Title
Age-Period-Cohort analysis: A design-based approach

Authors
Dinas, Elias
Stoker, Laura

Publication Date
2014-03-01

DOI
10.1016/j.electstud.2013.06.006

Peer reviewed

Age-Period-Cohort Analysis: A design-based approach

Elias Dinas & Laura Stoker

Dinas, Elias, and Laura Stoker. 2014. "Age-Period-Cohort Analysis: A Design-Based Approach." *Electoral Studies*, 33: 28-40.

Abstract

This paper develops a design-based approach to identifying cohort effects in APC analyses. Cohort effects arise when one cohort is treated by a unique set of formative socialization experiences, which causes it to differ from other cohorts in relevant outcomes. APC analyses typically compare treated and untreated cohorts from a single population. Our approach introduces a second group—a control group, in which no unit is treated but that is otherwise similar to the first—and adapts difference-in-difference estimation to the APC framework. The approach yields two identification strategies, each based on transparent and testable assumptions. We illustrate how the method works and what is to be gained through three examples.

Keywords: Age-Period-Cohort analysis; identification problem; cohort effects; control group; difference-in-difference

Highlights:

- A design-based solution to the identification problem in Age-Period-Cohort Analysis
- Rests on the inclusion of a control group to identify cohort effects
- Adapts difference-in-difference estimation to the APC framework
- Identified by assuming either parallel age or parallel period effects
- Allows for analysis of birth-year cohorts and fully-factored age and period effects

1. Introduction

Studies in sociology, demography and political science have often used data from repeated cross-sections of individuals to estimate the effects of aging, period shocks, and early socialization experiences on attitudes or behaviors. The starting point for estimation is typically the Age-

Period-Cohort (APC) model. In the APC model, outcomes can vary across individuals as they age (aging or life cycle effects), across time for all individuals (period effects), and across individuals depending upon the year of their birth (cohort or generation effects). The well-documented problem with these models is that only two of these effects can be identified. Age (years since birth), period (year), and cohort (year of birth) are exact linear functions of each other: Age = Period – Cohort (introduction of this issue; Winship and Harding 2004).

In order to estimate the relative contribution of age, period and cohort effects one must make one or more assumptions. The most common assumptions relate to the grouping of cohorts and the adoption of a polynomial function to model the effect of age or time. Analysts will group individuals born across adjacent years into cohorts (e.g., those born between 1965 and 1980 might be called Generation X) instead of working with annual birth cohorts. They will use continuous age and time variables, instead of dummy variables for each age or time point, and impose a functional form on the relationship between these variables and the outcome. Under such assumptions, the typical equation used to estimate these effects is of the following form:

$$\begin{aligned} PID_{it} &= a + b_1 Cohort_2 + b_2 Cohort_3 + \ldots + b_{k-1} Cohort_k + \\ & c_1 Time + c_2 Time^2 + d_1 Age + d_2 Age^2 + e_{it} \end{aligned} \qquad (1)$$

In this setting, b_1 through b_{k-1} represent the difference in the level of party support between each new cohort and the oldest cohort, used as the reference category. This between-cohorts comparison is meaningful only if the cutpoints defining the cohort boundaries are defensible and the model accurately represents how period and aging effects operate.

This paper develops and illustrates an alternative approach to solving the APC identification problem. The approach introduces a control group to aid in the identification of cohort effects while also accounting for age and period effects. As such, it is primarily useful for readers interested in studying cohort effects within an APC framework. The approach yields two strategies for identifying cohort effects, each based on explicit and testable assumptions. Moreover, it identifies cohort effects without requiring that cohorts be grouped across adjacent birth years and without imposing assumptions about the functional forms of the age and period effects. In what follows, we first present an example to illustrate the logic of the approach. We then develop the core elements of the method. Our arguments build on prior work in sociology (Firebaugh and Chen 1995), economics (Card and Krueger 1994, Nielsen and Nielsen 2012, Pischke 2007), and statistics (Rosenbaum 1987). Finally, we illustrate our arguments through three empirical examples.

2. Using a control group: an example

Over-time data on the fraction of children who repeat second grade, adapted from Pischke (2007), are depicted in Figure 1. Until the 1960s, most German children started school in the spring. This changed in 1967, when the beginning of the school term was moved to the fall. This transition required two short school years, where time in school was compressed from 37 to 24 weeks. The upper curve of Figure 1 presents the average level of grade repetition for all second grade cohorts from 1962 to 1973. There seems to be an upward spike in the average level of grade repetition among those cohorts who were affected by this policy change, those whose school year ended in 1967 or 1968. The 1969 school-ending cohort also seems to have slightly higher repetition rates than the preceding and following ones, even though they had not been directly affected by the policy. These differences, however, are more modest than those that differentiate the 1967 and 1968 cohorts from the posterior cohorts. The pattern is relatively flat from 1971 to 1973 (Pischke 2007, Angrist and Pischke 2009).

In this case, we know that the treatment of a shortened school year was assigned to only two of the examined cohorts so there is no ambiguity about which cohorts should be distinctive if the treatment had an effect. Moreover, age is held constant by design. Still, however, estimates of the between-cohort differences rest on assumptions about how period effects are operating. Are they operating across younger cohorts and older cohorts to the same extent? Would we observe this gap even without this policy change?

To answer this question, Pischke (2007), leverages the fact that not all German schools shifted their school terms at the same time. Specifically, Bavarian schools started in the fall throughout the period. In the analysis, Pischke (2007) uses Bavaria as a control group, estimating the impact of the policy change through a within-cohort comparison. Rather than comparing the affected cohorts to those who entered the same schools earlier or later, he compares the affected cohorts with their same-school-year counterparts in schools where no policy change occurred. If the policy change caused an increase in grade repetition rates, this gap should be larger than the comparable gaps formed for cohorts entering school earlier or later. Crucially, period effects are controlled in the analysis to the extent that they operate on children in each set of schools to the same extent.

Although our examples come from a different theoretical background, the logic is very similar. Much is to be gained, we will argue, by adding a control group in the APC model. Doing so helps to test and justify the identification assumptions made in the analysis and yields new avenues for testing the robustness of the evidence for cohort effects.

3. Adding a control group in APC models: an augmented difference-in-difference approach

Most APC analyses look only at between-cohort differences within a targeted sample—a sample that contains a treated cohort as well as those born earlier and/or later. We add a control sample, which helps with the identification of cohort effects through a combination of both between- and within-cohorts comparison. The approach requires the identification of untreated (control/placebo) and treated subjects within a given cohort. It supplements the traditional between-cohort comparison with a within-cohort comparison.

To develop the idea, it is useful to distinguish between a cohort, a generation, and a generational unit.[1] The defining property of a cohort is the year (or interval of years) of birth. For a cohort to be called a generation, important attributes must be common to its members and distinguish its members from those of other cohorts. As developed by socialization researchers, this distinctiveness is fueled by differences in the political, social, economic, and/or cultural influences that different cohorts experience in their formative or impressionable years—typically identified as the years from late adolescence through early adulthood. A political generation is formed if cohort members develop a distinctive set of political beliefs, attitudes, and/or behaviors as a result of the context they experience during their impressionable years, and if those outcomes tend to persist as the individuals age (Braungart and Braungart 1984, Jennings and Niemi 1981, Mannheim 1952, Niemi and Sobieszek 1977, Schuman and Rodgers 2004, Sears 1983).

If, however, there is variation within a cohort in how individuals experience these (political, social, economic, and/or cultural) forces, or if only some cohort members but not others are affected, then generational (or generation) units come into being. Mannheim (1952)

[1] Alwin and McCammon (2003) provide an extended discussion of these and related concepts, which also brings in examples from the literature. Our abbreviated discussion here is designed to show how the use of a control group can be related to these ideas.

considered it likely that young people's social location (e.g. social class) would affect how they experienced the Zeitgeist of their formative years, resulting in multiple generational units within the same cohort. Thus, for example, experiencing the Great Depression in one's formative years would affect those from poor homes (one generational unit) differently than those from wealthy homes (a second generational unit). In Jennings' (2002) treatment, those from the high school Class of 1965 who, as young adults, participated in protest activities on college campuses formed a unique generational unit because of how the character of the times was filtered through those experiences. In short, a generational unit is a subgroup whose formative experiences are distinct from those of other members of their cohort.

Our approach to studying cohort effects requires that one classify individuals along two dimensions, as with the identification of generational units. The first dimension is cohort, which, as usual, distinguishes individuals by year (or interval) of birth. The second dimension is a characteristic that we call eligibility. Together, and as explained below, these dimensions determine whether or not a treatment is received.[2] Specifically, the conjunction of cohort and eligibility determines treatment status; one only receives the treatment if one is in a specific cohort and is eligible.

Using these terms, the process generating cohort effects can be summarized as follows. At some point in time the socialization environment of people who are eligible shifts. Those going through their impressionable years in the new environment would end up distinctive relative to those going through their impressionable years in the old environment. If the later environment is thought of as the treatment, then the treated group is the cohort going through

[2] This presents the simplest case, where individuals are either treated or they are untreated. The approach generalizes to more complex situations, with multiple treatments.

their impressionable years during the later period (the treatment period). This cohort will end up distinctive when compared the cohort socialized earlier. Importantly, another group of people—those who are ineligible—can also be identified as having gone through their impressionable years either before or during the treatment period. However, the socialization experienced by this group did not shift across the time frame. Those who were in their impressionable years during the treatment period lack the attribute (eligibility) that is required for the treatment of a new socialization environment to have been received.

Some notation is needed to develop the idea further. Let $Y_{E, C}$ refer to the potential outcome conditional on eligibility (E) and on cohort (C). E=1 if *i* is eligible and E=0 if *i* is not eligible. C=1 if *i* went through the impressionable years (IYs) during the treatment period and C=0 if *i* did not—i.e., if *i* went through the IYs earlier. When taken together, E and C generate the following four groups, only one of which is actually treated:

Y_{00}: not eligible, not in IY cohort during treatment period (untreated)

Y_{10}: eligible, not in IY cohort during treatment period (untreated)

Y_{01}: not eligible, in IY cohort during the treatment period (untreated)

Y_{11}: eligible, in IY cohort during the treatment period (treated)

Of the six possible combinations two are most interesting, namely the between-cohorts comparison of the treated and untreated units (Y_{11}-Y_{10}) and the within-cohorts comparison of the treated and untreated units (Y_{11}-Y_{01}). The former considers only the eligible, comparing those who experienced the treatment during their impressionable years (Y_{11}) to those whose impressionable years did not coincide with the treatment period (Y_{10}). The latter compares eligible individuals who experienced the treatment during their IYs (Y_{11}) to ineligible individuals from the same cohort (Y_{01}). It is clear from this classification that without the inclusion of the control group, our inference would result from the between-cohort comparison. What is gained by adding a control group is the opportunity for within-cohort comparison.

An example will help us develop how within-cohort comparisons become useful when applied to the APC framework. Following Firebaugh and Chen (1995), imagine we are interested in whether women going through their impressionable years (IYs) before 1920, when the 19th amendment to the U.S. Constitution was ratified, subsequently voted at lower rates than those going through their IYs post-1920. The expectation of a turnout difference across these cohorts follows from a difference in their early socialization experiences. The older cohort learned of politics as a sphere where voting rights were the exclusive to men, while the latter acquired an understanding of politics and citizenship in which women and men were able to participate on an equal (or more equal) footing. A between-cohorts comparison, the only possible comparison without the control group, would ask the following question: Are women who came of age before 1920 (call these C_i=1; i.e., the treatment period is pre-1920) less likely to vote than women who came of age later (C_i=0)? For a single time point, this question would imply estimating the following quantity of interest:

$$E[Y_i|C=1] - E[Y_i|C=0] = E[Y_{1i}-Y_{0i}|C=1] + \{E[Y_{0i}|C=1] - E[Y_{0i}|C=0]\}. \qquad (2)$$

The right-hand side of (2) is comprised of the average treatment effect of the amendment on turnout[3] plus a term (in {}) denoting the average difference in the turnout rates between those socialized before and those socialized after 1920 in the absence of the amendment. This term represents the bias caused by the non-random selection into the treatment condition. Although many different factors could be involved, an obvious confound is age. In any given election, the

[3] Strictly speaking, the first term is the average treatment effect on the treated.

average turnout rates in the absence of the amendment are likely to be higher for those socialized before 1920 than those socialized after 1920, simply because the former are older than the latter. In this example, ignoring the potential confounding role of age would probably lead to a downward bias in the estimate, for age is known to be positively related to voting.

Another problem involves the vague terms used to denote cohort status, i.e. whether women experienced the treatment (here, pre-19th amendment political socialization) during their impressionable years. When do the IYs start and when do they end? These are crucial questions for APC analyses. Should we code women C=1 if they were aged 30 or more by 1920, or if they were aged 25 or more by then, or perhaps code them as C=1 as long as they had achieved adulthood (age 21) by 1920? The results one obtains may well vary according to how the cohort boundaries are defined,

Both problems are ameliorated by the addition of a control group—a group of ineligibles, in the terminology laid out above. Regardless of the cutpoints used to define the treated IY cohort, the focus turns to *differences*—specifically, (1) how the IY cohort differs from the non-IY cohort among the eligibles, (2) how the IY cohort differs from the non-IY cohort among the ineligibles—and to *differences in differences*—specifically, how (1) differs from (2).

If we consider men to be the control group in this example, we then define eligibility E=1 if women and E=0 if men. The first question we address as a result of including this control group is the following: what is the difference in the likelihood of voting between IY (C=1) and non-IY (C=0) women? This difference is calculated as follows for one single point in time:

$$E[Yi|C=1,E=1,T=t] - E[Yi|C=0,E=1,T=t] = E[Y_{1i}-Y_{0i}|C=1,E=1,T=t] + \{E[Y_{0i}|C=1,E=1,T=t] - E[Y_{0i}|C=0,E=1,T=t]\}. \tag{3}$$

We then ask the same question for men:

$$E[Y_i|C=1,E=0,T=t] - E[Y_i|C=0,E=0,T=t] = E[Y_{1i}-Y_{0i}|C=1,E=0,T=t] + \{E[Y_{0i}|C=1,E=0,T=t] - E[Y_{0i}|C=0,E=0,T=t]\}. \quad (4)$$

This leaves a final question: is the inter-cohort gap for women larger than the inter-cohort gap for men? To answer this question, we need to subtract (4) from (3). Doing so we get δ:

$$\delta = E[Y_{1i}-Y_{0i}|C=1,E=1,T=t] - E[Y_{1i}-Y_{0i}|C=1,E=0,T=t] + \{E[Y_{0i}|C=1,E=1,T=t] - E[Y_{0i}|C=0,E=1,T=t]\} - \{E[Y_{0i}|C=1,E=0,T=t] - E[Y_{0i}|C=0,E=0,T=t]\} \quad (5)$$

The first difference is the effect of the treatment—being socialized during the treatment period—on women.[4] The second difference is the effect of being socialized during the treatment period for men. Since men were not actually treated during the treatment period, this second difference is a placebo effect, equaling zero. Put differently, this term drops out under the assumption that the control group was untreated during the treatment period. The remaining terms represent the difference across cohorts that would be expected in the absence of the treatment among women (in first set of brackets) and men (in second set of brackets), respectively. Since each cohort difference is calculated using data from one point in time, inter-cohort differences cannot arise from period effects but could arise from aging effects. If, however, we are willing to assume that aging effects are the same between men and women, then the last two differences cancel out and we are left with:[5]

$$\delta = E[Y_{1i}-Y_{0i}|C=1,E=1,T=t] \quad (6)$$

This is precisely the quantity of interest—the effect of the treatment on the treated group, in this

[4] The counterfactual to having been socialized during the treatment period is having been socialized in the non-treatment period, which is when the non-IY cohort was socialized. Also, as noted previously, the difference-in-difference estimand is the average treatment effect on the treated.

[5] For convenience, we assume here that any difference in Y between cohorts in the absence of the treatment is due to age, the most obvious confound. There are, of course, other selection threats, as discussed below. Some other influential variable, besides age, could differentiate the two cohorts in a way that is not comparable across groups.

example women who experienced pre-19th amendment political socialization during their impressionable years.

Crucial to this result is the assumption that aging effects are the same for men and women. This assumption is analogous to what is called the parallel trends assumption in over-time differences-in-differences analysis. The parallel trends assumption postulates that in the absence of an intervention, both the units that would receive the treatment and the units that would not receive the treatment would move in a similar fashion (Angrist and Pischke 2009: 228). Figure 2 presents a visual display of this assumption. The two parallel lines denote the average trend in some outcome of interest between two time points, t_1 and t_2. Although we would expect the same declining trend for both groups, some shock that affects only group 2 disrupts the trend, resulting in a much smaller (indeed zero) gap between the two time points. The assumption is that, absent this shock, the counterfactual trend for group 2 would be similar to that of Group 1.

Although in this case the parallel trends assumption applies to two cohorts instead of two time points, the logic is the same. It is assumed that differences across cohorts due to age would be the same for men and women. If the IY vs. non-IY gap among women is different than the IY vs. non-IY gap among men, this difference-in-difference identifies a cohort effect.[6] One can generalize the simple two-cohort case to the more common situation where multiple cohorts (and, hence, multiple difference-in-difference estimates) are of interest. The assumption of parallel age effects remains key regardless of how many cohorts one distinguishes or how the cohort cutpoints are defined.

[6] For simplicity, we use the terminology cohort effects rather than generational effects or, even more appropriate given our conceptualization, generational unit effects. Below, we address the question of whether cohort effects can be understood as causal effects.

The difference-in-difference approach laid out above yields an estimated cohort effect given data at any one time point. With repeated cross-sectional data on T time points, the process of estimation could simply be repeated T times, yielding T estimates of cohort effects.[7] If variation in early socialization experiences is consequential, one should find persistent rather than episodic or ephemeral differences in outcomes across the adult years. Or, as we illustrate below, one could specify an APC model that yields an estimated difference-in-difference across the time points represented in the data. In setting up that model, one need not assume a particular functional form for period effects, nor need one assume that period effects are parallel for the two groups. Put in the terms of our example, one could specify gender-specific year fixed effects. Moreover, there is no need to seek leverage for identification by combining individuals from adjacent birth years into cohorts. Finally, one need not make any particular assumption about the form of the aging effects. Age fixed effects could be employed, though they might eat up a lot of degrees of freedom and a more parsimonious specification may well be preferable. What one cannot do is interact age fixed effects with gender.

Importantly, however, there is another way to think about this difference-in-difference approach and its relation to APC analysis. Above we described the comparison as between cohorts at a *single point in time*. Alternatively, we could describe it as between cohorts at a *single point in the life cycle*. The analog of equation (5), in this context, is equation (7):

$$\delta = E[Y_{1i}-Y_{0i}|C=1,E=1, A=a] - E[Y_{1i}-Y_{0i}|C=1,E=0,A=a] + \{E[Y_{0i}|C=1,E=1,A=a] - E[Y_{0i}|C=0,E=1,A=a\} - \{E[Y_{0i}|C=1,E=0,A=a] - E[Y_{0i}|C=0,E=0,A=a]\} \quad (7)$$

[7] We assume that all observations are post-treatment—i.e., that researchers are looking for evidence of differences in the political attributes of adults depending on the character of their early socialization. If one faces a situation where both pre-treatment and post-treatment observations are available, time can be exploited as with the ordinary difference-in-difference set-up.

The first term is the treatment effect among women at a particular point in the life cycle while the second term is the comparable placebo effect among men. The remaining four terms now represent inter-cohort differences in the absence of treatment among women (first pair) and among men (second pair) when the cohorts are compared at the same point in the life cycle. In this comparison, age is held constant but time, of necessity, is not.[8] If, however, we are now willing to assume that the period effects are the same for women and men (while again assuming a zero placebo effect), we arrive again at the estimand δ, though in this case for fixed ages (A=a) instead of fixed time points (T=t): $\delta = E[Y_{1i}-Y_{0i}|C=1,E=1,A=a]$.

In this case, everything we described, above, with regard to age and period effects become flipped. In this comparison, we need not assume anything about parallel age effects but instead must assume parallel period effects. Under this assumption and with this revised comparison, we again can identify cohort effects, even when using birth-year cohorts. And while we can perform this comparison at multiple, discrete points in the life-cycle, we can also use an APC model applied to repeated cross-sectional data to estimate the average of those life-cycle comparisons. In this case, the model can allow age to vary freely across groups in a fully factored fashion but must impose the assumption that period effects (either fully factored or represented in a more parsimonious fashion) are the same across groups.

Thus, we arrive at two strategies of identification—one that holds age effects constant and one that hold period effects constant across the two groups. If both strategies yield comparable estimates of cohort effects, one gains confidence that neither identification

[8] For example, if cohort A was born in 1940 and cohort B was born in 1960, and one wished to compare the two when each was age 35, one would compare data on A from 1975 with data on B from 1995.

assumption is responsible for them.[9] As illustrated later, interpretation can be aided by comparing the results from these specifications to one in which both parallel period effects and parallel age effects are assumed.

APC models using each identification strategy also provide direct information about whether the parallel effects assumptions are reasonable. A naive way to assess the assumption of parallel age effects would be to estimate local regression curves relating the dependent variable to age for the two groups. A visual speculation of the resulting trends could indicate whether the assumption holds. However, age is confounded with cohort and thus this examination is prone to find violations when cohort effects are present. A better test for the parallel age effects assumption would evaluate the pattern in the model that includes cohort effects, period effects constrained to be the same across groups, and age effects allowed to vary across groups. Summarizing age with a polynomial function would help in better evaluating whether the assumption holds, but an analysis with age fully factored would yield good visual evidence. The exact same procedure can be repeated for the inspection of group-specific (gender-specific in our example) period effects.

In sum, with the addition of a control group one can estimate cohort effects by focusing on how between-cohort differences in one sample (with one or more cohorts treated) compare to between-cohort differences in a control sample (with no cohort treated). Within the APC framework, one can identify cohort effects by assuming either parallel period effects or parallel age effects. As a robustness check, both specifications should be employed, with the results

[9] The logic here is analogous to the use of multiple group testing. Multiple groups, which are known to differ in their values with regard to an unobserved covariate (Rosenbaum 1987), are separately compared with the treated group. If the effect does not differ between the two comparisons, there is more confidence that the results are not driven by the unobserved confounder. Here, we have two confounds but can only control for one of them at a time. See also Nielsen and Nielsen (2012), who discuss the utility of working with data on two samples when identifying mortality models.

compared to each other and to those from a specification that assumes age and period effects are both parallel. All of these specifications can be estimated using cohorts defined by birth year instead of using cohorts defined by birth year intervals, and when using fully factored age and period variables, though data sparseness and power considerations may prompt the use of more parsimonious models. Finally, adding a control group may also yield efficiency gains. If the parallel age effects and/or parallel period effects assumption is valid, then adding the control group will increase efficiency in the estimation of age and/or period effects; more data are brought to bear in estimating those parameters. If age/period effects are present, this, in turn, will reduce the residual variation in the dependent variable and enhance the efficiency of estimation of other parameters, including those representing cohort effects.

3.1 A note on treatments

Thus far we have used the example of the 19th amendment to lay out the relevant terms of our approach. In this case, the treatment variable, i.e. the socialization experienced by women, changed only once and at a particular moment in time (1920). In such a case, older cohorts (IYs pre-1920) will differ from younger cohorts (IYs post-1920), but differences are not expected within the two groups. Alternatively, the crucial socializing environment could vary more continuously over time. One example comes from Gibson and Caldeira's (1992) study of support for the Supreme Court among African Americans, which shows that support levels depend on whether people were going through their IYs when the Supreme Court was progressive on civil rights. Those experiencing young adulthood when Earl Warren served as Chief Justice (1953-1969) showed more support for the Supreme Court later in life than those born earlier or born later. In such cases, multiple cohorts, not just two cohorts, will be of interest. The pairwise difference-in-difference framework we developed still applies, but now with multiple cohorts (and differences) compared in pairwise fashion.

Our attempt to link the ACP analysis with the potential outcomes language of causal inference necessitates a final clarification point. In the potential outcomes framework, it is problematic to talk about causal effects of non-manipulable variables (see Holland 1986, with discussion). Cohort, as such, is non-manipulable. However, in APC analysis cohort is a proxy for the true causal variable—the socialization environment experienced by individuals in their formative or impressionable years—which is manipulable (though is unobserved). More generally, one can think of the APC analysis as providing information as to whether differences-in-differences consistent with the hypothesized causal effect are present.

3.2 What makes a good control group?

The utility of this approach depends crucially on the quality of the control group. Generally speaking, a good control group is one in which no unit receives the treatment but that is otherwise similar to the treatment group. In our framework, likewise, a good control group is one that contains no treated units but that is otherwise similar to the group that does contain treated units (i.e., the eligible group). The premise that control group members are untreated is what allowed us to characterize the effect of their being socialized during the treatment period as a placebo effect, expected to be zero.

Beyond that consideration, a minimal criterion for a good control group is that it contain the same cohorts. This is of course essential to implementing either of the difference-in-difference solutions to the APC identification problem, described above. In addition, groups who were born at the same time and, thus, shared many experiences by virtue of experiencing the same sociopolitical context at the same point in the life cycle, can be expected to share unobservables. Moreover, the fact that the groups are made up of the same cohorts may make it more plausible that life-cycle processes or period effects are operating in parallel fashion. The former are playing out over the same stretch of history, while the latter are operating on individuals who are matched with regard to age. As such, unmodeled interactions between age and period are also controlled. If, for example, events taking place in the post-treatment period affect younger people more than they affect older people, this could confound an analysis that excludes a control group. The addition of a control group that is matched in terms of cohort removes the confound.

Furthermore, no confound is introduced by differences across the groups that are independent of cohort. Put in terms of the 19th amendment example, estimates of cohort effects would be not be biased by an unmeasured attribute that caused men's turnout to be higher than women's turnout so long as that attribute affected men in the IY cohort to the same extent as it affected men in the non-IY cohort. However, confounds could arise if the distribution or effect of some uncontrolled variable varies across groups *and* cohorts.[10] As such, the ideal control sample is one that shares as much in common with the treatment sample as possible, and specifications that control for observables through regression adjustment or matching should be evaluated. As we discuss below, one may have choices about which control group to utilize, or be able to use more than one control group.

[10] In this situation, the between-cohort gap in $Y_{T=t}$ for the two samples, in the absence of the treatment, will not be the same even when age effects are parallel, hence we cannot simplify equation (5) to yield equation (6). Likewise, the between-cohort gap in $Y_{A=a}$ for the two samples in the absence of treatment will not be the same even if period effects are parallel.

4. Empirical examples

In what follows we present three empirical applications. Each of these examples serves a different purpose. We start with the analysis of the effect of the 19th amendment (or the shift in the socialization environment it engendered) on female turnout. This example provides a straightforward framework to test and evaluate the parallel age and parallel period effects assumptions. The second example relates to the Southern partisan realignment, a case that stimulates a discussion about how to examine the effect of an intervention when the exact time point in which it has taken place is unclear. Without a clear point of transition in the socialization experienced by citizens, any grouping of cohorts by birth interval becomes arbitrary. This example illustrates how the addition of a control group facilitates estimation with year-of-birth cohorts. Last, we examine the effect of being socialized during and immediately after the transition from an authoritative regime. In particular, we will examine whether youth from post-communist countries locate themselves more to the right on the ideological spectrum as a result of the association between the left and communism. The added feature of this example is that it provides a difficult case for the selection of a control group, showing the gains from incorporating multiple control groups.

4.1 Example 1: The 19th amendment and women's subsequent turnout rates

Thus far we have used the 19th amendment case in laying out the key elements of our approach. At this point, we use data from the American National Election Studies cumulative file to actually examine whether being socialized in a period in which women did not have the right to vote left a long shadow on women's turnout profiles.[11] Here, and in our later examples, our focus is on questions of method rather than the theoretical basis of expectations about cohort effects or the implications of the findings for political science.

We use the cohort categorization proposed by Firebaugh and Chen (1995), which yields 8 cohorts with the following birth years: <1896, 1896-1905, 1906-1915, 1916-1925, 1926-1935, 1936-1945, 1946-1955, 1956-1965. The main hypothesis in this analysis is that women from the two oldest cohorts were less likely to vote even after the passage of the 19th amendment than were women born and socialized later. We use all available ANES studies until 1990. Both age and period are fully factored. We start by assuming both parallel period and parallel age effects. We thus estimate the following linear probability model:[12]

$$Vote_i = a + \sum_{j=1}^{7} b_i(Cohort_{i,j+1}) + c_1 Female_i + \sum_{j=1}^{7} d_k(Cohort_{i,j+1})(Female) + \sum_{k=1}^{19} g_m Period_{i,k+1} + \sum_{m=1}^{77} b_m Age_{i,m-1} + e_i \qquad (8)$$

The reference categories are: the first (oldest) cohort; Year 1948; Age 35 years. Thus, j indexes cohort coefficients, j=1,…,7; k indexes year fixed effects, k=1,…,19; m indexes age fixed effects: m=1,…,34,36,…,77. The coefficients on the cohort variables indicate cohort differences among men, while the interactions between the gender variable (Female) and the cohort variables tell us whether the cohort differences found for women are different from those found for men. These interaction coefficients are the key quantities of interest. Although they only depict the female-

[11] This question was examined by Firebaugh and Chen (1995), in an analysis that considered women's turnout in relation to men's turnout and that was a source of inspiration for this paper. Their rationale for the use of a control group differs from the difference-in-difference rationale we have developed here.

[12] We used OLS just to keep things simple and since the form of estimation is not an issue for us here. Clearly, however, a Logit or Probit analysis would be more appropriate with a dichotomous dependent variable such as vote turnout. For a discussion of difference-in-difference estimation applied to data on more than two time points, in relation to models like (8), see Autor 2003, Lechner 2011.

male difference for cohorts 2-8 relative to cohort 1, the other comparisons that are of interest (e.g., 3-7 vs. 2) are implied. According to the main hypothesis, we would expect a significantly greater difference for women than for men when comparing the first two cohorts to the later ones.

Next, we relax the parallel effects assumptions, first for period effects and then for age effects. This is done by adding interactions between the period/age dummies and gender. We then evaluate the extent to which our original cohort effects change as a result of relaxing either of the parallel effects assumptions. As will be seen, such a comparison is useful because it provides unambiguous evidence about the sensitivity of the findings to the parallel effects assumptions. Importantly, our results here seem to be intact to the inclusion of gender-specific period or age fixed effects. In a final step, we compare our results with what we would have found if we had only focused on women, hence using no control group in the analysis

The left panel in Figure 3 graphically depicts the turnout results from the first analysis. Shown are the gender gaps in turnout for each cohort. Cohort effects are found by comparing these gaps.[13] A largely monotonic but non-linear ascending pattern can be detected. It indicates a decline in the turnout gap between men and women as we move to more recent cohorts. Consistent with expectations, the gender gap is highest among the two oldest cohorts. Moreover, no difference between the two oldest cohorts is evident, as expected. The gender gap among younger cohorts diminishes gradually at first but soon approaches a ceiling. No differences (and not much of a gender gap) are evident among the four youngest cohorts, born in 1926 or later, though all are distinctive from the two oldest cohorts (and perhaps the 3rd and 4th as well; our

[13] The difference in the gender gap between, say, cohort 1 and cohort 8 is equal to the difference in the cohort 1 vs. cohort 8 gap between men and women. Presenting the gender gaps allows one to visualize all the potential cohort comparisons.

focus here is not on statistical significance). Hence, the findings seem to confirm prior evidence (Firebaugh and Chen 1995) about cohort-based differences in female turnout that resulted from the different socialization experiences generated by the change in electoral suffrage.

The center and rightmost panels of Figure 3 present results from analyses that relax, in turn, the parallel period and parallel age effects assumptions. In both cases, the results seem to be largely unaffected. Only minor differences are observed, none of which poses any significant qualifications to the conclusions drawn when neither age nor period was allowed to vary with gender.

By implication, the similarity across the three sets of results in Figure 3 indicates that the parallel effects assumption is probably valid for both age and period. Direct evidence on this question comes from the models that allow these assumptions to vary, albeit not at the same time. When age effects are allowed to vary only 4 of the 77 interaction coefficients are statistically significant at $p<0.05$ (not correcting for multiple testing) and plots of how age is related to turnout for men and women shows no systematic variation. Similarly, when period effects are allowed vary by gender, almost all of the interaction coefficients are small in magnitude, only 1 of 19 is statistically significant at $p<0.05$, and plots of the period effects for men and women look very similar. (These plots are available in the Appendix; see Figures A1 and A2)

Finally, we consider how the finding presented in Figure 3 compares to those obtained when not using the control group—that is, when only analyzing women. Figure 4 presents results from three specifications, each of which included the 7 cohort dummies but where one included neither age nor period effects (left panel), one included only period fixed effects (middle panel) and one included only age fixed effects (right panel). These are comparable to the left, middle, and rightmost panels of Figure 3, which, respectively, held period and age effects constant across genders, allowed period but not age effects to vary across genders, and allowed age but not period effects to vary across genders.[14] To be sure, it is unlikely that one would just estimate a model of cohort effects without trying to account for aging and period effects. These comparisons do, however, serve to illustrate how the control group gives one leverage in estimating cohort effects.

The variation in the results across specifications in Figure 4 is notable, as are the contrasts with those given in Figure 3. A pronounced curvilinear pattern emerges in the first set of results, with turnout rates among the two oldest cohorts on a par with those of the youngest cohorts, though lagging behind those for the middle cohorts. The pattern is similar when period, but not age effects are taken into account (middle panel). The implication is that uncontrolled aging effects are generating this curvilinear pattern. The estimates in Figures 4a and 4b did not yield this curvilinear pattern precisely because the age confound is also present among men, our control group. In the rightmost panel of Figure 4, where age but not period effects were included, cohort differences largely disappear. Visually, the only cohort that appears distinctive is the youngest. Here it is uncontrolled period effects that are responsible for results that diverge from those found in the rightmost panel of Figure 3.

As this exercise illustrates, adding a control group and estimating cohort effects through a difference-in-difference approach can provide real leverage in an APC analysis. Without a control group, the solution of excluding age effects or of excluding period effects is not a viable

[14] When period (and/or age) effects are held constant across the groups, including vs. excluding the period (and/or age) variables has no material consequence for the difference-in-difference estimates. Thus, for example, if we drop the period and age fixed effects from the model that generated the results presented in the leftmost panel of Figure 3, such that we only include the cohort variables, gender, and the interactions between the two, the results are almost identical.

one. As in this case, it can yield results for cohort effects that are sensitive to the choice of specification, that fail to conform to theoretical expectations, and that are vulnerable to rival interpretations invoking the excluded variable. The comparable solution in the control group design is not to exclude an effect (age or period) but to assume that at least one of those effects operates to the same extent on both groups. This is not a fail-proof solution to the APC identification problem, but can be of great value when the parallel effects assumption is warranted.

4.2 Example 2: The Southern realignment revisited

A key and well-documented feature of American politics is the gradual but steady movement of the previously solid Democratic South to the Republicans. Although the origins of this shift are debated (e.g., Carmines and Stanley 1990, Osborne, Sears and Valentino 2011, Stanley 1988), the bulk of this literature gives pride of place to racial attitudes and civil rights policies (see, illustratively, Valentino and Sears 2005). What is important for our purposes is that while the most prominent sources have been identified, the exact departure and closing points of this process are largely unknown. Most empirical analyses locate the beginning of this realigning trend in the early 1960s, as also manifested by George Wallace's strong presidential candidacy.

One of the debates about the Republican realignment in the South is whether this change has been driven by conversion or by generational replacement. Have Southerners switched to the Republicans in a rather homogenous fashion, responding in this way to the ever growing cross-pressure between their party and their racial and other issue attitudes? Or do new cohorts, whose political experience starts after the 1950s, largely drive this shift?

The aim here is not to challenge prior studies focusing on the issue of the Southern realignment. We employ this case because it provides an interesting example of how our approach works when it is not clear when the intervention takes place or how long it lasts. We again use the ANES data, but we now adopt a more agnostic attitude with regard to the construction of cohorts. In particular, we refrain from clustering individuals in more than one year-of-birth category. The only exception in this rule is the reference category, namely all individuals born before 1880.[15] Hence, a long series of dummies, each one denoting respondents born at any given year from 1881 to 1992, is used to capture cohort effects. Both time and age are also controlled in a fully factored fashion.

The important question, once again, is the choice of the control group. Here an obvious candidate for this role is the non-South, i.e. all ANES respondents residing in non-Southern states. Following the previous strategy, we first estimate a model interacting all year-of-birth dummies with South, a dummy that switches on for Southern respondents.[16] We plot differences in differences in this case, specifically, the difference in the probability of being a Democrat between a Southern and a non-Southern member of the cohort in question compared to the same South vs. Non-South difference for the oldest cohort, used as the reference category.[17] The results from this exercise, holding period and age effects (fully factored) constant across the regions, appear in the left-most graph of Figure 5. A local regression curve traces the conditional expectation across all cohorts, sorted from the oldest to the youngest. A largely monotonic but nonlinear pattern is found, with successive cohorts less Democratic in their party identification until we reach those born in the 1970s. The next two graphs (6b and 6c) relax the parallel periods

---

[15] This is done because until that year of birth, we had less than ten cases in each cell.

[16] South includes the following states: Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, Missouri, North Carolina, South Carolina, Tennessee, Texas, and Virginia.

[17] The dependent variable is a dummy that denotes Democrats (including leaners). For simplicity, OLS estimation was used.

and aging assumptions, respectively. Results not shown reveal sizeable period effects, which vary between the South and the North, but virtually no age effects in either region. Because of this, the results from relaxing the parallel period effects assumption (see middle panel) differ from the first set, while those from relaxing the parallel age effects assumption (see rightmost panel) do not. Once the parallel period effects assumption is relaxed, the cohort differences are attenuated, but we still find a monotonic trend with the greatest gap involving those who were born in the late 1940s and who came of age in the late 1960s.

Two points about this case should be underscored. First, and unlike the previous example, we do see variation in the results depending upon the assumption(s) employed. This, in turn, is due to the significant regional variation in period effects that is exposed when such variation is allowed by the model. Fortunately, no evidence of regional variation in age effects is found.[18] Thus, we can confidently ignore the first and third set of results and focus our interpretations on the second, which are generated using the model that allows regional variation in period but not age effects.

Second, as we have already pointed out, the inclusion of a control group gives one leverage for identifying cohort effects no matter how cohorts are defined. There is no need to gain leverage by using birth-interval rather than birth-year cohorts. As demonstrated here, analysis of year-of-birth cohorts can be useful when there is uncertainty about when the relevant causal variable changed or how that shift (or those shifts) affected individuals depending upon their age at the time. The results can be the basis of inductive reasoning about the formation of unique generations or generational units. At the same time, if one does have an explicit

[18] Theory and substance is not our focus, but this pattern is not surprising. There is no particular reason to believe that aging effects on party identification would vary across the regions (if one expected them to exist in the first place), while it is more plausible to believe that partisan forces play out differently across the regions.

hypothesis about where cohort boundaries should be drawn, year of birth cohorts can be used to put that hypothesis to a test.

4.3 Example 3: Ideological leanings in post-Communist democracies

Our third example is theoretically motivated by a common pattern found in the literature on ideological positioning in post-communist democracies. Most surveys show that people in Eastern European countries are on average located more to the right than their counterparts in established democracies.[19] One conjecture is that right-wing support in Eastern Europe is being fueled by right-leaning new (younger) cohorts who entered adulthood during or after the democratic transition. The central idea is that the newer cohorts would be more right-leaning than older cohorts because they were not socialized during an era of Communist ascendency (Pop-Eleches and Tucker 2010, Linek 2011). Again, we have no intention to contribute to this literature. We address this question because it poses an interesting challenge for our approach to APC analysis—namely the choice of a control group.

We use the European Election Study (EES) surveys of 1989, 1994, 1999 and 2004. With the exception of Eastern Germany, all other cases classified as new democracies are included only in 2004. This is not ideal, and it introduces estimation complications that an analyst would ordinarily need to take into account. We, however, set these issues aside, as they do not bear on the utility of the example for our purposes.

To investigate this question, we place respondents in six cohorts (birth years <1920, 1920-1934, 1935-1949, 1950-1964, 1965-1987, 1980>), for simplicity. Age is controlled for with a linear, a quadratic and a cubic term. Period effects are accounted for with the EES study

[19] In data from the European Election Study, described below, a regression of left-right placement (scored 0=most left to 1=most right) on a dummy denoting respondents coming from East European countries generates a coefficient of .367 (.051).

dummies. The results using only the East European cases appear in the first (left-most) graph of Figure 6. Contrary to expectations, each incoming cohort seems to be located more to the left than the oldest cohort, born before 1920.

If we now wish to examine the same pattern using a control group, the important question is which group to choose as a control. Intuition suggests the choice of respondents to the same surveys from established democracies. If there is a general shift to the right among younger cohorts in new democracies, this should be evident in inter-cohort comparisons that juxtapose respondents from new democracies with those from established democracies. The still unresolved question, however, is which exact country to use for this role. Here, theory and existing research is important. Prior knowledge of the available cases should be employed with the aim of approximating a most-similar design. That said, the opportunity to use more than one country should also be taken, as doing so will reveal the robustness of the results to the choice of a control group. This strategy constitutes a very valuable robustness check to potential treatment-group unobservables in difference-in-difference (Card and Krueger 1994) and multiple control-group designs (Lu and Rosenbaum 2004).

We first estimate cohort effects using all respondents from established democracies as the control group. The results appear in the second (right-most) graph of Figure 6.[20] The pattern gives credence to the idea that young cohorts in Eastern Europe are less enamored of leftish politics than were those who were socialized under communist rule. As a next step, we assess the robustness of these results to the choice of different control groups, using respondents from each established democracy in turn. If the findings arise because the younger cohorts were not

[20] Note that this analysis assumes parallel period and parallel age effects. Relaxing either assumption diminishes the magnitude of the inter-cohort differences.

socialized in an era of Communist dominance, or because the left was delegitimized in their eyes as a result of its connection to the prior, non-democratic regime, the choice of the control group should not matter very much in the final results. This is because none of the countries used as controls experienced an authoritative regime linked to the left.[21] Figure 7, which shows all possible tests, indicates that the findings seem to hold with almost no exception, regardless of the country chosen as a control group. In all cases, the youngest cohort registers the most extreme-right scores. The use of multiple control groups boosts confidence in the robustness of the findings.

5. Conclusion

Over the last decades a large variety of applications in social sciences, epidemiology, and other disciplines have employed Age-Period-Cohort models to disentangle the roots of change in various phenomena of interest. The identification problem has typically been addressed by grouping birth-year cohorts, constraining the functional form of period and/or age effects, or by applying other analytical techniques based on assumptions about how various quantities of interest work. The design-based approach that we have described and illustrated is a promising alternative when cohort effects are of special interest. The approach introduces a control group and adapts differences-in-differences estimation to the APC framework, yielding two alternative identification strategies. The assumptions underlying these strategies are transparent, testable, and in some applications may even be theoretically expected to hold. Selection threats remain to

[21] Spain, Portugal and Greece may be problematic as control groups because they have transitioned away from authoritarian regimes linked to the right. This could have produced cohort effects in the control group, with younger cohorts more left-leaning than older cohorts. In our analysis, this would generate the appearance of cohort effects in the eastern European countries, even if younger cohorts there did not especially lean to the right. The fact that our cohort effect estimates are consistent across all of the countries used as controls renders this possibility less worrisome.

hinder causal inference, but there are reasons to believe these are mitigated by the inclusion of the control group.

Cohort effects are expected to arise if exogenous events have led one or more cohorts within a population to receive the treatment of a new socialization context. Estimating these effects is hampered by the fact that cohort is fully confounded with age when cohorts are compared at any one point in time, and fully confounded with period when cohorts are compared at any one point in the life cycle. Adding a control group—a group in which no individual was treated but that is otherwise as similar to the first group as possible—provides a way out of this conundrum. Instead of looking at the difference between treated and untreated cohorts in a single sample, one looks at how this difference compares to a second difference—that between the same cohorts in the control sample. When estimating this quantity holding time constant, age effects are no longer confounding if one assumes that they affect each group in a parallel fashion. Likewise, when holding age constant, period effects are no longer confounding if one assumes they operate in parallel across the groups. In an APC framework, thus, cohort effects are identified either if age is constrained to have the same effect across groups or if period effects are constrained to be the same across groups. As we have illustrated, these assumptions can be tested directly, if imperfectly, by examining the evidence for non-parallel age or period effects when each assumption is relaxed in turn, and can be tested indirectly by comparing cohort effect estimates obtained from these specifications to those obtained when both parallel age and parallel period effects are assumed.

To implement this approach, researchers must find a control sample for which Y data are available that is comparable to the sample from the actual population of interest—save for the shock that is thought to have generated cohort effects in the latter. It is the introduction of the control group into the APC model that yields difference-in-difference estimation of cohort effects while accounting for period and age effects. Because of the within-cohort comparisons that are integral to the estimation strategy, it is essential that the control sample includes the same cohorts as the initial sample. A successful choice for the control group should otherwise be based on the logic governing Mill's most-similar-cases design: except for the fact that no cohort is treated, the control group should be as similar as possible to the target sample. Introducing a control group does more than provide solutions to the APC identification dilemma. Doing so can improve the specification of the APC model because the two groups are bound to share unobservables and because unmodeled age*period interactions are implicitly controlled.

An important criticism that can be made with regard to this approach is that it may be difficult to find a suitable control group. It is crucial that the treatment experienced by one or more cohorts in the target sample was not also experienced by the comparable control cohorts. Even so, unmodeled cohort effects operating within the control sample could confound the analysis, as could uncontrolled variables that vary across cohorts and groups. Given that explicit empirical evidence on these issues may often be lacking, we urge researchers to engage in multiple control group testing. Doing so, one can better rule out unobserved or unmodeled confounders that could be driving results.

References

Alwin, D.F., McCammon, R.J. 2003. Generations, Cohorts, and Social Change. In Mortimer, J.T. Shanahan, M. J. (Eds.), Handbook of the Life Course. New York: Kluwer Academic/Plenum Publishers, 2-49.

Angrist, J., Pischke J.-S., 2009. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton: Princeton University Press.

Autor, D., 2003. Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing. Journal of Labor Economics 21, 1-42.

Braungart, R., Braungart, M., 1984. Generational Politics. Micropolitics 3, 349-415.

Card, D., Krueger, A., 1994. Minimum Wages and Employment: A Case Study of the Fast Food Industry in New Jersey and Pennsylvania. American Economic Review 84, 772-84.

Carmines, E., Stanley, H., 1990. Ideological Realignment in the Contemporary South: Where Have All The Conservatives Gone? In Steed, R., Moreland, L., Baker, T. (Eds.) The Disappearing South? Studies in Regional Change and Continuity. Tuscaloosa, AL: University of Alabama Press, 21-33.

Firebaugh, G., Chen, K., 1995. Voter Turnout of Nineteenth Amendment Women: The Enduring Effect of Disenfranchisement. American Journal of Sociology 100, 972-96.

Gibson, J., Caldeira, G., 1992. Blacks and the United States Supreme Court: Models of Diffuse Support. Journal of Politics 54, 1120-45.

Holland, P., 1986. Statistics and Causal Inference. With Discussion. Journal of the American Statistical Association 81, 945-70.

Jennings, K., 2002. Generation Units and the Student Protest Movement in the United States: An Intra- and Inter-generational Analysis. Political Psychology 23, 303-24.

Jennings, K, Niemi, R., 1981. Generations and Politics. Princeton: Princeton University Press.

Lechner, M., 2011. The Estimation of Causal Effects by Difference-in-Difference Methods. Discussion Paper no. 2010-28, Department of Economics, University of St. Gallen.

Linek, L., 2011. The Impact of Past Events on Current Electoral Behaviour. Age-Period-Cohort Analysis of Czech Communist Party Voters. Paper presented at the ESRA Conference, Lausanne, July, 2011.

Lu, B., Rosenbaum, P., 2004. Optimal Pair Matching with Two Control Groups. Journal of Computational and Graphical Statistics 13, 422–434.

Mannheim, K., 1952 [1926]. Essays on the Sociology of Knowledge. London: Routledge and Kegan Paul.

Nielsen, B., Nielsen, Jens, 2012. Identification and Forecasting in Mortality Models. Working Paper. University of Oxford.

Niemi, R., Sobieszek, B., 1977. Political Socialization. Annual Review of Sociology 3, 209-33.

Osborne, D., Sears, D., Valentino, N., 2011. The End of the Solidly Democratic South: The Impressionable Years Hypothesis. Political Psychology 32, 81-108.

Pop-Eleches, G., Tucker, J., 2010. After the Party: Legacies and Left-Right Distinctions in Post-

Communist countries. Paper presented at the 2010 Annual Meeting of the American Political Science Association, Sept 2-5, 2010, Washington, DC.

Rosenbaum, P., 1987. The Role of a Second Control Group in an Observational Study. Statistical Science 2, 292-306.

Pischke, J.-S., 2007. The Impact of Length of the School Year on Student Performance and Earnings: Evidence from the German Short School Years. Economic Journal 117, 1216-42.

Schuman, H., Rodgers, W., 2004. Cohorts, Chronology, and Collective Memories. Public Opinion Quarterly 68, 217-254

Sears, D., 1983. The Persistence of Early Political Predispositions: the Roles of Attitude Object and Life Stage. In Wheeler, L., (Ed.) Review of Personality and Social Psychology, 4. Beverly Hills: Sage Publications, 79-116.

Stanley, H., 1988. Southern Partisan Changes: Dealignment, Realignment, or Both? Journal of Politics 50, 64-88.

Valentino, N., Sears, D., 2005. Old Times There Are Not Forgotten: Race and Partisanship in the Contemporary South. American Journal of Political Science 49, 672-88.

Winship, C., Harding, D., 2008. A Mechanism Based Approach to the Identification of Age-Period-Cohort Models. Sociological Methods and Research 36, 362-401.

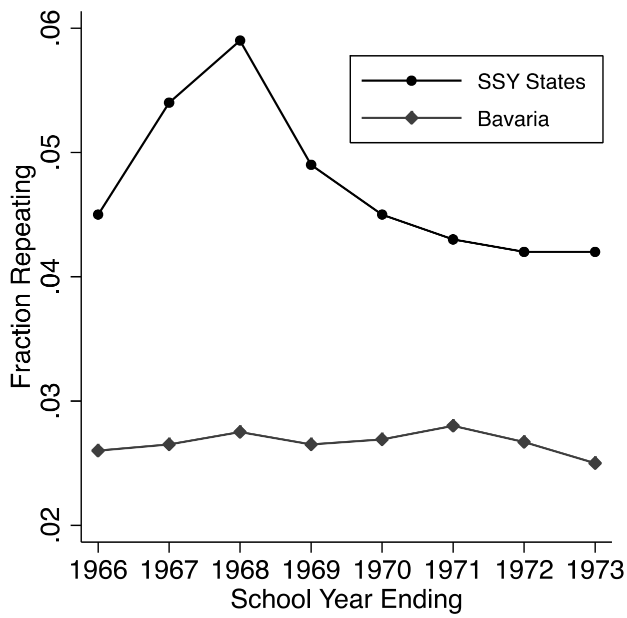

Figure 1: Use of a Control Group in Pischle's (2007) Study of Retention Rates

Note: SSY stands for Short School Years.

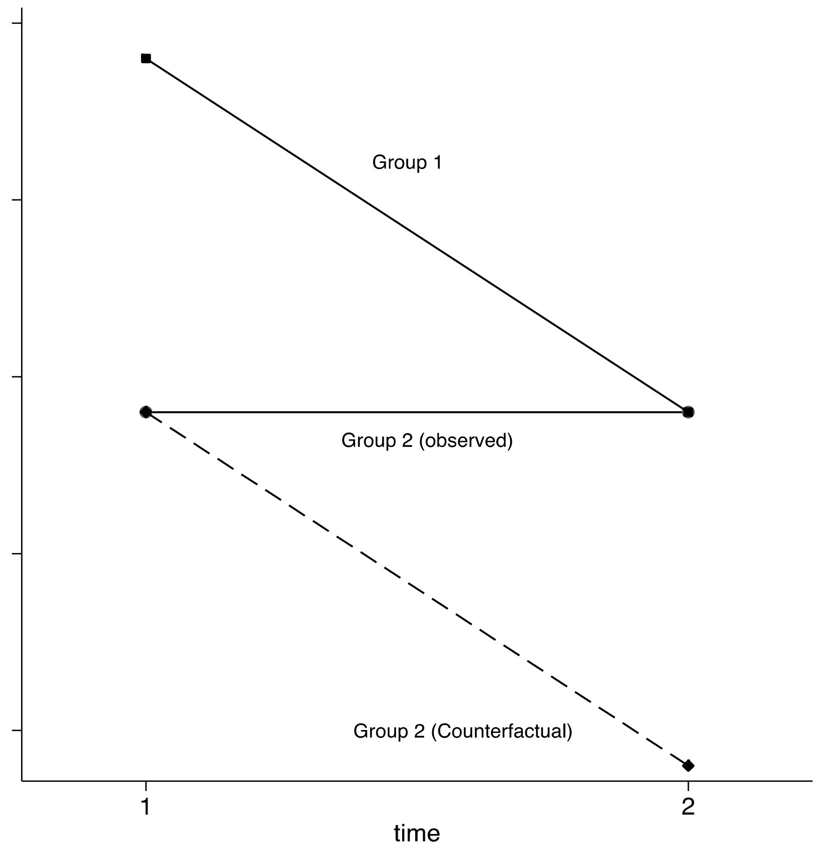

Figure 2: The parallel trends assumption in difference-in-difference analysis

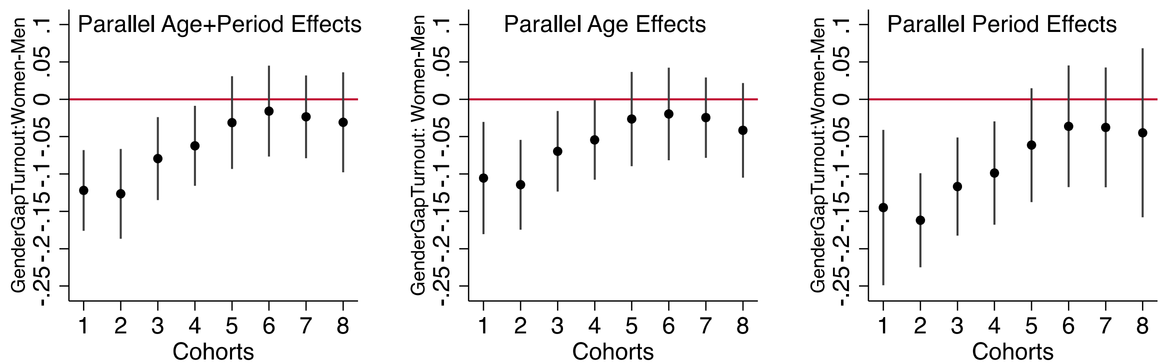

Figure 3: Cohort effects in 19th amendment example, estimated three ways.

Note: Shown in each panel is the estimated gender gap in turnout (male turnout – female turnout) for each of the eight cohorts introduced in the text. Confidence intervals are depicted by the vertical bars. Comparing the gaps across cohorts gives the cohort effects. The leftmost panel reports results from a model that included age and period fixed effects. The model generating results shown in the middle panel allowed interactions between gender and period but not age, while that for the rightmost panel allowed interactions between gender and age but not period.

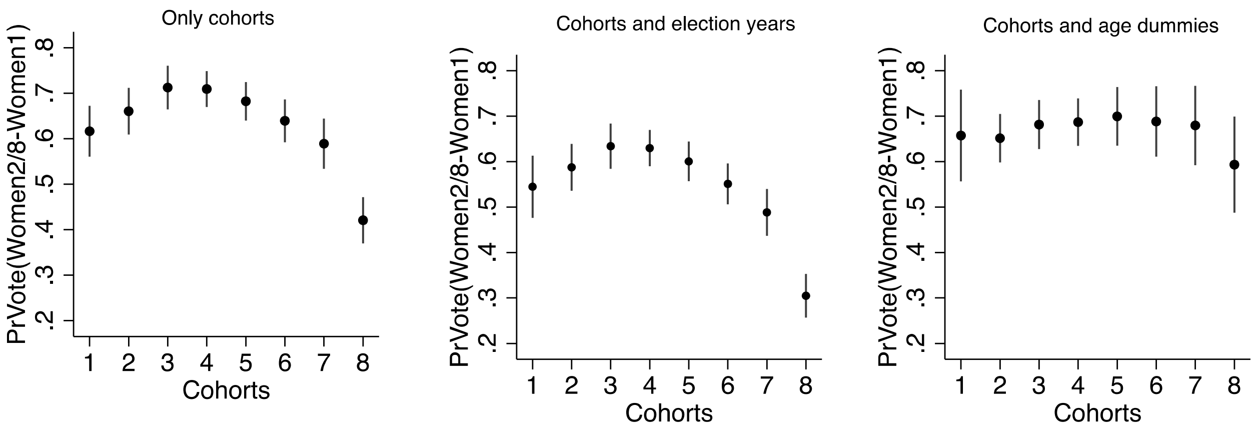

Figure 4: Cohort effects in the 19th Amendment example—analyzing women only

Note: Entries are predicted probability of voting, as estimated with OLS. The vertical bars show the estimated 95% confidence intervals. The model generating results shown in the leftmost panel included neither period nor age dummies. The model producing results shown in the middle panel included period but not age dummies, while that producing the rightmost results included age but not period dummies.

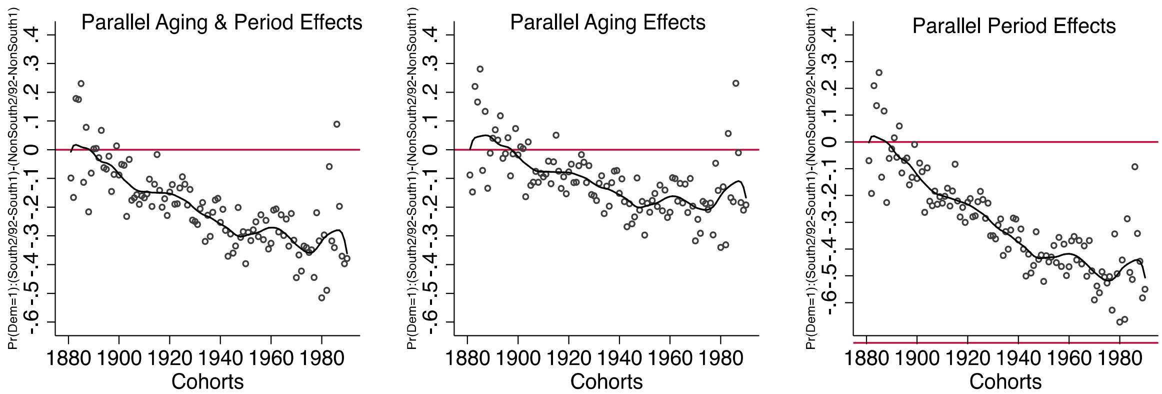

Figure 5: Cohort effects in the Southern realignment example

Note: The figure shows inter-cohort difference in the predicted probability of identifying as a Democrat among Southerners relative to the inter-cohort difference among non-Southerners. The inter-cohort difference is that between the cohort indicated on the X axis and the oldest cohort.

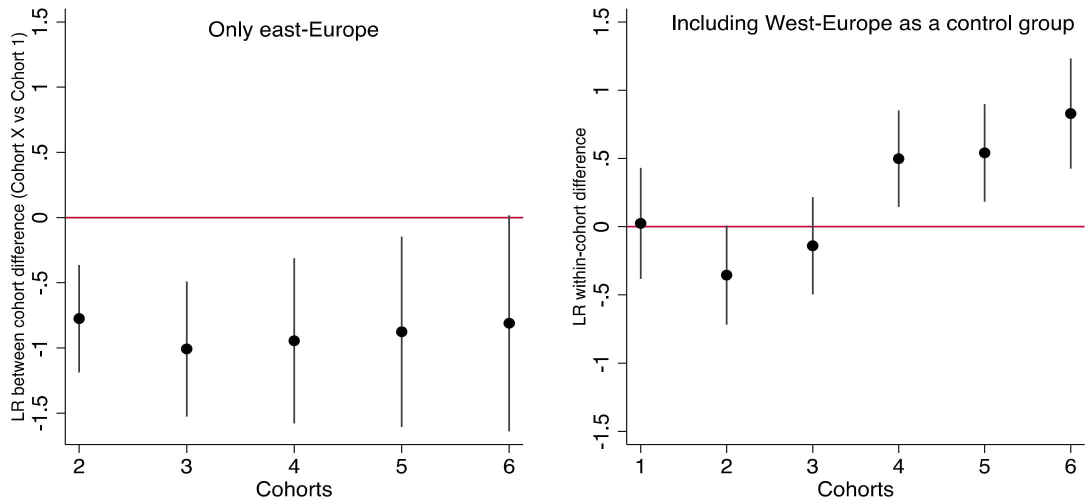

Figure 6: Left-Right placement in new democracies:
Cohort effect estimates with and without a control group

Note: Estimates in the left panel show the difference between the named cohort and the oldest cohort among those in new democracies. Estimates in the right panel show the difference-in-difference estimates—i.e., the gap between the named and oldest cohort among those in new democracies relative to that same gap among those in established democracies.

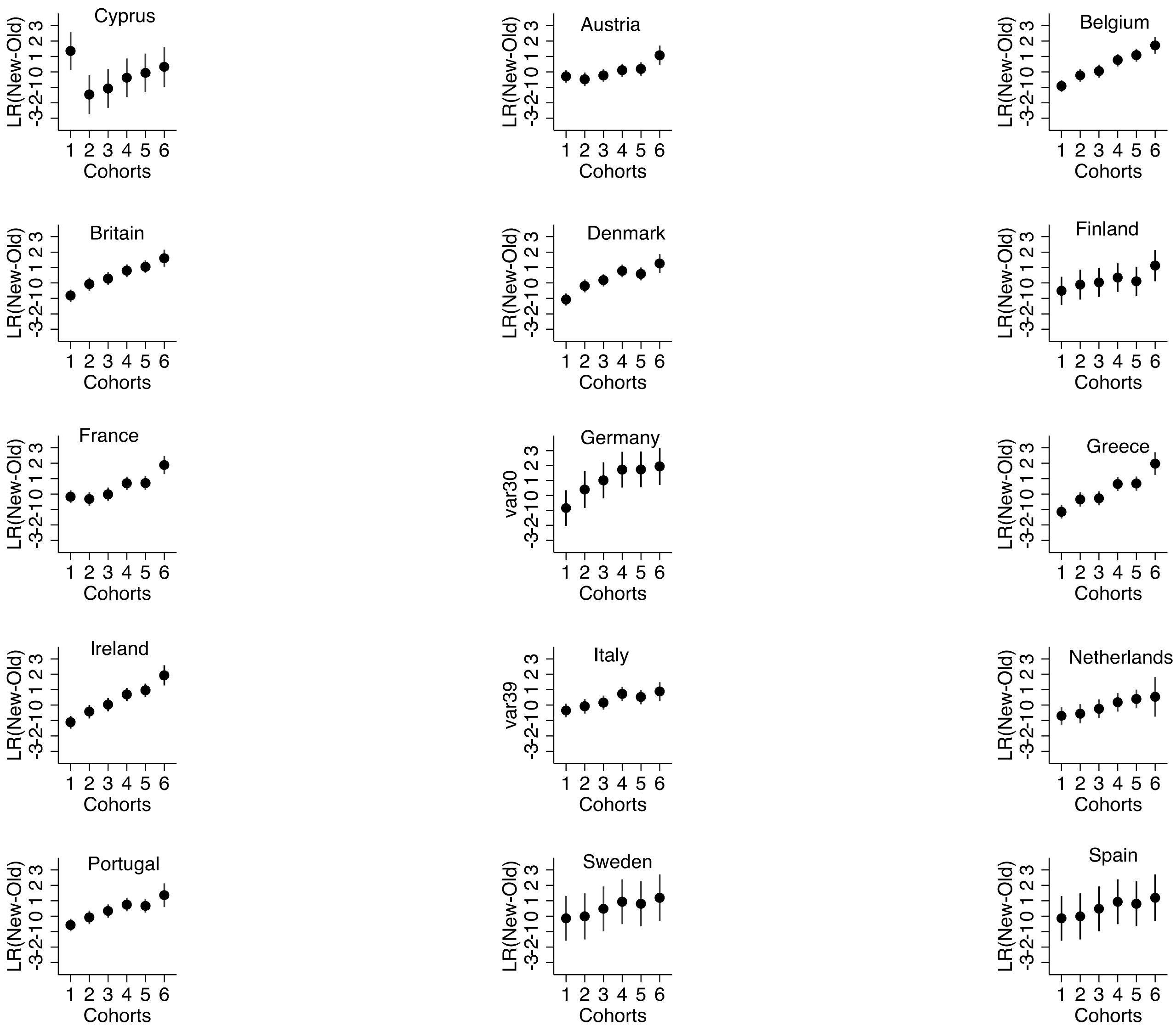

Figure 7: Left-right placement in new democracies: Varying the control group.

Figure A1: Estimated Period Effects for Men in Women (19th Amendment Example)

Note: Estimates come from the model that allowed period effects to vary by gender but constrained age effects to be parallel. Age (=35) and cohort (=oldest) are held constant.

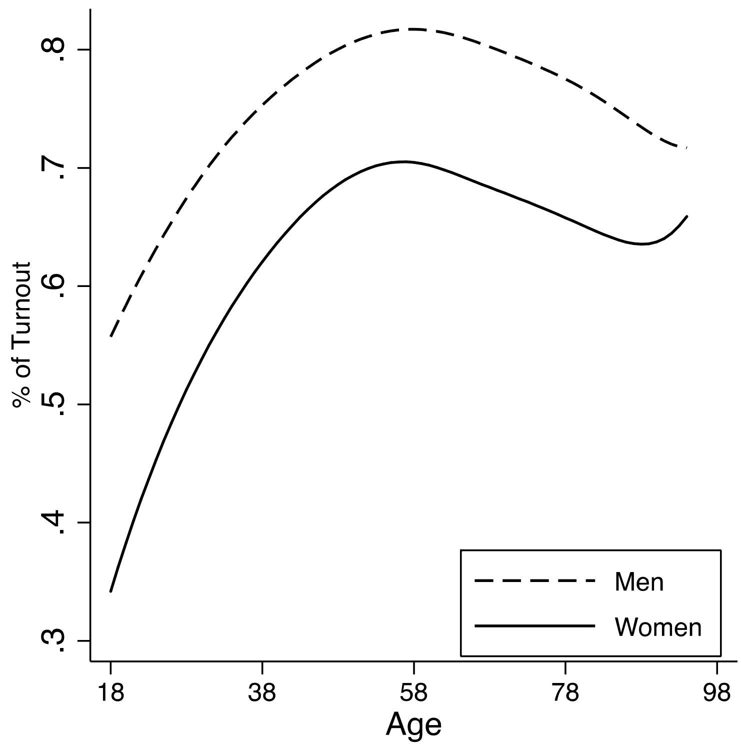

Figure A2: Estimated Age Effects for Men in Women (19th Amendment Example)

Note: Shown are smoothed curves summarizing how the predicted probability of voting varies with age among men and women. Estimates come from the model that represented age effects with dummy variables and allowed those age effects to vary by gender but constrained period effects to be parallel. Period (=0) and cohort (=oldest) were held constant.