Local Average Cohort Effects

Nicholas Reynolds
University of Essex *

October 18, 2024

Abstract

Many economic theories predict the existence of cohort or vintage effects: that workers, firms, or capital 'born' in a particular year have fixed characteristics that shape their outcomes throughout their lifecycle. The usual approach models cohort, age, and year effects as additively separable. This paper considers identification and inference in a nonseparable model which allows for unrestricted interactions between fixed features of cohorts and external factors which vary by age and year. I define Local Average Cohort Effects as the difference between the average outcomes for a cohort and what the *counterfactual* average outcomes of the cohort born one year earlier would have been had they experienced the same external age-by-year factors. I show that the estimator proposed by McKenzie (2006) for the additively-separable model identifies differences in Local Average Cohort Effects under the assumption that shocks to neighboring ages in neighboring years are exchangeable (as-if-randomly-assigned). This assumption also suggests a permutation-based test which provides inference with exact size in expectation. I apply the approach to show additional evidence of the (relative) decline in health and human capital of Americans born after 1947 documented in Reynolds (2024).

*Nicholas.Reynolds@essex.ac.uk. I thank Jesse Shapiro for originally pushing me to think more carefully and creatively about identification in this setting. I also received helpful comments from Ken Chay, Neslihan Sakarya, and Marco Francesconi. Portions of this work were previously included in the working paper "The Broad Decline of Health and Human Capital" (2023).

1. Introduction

Many economic theories predict the existence of cohort or vintage effects: that workers, firms, and capital 'born' in a given year have fixed characteristics that will shape their productivity (or other outcomes) throughout their lifecycle. Sometimes these cohort effects are of independent interest (eg. Hall, 1968; Card and Lemieux, 2001; Porzio et al., 2022). Other times age effects, such as lifecycle wage or consumption profiles, are the object of interest and cohort effects along with year effects are nuisance terms.[1]

The usual approach models cohort, age, and year effects as additively separable. A well-known identification problem arises because of the collinearity of cohort, age, and year (Hall, 1971). The linear trends in each of these effects cannot all be identified, but second-differences in each effect can, and therefore one additional linear restriction can pin down the full set of effects (Hall, 1971; Deaton, 1997; McKenzie, 2006). Perhaps less well appreciated is that allowing even basic parametric interactions between any of the two dimensions — such as allowing the impact of age to vary by cohort — introduces additional collinearity and exacerbates the identification problem (Heckman and Robb, 1985). To my knowledge, no work has considered the implications for identification and inference of introducing more general forms of heterogeneity and non-separability.

This paper introduces to the cohort effect setting the type of heterogeneity commonly considered for treatment effects. I specify a nonseperable model which allows for unrestricted interactions between fixed features of cohorts and external factors which vary by age and year. The model highlights that cohort effects answer counterfactual questions about what a given cohort would have earned had they faced different external factors. Identifying these generalized cohort effects is therefore akin to the "Fundamental Problem of Causal Inference" (Holland, 1986) in that it requires assumptions which allow identification of unobserved potential outcomes.

I define Local Cohort Effects as the difference between the observed outcomes for a cohort, C, in a given age and year, and what the outcomes of the cohort born one year earlier would have been if they had experienced the same external age-by-year factors as cohort C. I call the average of these across years Local Average Cohort Effects. As an example, consider testing the claim of my earlier paper Reynolds (2024) that there has been a decline in the health and human capital of Americans born after 1947, relative to the prior cohort trend. A relevant Local Average Cohort Effect in this setting is: the counterfactual difference in wages between what the 1948 cohort would have earned had they faced the set of external age-by-year factors that the 1947 cohort faced and what the 1947 did earn. If this effect is negative it implies that the 1948 cohort had more human capital than the 1947 cohort, at least along the dimensions which were relevant in the labor market

[1] See for example Mincer (1974); Heckman et al. (1998); Lagakos et al. (2018); Gourinchas and Parker (2002)

conditions which the 1947 cohort experienced. This definition therefore takes seriously the idea that the human capital of a cohort (or the fixed features of a particular vintage of capital) may be multidimensional and therefore it's impact may differ depending on the external factors prevailing in a given year for those of a given age.

I show that the estimator proposed by McKenzie (2006) for the additively-separable model identifies the Difference in Local Average Cohort Effects under the assumption that shocks to neighboring ages in neighboring years are exchangeable. In the context of the above example, this means one can identify the difference between i) the counterfactual difference in wages between what the 1948 cohort would have earned, had they faced the external factors the 1947 cohort faced, and what the 1947 cohort did earn, and ii) the counterfactual difference in wages between what the 1947 cohort would have earned, had they faced the external factors the 1946 cohort faced, and what the 1946 cohort did earn. If this is negative it implies that the change in human capital between the 1947 and 1948 cohort is smaller than that between the 1946 and 1947 cohort, again at least along the dimensions which were relevant in the particular "local" labor market conditions specified.

The assumption that shocks to neighboring ages in neighboring years are exchangeable deserves further comment. As an example, it would implies that the pair of external shocks affecting the wages of 30 year olds in 1970 and 1971 and the pair of external shocks hitting 31 year olds in 1970 and 1971 could be permuted or "swapped" without changing the joint distribution of any of the external shocks. It is "as-if-randomly-assigned" whether the set of shocks in 1970 and 1971 hit either 30 or 31 year olds. This allows for unrestricted dependence in the shocks over time hitting a given age, as well as non-random shocks to larger groups of ages in a given year. What is assumed random is the very local difference in external shocks impacting those one year apart in age.

This assumption also suggests a permutation-based test which provides inference with exact size in expectation. For example one can test whether the Difference in Local Average Cohort Effects described above is negative.

I apply this approach to test whether there has been a decline in health and human capital of Americans born after 1947, relative to the trend for prior cohorts. I fail to reject the null of a relative decline in health and human capital for all outcomes considered in Reynolds (2024): men's wages, maternal health as proxied by the birthweight of their infants, and men's and women's log mortality.

2. Literature review

The methodological literature on the "age-period-cohort" identification problem is large. The basic finding that the additively-separable age-period-cohort model is not identified seems to have been rediscovered a number of times. Perhaps the earliest example can be found in Hall (1968), Rodgers (1982) is a prominent example in sociology, and Deaton (1997) provides a useful textbook account. Heckman and Robb (1985) highlight that adding higher order terms exacerbates the identification problem. McKenzie (2006) suggest second difference estimators for the second difference in age, period, and cohort effects within the additively-separable model. Schulhofer-Wohl (2018) shows how to use these second difference estimators as an input to structural estimation.

These models have been used by economists to isolate "cohort effects" in a range of settings with important implications. For example, "vintage" effects in capital are important ingredients for growth accounting (Hall, 1968; Jorgenson, 1996), cohort effects in human capital have been studied in relation to the college-high-school wage gap (Card and Lemieux, 2001) and structural transformation (Porzio et al., 2022), and recently Sorkin and Wallskog (2023) estimates the contribution of firm cohort effects (related to year of firm entry) on inequality.

Given their wide use it is somewhat surprising that age-period-cohort models have escaped the wave of methodological contributions introducing nonseparablity and heterogeneity to common econometrics settings. The common approach in the treatment effect literature, exemplified by Imbens and Angrist (1994), is to introduce unrestricted treatment effect heterogeneity and ask if there are reasonable assumptions such that traditional estimators deliver an interpretable average causal effect of some kind, eg. the Local Average Treatment Effect. For example, De Chaisemartin and d'Haultfoeuille (2020) find that when treatment effects are allowed to be heterogeneous then the traditional estimator from two-way fixed-effects models with a treatment dummy does not deliver an interpretable average causal effect. They therefore suggest a new estimator. To my knowledge, no analogous heterogeneity has been studied in the cohort effects setting.

3. Model and definitions

This section presents a conceptual model which generalizes the idea of "cohort effects" and suggests a link to counterfactuals and potential outcomes. I then define "Local Cohort Effects" and "Local Average Cohort Effects", and the "Difference in Local Average Cohort Effects."

A. Model

Consider the following model:

$$Y_{apc} = g(\theta_c, \epsilon_{ap}) \tag{1}$$

where Y_{apc} denotes an outcome for "units" who are age a, in the year or "period" p, and who are members of the cohort c, ie. they were born in year c. The outcome is a function of: i) θ_c, underlying, fixed features of the cohort and ii) ϵ_{ap} external factors in the year p which affect individuals of the given age, a. There is no restriction (yet) on the structural function $g(\cdot, \cdot)$ or on the dimensionality of θ_c or ϵ_{ap}. Treat the cohort factors θ_c as fixed and the age-by-year factors, ϵ_{ap}, are random variables.

A leading example is the case where units are individuals and the outcome is their average or median earnings of workers born in a given year, at a given age, and in a particular year. In this case, conceptually, θ_c represents the underlying, fixed differences between individuals born in different years which impact their outcomes. It reflects broadly the health, human capital, cognitive ability and "skills" of the cohort. θ_c can be multidimensional to represent both the potentially multidimensional nature of health and "skill" (eg. Heckman, 2007), and to represent differences in the distribution of skills within cohorts. I do not specify explicitly when or how these cohort differences develop, just that they originate before the age one begins measuring the outcome. There is a long history in social science as well as neuroscience and human biology of studying differences between cohorts and of the suggestion that they are likely to differ due to different experiences at "critical periods" in their life.[2] For example, the cohort differences could date to labor market entry, schooling age, infancy, or in utero. The model treats cohort effects as fixed from the age at which the outcomes began to be measured, abstracting from investment in skills after that age and scarring effects in adulthood. In contrast, ϵ_{ap} represents the external factors such as technology or labor demand which will impact the outcome of a cohort who is age a in year p. The interaction between θ_c and ϵ_{ap} is unrestricted, so the model allows for example for a situation where the labor market for 30 year old workers would be a good match for the skills of one cohort but a poor match for those of another cohort.

The model can apply to settings with many other outcome variables or where the units are not individuals. For example, it could be applied to generalize the model in Sorkin and Wallskog (2023) which studies the impact of firms on inequality allowing firms which enter in a particular year to have different "cohort effects." Or it could be applied to the setting in Hall (1968) and a large subsequent literature (reviewed in Jorgenson, 1996) which studies whether capital built in a particular year (of a particular "vintage") has different productivity.

In the context of this model, the concept of "cohort effects" can be generalized and linked to the idea of counterfactuals and potential outcomes.[3] For example, a natural way to summarize a decline in the human capital of cohorts born between 1947 and 1960 would be to ask: had both cohorts faced the set of external age-by-year factors actually experienced by the 1947 cohort how would their earnings have differed? This is fundamentally a counterfactual question and involves the comparison of the observed outcomes for the 1947 cohort to a set of potential outcomes for the 1960 cohort which will never be observed. In particular, the outcomes Y_{apc} will only be observed

[2]See for example Ryder (1965); Easterlin (1987); Fogel and Costa (1997), and Cunha et al. (2006) and the citations therein.

[3]See for example Rubin (1974); Holland (1986); Heckman (2010).

for cohorts, ages, and periods such that $c = p - a$. We can define the structural function $g(\cdot, \cdot)$ however for all pairs of θ_c and ϵ_{ap} — defining "potential outcomes" which a cohort would have had if they had been exposed to different external factors. Identifying these generalized cohort effects therefore is akin to the "Fundamental Problem of Causal Inference" (Holland, 1986).

I will consider identification and inference in a setting where the cohort factors θ_c are fixed and uncertainty comes from the external factors ϵ_{ap}. So all expectations below are taken with respect to the unspecified distribution of the set of ϵ_{ap}.

Define Local Cohort Effects as the expected difference between the observed outcomes for a cohort, c, in a given age and year, and what the expected outcomes of the cohort born one year earlier would have been if they had experienced the same external age-by-year factors as cohort c:

$$\phi_{c,ap} = E\left[g(\theta_c, \epsilon_{ap}) - g(\theta_{c-1}, \epsilon_{ap})\right]$$

Note that while $g(\theta_c, \epsilon_{ap})$ is observed, $g(\theta_{c-1}, \epsilon_{ap})$ is an unobserved potential outcome.

Then define the Local Average Cohort Effect as the expected average Local Cohort Effects across a set of years $\mathcal{P}$:

$$\Phi_{c,\mathcal{P}} = E\left[\tfrac{1}{P}\textstyle\sum_{p\in\mathcal{P};a=p-c}\phi_{c,ap}\right]$$

As an example, consider testing the claim of my earlier paper Reynolds (2024) that there has been a decline in the health and human capital of Americans born after 1947, relative to the prior cohort trend. A relevant Local Average Cohort Effect in this setting is: the counterfactual difference in wages between what the 1948 cohort would have earned had they faced the set of external age-by-year factors that the 1947 cohort faced and what the 1947 did earn. If this effect is negative it implies that the 1948 cohort had more human capital than the 1947 cohort, at least along the dimensions which were relevant in the labor market conditions which the 1947 cohort experienced. This definition therefore takes seriously the idea that the human capital of a cohort (or the fixed features of a particular vintage of capital) may be multidimensional and therefore it's impact may differ depending on the external factors prevailing in specific years for those of a specific ages. This is the sense in which the cohort effects are "local" to particular ages and years.

Then define the Difference in Local Average Cohort Effects as the expected difference in local cohort effects between cohort $c+1$ and cohort c across a set of years $\mathcal{P}$:

$$\begin{aligned}\Psi_{c,\mathcal{P}} &= E\left[\tfrac{1}{P}\textstyle\sum_{p\in\mathcal{P};a=p-c}\phi_{c+1,a-1,p} - \phi_{c,ap}\right] \\ &= E\left[\tfrac{1}{P}\textstyle\sum_{p\in\mathcal{P};a=p-c}\left(g(\theta_{c+1}, \epsilon_{a-1,p}) - g(\theta_c, \epsilon_{a-1,p})\right) - \left(g(\theta_c, \epsilon_{ap}) - g(\theta_{c-1}, \epsilon_{ap})\right)\right]\end{aligned}$$

This can be thought of as a non-parametric generalization of the second-difference in cohort effects considered in McKenzie (2006). Consider the model with earnings by age-year-cohort as the outcome. $\phi_{c+1,ap} - \phi_{c,ap}$ will be negative if the counterfactual difference in earnings between what cohort $c+1$ would have earned and what c did earn, is larger than the difference between what cohort c earned and the counterfactual earnings of cohort $c-1$. If for a given c, one knew that for many years and ages that $\phi_{c+1,ap} - \phi_{c,ap}$ is on average large and negative — we might conclude that there is a "trend break" or "kink" in human capital at cohort c in the sense that the difference in counterfactual earnings is on average much smaller between cohort c and cohort $c+1$ than it is between cohort c and cohort $c-1$.

For example, if there is a sharp break in the cross cohort trend in human capital at the 1947 cohort, such that the change in human capital between the 1947 and 1948 cohort is much smaller than change in human capital between 1946 and 1947 cohort, as I argue in (Reynolds, 2024). Then $\Psi_{1947,\mathcal{P}}$ with earnings as the outcome should be large and negative, because:

$$\Psi_{1947,\mathcal{P}} = E\left[\tfrac{1}{P}\textstyle\sum_{p\in\mathcal{P};a=p-c}\left(g(\theta_{1948},\epsilon_{a-1,p}) - g(\theta_{1947},\epsilon_{a-1,p})\right) - \left(g(\theta_{1947},\epsilon_{ap}) - g(\theta_{1946},\epsilon_{ap})\right)\right]$$

B. Unbiased estimator under exchangeabililty of age-by-year shocks

This section shows that the estimator proposed by McKenzie (2006) for the additively-separable model is an unbiased estimator of Differences in Local Average Cohort Effects under the assumption that age-by-year shocks are exchangeable for neighboring ages.

The key assumption for identification is that the age-by-year shocks in neighboring years hitting an age a are "exchangeable" with the age-by-year shocks in the same years hitting age $a-1$. That is that, for all a, p:

$$F(\{\epsilon_{ap}, \epsilon_{a,p-1}\}, \{\epsilon_{a-1,p}, \epsilon_{a-1,p-1}\}, \boldsymbol{\epsilon}) = F(\{\epsilon_{a-1,p}, \epsilon_{a-1,p-1}\}, \{\epsilon_{ap}, \epsilon_{a,p-1}\}, \boldsymbol{\epsilon})$$

where $\boldsymbol{\epsilon}$ denotes the vector of all other $\epsilon_{a',p'}$ not listed, and $F()$ denotes the joint probability distribution of the entire sequence of shocks. The pairs of shocks are exchangeable in the sense that the joint distribution of all of the age-by-year shocks in neighboring ages and years are invariant to permuting the "label" of ages on the shocks.[4]

For example, it means that it is "as-if-randomly-assigned" whether the set of shocks in neighboring years hits 30-year-olds or 31-year-olds. This assumption is somewhat unusual but has a number of appealing properties. It allows for unrestricted dependence in the shocks over time hitting a given age a. It also allows for there to be non-random shocks to a larger groups of ages in a given year — for example the effect of the supply of workers of nearby ages on wages in

[4]For discussion of exchangeability see for example Draper et al. (1993); Bernardo (1996). See Brock and Durlauf (2001) for an economic application.

Card and Lemieux (2001). What is assumed random is the very-local difference in external factors impacting those one year apart in age.

Under this assumption, one can show easily that the second difference estimator suggested in McKenzie (2006) is an unbiased estimator of the Difference in Local Average Cohort Effects. Begin by defining the first-difference in outcomes of those who are the same age, in neighboring years: $\Delta^p Y_{apc} = Y_{apc} - Y_{a,p-1,c-1}$. And define the second difference in outcomes, which takes the difference in the above between neighboring ages:

$$\Delta^a \Delta^p Y_{apc} = \Delta^p Y_{a-1,p,c+1} - \Delta^p Y_{apc} = (Y_{a-1,p,c+1} - Y_{a-1,p-1,c}) - (Y_{apc} - Y_{a,p-1,c-1})$$

The idea is that the second difference in outcomes, can be written as the sum of two terms. The first term is similar to the Difference in Local Cohort Effects, but for the particular observed draw of the age-by-year shocks. The second term is the difference in age-by-year shocks for that year:

$$\Delta^a \Delta^p Y_{apc} = \underbrace{[g(\theta_{c+1}, \epsilon_{a-1,p}) - g(\theta_c, \epsilon_{a-1,p})] - [g(\theta_c, \epsilon_{a,p-1}) - g(\theta_{c-1}, \epsilon_{a,p-1})]}_{\sim\text{diff in local cohort effects}}$$
$$+ \underbrace{(g(\theta_c, \epsilon_{a-1,p}) - g(\theta_c, \epsilon_{a-1,p-1}) - (g(\theta_c, \epsilon_{a,p}) - g(\theta_c, \epsilon_{a,p-1})}_{\text{diff in age-by-year shocks}}$$

A potential estimator is the average of these second-differenced outcomes across a set of P years $\mathcal{P}$. This can be written as two terms as well. The first term is similar to the Difference in Local Average Cohort Effects, but for the particular observed draw of the age-by-year shocks. The second term is plus the average difference in age-by-year shocks:

$$\hat{\Psi}_c \equiv \frac{1}{P} \sum_{p \in \mathcal{P}} \Delta^a \Delta^p Y_{apc} = \underbrace{\frac{1}{P} \sum_{p \in \mathcal{P}} [(g(\theta_{c+1}, \epsilon_{a-1,p}) - g(\theta_c, \epsilon_{a-1,p})) - (g(\theta_c, \epsilon_{a,p-1}) - g(\theta_{c-1}, \epsilon_{a,p-1}))]}_{\sim\text{diff in Local Average Cohort Effects}}$$
$$+ \underbrace{\frac{1}{P} \sum_{p \in \mathcal{P}} [(g(\theta_c, \epsilon_{a-1,p}) - g(\theta_c, \epsilon_{a-1,p-1}) - (g(\theta_c, \epsilon_{a,p}) - g(\theta_c, \epsilon_{a,p-1}))]}_{\text{avg. diff in age-by-year shocks}} \quad (2)$$

Then under the assumption of exchangeability the average difference in age-by-year shocks goes to zero in expectation, and the estimator is equal in expectation to just the Difference in Local Average Cohort Effects:[5]

[5] Recall that only the age-by-year shocks are random variables, so expectation is taken with respect to their distribution.

$$E\left[\hat{\Psi}_c\right] = E\left[\frac{1}{P}\sum_{p\in\mathcal{P}}\left[\left(g(\theta_{c+1},\epsilon_{a-1,p}) - g(\theta_c,\epsilon_{a-1,p})\right) - \left(g(\theta_c,\epsilon_{a,p-1}) - g(\theta_{c-1},\epsilon_{a,p-1})\right)\right]\right] = \Psi_c$$

C. Permutation-based inference

The exchangeability assumption allows for a simple permutation test in the spirit of Fischer's exact test.[6]

The test will have correct size for a "sharp null" hypothesis that for all possible draws of the age-by-year shocks, a term like the difference in Local Cohort Effects is equal to zero in all years. Denote this hypothesis $H_{0,A}$. It can be written as:

$\frac{1}{P}\sum_{p\in\mathcal{P};a=p-c}\left(g(\theta_{c+1},\epsilon_{a-1,p}) - g(\theta_c,\epsilon_{a-1,p})\right) - \left(g(\theta_c,\epsilon_{ap}) - g(\theta_{c-1},\epsilon_{ap})\right) = 0$
for all possible draws ϵ_{ap} and all a, p s.t. $c = p - a$

Interestingly, the test will also have correct size for an, admittedly somewhat unusual, weak(er) null hypothesis that for all possible draws of the age-by-year shocks, a term like the Difference in Local Average Cohort Effects across all years is equal to zero. Denote this hypothesis $H_{0,B}$. It can be written as:

$\frac{1}{P}\sum_{p\in\mathcal{P};a=p-c}\left(g(\theta_{c+1},\epsilon_{a-1,p}) - g(\theta_c,\epsilon_{a-1,p})\right) - \left(g(\theta_c,\epsilon_{ap}) - g(\theta_{c-1},\epsilon_{ap})\right) = 0$
for all possible draws ϵ_{ap} and all a, p s.t. $c = p - a$

Under the stated conditions, the test will not in general have correct size for the more general null that the expected value (population mean) of the average second-difference in local cohort effects is equal to zero.[7]

I take the sequence of observed first difference in outcomes for the neighboring cohorts $\tilde{c}+1$ and $\tilde{c}$ as data: $\{\Delta^p Y_{a-1,p,\tilde{c}+1}, \Delta^p Y_{a,p,\tilde{c}}\}_{p\in\mathcal{P},a=p-\tilde{c}}$. Consider then all possible sequences which adjust this sequence by permuting any number of the neighboring first differences, and keeping the other first differences in the order observed. Ideally I would use this entire set of possible sequences to conduct the permutation test. As a computationally feasible approximation, I instead make K repeated draws where in each draw: for each pair of neighboring first differences I randomly either permute them or keep them as observed.

[6] See for example Ernst (2004), Imbens and Rubin (2015), and Young (2019).

[7] It seems possible that one could impose mixing conditions on the age-by-year shocks such that a Central Limit Theorem holds on the second-differenced outcomes under the null. Then, along the lines of Janssen (1997); Chung and Romano (2013); Young (2024) potentially a permutation test based on a studentized test statistic would have asymptotically correct size for the general null as the number of periods tends to infinity.

I will use as a test statistic the average second difference estimator, $\hat{\Psi}_c$, defined in Section 7. Denote this test statistic defined based on the observed data as $\hat{\Psi}_{\tilde{c}}$. Define the corresponding test statistics calculated from the permuted data from each draw 1 to K as $\hat{\Psi}^{(1)}$, $\hat{\Psi}^{(2)}$ ... $\hat{\Psi}^{(K)}$.

Under either null hypothesis $H_{0,A}$ or $H_{0,B}$ the first term in Equation 2, representing the second difference in local average cohort effects, is zero and the test statistic is therefore equal to:

$$\hat{\Psi}_c = \frac{1}{P}\sum_{p\in\mathcal{P}}\left[\left(g(\theta_c,\epsilon_{a-1,p}) - g(\theta_c,\epsilon_{a-1,p-1}) - \left(g(\theta_c,\epsilon_{a,p}) - g(\theta_c,\epsilon_{a,p-1})\right)\right)\right]$$

It then directly follows from the assumption of pairwise exchangeability that the sequence of test statistics calculated on the observed data and based on the K draws, $\{\hat{\Psi}_{\tilde{c}}, \hat{\Psi}^{(1)}, \hat{\Psi}^{(2)} \ldots \hat{\Psi}^{(K)}\}$, is a sequence of exchangeable random variables.

Therefore any ordering of the sequence from smallest to largest is equally likely. Assuming the test statistic is continuous and therefore one can ignore ties, then the probability that $\hat{\Psi}_{\tilde{c}}$ is greater than m of the permuted test statistics is simply $\frac{m+1}{K+1}$.[8] One can therefore calculate a 1-sided permutation p-value as:

$$\hat{p}_{\tilde{c},1} = \frac{1+\sum_{k=1}^{K}\mathbb{1}\left(\Psi_{\tilde{c}} \geq \hat{\Psi}^{(k)}\right)}{K+1}$$

And such a p-value has correct size:

$$\begin{aligned} P(\hat{p}_{\tilde{c},1} \geq \alpha) &= P\left(\frac{1+\sum_{k=1}^{K}\mathbb{1}\left(\Psi_{\tilde{c}} \geq \hat{\Psi}^{(k)}\right)}{K+1} \geq \alpha\right) \\ &= P\left(\sum_{k=1}^{K}\mathbb{1}\left(\Psi_{\tilde{c}} \geq \hat{\Psi}^{(k)}\right) \geq \alpha(K+1)-1\right) = \alpha \end{aligned}$$

In practice I use the following two-sided p-value:

$$\hat{p}_{\tilde{c}} = 2\cdot\min\left(\frac{1+\sum_{k=1}^{K}\mathbb{1}\left(\Psi_{\tilde{c}} \geq \hat{\Psi}^{(k)}\right)}{K+1}, \frac{1+\sum_{k=1}^{K}\mathbb{1}\left(\Psi_{\tilde{c}} \leq \hat{\Psi}^{(k)}\right)}{K+1}\right)$$

In the application below I also construct confidence intervals by inverting the permutation test following the approach in Imbens and Rubin (2015); Ganong and Jäger (2018). The confidence

[8] See Phipson and Smyth (2010) for a similar argument in a different setting.

interval can be interpreted as the set of second-differences in local average cohort effects which the test fails to reject. In particular it is the interval B such that for all $b \in B$ one would fail to reject null hypothesis that: $H_{0,B} : \frac{1}{P}\sum_{p\in\mathcal{P}} \Delta\Delta^{ap}\psi_c = b$ for for all possible draws ϵ_{ap} and a, p s.t. $c = p - a$

I follow a similar approach to that described in Ganong and Jäger (2018) to construct the confidence intervals. I test the above null for a given value b by reconstructing the data adding in a hypothetical effect b to one of the first differences. That is I replace the sequence listed above with the following reconstructed data: $\{\Delta^p Y_{a-1,p,\tilde{c}+1}, \Delta^p Y_{a,p,\tilde{c}} + b\}_{p\in\mathcal{P},a=p-\tilde{c}}$. I then construct a permutation p-value following the approach described above. I use a bisection algorithm to find the endpoints of the confidence interval.

4. Empirical application

A. Background

In Reynolds (2024), I argue that there has been a decline in the health and human capital of Americans born after 1947, relative to the prior cohort trend. I initially showed that the trend across cohorts in age-adjusted educational attainment, wages, maternal health (proxied by the birth weight of infants), and mortality all exhibited trend breaks near the 1947 cohort, such that each outcome declines for subsequent cohorts relative to the prior trend. I argued there that these simultaneous trend breaks, while striking, could in principle reflect differences in external factors which these cohorts were exposed to, rather than underlying differences in health and human capital. That is, cohorts born after 1947 may have been otherwise similar to earlier cohorts, but were merely unlucky to have experienced bad conditions throughout their lifetime.

I therefore used two methods to provide evidence under weaker assumptions about the nature of external factors, that these patterns reflect a decline in the underlying health and human capital of cohorts born after 1947, relative to the prior trend. First, I estimated traditional age-period-cohort models, which assume cohort, age, and year factors are each additively separable. Second, I estimated models with a trend break of unknown location in cohort effects while allowing for a separate polynomial in age in each year, adapting methods from the structural break literature. Both methods revealed strong evidence of a trend break at the 1947 or 1948 cohort in underlying health and human capital evident in each of the above outcomes.

However, doubts could remain depending on one's comfort with the assumptions. One possible concern is that both methods assume constant age, period, and cohort effects. Heterogeneity and complex interactions in these effects appear plausible: for example the impact that the human capital of cohort has on outcomes could depend on features of the labor market in a given year, or it may change throughout a cohort's lifecyle. Another potential concern is that while the controls based on separate polynomial in age in each year, while flexible, are fundamentally parametric and will only be guaranteed to be valid if the assumed parametric form is correct.

Below, I therefore apply the methods developed in this paper to test for a large second difference in local average cohort effects at the 1947 cohort. This can be seen as a generalization of testing for a "trend break" in cohort effects. These methods, as described above, allows for very unrestricted heterogeneity in the interaction between fixed features of cohorts and external factors which vary by year and age. A downside of course is that the inference is highly local to the 1946 to 1948 cohorts and the age-by-year factors they experienced.

B. Data

I study the same main outcomes and use the same data and sample restrictions as in Reynolds (2024). More detail is given in that paper and the associated replication file.

I study median hourly wages of men age 25 to 54, using the Current Population Survey, Merged Outgoing Rotation Group (CPS-MORG), from 1979 to 1993. I calculate the approximate birth year as the survey year minus the respondent's age. I restrict my analysis to cohorts born between 1930 and 1965. These restrictions lead to a sample of 970,479 men with non-missing earnings. I calculate the sample median separately for age-year-sex cells, using the survey weights. I adjust earnings using the CPI-U-RS.

I study the birthweight of infants based on the year of birth of the *mother* as a proxy for maternal health. I use the 1968 to 1995 Birth Data Files (National Center for Health Statistics, n.d.). I calculate the approximate birth year of each mother as the infant birth year minus the mother's age. I restrict my analysis to births occuring in years years 1968 to 1995, to mothers who were born between 1930 and 1970 and are ages 18 to 40. This results in a sample of more than 75 million births.I calculate mean birthweight in cells by year, single age, and mother's birth year.

To study mortality, I use data from the Human Mortality Database on number of deaths and population-at-risk by year and age. These data are derived from official vital statistics and census estimates. I again define cohort as year minus age. I restrict my analysis to the years 1975-2019, ages 25 to 85, and cohorts born between 1930 and 1965. I calculate the natural log of the mortality rate separately for men and women by single age and year.

C. Results

The results of applying the above strategy to estimate the second difference in local cohort effects reveal strong evidence evidence of the non-parametric equivalent of a trend break in cohort effects located at the 1947 cohort for wages, maternal health, and mortality of men and women.

For each outcome, Figure 1 shows the sequence of point estimates $\hat{\Psi}_c$ for the cohorts between 1937 and 1957. Panels A and B show that for both median log wage and mean birth weight (by mother's birth cohort) the estimated average second difference in local cohort effects for 1947 are negative and large in magnitude — around 2 or 3 times larger in magnitude than any other cohorts' estimates. For both outcomes the other cohorts' estimated average second differences in local cohort effects are near 0 — with the confidence intervals including zero for all but one other wage estimates and all except two other birth weight estimates.

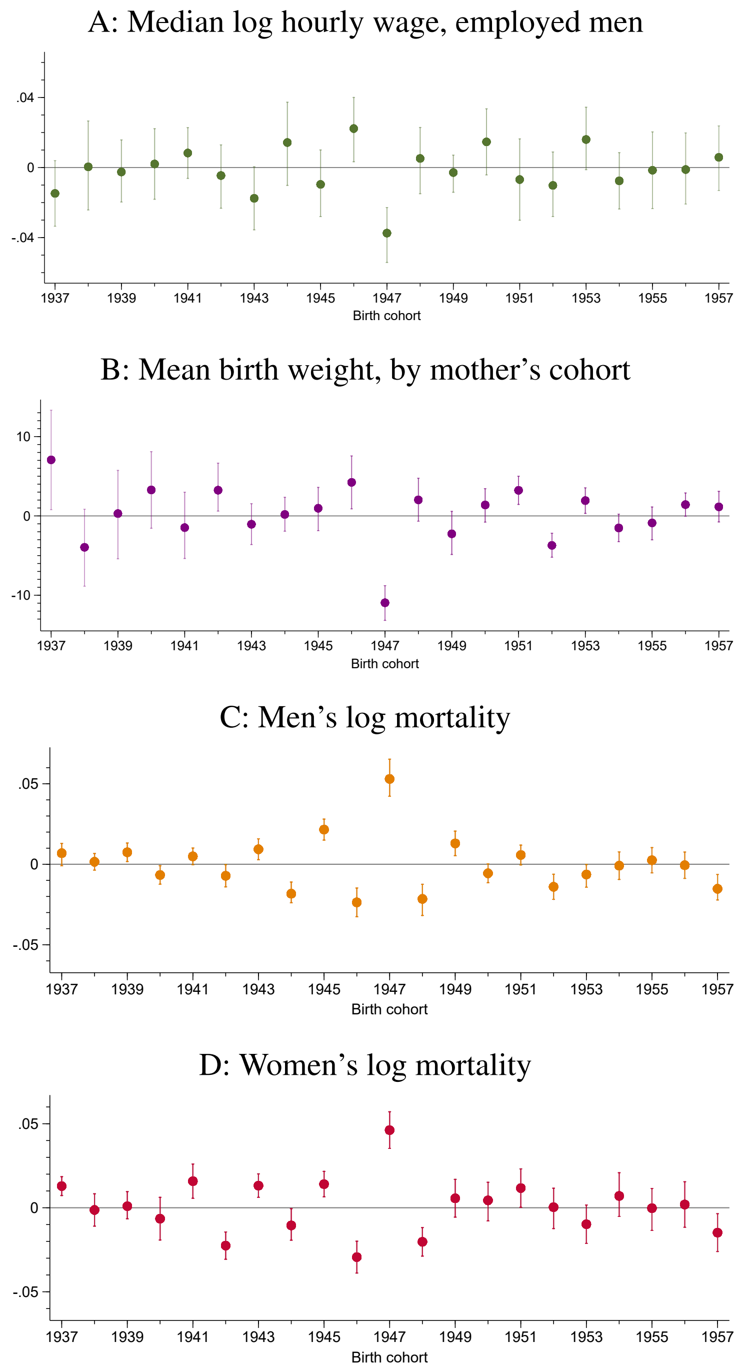
Panels C and D show results for men's and women's log mortality. For both outcomes, the estimated average second difference in local cohort effects for 1947 are positive and large in magnitude, again on the order of 2 times the magnitude of any other estimates. For these outcomes, however the other cohort's estimates are much less tightly centered around zero however — oscillating around zero with the confidence interval often not including zero.

Interestingly, the evidence of a relative decline in health and human capital beginning suddenly after the 1947 cohort using the nonparametric approach is *even stronger* than that shown in Reynolds (2024) using parametric approaches. This is true first with respect to the consistency of the timing of the relative decline. In Reynolds (2024), I estimated models with a trend break of unknown location in cohort effects while controlling for a separate polynomial in age in each year, adapting methods from the structural break literature. The estimated break locations from these parametric (but flexible) models, shown in Table 1 of Reynolds (2024), were concentrated very near 1947 for all outcomes. However, the exact cohort where the trend break was estimated to occur did vary across outcomes: between 1946 and 1950. In contrast, the nonparametric estimates of this paper point towards a sudden decline beginning precisely after the 1947 cohort in all outcomes. For *all* four outcomes the largest estimated difference in Local Average Cohort Effects shown in Figure 1 is for the 1947 cohort. These difference could suggest misspecification in the parametric cohort trend break model estimated in Reynolds (2024). In particular, the imposition that the size of the trend break in cohort effects is of the same magnitude in all years/ages is completely relaxed in the nonparametric approach presented here.

The magnitudes of the average second differences in local cohort effects are also all much *larger* than the estimated magnitude of the trend breaks for the same outcomes from the parametric trend break model in Reynolds (2024). This could again be due to misspecification in the additively-separable cohort trend break model. Although it is important to note that the nonparametric estimates are only informative about very local differences between the 1946, 1947, and 1948 cohorts; while the other models if correctly specified are informative about differences in cohort health and human capital more broadly for the cohorts born 1930 to 1965 (up to trend).

Figure 2 illustrates the permutation-based test of the sharp null that all of the second difference in local cohort effects for the 1947 cohort are 0. The dashed line shows the point estimate of $\hat{\Psi}_{1947}$ and the histogram shows the permutation distribution based on 10,000 simulations. The implied

Figure 1: Nonparametric estimation of average second difference in local cohort effects



This figure shows the results of implementing the estimator of the average second difference and in local cohort effects, described in the paper. For each outcome, the panels on the left show the sequence of point estimates $\hat{\Psi}_c$ for the cohorts between 1937 and 1957. They also show permutation-based, constant-effect, 95 % confidence intervals based on 1,000 simulations of the permutation distribution and a bisection algorithm. The confidence interval can be interpreted as the set of constant second-difference in local cohort effects which the test fails to reject.
Data and sample restrictions follow Reynolds (2024). Panel A is based on CPS-MORG data, 1979-1993, and includes men age 25-54. Panel B is based on vital statistics natality microdata, 1968-1990, mothers age 18-40. Panel C and D are based on data from the Human Mortality Database, and include 1975-2019, ages 25-85.

p-value of the "sharp null" hypothesis that the second-difference in cohort effects is equal to zero in all years for log wage is .0013, while those for the other three outcomes are .0002.

These results provide strong evidence evidence of the non-parametric equivalent of a trend break in cohort effects located at the 1947 cohort, from a quite unrestricted model with weak assumptions. The remaining threat to validity would be non-random changes in the impact of age across years between neighboring ages. For example, a large shock in a given year to the health of those age 30 and under, which did not impact those age 31 and over. Alternative external explanations, or more broadly external factors impacting each of the four outcomes, are generally thought to be smooth in age. For example, the effects of shifts in supply or demand on wages will be smooth as long as individuals who are close in age are substitutable (Card and Lemieux, 2001). Similarly, the effect of "biological aging" on health is generally thought to be a smooth, continuous process.

Discrete policy cutoffs based on age could seem to be a threat, but note that to generate the above results they would have to "follow" the same cohort over time. For example, moving from age 30 in 1997, to age 31 in 1998, to age 32 in 1999 and so on. A large shock in a single year to one age and not the neighboring age would not yield a statistically significant estimate because the permutation-based-inference procedure would correctly reveal that such a pattern is not particularly "unlikely" under the null.[9]

5. Conclusion

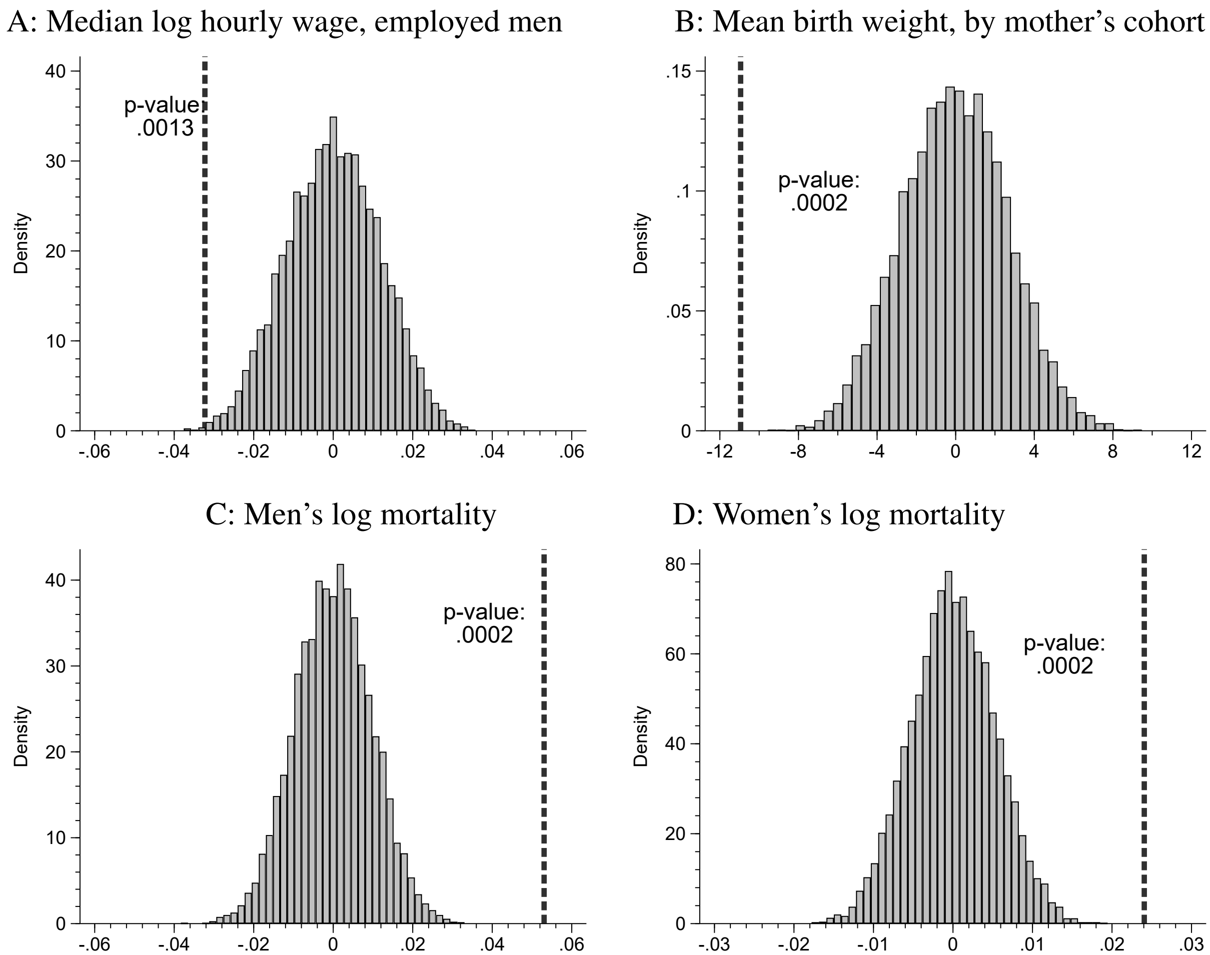
This paper bring heterogeneity into the "cohort effects" setting, by considering a nonseparable model which allows for unrestricted interactions between fixed features of cohorts and external factors which vary by age and year. I define Local Average Cohort Effects as the counterfactual average outcomes a cohort would have if they had experienced the external age-by-year factors of a cohort born one year earlier

I then study properties of the estimator proposed by McKenzie (2006) under this specified model. I show that this estimator identifies second differences in Local Average Cohort Effects under the assumption that shocks to neighboring ages in neighboring years are exchangeable. This assumption also suggests a permutation-based approach to inference.

Other estimators can also be studied in the setting specified here. For example, one could considers partial identification of particular linear functions of Local Cohort Effects under assumptions on the magnitude of external age-by-year shocks (similar in spirit to Petterson et al. 2023). For

[9] Consider a simple example where in one year the second difference in outcomes for the 1947 cohort is equal to Z, an economically meaningful magnitude, but in all other years the second difference is zero. The point estimate would be Z, and the resulting permutation distribution would have half it's mass at Z and half it's mass at $-Z$ — implying the outcome is quite likely under the null.

Figure 2: Permutation test of the sharp null that all of the second difference in local cohort effects for the 1947 cohort are 0



This figure illustrates the permutation-based test of the sharp null that all of the second difference in local cohort effects for the 1947 cohort are 0. The dashed line shows the point estimate of $\hat{\Psi}_{1947}$ and the histogram shows the permutation distribution based on 10,000 simulations. Data and sample restrictions follow Reynolds (2024) and Figure 2.

example one could impose an upper limit on how different the external shocks hitting workers age 25-27 can be from that hitting those age 28-30.

References

José M Bernardo. The concept of exchangeability and its applications. *Far East Journal of Mathematical Sciences*, 4:111–122, 1996.

William A Brock and Steven N Durlauf. Growth empirics and reality. *World Bank Economic Review*, 15(2):229–272, 2001.

David Card and Thomas Lemieux. Can falling supply explain the rising return to college for younger men? A cohort-based analysis. *The Quarterly Journal of Economics*, 116(2):705–746, 2001.

EunYi Chung and Joseph P Romano. Exact and asymptotically robust permutation tests1. *The Annals of Statistics*, 41(2):484–507, 2013.

Flavio Cunha, James J. Heckman, Lance Lochner, and Dimitriy V. Masterov. Interpreting the evidence on life cycle skill formation. volume 1 of *Handbook of the Economics of Education*, pages 697–812. 2006.

Clément De Chaisemartin and Xavier d'Haultfoeuille. Two-way fixed effects estimators with heterogeneous treatment effects. *American economic review*, 110(9):2964–2996, 2020.

Angus Deaton. *The analysis of household surveys: a microeconometric approach to development policy*. World Bank Publications, 1997.

David Draper, James S Hodges, Colin L Mallows, and Daryl Pregibon. Exchangeability and data analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 156(1):9–28, 1993.

Richard A Easterlin. *Birth and fortune: The impact of numbers on personal welfare*. University of Chicago Press, 1987.

Michael D Ernst. Permutation methods: a basis for exact inference. *Statistical Science*, pages 676–685, 2004.

Robert W Fogel and Dora L Costa. A theory of technophysio evolution, with some implications for forecasting population, health care costs, and pension costs. *Demography*, 34(1):49–66, 1997.

Peter Ganong and Simon Jäger. A permutation test for the regression kink design. *Journal of the American Statistical Association*, 113(522):494–504, 2018.

Pierre-Olivier Gourinchas and Jonathan A Parker. Consumption over the life cycle. *Econometrica*, 70(1):47–89, 2002.

Robert E Hall. Technical change and capital from the point of view of the dual. *The Review of Economic Studies*, 35(1):35–46, 1968.

Robert E Hall. The measurement of quality change from vintage price data. In *Price Indexes and Quality Change: Studies in New Methods of Measurement*, pages 240–272. Harvard University Press, 1971.

James J Heckman. The economics, technology, and neuroscience of human capability formation. *Proceedings of the National Academy of Sciences*, 104(33):13250–13255, 2007.

James J Heckman. Building bridges between structural and program evaluation approaches to evaluating policy. *Journal of Economic Literature*, 48(2):356–398, 2010.

James J. Heckman and Richard Robb. Using longitudinal data to estimate age, period and cohort effects in earnings equations. In *Cohort analysis in social research*, pages 137–150. Springer, 1985.

James J Heckman, Lance Lochner, and Christopher Taber. Explaining rising wage inequality: Explorations with a dynamic general equilibrium model of labor earnings with heterogeneous agents. *Review of Economic Dynamics*, 1(1):1–58, 1998.

Paul W Holland. Statistics and causal inference. *Journal of the American statistical Association*, 81(396):945–960, 1986.

Guido W. Imbens and Joshua D. Angrist. Identification and estimation of Local Average Treatment Effects. *Econometrica*, 62(2):467–475, 1994.

Guido W Imbens and Donald B Rubin. *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press, 2015.

Arnold Janssen. Studentized permutation tests for non-iid hypotheses and the generalized behrens-fisher problem. *Statistics & probability letters*, 36(1):9–21, 1997.

Dale W Jorgenson. Empirical studies of depreciation. *Economic inquiry*, 34(1):24–42, 1996.

David Lagakos, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman. Life cycle wage growth across countries. *Journal of Political Economy*, 126(2):797–849, 2018.

Max Planck Institute for Demographic Research, University of California, Berkeley, and French Institute for Demographic Studies. Human mortality database, n.d. Accessed at https://www.mortality.org as of 1 March 2024.

David J. McKenzie. Disentangling age, cohort and time effects in the additive model. *Oxford Bulletin of Economics and Statistics*, 68(4):473–495, 2006.

Jacob Mincer. *Schooling, Experience, and Earnings*. National Bureau of Economic Research, 1974.

National Center for Health Statistics. Birth data files, n.d. NBER. Accessed at https://www.nber.org/research/data/vital-statistics-natality-birth-data as of 1 March 2024.

Marco Stenborg Petterson, David Seim, and Jesse M Shapiro. Bounds on a slope from size restrictions on economic shocks. *American Economic Journal: Microeconomics*, 15(3):552–572, 2023.

Belinda Phipson and Gordon K Smyth. Permutation p-values should never be zero: calculating exact p-values when permutations are randomly drawn. *Statistical applications in genetics and molecular biology*, 9(1), 2010.

Tommaso Porzio, Federico Rossi, and Gabriella Santangelo. The human side of structural transformation. *American Economic Review*, 112(8):2774–2814, August 2022.

Nicholas Reynolds. The broad decline of health and human capital of Americans born after 1947. *American Economic Review: Insights*, 2024.

Willard L Rodgers. Estimable functions of age, period, and cohort effects. *American sociological review*, pages 774–787, 1982.

Donald B Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5):688, 1974.

Norman B. Ryder. The cohort as a concept in the study of social change. *American Sociological Review*, 30(6):843–861, 1965.

Sam Schulhofer-Wohl. The age-time-cohort problem and the identification of structural parameters in life-cycle models. *Quantitative Economics*, 9(2):643–658, 2018.

Isaac Sorkin and Melanie Wallskog. The slow diffusion of earnings inequality. Technical report, National Bureau of Economic Research, 2023.

Alwyn Young. Channeling fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results. *The Quarterly Journal of Economics*, 134(2):557–598, 2019.

Alwyn Young. Asymptotically robust permutation-based randomization confidence intervals for parametric ols regression. *European Economic Review*, 163:104644, 2024.