Demographic and Socioeconomic Inequality in a Highly Educated and Increasingly Immigrant Workforce: The Case of Biomedical Research

Misty L. Heggeness, MSW, MPP, Ph.D.
National Institutes of Health
Office of Extramural Research

Frances D. Carter-Johnson, MS, Ph.D.
National Institutes of Health
Center for Scientific Review

January 30, 2015

**This extended abstract was developed for submission to the Population Association of America (PAA) 2015 Annual Meeting. Do not cite without authors' permission.

Short Abstract:

The biomedical research workforce is highly educated with an increasing immigrant population who come to the United States with J-1 or H-1 visas to study or receive additional advanced training. Although highly educated, the workforce faces a wide range of issues related to livable wages, salaries, and benefits – particularly for those still in training. The level of inequality in wages between and among principal investigators, staff scientists, postdoctoral associates, and graduate students is important to analyze as it has the potential to hinder or accelerate innovation in biomedicine. This paper reports changing demographics of this population overtime using nationally representative census and household survey data to provide an innovative perspective. We also study individual-level income data within the workforce and calculate Gini coefficients to study inequality over time. Our results show the current status of inequality within the field with respect to demographics and income. Policy implications are discussed.

Long Abstract:

Introduction

The biomedical research workforce is highly educated with an increasing immigrant population who come to the United States with J-1 or H-1 visas to study or receive additional advanced training. Even though the workforce is highly educated, it struggles with a wide range of issues related to livable wages, salaries, and benefits – particularly for those still in training. The level of inequality in wages between and among principal investigators, staff scientists, postdoctoral associates, and graduate students is important to analyze as it has the potential to hinder or accelerate innovation in biomedicine. In this paper, we report the changing demographics of this population overtime using decennial Census data to provide an innovative perspective. We also study income data of individuals within the biomedical research workforce and calculate Gini coefficients to study inequality over time. Our results show the current status of inequality within the field with respect to demographics and income. Policy implications are discussed.

Background

In 2012, two National Institutes of Health (NIH) Advisory Committees to the Director released reports emphasizing the importance of the availability and analysis of comprehensive workforce data to maintaining the sustainability and diversity of the U.S. biomedical research workforce. The Biomedical Research Workforce Working Group report (2012) considered the biomedical workforce as a whole and included National Science Foundation (NSF) and NIH administrative data to estimate the number of individuals at each step of training and post-training employment. The second committee, the Working Group on Diversity in the Biomedical Research Workforce, focused on training and outcome data related to underrepresented groups in the biomedical workforce, used parallel data and served as NIH's most immediate response to Ginther and colleagues (2011) analyses suggesting racial disparities in NIH R01 grant funding.

Both reports proved important in providing an initial overview of and recommendations for increased understanding about the biomedical research workforce as an important segment of the labor workforce. However, both reports also identified various data gaps that limited detailed analysis and understanding of the current and future status of the biomedical workforce. While these reports contributed to the body of literature on the biomedical workforce, they also make it clear that the lack of comprehensive data is hindering accurate understanding and policymaking to sustain and diversify this important workforce.

More comprehensive studies on the biomedical research workforce are warranted to improve analyses by government agencies aimed at effectiveness and social equality as presented by the NIH's advisory committees. We address the necessity of historically broad studies to further understand the labor force and population changes for businesses and policymakers, who must consider the U.S. labor force, immigration and social equality simultaneously. Many previous studies focused on the STEM workforce in general and not the biomedical research workforce. However, with NIH's budget (approximately $30B) consistently exceeding the budgets of NSF (approximately $7 billion) and other federal agencies, it is important to increase understanding about the biomedical research workforce separate from other groups.

Historical and recent studies as well as the NIH reports that concentrated on the biomedical research workforce utilize data such as NSF's Survey of Earned Doctorates or American Association of Medical Colleges datasets (Stephan, 2012; Garrison & Gerbi, 1998). The former data set is limited to researchers who completed doctoral degrees in the United States and therefore does not account for researchers with degrees from foreign universities who enter the workforce as postdoctoral or established researchers. The latter dataset is restricted to individuals at medical schools.

Previous studies have illustrated increases in foreign born doctoral degree recipients on the order of more than 200% over time (Garrison & Gerbi, 1998); however, no studies have provided this information for foreign educated researchers working in the U.S. Increased understanding of the impact of the multi-faced foreign born scientist population is important to all levels of policy making related to the biomedical research workforce. This paper expands our understanding of the biomedical research workforce in the United States by using historical workforce data to close the gap on our understanding of the stock of biomedical researchers living and working in the U.S. each year, including both citizens and non-citizens.

Data and Methodology

We analyze data from the IPUMS decennial censuses from 1950 to 2000 and the IPUMS American Community Survey from 2010 (Ruggles et al. 2010). Using nationally representative census and household survey data allows for the identification of the entire workforce of biological and medical scientists, including those who were trained outside the United States. Another advantage of using a nationally representative dataset is it allows us to identify the stock of individuals in any given moment in time who identify themselves as biological and medical scientists. Having data on the stock of individuals over five decades, we are able to calculate Gini coefficients and Lorenz curves for income inequality within biomedical research for each decade.

Our data and analysis include respondents with five or more years of education and who work in biomedical research related occupations. Therefore, biomedical trained scientists in occupations unrelated to biomedical research are not included in our analyses. The data set includes demographic variables as well as those on birth year, citizenship status, country of birth, employment and income for each decennial census year such as household and individual salary. These rich variables allow for a detailed descriptive examination of trends for important and previously unanalyzed variables, which has the potential for increasing the quality of the data and analysis on the biomedical research workforce. We also show trends in biological and medical scientists by industry, socioeconomic status, gender, race/ethnicity, age, and citizenship status over the time period. The richness of the data also allowed creation of synthetic cohorts of biomedical scientists using the birth year variable that facilitated analysis of the representation of each cohort over time as well as median income by year.

Results

Figure 1 shows the exponential increase of biological and medical scientists in the past 100 years. Until 1970, the number of biomedical scientists was consistently less than 50,000. However, the population of biomedical scientists increased to approximately 80,000 in 1970, has consistently increased to more than 200,000 in 2010 and actually doubled between 1990 and 2000.

Figure 2 shows gender parity in the overall workforce by 2010. Consistent increases in female representation in the biomedical workforce occurred beginning as early as the data is available (1960-1970). However, analysis (not shown) of female representation by industry confirmed findings from previous studies of gender underrepresentation in both federal government and academic employment sectors.

In terms of race/ethnicity and citizenship status, our analyses are driven by calls for additional and comprehensive information on the representation of various subgroups in the biomedical research workforce. For example, Figure 3 reveals a trend of increasing representation for all racial groups over the entire span of the data. Our analyses are also the first opportunity to compare historical changes for citizens and non-citizens in the biomedical research workforce (Figure 4). Our study reveals increases over time for U.S. citizens with some fluctuations in 1990 where the number of U.S. citizens reported decreases. The ability to analyze citizenship status of the biomedical research workforce over time reveals consistent and large increases in naturalized and non-citizens during periods when U.S. citizen representation decreased (1980 to 1990) or experienced smaller growth (2000-2010).

These demographic trends provide the backdrop from which we will analyze inequality within the biomedical research workforce. Our examination of inequality by income (Figure 5) over time suggests increases in income inequality in the biomedical research workforce over time from 0.29 in 1950 to 0.37 in 2010. Income inequality jumped from 1950 to 1960 (0.29 to 0.35) and from 1960 to 1970 (0.35 to 0.39). Income inequality was the highest in 1980 at 0.41 before decreasing after 1980 (1990: 0.40, 2000:0.39, 2010: 0.37). However, the downward trends in income inequality since 1980 are small in magnitude compared to larger increases in inequality occurring between 1950 and 1970.

Stephan (2012) also analyzed income equality from 1973 to 2006 for academic faculty by rank for various STEM disciplines, including the life sciences. For life sciences faculty, income inequality was consistently lower (ranging from 0.12 to 0.25 at its largest) than in broader society (ranging from 0.31 to 0.42), suggesting more equal income distributions. Within academe, income inequality was consistently larger for full professors compared to associate and assistant professors, respectively, implying similar salaries amongst associate and assistant professors. Income inequality also steadily increased over the time period for all ranks of life sciences faculty. Comparing Stephan's findings with our income inequality estimations for the full biomedical research workforce for the same time frame, we find consistently larger income inequality within the full biomedical research workforce population (1970: 0.39) that is more comparable (actually larger) to income inequality in broader society (0.31). It is also important to note that the fluctuations in income inequality that we report for the full biomedical research workforce do not present to exist for life sciences faculty.

Discussion

We discuss potential social, economic and policy changes within the context of income inequality of individuals in the biomedical research field. For example, the trend of increases in overall employees in the biomedical research workforce overlaps with the NIH budget doubling from 1998 to 2003. Combining historical trends analysis with social, economic, and political context will not only highlight major changes in the characteristics of this workforce overtime but will also identify historic patterns and resulting policy changes. It also provides a relevant background from which to understand the relevant changes in inequality.

The ability to examine previous historical trends and the impacts of associated policy changes allows for a retrospective assessment of the effectiveness of each policy solution. By comparing the effectiveness of historical policies with the current state of biomedical research industry, we will increase the current understanding of the state and needs of the biomedical research workforce and all relevant stakeholders, such as government agencies, pharmaceutical businesses and educational institutions that both depend on and contribute to biomedical research. Increased understanding from our work will identify future areas for concern as well as provide critical and necessary direction for combining and analyzing alternative data sets to further increase understanding.

Figure 1.

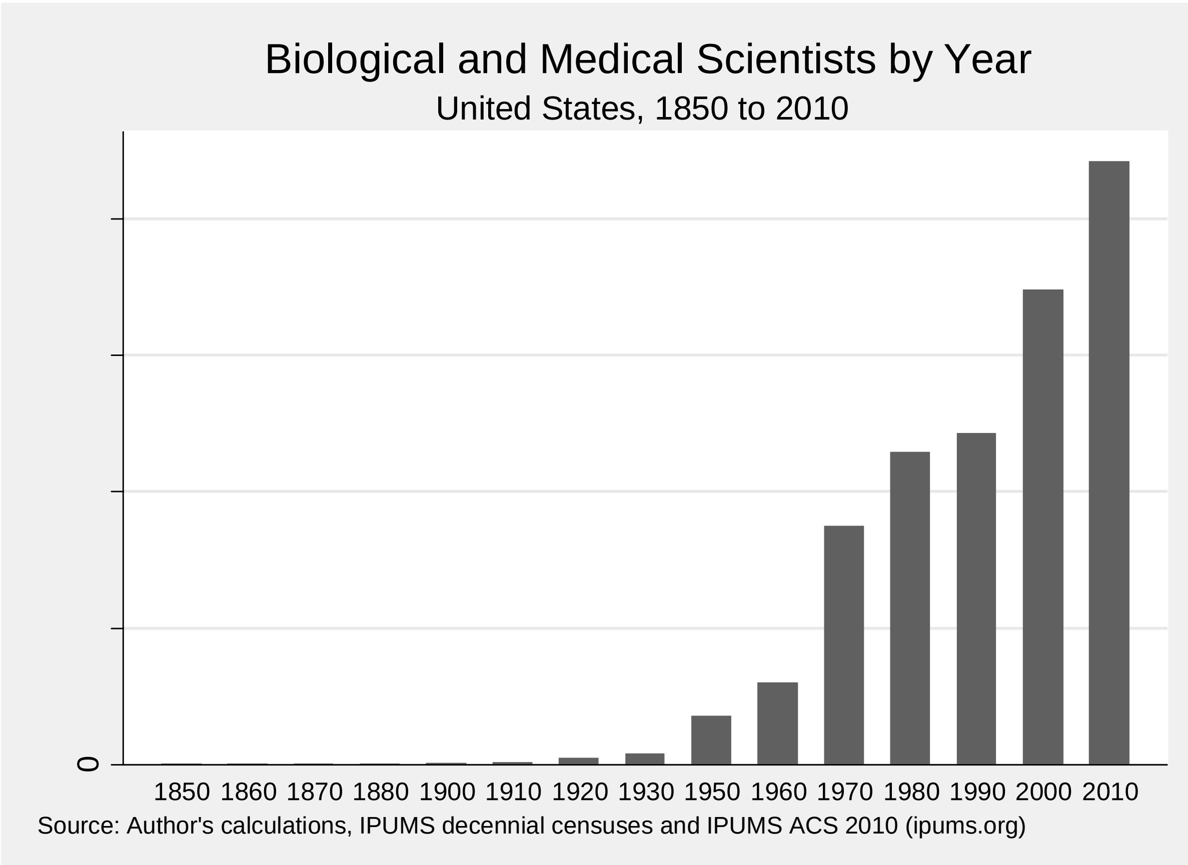
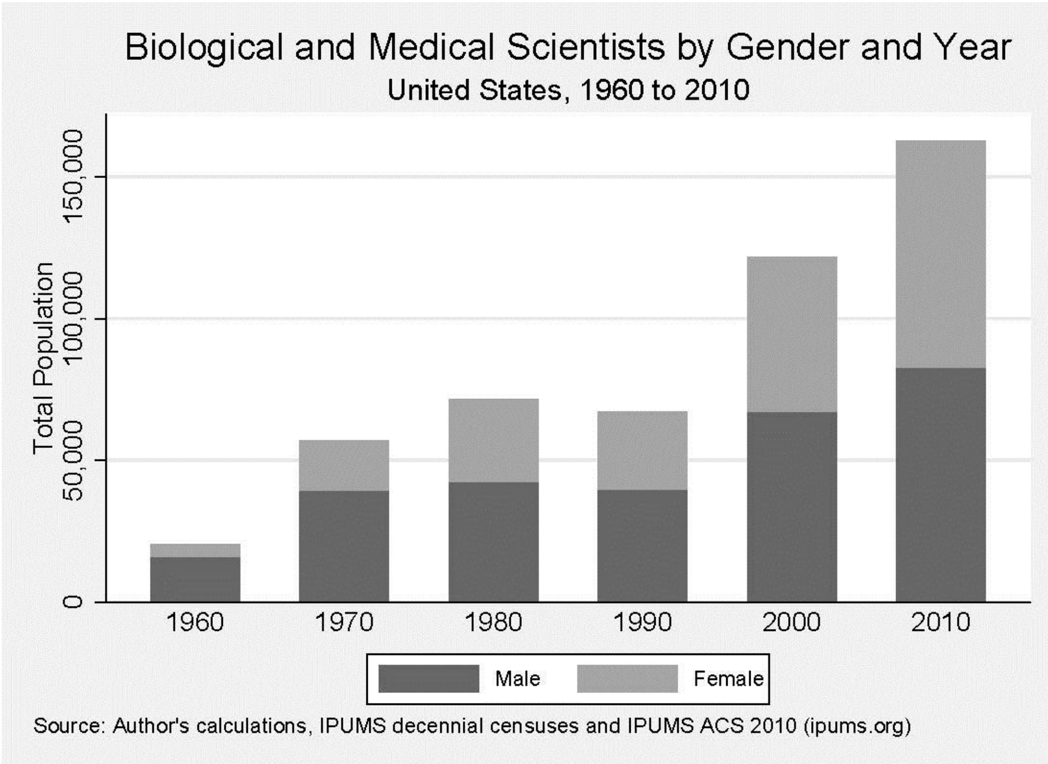
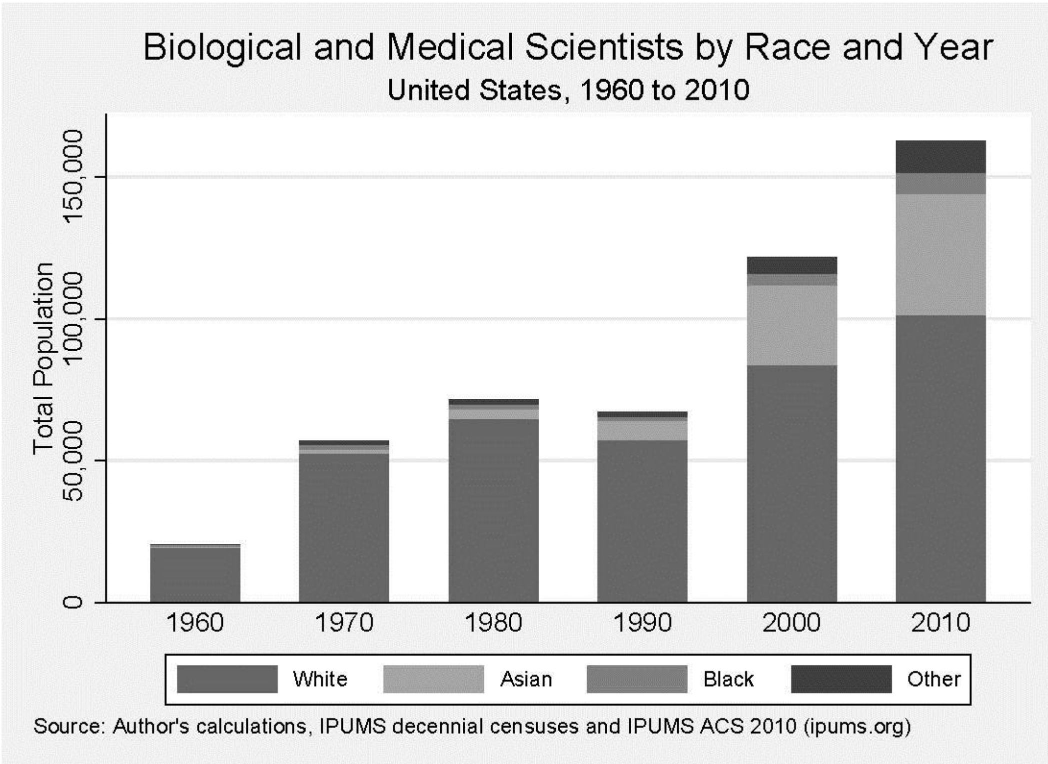

Biological and Medical Scientists by Year
United States, 1850 to 2010
0
1850 1860 1870 1880 1900 1910 1920 1930 1950 1960 1970 1980 1990 2000 2010
Source: Author's calculations, IPUMS decennial censuses and IPUMS ACS 2010 (ipums.org)

Figure 2.



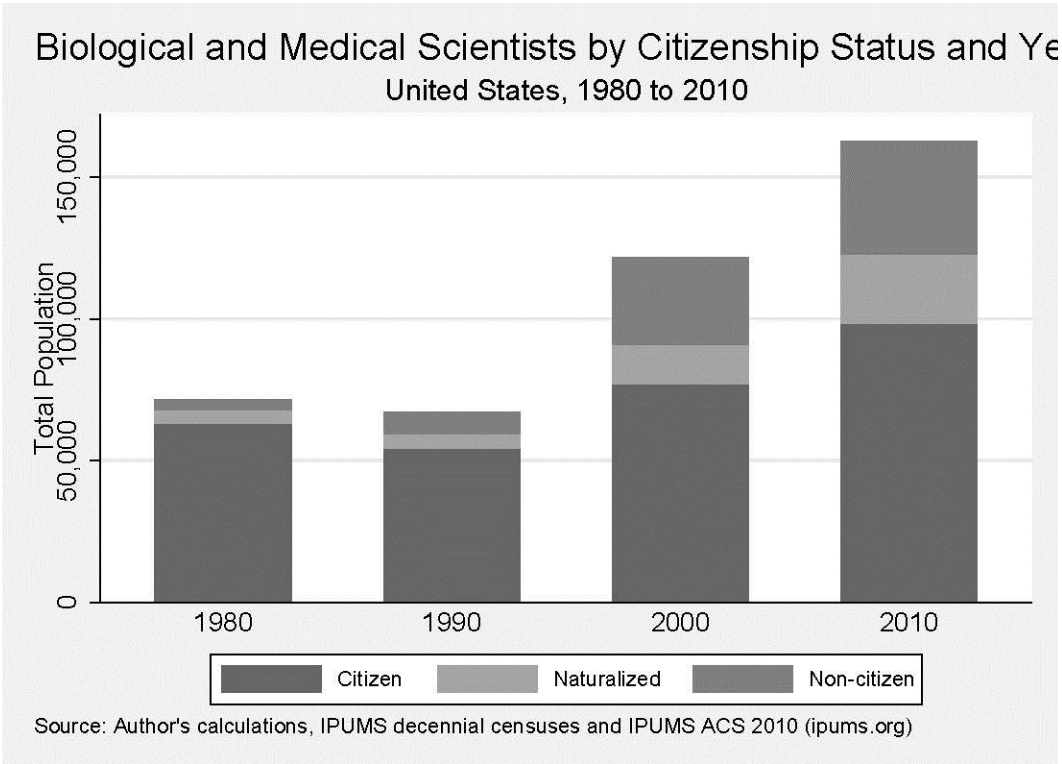
Biological and Medical Scientists by Gender and Year
United States, 1960 to 2010
Total Population
0
50,000
100,000
150,000
1960
1970
1980
1990
2000
2010
Male
Female
Source: Author's calculations, IPUMS decennial censuses and IPUMS ACS 2010 (ipums.org)

Figure 3.



Biological and Medical Scientists by Race and Year
United States, 1960 to 2010
Total Population
0
50,000
100,000
150,000
1960
1970
1980
1990
2000
2010
White
Asian
Black
Other
Source: Author's calculations, IPUMS decennial censuses and IPUMS ACS 2010 (ipums.org)

Figure 4.



Biological and Medical Scientists by Citizenship Status and Ye
United States, 1980 to 2010
Total Population
0
50,000
100,000
150,000
1980
1990
2000
2010
Citizen
Naturalized
Non-citizen
Source: Author's calculations, IPUMS decennial censuses and IPUMS ACS 2010 (ipums.org)

Figure 5.

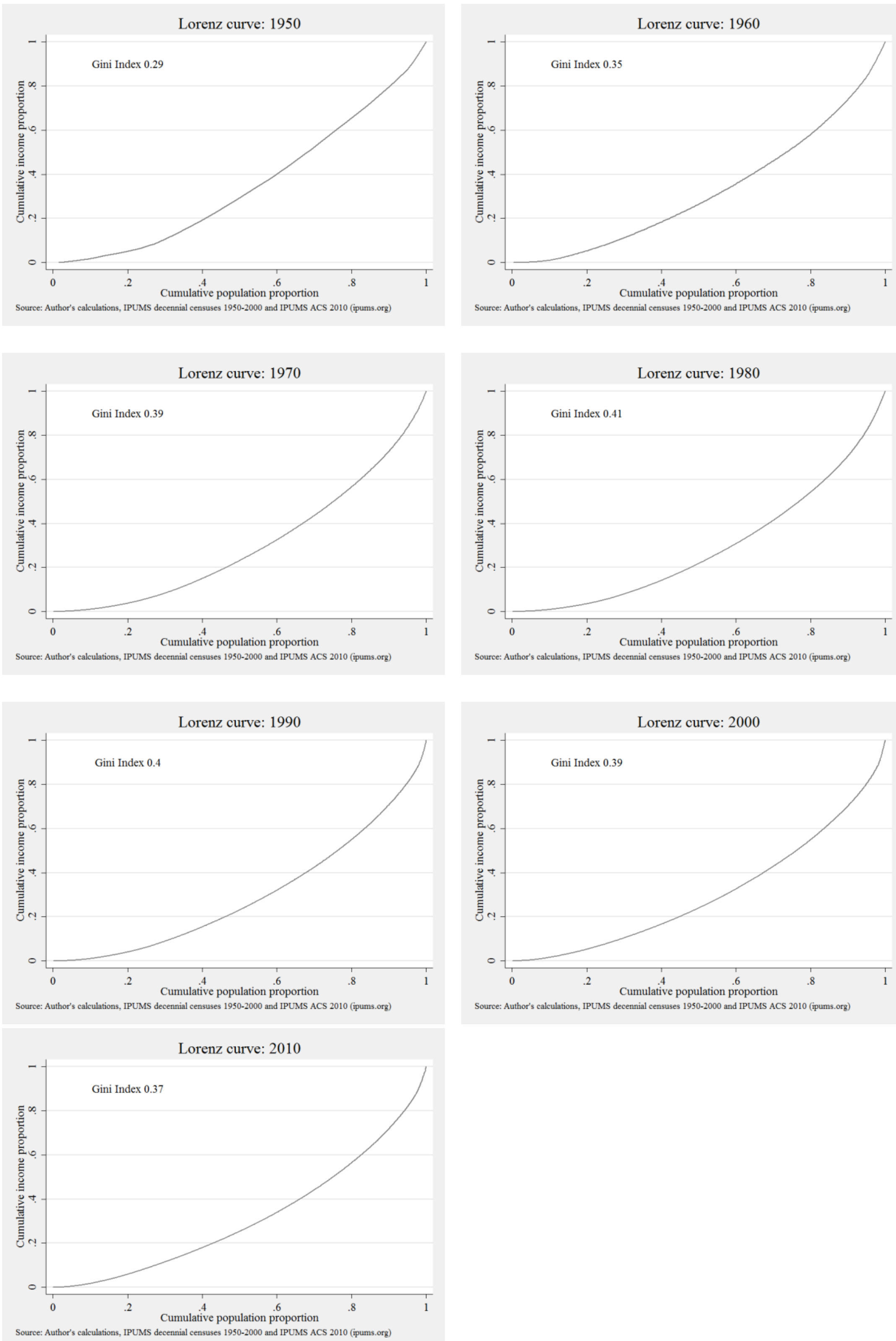

Lorenz curve: 1950
Gini Index 0.29
Cumulative income proportion
Cumulative population proportion
Source: Author's calculations, IPUMS decennial censuses 1950-2000 and IPUMS ACS 2010 (ipums.org)
Lorenz curve: 1960
Gini Index 0.35
Cumulative income proportion
Cumulative population proportion
Source: Author's calculations, IPUMS decennial censuses 1950-2000 and IPUMS ACS 2010 (ipums.org)
Lorenz curve: 1970
Gini Index 0.39
Cumulative income proportion
Cumulative population proportion
Source: Author's calculations, IPUMS decennial censuses 1950-2000 and IPUMS ACS 2010 (ipums.org)
Lorenz curve: 1980
Gini Index 0.41
Cumulative income proportion
Cumulative population proportion
Source: Author's calculations, IPUMS decennial censuses 1950-2000 and IPUMS ACS 2010 (ipums.org)
Lorenz curve: 1990
Gini Index 0.4
Cumulative income proportion
Cumulative population proportion
Source: Author's calculations, IPUMS decennial censuses 1950-2000 and IPUMS ACS 2010 (ipums.org)
Lorenz curve: 2000
Gini Index 0.39
Cumulative income proportion
Cumulative population proportion
Source: Author's calculations, IPUMS decennial censuses 1950-2000 and IPUMS ACS 2010 (ipums.org)
Lorenz curve: 2010
Gini Index 0.37
Cumulative income proportion
Cumulative population proportion
Source: Author's calculations, IPUMS decennial censuses 1950-2000 and IPUMS ACS 2010 (ipums.org)

Table 1. Descriptive Statistics of Biological and Medical Scientists, United States, 1960 to 2010

Year	1960	1970	1980	1990	2000	2010
Total Population	20,518	55,400	71,700	67,248	121,970	162,778
Sex						
Male	16,133	37,700	42,500	39,746	67,012	82,617
Female	4,385	17,700	29,200	27,502	54,958	80,161
Citizenship Status						
Citizen	--	--	63,100	54,579	77,158	98,425
Naturalized	--	--	4,800	5,081	13,923	24,309
Non-Citizens	--	--	3,800	7,588	30,889	40,044
Race						
White	19,325	51,200	64,900	57,481	83,944	101,464
Black	596	1,400	1,700	1,474	3,894	7,375
Other	597	2,800	5,100	8,293	34,132	53,939
Industry						
Private	--	8,700	25,700	32,579	76,863	113,304
Academic	--	42,800	38,400	28,662	30,255	27,881
Public	--	3,900	7,600	6,007	14,852	21,593

Source: Authors' calculations, IPUMS decennial censuses and IPUMS ACS 2010 (ipums.org)

Table 2. Descriptive Statistics of Biological and Medical Scientists (continued), United States, 1960 to 2010

Year	1960	1970	1980	1990	2000	2010
Age	41.4	39.3	39.5	41.3	41.2	42.8
Income	$47,619	$55,325	$50,345	$50,939	$47,818	$53,676

Source: Author's calculations, IPUMS decennial censuses and IPUMS ACS 2010 (ipums.org)

References:

Howard H. Garrison and Susan A. Gerbi. (1998). Education and Employment Patterns of U.S. Ph.D.'s in the Biomedical Sciences. The Journal of the Federation of American Societies for Experimental Biology, 12: 140-147.

Donna K. Ginther, Walter T. Schaffer, Joshua Schnell, Beth Masimore, et al. (2011). Race, ethnicity, and NIH research awards. Science, 333:1015-1019.

National Institutes of Health, Advisory Committee to the Director (2012). Biomedical Research Workforce Working Group Report, June 14, 2012. Accessed April 10, 2014 from, (http://acd.od.nih.gov/Biomedical_research_wgreport.pdf.

National Institutes of Health, Advisory Committee to the Director. (2012). Working Group on Diversity in the Biomedical Research Workforce Report, June 13, 2012. Accessed August 13, 2014 from, http://acd.od.nih.gov/Diversity%20in%20the%20Biomedical%20Research%20Workforce%20Report.pdf.

Steven Ruggles, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]. Minneapolis: University of Minnesota, 2010.

[1] For more details, see: http://report.nih.gov/investigators_and_trainees/ACD_BWF/index.aspx.