

LIS

Working Paper Series

No. 686

Anatomy of Income Inequality in the United States: 1979-2013

Aboozar Hadavand

March 2017



CROSS-NATIONAL
DATA CENTER
in Luxembourg

Luxembourg Income Study (LIS), asbl

Anatomy of Income Inequality in the United States: 1979–2013

Aboozar Hadavand

ahadavand@gradcenter.cuny.edu

Department of Economics

The Graduate Center, City University of New York

365 Fifth Avenue New York, NY 10016 USA

Abstract

This paper provides a novel analysis of the trend in income inequality in the United States between 1979–2013. There are two ways in which this paper contributes to the literature. First, I analyze how much of the existing inequality in the U.S. is due to the demographic changes that happened over this period. Using microdata from Luxembourg Income Study and after decomposing inequality into within- and between-age group components, I find that the within-group share of overall inequality in the U.S. is high and steady compared to other developed countries. I also find that about 17 percent of the rise in inequality in this period is due to the between-group component (life-cycle effects). Second, I provide a regression analysis to explain cross-group variations in inequality during the period. I estimate that most of the rise in inequality has happened among middle-aged men while inequality among women, especially among married women has, in fact, decreased. This more granular analysis of inequality can help us investigate the causes of inequality, which would be impossible if we only look at a single inequality statistic.

Keywords: Inequality Decomposition; Within-Group Inequality; Income Distribution

JEL Classification: D31, J11, D63

Acknowledgments

I would like to thank Sarah Thomas, Branko Milanovic, Wim Vijverberg, Janet Gornick, and students at the inequality seminar at the Graduate Center of City University of New York for their insights that greatly assisted this research. I would also like to express my gratitude to David Autor for sharing his workspace automation data. You can find the codes used in this paper here: <https://github.com/ahdvnd/Papers>.

1. Introduction

Many indicators tell us that income inequality in the United States has risen since the late 1970s. Data retrieved from the All the Ginis database show that over the past three decades the Gini coefficient has risen by 8-18 percent depending on what type of income

measures are used.¹ Data from the Luxembourg Income Study (LIS Inequality and Poverty Key Figures, 2016) show that between 1979 and 2013, the ratio of household incomes in the 90th percentile to those in the 10th percentile rose from 4.55 to 5.81. Similarly, the ratio comparing the 80th percentile to the 20th percentile rose from 2.66 to 3.15. One could also compare changes in the median equivalent household income to changes in the mean household income to demonstrate that there has been faster growth at the top of the income distribution relative to the rest of the distribution. Between 1979 and 2013, the median equivalent household income increased by 259 percent, while the mean household income increased by almost 300 percent.

What do all of these measures tell us? They certainly can signal a broad shift in the overall level of inequality, but they do not tell us much about the underlying mechanisms that contribute to such shifts. In order to better understand specific transformations of income distributions across space and time, we must look under the hood of the income distribution and observe what has been happening within and across various subgroups of the populations we are studying. When it comes to the recent rise in income inequality in the United States, it is relevant to first understand how changes in the overall age distribution have affected inequality and to understand the different degrees to which various subgroups of the population have been affected by the recent rise in inequality. This study attempts to analyze and explore beneath the surface of traditional inequality measures. I do this by calculating inequality within and between various cohorts of the American population using Pyatts (1976) decomposition method. I decompose inequality into within- and between-group components, where the within-cohort component is a weighted average of inequality within each cohort and the between component measures inequality that exists between cohorts. I find that the share of the within-cohort component of overall inequality in the United States has been steady and high compared to most developed countries. I also find that about 83 percent of the rise in inequality between 1979-2013 is due to changes in within-cohort inequality.

The importance of such decompositions are multifold, but a few points will help demonstrate why decompositions are central to deepening our understanding of income inequality. First, it is easily recognized that demographic changes are partially responsible for differences in inequality across space and time. The importance of demographic dynamic in inequality is mainly due to the Permanent Income Hypothesis (PIH) as it applies to inequality, which states that inequality within a specific age cohort increases as the cohort ages (Eden, 1980). Therefore, differences in measured inequality between two countries—one with a relatively young population and the other with an aging population for example—could, in fact, be explained by the differences in the demographic traits of each country. Likewise, a single country that ages over time may experience a *natural* rise or fall in inequality that is largely a result of the demographic shift. When looking at differences in inequality across space and time it would be useful to be able to estimate the difference that results from these types of basic demographic incongruities.

Related to this point is the fact that in every country, differences in income due to an age-income profile can partially explain the magnitude of inequality. Such a profile illustrates the typical evolution of income with respect to a person's experience and age. In

1. See <http://bit.ly/1pE3YeX>

most countries, this profile, when graphed, is an arched curve where income peaks around the late middle-aged years. It is considered reasonable that workers are rewarded somewhat proportionally according to their age and/or experience. Thus, we would expect that even in societies that many consider to be very equal, we would find some reasonable differences in income across individuals according to these traits. If this is the case, even under perfect equality of opportunity and a regime of social policies aimed at reducing inequality, the Gini coefficient would never be close to zero. In fact, I calculate that even if every American worker's income follows the cross-sectional trajectory of the life-cycle income in 2013, the Gini coefficient would still be about 13.0.² Many people do not believe in a policy scheme that equalizes the incomes of the young and the mature, and yet, such differences do show up in traditional inequality metrics. One could argue, therefore, that since individuals typically expect to receive higher incomes as they age, inequality due to differences in age can be justified to a degree and that this notion should be reflected in the way we measure and discuss income inequality.

Finally, decomposition analysis can extend to different subgroups of the population based on gender, race, education, and occupation. By measuring inequality within these various groups, we can learn about the effects that different policies or economic phenomena have had in reducing (or increasing) inequality within those cohorts. Experts in the field of inequality usually relate rising inequality to various factors such as financialization, weakening of anti-poverty measures, reductions in redistributive policies, globalization, the growing importance of technical skills, soaring compensations for the top 1 percent (and in general changes in pay norms), dynamic changes in the labor market, immigration, changes in household structure, and reduction in economic mobility (which itself is caused by multiple factors); however, our standard statistics for measuring inequality are single numbers that obscure our understanding of how these many factors combine and contribute to overall inequality. It is these three considerations that have motivated the decomposition analysis in this paper. A second contribution of this paper is a detailed analysis of cross-cohort differences in within-cohort inequality among cohorts categorized by gender, education, and race among other factors.

This study is the first attempt (to my knowledge) that uses a decomposition of the Gini coefficient to study inequality between and within different age, gender, occupational, and educational groups. My analysis reveals a great deal of interesting findings, some of which are consistent with previous inequality studies, and some of which have yet to be discussed in the inequality literature. There have been some attempts in the past to decompose inequality measures according to age and geographic regions. Deaton and Paxson (1994) find a dramatic increase in consumption inequality by age using data from the 1980s in the United States, the United Kingdom, and Taiwan. Heathcote et al. (2005) theorize the impacts cohorts have on the age profiles of inequality. They find that attributing the rising inequality in the United States solely to cohort effects is also misleading and one has to consider time effects as well. Juhn et al. (1993) study more narrowly-defined groups of male workers in the United States and attribute most of the increase in wage inequality for males to increased returns to skills. Osberg (2003) looks at inequality among different age groups of the population in the United States and some selected countries. He argues

2. This calculation is done by using cross-sectional income and age data in 2013. Calculations are not provided here but are available upon request.

that the decline in average family size in recent decades, which is the result of unequal changes at different points in the age distribution, is likely to be responsible for changes in the distribution of income.

In spite of all these studies, there is a lack of coherent analysis of the evolution of inequality in the United States across different subgroups. The main challenge in my analysis is in figuring out how to calculate inequality measures for and between the various subgroups of the population.

In Section 2 of the paper I discuss the data used in my analysis. I use data from the Luxembourg Income Study (LIS) Database, which provides a large set of micro-level data collected from household surveys conducted across several countries. The data has already been harmonized for the sake of cross-country comparisons. Section 3 reviews the academic literature in this area and highlights some of the main attempts that have been made at adjusting and decomposing inequality measures. In this section, I also provide an age-cohort analysis of inequality in the United States. In Section 4, I use regression analysis to study why inequality in some cohorts has increased significantly while in others had declined. I do this while controlling for age differences among these various groups. The paper ends with a brief summary of my findings and a discussion about further applications of this method.

2. Data

The data used in this study is from Luxembourg Income Study (LIS) database, which is one of the largest available income databases of micro-level data collected from multiple countries over a period of decades and are harmonized for cross-country comparisons. The data set contains income (among many other variables) at both the individual and household level. LIS data for the United States is a reconstructed and a harmonized version of the Current Population Survey (CPS). However, between 1979-2013 only years 1979, 1986, 1991, 1994, 1997, 2000, 2004, 2007, 2010, and 2013 are included in LIS.

Discussions around the choice of the unit of analysis should be an integral part of any inequality study. I will justify the use of personal-level data as opposed to household-level data due to the fact that the very purpose of this study is to look at age, gender-specific, educational, and racial groups, which are hard to define for a family or household.³ Furthermore, as Deaton and Paxson (1993) note, “unlike individuals, households form and dissolve over time.” This may lead to shaky results since when we track households through the age of the head of the household, it is hard to assume that the sampling population in successive years remains the same. This is a more binding problem with older households that are more prone to changes such as death, the departure of children from the household, etc.. Another important consideration when using the household as a consumption unit is the choice of the equivalence scale for calculating income per member of the family. For instance, one can simply divide the total household income by the number of household members. However, scholars often suggest to take into account the economies of scale in families. The most used scale by researchers is the square root of the number of household members.⁴ Still, while the choice of the scale can affect the inequality measures, it is

3. Other researches have, nonetheless, used the age, gender, education, and race of the head of the household, which seems irrelevant as it is hard to justify a society that comprises of only “heads of households.”

4. For a thorough analysis of the equivalence scale read Buhmann et al. (1988).

arbitrarily chosen and it is hard to defend that it should be the same across countries. Working with individual level data does not require such arbitrariness in calculation of individual-equivalent incomes. Gottschalk and Smeeding (1997) argue that “economic and demographic decisions within households are endogenous and so complex that empirical research is far from being able to sort out the linkages from individual earnings to household disposable income.”

My decision has its vices too. By using individual level data, I ignore the family structure in which each individual is situated (number of dependents, etc.). It is obvious that such considerations are important in order to understand how much a person needs in terms of annual income. On the other hand, equality of income does not imply equality of standard of living when family structure varies.

Since one of the problems with cross-country comparisons is the heterogeneity in standards of data collection and constituting variables, the LIS data is advantageous since it minimizes those discrepancies and harmonizes the surveys. I use after-tax incomes for individuals of age 20 to 79, where income is defined as the sum of monetary and non-monetary income from labor, and monetary income from capital. I define cohorts along 5-year intervals. Therefore, the youngest cohort in our sample contains individuals in the age range 20-24 and the oldest cohort includes individuals aged 75-79.

3. Inequality within and between age cohorts across the world

In this section, I begin my analysis of inequality in the United States by calculating “between” and “within” inequality measures that are based on age. I will briefly compare the within and between-age-cohort shares of inequality in the United States to the same shares calculated for a number of other countries between 1979 and 2013. To understand the importance of life-cycle effects in inequality calculations, let us first focus on the most widely used measure of inequality, the Gini coefficient, and how it is calculated. The Gini coefficient is based on the cumulative distribution of income and is calculated as the area between the cumulative distribution curve (called the Lorenz curve) and the perfect equality line (the 45-degree line). The Gini coefficient is simple, has an intuitive interpretation, and is nicely scaled between zero and one with a Gini of zero representing a perfectly equal society and a Gini of one representing a perfectly unequal one.

The simplicity of the coefficient, however, glosses over the fact that we may not, in reality, associate perfect equality with a completely equal distribution of incomes as measured at any particular moment in time. The basic intuition behind the 45-degree line of equality (henceforth the line of equality) is that, ideally, every person’s income should be equal to, or at the very least, compared against the average income in society. This would be misleading and could be considered a gross oversimplification if we accept that there are reasonable differences of income in society that are due, in particular, to life-cycle differences in income. If this is the case, even in societies we consider to be very equal, we would not expect to find that all individual incomes are closely aligned with the average societal income. If we want to adjust our inequality measures to account for reasonable life-cycle differences in income, we must start by looking at the overall age-income profile and the overall age distribution

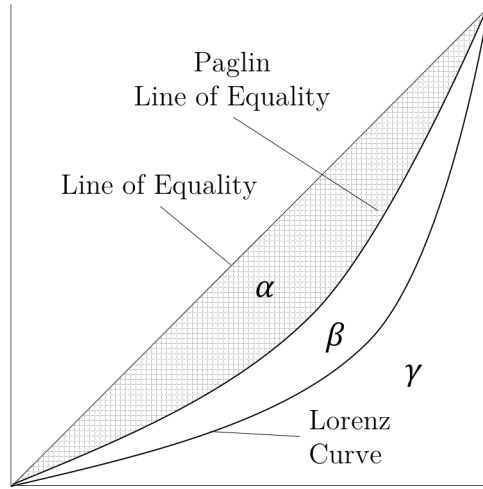


Figure 1: Lorenz Curve

of the societies or countries under our consideration. The overall age-income profile of a society is the aggregate of the individual age-income profiles of all adults.⁵

Paglin (1975) proposed a method to reconstruct the line of perfect equality. This new line, which he called the P-reference line (henceforth the P-line), was defined “in a way which conforms to what casual users of the [Lorenz] curve might infer is the meaning of equality: equal lifetime incomes but not with the added constraint of a flat age-income profile. The P-line is a breakdown of the line of equality for each age cohort. The construction of the new P-line is done by taking the average income in each age group and ranking these groups by their mean incomes. The next step is to calculate the cumulative share of the population and share of the total incomes according to the ranking in the previous step. This new curve is used as the line of equality against the standard Lorenz curve to calculate a modified Gini coefficient, which has been called the P-Gini. The line of equality, the P-line, and the Lorenz curve are depicted in Figure 1. The area α represents the portion of the Gini coefficient that is due to between-group inequalities and the area β represents the area that is due to within-group inequalities.

The improvement from the line of equality to the P-line can be best summarized in an example of two societies represented in Figure 2. Imagine two societies, one with an arched age-income profile A and one with a flat age-income profile B, in which individuals income does not change over the course of their lifetime. Assuming the same distribution of income at any point in time will result in exactly the same Gini coefficient for both societies. However, replacing the line of equality with the P-line yields a lower Gini for society B than society A.

5. In an ideal case, we would want to generalize the evolution of income of a sample of individuals over their lifetime to obtain the societal profile. Therefore, longitudinal data are needed in order to estimate the profile for a society and any attempt to derive the age-income profile in a country based on cross-sectional data suffers from biases due to cross-cohort discrepancies due to differences in years of schooling, post-graduation experiences, economic conditions, technological changes, etc. Obtaining age-income profiles based on longitudinal data is, however, an impossible task simply due to the fact that they are non-existent, at least for cross-country comparisons.

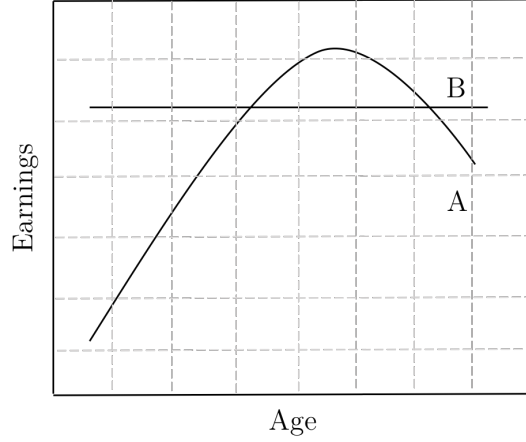


Figure 2: Typical age-income profile

Table 1: Pagin's Gini versus the Lorenzian Gini

Gini Coefficients	Area
Lorenzian Gini	$(\alpha + \beta)/(\alpha + \beta + \gamma)$
P-Gini (within Groups)	$\beta/(\alpha + \beta + \gamma)$
Age-Gini (Between Groups)	$\alpha/(\alpha + \beta + \gamma)$

Table 1 uses the notations on the graph to show the different versions of the Gini coefficient. It is clear that the Lorenzian Gini is the sum of the P-Gini or the within-cohort Gini and the Age-Gini or the between-cohort Gini coefficients. Applying this methodology to family income data from the CPR P-60 series, Paglin showed that the traditional Gini coefficient was 50 percent higher than the P-Gini in 1972.

Since the introduction of Paglin's decomposition approach, other methods have been introduced due to some major caveats⁶ that arise with Paglin's analysis. One of the most important contributions, is a method introduced by Pyatt (1976) which is the method used in this paper. Pyatt used this approach and suggested that the Gini coefficient can be written in terms of the expected value (in the statistical sense) of a game where individuals compare themselves to other randomly drawn individuals from the population. A detailed explanation of the Pyatt's method is explained in Appendix A. As previously mentioned, age cohorts in this analysis are defined by 5-year intervals with the youngest cohort consisting of individuals aged 20-24 and the oldest cohort consisting of individuals aged 75-79.

Using Pyatt's method, Figure 3 below shows the share of overall income inequality that is due to within-age-cohort income differences in selected countries. It is important to note that what I am comparing here is the share of overall inequality that can be explained by within-age-cohort differences in income. Looking at this figure, we can immediately see that the within-cohort share of inequality in the United States has stayed in the range of 71-74

6. Some of these issues are: (1) the P-Gini coefficient is sensitive to the boundaries of the age cohorts,⁷ (2) the Paglin's method ignores overlapping income differences. It is apparent that even the youngest and the oldest groups overlap. Nelson (1977) shows that if age-income distributions overlap, the P-Gini will be affected by cohorts' mean income and population weights.

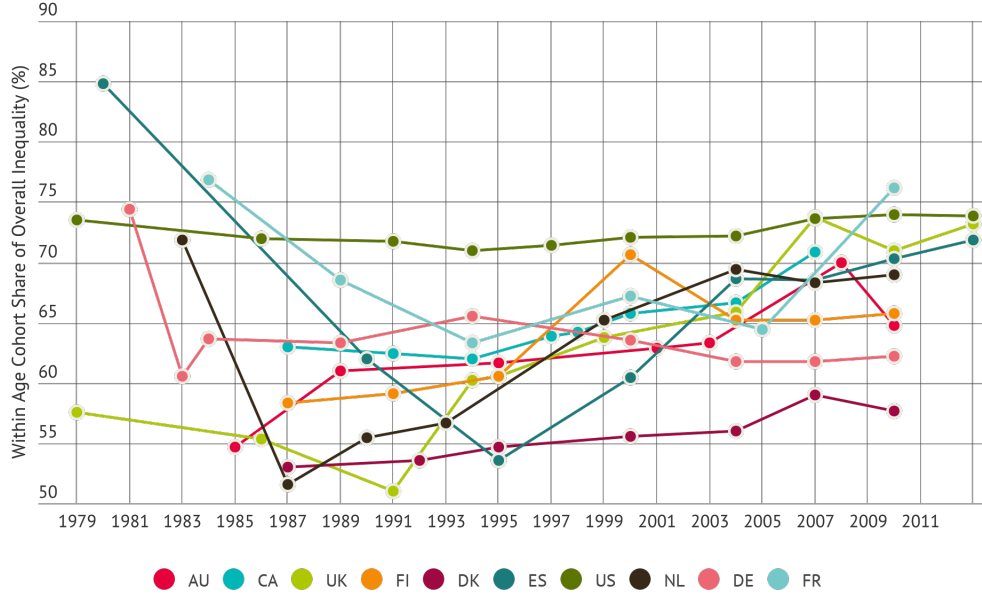


Figure 3: Within Age-Cohort Share of Inequality

percent of overall income inequality. This rate has been consistently higher than it has been in most other countries, especially the Scandinavian countries. In Denmark, for example, the within-age-cohort share of inequality has been between 50-60 percent. It appears that only Canada and France lead the United States in their share of within cohort inequality in 2010.⁸ The share of within-age-cohort inequality in the United Kingdom has risen from 57.5 percent in the late 1970s to almost as high as the United States (73.2 percent) in 2013.

The cross-country differences in Figure 3 leave us with two questions. First, what explains cross-country differences in the within-cohort share of overall income inequality? And second, why has this share remained at such a high and constant level in the United States, while it has increased (and in some cases fluctuated quite dramatically) in other countries? The full analysis of these two questions is left for future research, as their analysis requires an additional qualitative investigation into policy and other socioeconomic trends in these various countries. For now, we continue our investigation by looking at the age-income profile and demographic trends as potential explanations.

So far, I have observed that the within-age-cohort share of inequality has been both high and steady in the United States. To corroborate this observation, I first examine changes in the age income profile for some of the countries in our sample. Changes in the income profile will provide a sense of how the between-age-cohort component of inequality has changed over time, and will therefore, provide some validation for the trends we observe in Figure 3. An analysis of basic demographic shifts, on the other hand, will help us account for some

8. The within cohort share of inequality in Canada in 2010 reaches a surprisingly high value of 91.3 percent from 70.8 percent in 2007.

of the observed increases in both overall inequality and within-age-cohort inequality in the United States over time.⁹

3.0.1 AGE-INCOME PROFILE

Age-income profiles demonstrate the evolution of individual earnings over the course of a lifetime. The income profiles of most societies are typically inverse-U shaped curves with a positive relationship between earnings and age up until the point where incomes peak (usually around the age of 40-50, but varying greatly across societies) and a negative relationship between the two variables afterward.¹⁰

In the context of inequality, the income profile can be used to examine the difference in income between individuals of different age groups. In other words, the income profile can be understood as a reflection of the between-age-cohort component of inequality. In general, a steeper income profile suggests a high between-age-cohort share of inequality, and conversely, a flatter income profile indicates that the within-age-cohort share of inequality is higher. Additionally, if we see a dramatic change in the shape of the income profile over time, we would expect to see a corresponding change in the within-share of inequality.

Figure 4 below shows the income profile in 1979 and 2013 for all workers aged 20-79 for the United States, United Kingdom, Canada, and Denmark. In each case, the average income of each cohort is compared to the average income in the highest earning age group of that year in that country; the profiles are smoothed. What we see in Figure 4 is a very slight steepening of the first half of the income profile in the United States between 1979 and 2013 and a slight flattening in the second half of the life cycle. Overall, however, the shape of the income profile in the U.S. has not changed much. This is consistent with the observation that within-age-cohort inequality in the United States has not changed much over this period. Meanwhile, the income profiles in the United Kingdom and Canada have changed much more noticeably. In both countries, the income profile has flattened, which implies that the between-cohort component of inequality in both country has declined. We can confirm this trend by looking at Figure 3. For both countries, the within-cohort share of inequality has risen dramatically.

Figure 4 also demonstrates another finding that will become relevant to our discussion later on in this paper. In each of the four countries, the income profile has flattened for older workers and steepened for younger ones. For instance, the income ratio of the oldest group to the highest earning group has increased by roughly 100 percent from 0.3 to 0.6 in

-
9. The share of the overlap term tends to be relatively constant over time for most countries and very similar across countries, so the overlap term is not considered. For the sake of space, I do not report the overlap terms independently.
10. The explanations for the positive sloping part of the income profile are as follow: first, income goes up as age increases due to the augmentation of human capital and experience, which can be due to more on-the-job training. Promotions, professional networking effects, and psychological development are also responsible for the increase in income over a lifetime. Other factors such as paying off student loans, mortgages, and other forms of liabilities contribute to the increase in the incomes of workers before the tipping point. After the peak of the profile, the negative slope of the income profile can be explained by depreciation of human capital, cognitive and non-cognitive skills, and physical abilities, as well as, decline in hours of work before retirement and natural reductions in income after retirement. Mincer (1974) attributed most of the decline in earnings after the peak of income to the fall of working hours rather than a fall in hourly wages.

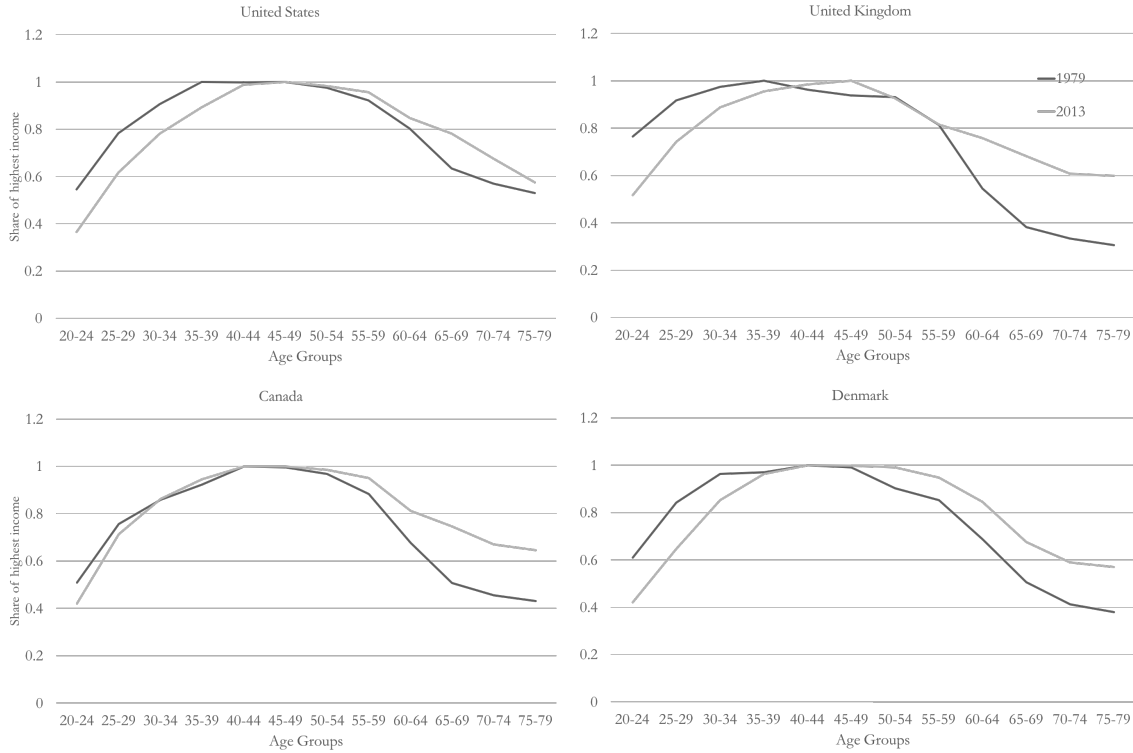


Figure 4: Age-Income Profile in the United States, United Kingdom, Canada, and Denmark

the United Kingdom. Changes in family structures, too, may impact the income profile. It is also noteworthy that in recent years, the ratio of the oldest income earners to the highest earners has been the same in all the four countries at around 0.6.

3.0.2 DEMOGRAPHIC DYNAMICS

Demographic differences among countries and changing demographics over time are responsible for part of the dynamics of income distribution in countries. The fact that median ages across countries ranges from 15 (in countries such as Niger, Uganda, and Mali) to 45 (as is the case in Monaco, Japan, and Germany) reflects these demographic disparities. As the share of younger individuals in low-income and lower-middle-income countries has increased over time, the opposite trend has been observed in more developed countries. In high-income countries, the share of individuals in the age group 45-49 and 50-54 combined rose from 11.1 to 14.6 percent between 1950 and 2015. In upper-middle-income countries, the share of the groups 40-44 and 45-49 increased by 2.3 and 2.8 percentage points, respectively. Since individuals in the middle-aged groups are supposedly among the medium to high earners in most countries, an increase in their share of the population can have significant effects on the income distributions in those countries.

In the United States, the median age has increased by 32 percent from 28 years to almost 37 years in the past four decades. The share of 45-64 year-old individuals has increased

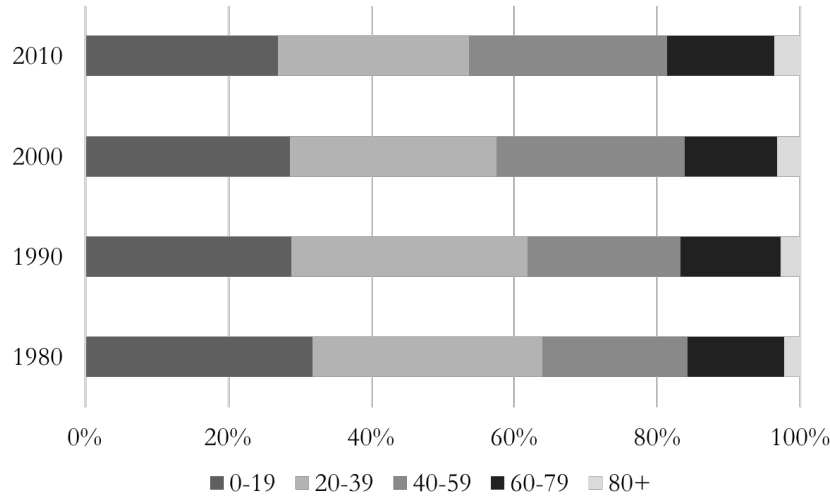


Figure 5: Age Distribution in the United States, *Source: Age and Sex Composition, U.S. Census Bureau, 2010*

from 20.6 to 26.4 percent, and the share of younger cohorts has decreased significantly. This is shown in Figure 5.

The effects of this demographic transition, namely an aging of the U.S. population, play out in two ways which I will discuss briefly here. Primarily, it impacts the inequality within each cohort and does so through the Permanent Income Hypothesis (PIH) as it has been applied to income inequality. The Permanent Income Hypothesis (PIH), which was originally applied to questions of inequality by Eden (1980) and further investigated by Deaton and Paxson (1993), suggests that inequality among individuals of the same age (or same age range) should increase as the cohort ages.¹¹ This observed trend implies that an aging society may experience rising inequality simply due to the fact that older cohorts tend to be less equal.

If this is the case, then moving to an older society is associated with a rise in income inequality within cohorts. Secondly, an aging population increases the population weights of older cohorts. As a result, the older groups that tend to be more unequal based on PIH contribute more to the overall within-cohort inequality. The combined effect of these two trends is to increase overall inequality by way of increasing the within-age-component of inequality.

My results based on LIS confirm the PIH hypothesis in the United States as shown in Figure 6. The graph shows each cohort tracked through time. The cohort that was 20-24 years old in 1974 is 40-44 in 1994 and so on.¹² As shown, inequality within cohorts of different ages in 1974 increases as cohorts age; however, the only difference between my results and those confirmed by other researchers is the decrease in inequality within the

11. The theory has been tested for various countries. For instance, see Blundell (2014) and Heathcote et al. (2005).

12. The years are chosen in 10-year intervals but are not exact since it jumps from 1974 to 1986 to 1994 to 2004 and finally 2013.

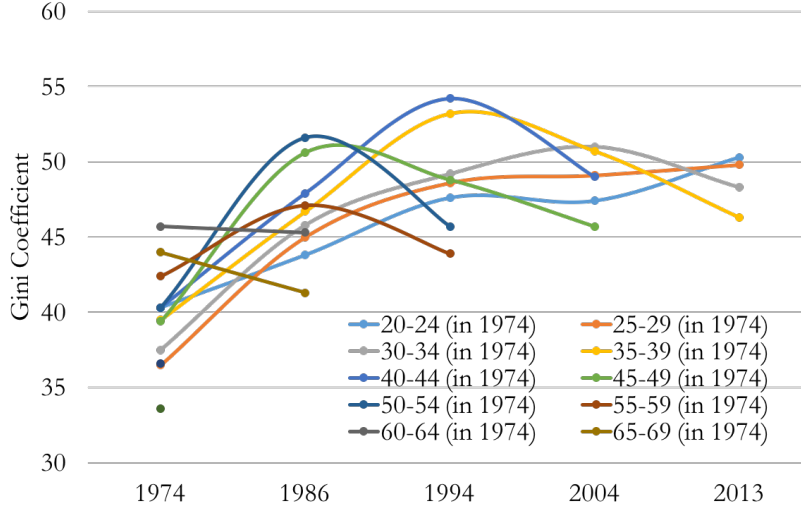


Figure 6: Inequality over lifetime for different American cohorts

same cohorts as they reach their 60s.¹³ In countries where the distribution of the pensions of the retired follows the same distribution as the pre-retirement income, inequality within the group of elderly does not change that much or might even increase, whereas in countries where incomes get boosted after retirement, inequality among the elderly might decrease. For instance, in Canada, the bottom decile is better off in retirement years than in their working years.

Multiple factors can explain why within-cohort inequality increases as the cohort ages. Among them, I believe accumulation of income, the role of credit constraint in access to higher education, assortative mating, better access to credit for high-income individuals, and better mental and physical health are the most important factors. Increase in the average age in a society can also push income inequality upward if there are strong intergenerational transfers where bequests are important (Deaton and Paxson, 1994). Higgins and Williamson (1999) argue that slower population growth that shifts the population age distribution toward older, more experienced cohorts, may potentially reduce the experience premium, which then lowers aggregate inequality.

Another impact of an aging population on income inequality is the burden that a higher share of elderly puts on younger workers. The social security taxes, levied on younger workers to support older retired members of society, depress the disposable income of the young. Many researchers have pointed out the shrinking ratio of workers to retirees. This ratio, also called the "support ratio," measures the number of individuals aged 20-64 divided by the number of individuals aged 65 and over. A higher ratio means that more workers share the burden of supporting the elderly. This ratio tends to be fairly low among countries in the developed world. However, the United States enjoys a higher ratio compared to other developed countries.¹⁴ Reznik et al. (2005) estimate that in 2005 with the scheduled tax

13. Prus (2000) finds similar results for Canada.

14. According to OECD, the support ratio in the U.S. was the fourth largest among the OECD countries and only Turkey, Israel, and Luxembourg had a higher support ratio than the U.S.

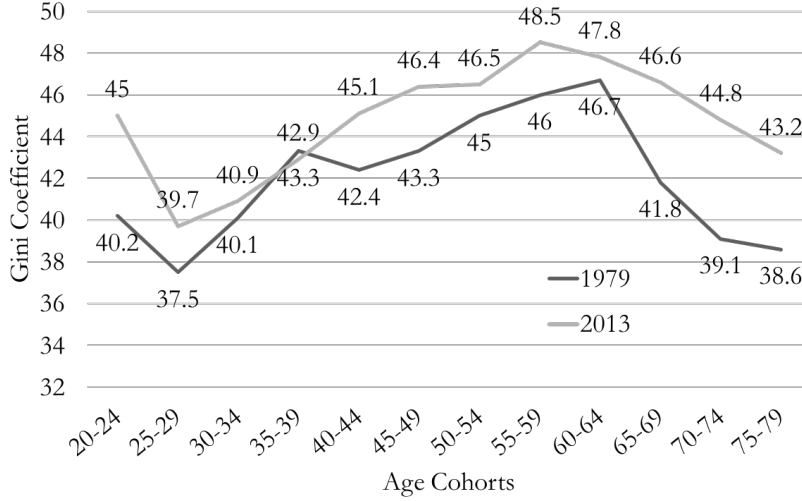


Figure 7: Inequality Measured in terms of Gini for each age cohort between 1979-2013

rates and benefits, the social security program needed a support ratio of about 2.8 to function at a pay-as-you-go level.¹⁵ It is projected that by 2040 this ratio will fall to only 2.1, putting even more burden on the young population and creating a larger age-gap, and therefore, causing overall income inequality to increase.

4. Cross-group empirical analysis

The analysis in previous section reflects the importance of inequality decomposition into within- and between-cohort components. A further step is to look at inequality within each age cohort. Figure 7 reveals that the inequality within each of these age cohorts has changed differently. Interestingly, the growth in within-cohort inequality ranges from -1 percent for 35 to 39-year olds to 14.5 percent for 70 to 74-year olds. Most of the increase in inequality in the period 1979-2013 is due to increases in within-cohort inequality among the elderly and middle-age workers. Increases in inequality among the young is not as drastic as among other age groups. We do see an increase within the youngest cohort (20 to 24-year olds), but this can be explained by differences in work and educational choices. Note that the overall within-cohort component of inequality calculated in the previous section is a weighted average of inequality within each age group.

What factors can explain the variations in terms of change in inequality across the age groups mentioned above? What has precipitated high inequality among older cohorts is different than the factors responsible for the inequality among younger cohorts. A useful framework for isolating different factors that have shaped inequality in the United States is to partially examine them through a simple wage equation. To do this, in the next section I find the effect on within-cohort inequality of factors across age, gender, and racial groups and the section that follows I look at additional factors such as marriage rate, variation in number of children, etc. as determinants of cross-group inequality differences.

15. So that the tax revenue roughly equals benefit payments.

4.1 Basic Model

I use each cohort as a unit of analysis. I follow Juhn et al. (1993) and Heathcote et al. (2005) and assume that those effects of cohort characteristics and time effects (year effects) are additively separable. To do this, I first group individuals in each year into age, gender, and racial groups. Since racial categories in 1979 only include whites, blacks, and Hispanics, I exclude other racial categories that are added in earlier years. I also use 12 age categories defined as 5-year intervals.¹⁶ As a result, there are 72 cohorts in each year. The data includes years 1979, 1986, 1991, 1994, 1997, 2000, 2004, 2007, 2010, and 2013. Table 2 summarizes the within groups inequality (WGI) in terms of Gini coefficient for different years. It is interesting to find that although average inequality across cohorts has increased over the period 1979-2013, cross-cohort variations has decreased during the same period. In the following section I review gender and race as the most important factors explaining cross-group differences in inequality.

Table 2: Inequality in different groups: 1979-2013

Year	Mean	S.D.	Minimum	Maximum
1979	37.9	6.7	27.3	51.7
1986	39.1	5.7	27.8	51.0
1991	40.0	4.4	30.7	50.8
1994	41.0	5.2	26.8	52.5
1997	40.9	4.8	31.5	50.8
2000	40.8	3.9	32.1	49.4
2004	41.6	3.9	33.5	54.0
2007	41.2	3.6	34.7	52.5
2010	41.3	3.3	33.9	48.0
2013	42.1	3.1	35.3	48.7
All Years	40.6	4.7	26.8	54.0

4.1.1 GENDER

The question we are facing is whether within cohort inequality is different for men and women and whether it has changed over time. To investigate this, I further decompose the Gini coefficient for male and female individuals. Inequality among cohorts of women that was much higher than men in 1979 seems to have declined to a level close to that of men in 2013 as shown in Figure 8. In fact, the within-cohort inequality within cohorts of women has become surprisingly similar to the shape of within-cohort inequality among men in 2013. The decline in inequality among women and the rise in inequality among men can explain this trend. If we look closer we see that the inequality within some cohorts of women has indeed fallen by as much as 14 percent. For men, the increase in inequality among some cohorts (for instance 30- to 34-year olds and 40- to 44-year olds) reaches almost

16. Thus cohorts will be 20-24, 25-29, 30-34, and so on. The results in this section and the following sections are robust to the length of the intervals.

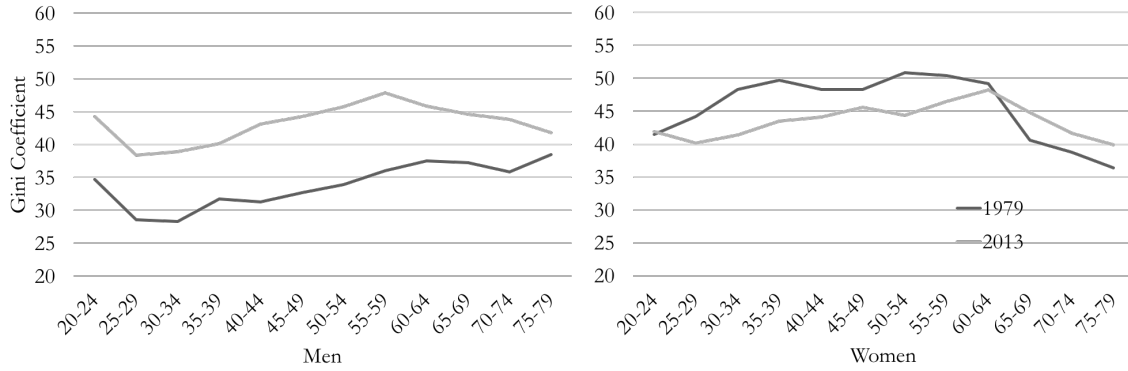


Figure 8: Change in inequality within cohorts of men and women between 1979-2013

37 percent. This has been noticed by few researchers¹⁷. However, the decline in inequality among cohorts of women has not been investigated as much as it deserves.

Note that we no longer see a large increase in inequality among the older men and women as we saw in Figure 7. The answer to this contradiction is the rise in inequality between men and women of older ages during the same period. The closing of the gender gap is not uniform across all age groups. Previous studies suggest that the gender gap increases by age, i.e. it is higher among older workers.¹⁸ As 9 shows are some interesting observations. First, the gender gap is larger for older cohorts in 2013 but not in 1979. The between gender inequality for 65- to 79-year olds is almost three times the between gender inequality between 20- to 24-year olds. Second, the gender gap has gone down for all cohorts between 1979 and 2013 with the exception of 70- to 74-year olds and 75- to 79-year olds.

4.1.2 RACE

Race plays a big role in determining income. A large body of academic work is devoted to racial differences in income. If income differences are equal across racial groups, how are inequality levels different within each group? Figure 10 shows WGI for different racial groups in 1979 and 2013. One can see the variations between these racial groups. For instance, in 1979 there is almost 4.0 Gini points between WGI among blacks and whites. Between 1979-2013, while inequality among blacks has gone up by about 86 percent, it has increased by 3.3 percent among Hispanics.

Race can also explain variations in inequality among age groups through life expectancy. Income inequality and life expectancy are closely tied to one another. On the one hand, life expectancy is lower in more unequal societies and on the other hand, the dynamic of life expectancy can affect the distribution of income in a society. The latter explanation is mainly important in our context. In the period 1979-2013, life expectancy at birth increased by almost 8 years for the population as a whole. The racial differences narrowed but never vanished. In 1970, the gap in life expectancy between a black male and white female was

17. For instance see Osberg (2003) and Jenkins (1995)

18. For instance, see Goldin (2014).



Figure 9: Inequality between men and women of different age groups in 1979 and 2013.
Source: LIS Data

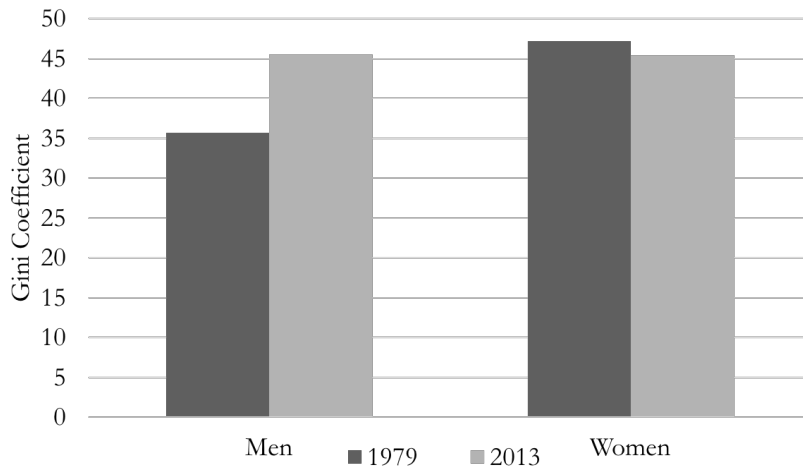


Figure 10: Inequality within racial groups between 1979-2013

as high as 15 years but narrowed to 9 years in 2013. These trends are shown in the Figure 11.

As depicted, while the average life expectancy for black females exceeded the age of 70 in 1974, for black males this only happened after 2007. This increase in average life expectancy for black Americans, who tend to be found in the lower end of the income distribution may explain the rise in inequality among the older cohorts.

To begin with our regression analysis, I first look at how the main factors explained above contribute to the cross-cohort inequalities both individually for each year and pooled for all years with dummy year effects.



Figure 11: Life expectancy at birth in the United States between 1970-2013. Source: Center for Disease Control

$$WGI_{it} = \beta_0 + \beta_1 \text{Age}_{it} + \beta_2 \text{Female}_i + \beta_3 \text{Race}_i + v_t + \epsilon_{it}$$

where i represents the i -th cohort, t represents year t , and v_t captures year effects. *Age*, *Female*, and *Race* are vectors of dummies for cohort i . The dependent variable is the inequality in terms of Gini coefficient in cohort i at time t . Table 3 shows the regression coefficients. Last column is aggregated data for all years but controls for year effects. The variable age is a dummy variable representing each age group.

As it appears most variables are significant predictors of within-group Gini coefficient in the aggregated model. The coefficient reflecting the additional inequality within women compared to men has declined over time. The coefficient is as high as 8.0 Gini points in 1979 but declines to about zero in 2013, pointing to the fact that women and men have become equally unequal in recent decades. The most equal age groups are the age groups 25-29, 30-34, and 75-79 year olds. On the other hand, the most unequal age groups are middle-aged Americans of age range 55-64. This is consistent across during the period. In terms of racial groups whites are the most unequal. In the aggregated sample of all years, whites tend to be 2.0 Gini points more unequal than blacks, while Hispanics, tend to stand in between. Over time, however, Hispanics have become as equal as blacks. The constant term has increased from 33.6 to 44.1 (by about 10.0 Gini points) pointing out the overall trend in inequality in the same period. The same analysis holds for within-group Theil index and P90/P10 income ratio as shown in Table 4 columns 1 and 2.

It is important to note that the coefficients represented in Table 3, are based on annual earnings. The variance of income is largest at either end of the working life (Blundell, 2014) due to work, school or work, and retirement decisions. The income differences among the youth have become more pronounced due to increases in college enrollment. On the other hand, health shocks and the dynamic of retirement decisions likely explain income inequality at the end of the working life. Therefore, it is very important to look at how much of the inequality among different cohorts is due to enrollment or retirement decisions.

Table 3: Cross Cohort Inequality Dynamic, Gini Coefficient

	1979	1986	1991	1994	1997	2000	2004	2007	2010	2013	All Years
Gender											
Female	8.54***	7.52***	5.01***	4.25***	3.08***	1.56*	0.72	2.00**	0.27	0.03	3.30***
Age											
25-29	-0.0493	-0.0328	-4.50*	-4.13*	-5.75**	-3.55	-6.43**	-2.88	-5.85***	-5.38***	-4.67***
30-34	-3.70	-2.60	-2.03	-3.43	-2.72	-2.57	-5.18*	-2.13	-5.57***	-4.77***	-3.47***
35-39	-1.17	-2.25	-0.87	-0.37	-0.02	-1.12	-3.58	-1.77	-3.13*	-4.23**	-1.85**
40-44	-2.35	-0.72	-1.65	-0.18	-0.93	-0.78	-1.58	-0.72	-2.78*	-2.48	-1.42*
45-49	-0.03	-1.32	-0.03	0.72	-0.18	2.25	-2.23	0.233	-1.42	-1.85	-0.39
50-54	1.08	1.02	1.32	2.63	0.00	1.55	-0.60	0.97	-0.13	-1.52	0.63
55-59	2.25	2.57	2.95	3.43	3.80	3.30	0.88	1.83	-1.10	1.33	2.12**
60-64	1.87	2.57	2.43	3.68*	3.18	2.65	1.27	2.85	0.77	0.58	2.18***
65-69	-3.03	-2.43	-0.30	-1.52	-0.28	0.53	-0.10	4.93**	1.40	-0.23	-0.10
70-74	-4.88	-5.90**	-1.92	-5.85**	-2.83	-0.80	-2.03	-0.38	-1.85	-2.32	-2.88***
75-79	-5.85*	-8.65***	-5.15**	-9.17***	-1.90	-2.62	-1.47	-1.23	-6.18***	-5.75***	-4.80***
Race											
White	3.53*	2.10*	1.06	1.85*	3.83***	3.37***	1.58	2.05*	2.04**	1.09	2.25***
Hispanic	1.65	1.26	1.29	0.00	3.34**	1.38	0.84	0.20	1.02	-0.69	1.03**
Constant	33.6***	36.0***	37.5***	39.4***	37.6***	38.6***	42.2***	39.3***	42.3***	44.1***	36.4***
Year Dummies	<i>N</i>	<i>N</i>	<i>N</i>	<i>N</i>	<i>N</i>	<i>N</i>	<i>N</i>	<i>N</i>	<i>N</i>	<i>N</i>	<i>Y</i>
Obs.	72	72	72	72	72	72	72	72	72	72	720
Adj. R^2	0.524	0.714	0.544	0.638	0.386	0.335	0.208	0.383	0.574	0.508	0.441

Standardized beta coefficients

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In terms of retirement age, too, it is important that older men have been retiring at ever-earlier ages starting the late nineteenth century. This could be due to the introduction of the social security programs, the rise of incomes and consequently savings, and lifestyle changes in older ages during the same period. This decline in labor force participation was steeper during the early 1970s and 1980s and it fell from 83.4 percent in 1969 to 67.2 percent in 1989 (Juhn and Potter, 2006). However, labor force participation among older men stabilized by later 1980s and early 1990s and has even slightly increased since then.

Therefore, it is useful to see the within-cohort dynamic of inequality among employed and full-time workers. I first replicate the analysis for employed workers controlling for the dynamic of labor market. Since this highly limits the number of people in the older groups, I limit the sample to individuals of age 69 and younger. Column 3 in Table 4 uses the within-group Gini coefficient for employed individuals only. Once I limit the sample to employed workers, the within-cohort inequality gap between men and women vanishes. Age does not seem to be a significant determinant of across-cohort variations of inequality and middle-aged workers are no longer as unequal as when we do not have an employment restriction. The constant term has decreased from 0.364 in the entire sample to 0.314 in the sample with only employed individuals, which reflects that inequality is partially driven by inequality between the employed and the unemployed.

In column 4, I limit the sample to only full-time workers (those working above 35 hours per week). The regular weekly hours are calculated as the sum of hours worked at first and second jobs including family work and overtime. It is interesting to note that in the full-time sample, it is men that are more unequal than women and inequality in hours worked is partially responsible for the high inequality among women. Workers of age 45-64 years old seem to constitute the most unequal groups.

In what follows, I review other factors that I believe can explain cross-cohort variations in inequality in the United States.

4.2 Other factors

4.2.1 FEMALE LABOR SUPPLY, MARRIAGE, AND BIRTH RATE

Inequality in terms of income among individuals regardless of their labor force or employment status is largely affected by their labor supply decisions. This is particularly important among women who traditionally have a more elastic labor supply rate. Factors such as cultural changes, and changes in divorce rate and household technology have been characterized as some of the possible explanations for the increase in female labor force participation rate. According to Johnson and Skinner (1986) divorced women are more likely to participate in the labor market and an increase in the risk of divorce may also increase married women's participation since they may choose to stay in the labor force to hedge against future income risk associated with divorce. It is important to note that the expansion in female labor supply has been mainly driven by changes that have occurred in job market decisions of married women, more specifically white married women who traditionally used to be mainly working at home. Researchers also associate the increase in labor force participation among this group to policy changes that mostly affected women of color in the United States. Among them are the transformation of the old Aid for Families with Dependent Children (AFDC) program into more temporary and conditional assistance

Table 4: Cross Cohort Inequality Dynamic, Other measures and Gini coefficient of employed and full-time workers

	(1) Theil	(2) P90/P10	(3) Gini (Employed)	(4) Gini (Full-time)
Gender				
Female	0.0386***	8.715***	0.45	-1.98***
Age				
25-29	-0.0527***	-0.208	-4.55***	-2.50***
30-34	-0.0344*	1.963	-3.35***	-1.10
35-39	-0.0058	2.444	-1.38*	0.72
40-44	-0.0026	0.290	-0.83	1.30
45-49	0.0200	1.138	-0.11	2.30***
50-54	0.0282*	1.725	0.24	2.39***
55-59	0.0500***	2.531	0.74	2.52***
60-64	0.0391**	-1.599	0.72	2.13**
65-69	0.0194	-5.739*	0.93	1.26
70-74	-0.0225	-6.514**		
75-79	-0.0464***	-7.023**		
Race				
White	0.0453***	6.600***	3.05***	2.79***
Hispanic	0.0230***	3.098**	1.97***	1.73***
Constant	0.218***	16.93***	31.4***	28.0***
Year Dummies	Y	Y	Y	Y
Obs.	720	720	600	600
Adj. R^2	0.285	0.253	0.404	0.371

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

in the Temporary Assistance to Needy Families (TANF) program that happened in 1996 as well as the expansion of the Earned Income Tax Credit program (EITC). On the other hand, women from high-income households were mainly motivated to join the labor market by the ever-increasing skill premium and the Tax Reform Act of 1986, which reduced the top marginal tax rate from 50 to 28 percent.¹⁹

There has been a battery of empirical studies on the effect of marriage on inequality. Numerous studies are conducted on the effect of assortative mating on inequality, especially those calculated based on household incomes. Greenwood et al. (2016) show that assortative mating, divorce, and female labor supply accounted for about one-third of the increase in income inequality in the United States from 1960 to 2005. However, there is not a great deal of analyses on the effect of marriage on inequality among workers. Most of the impact of marriage rate on inequality, especially among women, is through labor market participation. Juhn and Potter (2006) show that although general increases in wages have caused all women to supply more hours than before, women who marry to high-wage husbands, on average, increase their labor supply even faster than those married to low-wage husbands. They report that women married to men in the bottom quintile of earnings increased their labor

19. See Juhn and Potter (2006).

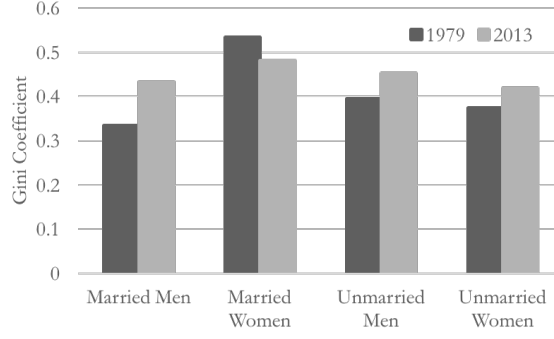


Figure 12: Inequality in terms of Gini coefficient for different marital and gender groups

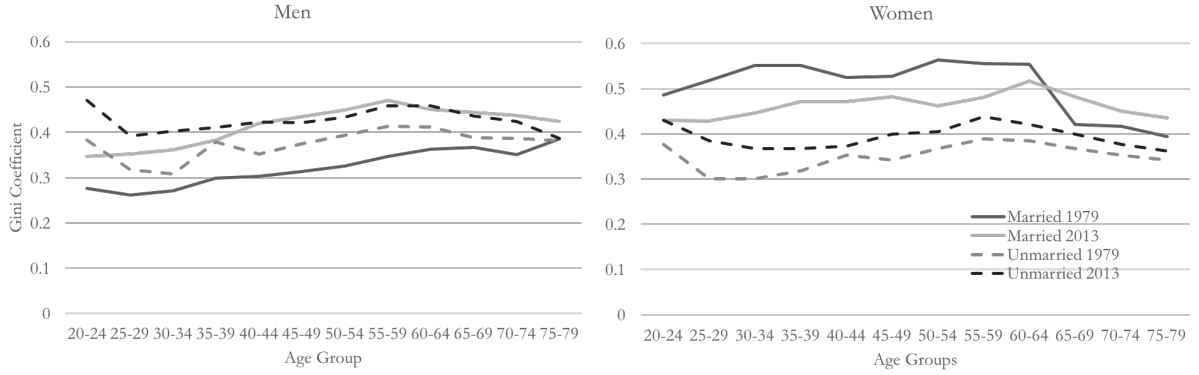


Figure 13: Within age-cohort inequality in terms of Gini coefficient for different marital and gender groups

force participation rate from 46 percent in 1969 to 66 percent in 1999. On the other hand, those married to men in the top quintile of earnings increased their labor force participation rate from 31 to 66 percent during the same period. Gottschalk and Smeeding (1997) argue that married women's labor force participation rates, hours, and wages have increased in almost all countries during the 1980s. They also find that the correlation between husbands and wives earnings has increased during the same time.

If we analyze the trend in inequality among married and unmarried workers by gender, we observe peculiar trends. Although, inequality within married and unmarried men as well as unmarried women have all gone up, inequality within married women has declined between 1979 and 2013. Figure 12 depicts the Gini coefficient for each group. While inequality within married men soared by almost 30 percent between the two years, it decreased by about 10 percent for married women. Figure 13 represent within age-cohort inequality for the same groups. A noticeable trend is again the decline in the inequality within almost all cohorts of married women. We could safely assume that it is married women that have driven down the overall inequality among women. Among men, most of the inequality is driven by married men.

This reduction in inequality among married women could be correlated with the decline in the number of children per woman during recent decades. If women without children have generally higher earnings (due to longer hours of work and higher labor force participation) than those with children, a general declining trend in childbirth per woman can reduce inequality. Goldin (2014) reports that women with children work around 24 percent fewer hours per week than women without children in 2012. Finding the link between wage and number of children is harder to establish due to selection, however, most causal estimates point to a negative relationship between the number of children and wages in most countries (Sigle-Rushton and Waldfogel, 2007). Numerous factors can be suggested for the existence of this negative relationship. If female wages are correlated with their market productivity, then reduction in work experience, loss of human capital, atrophy of market skills while not working, and reduced incentive to invest in training that may bring a payoff in the future during the childbearing and parenting period can ebb wages for mothers (Lundberg and Rose, 2000).

My calculations show that inequality in terms of number of children among married workers has decreased by about 8 percent between 1979-2013. In particular, the inequality of number of children per woman declined by 15 percent for 40- to 44-year olds, by 17 percent for 45- to 49-year olds, and by 14 percent for 50- to 54-year olds. Altogether, this suggests that childbirth rate partially explains the reduction in inequality among married women. The question that arises is why the reduction in the number of children per family contributed to a decline in wage inequality among women but not men. I hypothesize that the family gap and its uneven effect on women and men is a suspect. It has been argued that childbirth leads to salient reallocations of time and effort for married couples. It tends to impose a wage penalty on maternal wage while it may cause an increase in the paternal wage. Lundberg and Rose (2000) find that the birth of the first child is associated to a 5 percent decrease in the mother's wage while it is linked to a 9 percent increase in father's wage. This may be due to the fact that fathers may need to work longer hours to compensate for the mother's time spent at home and not at work. The decline in childbirth, especially among lower-income families, may have caused a reduction in fathers' hours of work and, therefore, their wages which may in turn increase inequality. On the other hand, the same factor may have caused women of lower income to spend more hours at work that consequently leads to closing the wage gap within women.

4.2.2 TECHNOLOGY AND AUTOMATION

It has been almost eight decades since Johan Maynard Keynes coined the term "technological unemployment." Since then the idea that machines will eventually displace workers and create "superstars" or "winners" has been popular both inside and outside academia. Brynjolfsson and McAfee (2012), in their book, *Race Against the Machine*, discuss how technological changes can lead to a rise in inequality among workers. The best place to start for understanding within-gender inequality trends in the United States is to look at how technology affects employment in different sectors.

David Autor argues that the effect of automation on employment is not uniform across all occupations. He divides jobs into three categories: a) *routine* jobs or jobs that follow an exhaustive set of rules such as bookkeeping, clerical work and repetitive production tasks, b)

manual jobs or jobs that require situational adaptability, visual and language recognition, and in-person interactions such as food preparation, serving jobs, cleaning, and maintenance, and c) *abstract* jobs or jobs that require problem-solving skills, intuition, creativity, and persuasion such as managerial, technical, and professional occupations (Autor, 2014). Autor then argues that manual and abstract skills that demand more flexibility, judgement, and common sense skills are the ones that are less likely to be replaced by machines. While computers are good substitutes for routine jobs, they mostly complement abstract jobs and may have ambiguous effect on manual jobs. As a result, if automation replaces routine jobs and increases productivity of workers in manual and abstract jobs, the result is a job polarization in which there is a growth in employment in high-education, high-income and low-education, low-income jobs and a decline in employment in middle-education, middle-income jobs. These findings are empirically supported by Goos and Manning (2007). This job polarization can have direct impact on the distribution of income in the economy. But we can only understand the consequences of this job polarization for wage inequality through the elasticity of demand and supply of jobs.

Let us start with abstract jobs. The fact that the accumulation of skills and human capital is slow makes the supply of workers in abstract jobs very inelastic, as a result an increase in demand for abstract jobs (potentially due to rises in productivity) is not usually followed by an influx of workers to supply those jobs. Evidence suggest that computerization has benefited all workers in abstract jobs by raising their wages (Autor, 2014). However, the story is different for women. The large number of idle women who have been recruited to the labor force makes supply of women (both skilled and unskilled) more elastic. Blundell and MaCurdy (1999) finds that the own-wage labor supply elasticity of women is almost ten-fold that of men. The inelastic supply of male abstract workers and more elastic supply of female abstract workers lead to an increase in wages in managerial jobs done by men and a decrease in wages in those done mainly by women.²⁰ Due to the already high wages in these sectors, the compound effect is higher inequality among men and lower inequality among women.

On the other hand, since computers do not necessarily complement (or substitute) manual jobs, the productivity gains in those jobs are negligible. I argue that the demand for manual tasks are relatively income elastic. Therefore, a growth-induced increase in the aggregate income can lead to a rise in demand for manual occupations. Now, due to the high elasticity of supply in those jobs a wage rise in manual jobs is naturally compensated with more supply of workers. Consequently, unlike abstract jobs, the effect of computerization on manual skills is not necessarily positive for both men and women.

Lastly, high substitution rate between routine jobs and computerization means lower wages for those working in the related sectors. Figure 14 shows the employment shares of occupational categories for both men and women. I have added two more categories: (1) cognitive routine jobs, which are jobs that require more cognitive skills, include sales and office occupations, and (2) manual routine jobs, which are routine jobs that require physical skills more than cognitive skills, include construction, transportation, production and repair

20. In fact, it seems that the rise in the demand for skilled labor is positively correlated with automation and and computerization. Weinberg (2000) finds evidence for a positive correlation between computer investment at the industry level and demand for female labor.

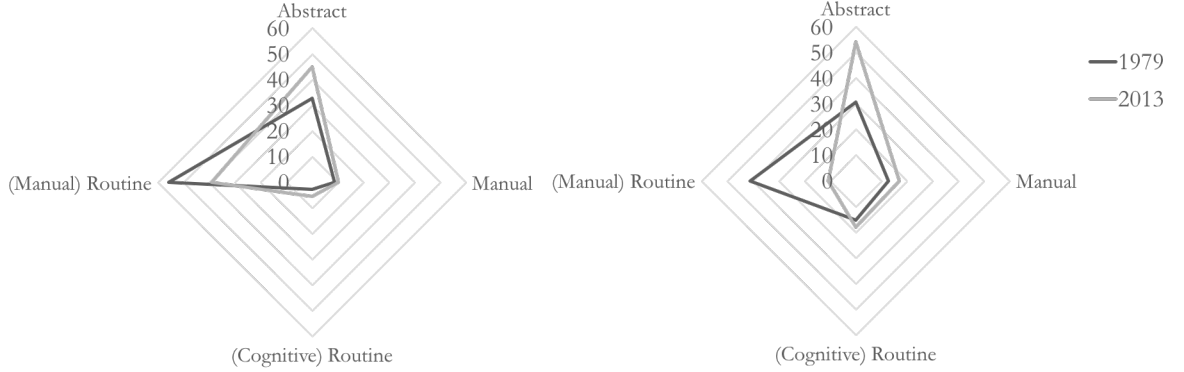


Figure 14: Changes in occupational distribution for major occupational groups by gender between 1979-2014 Source: data from LIS Note: Abstract jobs contain managerial, professional, and technical occupations. Cognitive routine jobs are sales and office-related occupations. Manual routine jobs are production, transportation, and material moving occupations as well as natural resources, construction, and maintenance occupations. Manual jobs are cleaning, personal, and protective services.

occupations. As figure shows, women has reduced their employment in manual routine jobs at a higher rate than men (30 versus 15 percentage points between 1979 and 2014).

In conjunction with the dynamic of employment shares, the ever decreasing decline in wages for routine jobs has led to an increase in inequality among men and a decrease in inequality among women between 1979 and 2013. Note that the share of women in abstract jobs has increased over the period. Also it is important to notice that for men the changes in employment shares are quite negligible. Figure 15 depicts the percentage change in mean annual wages by occupational categories for men and women. The only negative growth in wages for women is in manual routine jobs, whereas, the only positive growth in wages for men is seen in abstract jobs. The increase in annual wages in managerial and professional jobs for women is as high as 25 percent. Consistent with elastic supply of workers for manual jobs, as employment rose for this occupation category, wages declined for men. Overall, the wage growth in routine category has been very modest for women and highly negative for men.

Equation below shows the relationship between WGI and the factors discussed above controlling for gender, race, and age. I add another factor measuring the share of cohort with college or graduate degrees. Occupations are grouped into four exhaustive and mutually exclusive categories: (1) abstract, (2) manual, (3) cognitive routine, and (4) manual routine. The manual-routine group is the omitted category in the regression.

$$WGI_{it} = \beta_0 + \beta X_{it} + \gamma_1 \text{ShareMarr}_{it} + \gamma_2 \text{VarChild}_{it} + \gamma_3 \text{ShareHiEd}_{it} + \gamma_4 \text{ShareAbstract}_{it} + \gamma_5 \text{ShareManual}_{it} + \gamma_6 \text{ShareCogRoutine}_{it} + v_t + \epsilon_{it}$$

where X_{it} represents *Age*, *Female*, and *Race* dummy variables for cohort i at time t . ShareMarr_{it} is the share os married individuals in the cohort, VarChild_{it} is the variance of number of children in each cohort, ShareHiEd_{it} is the share of individuals with college degrees or above, and $\text{ShareAbstract}_{it}$, ShareManual_{it} , and $\text{ShareCogRoutine}_{it}$ are the share

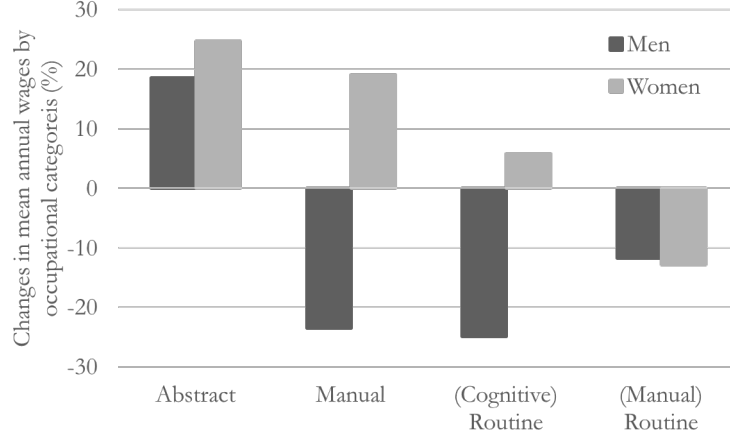


Figure 15: Changes in mean annual wages by occupational categories and gender, Note: data from LIS

of those in abstract, manual, and cognitive-routine jobs, respectively. By construction, $\gamma_4 + \gamma_5 + \gamma_6 = 1$. Table 3 shows the regression results for different measures of inequality.

Table 5 shows the regression results. After adding the variables, females remain to be more unequal. Looking at the Gini coefficient for all individuals in the sample, women are more unequal by almost 5.0 Gini points. Again, this gap in terms of within-cohort inequality between men and women vanishes when we only consider only employed and full-time workers. The inequality among the youngest cohort (20- to 24-year olds) is the highest followed by 55- to 64-year olds. Inequality among cohorts with higher share of individuals with college degree and above is slightly lower. A 10 percent increase in share of highly educated individuals is associated with a 2.0 point decrease in the Gini coefficient. Share of abstract workers does not seem to affect inequality much. Only in model 5, a 10 percent increase in the share of abstract workers increases the Gini coefficient by about 2.0 points. For every 1 point increase in the variance of number of children, inequality can decrease by 2.0 points in model 3.

5. Concluding Remarks

In this paper, I have analyzed the evolution of income inequality in American society between 1979-2013. I offer a more granular analysis by looking at inequality that exists within and between multiple demographic groups instead of looking at single statistics. I pursue this by decomposing the inequality into different sub-groups using a Gini decomposition method introduced by Pyatt (1976). I advocate decomposing inequality into smaller age and gender groups for the following reasons: (1) we can gain greater insights into the causes (and consequences) of inequality by looking at smaller groups of the population, (2) changes in inequality across space and time can be partially caused by the demographic dynamics of those countries, and (3) inequality measures do not reflect reasonable differences in income among groups of different ages that exist due to life cycle effects.

Table 5: Cross Cohort Inequality Dynamic for Different Inequality Measures

	(1) Theil Index	(2) P90/P10	(3) Gini	(4) Gini (Emp.)	(5) Gini (Full-time)
Gender					
Female	0.0802***	15.49***	5.31***	3.80***	-2.36*
Age					
25-29	-0.0777***	-0.479	-6.20***	-6.85***	-4.35***
30-34	-0.0816***	-2.077	-6.56***	-6.91***	-3.82**
35-39	-0.0621*	-3.301	-5.57***	-5.48***	-2.27
40-44	-0.0655*	-6.884	-5.60***	-5.24***	-2.01
45-49	-0.0475	-7.678	-4.89***	-4.60***	-1.04
50-54	-0.0459	-9.729*	-4.39***	-4.26**	-1.19
55-59	-0.0319	-12.25**	-3.49***	-3.76**	-1.09
60-64	-0.0499*	-19.33***	-3.99***	-3.74**	-1.70
65-69	-0.0724**	-25.10***	-6.42***	-3.47*	-2.29
70-74	-0.109***	-25.31***	-8.76***		
75-79	-0.1210***	-24.37***	-10.10***		
Race					
White	0.0173	1.305	-0.06	2.14**	0.21
Hispanic	-0.0214*	-8.235***	-2.25***	-0.14	0.85
Share Married	0.0021***	0.545***	0.16***	0.07**	0.06*
Var(No. of Children)	-0.0287	-12.22**	-2.22*	0.28	-1.11
Share High Ed	-0.0025**	-0.646***	-0.20***	-0.10	-0.16*
Share in Abstract	0.0002	-0.0768	0.03	0.00	0.17***
Share in Manual	-0.0003	0.0287	-0.02	-0.08*	-0.06
Share in Cog-Routine	-0.00100*	-0.0308	-0.01	-0.14***	0.00
Constant	0.213***	17.27***	35.3***	31.7***	27.7***
Obs.	720	720	720	600	600
Adj. R^2	0.312	0.329	0.510	0.419	0.396

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

I find that while in most countries the part of inequality that is due to differences within age groups has fluctuated over the period 1979-2013, in the United States, this share has remained relatively stable. In the U.S., the within-cohort proportion of inequality has not exceeded the range of 71-74 percent of overall inequality, but this range is among the highest compared to a sample of countries I used in this study. As a point of reference, the within-cohort share of inequality in the United Kingdom has soared from a low value of 57.5 percent in the late 1970s to a level as high as the United States in 2013. This is mainly due to three facts: (1) a relatively stable age-income profile, (2) the Permanent Income Hypothesis combined with demographic changes in the American society and the aging of

the baby boomers. I also find that about 17 percent of the increase in inequality during this period is due to between age-cohort inequality and the rest is due to the rise in inequality within age cohorts.

I further break-down income inequality into gender and racial groups in addition to age and find cross-group differences in WGI and change in WGI. I find that there are significant differences in terms of WGI across groups. For instance, while inequality among middle-aged Americans has remained the same or has changed little, it has increased by 9 and 5 percent respectively among 20- to 24-year olds and 25- to 29-year olds. Another finding is that men are the primary contributor to the increase in overall inequality. To investigate what factors determine these cross-group differences in inequality, I employ a regression model. The regression results are interesting. I find that whites are more equal compared to blacks and Hispanics. Moreover, men tend to be more equal than women but the difference fades out over the period. Middle-aged Americans of age range 55-64 are less equal than their younger and older counterparts. Other factors such as share of abstract workers, share of people with post-secondary degrees, share of married individuals, and variance of number of children can explain the cross-group WGI.

Of course, the analysis in this paper is not the complete picture. There are other factors to consider across time and groups. The decrease in the real value of minimum wage is coupled with a sharp decline in the share of female workers who earn at or below the wage floor and a relatively stable share for male workers. This can lead to a decline in inequality among women and a rise in inequality among men. One can also argue that the combination of the larger decline in unionization among men than women and arguably the larger negative impact of deunionization on wage inequality for men than women can explain differences in WGI between men and women. The dynamic of education policy during the period 1979-2013, too, can explain why some groups are more equal than others especially across age and time. Educational opportunities and inequality of income are two sides of the same coin. The gap in financial resources provided to children and youth and differences in the capacities of families to invest in their children has widened in recent decades. This trend has been coupled with a period of constant increase in the return to education and college premium. Therefore, unless offset by public policies that lessen such effects, the lower equality of opportunity and mobility rates get translated to higher degrees of inequality. Other factors such as the increase in Social Security's full retirement age, reduction in participation in retirement benefits, a transition from defined benefit plans to defined contribution plans, a fall in the share of employers that offer retiree health benefits, and an increase in average total out-of-pocket spending on services and premiums by Medicare beneficiaries could all be suspects in the rise in inequality among the American elderly in the past few decades.

All in all, I believe studying the evolution of income inequality in the United States, or any other country requires a detailed analysis of cohorts rather than looking at single numbers. The dissection of inequality can lead to more insights into what has caused inequality and what to do about it.

References

- David Autor. Polanyi’s paradox and the shape of employment growth. Technical report, National Bureau of Economic Research, 2014.
- Richard Blundell. Income dynamics and life-cycle inequality: Mechanisms and controversies. *The Economic Journal*, 124(576):289–318, 2014.
- Richard Blundell and Thomas MaCurdy. Labor supply: A review of alternative approaches. *Handbook of labor economics*, 3:1559–1695, 1999.
- Erik Brynjolfsson and Andrew McAfee. *Race against the machine: How the digital revolution is accelerating innovation, driving productivity, and irreversibly transforming employment and the economy*. Digital Frontier Press, 2012.
- Brigitte Buhmann, Lee Rainwater, Guenther Schmaus, and Timothy M Smeeding. Equivalence scales, well-being, inequality, and poverty: Sensitivity estimates across ten countries using the luxembourg income study (lis) database. *Review of income and wealth*, 34(2): 115–142, 1988.
- Angus Deaton and Christina Paxson. Intertemporal choice and inequality. Technical report, National Bureau of Economic Research, 1993.
- Angus Deaton and Christina Paxson. Intertemporal choice and inequality. *Journal of Political Economy*, 102(3):437–467, 1994.
- Benjamin Eden. Stochastic dominance in human capital. *The Journal of Political Economy*, pages 135–145, 1980.
- John P Formby, Terry G Seaks, and W James Smith. On the measurement and trend of inequality: a reconsideration. *The American Economic Review*, pages 256–264, 1989.
- Claudia Goldin. A grand gender convergence: Its last chapter. *The American Economic Review*, 104(4):1091–1119, 2014.
- Maarten Goos and Alan Manning. Lousy and lovely jobs: The rising polarization of work in Britain. *The review of economics and statistics*, 89(1):118–133, 2007.
- Peter Gottschalk and Timothy M Smeeding. Cross-national comparisons of earnings and income inequality. *Journal of economic literature*, 35(2):633–687, 1997.
- Jeremy Greenwood, Nezih Guner, Georgi Kocharkov, and Cezar Santos. Technology and the changing family: A unified model of marriage, divorce, educational attainment, and married female labor-force participation. *American Economic Journal: Macroeconomics*, 8(1):1–41, 2016.
- Jonathan Heathcote, Kjetil Storesletten, and Giovanni L Violante. Two views of inequality over the life cycle. *Journal of the European Economic Association*, 3(2-3):765–775, 2005.

- Matthew Higgins and Jeffrey G Williamson. Explaining inequality the world round: Cohort size, kuznets curves, and openness. Technical report, National Bureau of Economic Research, 1999.
- Stephen P Jenkins. Did the middle class shrink during the 1980s? UK evidence from kernel density estimates. *Economics letters*, 49(4):407–413, 1995.
- William R Johnson and Jonathan Skinner. Labor supply and marital separation. *The American Economic Review*, pages 455–469, 1986.
- Chinhui Juhn and Simon Potter. Changes in labor force participation in the united states. *The Journal of Economic Perspectives*, 20(3):27–46, 2006.
- Chinhui Juhn, Kevin M Murphy, and Brooks Pierce. Wage inequality and the rise in returns to skill. *Journal of political Economy*, pages 410–442, 1993.
- LIS Inequality and Poverty Key Figures. <http://www.lisdatacenter.org>, 2016. December, 2016, Luxembourg: LIS.
- Shelly Lundberg and Elaina Rose. Parenthood and the earnings of married men and women. *Labour Economics*, 7(6):689–710, 2000.
- Luxembourg Income Study (LIS) Database. <http://www.lisdatacenter.org>. Multiple Countries; 2016-2017. Luxembourg: LIS.
- Jacob A Mincer. Schooling, experience, and earnings. *NBER Books*, 1974.
- Eric R Nelson. The measurement and trend of inequality: Comment. *The American Economic Review*, pages 497–501, 1977.
- Lars Osberg. Long run trends in income inequality in the united states, uk, sweden, germany and canada: A birth cohort view. *Eastern Economic Journal*, 29(1):121–141, 2003.
- Morton Paglin. The measurement and trend of inequality: a basic revision. *The American Economic Review*, pages 598–609, 1975.
- Morton Paglin. The measurement and trend of inequality: Reply. *The American Economic Review*, pages 520–531, 1977.
- Steven G Prus. Income inequality as a canadian cohort ages an analysis of the later life course. *Research on Aging*, 22(3):211–237, 2000.
- Graham Pyatt. On the interpretation and disaggregation of gini coefficients. *The Economic Journal*, pages 243–255, 1976.
- Gayle Reznik, Dave Shoffner, and David A Weaver. Coping with the demographic challenge: fewer children and living longer. *Social Security Bulletin*, 66(4):37–45, 2005.
- Walter Garrison Runciman. Relative deprivation & social justice: Study attitudes social inequality in 20th century England. 1966.

Amartya Sen. *On economic inequality*. Oxford University Press, 1973.

Wendy Sigle-Rushton and Jane Waldfogel. The incomes of families with children: A cross-national comparison. *Journal of European Social Policy*, 17(4):299–318, 2007.

Bruce A Weinberg. Computer use and the demand for female workers. *Industrial & Labor Relations Review*, 53(2):290–308, 2000.

Appendix A. Pyatt's Decomposition Method

To understand Pyatt's method, first note that the inequality among a set of numbers $x_1, x_2, x_3, \dots, x_n$ can be expressed, in terms of the Gini coefficient, as

$$Gini = \frac{\left(\frac{1}{2n^2}\right) \sum_{i=1}^n \sum_{j=1}^n |x_j - x_i|}{\frac{1}{n} \sum_{j=1}^n x_i}$$

In other words, Gini can be written as the ratio of the mean absolute difference between each income pair in the society (x_i, x_j) over twice the average level of all incomes \bar{x} . Note that the Gini formula above can also be written as

$$Gini = \frac{2\left(\frac{1}{2n^2}\right) \sum_{i=1}^n \sum_{j=1}^n \max(0, x_j - x_i)}{\frac{1}{n} \sum_{j=1}^n x_i}$$

Now, imagine a game conducted for each individual in which individual i randomly picks person j from the population. If person j 's income is greater than his own income, he will switch to that income, otherwise he will keep his own income. Therefore, no individual can lose from participating in this game. As a result, the expected gain for individual i is

$$\frac{1}{n} \sum_{j=1}^n \max(0, x_j - x_i) \text{ for all } i$$

If now, we average the expected gains over all individuals, we obtain the following expression

$$\text{Average Expected Gain} = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \max(0, x_j - x_i)$$

Note that this is the numerator of the Gini formulation introduced above. This formula is similar to the calculation of the aggregate deprivation, where deprivation is defined as what the individual does not have but sees as feasible when comparing herself to other persons in the society.²¹ Now the Gini coefficient can be interpreted as the average expected gain (in the income comparison game), expressed in (or normalized by) the mean income. A higher expected gain means that the individual would be better off in someone else's shoes, therefore, it corresponds to a higher Gini coefficient. You can expect the poorest person in the population to have the highest expected gain and equal to the mean income, and the richest person to have an expected gain of zero.²²

Suppose now the population has been divided into m mutually exclusive and exhaustive groups, or in our case, age cohorts. We can now look at the average expected gain for individual in group i drawing incomes at random from group j

21. For a more detailed discussion on deprivation see Runciman (1966).

22. Amartya Sen has a similar interpretation of the Gini coefficient. He states that ■in any pair-wise comparison the man with the lower income can be thought to be suffering from some depression on finding his income to be lower. Let this depression be proportional to the difference in income. The sum total of all such depressions in all possible pair-wise comparisons takes us to the Gini coefficient■ (Sen, 1973).

$$\text{Average Expected Gain} = \sum_{i=1}^n \sum_{j=1}^n E(\text{gain}|i \rightarrow j) \Pr(i \rightarrow j)$$

$E(\text{gain}|i \rightarrow j)$ is the average expected gain, taken over all individuals in group i , when they draw a member of group j to compare within the game setup. It is easy to see that $\Pr(i \rightarrow j) = p_i p_j$ where p_i and p_j are population shares of groups i and j , respectively. Gini coefficient in matrix form can be written as

$$Gini = \frac{p' E p}{\left(\frac{1}{n}\right) \sum_{i=1}^n x_i} = \frac{p' E p}{s' p} = (s' p)^{-1} p' E p$$

Where s and p are vectors of income shares and population shares of each group, respectively. E is an $m \times m$ matrix with diagonal elements representing within-group expected gains and off-diagonal elements representing between-group expected gains. A reformulation of the Gini coefficient can be obtained by defining $\pi = (s' p)^{-1} \hat{s} p$ where π is a column vector with the i -th row being the proportion of group i 's income from the aggregate income. Therefore,

$$Gini = \pi' s^{-1} E p = \pi' E^* p$$

Where $E^* = \hat{s}^{-1} E$. Matrix E^* is simply a normalization of the matrix E by the mean income of each population group. Now, the decomposition of the Gini coefficient is done in two steps. In the first step, we calculate a matrix E_2^* , which is the normalized expected gains under the assumption that the members of each cohort have incomes equal to the mean income of their cohort. This should remind us of the corresponding P-equality line in Paglin (1975). As one would expect, the diagonal elements of E_2^* are all zero (since it is assumed that no inequality exists within each group), and so are those off-diagonal elements (i, j) for which mean income of group i is bigger than mean income of group j . The second step is to calculate matrix E_1^* as follows

$$E_1^* = E^* - E_2^*$$

The diagonal elements of E_1^* are the Gini coefficients of within-group inequality as in matrix E^* . We would expect the off-diagonal elements of E_1^* to be all zero in case the groups do not overlap. Those off-diagonal elements represent the inequality associated with those who are lagged behind. The overlapping terms can be thought of as the "across groups" contribution to the within-Gini coefficient.²³

In calculating the within-group inequality, it is a mistake to only focus on the diagonal elements elements of matrices E^* and E_1^* and toss everything else.²⁴ This approach in treating the overlap terms is "equivalent to putting blinders around members of each cohort as that they can only compare themselves with others in the same cohort" (Paglin, 1977). Based on this approach the within-cohort Gini can be calculated as follows

$$\text{Within-Gini} = \pi' E_1^* p$$

This version of the within-Gini coefficient is superior to the P-Gini coefficient since it takes into account the lagged individuals, a non-negligible part of the overall inequality.

23. A detailed explanation of the matrices can be found in Pyatt (1976).

24. This is what Nelson (1977) suggests.