Population Aging, Cohort Replacement, and the Evolution of Income Inequality in the United States*

Vesa-Matti Heikkuri[1] and Matthias Schief[2]

[1]Department of Economics, Brown University
[2]Department of Economics, Brown University

Abstract

We study how demographic change affects the evolution of income inequality in the United States both historically and prospectively. We emphasize the distinct roles of population aging and cohort replacement, and develop a methodology to study their joint compositional effect on income inequality. In the process, we also develop a novel methodology to aggregate sub-population Gini coefficients into a population-level Gini coefficient based on the principle of maximum entropy. We find that rising income inequality is embodied in birth cohorts born since the mid-20th century and that the increase in inequality over the past two decades can be fully accounted for by demographic change. Furthermore, we predict that demographic change over the next two decades will lead to further increase of the income Gini coefficient by one to four percentage points.

Keywords: Demographic change, income inequality, compositional effects
JEL classification: C46, D31, J11

*Address correspondence to matthias_schief@brown.edu and vesa-matti_heikkuri@brown.edu. We thank Joaquin Blaum, James Feyrer, Oded Galor, Suvi Heikkuri, Stelios Michalopoulos, Cosimo Petracchi, Devesh Rustagi, and David Weil for valuable comments and suggestions. We also thank participants at the Growth Lab at Brown University, Max Planck Institute for Demographic Research, PAA 2022 annual meeting, and the Stone Center on Socio-Economic Inequality at CUNY. We acknowledge support from the James M. and Cathleen D. Stone Wealth and Income Inequality Project.

I Introduction

Income inequality in the United States is much higher today than half a century ago. This increase in inequality reflects changes in the technological and institutional environment as well as changes in the composition of the population. In this paper, we study how demographic change affects income inequality by altering the composition of the population in terms of age groups and birth cohorts. Our goal is to explain how demographic change has affected the evolution of income inequality in the past and to project how income income inequality will evolve under future demographic change. While it is difficult to predict future changes in the technological and institutional environment, demographics can be projected reliably several decades into the future.

We study how demographic change affects income inequality by considering the following thought experiment: how does income inequality evolve over time if the economic environment is held fixed in a given base year and only demographic change is allowed to take place thereafter? In this thought experiment, the economic environment determines the mapping of characteristics, such as age or education, into incomes. The economic environment also shapes the characteristics of young cohorts entering the labor market. Demographic change affects income inequality by altering the composition of these characteristics via population aging and the gradual replacement of old birth cohorts by new ones.

We distinguish population aging and cohort replacement as separate channels of demographic change. Population aging is the change in the population shares of different age groups, whereas cohort replacement is the change in the population shares of different birth cohorts. Cohort replacement occurs even in the absence of population aging. Importantly, both of these demographic processes can affect aggregate income inequality.

Population aging affects income inequality because the distribution of income within a cohort changes with age. Heterogeneous returns to experience, persistent idiosyncratic shocks, and differential rates of wealth accumulation all imply that within-cohort income inequality is higher among older households. Similarly, on-the-job training, accumulation of experience over the working life, and retirement imply that mean income follows a hump-shaped path in age. Population-level income inequality therefore depends on the age structure of the population.

Cohort replacement affects income inequality because the distribution of incomes at a given age differs across cohorts. Cohort-specific characteristics such as the distribution of human capital, the allocation of talent across professions, the macroeconomic environment when entering the labor market, and the degree of positive assortative mating imply that income distributions differ across birth cohorts. Population-level income inequality therefore also depends on the cohort structure of the population.

As a benchmark, we first implement our thought experiment using a re-weighting method in the spirit of DiNardo et al. (1996). This exercise makes the strong assumption that differences in income distributions across birth cohorts can be fully explained by differences in age and the college share. Under this ignorability assumption, the re-weighting analysis finds that demographic change and rising income inequality are seemingly unrelated. We then argue that birth cohorts likely differ in unobserved characteristics in a way that renders the re-weighting analysis misleading.

In the main part of the paper, we develop a parametric methodology to implement our thought experiment in the presence of cohort differences in both observed and unobserved characteristics. Using household-level income data for the US, we estimate life-cycle profiles and cohort differences in mean incomes and income Gini coefficients. We document important cohort differences in income distributions that are not accounted for by differences in age and the college share. We then predict population shares and use the estimated profiles to construct counterfactual moments for sub-population income distributions under demographic change when the economic environment is held fixed.

Recovering the population-level Gini coefficient requires us to develop a novel methodology for aggregating sub-population Gini coefficients. We follow the principle of maximum entropy and parameterize sub-population income distributions using a maximum entropy distribution for given mean and Gini coefficient. This allows us to derive the population-level Gini coefficient using the fitted within-cohort income distributions together with the respective population shares. Our methodology is able to aggregate sub-population Gini coefficients with only limited loss of information.

We find that demographic change plays an important role in the evolution of population-level income inequality in the US – both in the past and in the future. In our preferred specification, the compositional effect of demographic change accounts for almost all of the increase in income inequality over the past three decades. Moreover, we predict that demographic change will further increase inequality in the near future, with our estimates suggesting an increase in the income Gini coefficient of between one and four percentage points by the year 2040.

We decompose the full effect of demographic change into the parts explained by population aging and cohort replacement. While we find that population aging contributed substantially to rising income inequality since the 1990s, our results also show that the compositional effect of cohort replacement is generally larger than the effect of population aging. Moreover, we find that the predicted increase in income inequality in the future is driven almost exclusively by cohort replacement.

The rest of the paper is structured as follows. In section I.1, we discuss how our paper relates to the existing literature, and we introduce our data sources in section I.2. Section II implements the re-weighting analysis and discusses its shortcomings. In section III, we develop our parametric method. In section IV, we present our main results. Section V concludes.

I.1 Related literature

The literature on demographic change and economic inequality typically studies the compositional effects of a changing population structure using re-weighting methods. Recent papers in this literature include Kuhn et al. (2020) and Auclert et al. (2021), who study the effects of population aging, and Eika et al. (2019), who study the impact of changing household characteristics[1]. Kuhn et al. (2020) assemble a new micro data set for household income and wealth in the US going back to 1949 and study, among other things, the effect of demographic change on income and wealth inequality in the past. They find a moderately positive effect of population aging on income inequality that is roughly constant across the sample period. Auclert et al. (2021), on the other hand, use population projections to predict the compositional effect of demographic change on the future evolution of the wealth-to-output ratio in the United States and a number of other countries. They predict that population aging will have a significant impact on the wealth-to-output ratio in the United States over the next decades. Eika et al. (2019) study the role of educational assortative mating on household income inequality. They find that educational assortative mating accounts for a non-negligible share of cross-sectional inequality but that the trend in sorting has hardly affected income inequality. They also find that the increase in college attendance and completion rate by women has slowed down the increase in household income inequality.

The compositional effects of a changing population structure on economic inequality have also been studied in the labor literature, where the focus has predominantly been on the skill composition of the population and the role of skill-biased technical change (Juhn et al., 1993; Lemieux, 2006; Autor et al., 2008; Hoffmann et al., 2020). Lemieux (2006), for example, studies how changes in the composition of the US population in terms of experience and educational attainment affect residual wage inequality using a re-weighting analysis. He finds that increases in within-group inequality are concentrated in the 1980s and that the increase in population-level wage inequality in the subsequent decade is driven by composition effects.

Cohort differences in income distributions play an important role in our paper. A source of cohort differences that has recently received increased attention is scarring. This literature has documented long-lasting negative effects on earnings and employment for cohorts entering the labor market in a bad economy, and often finds that these effects are heterogeneous and therefore affect inequality (Raaum and Røed, 2006; Kahn, 2010; Oreopoulos et al., 2012; Rothstein, 2019; Schwandt and Von Wachter, 2019). Outside the scarring literature, there is evidence of secular trends in cohort-specific characteristics. Card and Lemieux (2001) attribute the rising college premium to a slowdown in educational attainment for cohorts born after 1950. Hendricks and Schoellman (2014) explain the same phenomenon by growing ability gaps between high school and college-educated workers across different birth cohorts. More recently, Hsieh et al. (2019) argue that cohort-specific improvements in the allocation of talent have contributed significantly to US economic growth. Similarly, the literature on structural change has documented that a large share of labor reallocation can be accounted for by new cohorts entering growing industries (Lee and Wolpin, 2006; Hobijn et al., 2019; Porzio et al., 2021).

To construct counterfactuals, we estimate how income distributions depend on age and birth

[1]Older papers include Burtless (1999),Daly and Valletta (2006), Larrimore (2014) and Greenwood et al. (2014).

cohort using a standard age-period-cohort model[2]. In this respect, our paper is also related to the literature devoted to studying life-cycle profiles of economic inequality. In particular, we build on Deaton and Paxson (1994a,b) who estimate age profiles for within-cohort income and consumption variance in the US and propose a normalization for dealing with the linear dependence of age, period, and cohort effects. Heathcote et al. (2005) point out the importance of the choice of normalization in estimating the age profile of income inequality. To deal with this issue, we follow Lagakos et al. (2018) who suggest exploring the results under a range of different normalizations.

I.2 Data sources

We use data on household income for the years 1968-2020 from the Current Population Survey (CPS). Our measure of household-level income is the total money income during the previous calendar year of all adult household members[3]. Total money income is the sum of wages and salaries, income from professional practice and self-employment, rental income, interest, dividends, transfer payments, as well as business and farm income. We complement the CPS data with a longer series of harmonized repeated cross-sections based on archival data from historical waves of the Survey of Consumer Finances that was recently made available by Kuhn et al. (2020). This data set spans the time period 1949-2019 and reports household-level total income, which has the same definition as total money income in the CPS. We follow Kuhn et al. (2020) and refer to these data as the SCF+ data set[4]. The CPS and SCF+ data sets complement each other. While the CPS data set has larger sample size, the SCF+ data cover two more decades.

About 0.16 percent of all households in the CPS data and 0.31 percent of all households in the SCF+ data report negative total incomes[5]. We censor the income distribution by re-coding negative values as zeros[6]. To avoid problems associated with topcoding in the CPS and SCF+ data sets, we focus on income inequality among the bottom 99%.

For the CPS data, we use annual surveys between 1968 and 2020. For the SCF+ data, we use triennial waves constructed by Kuhn et al. (2020) which leave us with data for every third year between 1950 and 2019 with the exception of the years 1974, 1980, and 1986. We correspondingly aggregate the data into three-year age groups and birth cohorts. We assign households to their respective birth cohort based on the age of the household head. Furthermore, we restrict our attention to households in which the household head is between 26 and 79 years old in the CPS

[2]Other papers that also use an age-period-cohort model to study life-cycle behavior include Attanasio (1998), Storesletten et al. (2004), Low et al. (2010), Huggett et al. (2011), Aguiar and Hurst (2013), and Heathcote et al. (2014).

[3]Specifically, we use the variable HHINCOME from the IPUMS CPS harmonized microdata (Ruggles et al., 2020).

[4]The data set in Kuhn et al. (2020) covers the time period 1949-2016. We added to this data set the 2019 Survey of Consumer Finances.

[5]After applying sampling weights, households with negative income make up 0.13 percent of the population in both the CPS and the SCF+ data sets.

[6]Negative income levels pose a challenge for the interpretation of differences in income inequality across different subgroups of the population, as they can inflate the Gini coefficient even if the dispersion of income is low.

data and between 26 and 80 in the SCF+ data[7]. The median age in the US population has increased from 27 in 1970 to 38 in 2019, and is predicted to increase to 42 by the year 2060[8] The leftmost panel in figure 1 shows the age distribution in the US population in the years 1970 and 2010, as well as projections for the year 2050. Over this time period, old people progressively make up a higher fraction of the US population.

The middle and the right panels show the corresponding age distribution among household heads in the CPS and the SCF+ data sets. As we want to study the role of demographic change not only in the past but also in the future, we need to translate the predicted changes in the age structure of the US population into corresponding changes in the age structure of household heads in the survey data. We do this by using the constant headship rate method. That is, we compute the probability that an individual of a given age in the latest survey wave is recorded as the household head, and we assume that these probabilities remain fixed in the future.

II A re-weighting analysis

As a benchmark, we first implement our thought experiment using a re-weighting analysis. This method is commonly used in the literature studying compositional effects of demographic change (see e.g. Lemieux (2006), Kuhn et al. (2020), Auclert et al. (2021)). The re-weighting method relies on an ignorability assumption which posits that the distribution of unobserved characteristics conditional on observed characteristics is time-invariant (Fortin et al., 2011). Hence, this method is valid only if cohort differences can be fully captured by differences in observed characteristics.

In our thought experiment, we hold the economic environment fixed in the base year and only allow for demographic change to take place thereafter. People who have entered the economy before the base year maintain their age-invariant characteristics while the people who enter the economy after the base year are set to have the same distribution of characteristics as the youngest cohort in the base year. Cohort sizes evolve as observed in the data.

In this exercise, we focus on age and education as observed characteristics. These are the most common characteristics used in studies of compositional effects on inequality (see e.g. Lemieux

[7]The definition of the household head in the SCF+ data is the male partner in mixed-sex couples, the older partner in same-sex couples, and the single core individual in households without a core couple. In the CPS data, the definition of household head changed in the year 1980. While it was similar to the SCF+ in the years prior to 1980, the CPS has since discontinued the use of the name "household head" and has replaced it with the name "householder". A householder is the person in whose name the housing unit is owned or rented or, in the case of a married couple jointly owning or renting the house, it is either of the spouses. In our main results, we assign households to their respective age groups based on the age of the individual who is called household head/householder in the respective data set. As a robustness check, we re-define household heads in the CPS to be always the male partner in mixed-sex couples and the older person in same-sex couples to be consistent with the definition in the SCF+ data. The results hardly change under this alternative definition. Finally, we drop household heads aged 80 in the CPS data, because for some waves individuals that are older than 80 are coded as 80.

[8]United States Census Bureau (2017).

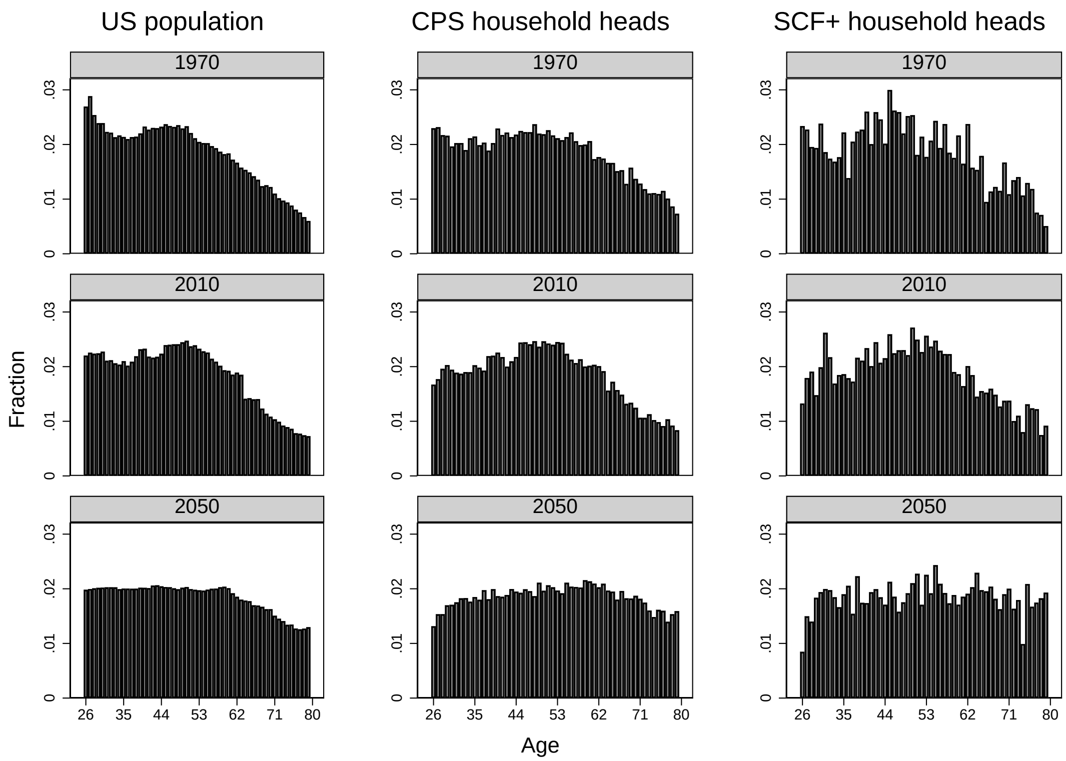


Figure 1: The evolution of the age composition of the US population and the implied age distributions of household heads in the CPS and the SCF+ data sets.

(2006), Kuhn et al. (2020),Hoffmann et al. (2020)), and we observe these characteristics in both surveys. Moreover, educational attainment is fairly constant after age 26 which simplifies the implementation of the thought experiment.

Let $F_{YX,p}$ be the cumulative distribution function for the joint distribution of income Y and characteristics X in year p, and let $F_{X,p}$ be the cumulative distribution function for the marginal distribution of characteristics X in year p. Given a pair of base year $\bar{p}$ and target year p' with $p' > \bar{p}$, our goal is to compute a counterfactual income distribution for the target year, $\tilde{F}_{YX,\bar{p},p'}$, by fixing the conditional distribution of income in the base year, $F_{Y|X,\bar{p}}$, and implementing a distribution of characteristics, $\tilde{F}_{X,p'}$, as implied by our thought experiment. This counterfactual income distribution can be obtained under the re-weighting approach by suitably re-weighting the cross-sectional data in the base year. In our implementation, the characteristics X include age and a dummy for college education.

We construct two sets of re-weighting factors. The first set of factors re-weigh the data in the base year to match the marginal distribution of age in the target year. The second set of re-weighting factors adjusts the share of college-educated household heads in each age group to account for the fact that any given age corresponds to a more recent birth cohort in the target year compared to the base year. For birth cohorts older than 26 in the base year, we assume that their college share remains fixed. For birth cohorts that turn 26 only after the base year, we assume that they have the same college share as the youngest cohort in the base year.

For each pair of base and target year $(\bar{p}, p')$, the first set of age-specific re-weighting factors, $\phi_{\bar{p},p'}(a)$, is given by the ratio of the marginal density of age in the target year to the density in the base year

$$\phi_{\bar{p},p'}(a) = \frac{dF_{A,p'}(a)}{dF_{A,\bar{p}}(a)}. \tag{1}$$

Using the fact that birth cohort equals year minus age, the second set of age and education-specific re-weighting factors, $\psi_{\bar{p},p'}(a,e)$, can be written as

$$\psi_{\bar{p},p'}(a,e) = \begin{cases} \frac{dF_{E|A,\bar{p}}(e|a-(p'-\bar{p}))}{dF_{E|A,\bar{p}}(e|a)} & \text{for } a \geq a_0 + p' - \bar{p} \\ \frac{dF_{E|A,\bar{p}}(e|a_0)}{dF_{E|A,\bar{p}}(e|a)} & \text{for } a < a_0 + p' - \bar{p} \end{cases} \tag{2}$$

where $F_{E|A,\bar{p}}(e|a)$ denotes the conditional distribution of education, E, given age, A, in the base year, $\bar{p}$, and a_0 is set to 26. Since in our data sets age and education take discrete values, we can use relative frequencies as estimators for densities[9].

We can study the compositional effects of demographic change by multiplying the sample weight of each observation in the base year by the product of the two re-weighting factors, which results in the desired counterfactual distribution of income for the target year. In particular, we can compute

[9] DiNardo et al. (1996) suggest estimating densities using a regression analysis and Bayes theorem. Their approach is equivalent to ours when densities are estimated using a fully saturated model. Estimating a fully saturated model is feasible in our context, because the common support assumption holds.

the counterfactual Gini coefficient $\tilde{G}_{\bar{p},p'}$ as

$$\tilde{G}_{\bar{p},p'} = \frac{\sum_{i}^{N^{\bar{p}}} \sum_{j=i+1}^{N^{\bar{p}}} w_i w_j |y_i - y_j|}{\sum_{i}^{N^{\bar{p}}} w_i \sum_{i}^{N^{\bar{p}}} w_i y_i}, \tag{3}$$

where $(y_i)_{i=1}^{N^{\bar{p}}}$ is the vector of income observations from the base year $\bar{p}$ and $w_i = \omega_i \phi_{\bar{p},p'}(a_i)\psi_{\bar{p},p'}(a_i, e_i)$, where ω_i is the sample weight for observation i in the base year.

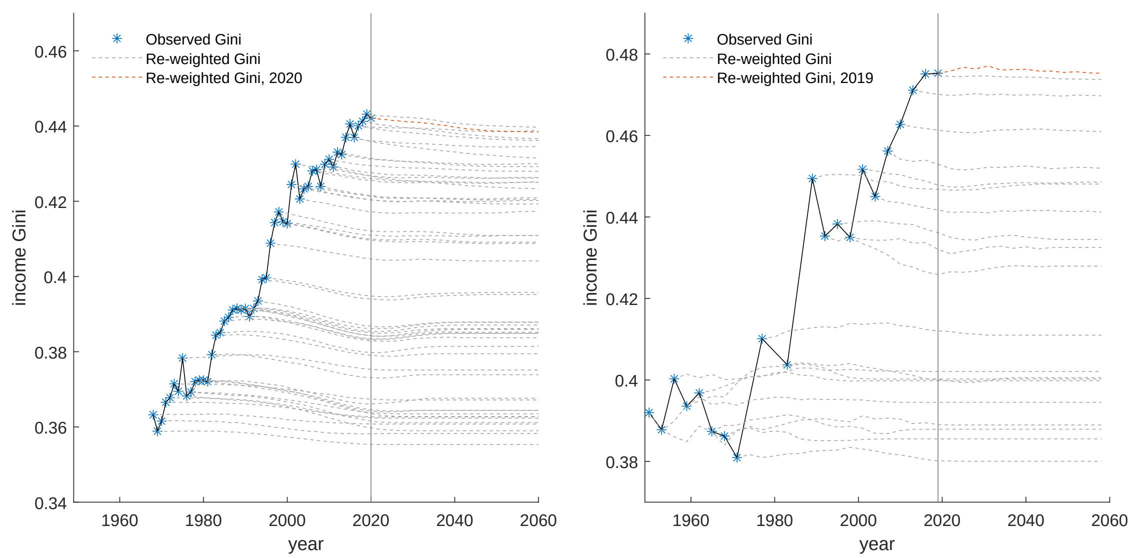
Because we only change the weights and not the underlying observations in the base year, the income distributions within population subgroups are held fixed and the overall income Gini coefficient can only change as a result of changing shares of different subgroups. Hence, as long as cohort differences in income distributions are captured by age and education of the household head, this exercise corresponds to the thought experiment of fixing the economic environment in its state in the base year and only allowing the compositional effects of demographic change to shape income inequality in the subsequent years.

In figure 2, we plot the actual evolution of the income Gini coefficient together with counterfactual evolutions for different base years. We show the results for both the CPS and the SCF+ data for comparison. The blue stars in panel (a) show the observed evolution of income inequality in the CPS data between 1968–2020. The dashed lines starting from different base years show the evolution of income inequality driven by demographic change. It appears that demographic change has contributed little to the observed increase in income inequality. We find similar results for the SCF+ data as shown in panel (b).

In both panels of figure 2, we also plot the predicted evolution of income Gini coefficient until the year 2060 by using population projections to infer the predicted age structure of household heads in future waves of the survey data. The red dashed lines show the prediction based on the most recent survey year in the corresponding data set. While income inequality has risen sharply in the last five decades, the re-weighting analysis suggests that demographic change will not further increase inequality in the future.

II.1 Shortcomings of the re-weighting analysis

The re-weighting method gets the effect of demographic change right only if the ignorability assumption is satisfied, that is, if the conditional distribution of unobserved characteristics does not change over time. In our case, this means that all characteristics which affect a cohort's income distribution can be summarized by that cohort's share of college-educated household heads. There are many reasons to doubt the validity of the ignorability assumption. Hsieh et al. (2019) argue that the last century has seen important improvements in the degree of positive sorting into occupations based on talent. They find that improved sorting explains a large fraction of observed income growth. Similarly, Hendricks and Schoellman (2014) provide direct evidence of improved sorting into education based on ability, and argue that this can explain all of the observed increase in the college premium between the years 1950 and 2000. Because we do not observe ability, the



(a) CPS

(b) SCF+

Figure 2: The effect of demographic change as suggested by the re-weighting analysis. The blue stars show the observed evolution of income inequality. The dashed lines starting from different base years show the evolution of income inequality driven by demographic change. The dashed red lines highlight the counterfactuals following the most recent survey wave in each data set.

ignorability assumption requires that changes in the degree of sorting are fully accounted for by changes in the share of college-educated household heads in each birth cohort.

Unlike the share of college-educated households heads in the population, which has been rising steadily for the past 60 years, growth in the college share by cohort has been very slow since the late 1960s. As a result, birth cohorts born in the second half of the 20^{th} century are fairly similar in their share of college-educated households heads. Hence, changes in the degree of positive sorting over the most recent decades cannot be accounted for by changes in college graduation rates. The same arguments applies to curriculum reforms, improvements in the health environment, or trends in the degree of assortative mating, which all imply that the household of a middle-aged college graduate in the year 1980 may not be a good stand-in for the household of a middle old college graduate in the year 2020.

To make things worse for the re-weighting method, population aging requires the researcher to increase the weights on older households in order to match the evolution of the population age structure. This is problematic because older individuals in any given cross-section belong to earlier birth cohorts. Any attempt to match the evolution of the age structure therefore implies up-weighting households whose unobserved characteristics are furthest from those of the typical household in the target year. To the extent that not all relevant cohort-level characteristics can be summarized by the share of college-educated households heads, the results from our non-parametric re-weighting analysis may be severely misleading.

III A parametric method

In this section, we develop a novel parametric approach to implement our thought experiment of holding the economic environment fixed while allowing demographic change to take place. We proceed in three steps. First, we estimate the life-cycle profiles of average income and income inequality for different birth cohorts and education groups. Second, we use our estimates to construct counterfactual moments for income distributions at the subpopulation level according to our thought experiment. In particular, we predict how income distributions evolve as cohorts age when the economic environment is held fixed. In line with our thought experiment, we assume that cohorts entering the economy after any given base year have the same income distribution and college share as the youngest cohort in the base year. Third, we use predicted population shares to construct a population-level income distribution and plot the evolution of income inequality. In the remainder of this section, we discuss each of these steps in detail.

III.1 Estimating age, period, and cohort profiles

We assume that two key moments of the income distribution, the logarithms of the mean and the Gini coefficient, are described by an additively-separable age-period-cohort model. This allows us to use repeated cross-sections in the CPS and SCF+ data sets to estimate how the distribution of income differs across birth cohorts and how it evolves within cohorts as they age. We motivate the age-period-cohort model by showing that a simple income process leads to additively separable age, period, and cohort profiles in the logarithms of mean income and the income Gini coefficient.

III.1.1 A simple income process

In each period, each household with observable characteristics $X = \{\text{age}, \text{education}\}$ experiences an income shock that has two components: a level component and an inequality component. The level component increases or decreases all incomes by a given factor while leaving inequality between the households unchanged. The inequality component stretches or compresses the income distribution while leaving the average income unchanged. In particular,

$$y_t = (1+\alpha)\Big(y_{t-1} + \beta\big(y_{t-1} - \mathbb{E}[y_{t-1}]\big)\Big)$$

where $\alpha \in [-1, \infty)$ and $\beta \in [-1, \infty)$ denote the level and inequality component respectively, and $\mathbb{E}[y_{t-1}]$ is the mean income in the demographic group. A positive α means that the shock increases average income, while a positive β means that the shock increases inequality. Negative value of α and β achieve the opposite.

We assume that income shocks are education-specific and separable in age and period, so that

$$\alpha(e, a, t) = \alpha_a(e, a)\alpha_t(e, t),$$
$$\beta(e, a, t) = \beta_a(e, a)\beta_t(e, t).$$

Finally, we allow income distributions to differ between birth cohorts due to the distribution of unobserved but fixed characteristics like assortative mating, sorting into occupation, or the quality of education.

With this income process, we can model how changes in the technological and institutional environment affect the distribution of income in the economy. For example, skill-biased technical progress that disproportionately increases the incomes of high earners corresponds to an income shock where α_t and β_t are both greater than zero. A recession on the other hand corresponds to an income shock where α_t is negative. This type of income shocks also allows us to model life-cycle dynamics. For example, accumulation of labor market experience that complements skills is captured by positive α_a and β_a.

The income of a household headed by an individual with education e at age a in time t is given by

$$\begin{aligned} y^e_{i,a,t} = &\prod_{k=1}^{a}\Big(1+\alpha_a(e,k)\Big)\prod_{m=t-a+1}^{t}\Big(1+\alpha_t(e,m)\Big) \\ &\times\left[y^e_{i,0,t-a} + \prod_{k=1}^{a}\Big(1+\beta_a(e,k)\Big)\prod_{m=t-a+1}^{t}\Big(1+\beta_t(e,m)\Big)\big(y^e_{i,0,t-a} - \mathbb{E}[y^e_{a,t}]\big)\right]. \end{aligned}$$

This income process implies education-specific and additively separable age, period, and cohort profiles for the logarithms of mean income and the income Gini coefficient[10], so that

$$\ln\left(\mathbb{E}[y^e_{a,t,c}]\right) = \sum_{k=1}^{a} \ln\left(1+\alpha_a(e,k)\right) + \sum_{m=1}^{t} \ln\left(1+\alpha_t(e,m)\right) + \ln \mu_0^{e,c},$$

$$\ln\left(G(y^e_{a,t,c})\right) = \sum_{k=1}^{a} \ln\left(1+\beta_a(e,k)\right) + \sum_{m=1}^{t} \ln\left(1+\beta_t(e,m)\right) + \ln G_0^{e,c},$$

where $\mu_0^{e,c}$ and $G_0^{e,c}$ are the initial mean and the Gini coefficient of birth cohort c with education e before receiving any income shocks[11]. Hence, we can estimate these profiles using a standard age-period-cohort model.

III.1.2 Age-period-cohort model

We partition our main sample into year-by-age-by-education sub-samples and compute the mean and Gini coefficient in each sub-sample. As a result, we obtain two balanced panels in age and survey waves – one for households with a college-educated household head and another for households without a college-educated household head. We model the income moments as being generated by additive age, period, and cohort effects.

The model can be written as

$$M_{apc} = \alpha_a + \pi_p + \kappa_c + \varepsilon_{apc}, \tag{4}$$

where M_{apc} is the observed moment at age a, in period p, and in cohort c. The age, period, and cohort effects are captured by α_a, π_p, and κ_c, respectively. There is also a mean zero error term, ε_{apc}, which captures both sampling variance and unmodeled noise[12]. We estimate this model separately for college and non-college educated households. Note that in this model we are not imposing any functional form on the age, period and cohort profiles. The additive separability of the model together with repeated cross-sectional data allows us to distinguish the effects of age,

[10] The additive separability for the logarithm of the Gini coefficient follows from the fact that the inequality shock multiplies the Gini coefficient by $1+\beta$, see Heikkuri and Schief (2022).

[11] Note that additive separability does not strictly require that the income of each household evolves according to this income process as long as the percentiles of the income distributions follow this process. Put differently, only the distribution of incomes but not the positions of individual households in it matters.

[12] The average sample size for a given survey year in the SCF+ data is only about 43% of the average sample size in the CPS data, causing larger sampling variation in the estimated mean income levels and income Gini coefficients at the cohort-year level. Sampling variation induces classical error in our left-hand side variables and does not bias our results. However, the lack of precision in the SCF+ data still leads to unwelcome uncertainty in our estimation results, especially for very early and very late birth cohorts that are observed less often in our sample. We address this problem by including neighboring age groups when estimating average income levels and income Gini coefficients at the cohort-year level in the SCF+ data. For example, when we compute the income Gini coefficient for birth cohort 1960 in the year 2001, we include not only households with household heads aged 40-42, but also those with households heads aged 37-40 and 43-46.

period and cohort. The assumption of additive separability seems reasonable for log mean income since multiplicative age, period and cohort effects on income imply additive separability for log mean income. Moreover, if changes in inequality occur through multiplicative shocks to incomes that are linearly related to income percentiles, then it can be shown numerically that the age, period, and cohort effects on the income Gini coefficient are also approximately additively separable[13].

Unfortunately, equation (4) is not identified and cannot be estimated from the data. An obvious problem is that we need to normalize at least one each of the age, period, and cohort effects. A more fundamental identification problem, however, arises because of the linear dependency between age, period, and cohort. The nature of the identification problem can be seen more clearly when we decompose age, period, and cohort effects into two parts – linear trends in age, period, and cohort, and fixed effects that capture deviations from these trends. In particular, if we require that the fixed effects sum to zero and be orthogonal to a trend[14], then the model can be rewritten as

$$M_{apc} = \theta + \alpha\, a + \pi\, p + \kappa\, c + \check{\alpha}_a + \check{\pi}_p + \check{\kappa}_c + \varepsilon_{apc} \tag{5}$$

with the following restrictions on the parameters,

$$\sum_a \check{\alpha}_a\, a = 0 \quad \text{and} \quad \sum_a \check{\alpha}_a = 0, \tag{6}$$

$$\sum_p \check{\pi}_p\, p = 0 \quad \text{and} \quad \sum_p \check{\pi}_p = 0, \tag{7}$$

$$\sum_c \check{\kappa}_c\, c = 0 \quad \text{and} \quad \sum_c \check{\kappa}_c = 0. \tag{8}$$

In this formulation, the overall trend in the age, period, and cohort profiles are captured by the coefficients α, π, and κ, respectively, and θ is a constant. While we still cannot separately estimate all three linear trends in the age, period, and cohort profiles, the deviations from the linear trends are identified and can be estimated from the data even if we do not know what the linear trends are. Moreover, the relationship between the linear trends in the age, period, and cohort profiles is known, so that only a single identifying assumption on any of the linear trends in equation (5) is enough to estimate this model. For example, if we set $\pi = \kappa$, then our estimated parameters are

$$\hat{\alpha} = \alpha^* + \frac{\pi^* - \kappa^*}{2} \tag{9}$$

$$\hat{\pi} = \hat{\kappa} = \frac{\pi^* + \kappa^*}{2} \tag{10}$$

where $\hat{\alpha}$, $\hat{\pi}$, and $\hat{\kappa}$ are the estimated linear trends for age, period, and cohort, and α^*, π^*, and κ^* are the true linear trends in the data generating process.

[13]see Appendix **??** for a more detailed discussion of additive separability.

[14]This normalization of the age, period, and cohort effects is due to Deaton and Paxson (1994b).

Panel A: CPS	(Log) mean income		(Log) income Gini	
	College	Non-college	College	Non-college
N	2,862	2,862	2,862	2,862
R^2	0.95	0.97	0.92	0.90
Shapley decomposition of R^2				
Linear trends	0.23	0.32	0.80	0.74
Nonlinear age effects	0.55	0.43	0.05	0.07
Nonlinear period effects	0.03	0.03	0.03	0.02
Nonlinear cohort effects	0.19	0.22	0.11	0.17
Panel B: SCF+	(Log) mean income		(Log) income Gini	
	College	Non-college	College	Non-college
N	399	399	399	399
R^2	0.94	0.97	0.90	0.91
Shapley decomposition of R^2				
Linear trends	0.15	0.36	0.54	0.46
Nonlinear age effects	0.41	0.23	0.03	0.09
Nonlinear period effects	0.24	0.17	0.19	0.15
Nonlinear cohort effects	0.21	0.24	0.25	0.30

Table 1: Sample sizes and explained variation in the age-period-cohort model.

III.1.3 Estimation results

Our additively separable model is able to account for a large share of the variation in mean incomes and income Gini coefficients at the cohort-year level. Moreover, the estimated model finds important cohort effects in both mean income and income inequality that are not driven by differences across cohorts in the share of college-educated households. Table 1 reports the coefficient of determination, together with a Shapley decomposition of the share of explained variation into the respective parts accounted for by the linear trends in age, period, and cohort, and the deviations from these trends in the age, period, and cohort profiles[15]. Nonlinear cohort effects account for between 10% and 30% of the explained variation in the estimated models, underscoring the importance of cohort differences in the process of demographic change.

In figures 3 and 4, we plot the age, period, and cohort profiles of mean income and income

[15]The table also reports the number of observations in each model, which is given by the number of year-age combinations in each data set.

inequality estimated separately on the CPS and the SCF+ data sets. The profiles are plotted after normalizing the linear trends in cohort and period effects to be of equal magnitude. As discussed in section III.1, nonlinearities in the profiles are identified from the data and do not depend on the normalization of the linear trends[16].

We find that average income is increasing in age until about age 50, and decreasing thereafter. Similarly, income inequality also increases over the working life and flattens out after retirement age. Compared to households without a college-educated household head, households with a college-educated household head experience a steeper increase in income over the working life and a sustained increase in income inequality even at older ages. We estimate very similar age profiles in both data sets.

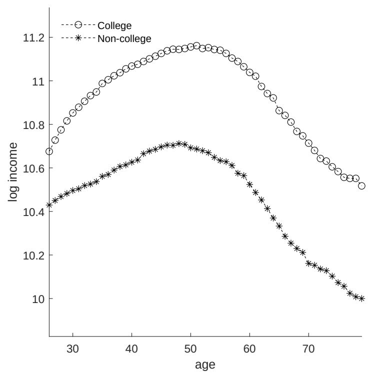
In the period profiles, we document cyclical movement in log mean income reflecting business cycles, which can be seen more clearly in the annual CPS data. The period profiles also document a stark increase in income inequality in the early 1980s among non-college educated households that is followed in the 1990s by a similar increase for college-educated households. Interestingly, while income inequality at the population-level has increased substantially in the past two decades, we find much less of an increase or even a decrease in period effects for income inequality over this time period. Finally, in the SCF+ data where we estimate period effects going back to 1950, we find that period effects for the income Gini coefficient show little trend before the 1980s.

Most strikingly, we find pronounced nonlinear cohort profiles for both average income and income inequality. The cohort effects for log mean income are increasing up to approximately the birth cohort 1947 and become flat or decreasing thereafter. The cohort effects for income inequality, on the other hand, describe a U-shaped profile and are increasing during the second half of the 20^{th} century. These patterns are found in both the CPS and the SCF+ data.

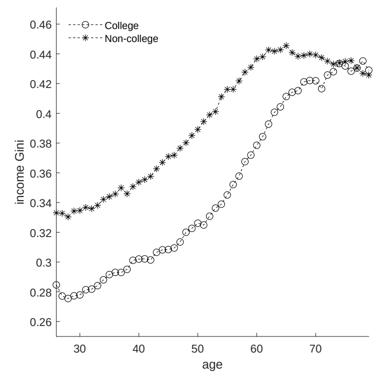
The estimated cohort profiles cast doubt on the validity of our earlier re-weighting analysis. To the extent that our additively separable age-period-cohort models serve as an adequate description of changes in income distributions over time, the estimated cohort profiles clearly show that differences in age and educational attainment alone cannot satisfactorily explain cohort-level differences in income distributions. In this sense the cohort profiles reject the ignorability assumption that was required to interpret the results of the re-weighting analysis.

A potential concern is that the estimated profiles are affected by the fact that we observe different cohorts in different time periods and at different ages, and that we observe some cohorts more often than others. We address this concern by re-estimating the age-period-cohort model on different restricted time windows and comparing the resulting cohort profiles. For both data sets, we find that the estimated profiles from restricted samples align well with the ones from the full sample. We describe the procedure in more detail and show the estimated profiles in appendix B.

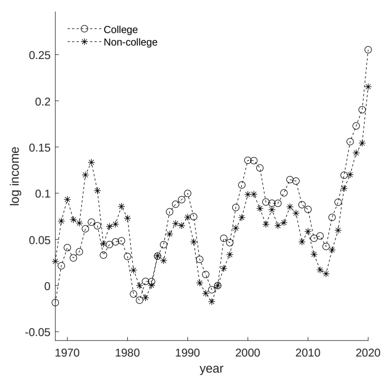
[16]In appendix A, we plot the profiles for alternative normalizations of the linear trends.



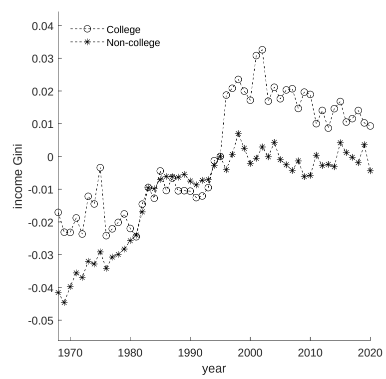
(a) Age profile: log mean income



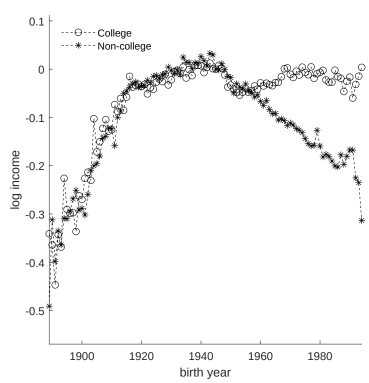
(b) Age profile: income Gini



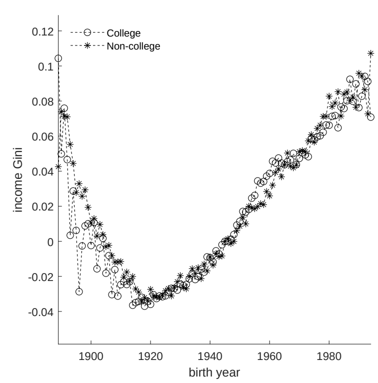
(c) Period profile: log mean income



(d) Period profile: income Gini

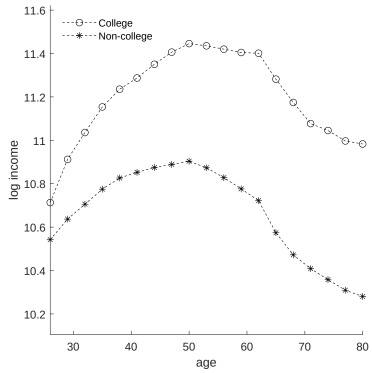


(e) Cohort profile: log mean income

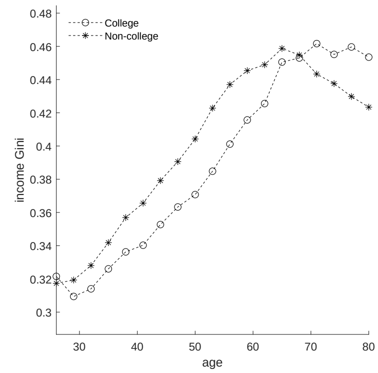


(f) Cohort profile: income Gini

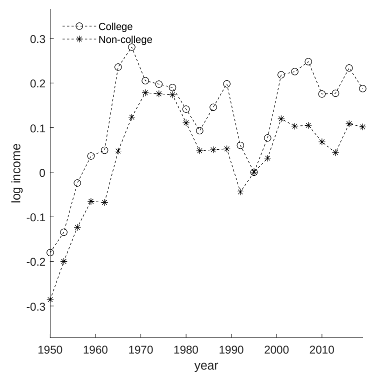
Figure 3: Age-, period-, and cohort profiles of log mean income and income Gini coefficients in the CPS data. The first element in each profile is normalized to zero.



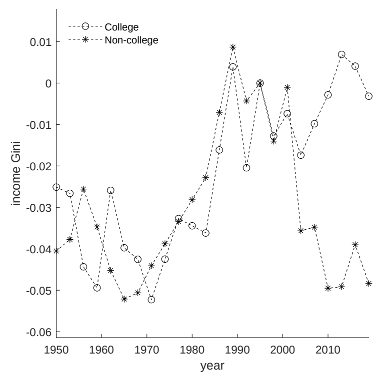
(a) Age profile: log mean income



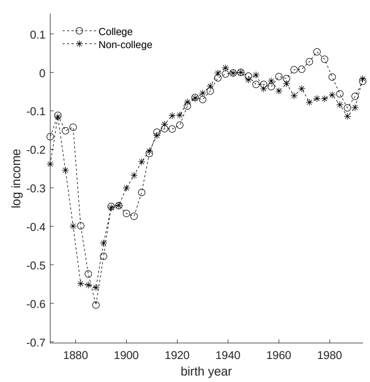
(b) Age profile: income Gini



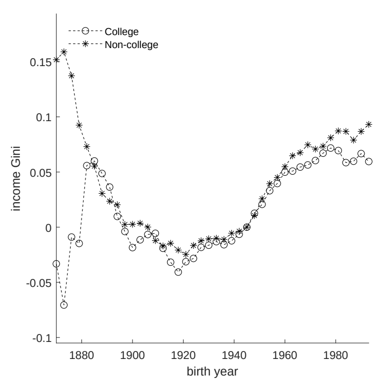
(c) Period profile: log mean income



(d) Period profile: income Gini



(e) Cohort profile: log mean income



(f) Cohort profile: income Gini

Figure 4: Age-, period-, and cohort profiles of log mean income and income Gini coefficients in the SCF+ data. The first element in each profile is normalized to zero. The period fixed effects for the years 1974, 1980, and 1986 are linearly interpolated.

III.2 Constructing counterfactual moments

We use the results from the age-period-cohort model to construct counterfactual income distributions at the cohort level by assigning to each cohort in each year their education-specific cohort and age effects while keeping period effects fixed at the base year level. We set the cohort effects for newly entering cohorts equal to the cohort effect of the youngest cohort observed in the base year.

III.2.1 Normalizing the linear trends

We use both the linear and nonlinear age, period and cohort effects to construct counterfactual moments. While the estimated nonlinear effects are invariant to the chosen normalization, the estimated linear trends depend on it. Hence, the choice of normalization matters for the construction of counterfactual moments.

The normalization does not affect the predicted values of the model, and it can therefore not be estimated from the data. To address this issue, we follow Lagakos et al. (2018) and derive our results under three different normalizations. As extreme cases, we set either the period or the cohort trend in the income Gini coefficient to zero. As an intermediate case, we assume that the trends in period and cohort effects are of equal magnitude and we let the data tell us what these trends are. We treat the intermediate case as our baseline normalization. In the case of log mean income, we always use the baseline normalization. While our results depend quantitatively on the normalization for income Gini, our qualitative findings are invariant to the choice of normalization[17].

It is likely that income growth is due to both period and cohort effects. Technological progress which increases incomes independently of age and education naturally constitutes a trend in period effects. Increasing levels of human capital, to the extent they are generated by more schooling and higher quality of education or cohort-specific improvements in health, on the other hand, are cohort effects. Jones (2002) calculates that almost a third of growth in GDP per capita between 1950 and 1993 can be accounted for by rising educational attainment while roughly two thirds is accounted for by improved productivity[18]. Moreover, Hsieh et al. (2019) shows that a significant share of productivity growth is explained by improved sorting of talent, which occurs between cohorts.

Assessing trends in income inequality is more difficult. To the extent that rising inequality is driven by technology or policies that increase inequality independent of household characteristics, it should be captured by period effects. A reduction in redistribution from rich to poor, for example, is a period effect that increases inequality. On the other hand, increasing inequality due to secular trends in cohort-specific characteristics such as the distribution of human capital, sorting to education and occupations, or positive assortative mating should be attributed to the trend in cohort effects. Since a trend in both period and cohort effects is plausible, we consider the inter-

[17] We have also derived our results under different normalization of the trends in log mean income. The results do not vary meaningfully.

[18] Jones (2021) updates these numbers to one quarter and two thirds for years between 1950 and 2007. The remainder is explained by the increase in the ratio of labor force to population.

mediate normalization of equal period and cohort trends in income inequality a sensible baseline normalization[19].

III.2.2 Implementing the thought experiment

We use the estimated age, period, and cohort effects for log mean income and the income Gini coefficient to construct counterfactual mean, $\tilde{\mu}_{a,p,c,e}$, and Gini coefficient, $\tilde{g}_{a,p,c,e}$, for each sub-population as implied by our thought experiment. We give each age-education-group their corresponding age effect, the period effect of the base year, and the estimated cohort effect for cohorts that are present in the base year, and the cohort effect of the youngest age-group a_0 in base year $\bar{p}$ for cohorts that enter the economy after the base year. In particular, for base year $\bar{p}$ and target year $p' > \bar{p}$, we set the sub-group moments as

$$\tilde{\mu}_{a,p',c,e} = \begin{cases} \exp\left(\theta_e^\mu + \alpha_e^\mu a + \pi_e^\mu \bar{p} + \kappa_e^\mu c + \check{\alpha}_{a,e}^\mu + \check{\pi}_{\bar{p},e}^\mu + \check{\kappa}_{c,e}^\mu + \frac{\sigma_e^2}{2}\right) & \text{if } \; c < \bar{c}_0 \\ \exp\left(\theta_e^\mu + \alpha_e^\mu a + \pi_e^\mu \bar{p} + \kappa_e^\mu \bar{c}_0 + \check{\alpha}_{a,e}^\mu + \check{\pi}_{\bar{p},e}^\mu + \check{\kappa}_{\bar{c}_0,e}^\mu + \frac{\sigma_e^2}{2}\right) & \text{if } \; c \geq \bar{c}_0, \end{cases} \tag{11}$$

$$\tilde{g}_{a,p',c,e} = \begin{cases} \theta_e^g + \alpha_e^g a + \pi_e^g \bar{p} + \kappa_e^g c + \check{\alpha}_{a,e}^g + \check{\pi}_{\bar{p},e}^g + \check{\kappa}_{c,e}^g & \text{if } \; c < \bar{c}_0 \\ \theta_e^g + \alpha_e^g a + \pi_e^g \bar{p} + \kappa_e^g \bar{c}_0 + \check{\alpha}_{a,e}^g + \check{\pi}_{\bar{p},e}^g + \check{\kappa}_{\bar{c}_0,e}^g & \text{if } \; c \geq \bar{c}_0, \end{cases} \tag{12}$$

where $\bar{c}_0 \coloneqq \bar{p} - a_0$ is the youngest cohort present in the base year, superscripts μ and g indicate the statistical moment and subscript e the education group for which the parameters have been estimated, and σ_e^2 is the estimated variance of the error term for log mean income[20].

III.3 An aggregation methodology for the Gini coefficient

To study the evolution of income inequality at the population level, we need to aggregate the predicted sub-population means and Gini coefficients into a population-level Gini coefficient. Unfortunately, the Gini coefficient is not an aggregative inequality measure. That is, it is not sufficient to know the mean, Gini coefficient, and the population share of each subgroup to reconstruct the

[19] Huggett et al. (2011) find that more than 60% of lifetime earnings and wealth inequality is due to characteristics that are fully formed by early adulthood. It is therefore plausible that a significant share of the rising income inequality can be explained by changing cohort-level characteristics.

[20] We estimate the age-period-cohort model on log mean income rather than mean income, so that multiplicative age, period, and cohort effects on income show up as additive effects on log income. To convert predicted levels of log mean income to predicted levels of mean income, we take into account that the expected value of mean income is approximately given by $\exp\left(\lambda + \frac{\sigma^2}{2}\right)$ where λ is the predicted value of log mean income and σ^2 is the variance of the expected value of log mean income, which corresponds to the variance of the error term in the age-period-cohort model. This approximation is exact if the error term is normally distributed.

population-level Gini coefficient (Bourguignon, 1979)[21]. Nor is the population-level Gini coefficient necessarily monotonic in the subgroup Gini coefficients. Cowell (1988) constructs an example of an income distribution and a set of transfers which decrease all subgroup Gini coefficients but increase the population-level Gini coefficient while holding the subgroup means constant.

To overcome this issue, we propose a method to map the moments of sub-population distributions into the population-level Gini coefficient. The idea is to fit a parametric distribution for each set of sub-population moments and aggregate these distributions to generate a population-level income distribution.

We follow the principle of maximum entropy[22] by Jaynes (1957) and assume that income distributions at the sub-population level follow the parametric distribution that maximizes entropy subject to being supported on the positive real line and having given mean and Gini coefficient. This distribution was derived by Eliazar and Sokolov (2010), and we refer to it as the Maximum Entropy (ME) distribution[23].

The cumulative distribution function of the ME distribution is given by

$$F^{\mathrm{ME}}(y;\sigma,\rho) = 1 - \frac{1}{\sigma\exp(\rho y) + (1-\sigma)} \tag{13}$$

where $y \geq 0$ and the parameters σ and ρ are related to the mean income, μ, and the income Gini

[21] It is possible to write the Gini coefficient as a function of sub-population Gini coefficients, means, population shares and a residual term which relates to the overlap between the income levels in different subgroups (see, for example, the discussion in Mookherjee and Shorrocks (1982)). If the subgroups can be strictly ordered by income levels, that is, if there is no overlap between subgroups, then the residual term vanishes. In this special case, the Gini coefficient would be aggregative. However, there is significant overlap in income levels across sub-groups defined by age and education. Although on average the middle-aged households have higher incomes than the young and the old, the high-income households among the young and the old have much higher incomes than the low-income households among the middle-aged. Similarly, college-educated households have higher incomes on average, but the richest non-college educated households are much richer than the poorest college educated households. In fact, applying the decomposition formula in Mookherjee and Shorrocks (1982), the observed magnitude of the residual term varies between 56% and 60% of the income Gini across the waves in the CPS data and between 44% and 57% of the income Gini across the waves in the SCF+ data. Hence, we cannot approximate the population-level Gini coefficient by using the decomposition and setting the residual term to zero.

[22] The principle of maximum entropy states that "in making inferences on the basis of partial information we must use that probability distribution which has maximum entropy subject to whatever is known". Entropy is defined as $-\mathbb{E}\log(p(x_i))$, where $p(x_i)$ is the probability/density of outcome x_i. Intuitively, entropy is the expected value of uncertainty in a random variable's outcomes or the average level of information gained from observing the variable's outcomes. By using a maximum entropy distribution to model within-cohort income distributions, we make the least amount of additional assumptions on the functional form of the income distribution after imposing the mean and the Gini coefficient. This approach therefore minimizes bias resulting from making assumptions that may not be true.

[23] This is a slight abuse of language as the distribution we use here is a particular member of the class of maximum entropy distributions. In appendix E, we use lognormal and gamma as alternative distributions, which are also maximum entropy distributions but for different information constraints.

coefficient, g, as follows:

$$\mu = \frac{\log \sigma}{(\sigma - 1)\rho} \tag{14}$$

$$g = 1 + \frac{1}{\sigma - 1} - \frac{1}{\log \sigma}. \tag{15}$$

Since equations (14) and (15) are invertible for $\mu > 0$ and $0 < g < 1$, we can write the sub-population CDF as a function of the sub-population moments,

$$F_{a,p,c,e}(y) = F^{\mathrm{ME}}(y; \mu_{a,p,c,e}, g_{a,p,c,e}). \tag{16}$$

After fitting all income distributions at the cohort-year level using the observed cohort-level means and Gini coefficients, we sum over the CDFs to construct the population-level income distribution

$$\Phi_p(y) = \sum_{a,e} s_{a,p,e} F_{a,p,c,e}(y), \tag{17}$$

where $s_{a,p,e}$ is population share of age-education-group (a, e) in period p. The population-level Gini coefficient can then be computed as

$$G_p = 1 - \frac{1}{\mu_p} \int_0^\infty \big(1 - \Phi_p(y)\big)^2 dy, \tag{18}$$

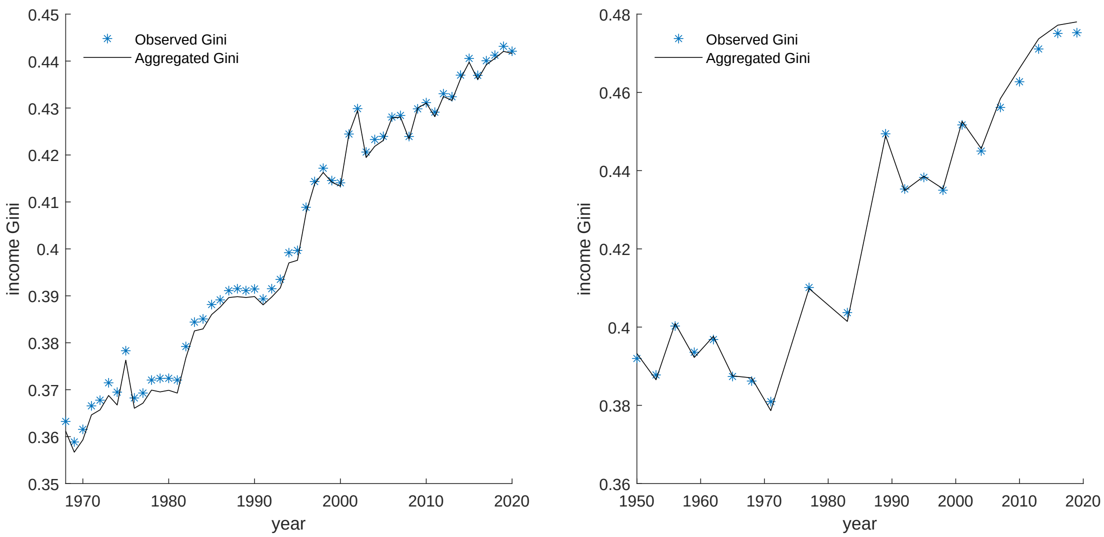
where μ_p is the population-level mean income in period p.

To test our methodology we compute the population-level income Gini coefficient for each survey year by applying our aggregation method to the observed sub-population moments, and compare them with the population-level Gini coefficients computed directly from the data. Figure 5 depicts the results of this comparison. The solid black line shows the aggregated Gini coefficients while the blue stars depict the population-level Gini coefficients computed directly from survey data. The aggregated Gini coefficients follows closely the path of the true Gini coefficients in both the CPS and SCF+ data. This observation lends confidence that our aggregation method is able to aggregate the Gini coefficients with only limited loss of information.

To construct our main counterfactuals, we apply this aggregation methodology to the counterfactual sub-population moments and predicted population shares. Counterfactual sub-population moments are constructed as in equations (11) and (12). Similarly, we construct predicted population shares as

$$\tilde{s}_{a,p',c,e} = \begin{cases} \phi_{a,p'}\, \psi_c & \text{if } \; c < \bar{c}_0 \\ \phi_{a,p'}\, \psi_{\bar{c}_0} & \text{if } \; c \geq \bar{c}_0, \end{cases} \tag{19}$$

where $\phi_{a,p}$ denotes the population share of age group a in year p, which is either observed in the survey data or taken from the census forecasts, ψ_c is the college share of cohort c, which is assumed to be constant after age 26, and $\bar{c}_0$ is the youngest cohort present in the base year.



(a) CPS **(b)** SCF+

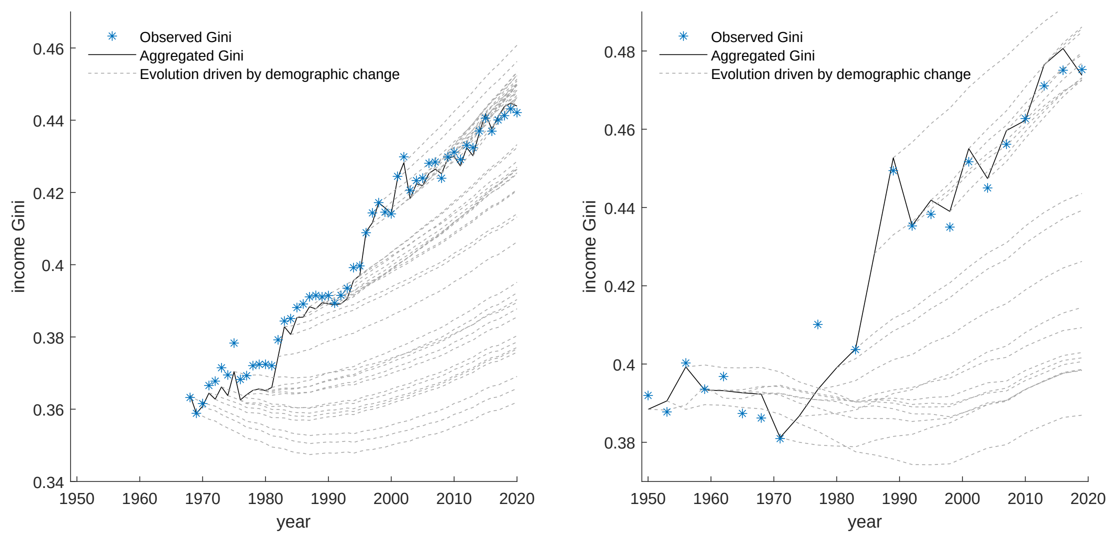
Figure 5: Aggregated Gini coefficients using ME distribution.

IV Results

IV.1 The role of demographic change in the past

Figure 6 shows how demographic change in the past has affected the evolution of income inequality. The blue stars show the actual evolution of income inequality in the CPS and SCF+ data. The solid black line shows the aggregated Gini coefficients using the predicted values from the age-period-cohort model. In contrast to figure 5, differences between our aggregated time series and the observed population-level Gini coefficients now stem from two sources. First, like in figure 5, using parametric income distributions introduces error if incomes at the sub-population level do not follow exactly the assumed parametric distribution. Second, using predicted values from our age-period-cohort models to fit the parametric distributions introduces additional error if the estimated model does not explain all the variation in the data. Overall, we match the shape of the time series well and the differences between the aggregated and the observed population-level Gini coefficients are small.

Starting from each possible base year, a dashed gray line plots how income inequality would have evolved if demographic change unfolded as it actually did but the economic environment was



(a) CPS

(b) SCF+

Figure 6: Counterfactual evolution of income inequality in the past. The blue stars show the observed evolution of income inequality. The dashed lines starting from different base years show the evolution of income inequality driven by demographic change.

held fixed in the base year[24]. In contrast to the results from the re-weighting analysis, we now find that demographic change has had an important effect on the evolution of income inequality in the past. Moreover, it turns out that demographic change has become more important over time. In particular, we find that demographic change has little effect on income inequality if the economic environment is held fixed in the 1950s, 1960s, or 1970s. The counterfactual trajectories of income inequality for those base years are approximately flat and there is no change in the income Gini coefficient over the sample period. The slopes of the counterfactual trajectories are steeper, however, for more recent base years. For example, if the economic environment is held fixed after the mid-1990s, then our counterfactuals not only show an increase in the income Gini coefficient, but demographic change can actually account for all of the observed increase in income inequality. These results are consistent across both data sets.

How much of the actual increase in the income Gini coefficient can be accounted for by demographic change depends on the normalization of the linear age, period, and cohort trends. As we show in appendix C, however, a large share of the observed increase in income inequality since the 1990s can be accounted for by demographic change even if we impose that there is no trend in the cohort profile and we instead allow the period effects to exhibit a strong positive trend.

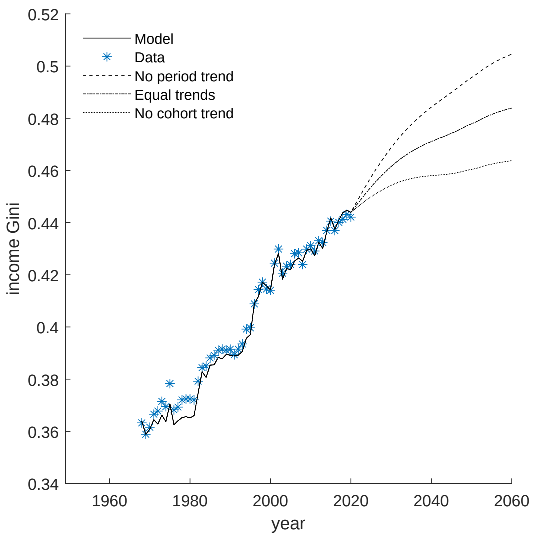
In figure 6, we use age, period, and cohort effects that are estimated on the full sample. As an additional exercise, we compute vintage predictions in appendix D, in which for each base year we only use data up until that year. We obtain similar results, especially for the SCF+ data for which the vintage predictions are almost identical to the past counterfactuals depicted in figure 6.

IV.2 The role of demographic change in the future

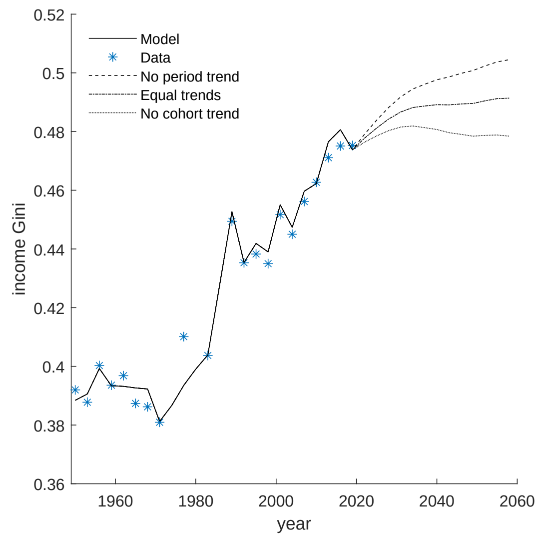
The important role of past demographic change, especially in the recent decades, raises the question whether demographic change will further increase income inequality in the future. To address this question, we choose the most recent survey wave as base year and plot in figure 7 the evolution of income inequality under predicted demographic change until the year 2060. Up until the year of the most recent survey wave, the Gini coefficient is computed using the predicted values for the sub-population moments from our age-period-cohort regressions in section III.1.3. For all years in the future, we use predicted within-cohort income distributions. We plot the evolution of the Gini coefficient in the future under three different normalization, depicted as the dashed gray lines.

In contrast to the re-weighting analysis, we find that demographic change will lead to an increase in income inequality over the next four decades. This is the case in both the CPS as well as the SCF+ data irrespective of whether we attribute trends to cohort effects or period effects, although the increase is more dramatic in the former case. The difference between the three specifications is due to two interlinked factors. First, assuming no period trend in income Gini coefficient in the

[24] To compute the counterfactual evolutions starting from base years 1974, 1980, and 1986, for which we do not have survey waves in the SCF+ data, we linearly interpolate the period fixed effects and the age-education-specific population shares.



(a) CPS



(b) SCF+

Figure 7: Counterfactual evolution of income inequality in the future.

past implies a stronger positive cohort trend. Thus, cohort replacement will put a stronger upward pressure on overall income inequality. Second, the estimated age profile of within-cohort income Gini coefficient is steeper if we assume no period trends, which implies that projected population aging has a greater effect on income inequality.

IV.3 Decomposing the effects of demographic change

We have shown above that demographic change has mattered for the evolution of income inequality in the recent past and will likely further increase income inequality in the future. An interesting question is whether projected demographic change will increase income inequality predominantly through the effect of population aging or cohort replacement. To investigate the respective contributions of these two channels, we construct additional counterfactuals where we shut down either the population aging or the cohort replacement channel. These counterfactuals are computed for the baseline normalization where we assume equal trends in period and cohort effects in both log mean income and income Gini coefficient.

To isolate the effect of cohort replacement, we fix the marginal age distribution of the population. In particular, for each target year p', we construct the counterfactual population-level CDF as

$$\tilde{\Phi}_{p'}(y) = \sum_{a,e} \tilde{s}_{a,p',e} \tilde{F}_{a,p',c,e}(y), \tag{20}$$

where counterfactual income distributions, $\tilde{F}_{a,p',c,e}(y)$, are constructed as before, and population shares, $\tilde{s}_{a,p',e}$, are constructed as

$$\tilde{s}_{a,p',c,e} = \begin{cases} \phi_{a,\bar{p}}\,\psi_c & \text{if } \; c < \bar{c}_0 \\ \phi_{a,\bar{p}}\,\psi_{\bar{c}_0} & \text{if } \; c \geq \bar{c}_0, \end{cases} \tag{21}$$

where $\phi_{a,\bar{p}}$ is the population share of age group a in the base year $\bar{p}$, and ψ_c is the college share of cohort c.

To isolate the effect of population aging, we remove all cohort differences and allow only population shares of different age groups to change. In particular, we first equalize college shares across birth cohorts by setting the college share in each cohort equal to the aggregate college share in the base year. We then equalize cohort effects by setting cohort effects equal to a common cohort effect, $\bar{\kappa}_{\bar{p},e}$, in each education group e. The common cohort effects are chosen such that the predicted aggregate Gini coefficient in the base year remains unchanged. We thus set means and Gini coefficients as

$$\tilde{\mu}_{a,p',c,e} = \exp\left(\theta_e^{\mu} + \alpha_e^{\mu} a + \pi_e^{\mu}\bar{p} + \check{\alpha}_{a,e}^{\mu} + \check{\pi}_{\bar{p},e}^{\mu} + \bar{\kappa}_{\bar{p},e} + \frac{\sigma_e^2}{2}\right) \tag{22}$$

$$\tilde{g}_{a,p',c,e} = \theta_e^{g} + \alpha_e^{g} a + \pi_e^{g}\bar{p} + \check{\alpha}_{a,e}^{g} + \check{\pi}_{\bar{p},e}^{g} + \bar{\kappa}_{\bar{p},e} \tag{23}$$

and form sub-population CDFs by plugging these moments into equation (16). Finally, we construct population-level CDF using

$$\tilde{\Phi}_{p'}(y) = \sum_{a,e} \tilde{s}_{a,p',e}\tilde{F}_{a,p',c,e}(y), \tag{24}$$

where $\tilde{s}_{a,p',e}$ are constructed as

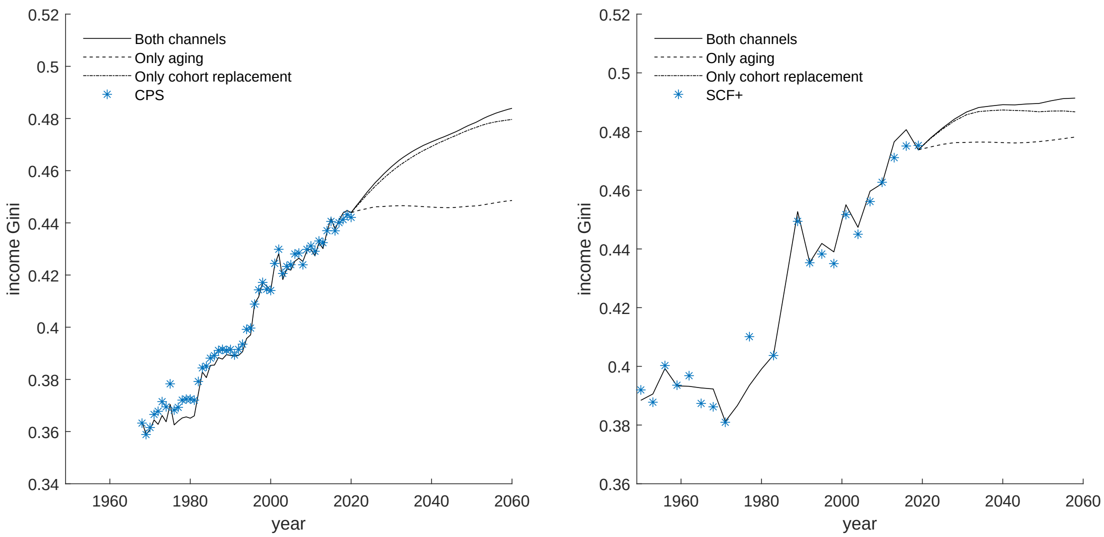
$$\tilde{s}_{a,p',c,e} = \phi_{a,\bar{p}}\,\psi_{\bar{p}}, \tag{25}$$

where $\phi_{a,p}$ is the population share of age group a in year p and $\psi_{\bar{p}}$ is the college share in the population in the base year $\bar{p}$.

Figure 8 plots the future evolution of income inequality under these counterfactuals. The solid line shows the full effect of demographic change under the baseline normalization. The two dashed lines show the isolated effects of population aging and cohort replacement. The main observation from this figure is that cohort replacement is driving most of the increase in income inequality in the future while population aging has a small positive effect.

Having found that population aging will hardly affect income inequality in the future raises the question whether this was also the case in the past. In figure 9 we compare the increase in income inequality over different 21 year periods in our baseline counterfactual to the obtained increase if we shut down either the cohort replacement or the population aging channel.

We find that population aging increased income inequality in the 1950s and in the time period after year 1980. The effect of population aging peaked in the 21 year period starting around the



(a) CPS

(b) SCF+

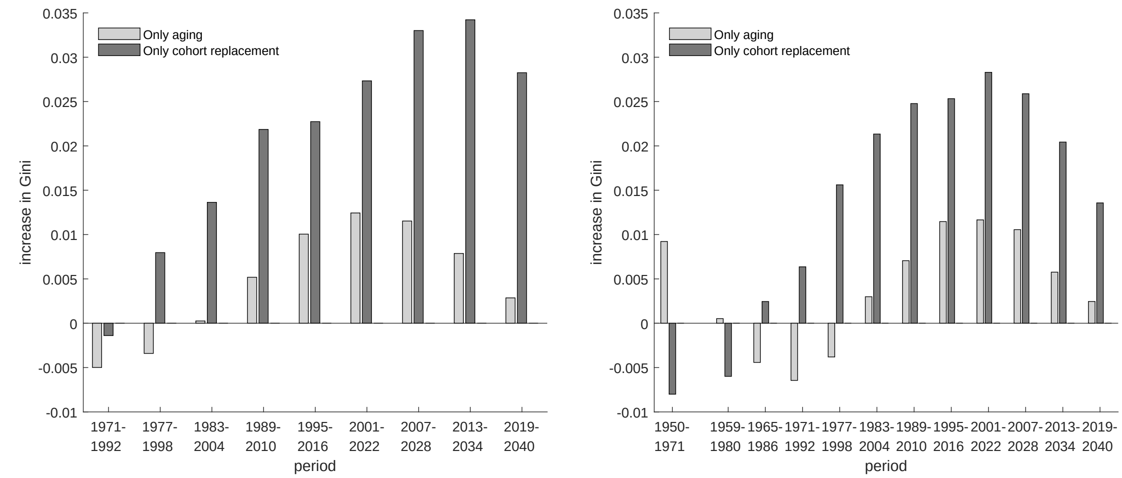
Figure 8: Decomposition of the effect of demographic change in the future.

year 2000 when it explained about one third of the full effect of demographic change[25]. We can also see that the effect of population aging has returned to almost zero by the time period starting with the latest survey wave, which is consistent with the findings in figure 8. Overall, cohort replacement accounts for most of the increase in income inequality driven by demographic change.

V Discussion

In this paper, we study the relationship between demographic change and the evolution of income inequality in the United States. Our approach is to consider a thought experiment where we hold the economic environment fixed and trace out the evolution of income inequality under demographic change. This thought experiment highlights not only the compositional effects of population aging, but also informs how past changes in the economic environment that engender cohort differences can have delayed effects on the evolution of aggregate inequality. A further advantage of this thought experiment is that it can be combined with population projections to study the effect of demographic change not only in the past but also in the near future. The main limitation of our

[25]As the aggregation of age, period, and cohort effects into a population-level Gini coefficient is highly nonlinear, the individual effects of population aging and cohort replacement need not sum up exactly to the full effect of demographic change in our counterfactual. In practice, however, we find that the sum of the effects of the individual channels is not too far from the full effect of demographic change.



(a) CPS

(b) SCF+

Figure 9: Decomposition of the effect of demographic change in the past and future.

approach is that it abstracts from general equilibrium effects of demographic change.

A key challenge in simulating the evolution of income inequality under demographic change is to understand how income distributions change with age and differ between birth cohorts. The re-weighting approach infers age profiles from a single cross-section. This requires that cohort differences can be fully captured by characteristics that are observed in the data. If this is not possible, then the conditional age profile in the cross-section is confounded and the re-weighting analysis will be misleading.

Our parametric methodology estimates age profiles and cohort differences from repeated cross-sections using a flexible age-period-cohort model. This model comes with its own strong assumptions. First, we assume additive separability for both log mean income and the income Gini coefficient. Second, we assume that the estimated age-profile is time-invariant. Third, the results depend on a choice of normalization for the linear trends.

If the assumptions of the age-period-cohort model are met, our parametric methodology has important advantages over the re-weighting analysis. The parametric methodology can handle arbitrary cohort differences in both observed and unobserved characteristics. Moreover, the estimated age profiles are more robust to sampling variation and temporary shocks. Finally, we can separately analyze the effects of population aging and cohort replacement.

An important insight from our paper is that changes in the economic environment are not immediately reflected in the degree of inequality at the population level. In our preferred specification, we find that there has been an inequality-inducing trend in cohort-level characteristics for much of the 20^{th} century, and that the increase in income inequality over the past three decades can be almost entirely accounted for by demographic change. Researchers interested in explaining the rise in income inequality in the United States may read our paper as an invitation to study changes in cohort-level characteristics several decades before the 1970s when income inequality began to rise at the aggregate level.

References

Aguiar, Mark and Erik Hurst, "Deconstructing Life Cycle Expenditure," *Journal of Political Economy*, 2013, *121* (3), 437–492.

Attanasio, Orazio P., "Cohort Analysis of Saving Behavior by U.S. Households," *Journal of Human Resources*, 1998, *33* (3), 575–609.

Auclert, Adrien, Hannes Malmberg, Frederic Marténét, and Matthew Rognlie, "Demographics, Wealth, and Global Imbalances in the Twenty-First Century," *Working paper*, 2021.

Autor, David H, Lawrence F Katz, and Melissa S Kearney, "Trends in US wage inequality: Revising the revisionists," *The Review of Economics and Statistics*, 2008, *90* (2), 300–323.

Bourguignon, Francois, "Decomposable Income Inequality Measures," *Econometrica*, 1979, *47* (4), 901–920.

Burtless, Gary, "Effects of Growing Wage Disparities and Changing Family Composition on the US Income Distribution," *European Economic Review*, 1999, *43* (4-6), 853–865.

Card, David and Thomas Lemieux, "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis," *Quarterly Journal of Economics*, 2001, *116* (2), 705–746.

Cowell, Frank A, "Inequality Decomposition: Three Bad Measures," *Bulletin of Economic Research*, 1988, *40* (4), 309–312.

Daly, Mary C and Robert G Valletta, "Inequality and Poverty in United States: the Effects of Rising Dispersion of Men's Earnings and Changing Family Behaviour," *Economica*, 2006, *73* (289), 75–98.

Deaton, Angus and Christina Paxson, "Intertemporal Choice and Inequality," *Journal of Political Economy*, 1994, *102* (3), 437–467.

Deaton, Angus S and Christina Paxson, "Saving, growth, and aging in Taiwan," in "Studies in the Economics of Aging," University of Chicago Press, 1994, pp. 331–362.

DiNardo, John, Nicole M Fortin, and Thomas Lemieux, "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semi-Parametric Approach," *Econometrica*, 1996, *64*, 1001–1044.

Eika, Lasse, Magne Mogstad, and Basit Zafar, "Educational Assortative Mating and Household Income Inequality," *Journal of Political Economy*, 2019, *127* (6), 2795–2835.

Eliazar, Iddo and Igor M Sokolov, "Maximization of Statistical Heterogeneity: From Shannon's Entropy to Gini's index," *Physica A: Statistical Mechanics and its Applications*, 2010, *389* (16), 3023–3038.

Fortin, Nicole, Thomas Lemieux, and Sergio Firpo, "Decomposition Methods in Economics," in "Handbook of Labor Economics," Vol. 4, Elsevier, 2011, pp. 1–102.

Greenwood, Jeremy, Nezih Guner, Georgi Kocharkov, and Cezar Santos, "Marry Your Like: Assortative Mating and Income Inequality," *American Economic Review*, 2014, *104* (5), 348–53.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L Violante, "Two Views of Inequality over the Life Cycle," *Journal of the European Economic Association*, 2005, *3* (2–3), 765–775.

_ , _ , and _ , "Consumption and Labor Supply with Partial Insurance: An Analytical Framework," *American Economic Review*, 2014, *104* (7), 2075–2126.

Heikkuri, Vesa-Matti and Matthias Schief, "Subgroup Decomposition of the Gini Coefficient: A New Solution to an Old Problem," *working paper*, 2022.

Hendricks, Lutz and Todd Schoellman, "Student Abilities during the Expansion of US Education," *Journal of Monetary Economics*, April 2014, *63*, 19–36.

Hobijn, Bart, Todd Schoellman, and Alberto Vindas, "Structural transformation by cohort," Technical Report, Technical Report, Arizona State University 2019.

Hoffmann, Florian, David S Lee, and Thomas Lemieux, "Growing income inequality in the United States and other advanced economies," *Journal of Economic Perspectives*, 2020, *34* (4), 52–78.

Hsieh, Chang-Tai, Erik Hurst, Charles I Jones, and Peter J Klenow, "The Allocation of Talent and US Economic Growth," *Econometrica*, 2019, *87* (5), 1439–1474.

Huggett, Mark, Gustavo Ventura, and Amir Yaron, "Sources of Lifetime Inequality," *American Economic Review*, 2011, *101* (7), 2923–54.

Jaynes, Edwin T, "Information Theory and Statistical Mechanics," *Physical Review*, 1957, *106* (4), 620.

Jones, Charles I, "Sources of US economic growth in a world of ideas," *American Economic Review*, 2002, *92* (1), 220–239.

_ , "The past and future of economic growth: A semi-endogenous perspective," Technical Report, National Bureau of Economic Research 2021.

Juhn, Chinhui, Kevin M Murphy, and Brooks Pierce, "Wage inequality and the rise in returns to skill," *Journal of Political Economy*, 1993, *101* (3), 410–442.

Kahn, Lisa B, "The Long-Term Labor Market Consequences of Graduating from College in a Bad Economy," *Labour Economics*, 2010, *17* (2), 303–316.

Kuhn, Moritz, Moritz Schularick, and Ulrike I Steins, "Income and Wealth Inequality in America, 1949–2016," *Journal of Political Economy*, 2020, *128* (9), 3469–3519.

Lagakos, David, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman, "Life Cycle Wage Growth across Countries," *Journal of Political Economy*, 2018, *126* (2), 797–849.

Larrimore, Jeff, "Accounting for United States Household Income Inequality Trends: The Changing Importance of Household Structure and Male and Female Labor Earnings Inequality," *Review of Income and Wealth*, 2014, *60* (4), 683–701.

Lee, Donghoon and Kenneth I Wolpin, "Intersectoral labor mobility and the growth of the service sector," *Econometrica*, 2006, *74* (1), 1–46.

Lemieux, Thomas, "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?," *American Economic Review*, 2006, *96* (3), 461–498.

Low, Hamish, Costas Meghir, and Luigi Pistaferri, "Wage Risk and Employment Risk over the Life Cycle," *American Economic Review*, 2010, *100* (4), 1432–1467.

Mookherjee, Dilip and Anthony Shorrocks, "A Decomposition Analysis of the Trend in UK Income Inequality," *The Economic Journal*, 1982, *92* (368), 886–902.

Oreopoulos, Philip, Till Von Wachter, and Andrew Heisz, "The Short- and Long-Term Career Effects of Graduating in a Recession," *American Economic Journal: Applied Economics*, 2012, *4* (1), 1–29.

Porzio, Tommaso, Federico Rossi, and Gabriella V Santangelo, "The human side of structural transformation," Technical Report, National Bureau of Economic Research 2021.

Raaum, Oddbjørn and Knut Røed, "Do Business Cycle Conditions at the Time of Labor Market Entry Affect Future Employment Prospects?," *Review of Economics and Statistics*, 2006, *88* (2), 193–210.

Rothstein, Jesse, "The Lost Generation? Scarring after the Great Recession," *Working paper*, 2019.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek, "IPUMS USA: Version 10.0," *Minneapolis, MN: IPUMS*, 2020.

Schwandt, Hannes and Till Von Wachter, "Unlucky Cohorts: Estimating the Long-Term Effects of Entering the Labor Market in a Recession in Large Cross-Sectional Data Sets," *Journal of Labor Economics*, 2019, *37* (S1), 161–198.

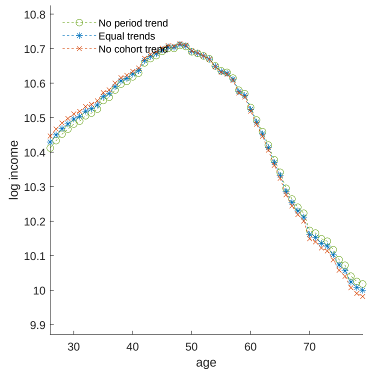
Storesletten, Kjetil, Christopher I Telmer, and Amir Yaron, "Consumption and Risk Sharing over the Life Cycle," *Journal of Monetary Economics*, 2004, *51* (3), 609–633.

United States Census Bureau, "2017 National Population Projections Tables: Main Series," 2017. Data retrieved from the United States Census Bureau, `https://www.census.gov/data/tables/2017/demo/popproj/2017-summary-tables.html`.

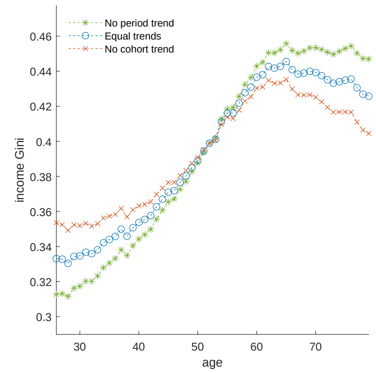
Appendix

A Age, period, and cohort profiles under different normalizations

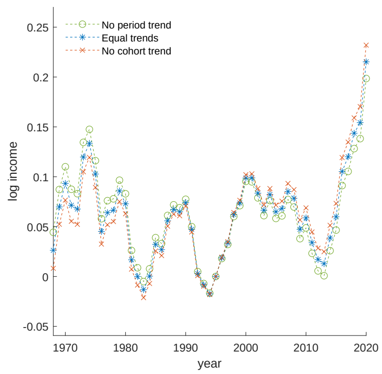
Figures 10 to 13 show the estimated age, period, cohort profiles under alternative normalizations suggeted by Lagakos et al. (2018). The green lines show the profiles under no trend in period effects normalization, the red line shows the profiles under no cohort trend normalization and the blue line shows the profiles under the intermediate case of assuming equal trends in period and cohort profiles. Hence, the profiles plotted in blue correspond to those shown in section III.1.3.



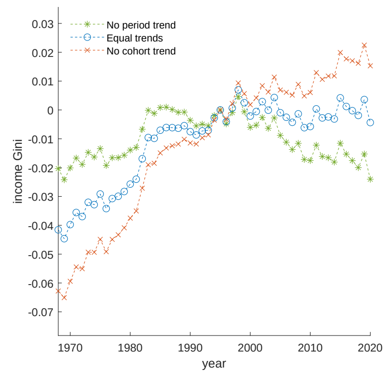
(a) Age profile: log mean income



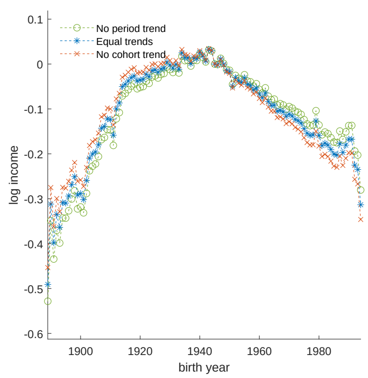
(b) Age profile: income Gini



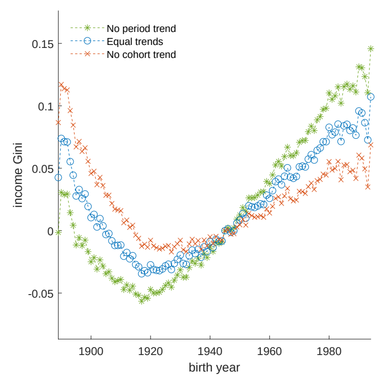
(c) Period profile: log mean income



(d) Period profile: income Gini

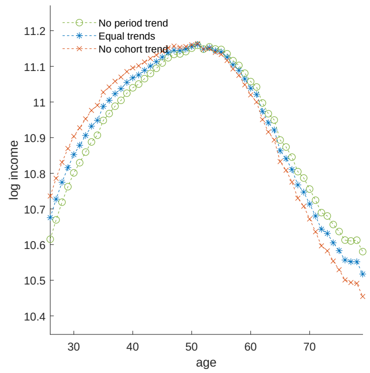


(e) Cohort profile: log mean income

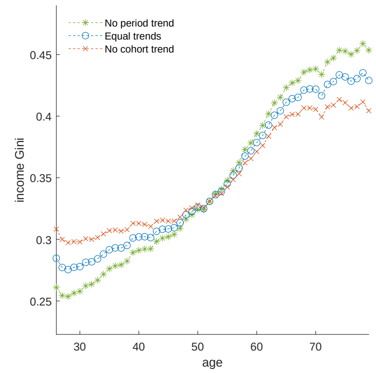


(f) Cohort profile: income Gini

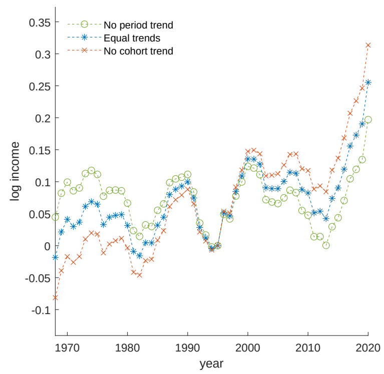
Figure 10: Age-, period-, and cohort profiles of log mean income and income Gini coefficients for households with non-college educated household head in the CPS data under different normalizations for the linear trends. The first element in each profile is normalized to zero.



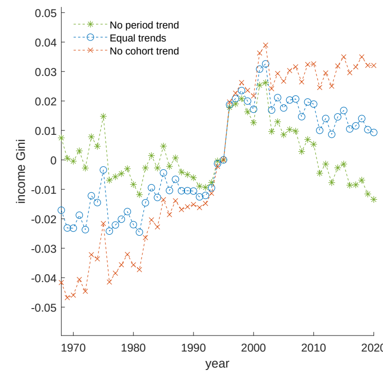
(a) Age profile: log mean income



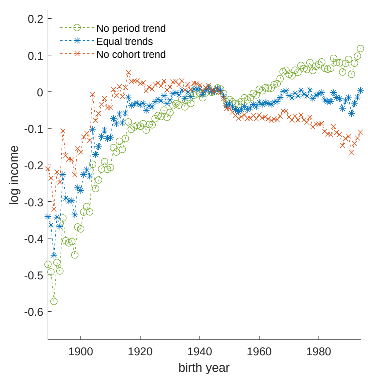
(b) Age profile: income Gini



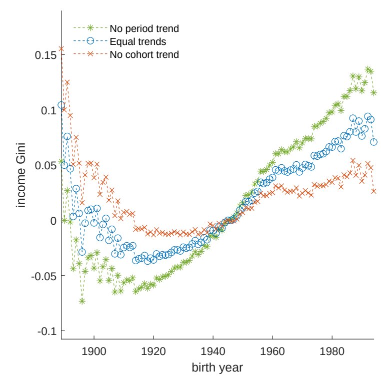
(c) Period profile: log mean income



(d) Period profile: income Gini

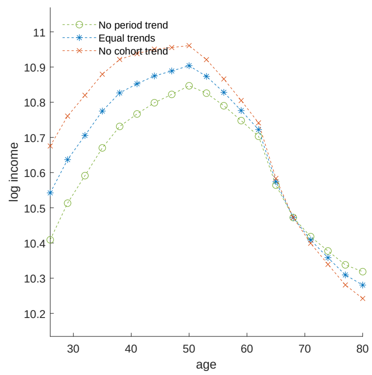


(e) Cohort profile: log mean income

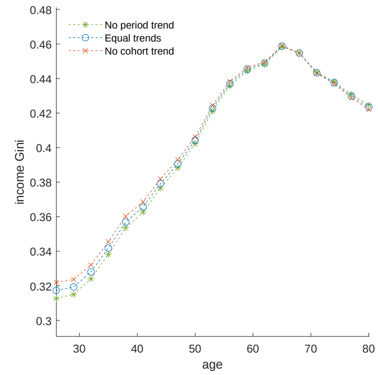


(f) Cohort profile: income Gini

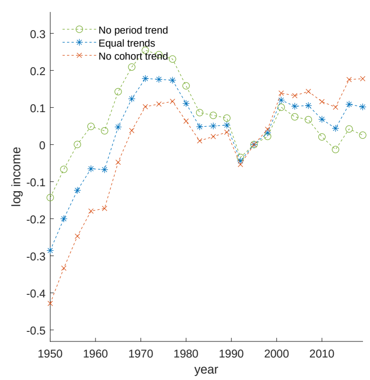
Figure 11: Age-, period-, and cohort profiles of log mean income and income Gini coefficients for households with college educated household head in the CPS data under different normalizations for the linear trends. The first element in each profile is normalized to zero.



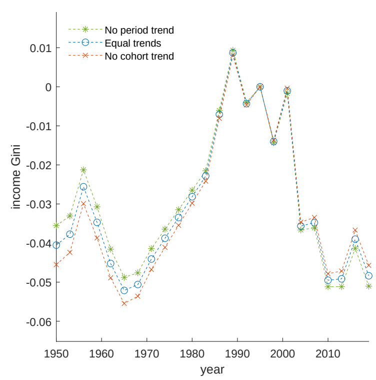
(a) Age profile: log mean income



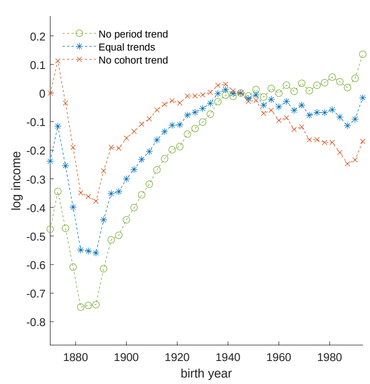
(b) Age profile: income Gini



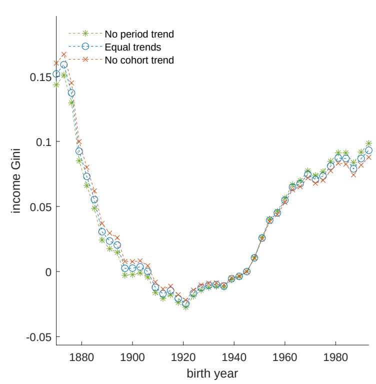
(c) Period profile: log mean income



(d) Period profile: income Gini

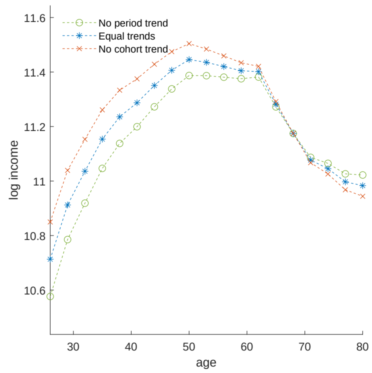


(e) Cohort profile: log mean income

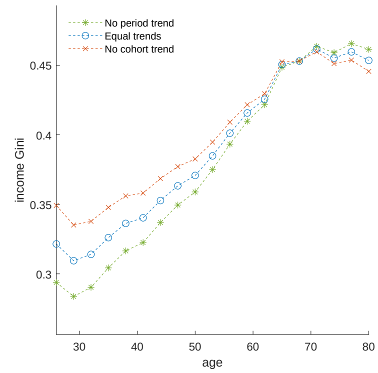


(f) Cohort profile: income Gini

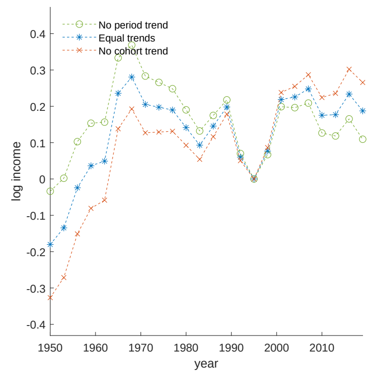
Figure 12: Age-, period-, and cohort profiles of log mean income and income Gini coefficients for households with non-college educated household head in the SCF+ data under different normalizations for the linear trends. The first element in each profile is normalized to zero.



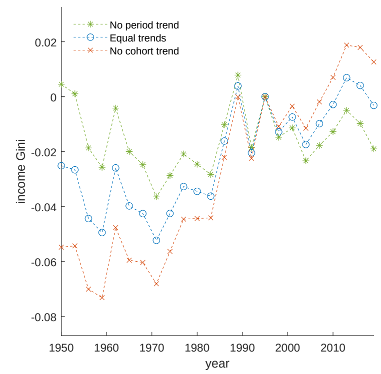
(a) Age profile: log mean income



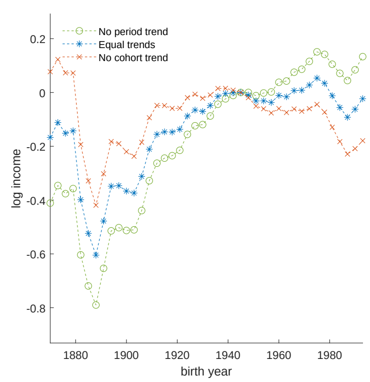
(b) Age profile: income Gini



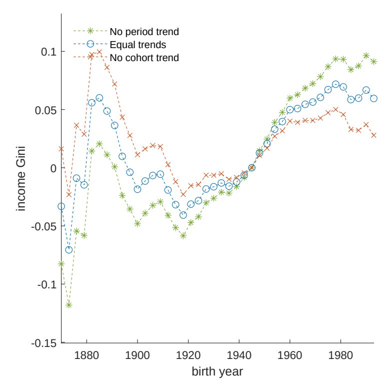
(c) Period profile: log mean income



(d) Period profile: income Gini



(e) Cohort profile: log mean income



(f) Cohort profile: income Gini

Figure 13: Age-, period-, and cohort profiles of log mean income and income Gini coefficients for households with college educated household head in the SCF+ data under different normalizations for the linear trends. The first element in each profile is normalized to zero.

B Robustness of estimation results from age-period-cohort model

We estimate our age-period-cohort model on a balanced sample in age and survey waves. As a consequence, only a fraction of birth cohorts are observed at all ages, while most cohorts are observed when they are either relatively young or relatively old. It also means that we for some cohorts we have much less observations that for others. One may thus be concerned that this panel structure affects our estimation results. For example, if the age profile of income inequality differed across decades, the age-period-cohort model would partially attribute the induced variation to cohort effects. In this case, estimating the age-period-cohort models on different sub-periods would result in different age and cohort profiles.

We address this concern by re-estimating the age-period-cohort model on different restricted time windows and by comparing the resulting age and cohort profiles. Reassuringly, we find that profiles estimated on restricted samples are similar to the profiles estimated on the full sample. In figure **??**, we plot the cohort profiles from estimating the age-period-cohort model on different sub-periods consisting of 10 consecutive waves in the CPS and of 5 consecutive waves in the SCF+ data. For each sub-period, we re-estimate all parameters of the model such that the age and cohort profiles can take different shapes. In each case, we normalize the trend in the age profile to be equal to the trend estimated in the full sample under the normalization of equal cohort and period trends. This normalization, however, does not force the linear slopes of the period and cohort profiles in the models estimated on the sub samples to be identical to the ones in the full model, nor does it constrain the nonlinear age, period, and cohort effects. Reassuringly, we find that the overall shapes of the cohort profiles do not depend on the sub-period used to estimate the age-period-cohort model.

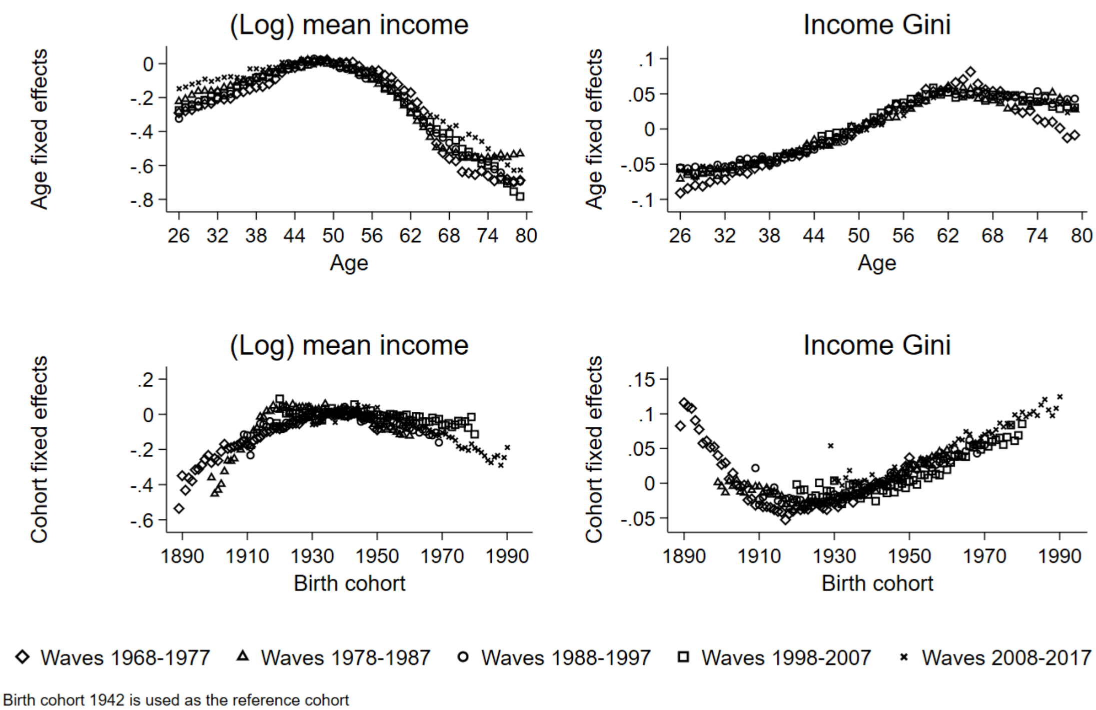


Figure 14: CPS, Non-college

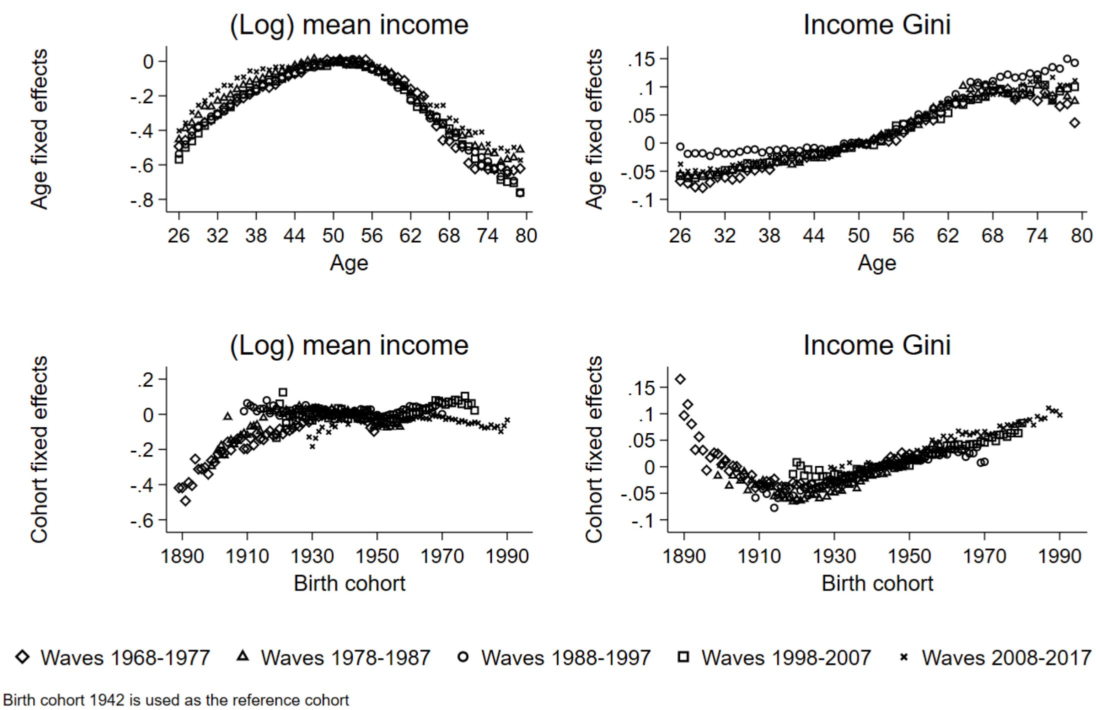


Figure 15: CPS, College

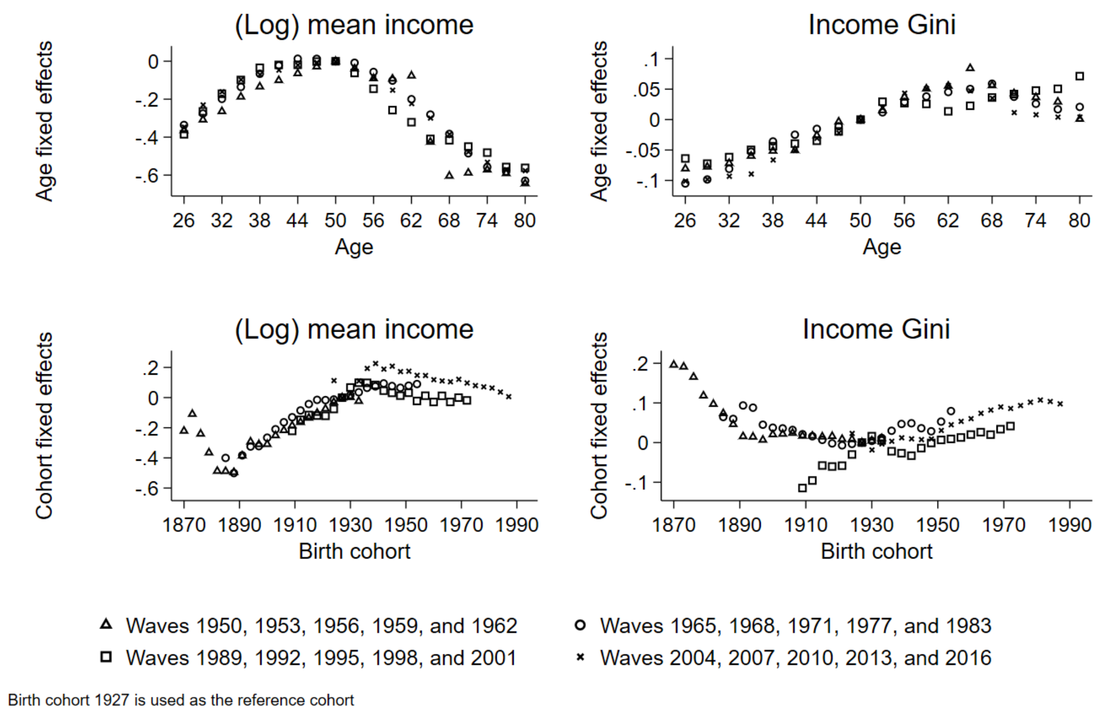
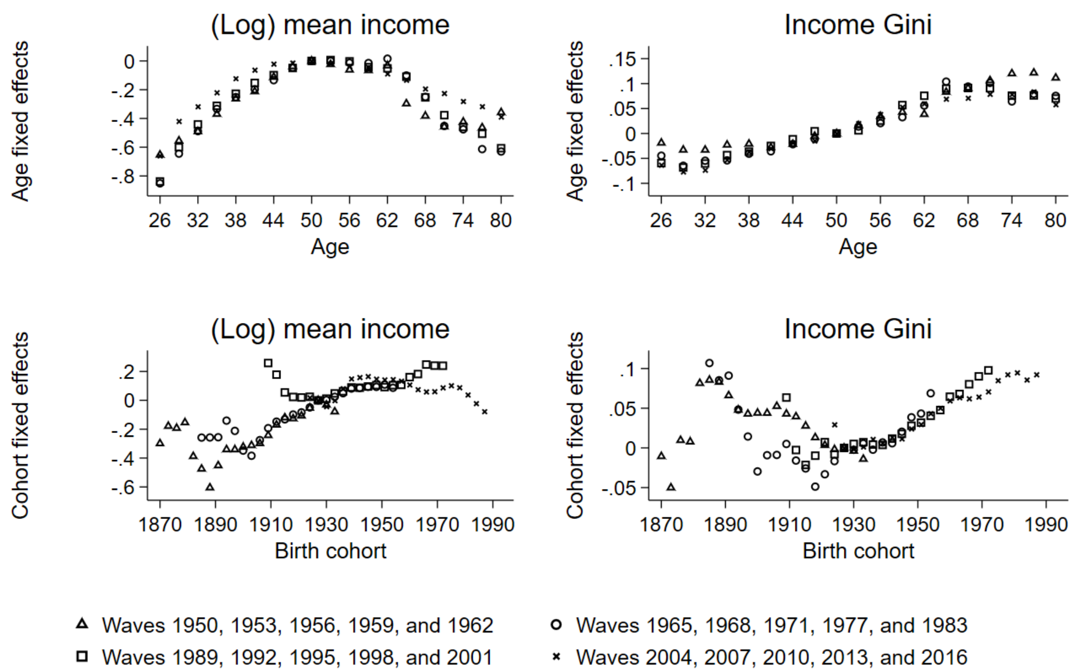


Figure 16: SCF+, Non-college

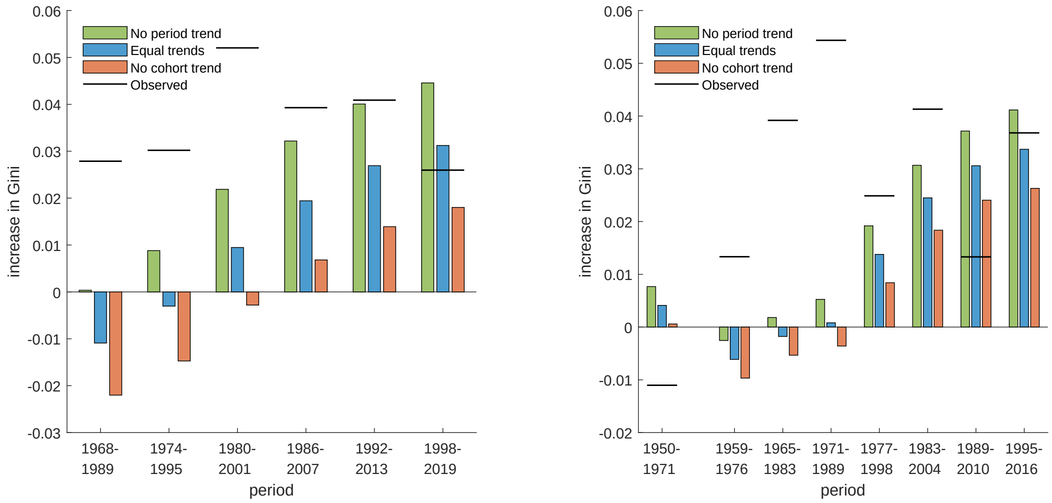


Birth cohort 1927 is used as the reference cohort

Figure 17: SCF+, College

C The role of demographic change in the past under alternative normalizations

In figure 6 in the main text, we show that under our baseline normalization demographic change explains a large share of the observed increase in income inequality in the past – especially since the 1990s. How much of the actual increase in the income Gini coefficient following a given base year can be accounted for by demographic change depends on the normalization of the linear age, period, and cohort trends. In figure 18 we plot the actual increase in the income Gini over an 21 year period following each base year together with the increase in the counterfactuals over the same time period. We again see that the role of demographic change has become more important starting in the 1970s. The role of demographic change is always stronger if we assume that there are no period trends in income inequality over the sample period and the cohort profile is therefore estimated to have a positive trend. Nevertheless, we find that a large share of the observed increase in income inequality since the 1990s can be accounted for by demographic change even if we impose that there is no linear trend in cohort effects and we instead allow the period effects to exhibit a strong positive linear trend.



(a) CPS **(b)** SCF+

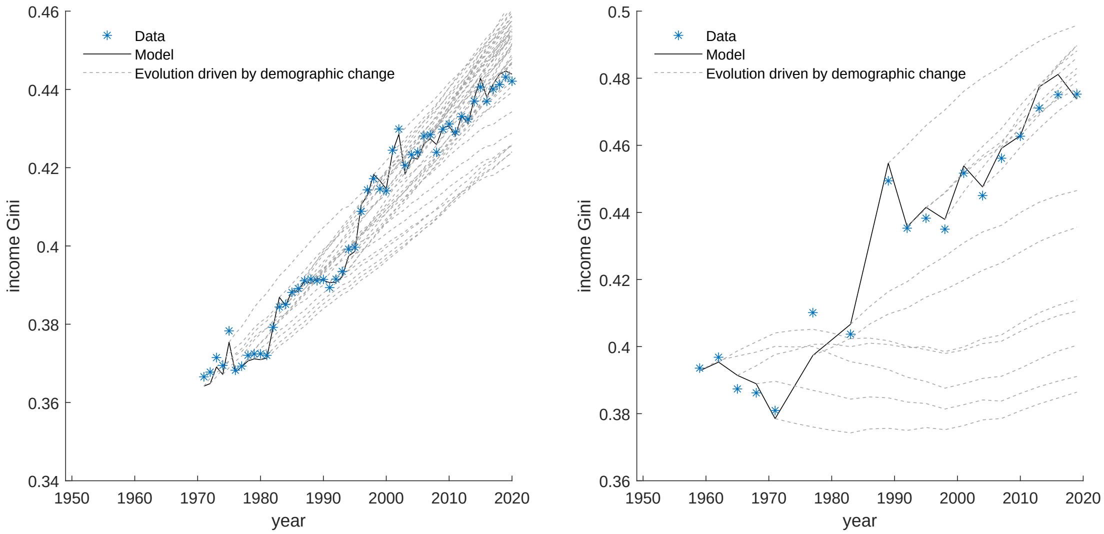
Figure 18: Increase in income Gini driven by demographic change over periods of 21 years.

D Vintage predictions

In the baseline counterfactuals, we use age, period, and cohort effects that are estimated on the full sample. In particular, even when we consider a base year in the past, we estimate the relevant cohort fixed effects from all available years in our data set – including all years after the base year. If the age-period-cohort model describes the data generating process well, this approach is innocent and will increase precision while not biasing the estimates. In appendix B we show that estimating the age-period-cohort model on restricted sub-periods does not appear to affect the estimated nonlinear cohort effects by much.

The baseline counterfactuals do not, however, necessarily correspond to the predictions that we would have made, had we written the paper in the respective base year. Besides the fact that we would not have obtained exactly the same estimates, the normalization of the linear trends also depends on the time period covered in the data. In figure 19, we recompute figure 6 from the main text with the exception that for each base year, we now we only use data up until that year to estimate the age, period, and cohort effects that we use for the predictions. Moreover, for each base year, we force the linear trends in period and cohort profiles to be of equal size. This exercise corresponds to computing vintage predictions, which show the predictions that we would have made, had we written the paper in the respective base year.

Panel (b), which shows the vintage predictions for the SCF+ data, looks remarkably similar to the corresponding panel in figure 6 in the main text. While we would not have predicted any increase in income inequality due to demographic change had we written the paper in the early



(a) CPS data

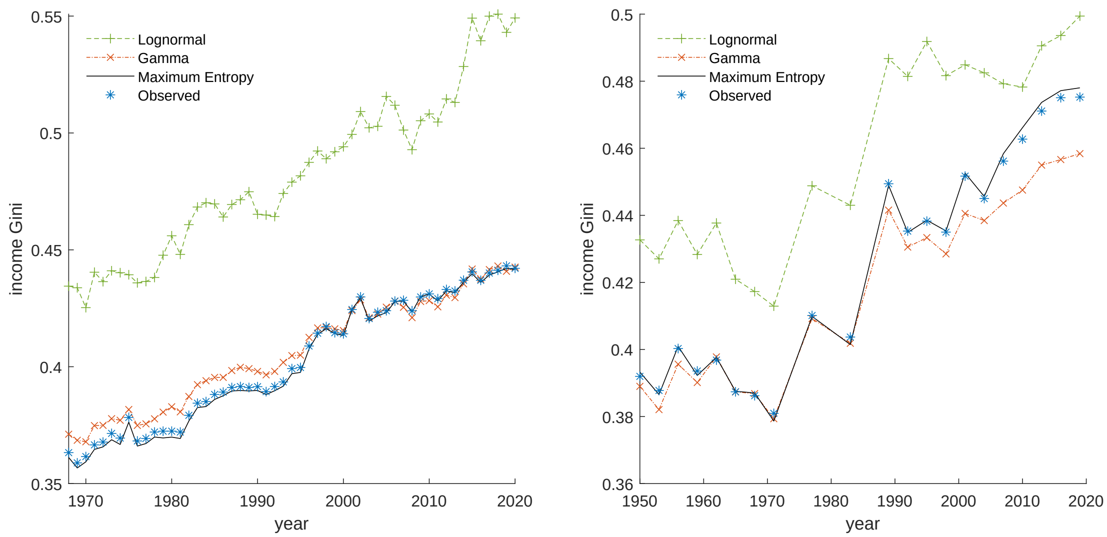
(b) SCF+ data

Figure 19: Vintage predictions

1970s, we would have predicted a steep increase had we written the paper in the early 1990s. In fact, as we find in figure 6 in the main text, the predicted increase in income inequality since the 1990s accounts for all of the observed increase since then.

The findings are somewhat different for the CPS data. While we also would have predicted all of the observed increase in income inequality had we written the paper in the 1990s, we find that we would have predicted a significant increase even if we had written the paper in the early 1970s. This stands in contrast to the findings in the main text and the vintage predictions derived from the SCF+ data. This discrepancy stems from the normalization of the linear trends.

Because the SCF+ data covers a longer time period that includes the 1950s and 1960s during which income inequality did not increase, the linear trends in period and cohort profiles are comparatively small and the choice of normalization matters little. Hence, restricting the data set by dropping later years hardly affects how we normalize the linear trends. The CPS data on the other hand covers a time period throughout which income inequality has increased and the linear trends in period and cohort profiles, which we force to be equal, are therefore steeper. Moreover, we estimate the profile of period effects to be steeper at the beginning of the sample period relative to the end. Hence, restricting the data set by dropping later years combined with the equal trends assumption results in steeper cohort profiles for earlier base years. Consequently, demographic change drives up income inequality more strongly for earlier base years.



(a) CPS

(b) SCF+

Figure 20: Aggregated income Gini coefficients using different parametric distributions.

E Aggregation of Gini coefficients with Maximum Entropy distribution

Figure 5 in the main text shows that we can match the population-level Gini coefficients extremely well by using a parametric distribution to describe sub-population income distributions. We choose the distribution that is supported on the positive real line and maximizes entropy given our estimated moments – mean income and income Gini. In this appendix, we consider alternative two-parameter distributions, lognormal and gamma distribution, to describe incomes at the cohort-year level and compare them to the distribution used in the main text.

The lognormal and the gamma distributions are maximum entropy distributions for given mean and variance of log income, and mean income and mean logarithmic deviation of income, respectively. In figure 20 we show that using these alternative income distributions and targeting their respective characterizing moments at the cohort-year level results in a worse fit for the population-level Gini coefficient, especially in the case of the lognormal distribution.