dataset. Therefore the period index starts not at 1, but at $L+1$, where L is the number of periods in the age-cohort array not observed in the dataset. The general form of the classical APC model accounts for this by allowing for L in the indexing of period parameters:

$$\begin{aligned}\mu_{ac} = \delta &+ \alpha_1 \mathbb{1}(a=1) + \cdots + \alpha_A \mathbb{1}(a=A) \\ &+ \beta_{L+1}\mathbb{1}(p=L+1) + \cdots + \beta_{L+P}\mathbb{1}(p=L+P) \\ &+ \gamma_1 \mathbb{1}(c=1) + \cdots + \gamma_C \mathbb{1}(c=C). \end{aligned} \tag{7}$$

This encompasses the model for age-cohort data in equation (5), which is obtained by setting $L=0$.

The parameters of the classical model are collectively referred to as θ, which is of dimension $q = 1 + A + P + C$:

$$\theta = \{\delta, \alpha_1, \ldots, \alpha_A, \beta_{L+1}, \ldots, \beta_{L+P}, \gamma_1, \ldots, \gamma_C\}. \tag{8}$$

Two final considerations should be noted before we address the question of identification in the classical APC model. First, a feature of the classical APC model is that it is linear in age, period, and cohort; that is, it assumes no interaction effects between age, period, and cohort. Provision for testing this assumption, with repeated cross section data, is made in §??. It is also possible to test this assumption in aggregate data. Second, it is a feature of all the datasets presented above that they are contiguous, i.e. there are no empty age-cohort cells within the boundaries of the portion of the age-cohort array occupied by the data. Throughout this thesis, I assume all datasets worked with are contiguous. Further research is necessary to determine what violations of contiguity can be permitted without compromising identification.

3.2.2 The classical APC model is not identified

The full set of parameters θ in the classical APC model (7) cannot be identified because the model is overparametrized due to the relationship $p = a + c - 1$.

To explain the identification problem it is useful to define a shorthand for model (7). Let model (7) be written as

$$\mu_{ac} = \delta + \alpha_a + \beta_p + \gamma_c. \tag{9}$$

Only those parameters associated with indicator variables which equal 1 in age-cohort cell $\{a, c\}$ appear.

The identification problem can be characterized by a group of transformations of model (9), defined by Carstensen (2007) as follows. For any constants $\{\mathfrak{a}, \mathfrak{b}, \mathfrak{c}, \mathfrak{d}\} \in \mathbb{R}$ the predictor μ_{ac} satisfies

$$\begin{aligned} \mu_{ac} = \{\alpha_a + \mathfrak{a} + (a-1)\mathfrak{d}\} + \{\beta_p + \mathfrak{b} - (p-1)\mathfrak{d}\} \\ + \{\gamma_c + \mathfrak{c} + (c-1)\mathfrak{d}\} + \{\delta - \mathfrak{a} - \mathfrak{b} - \mathfrak{c} - \mathfrak{d}\}. \end{aligned} \tag{10}$$

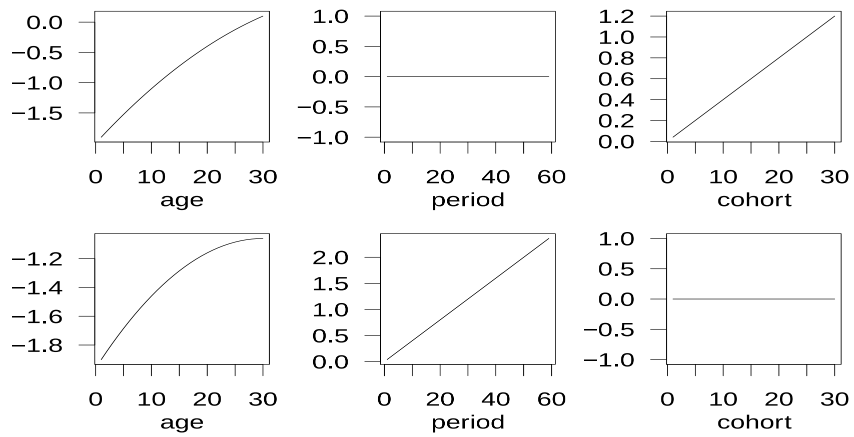
The constants $\{\mathfrak{a}, \mathfrak{b}, \mathfrak{c}, \mathfrak{d}\}$ cancel on the right hand side of (10), so that μ_{ac} does not depend on $\{\mathfrak{a}, \mathfrak{b}, \mathfrak{c}, \mathfrak{d}\}$. We say that μ_{ac} is invariant with respect to the transformations in (10). The individual effects α_a, β_p, γ_c, are not invariant to the transformation; they are only identified up to linear trends. For instance, the age effect α_a is observationally equivalent to $\alpha_a + \mathfrak{a} + (a-1)\mathfrak{d}$ for any $\{\mathfrak{a}, \mathfrak{d}\}$.

The implication of the identification problem is that the fit of model (9) with parameter vector θ is equal to that of model (9) with parameter vector θ^*, defined

$$\begin{aligned} \theta^* &= \{\delta^*, \alpha_1^*, \ldots, \alpha_A^*, \beta_{L+1}^*, \ldots, \beta_{L+P}^*, \gamma_1^*, \ldots, \gamma_C^*\} \\ \alpha_a^* &= \alpha_a + \mathfrak{a} + \mathfrak{d} \times a \\ \beta_p^* &= \beta_p + \mathfrak{b} - \mathfrak{d} \times p \\ \gamma_c^* &= \gamma_c + \mathfrak{c} + \mathfrak{d} \times c \\ \delta^* &= \delta - \mathfrak{a} - \mathfrak{b} - \mathfrak{c} - \mathfrak{d}, \end{aligned} \tag{11}$$

As the fit of θ and θ^* are equal there is no unique θ which solves the model, i.e. the model is not identified. The following example is given in Fannon & Nielsen

Figure 2: APC effects of equal-likelihood classical APC model parameters



(2019): let $\mu_{ac} = -2 + 0.1a - 0.001a^2 + 0.04c$ as in Bell & Jones (2014). This could arise from θ defined by $\alpha_a = -2 + 0.1a - 0.001a^2$, $\gamma_c = 0.04c$, $\beta_p = \delta = 0$; or equally from θ^* defined by $\alpha^*_a = -1.96 + 0.06a - 0.001a^2$, $\beta^*_p = 0.04p$, $\gamma^*_c = \delta^* = 0$. In this example, θ^* arises from θ when choosing $\mathfrak{a} = 0$, $\mathfrak{b} = -\mathfrak{c} = -\mathfrak{d} = 0.04$. Thus, the vector θ is not identified.

The severity of the identification problem is clear from Figure 2. The top row plots the age, period, and cohort effects implied by the parameter vector θ in the example from Bell & Jones (2014), while the bottom row plots the age, period, and cohort effects implied by the parameter vector θ^*. Under θ, there is a linear effect in cohort but not in period; under θ^* it is the opposite. Switching that linear effect from cohort, under θ, to period, under θ^*, also influences the age effect; the slope in age is much less steep in the bottom row. Clearly the choice between θ and θ^* will substantially impact interpretation, but both parameter vectors are an equally good fit to the data.

Kuang et al. (2008) show that the Carstensen transformation fully characterises the lack of identification in the model. There are not more than these four dependencies. Equation (10) therefore summarizes the $q - 4$ dimensional variation of the linear function from θ to μ.

3.3 Towards identification: Reparametrization of the classical APC model in terms of accelerations

The APC acceleration framework for aggregate data relies on a reparametrization of the classical APC model which addresses the lack of identification outlined in §3.2.2. The idea of reparametrization is to combine parameters in the original model (7) such that the number of parameters is reduced by the amount needed to resolve the under-identification. In the context of the classical APC model, it is desirable to find a reparametrization that is invariant to the constants $\{\mathfrak{a}, \mathfrak{b}, \mathfrak{c}, \mathfrak{d}\}$ defined in §3.2.2. Such a reparametrization does not constrain the original parameters θ of the classical APC model, thus avoiding the problems associated with constraint-based approaches to the APC identification problem. Ideally, the reparametrized model should also be easy to interpret.

The reparametrization used in the APC acceleration framework for aggregate data is expressed in terms of a parameter vector ξ, which has four fewer elements than the original parameter vector θ, exactly addressing the under-identification. The reparametrization in terms of ξ is invariant to the constants $\{\mathfrak{a}, \mathfrak{b}, \mathfrak{c}, \mathfrak{d}\}$ defined in §3.2.2. The parameter ξ is also straightforward to interpret: it combines accelerations in age, period, and cohort, which have long been recognised as identified in the classical APC model, with a single linear plane.

3.3.1 Defining accelerations in the classical model

Accelerations in age, period, and cohort can be defined in terms of the parameters of the classical APC model. Recall that an acceleration in age captures how the effect of aging one year changes with age. In terms of the classical APC model the effect of aging one year at age a is given by $\Delta\alpha_a = \alpha_a - \alpha_{a-1}$. The change in that effect from the previous age is given by

$$\Delta^2\alpha_a = \Delta\alpha_a - \Delta\alpha_{a-1} = (\alpha_a - \alpha_{a-1}) - (\alpha_{a-1} - \alpha_{a-2}). \tag{12}$$

This term $\Delta^2\alpha_a$ is the acceleration at age a.

To show that the acceleration $\Delta^2\alpha_a$ is identified in the classical model, it must be shown to be invariant to the arbitrary constants $\{\mathfrak{a}, \mathfrak{b}, \mathfrak{c}, \mathfrak{d}\}$. That is, we must show that $\Delta^2\alpha_a = \Delta^2\alpha_a^*$ for α_a^* defined in (11). The proof is as follows:

$$\begin{aligned}
\Delta^2\alpha_a^* &= \alpha_a^* - 2\alpha_{a-1}^* + \alpha_{a-2}^* \\
&= \alpha_a + \mathfrak{a} + \mathfrak{d} \times a - 2[\alpha_{a-1} + \mathfrak{a} + \mathfrak{d} \times (a-1)] + \alpha_{a-2} + \mathfrak{a} + \mathfrak{d} \times (a-2) \\
&= \alpha_a - 2\alpha_{a-1} + \alpha_{a-2} + \underbrace{\mathfrak{a} - 2\mathfrak{a} + \mathfrak{a}}_{=0} + \mathfrak{d} \times [\underbrace{a - 2(a-1) + a - 2}_{=0}] \\
&= \Delta^2\alpha_a.
\end{aligned}$$

Similar proofs exist for period and cohort accelerations, $\Delta^2\beta_p$ and $\Delta^2\gamma_c$.

These accelerations have long been known to be identified from the classical model, and have been used in several applications. An early example is Clayton & Schifflers (1987). McKenzie (2006) identified period accelerations in consumption from aggregate data and attributed the discontinuities they revealed to the Mexican peso crisis. He also identified age accelerations in consumption and used them to test the life cycle consumption hypothesis. Van Landeghem (2012) identified age accelerations in self-reported well-being from individual-level data and used them to evaluate the claim that well-being is U-shaped in age.

In previous applications using accelerations, the accelerations were constructed in multi-step procedures rather than being identified from a regression. The reliance on such cumbersome procedures may explain the limited adoption of the use of accelerations by researchers. The alternative approach of embedding the accelerations in a reparametrized version of the classical model, which can be estimated using standard linear regression, is expected to expand the use of accelerations. This is the approach of the framework developed for aggregate data in Nielsen (2015), and for individual-level data in this thesis.

3.3.2 The idea of reparametrization

Reparametrization is a general approach to dealing with overparametrized models which preserves all variation present in the model while still permitting the identification of a unique set of parameters. The number of parameters to be estimated is reduced by combining parameters of the original model into a new parameter vector which is identified and freely varying. In the context of the classical APC model, the original parameter vector θ, of dimension q, is combined to produce a new parameter vector ξ, which is selected to have the following properties. First, ξ should be of dimension $q-4$. Second, it should be uniquely identified from the data. Third, ξ should be invariant to the arbitrary constants $\{\mathfrak{a}, \mathfrak{b}, \mathfrak{c}, \mathfrak{d}\}$. This invariance indicates that the variation present in the original model is preserved; the reparametrization to ξ does not constrain θ in any way.

To illustrate the idea of reparametrization, consider a simple example of estimating the effect of eye colour on vitamin D absorption for a population of individuals indexed by i. The proposed model is

$$D_i = \psi_0 + \psi_1 \mathbb{1}(eye_i = blue) + \psi_2 \mathbb{1}(eye_i = brown) + \psi_3 \mathbb{1}(eye_i = green). \tag{13}$$

Here ψ_0 is the baseline level of vitamin D absorption and the coefficients ψ_1 through ψ_3 capture the additional effect on vitamin D absorption of having blue, brown, or green eyes respectively.

Unfortunately it is not possible to identify all four parameters. The model is over-parametrized by one degree. The vector $\{\psi_0, \psi_1, \psi_2, \psi_3\}$ is observationally equivalent to an alternative parameter vector defined in terms of an arbitrary constant $\mathfrak{g}$: $\{\psi_0 + \mathfrak{g}, \psi_1 - \mathfrak{g}, \psi_2 - \mathfrak{g}, \psi_3 - \mathfrak{g}\}$.

To address the under-identification the model can be reparametrized:

$$D_i = \phi_1 \mathbb{1}(eye_i = blue) + \phi_2 \mathbb{1}(eye_i = brown) + \phi_3 \mathbb{1}(eye_i = green)$$
$$\phi_1 = (\psi_0 + \psi_1) \quad ; \quad \phi_2 = (\psi_0 + \psi_2) \quad ; \quad \phi_3 = (\psi_0 + \psi_3)$$

The parameter ϕ_1 represents the combined, identified effect of baseline plus blue-eye-specific vitamin D absorption. The separate effects ψ_0 and ψ_1 cannot be identified, but the reparametrization does not limit their variation; that, is the parameter vector $\{\phi_1, \phi_2, \phi_3\}$ is invariant to the choice of $\mathfrak{a}$. This is a very simple example of the reparametrization approach taken by Nielsen (2015) to address the under-identification in the classical APC model.

An alternative approach to such identification problems that is sometimes used involves the imposition of untestable identifying constraints. For example, in model (13), I could impose the constraint that $\psi_1 = 0$, i.e. there is no additional effect of having blue eyes on vitamin D absorption. Under the assumption that this constraint is valid, the remaining original parameters are identified: $\phi_1 = \psi_0, \phi_2 = \psi_2, \phi_3 = \psi_3$. However this approach is not invariant to $\mathfrak{g}$; it relies on $\mathfrak{g} = 0$. The implication is that this approach constrains the variation of the underlying parameter. This creates two problems. First, there is no way to test if the constraint is valid. Second, there is a risk that in interpretation it may be forgotten that $\phi_1 = \psi_0, \phi_2 = \psi_2, \phi_3 = \psi_3$ if and only if the constraint is valid. In the example above, the risk would be that a large estimated value of ϕ_1 is taken to imply a large baseline level of vitamin D absorption, where in fact this interpretation relies on the assumption that $\psi_1 = 0$. Constraints-based approaches to identification of the classical APC model are considered in §5.1.

3.3.3 APC reparametrization: overview

Kuang et al. (2008) developed a reparametrization of the classical APC model in (7) which combines the original q parameters of θ into a new parameter vector ξ of dimension $q-4$, which contains the accelerations defined in §3.3.1. This reparametrization has two advantages. First, it preserves the full variation present in the classical model but is uniquely identified. Second, it incorporates the accelerations, which had been discussed in previous literature.

The first insight used in developing this reparametrization is that any single age, period, or cohort parameter in the classical APC model can be represented as a telescopic sum of accelerations as well as an initial parameter and slope parameter. Consider for example the age parameter, α_5, which can be represented as follows

$$\alpha_5 = \alpha_1 + 4\Delta\alpha_2 + 3\Delta^2\alpha_3 + 2\Delta^2\alpha_4 + \Delta^2\alpha_5. \tag{14}$$

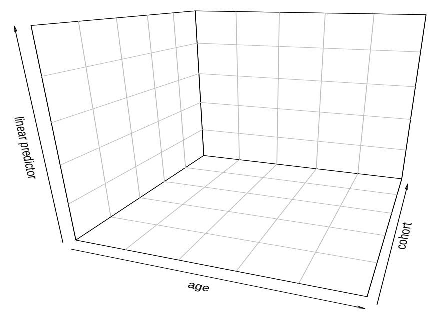
Here α_1 is the initial parameter, $\Delta\alpha_2 = \alpha_2 - \alpha_1$ is a slope in age, and the parameters $\{\Delta^2\alpha_3, \Delta^2\alpha_4, \Delta^2\alpha_5\}$ are accelerations. In fact each age indicator α_a can be represented in terms of the same initial parameter α_1, slope parameter $\Delta\alpha_2$, and a telescopic sum of accelerations, as follows:

$$\alpha_a = \alpha_1 + \sum_{r=2}^{a} \Delta\alpha_r \qquad\qquad \Delta\alpha_r = \Delta\alpha_2 + \sum_{s=3}^{r} \Delta^2\alpha_s. \tag{15}$$

It is straightforward to see that the application of this decomposition to α_5 results in equation (14). Similar representations could be created for an alternative choice of slope or initial parameter; the telescopic sum would then be replaced by an alternative procedure for summation of accelerations. Such representations, in terms of a single initial parameter, slope, and accelerations, can also be created for period and cohort parameters.

The second insight is that the initial parameters and slope parameters in each of age, period, and cohort can be combined into three parameters which define a single linear plane in age-cohort space. Age-cohort space is a three-dimensional space as seen in Figure 3. The horizontal axes of the space describe an age-cohort array like those seen in Figure 1. The vertical axis records the value of the linear predictor μ_{ac} associated with a given age-cohort cell. The single linear plane in age-cohort space, constructed by combining the initial parameters and slope parameters, can be used in conjunction with the sums of accelerations to study the shape of the relationship between the linear predictor and age, period, and cohort. It can also be used for forecasting. The initial parameters in age, period, and cohort are combined with the general intercept δ from parameter vector θ to form the origin point of the linear plane. The three slopes in age, period, and cohort can be reduced to the two slopes of a linear plane due to the relationship $p = a + c - 1$.

Figure 3: Age-cohort space



There are many possible linear planes and associated procedures for summing accelerations, because the sums in (15) could be constructed for any choice of initial parameter and slope. Indeed it is a general property of reparametrization as a technique that each model admits multiple reparametrizations. However, each reparametrization is unique in the sense that once a reparametrization is chosen, only one solution for the reparametrized vector is consistent with the data.

Kuang et al. (2008) and Nielsen (2015) focus on a particular choice of linear plane which is useful for estimation. This linear plane allows age and cohort accelerations to be summed symmetrically, facilitating construction of a design matrix. It is described in §3.3.4.

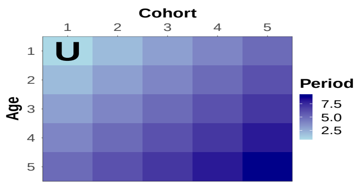
3.3.4 APC reparametrization: detail

The procedure for constructing the reparametrized classical model, in terms of the parameter vector ξ, has five steps. The first step is to select an appropriate linear plane. The second step is to express all parameters in the classical APC model in terms of the initial parameters and slopes of that linear plane, plus accelerations, using the approach in (15). The third step is to combine these re-expressed parameters to get an expression for the linear predictor μ_{ac} in terms of the linear plane and accelerations. The fourth step is to verify that the new parameter ξ has dimension $q-4$. In the fifth step, a concise representation of both the full parameter vector ξ and associated design vector x_{ip} is defined.

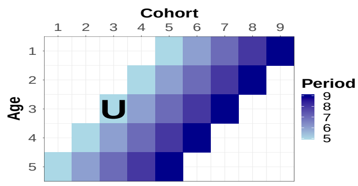
The first step is to select a linear plane which results in a simple design matrix. Simplicity is achieved by selecting a plane in which age and cohort are treated symmetrically and the origin point of the linear plane is close to the origin corner of the age-cohort space. The linear plane is defined by three parameters: an origin point and two slopes. I first consider the origin point, then the slopes.

The origin point of the linear plane is chosen to be the first point where the age and cohort indices are equal in the portion of the age-cohort space occupied by the dataset. This origin point is shown for two data structures displayed in age-cohort space in Figure 4. The origin point is marked by the letter U. Figure 4a shows data that has an age-cohort structure, while Figure 4b shows data that has an age-period structure; the different data structures were defined in §3.2.1.

Figure 4: Location of origin point in data in age-cohort space



(a) Age-cohort



(b) Age-period

Visualisations like those in Figure 4 are not necessary to find the origin point of the linear plane; it can be calculated from the data-determined constant L. Recall from §3.2.1 that L is the number of periods in the age-cohort space that predate data collection, and is used to index the period parameters in the classical APC model: $\{\beta_{L+1}, \ldots, \beta_{L+P}\}$. For data with an age-cohort structure as in Figure 4a, $L = 0$. Define U to be the integer part of $(L+3)/2$, that is, $U = \lfloor (L+3)/2 \rfloor$. The origin of the linear plane is the point where $a = c = U$. The period in which that point lies is $p = 2U - 1$. The value of the linear predictor μ_{ac} from the classical APC model at the origin point is given by

$$v_o = \delta + \alpha_U + \beta_{2U-1} + \gamma_U. \tag{16}$$

This value v_o is the intercept, or origin point, parameter of the linear plane.

The slopes of the linear plane are constructed symmetrically along the age and cohort axes of the age-cohort space, with respect to the origin point of the linear plane $\{U, U\}$. The first slope is obtained by taking a unit increment along the age axis of the age-cohort space away from this point. The slope associated with this unit increment along the age axis is the difference between the value of the linear predictor at $\{U, U\}$, and the value of the linear predictor at the adjacent cell $\{U+1, U\}$. This difference,

$$\begin{aligned} v_a =& \mu_{U+1,U} - \mu_{U,U} \\ =& \alpha_{U+1} + \beta_{2U} - \alpha_U - \beta_{2U-1} \\ =& \Delta\alpha_{U+1} + \Delta\beta_{2U}, \end{aligned} \tag{17}$$

is the first of the two slopes of the linear plane. The second slope is based on taking a unit increment away from the origin point along the cohort axis of the age-cohort space, i.e. the difference between the value of the linear predictor at $\{U, U\}$, and the value of the linear predictor at the adjacent cell $\{U, U+1\}$:

$$\begin{aligned} v_c =& \mu_{U+1,U} - \mu_{U,U} \\ =& \beta_{2U} + \gamma_{U+1} - \gamma_U - \beta_{2U-1} \\ =& \Delta\gamma_{U+1} + \Delta\beta_{2U}. \end{aligned} \tag{18}$$

Having defined the linear plane, the second step is to express each of the parameters of the classical APC model (7) in terms of APC accelerations plus the parameters of the linear plane. This is done using telescopic sums, such as the following for age parameters of the form α_a

$$\alpha_a = \alpha_U + \sum_{r=U+1}^{a} \Delta\alpha_r \qquad\qquad \Delta\alpha_r = \Delta\alpha_{U+1} + \sum_{s=U+2}^{r} \Delta^2\alpha_s. \tag{19}$$

For example, α_{U+4} can be decomposed into

$$\alpha_{U+4} = \alpha_U + 4\Delta\alpha_{U+1} + 3\Delta^2\alpha_{U+2} + 2\Delta^2\alpha_{U+3} + \Delta^2\alpha_{U+4}. \tag{20}$$

Note the similarity to equations (14) and (15). Similar decompositions can be defined for period and cohort parameters.

The third step is to use these decompositions to express the linear predictor of the classical APC model, μ_{ac}, in terms of accelerations and parameters of the linear plane. Insert the decompositions for the age, period, and cohort parameters into the linear predictor $\mu_{ac} = \alpha_a + \beta_p + \gamma_c + \delta$. The result is

$$\begin{aligned} \mu_{ac} =& (\delta + \alpha_U + \gamma_U + \beta_{2U-1}) \\ &+ \Delta\alpha_{U+1}(a-U) + \Delta\gamma_{U+1}(c-U) + \Delta\beta_{2U}[p-(2U-1)] \\ &+ \mathbb{A}_a + \mathbb{P}_p + \mathbb{C}_c \end{aligned} \tag{21}$$

where $\mathbb{A}_a, \mathbb{P}_p$, and $\mathbb{C}_c$ are sums of accelerations in each of age, period, and cohort:

$$\begin{aligned} \mathbb{A}_a &= \mathbb{1}(a<U)\sum_{r=a+2}^{U+1}\sum_{s=r}^{U+1}\Delta^2\alpha_s + \mathbb{1}(a>U+1)\sum_{r=U+2}^{a}\sum_{s=U+2}^{r}\Delta^2\alpha_s \\ \mathbb{P}_p &= \mathbb{1}(L\ odd\ \&\ p=2U-2)\Delta^2\beta_{2U} + \mathbb{1}(p>2U)\sum_{r=2U+1}^{p}\sum_{s=2U+1}^{r}\Delta^2\beta_s \\ \mathbb{C}_c &= \mathbb{1}(c<U)\sum_{r=c+2}^{U+1}\sum_{s=r}^{U+1}\Delta^2\gamma_s + \mathbb{1}(c>U+1)\sum_{r=U+2}^{c}\sum_{s=U+2}^{r}\Delta^2\gamma_s \end{aligned}$$

These definitions are drawn from p.62 of Nielsen (2015). Recall that for data with an age-cohort structure, $L = 0$ so $U = 1$ and

$$\begin{aligned} \mathbb{A}_a =& \mathbb{1}(a>2)\sum_{r=3}^{a}\sum_{s=3}^{r}\Delta^2\alpha_s \\ \mathbb{P}_p =& \mathbb{1}(p>2)\sum_{r=3}^{p}\sum_{s=3}^{r}\Delta^2\beta_s \\ \mathbb{C}_c =& \mathbb{1}(c>2)\sum_{r=3}^{c}\sum_{s=3}^{r}\Delta^2\beta_s. \end{aligned}$$

The three difference terms in equation (21), $\Delta\alpha_{U+1}$, $\Delta\beta_{2U}$, and $\Delta\gamma_{U+1}$, are combined to produce the two slopes of the linear plane, v_a and v_c. This is done using the relation $a + c = p + 1$. Line 2 of (21) becomes

$$(\Delta\alpha_{U+1} + \Delta\beta_{2U})(a-U) + (\Delta\gamma_{U+1} + \Delta\beta_{2U})(c-U) \tag{22}$$

where $(\Delta\alpha_{U+1} + \Delta\beta_{2U}) = v_a$ and $(\Delta\alpha_{U+1} + \Delta\beta_{2U}) = v_c$. Therefore the new parameter vector ξ is

$$\xi = \{v_o, v_a, v_c, \Delta^2\alpha_3, \ldots, \Delta^2\alpha_A, \Delta^2\beta_{L+3}, \ldots, \Delta^2\beta_{L+P}, \Delta^2\gamma_3, \ldots, \Delta^2\gamma_C\}. \tag{23}$$

The fourth step is to show that this new representation in terms of ξ contains exactly $q-4$ parameters. Recall that q is the dimension of θ, i.e. one general intercept plus an indicator for each age, period, and cohort: $q = 1 + A + P + C$. The new representation has three parameters, v_o, v_a, and v_c, describing a single linear plane. It also has a complete set of accelerations in age, period, and cohort. The total number of accelerations is $A - 2 + P - 2 + C - 2$ (subtract 2 because accelerations cannot be defined for the first two ages, periods, or cohorts in the sample). Thus the total number of parameters in ξ is $3 + A + P + C - 6 = q - 4$.

The final step of the procedure is to write the reparametrized linear predictor in terms of a design vector x_{ac} and a parameter vector ξ: $\mu_{ac} = x'_{ac}\xi$. The design vector x_{ac} summarizes the linear plane and all cumulations of accelerations performed in $\mathbb{A}_a, \mathbb{P}_p$, and $\mathbb{C}_c$. It is defined as follows, with $m(r,s) = max(r - s + 1, 0)$:

$$x_{ac} = \{1, (a-U), (c-U), x^{\mathbb{A}}_{ac}, x^{\mathbb{P}}_{ac}, x^{\mathbb{C}}_{ac}\} \tag{24}$$

where $x^{\mathbb{A}}_{ac}$, $x^{\mathbb{P}}_{ac}$, and $x^{\mathbb{C}}_{ac}$, cumulate accelerations in age, period, and cohort:

$$x^{\mathbb{A}}_{ac} = m(1,a), \ldots, m(U-1,a), m(a, U+2), \ldots, m(a, A), \tag{25}$$

$$x^{\mathbb{P}}_{ac} = \begin{cases} \mathbb{1}(p = 2U-2), m(p, 2U+1), \ldots, m(p, 2U-3+P) & \text{for } L \text{ odd} \\ m(p, 2U+1), \ldots, m(p, 2U-2+P) & \text{for } L \text{ even} \end{cases} \tag{26}$$

$$x^{\mathbb{C}}_{ac} = m(1,c), \ldots, m(U-1,c), m(c, U+2), \ldots, m(c, C) \tag{27}$$

The parameter vector is

$$\xi = \{v_o,\ v_a,\ v_c, \xi^{\mathbb{C}}, \xi^{\mathbb{P}}, \xi^{\mathbb{C}}\}, \tag{28}$$

where $\xi^{\mathbb{A}}$, $\xi^{\mathbb{P}}$, and $\xi^{\mathbb{C}}$ collect accelerations in age, period, and cohort respectively:

$$\begin{aligned} \xi^{\mathbb{A}} &= \Delta^2\alpha_3,\ ...,\ \Delta^2\alpha_A \\ \xi^{\mathbb{P}} &= \Delta^2\beta_{L+3},\ ...,\ \Delta^2\beta_{L+P} \\ \xi^{\mathbb{C}} &= \Delta^2\gamma_3,\ ...,\ \Delta^2\gamma_C. \end{aligned} \tag{29}$$

The APC acceleration framework for aggregate data embeds this reparametrized linear predictor $\mu_{ac} = x'_{ac}\xi$ in a regression model. The regression model links the linear predictor to the outcome of interest y_{ac} and permits estimation of the parameters in ξ. Thus, it is possible to obtain estimates of the accelerations without needing to use the multi-stage procedures of previous studies.

3.4 Analysis with the reparametrized classical APC model

The APC acceleration framework, which estimates the APC accelerations as elements of the parameter vector ξ, can be used in at least four ways.

First, estimates of the accelerations alone can be used to evaluate policy changes or shocks. I explain this usage in §3.4.1.

Second, the sum of accelerations provided for in the design vector x_{ac} can be used to examine the shape of the non-linear relationship between an outcome of interest and age, period, or cohort. To isolate the non-linear part of the relationship, the sum of accelerations must be detrended. §3.4.2 presents the procedure for summing and detrending accelerations, and subsequently constructing a visual representation of the shape of the non-linear relationship between an outcome of interest and age, period, or cohort.

Third, restrictions on the reparametrized classical model can be tested, either with the goal of achieving a more parsimonious representation of the data or with the goal of testing some functional form restriction. The functional form restriction might be implied by economic theory or imposed for convenience. I discuss some common restrictions in §3.4.3.

Fourth, projections based on the reparametrized model can be used for forecasting. I give examples in §3.4.4.

3.4.1 Accelerations identify discontinuities

Estimates of the accelerations can be used to identify discontinuities in the relationship between an outcome of interest and age, period, or cohort. For example, McKenzie (2006) identifies a negative period acceleration in Mexican consumption in 1996, using data recorded at two-year intervals. This means that between 1994 and 1996, consumption either grew more slowly than it had between 1992 and 1994, or declined more rapidly. Either one indicates depressed consumption, which McKenzie attributes to the 1995 peso crisis.

McKenzie (2006) use a multi-step procedure to estimate accelerations, but the general framework described in §3.3 allows all accelerations to be estimated from any generalized linear model with linear predictor $\mu_{ac} = x'_{ac}\xi$. McKenzie's procedure is restricted to linear models of the form $y_{ac} = \mu_{ac} + \varepsilon_{ac}$ and is not generalizable to other linear models such as the logit. McKenzie's procedure for estimating accelerations is as follows: first, differences of the outcome variable between adjacent cells by age, period, and cohort are constructed. Second, differences of those differences along age, period, and cohort are constructed. An average of the appropriate sub-set of these differences is constructed to estimate a single acceleration. This procedure is repeated for each acceleration. In the general framework described in §3.3, each acceleration is an element of the parameter vector ξ and so is estimated by regression. Relative to the procedure used by McKenzie, this framework is faster and reduces the potential for error. Furthermore, this framework is suitable for all generalized linear models, whereas McKenzie's framework is suitable only for linear models that can be estimated by ordinary least squares.

3.4.2 Sums of accelerations describe relationships

The sums of accelerations implied by the combination of the design vector x_{ac} and the parameter ξ can be used to study the shape of the relationship between an outcome and age, period, and cohort. Consider the sum of age accelerations at age a, given by $\mathbb{A}_a = x^{\mathbb{A}}_{ac}\xi^{\mathbb{A}}$. By plotting the value of this sum on the Y-axis against age a on the X-axis, one can visualise how the relationship between age and the outcome of interest evolves with age. One can determine whether the relationship is uniformly concave or convex, or some other shape. For example, Nielsen (2015) finds a concave relationship between age and the number of mesothelioma deaths in Belgium, whereas Fannon & Nielsen (2019) find a convex relationship between age and the proportion of the US labour force that is unemployed. Similarly the shape of the relationship between an outcome and cohort can be studied by plotting $\mathbb{C}_c = x^{\mathbb{C}}_{ac}\xi^{\mathbb{C}}$ against cohort, and the shape of the relationship between an outcome and period can be studied by plotting $\mathbb{P}_p = x^{\mathbb{P}}_{ac}\xi^{\mathbb{P}}$ against period.

It is important here to only consider the non-linear part of the shape. The non-linear part of the shape, which describes deviations from linearity such as convexity or concavity, is identified through the accelerations. The linear part of the shape, indicating whether an outcome increases or decreases with age, period, or cohort, is not identified.

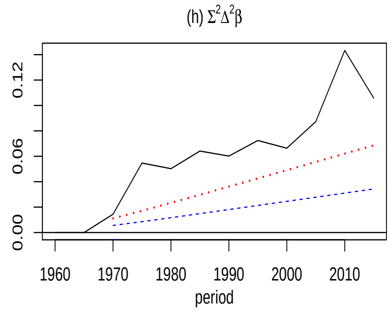
The reparametrization in terms of ξ, with sums of accelerations given by $\{\mathbb{A}_a, \mathbb{P}_p, \mathbb{C}_c\}$, as described in §3.3.3, is not optimal for visualising the non-linear part of the shape. Instead, it is optimized for ease of construction of the design vector by allowing age and cohort to be. As a consequence, the estimated linear plane is highly dependent on the age, period, and cohort effects in the vicinity of U. This may lead to a problem where there is the appearance of a slope in the summed accelerations, as seen in Figure 5a.

A slightly different parametrization of the classical APC model, which selects a different linear plane and summation of the accelerations, is preferred for isolating the non-linear part of the shape. It is important to note that the accelerations are not changed by this different parametrization; it is only the design vector x_{ac} and the linear plane parameters $\{v_o, v_a, v_c\}$ that change. The linear plane in this parametrization is chosen to ensure that the sum of accelerations in each of age, period, and cohort is anchored to begin and end in zero. The parameter vector associated with this parametrization is ξ^{dt}, with the dt referring to the "detrending" implied by anchoring the sums of accelerations at zero. This is defined

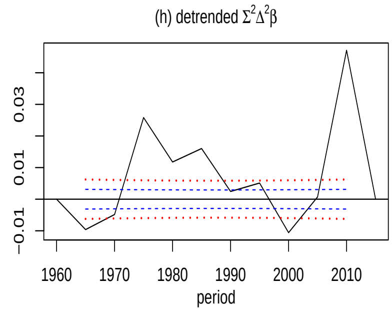
$$\xi^{dt} = \{v^{dt}_o, v^{dt}_a, v^{dt}_c, \xi^{\mathbb{A}}, \xi^{\mathbb{B}}, \xi^{\mathbb{C}}\}. \tag{30}$$

The sum of period accelerations resulting from this detrended representation, $\mathbb{P}^{dt}_p$, is seen in Figure 5b. It uses the same data and accelerations as Figure 5a but clearly isolates the non-linear part of the shape.

Figure 5: US unemployment rate, summed period accelerations



(a) $\mathbb{P}_p$ as in §3.3.3



(b) Detrended representation $\mathbb{P}_p^{dt}$

Data for these figures is from the OECD online database, see Fannon & Nielsen (2019).
Blue dashed, red dotted lines = 1, 2 standard deviations from zero.

The APC detrended acceleration parametrization in terms of ξ^{dt} can be constructed as a bijective mapping from the ξ parametrization. This is useful because the design matrix associated with ξ^{dt} is difficult to construct from scratch. The new linear plane parameters are constructed as follows:

$$
\begin{aligned}
v_o^{dt} =& v_o - (U-1)(v_a + v_c) - \mathbb{A}_1 - \mathbb{P}_{P+1} - \mathbb{C}_1 - \frac{L}{P-1}(\mathbb{P}_{L+P} - \mathbb{P}_{L+1}) \\
v_a^{dt} =& v_a + \frac{1}{A-1}(\mathbb{A}_A - \mathbb{A}_1) + \frac{1}{P-1}(\mathbb{P}_{L+P} - \mathbb{P}_{L+1}) \\
v_c^{dt} =& v_c + \frac{1}{C-1}(\mathbb{C}_C - \mathbb{C}_1) + \frac{1}{P-1}(\mathbb{P}_{L+P} - \mathbb{P}_{L+1}).
\end{aligned}
$$

The sums of accelerations are given by

$$
\begin{aligned}
\mathbb{A}_a^{dt} &= x_{ac}^{\mathbb{A}dt\prime}\xi^{\mathbb{A}} = \mathbb{A}_a - \mathbb{A}_1 - \frac{a-1}{A-1}(\mathbb{A}_A - \mathbb{A}_1) \\
\mathbb{P}_p^{dt} &= x_{ac}^{\mathbb{P}dt\prime}\xi^{\mathbb{P}} = \mathbb{P}_p - \mathbb{P}_1 - \frac{p-L-1}{P-1}(\mathbb{P}_{L+P} - \mathbb{P}_{L+1}) \\
\mathbb{C}_c^{dt} &= x_{ac}^{\mathbb{C}dt\prime}\xi^{\mathbb{C}} = \mathbb{C}_c - \mathbb{C}_1 - \frac{c-1}{C-1}(\mathbb{C}_C - \mathbb{C}_1).
\end{aligned}
$$

The model with the new ξ^{dt} parametrization and design vector x_{ac}^{dt} is then

$$
\mu_{ac} = x_{ac}^{dt\prime}\xi^{dt} = v_o^{dt} + (a-1)v_a^{dt} + (c-1)v_c^{dt} + \mathbb{A}_a^{dt} + \mathbb{P}_p^{dt} + \mathbb{C}_c^{dt}. \tag{31}
$$

These relationships are derived in Nielsen (2015).

This detrended parametrization is used exclusively for exploring the shape of the relationship between age, period, or cohort and some outcome. It was used to identify concavity and convexity in Nielsen (2015) and Fannon & Nielsen (2019), described above. For other applications using accelerations, the parametrization in terms of ξ is used because it is easier to estimate. I provide illustrations of how the detrended parametrization adapted to individual-level data may be used in §**??** and §**??** of this thesis.

3.4.3 Testing restrictions on the APC acceleration model

The regression framework makes it easy to test restrictions on the APC model, via likelihood ratio or Wald tests. The main APC model can be compared against sub-models, which restrict certain elements of the parameter vector ξ to zero. Seven categories of sub-model are considered (Nielsen, 2015; Oh & Holford, 2015).

AC/PC/AP First, the absence of accelerations in one of the APC effects can be tested. For instance, the absence of period accelerations is tested by imposing $\Delta^2\beta_{L+3} = \cdots = \Delta^2\beta_{L+P} = 0$. This gives an Age-Cohort (AC) model. In terms of the unidentified original parametrization θ, this is written as $\beta_{L+1} = \ldots \beta_{L+P} = 0$. The two formulations of the hypothesis are in fact equivalent (Nielsen & Nielsen, 2014). The latter formulation obscures the degrees of freedom and hides that the hypothesis does not constrain the linear period effects. Period-Cohort (PC) and Age-Period (AP) models are analogous to AC models.

Ad/Pd/Cd Second, the absence of accelerations in two components can be tested. For example, the absence of period and cohort accelerations is tested by imposing $\Delta^2\beta_{L+3} = \cdots = \Delta^2\beta_{L+P} = 0$ and $\Delta^2\gamma_3 = \cdots = \Delta^2\gamma_C = 0$, while leaving the linear plane unrestricted. This is the Age-drift (Ad) model (Clayton & Schifflers, 1987). Analogous models are Cohort-drift (Cd) and Period-drift (Pd).

A/C/P Third, a model with only one slope in the linear plane can be tested. For instance, the Age (A) model is obtained by imposing $v_c = 0$ on the Ad model. The constraint is that $v_c = \Delta\gamma_{U+1} + \Delta\beta_{2U} = 0$. In this case, $v_a = \Delta\alpha_{U+1} + \Delta\beta_{2U}$ identifies a combined age-period slope. The linear slope of the age effect, $\Delta\alpha_{U+1}$, remains unidentifiable. Analogous models are Cohort (C) and Period (P).

t Fourth, a pure linear plane model can be tested, where all double differences $\Delta^2\alpha_a$, $\Delta^2\beta_p$, $\Delta^2\gamma_c$ are set to zero. This is the trend (t) model.

tA/tP/tC Fifth, the trend can be further restricted. The tA model is constrained to have $v_c = 0$, while the tC model is constrained to have $v_a = 0$. The tP model is constrained to have $v_a = v_c$.

1 Sixth, the intercept (1) model has neither slope nor accelerations. All parameters are constrained to equal zero except v_o.

Finally, functional form restrictions may be of interest. For instance, a quadratic age effect $\alpha_a = \lambda_0 + \lambda_1 a + \lambda_2 a^2$ arises when $\Delta^2\alpha_3 = \cdots = \Delta^2\alpha_A = 2\lambda_2$ is imposed (Fannon & Nielsen, 2019; McKenzie, 2006).

An example of testing model restrictions is seen in Martínez Miranda et al. (2015), where restriction to the "AC" model is considered. Testing of model restrictions in the framework for individual-level data is explored in §?? and §??.

3.4.4 Forecasting

Estimates of the acceleration parameters can be used in conjunction with the linear plane to produce forecasts. First, the age-cohort space is extended to cover the future periods which it is desired to forecast. Then, the APC acceleration model is extended to incorporate the new age, period, and cohort accelerations required. These additional acceleration terms will need to be forecast, perhaps via an AR(1) or other time-series model for the accelerations in that time dimension. The linear plane is projected forward to cover the new age-cohort cells. Note that an implicit assumption of no structural change underpins any forecasts generated in this fashion. An example is the forecasting of mesothelioma mortality in Martínez Miranda et al. (2015). Further examples of epidemiological forecasting using the ξ reparametrization framework are Zhao et al. (2019), Ji et al. (2019), and Oddone et al. (2020).

4 Advantages of the reparametrized classical APC model

There are at least three benefits to using the APC acceleration framework to study the effects of age, period, and cohort. First, the parameter ξ is uniquely identified. Second, the parametrization is invariant to θ, meaning that it does not impose any constraints on the classical APC model. Third, the reparametrization is in terms of accelerations, which are familiar objects in the APC literature and have a natural interpretation. The APC acceleration framework permits these accelerations to be estimated from regression.

The parameter ξ is unique with respect to a particular set of linear predictors μ_{ac}. This is in contrast to the parametrization in terms of θ, where multiple values of θ fit the data equally well. For a given design vector x_{ac} and set of linear predictors μ_{ac}, there is a single ξ which provides the best fit.

The parameter ξ is invariant to the transformations in terms of $\{\mathfrak{a}, \mathfrak{b}, \mathfrak{c}, \mathfrak{d}\}$ described in equation (10) of §3.2.2. Corollary 2 in Kuang et al. (2008) shows that ξ is a maximal invariant function of θ with respect to (10). This means that for any θ which satisfies the linear predictors μ_{ac}, when it is reparametrized in terms of ξ, the same values of ξ will be obtained. The implication is that the reparametrization in terms of ξ does not limit the variation of θ. This stands in contrast to other approaches to identification that are discussed in §5. Requiring invariance draws focus to the identifiable part of θ, avoiding the part of θ that cannot be identified. While not exactly solving the APC identification problem, this approach makes the problem ignorable. Here, ξ represents what can be learned in an APC model, while the transformations (10) describe what cannot be learned.

The fact that the ξ parametrization is built around accelerations, which are well-known to be identified, means that it can be more readily used and interpreted. There are previous examples of successful use of accelerations in forecasting (Martínez Miranda et al., 2015); evaluating the impact of shocks (McKenzie, 2006); and testing theories (McKenzie, 2006; Van Landeghem, 2012). Two additional use cases, finding of new stylized facts and testing convenience functional form restrictions, are illustrated in §**??** and §**??** respectively.

The embedding of the accelerations in a linear regression framework via the ξ reparametrization is an improvement on the previous literature using accelerations. It makes the accelerations more accessible to researchers who are familiar with regression but may be put off by the multi-step procedures previously used to estimate accelerations. It simplifies model selection, because it is easy to test restrictions on the model. Finally it facilitates additional applications, including forecasting and using the non-linear shape of the APC effects to find new stylized facts. For forecasting, the combined linear plane provided for in ξ is essential. The design matrix x_{ac}^{dt} described in §3.4.2 provides for summation of the accelerations in a way that visually isolates the non-linear part of the relationship between an outcome of interest and age, period, and cohort.

5 Other approaches to analysis of APC effects

Several other approaches to APC analysis are commonly used in economics and other social sciences. The most common approach in economics uses the classical APC model in a regression framework; but instead of reparametrization, constraints are imposed to identify the model. In the best examples of this approach, the identifying constraints are carefully chosen based on the application at hand (e.g. Lagakos et al., 2018); but in many instances little justification is provided for the choice of constraints (Bardazzi & Pazienza, 2018; Bíró, 2017). The imposition of untestable and weakly-justified constraints creates problems for interpretation, since the estimated APC effects from these models depend upon the constraints.

A second approach common in economics is to replace the APC terms in the classical model with latent variables that reflect the phenomena the APC effects are expected to capture (Heckman & Robb, 1985). Again, problems of interpretation arise when this approach is not used carefully.

A third approach, less common in economics, is to use a model that is not derived from the classical APC model. A well-known example is the Lee-Carter model (Lee & Carter, 1992), used in actuarial science. A thorough analysis of such models is beyond the scope of this paper, but they are also subject to the APC identification problem and often implicitly impose identifying constraints, creating the potential for problems with interpretation analagous to those arising when constraints are imposed on the classical APC model.

A fourth approach is to abandon the effort to get precise estimates of age, period, and cohort effects and instead focus on careful inspection of plots of the data. This is useful in cases where the data is such that the effect of interest is clearly visible from the plots, as in Almond (2006), but that is not always the case.

A final phenomenon seen in economics, although not a solution to the APC problem, is the failure to recognise the APC identification problem at all, for example in Rosenthal (2014). In the following I consider the relationship between these approaches and the reparametrization framework outlined in §3. I discuss these approaches in more detail in this section.

5.1 Constraints on the classical APC model

The imposition of constraints to identify the classical APC model is common in economics. These constraints may be just-identifying or over-identifying. All constraints approaches carry a risk of mis-interpretation if the sensitivity of the estimated effects to the constraints is not fully appreciated.

Just-identifying constraints impose four constraints on the original parameter vector θ of the classical APC model. The new constrained parameter vector θ_c has a dimension of variation equal to $q-4$ and is therefore identifiable.

The use of just-identifying constraints can result in mis-interpretation if the dependence of the resulting estimates on the constraints is not fully understood. Essentially, a problem arises if the solution for θ_c is mistaken as the solution for θ. It is only true that $\theta_c = \theta$ when the imposed just-identifying constraints are true; and it is not possible to test these constraints. A failure to recognise this may lead the applied researcher to make statements about θ when in fact those statements only apply to θ_c. A typical example would be a claim about the linear relationship between age, period, or cohort and some outcome of interest implied by the estimated θ_c, which fails to recognise that the linear relationship is only identified by the constraints.

A simple example of a just-identifying constraint is the approach suggested by Mason et al. (1973) and used by Ejrnæs & Hochguertel (2013). In this approach, four elements of θ are constrained to equal zero: one each in age, period, and cohort, plus one further parameter. For example, one could set $\alpha_1 = \beta_{L+1} = \gamma_1 =$

$\beta_{L+2} = 0$. The first three constraints (one parameter in each of age, period, and cohort) are standard normalizing constraints used when working with complete sets of dummy variables. The constraint on the additional parameter is needed as a consequence of the relationship $p = a + c - 1$, and it is this constraint that creates problems for interpretation. This constrains the linear slope in one of the three of age, period, and cohort, so that any remaining linear effect due to that variable must be distributed between the other two.

Another set of just-identifying constraints, common in economics, is that used by Deaton & Paxson (1994). Again, there are four constraints in total: three normalizing constraints and one linear constraint. The normalizing constraints restrict the sums of age, period, and cohort parameters to zero:

$$\sum_a \alpha_a = \sum_p \beta_{L+p} = \sum_c \gamma_c = 0. \tag{32}$$

The additional constraint is that there is no linear effect of period on consumption. The constraint takes the form

$$\sum_p p\beta_p = 0. \tag{33}$$

One could alternatively set the linear effect in age or cohort to zero. The sensitivity of the estimated effects to the fourth constraint is illustrated by Figure 4 of Lagakos et al. (2018), where one can see how the estimated trends in wages depend on whether the linear effects are constrained to be zero in period or in cohort.

A range of other just-identifying constraints exist. One class of just-identified models impose the fourth constraint by selecting a generalized inverse with particular properties. This class includes the intrinsic estimator of Yang et al. (2004) and the procedure described in Schulhofer-Wohl (2018). O'Brien (2015) provides a discussion of these models, and a refutation of the claim that the constraint implied by the generalized inverse results in $\theta_c = \theta$. There are also Bayesian versions of the constraints-based approach, see the discussion in Fannon & Nielsen (2019).

It is difficult to envision a situation where the use of just-identifying constraints could be preferred to the reparametrization outlined in §3. Interpretation of a just-identified model requires the researcher to keep track of how the imposed constraints affect the estimates, which is not straightforward (see the discussion in Nielsen & Nielsen, 2014). The APC acceleration model does not have this requirement. The just-identified model may be preferred where the researcher has complete confidence in the imposed constraints, so that $\theta_c = \theta$; then the just-identified model permits inference on the linear effects, which the APC acceleration model does not. However, in most applications in economics the underlying phenomena are too complex, and our understanding of them too limited, for such confidence in untestable identifying constraints to be warranted.

Models which involve over-identifying constraints are also common in economics. These approaches share the problem of the just-identifying constraints approach, that the impact of the constraints on the estimated effects is often not fully appreciated. There are many types of over-identifying constraint. One is to assume that there are no effects attributable to one of the three of age, period, and cohort; for example Méndez & Sepúlveda (2012) assume that there are no cohort effects on skill acquisition. A second type involves restricting large numbers of age, period, or cohort effects to be equal, by creating "bands" within which there is no change in the relationship to the outcome of interest. For example, Bíró (2017) uses 10-year cohort bands in his analysis of durable good consumption.

Another over-identifying constraints approach, more common in sociology and demography than economics, is the hierarchical APC (HAPC) model (Yang & Land, 2006). This is model includes a polynomial in age, while period and cohort are modelled as normal, mean-zero random effects. The random effects approach effectively constraints the period and cohort effects, as is discussed in §3.1.2.2 of Fosse & Winship (2019) and §5.4.6 of Nielsen & Nielsen (2014).

The reparametrization framework is preferable to these constraints approaches because it can yield useful insights but is not vulnerable to mis-interpretation. While some of those who use constraints-based methods do so with appropriate care and recognise the sensitivity of their estimates to the identifying assumption (an excellent example is Lagakos et al., 2018), studies which do not take appropriate care are common. There is a particular risk if the constraints are believed to "solve" the identification problem, i.e. if θ_c is confused with θ. It is therefore important to have a general method for APC analysis which does not carry risks of mis-interpretation. The APC acceleration framework satisfies this need.

5.2 Latent variables

The latent variables approach treats age, period, and cohort as "proxies" for underlying latent variables which affect the outcome of interest (Heckman & Robb, 1985). A good latent variables analysis is time-consuming but is more informative about the determinants of the outcome of interest than an APC model. The APC acceleration framework is complementary to such an analysis; it can be used to guide the selection of latent variables. However, mis-application of the latent variables approach shares many of the problems that affect constraints-based analysis.

Good latent variables analysis adheres to the original idea of Heckman & Robb (1985), that all three of age, period, and cohort should be replaced by the latent variables and thus a better understanding of the outcome variable would be gained. An example is Kapteyn et al. (2005), where the focus of the paper is on determining which latent variables drive the cohort effects. The authors use economic theory to identify two candidate latent variables, and then test for the additional explanatory power of non-linear cohort effects while also accounting for age and period non-linearities. The APC acceleration framework can be used for this sort of testing. It can also be used to identify candidate latent variables; for example, in §?? my examination of the sums of age accelerations reveals the importance of accounting for childbirth in the model for hospital in-patient stays.

Problems arise when it is assumed that replacing one of the three of age, period, or cohort with a latent variable permits identification of the effects of the remaining two of age, period, and cohort. For example, in Jensen et al. (2001) industry-wide labour productivity and output variables are used to replace period so that the firm age and vintage effects on productivity can be estimated. This is in effect similar to the constraints approaches: the period effect is constrained to be equal to the effect of the latent variable. The estimated effects of age and cohort will be sensitive to this constraint, which cannot be tested.

5.3 Models other than the classical APC model

A range of alternative APC models exist, which do not inherit their structure from the classical APC model. Extensive analysis of these is beyond the scope of this paper. However, it is worth noting that they frequently suffer from the same issues that arise with constraints on the classical APC model. A well-known example is the Lee-Carter model (Lee & Carter, 1992), typically used for mortality modelling. This differs from the classical APC model because it is not linear in age, period, and cohort effects. Instead, the cohort and period effects are interacted, while the age effects enter additively. Nielsen & Nielsen (2014) discuss how the identification of this model is subject to *a priori* constraints like those imposed on the classical APC model in §5.1.

5.4 Visual analysis

In some studies, no effort is made to statistically separate the effects of age, period and cohort. Instead the focus is on graphical analysis. This works well when the goal is simply to know whether APC effects exist, rather than to quantify them. An example is the work of Almond (2006), who examines the long-term effects of being *in utero* during the 1918 influenza pandemic. There are clear discontinuities in the data associated with the 1918 cohort, see for example his Figures 2 and 3. Another example is Voas & Chaves (2016), who consider the relationship between religious affiliation and the passage of time. They find that for each cohort, the relationship between the passage of time (reflecting either age or period or both) and religious affiliation is flat, but is shifted down compared to the previous cohort. Such a graph implies either pure cohort effects or perfectly balanced age and period effects; the authors argue that the latter is unlikely.

The limitation of this approach is that it is not widely applicable. The data often does not exhibit patterns that are clearly attributable to one of age, period, or cohort. Visual analysis breaks down in such situations. Additionally, it is often important to quantify the magnitude of effects, which is not possible with visual analysis. This is why Almond (2006) follows up his graphical analysis with estimation of deviations from linearity in cohort.

6 Conclusion

By reparametrizing the classical age-period-cohort (APC) model, Kuang et al. (2008) and Nielsen (2015) developed a general regression framework for estimating the identifiable effects of age, period, and cohort from aggregate data. The identifiable effects are accelerations in each of age, period, and cohort, as well as a combined linear plane. These effects have been used in economic applications for data exploration, forecasting, evaluating the impact of policy changes and exogenous shocks, and testing theoretical claims. This thesis develops a complementary general regression framework for estimating age, period, and cohort accelerations, and a combined linear plane, from individual-level data.

The APC acceleration framework has several advantages relative to other approaches which are used to estimate the effects of age, period, and cohort. Other approaches which focus on the identifiable effects, in particular those focused on estimating accelerations, rely on multi-step procedures which are more cumbersome to implement than the APC acceleration framework. Approaches which do not focus on the identifiable effects rely on untestable identifying constraints. The extent to which the estimated APC effects are dependent on the constraints in these models is often not appreciated, leading to inaccurate interpretation of the estimates. The APC acceleration framework focuses on effects which are invariant to these constraints and so is less vulnerable to inaccurate interpretation.

The parameters identified from the APC acceleration framework have been used in a range of applications. The identified parameters are a single linear plane, which combines the linear effects of age, period, and cohort, and accelerations in each of age, period, and cohort. An acceleration in age, for example, captures how the effect of aging one year changes as a person ages. Accelerations have been used with aggregate data to estimate discontinuities arising from currency shocks and to test economic theories such as the life-cycle hypothesis (McKenzie, 2006). Accelerations in conjunction with the linear plane have been used to explore the age, period, and cohort effects in unemployment data (Fannon & Nielsen, 2019) and to forecast mesothelioma mortality (Martínez Miranda et al., 2015).

Despite the range of applications for which the APC acceleration framework is suitable, it has been relatively under-utilized in economics. I suspect four main factors are driving this. First, there may be a lack of awareness of the problems associated with other approaches to APC identification. This hypothesis is supported by the continued use of these approaches in recent papers, such as Bardazzi & Pazienza (2018). Second, there may be an incompatibility issue: many of the applied studies in economics where APC effects are of interest rely on individual-level data (see §2), but the reparametrization framework was developed for aggregate data. Third, there may be a lack of understanding of how accelerations can be effectively used in economic applications. This hypothesis is supported by the fact that applied papers often cite methodological papers on accelerations, but do not discuss them in detail (e.g. Fukuda, 2013). Fourth, researchers may be deterred by the cumbersome multi-step procedures previously used to estimate accelerations.

In this thesis, I address each of the four factors that I suspect are driving the under-utilization of the APC acceleration framework. First, I have already drawn attention to the problems associated with other approaches to APC identification in §5. Second, in next two chapters of the thesis I develop a theoretical framework for using the APC acceleration framework with individual-level data, considering repeated cross section data in §**??** and panel data in §**??**. Third, in each of §**??**, §**??**, and §**??** I provide a different application or example which illustrates how accelerations may be used. Fourth, in §**??** I develop an R package which implements the theoretical framework developed in §**??** and §**??**, enabling accelerations to be estimated from regression in a single command.

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*, 672–712.

Attanasio, O. P. (1998). Cohort analysis of saving behavior by US households. *Journal of Human Resources*, *33*, 575–609.

Bardazzi, R. & Pazienza, M. G. (2018). Ageing and private transport fuel expenditure: Do generations matter? *Energy Policy*, *117*, 396–405.

Bell, A. & Jones, K. (2014). 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*, *101*, 176–180.

Bíró, A. (2017). Effect of ageing on the ownership of durable goods. *Scottish Journal of Political Economy*, *64*, 501–529.

Browning, M., Crossley, T. F., & Lührmann, M. (2016). Durable purchases over the later life cycle. *Oxford Bulletin of Economics and Statistics*, *78*, 145–169.

Carstensen, B. (2007). Age-period-cohort models for the Lexis diagram. *Statistics in Medicine*, *26*, 3018–3045.

Clayton, D. & Schifflers, E. (1987). Models for temporal variation in cancer rates. II: age–period–cohort models. *Statistics in Medicine*, *6*, 469–481.

Deaton, A. S. & Paxson, C. (1994). Saving, growth, and aging in Taiwan. In *Studies in the Economics of Aging* (pp. 331–362). University of Chicago Press.

Dickerson, A., Hole, A. R., & Munford, L. A. (2014). The relationship between well-being and commuting revisited: Does the choice of methodology matter? *Regional Science and Urban Economics*, *49*, 321–329.

Ejrnæs, M. & Hochguertel, S. (2013). Is business failure due to lack of effort? Empirical evidence from a large administrative sample. *Economic Journal*, *123*, 791–830.

Fannon, Z. & Nielsen, B. (2019). Age-period-cohort models. In *Oxford Research Encyclopedia of Economics and Finance*. Oxford University Press.

Fitzenberger, B., Schnabel, R., & Wunderlich, G. (2004). The gender gap in labor market participation and employment: A cohort analysis for West Germany. *Journal of Population Economics*, *17*, 83–116.

Fosse, E. & Winship, C. (2019). Analyzing age-period-cohort data: A review and critique. *Annual Review of Sociology*, *45*, 467–92.

Fukuda, K. (2013). Decomposition of new venture growth into firm age, survey period and vintage effects. *Applied Economics*, *45*, 85–97.

Hanoch, G. & Honig, M. (1985). "True" age profiles of earnings: Adjusting for censoring and for period and cohort effects. *Review of Economics and Statistics*, *67*, 383–394.

Heckman, J. & Robb, R. (1985). *Using longitudinal data to estimate age, period and cohort effects in earnings equations*, (pp. 137–150). New York, NY: Springer New York.

Jensen, J. B., McGuckin, R. H., & Stiroh, K. J. (2001). The impact of vintage and survival on productivity: Evidence from cohorts of US manufacturing plants. *Review of Economics and Statistics*, *83*, 323–332.

Ji, W., Xie, N., He, D., Wang, W., Li, H., & Wang, K. (2019). Age-period-cohort analysis on the time trend of Hepatitis B incidence in four prefectures of southern Xinjiang, China from 2005 to 2017. *International Journal of Environmental Research and Public Health*, *16*.

Kapteyn, A., Alessie, R., & Lusardi, A. (2005). Explaining the wealth holdings of different cohorts: productivity growth and social security. *European Economic Review*, *49*, 1361–1391.

Krueger, A. B. & Pischke, J. (1992). The effect of social security on labor supply: a cohort analysis of the notch generation. *Journal of Labor Economics*, *102*, 412–437.

Kuang, D., Nielsen, B., & Nielsen, J. P. (2008). Identification of the age-period-cohort model and the extended chain-ladder model. *Biometrika*, *95*, 979–986.

Künn-Nelen, A. (2016). Does commuting affect health? *Health Economics*, *25*, 984–1004.

Lagakos, D., Moll, B., Porzio, T., Qian, N., & Schoellman, T. (2018). Life cycle wage growth across countries. *Journal of Political Economy*, *126*, 797–849.

Lee, R. D. & Carter, L. R. (1992). Modelling and forecasting U.S. mortality. *Journal of the American Statistical Association*, *87*, 659–671.

Low, H., Meghir, C., & Pistaferri, L. (2010). Wage risk and employment risk over the life cycle. *American Economic Review*, *100*, 1432–1467.

Martínez Miranda, M. D., Nielsen, B., & Nielsen, J. P. (2015). Inference and forecasting in the age-period-cohort model with unknown exposure with an application to mesothelioma mortality. *Journal of the Royal Statistical Society, Series A*, *178*, 29–55.

Mason, K. O., Mason, W. M., Winsborough, H. H., & Poole, K. (1973). Some methodological issues in cohort analysis of archival data. *American Sociological Review*, *38*, 242–258.

McKenzie, D. J. (2006). Disentangling age, cohort, and time effects in the additive model. *Oxford Bulletin of Economics and Statistics*, *68*, 473–495.

Meghir, C. & Whitehouse, E. (1996). The evolution of wages in the United Kingdom: Evidence from micro data. *Journal of Labor Economics*, *14*, 1–25.

Méndez, F. & Sepúlveda, F. (2012). The cyclicality of skill acquisition: Evidence from panel data. *American Economic Journal: Macroeconomics*, *4*, 128–52.

Nielsen, B. (2015). apc: An r package for age-period-cohort analysis. *The R Journal*, *7*, 52–64.

Nielsen, B. & Nielsen, J. P. (2014). Identification and forecasting in mortality models. *The Scientific World Journal*, *2014*, Article ID 347043, 24 pages.

O'Brien, R. M. (2015). Age–period–cohort models and the perpendicular solution. *Epidemiologic Methods*, *4*, 87–99.

Oddone, E., Bollon, J., Nava, C. R., Bugani, M., Consonni, D., Marinaccio, A., Magnani, C., & Barone-Adesi, F. (2020). Predictions of mortality from pleural mesothelioma in Italy after the ban of asbestos use. *International Journal of Environmental Research and Public Health*, *17*.

Oh, C. & Holford, T. R. (2015). Age-period-cohort approaches to back-calculation of cancer incidence rate. *Statistics in Medicine*, *34*, 1953–1964.

Roberts, J., Hodgson, R., & Dolan, P. (2011). "It's driving her mad": Gender differences in the effects of commuting on psychological health. *Journal of Health Economics*, *30*, 1064–1076.

Rosenthal, S. S. (2014). Are private markets and filtering a viable source of low-income housing? Estimates from a "repeat income" model. *American Economic Review*, *104*, 687–706.

Schulhofer-Wohl, S. (2018). The age-time-cohort problem and the identification of structural parameters in life-cycle models. *Quantitative Economics*, *9*, 643–658.

Van Landeghem, B. (2012). A test for the convexity of human well-being over the life cycle: Longitudinal evidence from a 20-year panel. *Journal of Economic Behavior and Organisation*, *81*, 571–582.

Voas, D. & Chaves, M. (2016). Is the United States a counterexample to the secularization thesis? *American Journal of Sociology*, *121*, 1517–1556.

Yang, Y., Fu, W. J., & Land, K. C. (2004). A methodological comparison of age-period-cohort models: The intrinsic estimator and conventional generalized linear models. *Sociological Methodology*, *34*, 75–110.

Yang, Y. & Land, K. C. (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. & Land, K. C. (2013). *Age-period-cohort analysis: New models, methods, and applications.* Boca Raton, FL: CRC Press.

Zhao, S., Dong, H., Qin, J., Liu, H., Li, Y., Chen, Y., Molassiotis, A., He, D., Lin, G., & Yang, L. (2019). Breast cancer mortality in Chinese women: does migrant status play a role? *Annals of Epidemiology*, *40*, 28–34.