

# The Glass Ceiling and The Paper Floor: Gender Differences among Top Earners, 1981–2012\*

Fatih Guvenen      Greg Kaplan      Jae Song

February 21, 2020

## Abstract

We analyze changes in the gender structure at the top of the earnings distribution in the United States over the last 30 years using a 10% sample of individual earnings histories from the Social Security Administration. Despite making large inroads, women still constitute a small proportion of the top percentiles: the glass ceiling, albeit a thinner one, remains. We measure the contribution of changes in labor force participation, changes in the persistence of top earnings, and changes in industry and age composition to the change in the gender composition of top earners. A large proportion of the increased share of women among top earners is accounted for by the mending of, what we refer to as, the *paper floor* – the phenomenon whereby female top earners were much more likely than male top earners to drop out of the top percentiles. We also provide new evidence on the top of the earnings distribution for both genders: the rising share of top earnings accruing to workers in the Finance and Insurance industry, the relative transitory status of top earners, the emergence of top earnings gender gaps over the life cycle, and gender differences among lifetime top earners.

**JEL Codes:** E24, G10, J31.

**Keywords:** Top earners, Glass ceiling, Gender gap, Paper floor, Industry

---

\*Guvenen: University of Minnesota, Federal Reserve Bank of Minneapolis, and NBER, e-mail: guvenen@umn.edu; Kaplan: University of Chicago and NBER, e-mail: gkaplan@uchicago.edu; Song: Social Security Administration, e-mail: jae.song@ssa.gov. This paper uses confidential data supplied by the Social Security Administration. The views expressed herein are those of the authors and not necessarily those of the Social Security Administration, the Federal Reserve Bank of Minneapolis, or the Federal Reserve System.

# 1 Introduction

The last three decades have seen tremendous changes in the distribution of earnings in the United States. Among these changes, two of the most well known are the increasing share of total earnings that accrues to top earners (i.e., individuals in the top 1 percent or top 0.1 percent of the earnings distribution) and the continued relative absence of women from this top earning group.<sup>1</sup> This latter phenomenon is commonly referred to as the *glass ceiling*, the emergence of which has spurred both debate over the appropriate policy response, as well as active research into its primary causes.<sup>2</sup> However, progress on both fronts has been hampered by a relative lack of good evidence on the gender structure at the top of the earnings distribution.<sup>3</sup> Our goal in this paper is to provide this necessary empirical evidence on the glass ceiling, using newer and better data than has been previously available. In doing so, we also revisit several important questions about top earners of both genders: the dynamics of their earnings, their industry composition, their age and cohort composition, and the evolution of earnings for lifetime top earners.

Our interest in top earners is motivated by their disproportionately large influence on the aggregate economy. This influence operates through at least three channels. First, top earners are crucial economic actors. In the United States, individuals in the top 1 percent of the income distribution earn approximately 15% of aggregate before-tax income and pay about 40% of individual income taxes – more than one and a half times the amount paid by the bottom 90 percentiles – and 50% of all corporate income tax.<sup>4</sup> Since this group includes virtually all high-level managers and executives of U.S. businesses (both public and private), top earners play a pivotal role in decisions about business investment, employment creation, layoffs, and international trade. Second, top earners are key political actors. Political scientists have argued that the increasing polarization of political discourse in the United States can be partly attributed to the rising influence of top earners, through political contributions that have in part been made possible by changes in campaign finance regulations since the 1970s.<sup>5</sup> Third, since the group of top earners includes a large fraction of the economy’s top talent, understanding the distribution of top earners across gender, industries, and cohorts helps us to better understand the allocation of human capital in the

---

<sup>1</sup>See [Bertrand et al. \(2010\)](#) and [Gayle et al. \(2012\)](#) for recent attempts to measure the gender composition of top earners.

<sup>2</sup>The term “glass ceiling” was coined in the 1980s and is typically defined as an “unseen, yet unbreachable barrier that keeps minorities and women from rising to the upper rungs of the corporate ladder, regardless of their qualifications or achievements” (see, for example, [Federal Glass Ceiling Commission \(1995\)](#)).

<sup>3</sup>Existing evidence is based almost exclusively on non-random sub-samples of top earners (CEOs and other top executives, billionaires from Forbes 400 lists, MBA graduates). We review this evidence below.

<sup>4</sup>Statistics are for 2010 from the [Congressional Budget Office \(2013, Table 3\)](#).

<sup>5</sup>See, for example, [Barber \(2013\)](#); [Baker et al. \(2014\)](#).

economy.

The pivotal role of top earners has led to a burgeoning literature whose goal is to explicitly model the thick Pareto tail at the top end of the earnings distribution and then either evaluate alternative mechanisms that could give rise to top earners (e.g., [Gabaix and Landier \(2008\)](#), [Jones and Kim \(2014\)](#)), study the allocation of top talent across occupations (e.g., [Hsieh et al. \(2013\)](#)), or ask how to best design fiscal policy in the presence of influential top earners (e.g., [Saez \(2001\)](#); [Badel and Huggett \(2014\)](#); [Guner et al. \(2014\)](#); [Kindermann and Krueger \(2014\)](#)). Therefore, one goal of this paper is to provide the empirical evidence that this literature requires in order to address these issues – on gender differences, persistence, mobility, age, and industry composition, and on the life-cycle dynamics of top earners. The literature on optimal taxation of top earners has so far only considered the taxation of individuals; as this literature moves toward studying the taxation of families, evidence on gender differences among top earners of the type we provide will become essential.

Our data set is a 10% representative sample of individual earnings histories from the U.S. Social Security Administration. Several features of these data are well suited for our goals. The large number of observations enables us to study earnings within the top 1 percent, including the earnings of those at the very top, the 0.1 percent, as well as the characteristics of female top earners, who constitute only a small subset of top earners. The panel nature of the data set enables us to track the same individuals over time and, hence, to perform our analysis using both five-year average earnings as well as annual earnings. This is important because of the relatively low probabilities of top earners remaining in the top percentiles from year to year, as shown by [Auten et al. \(2013\)](#), and which we confirm and expand on. The presence of Employer Identification Numbers (EIN) from W2 forms enables us to obtain detailed industry information about each worker’s jobs, which we use to construct an industry breakdown that is informative about the types of jobs held by top earners. In particular, we separate workers in Finance and Insurance, Health services, Legal services, and Engineering from executives in other service industries. The 32-year time span of our data and the absence of attrition both enable us to paint a sharper picture of how top earners’ earnings evolve over their life cycles than has been possible in previous work.

Our findings on gender differences speak to three broad themes: (i) trends in top earnings over the last three decades; (ii) the persistence and mobility of top earners; and (iii) the characteristics of top earners.

First, regarding recent trends in top earnings, we find that although large strides have been taken toward gender equality at the top of the distribution, very large differences between men and women still remain. Since 1981, the share of women among top earners has increased by more than a factor of 3. Yet in 2012, the earnings share of women still

comprised only 11% of the earnings of all individuals in the top 0.1 percent, and only 18% of the earnings of the top 1 percent. The glass ceiling is still there, but it is thinner than it was three decades ago. Moreover, among the top 0.1 percent, virtually all of the increase came in the 1980s and 1990s; the last decade has seen almost no further improvement. We decompose the rise in the share of women among top earners into a component that is due to changes in female participation in all parts of the distribution and find that these compositional effects play little role in explaining the observed trend. This finding reflects the fact that gender differences have narrowed much less in the bottom 99 percent of the distribution than in the top percentiles – the fraction of women in the bottom 99 percent increased from 43% in 1981 to 49% in 2012.

For top earners of both genders, after several decades of rising earnings, a leveling off has taken place during the last decade. Both the thresholds for membership and the average earnings of workers in the top percentiles have remained relatively flat since 2000. It is too soon to tell whether this represents a change in the increasing trend that has dominated the last half century (Kopczuk et al. (2010)), or whether it is a temporary flattening due to top earners suffering disproportionately large temporary falls in earnings during the 2000–2 and 2008–9 recessions (Güvenen et al. (2014b)).<sup>6</sup>

Second, regarding persistence and mobility at the top of the earnings distribution, we find substantial turnover among top earners. The frequency with which workers enter and exit the top earnings groups sounds a cautionary note to analyses of top earners that use only data from annual cross sections. This high tendency for top earners to fall out of the top earnings groups was particularly stark for women in the 1980s – a phenomenon we refer to as the *paper floor*. But the persistence of top earning women has dramatically increased in the last 30 years, so that today the paper floor has been largely mended. Whereas female top earners were once around twice as likely as men to drop out of the top earning groups, today they are no more likely than men to do so. Moreover, this change is not simply due to women being more equally represented in the upper parts of the top percentiles; the same paper floor existed for the top percentiles of the female earnings distribution, but this paper floor has also largely disappeared. We use a decomposition to show that this change in persistence accounts for a substantial fraction of the increase in the share of women

---

<sup>6</sup>See Güvenen and Kaplan (2017) for a reconciliation of this finding with results from other data sources and different samples that show a continued increase in the income share of the top 0.1 percent over this period. The difference in these findings is not due to our focus on wage and salary income as opposed to a broader measure. In Güvenen and Kaplan (2017) we show that the slowdown in the growth of top incomes is present for total income (including capital gains), except at the very top of the distribution (above the 99.99th percentile). Instead the difference in findings is mainly due to differences in the implied trends for the bottom 99 percent that arise because of the different units of analysis: individuals who satisfy a minimum earnings and age restriction, versus all tax units.

among top earners that we observe during the last three decades.

As the persistence of top earning women was catching up with men during this period, the persistence of top earning men was itself increasing, particularly after 2000. Throughout the 1980s and 1990s, the probability that a male in the top 0.1 percent was still in the top 0.1 percent one year later remained at around 45%, but by 2011 this probability had increased to 57%. When combined with our finding that the share of earnings accruing to the top 0.1 percent has leveled off since 2000, this implies a striking observation about the nature of top earnings inequality: despite the total share of earnings accruing to the top percentiles remaining relatively constant in the last decade, these earnings are being spread among a decreasing share of the overall population. Top earner status is thus becoming more persistent, with the top 0.1 percent slowly becoming a more entrenched subset of the population.

Third, regarding the industry composition of top earners, we find that the finance industry dominates for both men and women. In 2012, finance and insurance accounted for around one-third of workers in the top 0.1 percent. However, this was not the case 30 years ago, when the health care industry accounted for the largest share of the top 0.1 percent. Since then, top earning health care workers have dropped to the second 0.9 percent where, along with workers in finance and insurance, they have replaced workers in manufacturing, whose share of this group has dropped by roughly half. Perhaps surprisingly, these changes in industry structure do not play much of a role in explaining either the level or the change in the share of women among top earners, because the industry composition of the top percentiles is very similar for men and women.

Fourth, in order to gain some insight into possible future trends for the glass ceiling, we also examine the age and cohort composition of top earners. Top earners are older than average and have become more so over time. In contrast with analyses of the gender structure of corporate boards (e.g., [Bertrand et al. \(2012\)](#)), we do not find that female top earners are younger than male top earners. Entry of new cohorts, rather than changes within existing cohorts, account for most of the increase in the share of women among top earners. These new cohorts of women are making inroads into the top 1 percent earlier in their life cycles than previous cohorts. If this trend continues, and if these younger cohorts exhibit the same trajectory as existing cohorts in terms of the share of women among top earners, then we might expect to see further increases in the share of women in the overall top 1 percent in coming years. However, this is not true for the top 0.1 percent. At the very top of the distribution, young women have not made big strides: the share of women among the top 0.1 percent of young people in recent cohorts is no larger than the corresponding share of women among the top 0.1 percent of young people in older cohorts.

All of the above findings pertain to a relatively short-run perspective on identity of top earners, based either on annual earnings or five-year average earnings. But for many questions about top earners, such as human capital accumulation or optimal taxation, a longer-run perspective, based on lifetime earnings, is more relevant. However, due to the previous lack of large, long panel data sets on earnings, little is currently known about lifetime top earners. We end the paper by examining how male and female lifetime top earners differ over the life cycle, where in the distribution these individuals start their working lives, and in which parts of the distribution they spend the majority of their careers. We find that within the top 1 percent of lifetime earnings, men and women display distinct lifecycle patterns, so that the gender gap between these groups is inverse U-shaped over the life cycle, increasing substantially in the 30s (presumably when some females' careers are interrupted for family reasons) and then declining toward retirement.

Our results on the glass ceiling relate to a large and active literature. However, the bulk of the existing empirical evidence has been relatively indirect and pertains to somewhat specialized subsets of top earners, such as CEOs and other executives, members of corporate boards, the list of billionaires compiled by *Forbes* magazine, or MBA graduates from a top U.S. business school (e.g., [Bell \(2005\)](#), [Wolfers \(2006\)](#), [Bertrand et al. \(2010\)](#), [Gayle et al. \(2012\)](#)). Although these analyses have revealed a wealth of interesting information, the extent to which their conclusions carry over to other top earning women is unknown. For example, [Wolfers \(2006\)](#) reports that over a 15-year period starting in the early 1990s, only 1.3% of ExecuComp CEOs were women. This is about about 10 times smaller than the share of women we find among the top 0.1 percent of earners in the 2000s.

Finally, this paper is also related to the literature initiated by [Piketty and Saez \(2003\)](#) that aims to understand the evolution of top earnings. More recently, [Parker and Vissing-Jørgensen \(2010\)](#) and [Guvenen et al. \(2014a\)](#) have studied the cyclicity of top earnings. Our focus is on long-run trends rather than the cycle. [Kopczuk et al. \(2010\)](#), [Bakija et al. \(2012\)](#), and [Auten et al. \(2013\)](#) are related papers that also use large representative samples of individual-level data to study the trends and characteristics of top earners. [Brewer et al. \(2007\)](#) is a complementary paper that analyzes the characteristics of high income individuals in the United Kingdom. However, these papers do not focus on the glass ceiling or the paper floor.

## 2 Data

### 2.1 Data Source

We use a confidential panel data set of earnings histories from the U.S. Social Security Administration (SSA) covering the period 1981 to 2012.<sup>7</sup> The data set is constructed by drawing a 10% representative sample of the U.S. population from the SSA’s Master Earnings File (MEF). The MEF is the main record of earnings data maintained by the SSA and contains data on every individual in the United States who has a Social Security number (SSN). The data set contains basic demographic characteristics, including date of birth, sex, race, type of work (farm or non-farm, employment or self-employment), employee earnings, self-employment taxable earnings, and the Employer Identification Number (EIN) for each employer, which we use to link industry information. Employee earnings data are uncapped (i.e., there is no top-coding) and include wages and salaries, bonuses, and exercised stock options as reported on the W-2 form (Box 1). The data set grows each year through the addition of new earnings information, which is received directly from employers on the W-2 form.<sup>8</sup> For more information on the MEF, see [Panis et al. \(2000\)](#) and [Olsen and Hudson \(2009\)](#). We convert all nominal variables into 2012 dollars using the personal consumption expenditure (PCE) deflator. For an individual born in year  $c$ , we define their age in year  $t$  as  $t - c$ , which corresponds to their age on December 31 of that year.

To construct the 10% representative sample from the MEF, we select all individuals with the same last digit of (a transformation of) their SSN. Since the last four digits of the SSN are randomly assigned to individuals, this generates a nationally representative panel. The panel tracks the evolution of the U.S. population in the sense that each year, 10% of new individuals who are issued SSN numbers enter our sample, and those who die each year are eliminated (determined through SSA death records).

---

<sup>7</sup>The data set contains earnings information going back to 1978. However, prior to 1981 the data are of poorer quality due to inconsistencies in complying with the switch from quarterly to annual wage reporting by employers, mandated by the SSA (see [Olsen and Hudson \(2009\)](#) and [Leonesio and Del Bene \(2011\)](#)). For large parts of the population, most of these reporting errors can be corrected. However, these methods do not work well for very high-earning individuals, who are the focus of this paper.

<sup>8</sup>The MEF also contains earnings information on self-employment income for sole proprietors (i.e., income reported on form Schedule SE; see [Olsen and Hudson \(2009\)](#) for more information); however, these data are top-coded at the taxable limit for Social Security contributions prior to 1994. Because of this top-coding, we focus our main analysis on wage and salary data. In [Appendix G](#), we verify the robustness of our findings to the inclusion of self-employment income for the period 1994-2012.

## 2.2 Sample Selection

For the analyses in Sections 3, 4, 5, and 6, in each year  $t$  we select all individuals in our baseline 10% sample who satisfy the following two criteria:

1. The individual is between 25 and 60 years old.
2. The individual has annual earnings that exceed a time-varying minimum threshold. This threshold is equal to the earnings one would obtain by working for 520 hours (13 weeks at 40 hours per week) at one-half of the legal minimum wage for that year. In 2012, this corresponded to annual earnings of \$1,885.

We impose these selection criteria in order to focus on workers with a reasonably strong attachment to the labor market and to avoid issues that arise when taking the logarithm of small numbers. These criteria also make our results comparable to the literature on earnings dynamics and inequality, where imposing age and minimum earnings restrictions is standard (see, e.g., [Abowd and Card \(1989\)](#), [Juhn et al. \(1993\)](#), [Meghir and Pistaferri \(2004\)](#), [Storesletten et al. \(2004\)](#), and [Autor et al. \(2008\)](#)).

The MEF contains a small number of extremely high earnings observations each year. To avoid potential problems with outliers, we cap (winsorize) observations above the 99.999th percentile of the distribution of earnings for individuals who satisfy the above two selection criteria in a given year. From 1981 to 2012, the mean and median 99.999th percentiles across years were both \$11.5 million, and the maximum, which was in 2000, was \$25.4 million.

We report results using two definitions of earnings: (i) annual earnings and (ii) five-year average earnings. Annual earnings provide us with a snapshot of top earners in each year. But top earners in a given year include some workers whose high earnings were a one-off event such as the receipt of a large one-time bonus or other windfall. Such workers would not be considered as top earners when using five-year average earnings, which focuses on individuals with more stable membership of the top earnings percentiles.

For the analyses using annual earnings, in each year  $t = 1981, \dots, 2012$  we assign all individuals who satisfy the two selection criteria to a percentile group based on their earnings in year  $t$ . We focus mostly on the top 0.1 percent, second 0.9 percent and bottom 99 percent, but we also report some results from finer groupings within the bottom 99.9 percent. For the analysis using five-year average earnings, we construct a rolling panel for each year  $t = 1983, \dots, 2010$  that consists of all individuals who satisfy the two selection criteria in at least three of the years from  $t - 2, \dots, t + 2$ , including the most recent year  $t + 2$ . For each



of these individuals, we compute their average annual earnings over the years  $t - 2, \dots, t + 2$  that they satisfy the selection criteria. We then assign individuals to percentile groups based on these five-year average earnings. For both definitions of earnings, we keep all individuals in the top 1 percent of the distribution, and we take a 2% random sample of individuals in the bottom 99 percent. For brevity, we will hereafter abbreviate 0.1 percent as 0.1pct, and similarly for other percentiles.

In Section 7, we also analyze 30-year average earnings, which we refer to as lifetime earnings. For the analysis in that section, we restrict attention to cohorts of individuals from ages 25 to 54 and include individuals in the sample if they satisfy the two selection criteria above for a minimum of 15 years during that 30-year period. In Appendix F, we also report results for cohorts of individuals from ages 30 to 59.<sup>9</sup>

### 3 Trends in Top Earnings

In this section, we study trends in top earners from three related angles. In Section 3.1 we analyze trends for top earners in the overall earnings distribution, without distinguishing by gender. Then, in Section 3.2, we turn to gender-specific earnings distributions and define top earning men and women relative to their ranking in the distribution of workers of the same gender. Finally, in Section 3.3, we return to top earners in the overall distribution and analyze the gender composition of this group and how this composition has changed over time.

#### 3.1 Top Earners in the Overall Earnings Distribution

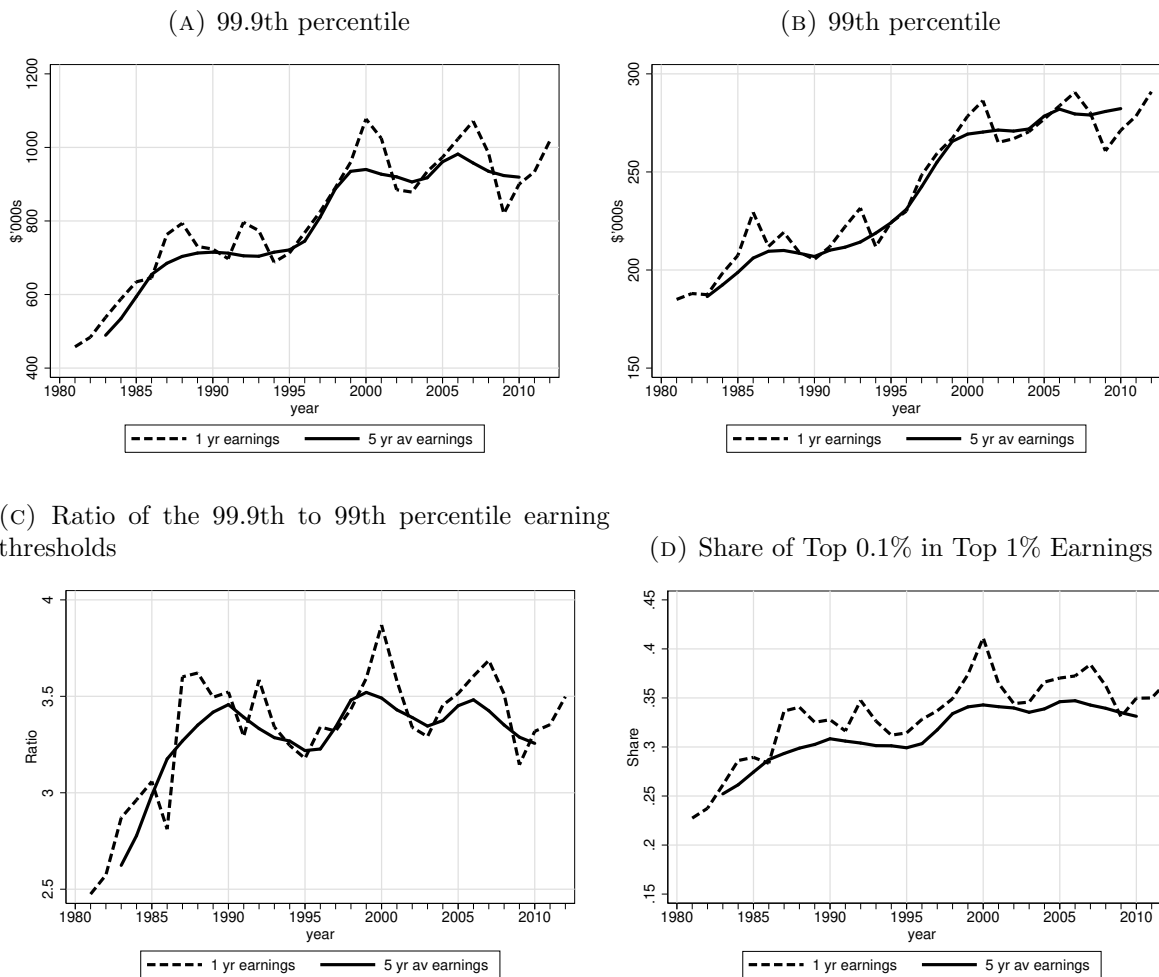
In 2012, a worker had to earn at least \$1,018,000 to be included in the top 0.1pct of the overall earnings distribution and at least \$291,000 to be included in the top 1pct. During the last five years of our sample (2008–2012), the analogous thresholds for being included in the top 0.1pct and 1pct based on five-year average earnings were \$918,000 and \$282,000 respectively.<sup>10</sup> These thresholds are only five percent to ten percent lower than the annual thresholds, which suggests that top earnings are quite persistent. As we will see, this

---

<sup>9</sup>To avoid possible privacy issues, we do not report any statistics for demographic cells (for example, a given industry/gender/year/income group) with fewer than 30 individuals. Thanks to the large sample size, such cells are rarely encountered.

<sup>10</sup>For comparison, mean (median) annual earnings in our data were \$51,000 (\$35,000) in 2012, and mean (median) five-year average earnings were \$53,000 (\$38,000), which illustrates the well-known vast gap between top earners and the average worker.

FIGURE 1 – Top Earnings Thresholds



persistence is a recurring theme in our findings.<sup>11</sup>

The top two panels of Figure 1 show how these top-earning thresholds have changed over our sample period. We emphasize four points. First, the thresholds have risen substantially, reflecting the well-documented rise in top earnings. Second, the rise has not been in the form of a secular trend but was more episodic, with two large bursts (from 1981 to 87 and from 1994 to 2000) interrupted by two periods when the thresholds were flat. The lack of an

<sup>11</sup>As explained in Section 2, our earnings data comprise only wage and salary income reported on W-2 forms. According to Statistics of Income (SOI) data from the Internal Revenue Service (IRS), wage and salary income in 2011 accounted for 45.6% of total income (excluding capital gains) for the top 0.1pct of taxpaying units, 62.3% for the second 0.4pct, and 77.0% for the next 0.5pct. The next biggest component of income is entrepreneurial income, which consists of profits from S corporations, partnerships, and sole proprietorships (Schedule C income). In 2011, this accounted for 28.6% of income for the top 0.1pct of tax units, 28.4% for the second 0.4pct, and 16.7% for the next 0.5pct.

upward trend after the turn of 21st century is especially noteworthy against the general perception that top earnings have been continuing to rise at a very fast pace. Third, Figure 1c, which plots the ratio of thresholds, shows that the thresholds have evolved almost in parallel fashion since 1987, suggesting that inequality *within* the top 1% has been largely stable in recent decades. This is also evident in Figure 1d which shows that the share of total earnings of the top 1pct earned by top 0.1pct has risen only slightly during this period.<sup>12</sup> Fourth, the thresholds for five-year average earnings (solid black lines) are not only smoother than the annual thresholds, but are also only slightly lower, reflecting the persistence of top earnings.<sup>13</sup>

These findings are not sensitive to the focusing on top-earnings percentile thresholds. In Appendix B we report the trends in the share of total earnings accruing to workers in various top percentiles (Figure B.1a), and trends in average earnings within each percentile group (Figures B.1b to B.1d). These figures confirm the episodic nature of the rise in top earnings, the tapering off in the rise post-2000, and the parallel trends in the top 0.1 pct and second 0.9pct.

Although the timing of earnings growth over this period was similar for other income groups, in particular the surge in earnings in the late 1990s with little-to-no growth after, the magnitude of this growth was much larger for top earners than the rest of the distribution. For example, focusing on five-year averages, the growth in average earnings from the 1981–85 period to the 2008–12 period was 139% for the top 0.1pct (Figure B.1b), 63% for the top 1pct (Figure B.1c), and only 22% for the bottom 99 percent (Figure B.1d).

### 3.2 Top Earners in Gender-Specific Earnings Distributions

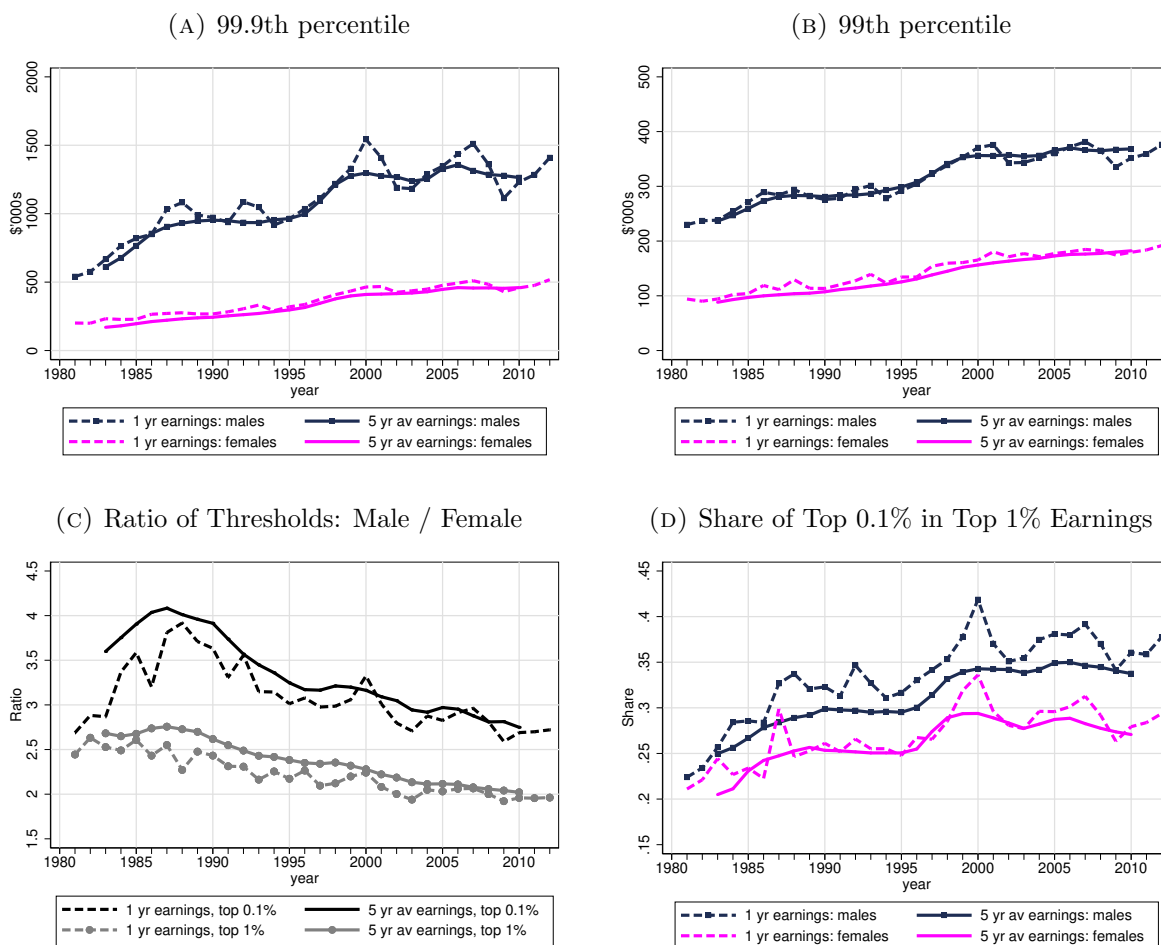
How different are these trends for top earning men and top earning women? To answer this question we split the overall sample by gender and define top earners of each gender relative to the gender-specific earnings distribution. Figure 2 shows the thresholds for membership of the top percentiles of these gender-specific earnings distributions. In 2012, men had

---

<sup>12</sup>For annual earnings there was an isolated peak in 2000, most likely due to payouts related to the information technology boom pushing up earnings at the very top of the distribution. Consistent with this hypothesis, the 2000 peak in annual earnings for the 99.9th percentile is particularly prominent in the engineering sector (which, according to our definition, includes technology companies; see Section 5) and is much less prominent in other sectors.

<sup>13</sup>The lack of a continued increase in top-earnings thresholds post-2000 is not specific to the particular measure of income (wage and salary earnings) or sample we use. In Guvenen and Kaplan (2017), we use data from aggregate tax records and show that the same tapering off happens for various measures of income, including capital and private business income. The only exception are the incomes above the top 0.01% threshold, which shows an upward trend driven almost entirely by private business income.

FIGURE 2 – Top Earnings Thresholds



to earn roughly twice as much as women in order to be included in the top 1pct of their respective gender-specific earnings distributions and nearly three times as much in order to be included in the top 0.1pct of their distributions.

Figure 2c shows the ratio of the top earnings threshold for men to the top earnings threshold for women. For five-year average earnings, this ratio for the top 0.1pct peaked in the late 1980s at around 4.1 and has declined monotonically since then to reach a level of 2.75 for the period 2008-12. This means that whereas two decades ago, a man at the 99.9th percentile of the male distribution earned over four times as much as a woman at the same percentile of the female distribution, today such a man earns less than three times as much as such a woman.

Although the gender differences in top earnings thresholds have narrowed in recent years, the gap between the average earnings of top male earners and top female earners has actually

widened. This can be seen in Figures B.2a in Appendix B, where we plot average earnings for the top 0.1pct of men and the top 0.1pct of women, and in Figure B.2b, where we plot average earnings for the second 0.9pct of men and the second 0.9pct of women. These two seemingly contradictory views of trends in the gap between the top ends of the gender-specific distributions – thresholds versus average earnings – can be reconciled by observing that inequality within the top 1pct, as measured by the earnings share of the top 0.1pct in the top 1pct, is higher for men than for women and has remained relatively constant since the late 1990s (Figure 2d).

### 3.3 Gender Composition at the Top: Cracks in the Glass Ceiling?

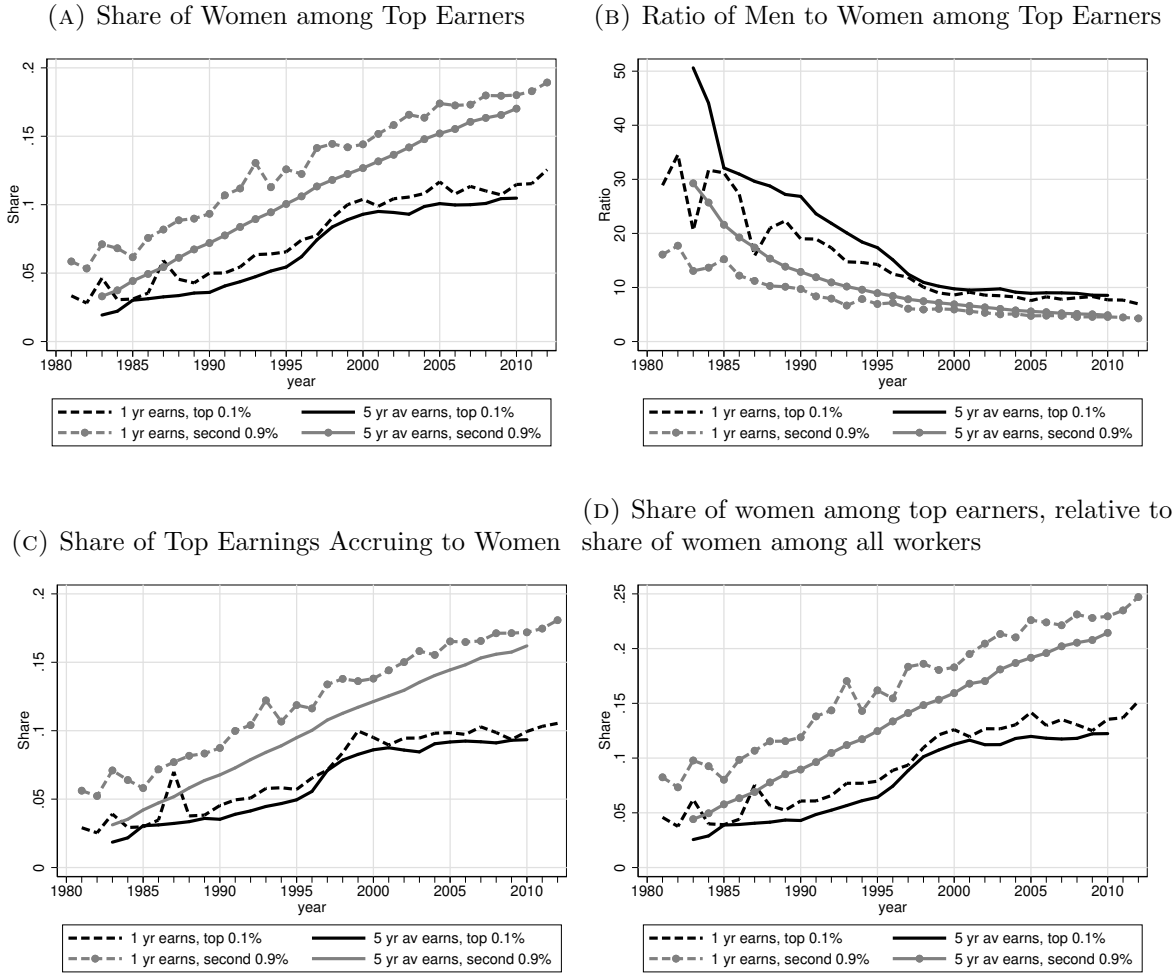
We now return to top earners in the overall earnings distribution and analyze the gender composition of this group. Our findings are displayed in Figure 3. The top left panel (Figure 3a) shows that the share of women among top earners has increased substantially since the early 1980s. For example, during the 1981–85 period, women constituted just 1.9% of the top 0.1pct group and just 3.3% of the second 0.9pct group based on 5-year average earnings (the solid lines). By 2008–12, the corresponding shares of women had risen to 10.5% and 17.0%, respectively.

The magnitude of this change is even more striking when expressed in terms of the number of men for every woman in the top percentiles, shown in the top right panel (Figure 3b). During the 1981–85 period, there were 50.6 men for every one woman in the top 0.1pct group, whereas in 2008–12 this number had fallen to 8.5 men for every woman. A similar decline happened for the second 0.9pct of earners, with the number of men per woman falling from 29.3 to 4.9 during the same period.

The rising fraction of women among top earners has also translated into a corresponding rise in the share of top earnings that accrues to women, shown in the bottom left panel (Figure 3c). In fact, the share of earnings has risen almost as rapidly as the rise in the population share of women in these groups, suggesting that the women who have entered the top percentiles are not disproportionately concentrated toward the bottom of the top earner groups.

When interpreting these trends, it is also important to consider that the gender composition of the overall labor force shifted toward women during this time, due in to the rise in female labor force participation, which raised the female employment share in the lower 99 percent of the earnings distribution from 44% to 49.2% from the 1981–85 period to the 2008–12

FIGURE 3 – Gender Composition of Top Earners



period (see Figure B.6 in B for the full time series). This means that part of the trends in Figures 3a and 3b might be due to this broader trend. But comparing the share of women among top earners with the share of women among *all* workers, suggests that this effect is small (Figure 3d).<sup>14</sup> The time-series for the share of women among top earners is almost unchanged by adjusting for the increase in the share of women among all workers.

This conclusion is confirmed by a formal decomposition of the change in the share of women in top percentiles into a component that is due to the changing gender composition of the overall labor force and a component that is due to the changing gender composition of top percentiles beyond the change in the overall distribution. The equations underlying the decomposition are contained in Appendix A. The results of the decomposition (Table 1)

<sup>14</sup>We define individuals to be working if they satisfy the age and minimum earnings criteria described in Section 2.

TABLE 1 – Decomposition of change in share of women among top earners

	<u>Annual earnings</u>		<u>Five-year earnings</u>	
	Top 0.1%	Second 0.9%	Top 0.1%	Second 0.9%
Total change in share	0.09	0.13	0.09	0.14
Fraction due to:				
– gender comp. of labor force	7%	9%	7%	7%
– gender comp. of top percentiles	93%	91%	93%	93%

Notes: Change for annual earnings is from 1981 to 2012. Change for five-year earnings is from the period 1981–85 to the period 2008–12. See Appendix A for details of decomposition.

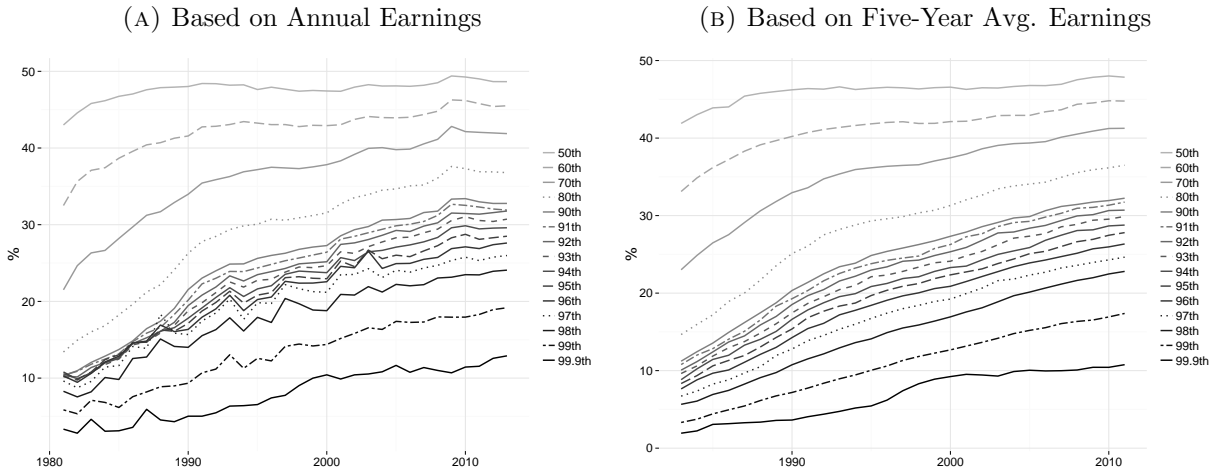
imply that only 7% to 9% of the total increase is due to changes in the overall female share of workers.

So far, the description of our empirical findings has painted a glass-half-full picture: women have made substantial inroads toward gender equality at the top. Today a working female is over four times more likely to be in the top 0.1pct of the earnings distribution than a working female was three decades ago (Figure 3d). Yet, with the same data, it is also easy to paint a glass-half-empty picture of these trends: despite this dramatic transformation, women are still vastly underrepresented at the top. There has been almost no increase in the share of women among the top 0.1pct of earners in the first decade of the 21st century (Figure 3a). Even in 2012, a working woman was only 12.2% as likely to be in the top 0.1pct as a working man was (Figure 3d), and the shares of women in the top percentiles were below 15% for the top 0.1pct, and below 20% for the second 0.9pct (Figure 3a).

### 3.4 Changing Gender Composition Outside of Top 1 Percent

We have so far focused on the gender composition inside the top 1pct group, but how has the gender composition changed for other high earnings percentiles outside of the top 1 percent? Figure 4 shows the time series of the share of women in selected percentiles above the median of the overall earnings distribution. It is clear from this figure that the share of women has increased across the entire upper half of the earnings distribution, and especially strongly inside the top 10 pct. This is true not only for annual earnings (left panel) but also for five-year earnings (right panel).

FIGURE 4 – Changing Gender Composition in Various Earnings Percentiles

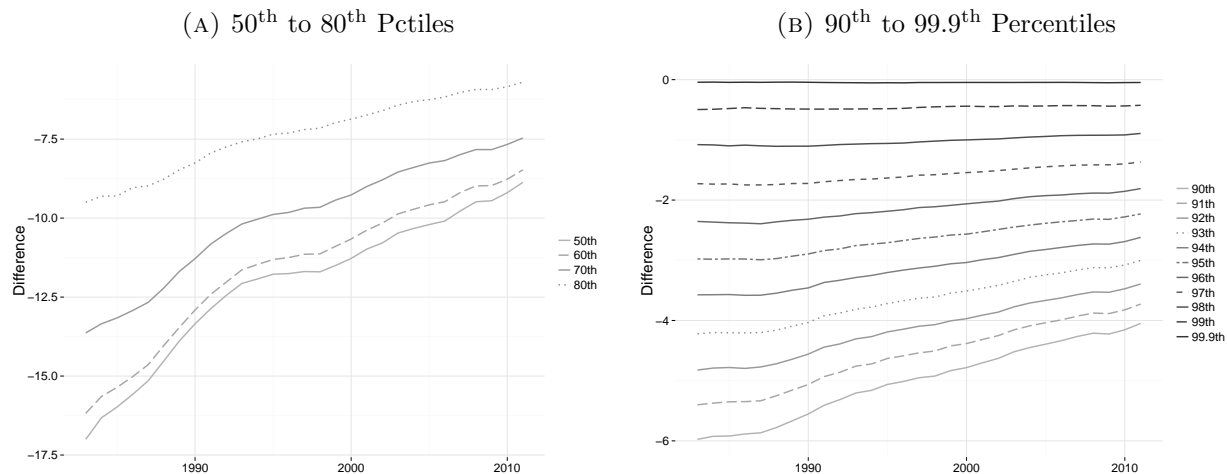


A related perspective on the convergence in labor market outcomes is to look at where female workers at a certain percentile of the female earnings distribution rank in the male earnings distribution. To answer this question, Figure 5 shows the difference between the two gender-specific five-year distributions (male's minus female's) for select percentiles above the median. The left panel shows the four deciles from the median to the 90th percentile, and the right panel shows the percentiles within the top 10 percent. So for example, the dotted line in the left figure shows that women in the 80th-90th percentiles (9th decile) of the female distribution in the early 1980s would have been around 9 percentiles lower in the male distribution, whereas they were only around 5 percentiles lower in the male distribution by 2011.

Similarly, the bottom panel in the right figure shows that women in the 90th-91st percentile of the female distribution would have been around 6 percentiles lower in the male distribution in the early 1980s (i.e. in the 84th percentile), and 4 percentiles lower in the male distribution by 2011 (i.e. in the 86th percentile). A similar pattern emerges for annual earnings distributions (see Figure B.3 in Appendix B.3.) Both Figure 4 and Figure 5 suggest that the convergence between genders was pervasive across the earnings distribution and not confined to the very top earnings groups only.



FIGURE 5 – Females in Male Earnings Distribution: Difference from Male Percentiles, 5-Year Earnings



## 4 A Paper Floor? Gender Differences in the Likelihood of Staying at the Top

In any given year, the members of a given earnings percentile are composed of newcomers (those who moved in since last year) and stayers (those who were in the that percentile in the previous year). Hence, to understand the changes in the gender composition of top earners, it is important to understand how mobility patterns differ for men and women and how these patterns have changed over time. The persistence of top earner status is also relevant for a range of other economic questions, including the determinants of wealth concentration at the top, the earnings risk faced by top earners, and the optimal taxation of their labor earnings. The scarcity of representative panel datasets covering top earners—which is necessary to measure earnings mobility—has largely prevented the analysis of these dynamics in the existing literature.<sup>15</sup> In this section, we fill this gap. We start by examining the mobility of overall top earners and we then examine gender differences in persistence and how these differences contribute to the trends shown in Section 3.

<sup>15</sup>Two notable exceptions are [Kopczuk et al. \(2010\)](#) and [Auten et al. \(2013\)](#) who also document transition rates among top percentiles. However, neither of these papers studies gender differences in mobility nor mobility within the top 1pct.

TABLE 2 – Transition Probabilities across Top Earnings Groups, Post-2000s

Panel A: Annual Earnings, One-Year Transitions				
	Top 0.1%	Second 0.9%	Bottom 99%	Exit Sample
Top 0.1%	0.57	0.31	0.07	0.05
Next 0.9%	0.04	0.65	0.27	0.04
Second 99%	< 0.01	< 0.01	0.91	0.08
Panel B: Five-Year Earnings, Five-Year Transitions				
	Top 0.1%	Second 0.9%	Bottom 99%	Exit Sample
Top 0.1%	0.40 (0.59)	0.22 (0.32)	0.06 (0.09)	0.32
Second 0.9%	0.05 (0.07)	0.46 (0.61)	0.24 (0.32)	0.25
Bottom 99%	< 0.01 (< 0.01)	< 0.01 (< 0.01)	0.72 (0.96)	0.27

Notes: One-year transition probabilities refer to the period 2011–12. Five-year transition probabilities refer to the period 2003–7 to the period 2008–12. In Panel B, numbers in parentheses are transition rates conditional on remaining in the sample (i.e., normalized by one minus exit rate).

## 4.1 Mobility of Overall Top Earners

We begin with a broad measure of mobility based on the transition probabilities in and out of three earnings groups—top 0.1pct, second 0.9pct and bottom 99 percent—over one-year and five-year periods. Below we will also analyze transition between groups outside of the top 1pct. We first analyze the mobility patterns in the most recent time period and then turn to how these patterns have changed since the early 1980s.

Table 2 reports the transition matrices for the most recent periods covered by our data. For annual earnings, these are one-year transition probabilities between 2011 and 2012, and for five-year average earnings these are transition probabilities between the 2003–7 period and the 2008–12 period.

The first question we are interested in is the mobility of top earners—the rates at which they enter, stay in, or exit a given top earner group. The top panel of Table 2 reports the annual transition rates, which reveal substantial year-to-year mobility and suggest that top earnings status is far from a permanent state. For example, of all the workers in the top 0.1pct in 2011 only 57% were still in the top 0.1pct one year later, 31% had dropped to the

second 0.9pct and 7% had dropped out of the top 1pct altogether (Panel A of Table 2). In addition, 5% of workers left the sample, either through aging (turning 61) or by failing to meet the minimum earnings criteria. For workers in the second 0.9pct in 2010, 69% were still in the top 1pct (of which 4% had moved up to the top 0.1pct) and 27% had dropped down to the bottom 99 percent.

The transition rates for five-year earnings are very similar to those for annual earnings once the higher exit rate is accounted for. The probability of exiting through aging is higher because of the five-year horizon and the fact that the top earners' age distribution skews older for five-year earnings than for annual earnings. In Panel B, the numbers in parentheses report the transition rates conditional on remaining in the sample (i.e., normalized by one minus the exit rate). Comparing these numbers to the annual rates in Panel A makes the similarities clear. For example, the probabilities of staying in the top two earnings groups after 5 years were 59% and 61%, respectively (versus 57% and 61%, respectively, in Panel A). Thus, both annual and five-year transitions reveal a significant amount of turnover at the top. An important corollary of this turnover is that drawing conclusions about top earners from cross-sectional data is fraught with danger since one-third of the individuals in these groups are different from one year to another. We return to this point in Section 7, where we study top earners over 30-year periods.

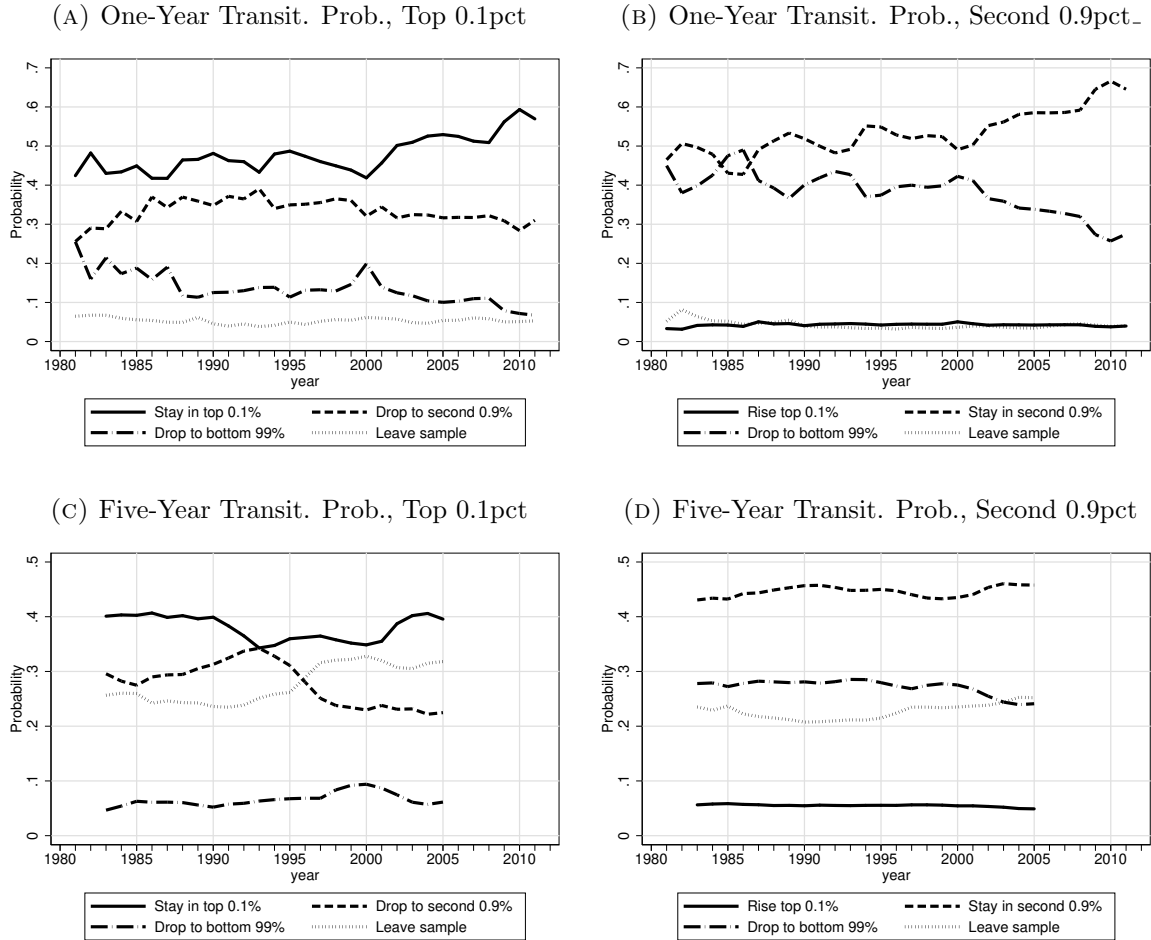
Before moving forward, a cautionary remark is in order. Although it might seem plausible at first blush, this large turnover at the top does *not* imply that the earnings of top earners display a lot of mean reversion. Nor is there a straightforward mapping between the transition probabilities reported in Table 2 and the persistence parameter of a first-order autoregressive (AR(1)) earnings process.<sup>16</sup>

We next turn to the evolution of these mobility patterns over time. Figure 6 plots the time series of the transition probabilities between the same groups as in Table 2. The overall pattern we see here is that mobility was higher—alternatively, the top earner status a less

---

<sup>16</sup>There are three reasons for this. First, holding the variance of earnings fluctuations fixed, the transition probabilities discussed here depend on the size of the earnings group we define (e.g., a 0.1pct group versus a 10-percentile wide group). The smaller the group, the higher the mobility in and out. Second, an empirically plausible earnings process also features permanent differences across workers, as well as transitory shocks. The transition rate is a function of the variances of these different disturbances. So, for example, simulating an AR(1) process for earnings with a persistence parameter of  $\rho = 0.99$  and Gaussian shocks generates annual and five-year transition rates for the top 0.1pct of around 0.7 and 0.5, respectively. However, we can generate the same transition rates by adding a fixed effect whose unconditional variance matches that of the AR(1) process and reducing  $\rho$  to 0.85. Third, and less obviously, the right tail of the earnings distribution is closer to Pareto than Gaussian. This implies that the gap between two percentiles of *log* earnings at the top end widens as we move up in the distribution. So if the variance of shocks to log earnings is the same for all workers, it will move fewer workers down to lower percentiles by virtue of this widening gap between earnings levels.

FIGURE 6 – Transition Probabilities In and Out of Top Percentiles, All Sample



Notes: These figures show the probability that a top earner based on average earnings over the period  $t - 2, \dots, t + 2$  is a top earner based on average earnings over the period  $t + 3, \dots, t + 7$ .

stable state—in previous decades. For example, during the 1980s and 1990s, the annual probability of staying in the the top 0.1pct group was fairly stable at about 45%. Since 2000, this probability has steadily risen, reaching over 57% in 2011 (Figure 6a).

The pattern is similar for the second 0.9pct group, with a fairly stable probability of staying put of around 50% during the first two decades, a 5% probability of moving upward into the top 0.1pct and a 40% chance of falling out of the top 1pct group. Since then, the probability of staying has risen to nearly 70%, mostly accounted for by a large reduction in the probability of moving down to the bottom 99pct group, from 40% to under 30% (Figure 6b). Transition probabilities for five-year average earnings over five-year horizons, which are displayed in Figures 6c and 6d, show similar qualitative trends to the one-year transition probabilities based on one-year earnings over the second half of the sample, but

the magnitude of the changes is smaller. In Figure B.7 in Appendix B, we report analogous figures for the time path of transition rates conditional on not leaving the sample, for which the trends in one-year and five-year transition probabilities are even more similar.

The rising stability of top earner status after the turn of the 21st century sheds light on the leveling off of the annual top earning thresholds and annual top earner shares during this period, shown in Figures 1 and B.1. Although the share of earnings accruing to the 0.1pct and top 1pct groups has not shown an upward trend during this period, the declining mobility at the top implies that top earners are slowly being entrenched, since the group shows less turnover than before. This observation highlights the benefits of studying top earners through the lens of individual panel data.

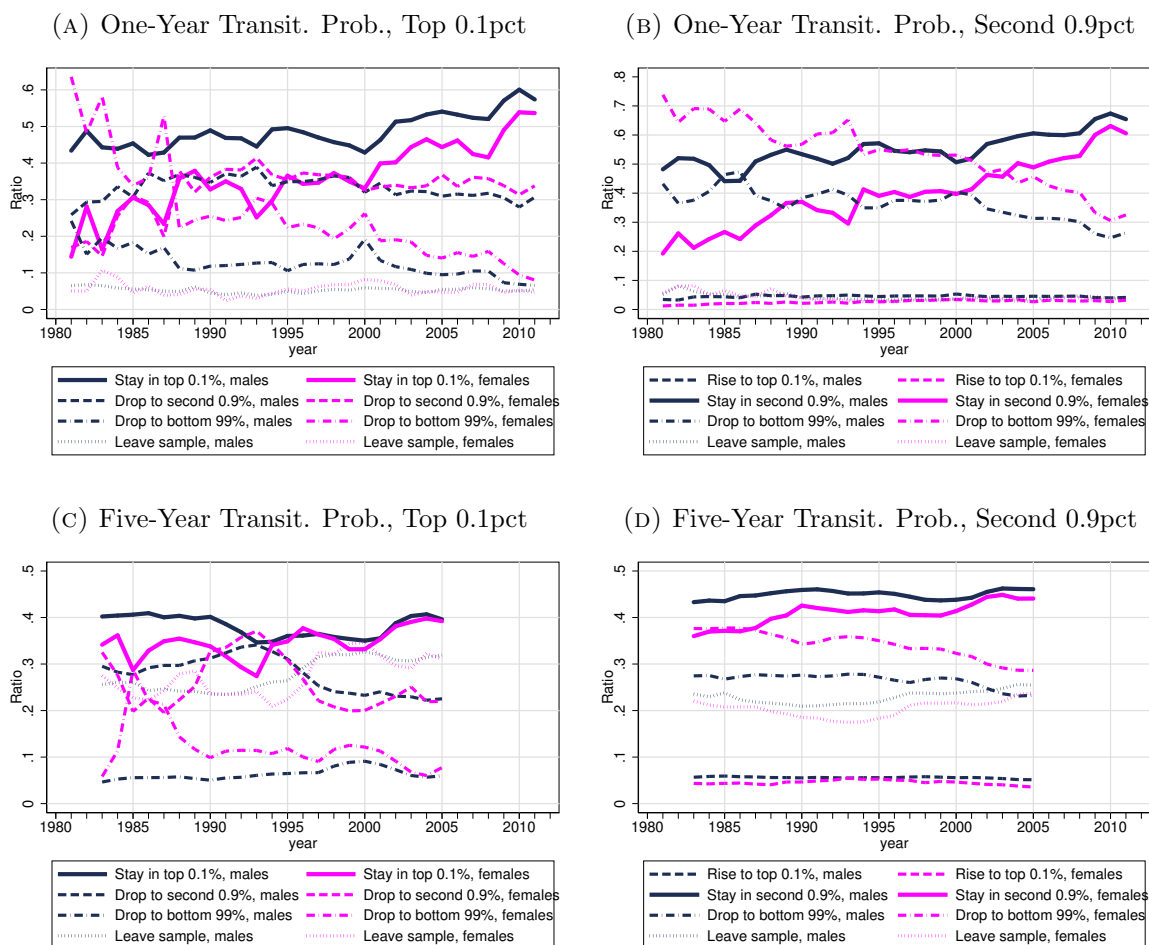
## 4.2 Gender Differences in Mobility

Behind these mobility patterns for top earners in the overall population are important differences by gender. In the early 1980s, there was a distinctive *paper floor* for top earning women, by which we mean that they faced a very high probability of dropping from either of the two top earning groups to the bottom 99 percent from one year to the next. For example, in 1981, this probability was 64% and 74% respectively, for women in the top 0.1pct and second 0.9pct groups. By comparison, these probabilities were much lower for men: 24% and 43%, respectively, for the same two groups. This is the essence of the paper floor: not only were women vastly underrepresented among top earners in a given year, but even those who did have high earnings were much more likely than men to drop out of the top earnings groups within a year.

However, the last three decades have seen a steady mending of the paper floor. This mending can be seen clearly in Figure 7, which shows the time path of transition probabilities separately for men and women in the top percentiles of the overall earnings distribution. The overall picture that emerges from the four panels in this figure is that the gender gap in persistence has almost disappeared. For example, in 2011, the annual probability of dropping from the top 0.1pct to the bottom 99 percent 8.1% for women compared with 6.6% for men; Similarly, the analogous probability of dropping down from the second 0.9pct was 32% for women compared with 26% for men. During the same time frame, there have been similar improvements in upward mobility. For example, the transition rate of women from the second 0.9pct up to the top 0.1pct has more than doubled, from 1.2% in 1981 to 3.2% in 2011, but has increased much less for men, from 3.4% to 4.1%.

One potential explanation for the mending of the paper floor is that top earning women may have become more evenly distributed within the top 0.1pct—rather than being bunched

FIGURE 7 – Transition Probabilities In and Out of Top Percentiles Over Time, by Gender



Notes: These figures show the probability that a top earner based on average earnings over the period  $t - 2, \dots, t + 2$  is a top earner based on average earnings over the period  $t + 3, \dots, t + 7$ , separately for male top earners (blue) and female top earners (pink).

just above the top 0.1pct threshold. If true, then mean reversion in earnings would be pushing fewer of them below the top 0.1pct threshold, increasing the persistence of top earnings status for women. However, two pieces of evidence suggest that this was not a major factor. First, the persistence of top earner status has also risen significantly when we define it relative to the gender-specific earnings distributions (see Appendix C). Persistence defined in this way has risen significantly for women but not for men, both at annual and five-year horizon. Second, the ratio of the earnings share of women in the top 0.1pct to the population share of women in top 0.1pct has barely changed over this period (see Figure 3c), suggesting that there has not been a dramatic change in the average position of women within the top 0.1pct. In fact, for both annual and five-year average earnings, the gender gap within the top 0.1pct (as measured by the ratio of average earnings of women to men)

TABLE 3 – Decomposition of Change in Share of Women among Top Earners

	Annual earnings	
	Top 0.1pct	Second 0.9pct
Change in female share	0.097	0.136
Fraction due to:		
Existing differences and persistence	56.5%	42.4%
Change in transition probabilities	43.5%	57.6%
due to change in inflow	10.2%	16.3%
due to change in outflow	33.3%	41.3%

Notes: The change in female share is for the period 1982–2012 rather than 1981–2012 because the decomposition requires an initial period to compute the initial transition probabilities. See Appendix A for details.

has been flat over the last 30 years, whereas the same ratio calculated within the second 0.9pct group has declined.

The dramatic increase in the persistence of female top earners has been an important factor in accounting for the rise in the share of women among top earners. To understand the contribution of changes in transition rates, we decompose the change in the gender composition of each top earning group into a component that is due to *different trends* in the transition probabilities in and out of the top percentiles for men versus women, and a component that is due to *pre-existing* differences in the same transition probabilities. We describe our procedure for implementing this decomposition in Appendix A. The former component measures the contribution of changes in persistence to the overall change in gender composition, whereas the latter component measures the change in gender composition that would have taken place absent any changes in the transition probabilities over this period.<sup>17</sup>

The decomposition, which is reported in Table 3, shows that 33% of the increase in the share of women among the top 0.1pct, and 41% of the increase among the second 0.9pct, is due to the fact that women are now less likely to drop out of the top percentiles than they were in the past and so receive high earnings for longer periods of time. The remainder

<sup>17</sup>Conceptually, the fraction of women in the top percentiles can change even if the transition matrix stayed constant, simply because of an earlier change in the transition matrix and the fact that it takes time for the implied Markov process to reach its new stationary distribution. Additionally, the fraction of women can change because of further changes in the transition matrix relative to the transition matrix for men. We perform this decomposition only for one-year transition probabilities using annual earnings, because the overlapping nature of the five-year analysis makes an analogous decomposition for five-year earnings difficult.

of the increase is due to pre-existing differences in the fraction of men and women in the top percentiles, and changes in the probability of new women entering the top earning percentiles.

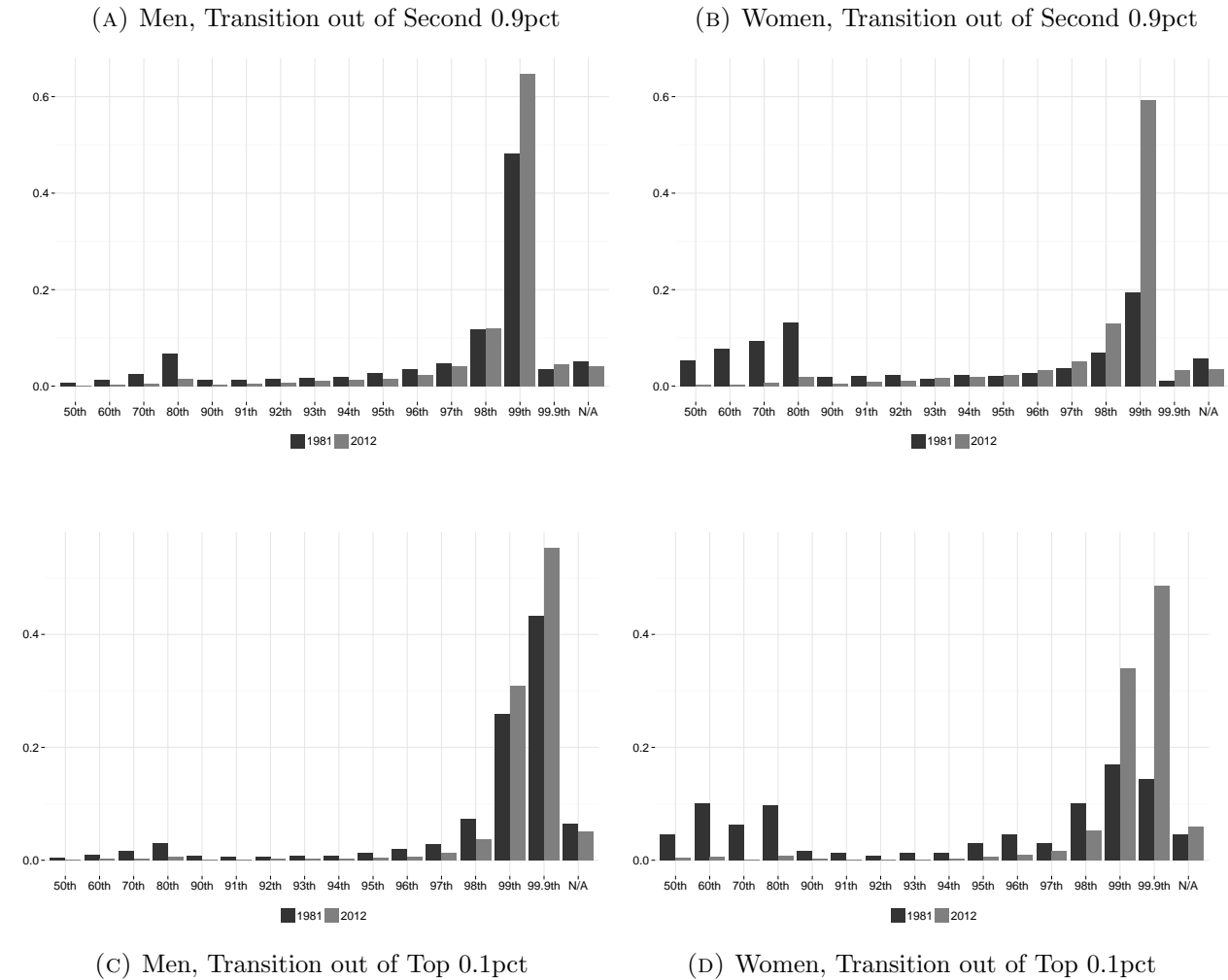
So far we have focused only on transition rates out of the top 1pct into the bottom 99pct, but it is also useful to know where in the bottom 99pct those workers dropping out of the top 1pct are actually dropping to. Figure 8 offers an answers to this question. The top two panels show annual transition rates out of the second 0.9pct group into each of the four deciles from the median to the 90th percentile, and into each percentile within the top decile, in 1981 and 2012 for men (left panel, Figure 8a) and women (right panel, Figure 8b). The figures illustrate clearly the large increase in the probability of staying in the second 0.9pct from one year to the next, with a much bigger increase in this probability for women so that the probability of staying in this two groups has nearly converged to the probability of staying for men.

However, for both genders this increase in staying in the top percentile groups is not due to a decline in transition rates to nearby percentiles, but rather is a due to a decline in transition rates to percentile groups much lower down the distribution. This finding is particularly stark for women. In 1981, had only had a 20% chance of staying in the second 0.9pct and those who fell out of the top 1% had an almost 40% chance of falling out of the top decile, and more than a 25% chance of dropping of the top 20pct. Thus, for women in 1981, dropping out of the top 1pct meant a very large drop in earnings. This has changed dramatically. By 2012, the probability of staying in the second 0.9pct had increased from about 20% to 60% (as seen above in Figure 7b), and importantly, this increase came as a result of a decline in the probability of large falls in earnings. For example, the 40% chance of falling out of the top decile declined to about 5% and the 25% chance of dropping out of the top 20pct declined to less than 2%.

The bottom two panels of Figure 8 show that a similarly dramatic change has taken place for transition probabilities out of the top 0.1pct. In Figure B.9 in Appendix B, we report analogous five-year transition rates, based on five-year earnings. The changes are smaller, but the conclusion that the increase in the staying probability for women is due primarily to a decline in transitions to lower parts of the distribution remains true. This analysis of finer transition rates suggests that the mending of the paper floor is a more robust phenomenon than one could infer from looking at only overall transition rates out of the top percentiles.



FIGURE 8 – Changes in Annual Transition Probabilities out of Top Percentiles into Finer Percentile Groups, by Gender



## 5 Where Is the Glass Thinner? Industry Composition of Top Earners

The trends in the gender composition of top earners, as well as the changes in the mobility of women in and out of the top percentiles that have partly fueled these trends, may in part be due to differential changes in the observable characteristics of male workers versus female workers over this period. In this section and the next, we examine gender differences in two potentially important characteristics that are observed in our data: the industry in which individuals work and the individual’s age. Our goal is to ascertain whether there are certain industries in which women have made greater inroads into the top percentiles and how much of the increased female share of top earners is due to an increased presence

TABLE 4 – Aggregating industries

Aggregated Industry	Included SIC Codes
1 Engineering	7370-7379, 3570-3579 (computers) 8711 (engineering services)
2 Health services	80
3 Legal services	81
4 Management, Accounting, Business Consulting	3660-69, 8700, 8712-8729, 8741-8749
5 Other Services	7000–8999 except 737, 781-84, 80, 81, 87
6 Finance, Insurance	60, 61, 62, 63, 64, 66, 67
7 Wholesale trade	50, 51
8 Retail trade	52–59
9 Transportation, Communication	40, 41, 42, 43, 44, 45, 47, 48
10 Durable Manufacturing	24, 25, 30, 31, 32, 33, 34, 37, 38, 39 35 (except 357), 36 (except 366)
11 Nondurable Manufacturing	20, 21, 22, 23, 26, 27, 28
12 Construction and Real Estate	15, 16, 17, 65
13 Commodities and Mining	2911, 46, 49, 10, 11, 12, 13, 14

Notes: SIC codes 781-784 and 79 correspond to “Hollywood, Artists, and Professional Sportsmen.” We include workers in this category as part of “Other Services” in order to avoid privacy issues.

of women in industries that have higher representation at the top of the distribution, as opposed to an increased share of women among the top earners within given industries.

To address these questions, we use the SIC code assigned to the EIN that is associated with each worker’s main source of earnings. Based on these SIC codes, we construct the 13 industry groups listed in Table 4. Our logic in combining SIC codes into industry groups in this way is to group together businesses in which top-earning workers are likely to perform similar tasks, despite their potentially disparate SIC codes. Typically, SIC codes are grouped based on 1-digit or 2-digit classifications. But such classifications are intended to group industries by the type of goods they produce, rather than by the type of work that their employees do. For example, the 1-digit SIC classification places a computer hardware company (such as Apple, Dell, or Hewlett-Packard) under Durable Manufacturing (SIC 357: Industrial machinery and equipment), while placing a computer software company (such as Google, Microsoft, or Oracle) under Business Services (SIC 737: Computer programming, data processing, and other data related services) and an engineering consulting company under Engineering, Accounting, Research Management, and Related Services (SIC 8711: Engineering services). Under our classification, workers at the businesses listed above are

all included as part of Engineering, top-earning workers at these firms likely have similar roles. Thus, our industry grouping should be interpreted as lying somewhere between an industry and an occupational classification, when compared with the typical SIC industry classification. Table 4 contains a full crosswalk between SIC codes and our 13 industry groups. In Appendix D we report the SIC codes of selected large U.S. companies.

We assign each individual to the industry that corresponds to the SIC code of their main employer in year  $t$  (i.e., the employer that contributes to the largest share of their annual earnings). For five-year average earnings, we define their industry as the SIC code of their main employer in the most recent year  $t + 2$ . To minimize the number of figures in the main text, in this section we only report results based on five-year average earnings. The analogous figures using annual earnings can be found in Appendix D, and they yield similar conclusions.

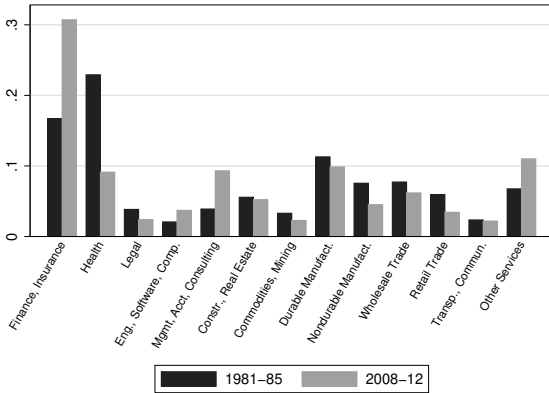
## 5.1 Industry Composition of All Top Earners

Finance and Insurance is by far the most highly represented industry among the highest earners. For the five-year period 2008–12, 31% of individuals in the top 0.1pct worked for employers in the Finance and Insurance industries, and these workers received 32% of the earnings of all individuals in the top 0.1pct. Among the second 0.9pct of workers, Health services is the most highly represented industry, in terms of both numbers of workers and share of earnings, with Finance and Insurance a close second. Together these two industries accounted for 33% of workers in the second 0.9pct in 2008–12 and accounted for 34% of the earnings of the second 0.9pct. The population shares and earning shares of each of the 13 industry groups among top earners in 2008–12 can be seen as the grey bars in Figure 9a and Figure 9c (top 0.1pct), and in Figure 9b and Figure 9d (second 0.9pct).

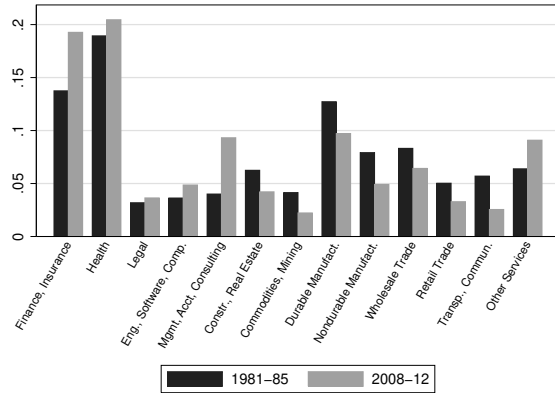
Interestingly, the industry share of top earners has not always looked his way. In the early 1980s, employers in the Health services industry represented a larger share of top earners than Finance and Insurance, in terms of both number of workers and total earnings. In addition, employers in Manufacturing, particularly those in Durable Manufacturing, had a very strong presence at the top of the earnings distribution. Hence, over the last three decades, the major change in the industry composition of top earners has been the rise in earnings in the Finance and Insurance industry, offset by a relative decline in the earnings of the highest paid doctors and, to a lesser extent, a relative decline for the highest earners employed by manufacturing firms. These changes can be seen in the panels of Figure 9 by comparing the solid black bars, which show the population and earnings shares of each industry group among top earners in 1981–85, with the grey bars, which show the

FIGURE 9 – Industry composition of top earners, five-year average earnings

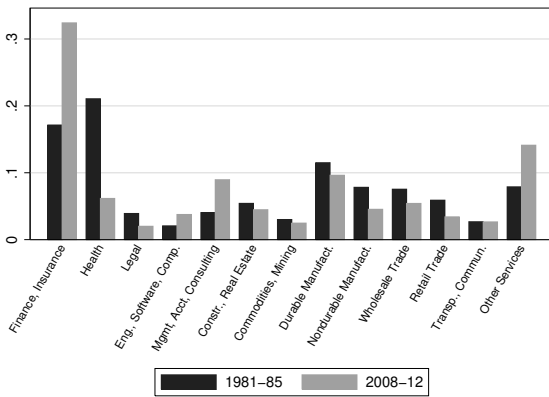
(A) Population shares, top 0.1pct



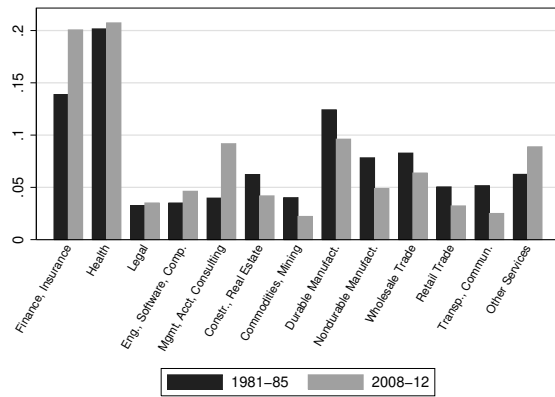
(B) Population shares, second 0.9pct



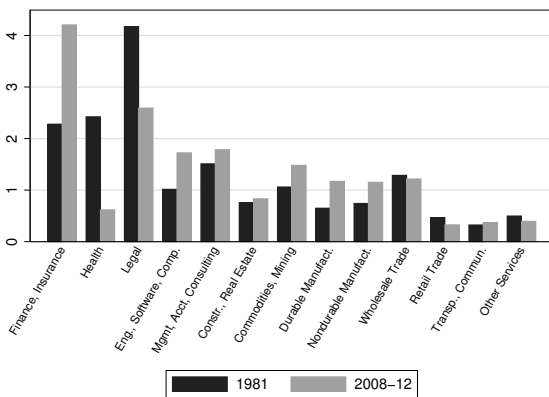
(C) Earnings shares, top 0.1pct



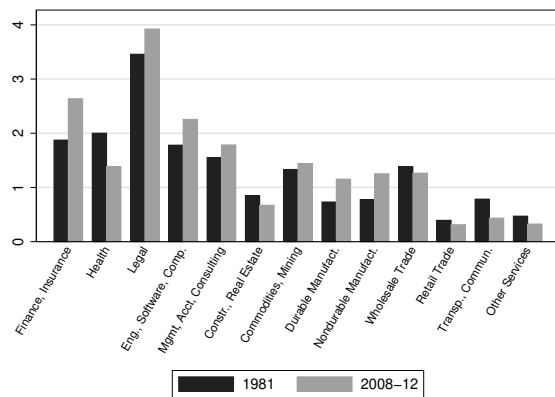
(D) Earnings shares, second 0.9pct



(E) Population shares, top 0.1pct relative to bottom 99 percent



(F) Population shares, second 0.9pct relative to bottom 99 percent



corresponding shares in 2008–12.

Finance and Insurance not only is the industry in which top earners are most likely to work, but also is the industry that is most heavily composed of top earners. For example, in 2008–12, a worker in the top 0.1pct of the earnings distribution was over four times as likely to be working in Finance and Insurance as a worker in the bottom 99 percent of the earnings distribution. This too was not always the case: in the early 1980s, a worker in the top 0.1pct was only around twice as likely to be working in Finance and Insurance as one in the bottom 99 percent. Instead, in the 1980s the industry with the highest relative likelihood of being in the top 0.1pct was Legal services, for which the ratio has dropped from 4.2 to around 2.6. These changes can be seen in Figure 9e, which shows how the share of each industry in the top 0.1pct relative to the share of that industry in the bottom 99 percent, has changed between the period 1981–85 and the period 2008–12. For the second 0.9pct, Legal services was, and still are, the industry with the highest representation relative to its representation in the bottom 99 percent (Figure 9f).

## 5.2 Gender Differences in Industry Composition

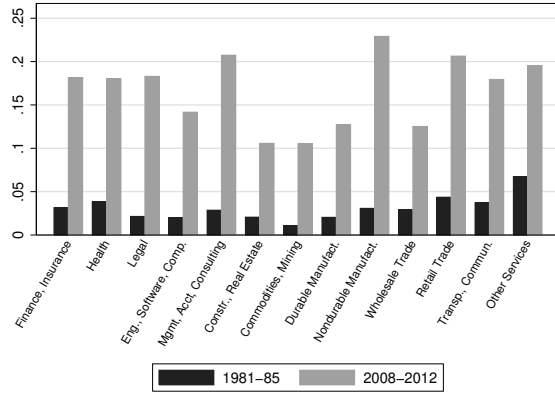
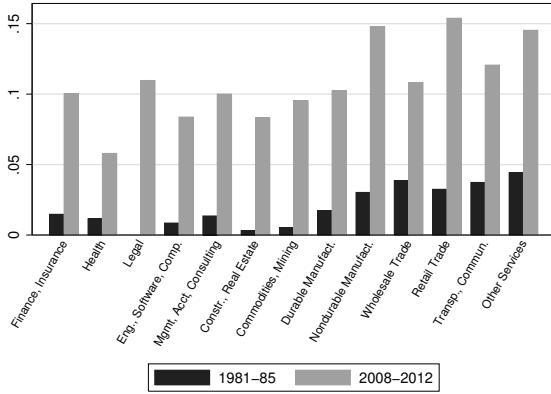
Surprisingly little variation can be seen across industries in the gender composition of overall top earners. In 2008–12, the share of women varied from 6% in Health services to just under 15% in Nondurable Manufacturing and Retail Trade for the top 0.1pct (Figure 10a), and from just over 10% in Construction and Real Estate to 24% in Nondurable Manufacturing for the second 0.9pct (Figure 10b). Thus, although some variation can be seen across industries, today there is no single industry, or subset of industries, in which top earning women are disproportionately absent. Thirty years ago, however, the share of women among top earning workers in Retail Trade and Other Services was substantially higher compared with other industries.

The similarity across industries in terms of the gender composition of top earners suggests that the large increase in the overall representation of women at the top of the earnings distribution is not due to women disproportionately moving into high earning industries like Finance and Health services. Moreover, the industry composition of top earners in the most recent five-year period 2008–12, shown in Figures 10c and 10d, is almost identical for men and women, suggesting that the remaining gender differences among top earners are not due to an underrepresentation of women at the top of any one industry, but rather are an across-the-board phenomenon.

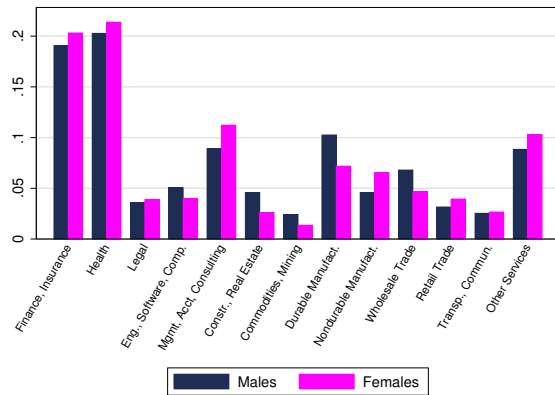
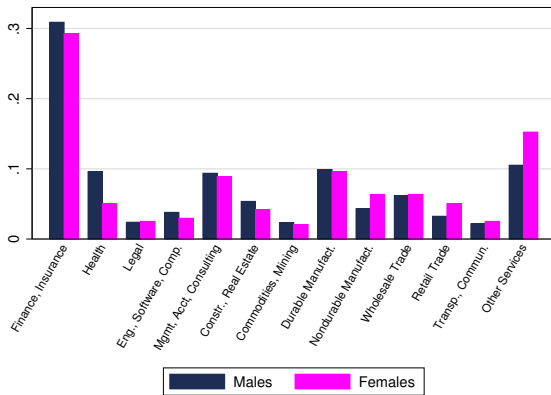
The conclusion that gender differences in industry shares do not play a role in understanding changes in the gender composition of top earners is confirmed by a formal decomposition.

FIGURE 10 – Top earners by industry and gender, five-year average earnings

(A) Share of women by industry within top 0.1pct cent (B) Share of women by industry within top 0.9pct



(C) Industry shares by gender within top 0.1pct, 2008–12 (D) Industry shares by gender within second 0.9pct, 2008–12



In Appendix A, we explain our procedure for decomposing the change in the gender composition of top earners into (i) a component that is due to changes in the industry composition of working women across the entire earnings distribution; (ii) a component that is due to changes in the industry composition of top earners of both genders; and (iii) a component that is due to changes in the gender composition of top earners within industries. The results of the decomposition, which are reported in Table 5, show that industry composition plays no role whatsoever in accounting for the increased representation of women at the top of the distribution. In fact, the contribution of the first two components is negative, suggesting that on average over this period there was a shift of the industry composition of working women toward industries that are slightly underrepresented in the top percentiles, and a shift of the industry composition of top earners toward industries with less female

TABLE 5 – Decomposition of change in share of women among top earners

	<u>Annual earnings</u>		<u>Five-year earnings</u>	
	Top 0.1%	Second 0.9%	Top 0.1%	Second 0.9%
Change in share	0.09	0.14	0.09	0.14
Fraction due to:				
- ind. comp. of women in labor force	-3%	-2%	-1%	0%
- ind. comp. of top earners	5%	5%	5%	6%
- female share of top earners within ind	98%	97%	96%	94%

Notes: Change for annual earnings is 1981-2012, and change for five-year earnings is 1983-2010 (centered five-year groups).

representation.

Although the last three decades have seen significant changes in the industry composition of top earners overall, these changes have been relatively similar for men and women and do not account for the changes in the gender structure of top earners over this period.

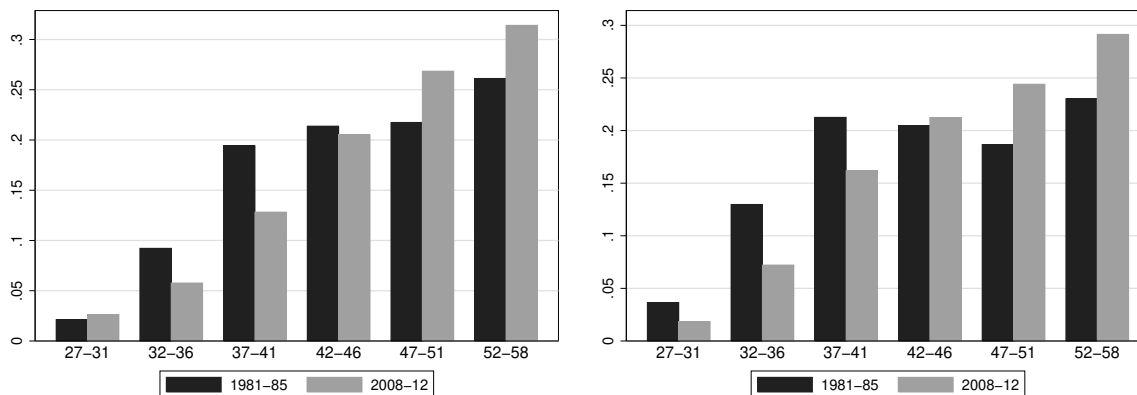
## 6 Looking Upward? Gender Differences in Top Earnings by Age

Since earnings growth at young ages is a key driver of earnings later in life, we can gain insight into possible future paths for gender differences among top earners by examining how top earnings gender gaps vary by age. Hence, in this section we divide the population of 25- to 60-year-olds into five-year age groups and study both the age distribution of top earners as well as the composition of top earners among individuals of a given age. Our goals are to ascertain whether top earning gender gaps already exist at the time of entry into the labor market or whether they emerge slowly as a cohort ages, and to measure how much of the increase in the share of women among top earners is due to shifts in the gender composition of recently entered cohorts versus changes that have occurred within older cohorts. To keep the presentation of the results manageable, in this section we only report our findings for five-year average earnings. The analogous figures for annual earnings lead to similar conclusions and are contained in Appendix E.<sup>18</sup>

<sup>18</sup>We analyze five-year earnings for five-year age groups. We denote each group by their age in the middle of the five-year period. So, for example, the 27-31 age group over the period 2008-12 refers to average earnings during this period for individuals who were ages 27-31 in 2010.

FIGURE 11 – Age distribution of workers, five-year average earnings

(A) Age distribution of individuals in top 0.1pct (B) Age distribution of individuals in second 0.9pct



Notes: Age labels in figures refer to center year of five-year age ranges. For example, 27-31 refers to the five-year age ranges 25-29, 26-30, ... 29-33.

## 6.1 Age and Gender Composition of Top Earners

Relative to the average age of the workforce, top earners are old and have become more so since the 1980s. For earnings over the five-year period 2008–12, 58% of the individuals in the top 0.1pct were ages 47 to 58 in 2010 (we measure an individual’s age in the middle of the five years used to construct average earnings), and 21% were ages 27 to 41. By contrast, for earnings over the five-year period 1981–85, only 48% of individuals in the top 0.1pct of the earnings distribution were ages 47 to 58 in 1983, and more than 31% were below 40. This aging of top earners can clearly be seen in Figure 11a, which shows the fraction of top earners in each five-year age bin for these two five-year periods. A similar pattern and trend is evident among the second 0.9pct of earners (Figure 11b), as well as for the share of top earnings that accrues to individuals of different age groups.

This large change in the age composition of top earners is a possible source of changes in the gender composition of top earners. Since top earners overall have become older, if female workers were on average initially older than male workers, then this would generate an increase in the share of women among top earners. In 1981–85, working women were indeed slightly older than men, although the difference is small. The aging of the labor force over this period was more pronounced for women than men, which could also generate an increase in the share of women among top earners (see Appendix E for figures illustrating these features of the data). Thus, a plausible conjecture is that part of the increased share of women among top earners is due to compositional effects related to the aging of the workforce.



TABLE 6 – Decomposition of change in share of women among top earners

	<u>Annual earnings</u>		<u>Five-year earnings</u>	
	Top 0.1%	Second 0.9%	Top 0.1%	Second 0.9%
Change in share	0.09	0.13	0.09	0.14
Fraction due to:				
- age comp. of top earners	10%	6%	11%	6%
- age comp. of women in labor force	6%	8%	6%	6%
- top earners within women given age	84%	86%	83%	88%

Notes: Change for annual earnings is 1981-12, and change for five-year earnings is 1983-10 (centered five-year groups).

Indeed, a formal decomposition confirms that around 12% to 17% of the increased share of women among top earners is due to age differences across men and women. In Appendix A, we explain our procedure for decomposing the change in the gender composition of top earners into (i) a component that is due to changes in the age composition of top earners relative to the age composition of the bottom 99 percent; (ii) a component that is due to differential changes in the age composition of men and women among workers in all parts of the earnings distribution; and (iii) a component that is due to changes in the fraction of women among top earners in a given age range. The results of the decomposition, which are reported in Table 6, indicate that 6% to 11% of the increase is due to the first component, 6% to 8% is due to the second component and 83% to 88% is due to the third component.

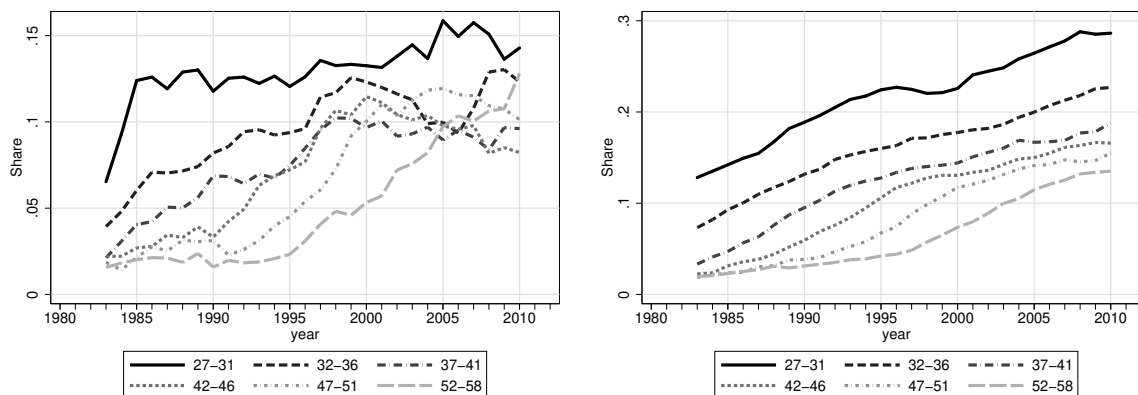
## 6.2 Gender Composition of Age-Specific Top Earners

Since there are so few young workers among the overall top earners, it is perhaps more informative to study the gender composition of top earners at each age, both for learning about the way in which top earnings gender gaps evolve over the working life and for making guesses at the future path of top earnings gender gaps.

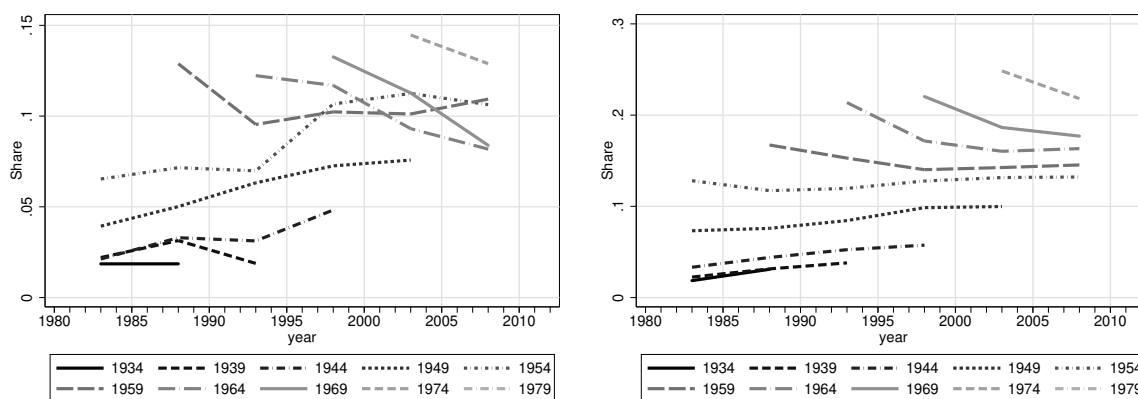
The thresholds for membership in the top percentiles of each age-specific earnings distribution are much higher for older workers than for younger workers. For the five-year period 2008–12, workers who were ages 27-31 in 2010 would have needed to earn an average of at least \$303,000 per year in order to be included in the top 0.1pct of their age group, and workers who were ages 52-58 in 2009 would have needed to earn an average of at least \$1,153,000 over the same period to be included in the top 0.1pct of their age group. For membership in the top 1pct, these thresholds were \$136,000 for ages 22-31 and \$342,000 for

FIGURE 12 – Female population shares by age group and cohort, five-year average earnings

(A) Share of women among top 0.1pct, by age groupgroup (B) Share of women among second 0.9pct, by age



(C) Share of women among top 0.1pct by cohort (D) Share of women among second 0.9pct by cohort



ages 52-58.

Since the early 1980s, these thresholds for membership in the age-specific top percentiles have increased for all age groups but have increased more sharply for older workers. The ratio of the 99.9th percentile of the five-year average earnings distribution for workers ages 52-58 to the 99.9th percentile for workers ages 27-31 was 3.0 in 1981-85 and had increased to 3.8 by 2008-12. See Appendix E for the full time series of top-earning thresholds for each age group.

The share of women among the top 0.1pct and second 0.9pct of earners in a given age group, shown in Figure 12a and Figure 12b respectively, is substantially higher for younger workers. However, in recent years, the share of women among the top 0.1pct of young workers has increased substantially less than the share of women among the top 0.1pct of older workers.

This is in contrast to the second 0.9pct, for which there has been a steady increase in the share of women among top earners of all age groups. Hence, the data show very different trends for the gender composition of young workers in the second 0.9pct versus those in the top 0.1pct of the earnings distribution. Whereas the share of women among the second 0.9pct of workers ages 27-31 increased by more than one-third between the period 1993–97 and the period 2008–12 (from 22% to 29%), the share of women among the top 0.1pct of workers ages 27-31 barely changed over the same period (from 12% to 14.0%).

Viewing these same trends from a cohort perspective rather than an age perspective reveals a striking observation about the source of the increased female share of top earners: almost all of the increase has come from the entry of successive cohorts with higher proportions of women among top earners at all ages, rather than from an increase in the female share within existing cohorts. This can be seen most clearly for the second 0.9pct in Figure 12d, which plots the same data as in Figure 12b, but connects the data for individuals from the same birth cohorts rather than individuals of the same age. Almost no increase has occurred in the share of women among the second 0.9pct of workers within cohorts, and the female shares for the more recent cohorts have actually declined as these cohorts have aged. However, a striking increase has occurred in the gender composition of top earners across cohorts. The same trends are evident for the top 0.1pct (Figure 12c), with the exception of the 1948 and 1953 birth cohorts, which were unique in that the female share increased as these cohorts aged.

If new cohorts continue to follow life-cycle trends for top earner gender shares that are similar to those cohorts just older than them, then these figures imply that we may expect to see a continued increase in the share of women in the second 0.9pct of earners in the next decade, but perhaps a leveling off of the share of women in the top 0.1pct. On the other hand, if these younger cohorts turn out to have trajectories for top earnings shares that mirror more closely those of the baby boomer cohorts, the share of women may continue to rise even at the very top of the earnings distribution.

## **7 Top Earners for Life? Gender Differences among Lifetime Top Earners**

Our analysis has so far focused on gender differences among top earners in a given one-year or five-year period. In this section, we turn our attention to gender differences among top earners over a longer thirty-year period, whom we refer to as lifetime top earners. Our main reason for adopting a lifetime perspective is the sizable transitory component in top

earnings implied by the mobility analysis in Section 4. Moreover, our reliance on first-order Markov transition matrices, which is standard in the literature, may mask richer life-cycle effects and longer-run dynamics that characterize the earnings trajectories of top earners.

One solution would be to explicitly model the earning dynamics for workers in the top percentiles. However, this is beyond the scope of this paper and would take us too far from our main goal of understanding gender differences in top earners. Instead, by measuring lifetime earnings directly, we can observe the cumulative impact of these earnings dynamics and life-cycle trends with a single statistic. Our goals in this section are thus (i) to measure the fraction of lifetime top earners that are female, (ii) to understand how lifetime top earners differ from others in terms of the life-cycle evolution of their earnings, and (iii) to examine how the timing of earnings over the life cycle differs between male and female lifetime top earners. Because of the need for data on the full earnings histories of top earners for this type of analysis, the existing literature offers little in the way of answers to these questions.

We categorize people based on their earnings over the 30 years between ages 25 and 54. Since our data cover the period 1981 to 2012, we have lifetime earnings information for three cohorts of workers.<sup>19</sup> We have chosen to focus on 30-year earnings, since this length balances the objectives of a long horizon that approximates a working life with the need to combine multiple cohorts in order to have a sufficiently large number of individuals in the top 0.1pct of lifetime earners. To construct lifetime earnings for the 25 to 54 age range, we first select all individuals from the 1956, 1957, and 1958 birth cohorts who satisfy the minimum earnings criteria described in Section 2 for a minimum of 15 years.<sup>20</sup> We then compute each individual’s total earnings over this age range and classify these individuals as in either the top 0.1pct, the second 0.9pct, or the bottom 99 percent of the distribution of lifetime earnings for individuals in these cohorts.

## 7.1 Lifetime Top Earners Overall

For the 1956–58 cohorts, the threshold for membership in the lifetime top 0.1pct was just over \$19.1 million (see Table 7). This is equivalent to average annual earnings of around \$635,000, which is smaller than the average threshold for membership in the top 0.1pct based on annual earnings over the same period, \$812,000. The threshold for membership

---

<sup>19</sup>In Appendix F, we report analogous figures and tables for average earnings over the 30 years from ages 30 to 59. Those results yield essentially the same conclusions as those for the 25- to 54-year age range.

<sup>20</sup>Recall that the threshold for satisfying the minimum earnings criterion is equal to the earnings one would obtain by working for 520 hours (13 weeks at 40 hours per week) at one-half of the legal minimum wage in that year.

TABLE 7 – Lifetime earnings top earnings statistics

	Top 0.1%	Second 0.9%	Bottom 99%
<i>30-year earnings thresholds:</i>			
99.9th percentile (\$'000s)	19,052		
99th percentile (\$'000s)		6,459	
Mean 30-year earnings (\$'000s)	33,874	9,552	1,197
Median 30-year earnings (\$'000s)	26,524	8,625	969
Mean no. working years	27.9	28.3	25.7
<i>Mean fraction of working years in age-specific:</i>			
Top 0.1pct	33%	4%	0%
Second 0.9pct	36%	38%	0%
Bottom 99 percent	31%	58%	99%

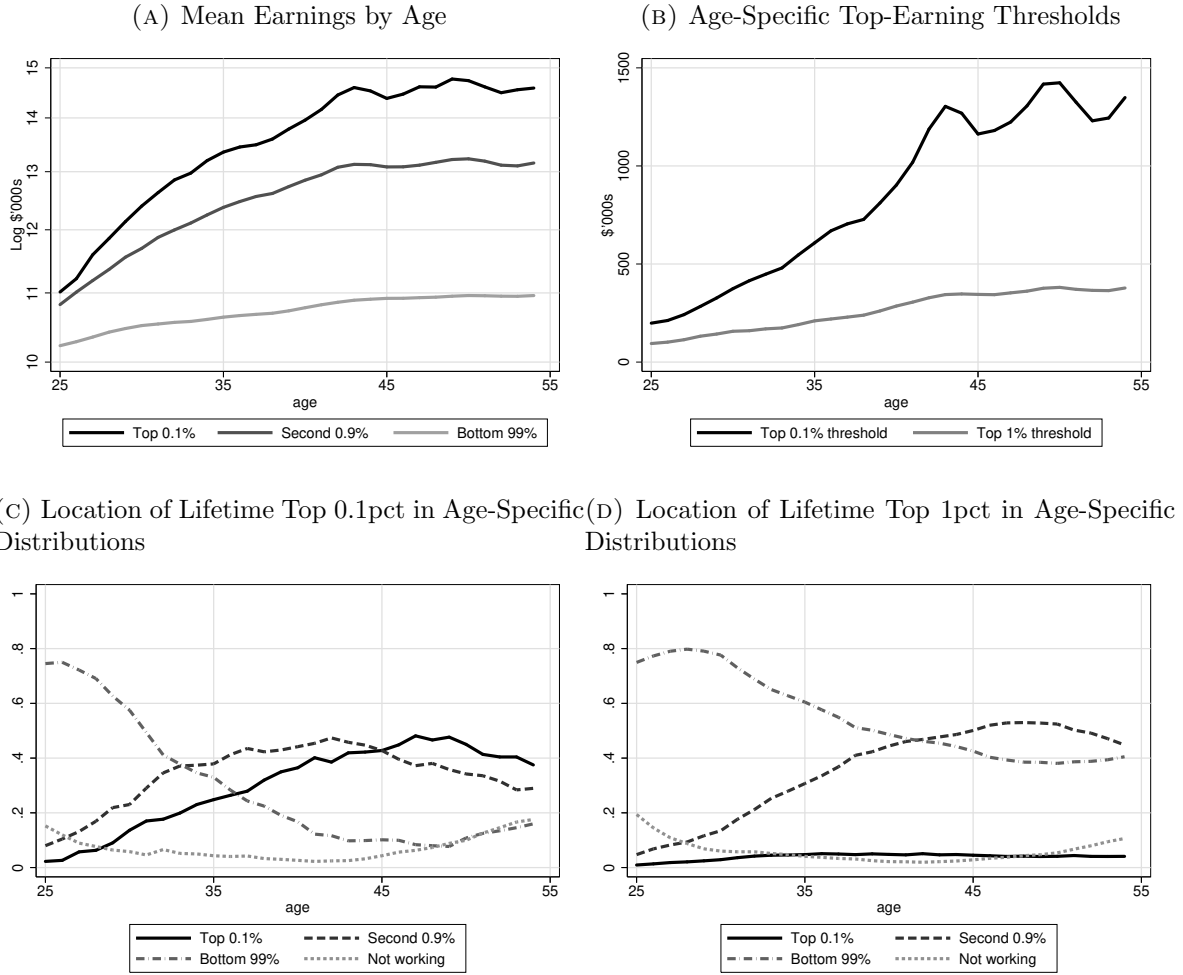
in the lifetime top 1pct was \$6.5 million, equivalent to average annual earnings of \$215,000, which is smaller than the average annual threshold of \$242,000.

Lifetime top earners have high total earnings both because they work for a greater number of years and because they have faster earnings growth than workers in the bottom 99 percent. The top 1pct of lifetime earners work an average of 2.5 years longer than the bottom 99 percent, but those in the top 0.1pct work on average half a year less than those in the second 0.9pct (Table 7).<sup>21</sup> However, these differences in the number of years worked are insignificant when compared with the differential average earnings growth experienced by the three earnings groups conditional on working, shown in Figure 13a. The higher average earnings growth for individuals in the top percentiles takes place entirely between the ages of 25 and 43, after which average earnings are constant for all three groups of workers. Hence, lifetime top earners tend to be workers who experience particularly high earnings growth over the first half of their careers.<sup>22</sup>

<sup>21</sup>Here, we define individuals as working in a given year if they meet the minimum earnings criterion in that year. Due to our imposed selection criteria, all individuals in the sample worked for a minimum of 15 out of the 30 years.

<sup>22</sup>Since we follow workers from only three cohorts, the age patterns that we document naturally confound time and age effects. We have also examined gender gaps in lifetime earnings for individuals in the full distribution of earnings using a smaller 1% sample that goes back to 1957. In those data, we can observe multiple cohorts and hence separate out time and age effects. The results of that analysis lead us to believe that these patterns are more likely to reflect age effects than time effects.

FIGURE 13 – Age Profiles by 30-Year Top Earning Groups



Notes: Figures refer to individuals from the 1956, 1957, and 1958 birth cohorts. Age-specific top-earning thresholds and groups are computed using only these three cohorts.

How closely related are lifetime top earners to annual top earners? This question is important, since although cross-sectional earnings data are more readily available than data on lifetime earnings, for many economic questions lifetime earnings are a more relevant statistic. As we explained in Section 6, the age distribution of top earners is strongly skewed toward older ages (see Figure 11a and Figure 11b). This means that very few lifetime top earners have earnings in the top 1pct of the annual earnings distribution during the first half of their careers. Hence, in order to track the earnings paths of lifetime top earners, it is useful to ask whether they are top earners with respect to their own cohort in a given year, rather than with respect to all workers in that year.

To this end, in each year that he/she is working, we categorize each worker as in either the

top 0.1pct, second 0.9pct, or bottom 99 percent of the age-specific distribution of earnings for workers in these three cohorts. The thresholds for membership in each of these groups at each age are displayed in Figure 13b, and show that the gap between the top 0.1pct and the rest of the top 1pct starts out relatively small and then widens from ages 25 to 43.

When compared with members of their own cohort, lifetime top earners and annual top earners are two very different groups. Typical members of the lifetime top 0.1pct spend nearly one-third of their working years in the bottom 99 percent of their cohort’s annual earnings distribution. The remaining two-thirds of their working years are on average split evenly between the top 0.1pct and the second 0.9pct of earners. The second 0.9pct of lifetime earners spend over half of their working years as members of the bottom 99 percent of annual earnings and only 4% of their time in the top 0.1pct. The average breakdown of working years for lifetime top earners in each annual earnings group is shown in the bottom three rows of Table 7.

The disconnect between annual top earners and lifetime top earners is particularly salient early in the working life. This can be seen in Figure 13c and Figure 13d, which show, respectively, the fraction of the lifetime top 0.1pct and second 0.9pct at each age that are in the within-cohort annual top 0.1pct, second 0.9pct, and bottom 99 percent, as well as the fraction that are not working. At young ages, well over half of both groups of lifetime top earners are in the bottom 99 percent, and even during the peak earnings years during the mid-40s, around 40% of the second 0.9pct of lifetime top earners are in the bottom 99 percent of their within-cohort distribution. This pattern of earnings growth – starting low and rising rapidly – is consistent with the predictions of models of human capital accumulation in the presence of heterogeneity in abilities (see, e.g., Ben-Porath (1967), Guvenen and Kuruscu (2010), and Huggett et al. (2011)). Consequently, identifying individuals as annual top earners may give at best a very noisy signal about their long-term prospects as lifetime top earners.

## 7.2 Gender Differences in Lifetime Top Earners

Since the individuals in the top percentiles of the earnings distribution based on short horizons are possibly a very different group of individuals compared with those that are in the top percentiles based on lifetime earnings, gender differences among annual or five-year top earners may or may not be informative about gender differences among lifetime top earners. In this section, we investigate these differences by measuring gender differences among lifetime top earners directly. Analogously to our analysis of gender differences in Section 3, we approach the measurement of lifetime top earner gender gaps from two perspectives.

TABLE 8 – Gender differences among lifetime top earners

	Top 0.1%	Second 0.9%	Bottom 99%
<i>Panel A: Overall top earners</i>			
Female worker share	12%	13%	48%
Female earnings share	10%	12%	38%
Log mean gender gap	0.17	0.05	0.43
Log p50 gender gap	0.07	0.06	0.46
No. working years gender gap	0.45	-0.05	1.09
<i>Panel B: Gender-specific top earners</i>			
Male threshold (\$'000)	24,471	8,387	
Female threshold (\$'000)	9,324	3,838	
Log mean gender gap	1.08	0.85	0.48
Log p50 gender gap	0.99	0.83	0.47
No. working years gender gap	-0.19	-0.21	1.13

First, we compare men and women in the top percentiles of the overall lifetime earnings distribution. Second, we compare men and women classified as top earners with respect to their gender-specific lifetime earnings distribution.

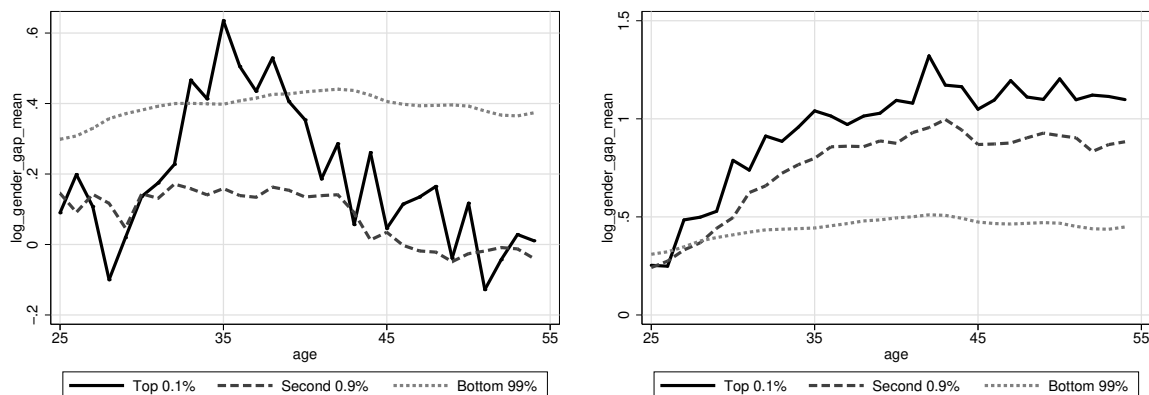
For the 1956–58 cohorts, about 12% of the top 0.1pct of lifetime earners were women (Panel A of Table 8). This compares with an average female share of the top 0.1pct of annual earners for this period of 8%. The fraction of the second 0.9pct of lifetime earners who were women was 13%, which was also the average female share of the second 0.9pct of earners over this period.

Within the top 0.1pct, average lifetime earnings are higher for men than women: there is a 17 basis point difference in the log mean and a 7 basis point difference in the log median. For the second 0.9pct, these differences are both around 5 basis points. In Figure 14a, we plot the gender gap based on annual earnings at each age for the overall lifetime earning groups. For example, the solid line shows the difference between the log of mean annual earnings for men in the top 0.1pct of the lifetime earnings distribution and the log of mean annual earnings for women in the top 0.1pct of the lifetime earnings distribution. The figure clearly illustrates that the gender gap among top earners is largest during the 30s. This finding is consistent with the hypothesis explored in [Bertrand et al. \(2010\)](#), that career interruptions for family reasons explain a substantial portion of the top earnings gender



FIGURE 14 – Gender gap among 30-year top earners by age

(A) Mean gender gap, overall lifetime top earners  
 (B) Mean gender gap, gender-specific lifetime top earners



Notes: Figures refer to individuals from the 1956, 1957, and 1958 birth cohorts. Age-specific top-earning thresholds and groups are computed using only these three cohorts.

gap.

The thresholds for membership in the top 0.1pct and top 1pct of male lifetime earners are over twice as large as those for membership in the corresponding percentiles of female lifetime earners (Panel B of Table 8). Moreover, the gender gaps between the tops of the respective lifetime earnings distribution are large: around 100 basis points in the log mean for the top 0.1pct, and around 85 basis points in the log mean for the second 0.9pct. These compare with a gap of under 0.5 for the bottom 99 percent. These large gender gaps at the top are not driven by a few top earning men, since the gaps in the log median lifetime earnings are very similar to the gaps in the means. Nor are the gaps being driven by women spending more time not working: in fact, on average the lifetime top earning women have a slightly higher number of working years than the lifetime top earning men. Figure 14b shows that these large gender gaps evolve gradually over the first half of the working life. At age 25 the average gender gap in the top percentiles is actually slightly lower than in the bottom 99 percent, but the gap gradually rises and remains constant after around age 35.

## 8 Conclusions

Although we have intentionally remained relatively descriptive in this paper, our findings potentially have important implications for a number of aspects of the U.S. economy. There-

fore, rather than concluding with a summary of our findings (for that, we refer readers to the introduction), we will conclude by mentioning some areas in which our empirical observations suggest the need for complementary theoretical work and further empirical analysis using other data sources.

We found that although the share of women among the top 1pct has increased steadily over the last 30 years, the fraction of women in the top 0.1pct has barely increased during the last decade, and the gender composition of both top earning groups is still very different from the composition of the bottom 99 percent. These findings reinforce the need for research into the factors that can account for both the glass ceiling and the paper floor. Our analysis of lifetime top earners revealed that the timing of the emergence of the top earnings gender gap is consistent with the hypothesis that career interruptions may be an important consideration. Our finding that industry composition plays very little role in explaining either the level or the change in the top earnings gender gap suggests that in this respect, selection into particular firms or jobs may be more important than selection into particular industries. Unfortunately, the SSA data lack many of the important variables that would be required for a more complete answer to this question: children, marital status, and work hours.

The large temporary component in top earners' earnings, the increasing persistence of top earner status, and the relatively weak relationship between annual top earners and lifetime top earners, all suggest the need for a comprehensive analysis of the dynamics of top earnings. Although an extensive literature has proposed and estimated various statistical models that provide a good fit to the dynamics of earnings for the bulk of the distribution, little is currently known about how well this class of models fits earnings dynamics for top earners.

On the theory side, our findings suggest the need for progress in at least two areas. First, there is the need to understand how and why the earnings distribution is characterized by a Pareto tail. Most existing theories of Pareto-generating mechanisms, such as the accumulation of random returns over long periods of time, can be adapted to explaining right-tail inequality in the wealth distribution, which is accumulated over time and passed down across generations. But for explaining right-tail inequality in earnings, new theories are necessary, since human capital is less easily transmitted across generations and, as we have shown, a large fraction of top earnings is accrued within a lifetime, often in just a few years.

Second, the rise of the Finance and Insurance industry in accounting for top earners of both genders suggests a need for better theories of labor compensation in this sector. Why is such a large share of labor earnings concentrated in a single industry? Does this reflect the

extraordinarily high productivity of this industry? Or do these earnings reflect rents? And if so, rents to what? A useful starting point would be to study the top of the distribution of earnings across and within firms, within industries.

## References

- Abowd, J. M. and Card, D. (1989). On the covariance structure of earnings and hours changes. *Econometrica*, 57(2):411–445. [2.2](#)
- Auten, G., Gee, G., and Turner, N. (2013). Income inequality, mobility, and turnover at the top in the US, 1987-2010. *American Economic Review, Papers & Proceedings*, 103(3):168–172. [1](#), [15](#)
- Autor, D. H., Katz, L. F., and Kearney, M. S. (2008). Trends in U.S. wage inequality: Revising the revisionists. *Review of Economics and Statistics*, 90(2):300–323. [2.2](#)
- Badel, A. and Huggett, M. (2014). Taxing top earners: A human capital perspective. Working paper, Georgetown University. [1](#)
- Baker, S. R., Bloom, N., Canes-Wrone, B., Davis, S. J., and Rodden, J. A. (2014). Why has U.S. policy uncertainty risen since 1960? *American Economic Review, Papers and Proceedings*, 104(5):56–60. [5](#)
- Bakija, J., Cole, A., and Heim, B. (2012). Jobs and income growth of top earners and the causes of changing income inequality: Evidence from U.S. tax return data. Working paper, Williams College. [1](#)
- Barber, M. (2013). Ideological donors, contribution limits, and the polarization of state legislatures. Technical report, Brigham Young University. [5](#)
- Bell, L. A. (2005). Women-led firms and the gender gap in top executive jobs. Discussion Paper 1689, Institute for the Study of Labor (IZA). [1](#)
- Ben-Porath, Y. (1967). The production of human capital and the life cycle of earnings. *Journal of Political Economy*, 75(4):352–365. [7.1](#)
- Bertrand, M., Black, S., Jensen, S., and Lleras-Muney, A. (2012). Breaking the glass ceiling? The effect of board quotas on female labor market outcomes in Norway. Working paper, University of Chicago. [1](#)
- Bertrand, M., Goldin, C., and Katz, L. F. (2010). Dynamics of the gender gap for young professionals in the financial and corporate sectors. *American Economic Journal: Applied Economics*, 2(3):228–255. [1](#), [7.2](#)
- Brewer, M., Sibieta, L., and Wren-Lewis, L. (2007). Racing away? income inequality and the evolution of high incomes. [1](#)

- Congressional Budget Office (2013). The distribution of household income and federal taxes, 2010. [4](#)
- Federal Glass Ceiling Commission (1995). A solid investment: Making full use of the nation's human capital. Technical report, U.S. Department of Labor, Washington, D.C. [2](#)
- Gabaix, X. and Landier, A. (2008). Why has CEO pay increased so much? *Quarterly Journal of Economics*, 123(1):49–100. [1](#)
- Gayle, G.-L., Golan, L., and Miller, R. A. (2012). Gender differences in executive compensation and job mobility. *Journal of Labor Economics*, 30(4):829–872. [1](#)
- Guner, N., Lopez-Daneri, M., and Ventura, G. (2014). Heterogeneity and government revenues: Higher taxes at the top? Working paper, Arizona State University. [1](#)
- Guvenen, F. and Kaplan, G. (2017). Top income inequality in the 21st century: Some cautionary notes. *Federal Reserve Bank of Minneapolis Quarterly Review*. [6](#), [13](#)
- Guvenen, F., Kaplan, G., and Song, J. (2014a). How risky are recessions for top earners? *American Economic Review, Papers and Proceedings*, 104(5):148–153. [1](#)
- Guvenen, F. and Kuruscu, B. (2010). A quantitative analysis of the evolution of the u.s. wage distribution: 1970-2000. In *NBER Macroeconomics Annual 2009, Volume 24*, pages 227–276. University of Chicago Press. [7.1](#)
- Guvenen, F., Ozkan, S., and Song, J. (2014b). The nature of countercyclical income risk. *Journal of Political Economy*, 122(3):621–660. [1](#)
- Hsieh, C.-T., Hurst, E., Jones, C. I., and Klenow, P. J. (2013). The allocation of talent and U.S. economic growth. Working paper, Stanford University. [1](#)
- Huggett, M., Ventura, G., and Yaron, A. (2011). Sources of lifetime inequality. *American Economic Review*, 101(7):2923–2954. [7.1](#)
- Jones, C. I. and Kim, J. (2014). A Schumpeterian model of top income inequality. Working paper, Stanford University. [1](#)
- Juhn, C., Murphy, K. M., and Pierce, B. (1993). Wage inequality and the rise in returns to skill. *Journal of Political Economy*, 101(3):410–442. [2.2](#)
- Kindermann, F. and Krueger, D. (2014). The redistributive benefits of progressive labor and capital income taxation. Working paper, University of Pennsylvania. [1](#)

- Kopczuk, W., Saez, E., and Song, J. (2010). Earnings inequality and mobility in the United States: Evidence from Social Security data since 1937. *Quarterly Journal of Economics*, 125(1):91–128. [1](#), [15](#)
- Leonesio, M. V. and Del Bene, L. (2011). The distribution of annual and long-run US earnings, 1981-2004. *Social Security Bulletin*, 71(1):17–33. [7](#)
- Meghir, C. and Pistaferri, L. (2004). Income variance dynamics and heterogeneity. *Econometrica*, 72(1):1–32. [2.2](#)
- Olsen, A. and Hudson, R. (2009). Social Security Administration’s master earnings file: Background information. *Social Security Bulletin*, 69(3):29–46. [2.1](#), [7](#), [8](#)
- Panis, C., Euler, R., Grant, C., Bradley, M., Peterson, C. E., Hirscher, R., and Steinberg, P. (2000). *SSA Program Data User’s Manual*. Social Security Administration, Baltimore, MD. [2.1](#)
- Parker, J. A. and Vissing-Jørgensen, A. (2010). The increase in income cyclicality of high-income households and its relation to the rise in top income shares. *Brookings Papers on Economic Activity*, 41(2):1–70. [1](#)
- Piketty, T. and Saez, E. (2003). Income inequality in the United States, 1913-1998. *Quarterly Journal of Economics*, 118(1):1–39. [1](#)
- Saez, E. (2001). Using elasticities to derive optimal income tax rates. *Review of Economic Studies*, 68(1):205–229. [1](#)
- Storesletten, K., Telmer, C. I., and Yaron, A. (2004). Cyclical dynamics in idiosyncratic labor market risk. *Journal of Political Economy*, 112(3):695–717. [2.2](#)
- Wolfers, J. (2006). Diagnosing discrimination: Stock returns and CEO gender. *Journal of the European Economic Association*, 4(2-3):531–541. [1](#)

# A Details of Decompositions

In this appendix, we provide details of the methodology underlying the decompositions presented in Table 1, Table 3, Table 5 and Table 6.

We start by establishing some notation. Let  $G_{it}$  be the gender of individual  $i$  who is included in our sample in year  $t$ , with the convention that  $G_{it} = 1$  for a female and  $G_{it} = 0$  for a male. Let  $p$  denote a percentile range (e.g. top 0.1pct, second 0.9pct or bottom 99 percent) and let  $D_{it}^p$  be an indicator variable that takes the value 1 if individual  $i$  is in the percentile range  $p$  of the earnings distribution in year  $t$ . Let  $\sigma_t^p$  be the fraction of top earners that are female.

$$\sigma_t^p = E_t [G|D^p = 1] \quad (1)$$

Let  $E_t$  denote a moment of a time  $t$  distribution and let  $P_t$  denote a probability based on the time  $t$  distribution.

## A.1 Decomposition for changing gender composition of the labor force (Table 1)

The goal is to measure how much of the observed change in  $\sigma_t^p$  is due to a changes in the share of women in the labor force  $E_t [G]$ . Using Bayes' rule we can decompose  $\sigma_t^p$  as

$$\sigma_t^p = \frac{P_t [D^p = 1|G = 1] P_t [G = 1]}{P_t [D^p = 1]} \quad (2)$$

$$\sigma_t^p P_t [D^p = 1] = E_t [D^p|G = 1] E_t [G] \quad (3)$$

$$\Delta (\sigma_t^p P_t [D^p = 1]) = E_t [D^p|G = 1] (\Delta E_t [G]) + (\Delta E_t [D^p|G = 1]) E_{t-1} [G] \quad (4)$$

The term on the LHS of (4) is the change in the fraction of the workforce that are female and in percentile group  $p$ . The first term on the RHS of (4) is the component of this change that is due to changes in the share of women in the labor force. The second term on the RHS is the component that is due to changes in the fraction of women that are in percentile group  $p$ . We implement this decomposition for each pair of consecutive years using sample analogues of the moments in (4) and then summing the components over all years to get the total decomposition.

In principal  $P_t [D^p = 1]$  is constant for all  $t$ , since it is simply the fraction of the population in percentile group  $p$ . However, since we take different size random samples for the top percentile groups compared with the bottom 99 percent, in practice there are small year-to-year fluctuations in our sample estimates of this moment. If  $P_t [D^p = 1]$  were constant then the fraction of  $\Delta \sigma_t^p$  that is due to changes in the gender composition of the labor force would be given by

$$\frac{E_t [D^p|G = 1] \Delta E_t [G]}{P_t [D^p = 1] \Delta \sigma_t^p} \quad (5)$$

With our decomposition the fraction is given by

$$\frac{E_t [D^p|G = 1] \Delta E_t [G]}{P_t [D^p = 1] \Delta \sigma_t^p + \sigma_{t-1}^p \Delta P_t [D^p = 1]} \quad (6)$$

Since the term  $\sigma_{t-1}^p \Delta P_t [D^p = 1]$  is very small relative to  $P_t [D^p = 1] \Delta \sigma_t^p$ , this sampling variation has a negligible effect on the results of the decomposition.

## A.2 Decomposition for changing for age and industry composition (Table 5, Table 6)

The goal is to measure how much of the observed change in  $\sigma_t^p$  is due to a changes in the distribution of an observable characteristic  $X_{it}$ . We consider only characteristics that which take a discrete set of values such as age and industry. Analogously to the decomposition above we can write

$$\begin{aligned}
\sigma_t^p P_t [D^p = 1] &= E_t [D^p | G = 1] E_t [G = 1] \\
&= \sum_x E_t [D^p | G = 1, X = x] P_t [X = x | G = 1] E_t [G] \\
&= \sum_x E_t [D^p | G = 1, X = x] E_t [G | X = x] P_t [X = x] \tag{7} \\
\Delta (\sigma_t^p P_t [D^p = 1]) &= \sum_x E_t [D^p | G = 1, X = x] \Delta E_t [G | X = x] P_t [X = x] \\
&\quad + \sum_x \Delta E_t [D^p | G = 1, X = x] E_{t-1} [G | X = x] P_t [X = x] \\
&\quad + \sum_x E_{t-1} [D^p | G = 1, X = x] E_{t-1} [G | X = x] \Delta P_t [X = x] \tag{8}
\end{aligned}$$

The term on the LHS of (8) is the change in the fraction of the workforce that are female and in percentile group  $p$ . The first term on the RHS is the component of this change that is due to changes in the gender composition of different categories (i.e. industries or age groups). The second term on the RHS is the component that is due to changes in the fraction of women in each category that are in percentile group  $p$ . The third term on the RHS is the component that is due to changes in the fraction of the overall labor force in each category of  $X$ .

## A.3 Decomposition for changes in mobility (Table 3)

The goal is to measure how much of the observed change in  $\sigma_t^p$  is due to changes in the transition probabilities in and out of the percentile group  $p$ . Let  $D_+^p$  be an indicator variable that takes the value 1 if an individual was in percentile group  $p$  in year  $t + 1$ . Since gender is constant over time,  $G_t = G_{t-1}$ , we can decompose  $\sigma_t^p$  using the relationship that

$$\begin{aligned}
\sigma_t^p P_t [D^p = 1] &= E_t [D^p | G = 1] E_t [G = 1] \\
&= \sum E_{t-1} [D_+^p | G = 1, D^q = 1] E_{t-1} [D^q | G = 1] E_{t-1} [G = 1] \\
&= \sum E_{t-1} [D_+^p | G = 1, D^q = 1] E_{t-1} [G | D^q = 1] E_{t-1} [D^q] \tag{9}
\end{aligned}$$

Then taking first differences yields

$$\begin{aligned}
\Delta (\sigma_t^p P_t [D^p = 1]) &= \sum_q E_{t-1} [D_+^p | G = 1, D^q = 1] \Delta E_{t-1} [G | D^q = 1] E_{t-1} [D^q] \\
&\quad + \sum_q \Delta E_{t-1} [D_+^p | G = 1, D^q = 1] E_{t-2} [G | D^q = 1] E_{t-1} [D^q] \\
&\quad + \sum_q E_{t-2} [D_+^p | G = 1, D^q = 1] E_{t-2} [G | D^q = 1] \Delta E_{t-1} [D^q] \tag{10}
\end{aligned}$$

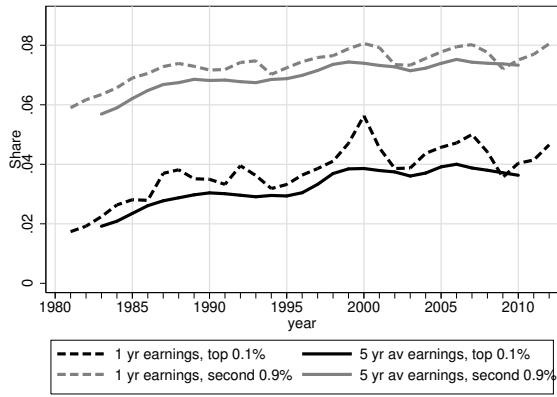


The term on the LHS of (10) is the change in the fraction of the workforce that are female and in percentile group  $p$ . The first term on the RHS is the component of the change that is due to changes in the female share of top percentiles in the previous period at the prevailing levels of persistence. The second term on the RHS is the component of this change that is due to changes in the transition probabilities into the top  $p$ -the percentile. The third term is due to sampling variation and is a negligible component of the overall change; we present the decomposition for the change net of the effects of this term.

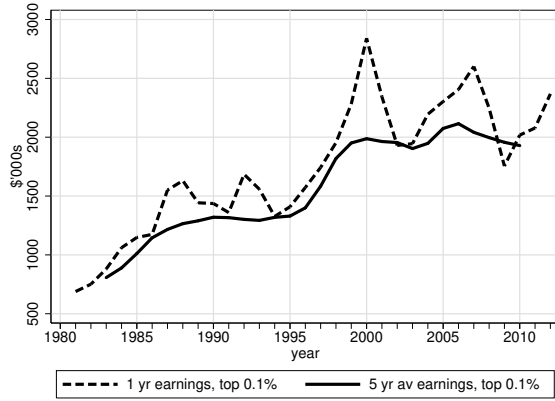
The idea behind this decomposition is that any one-time change in transition probabilities will lead to continued changes in the fraction of women in the top percentiles in subsequent years, even if there are no further changes in the transition probabilities. Hence any observed change is partly due to the effects of changes in the transition probabilities in the past as the system moves towards its new stationary distribution, and is partly due to new changes in the transition probabilities. The first term captures the former effect, the second term captures the latter effect.

FIGURE B.1 – Average earnings among top earners

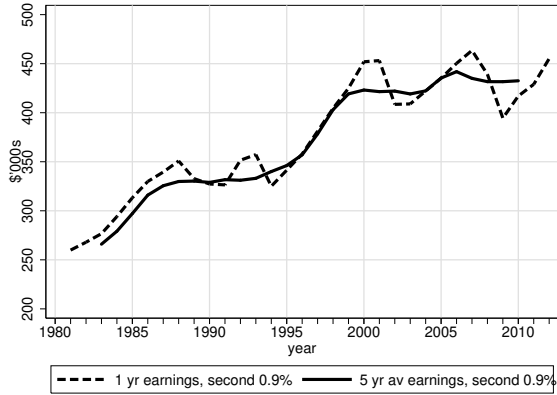
(A) Top earnings shares



(B) Average earnings in top 0.1pct



(C) Average earnings in second 0.9pct



(D) Average earnings in bottom 99 percent

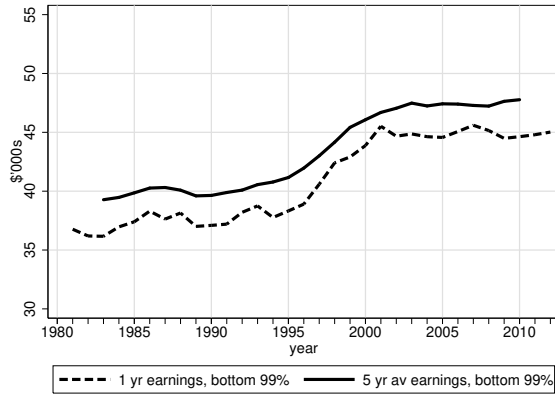
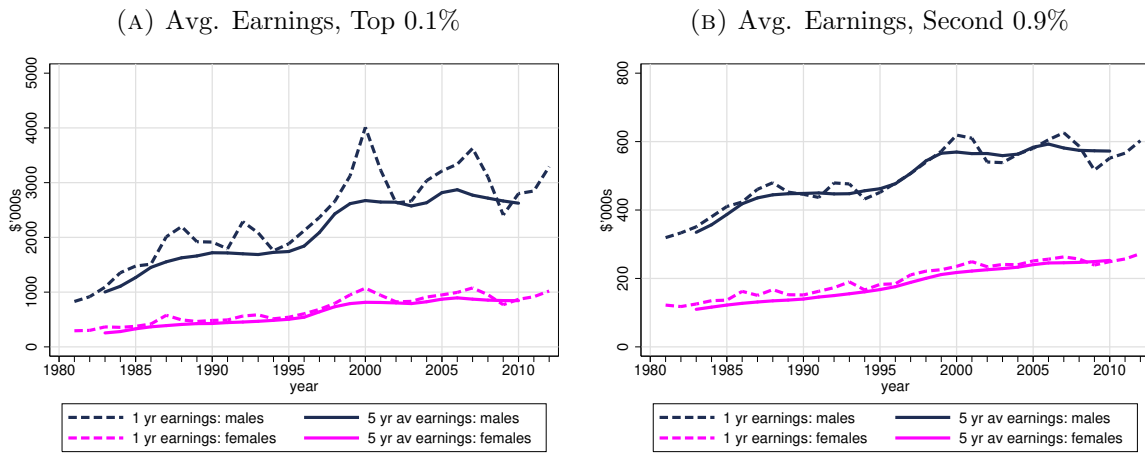


FIGURE B.2 – Top Earning Men versus Top Earning Women (Based on Gender-Specific Earnings Distributions)



## B Top Earner Trends and Mobility Additional Figures

FIGURE B.3 – Females in Male Earnings Distribution: Counterpart of Figure 5 for Annual Earnings

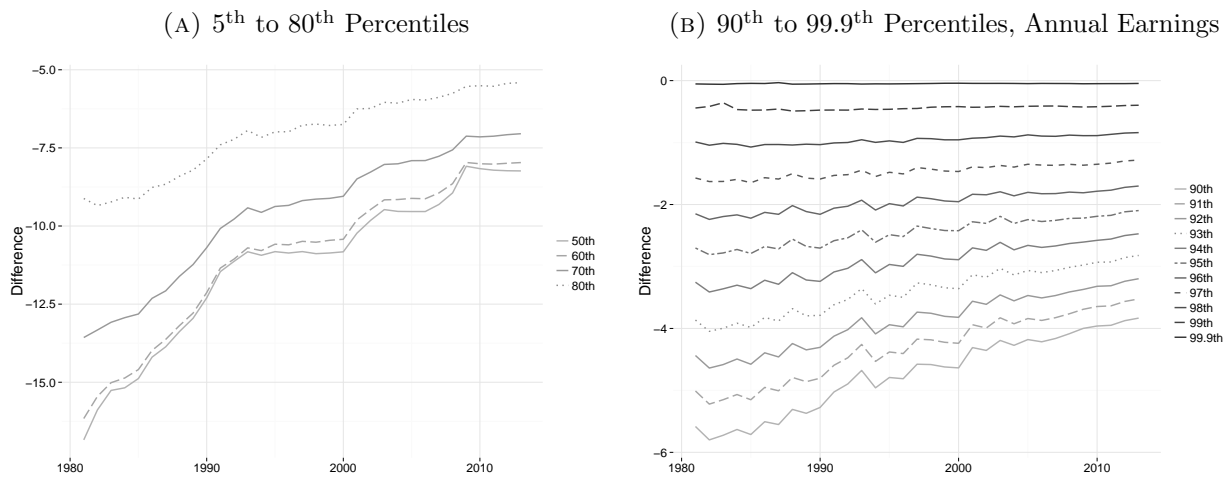
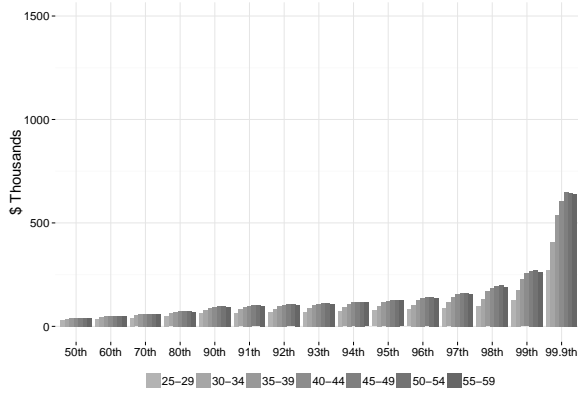
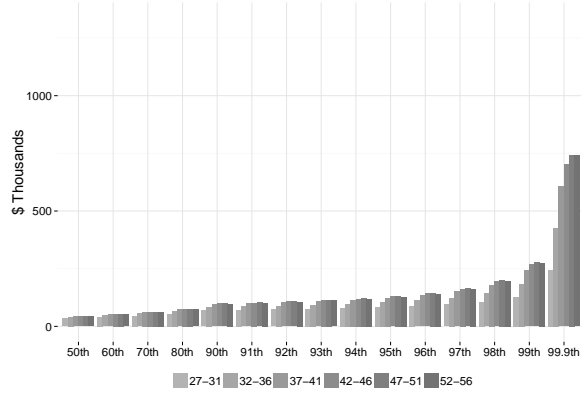


FIGURE B.4 – Top Earning Thresholds by Age

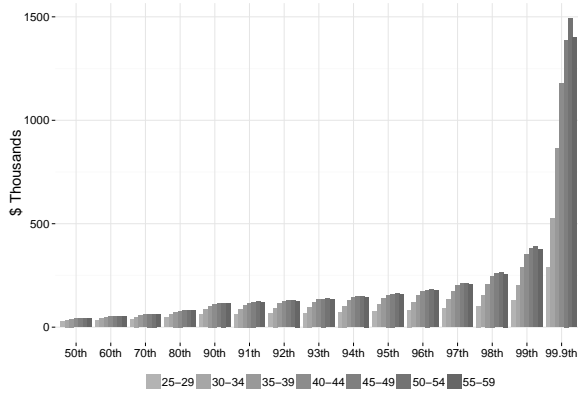
(A) Annual Threshold, 1981



(B) Five-Year Threshold, 1981–85



(C) Annual Threshold, 2012



(D) Five-Year Threshold, 2008–12

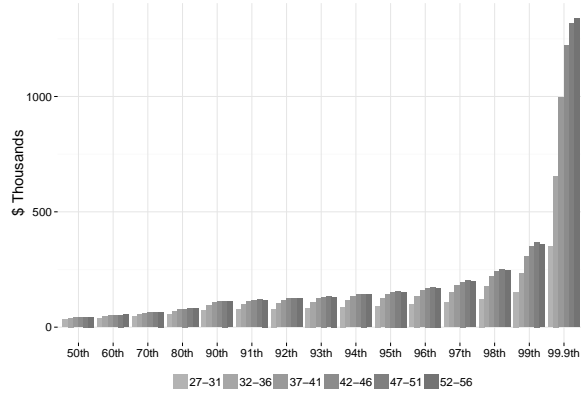


FIGURE B.5 – Top Earning Thresholds by Age and Gender, 5-Year Earnings

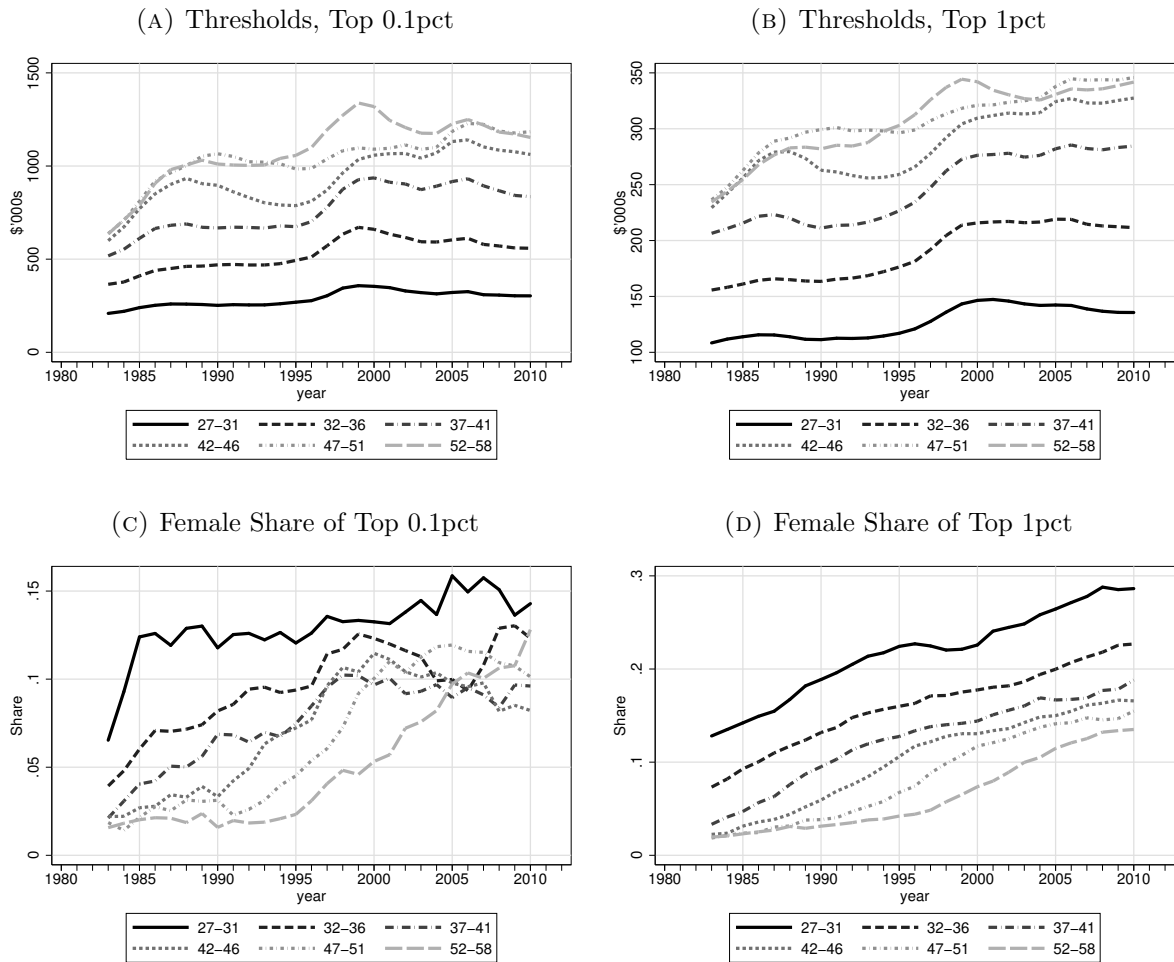
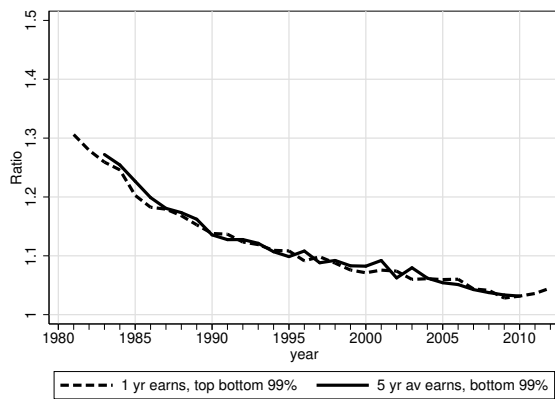


FIGURE B.6 – Gender composition of overall top earners, bottom 99%

(A) Female population share

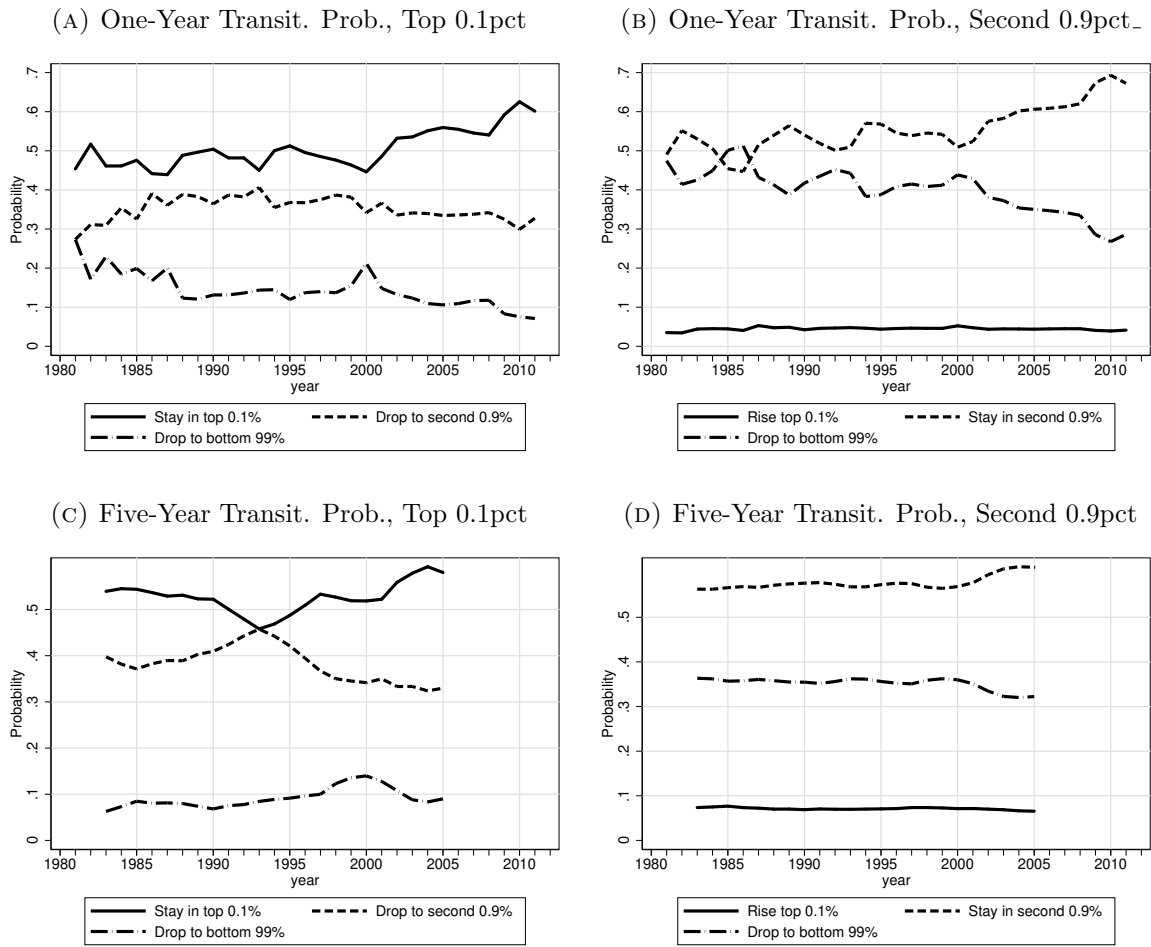


(B) Male-female population ratio



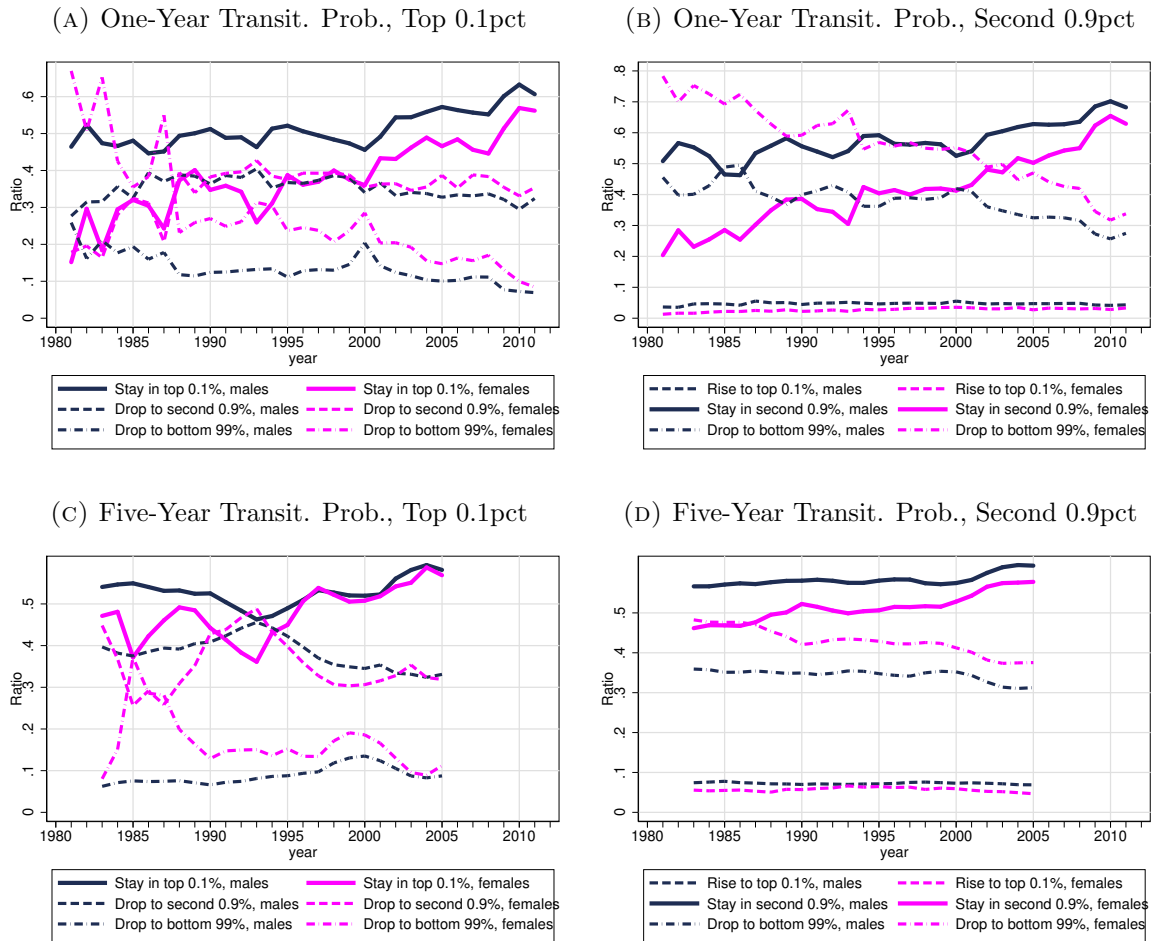
Notes: Time trend for the female population share and the male-female population ratio, for the bottom 99 percent of the earnings distribution.

FIGURE B.7 – Transition Probabilities In and Out of Top Percentiles, Conditional on Not Leaving Sample



Notes: These figures show the probability that a top earner based on average earnings over the period  $t - 2, \dots, t + 2$  is a top earner based on average earnings over the period  $t + 3, \dots, t + 7$ .

FIGURE B.8 – Transition Probabilities In and Out of Top Percentiles Over Time Conditional on Not Leaving Sample, by Gender

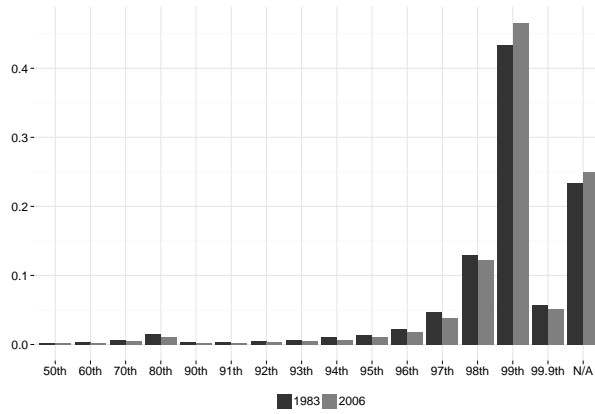


Notes: These figures show the probability that a top earner based on average earnings over the period  $t - 2, \dots, t + 2$  is a top earner based on average earnings over the period  $t + 3, \dots, t + 7$ , separately for male top earners (blue) and female top earners (pink).

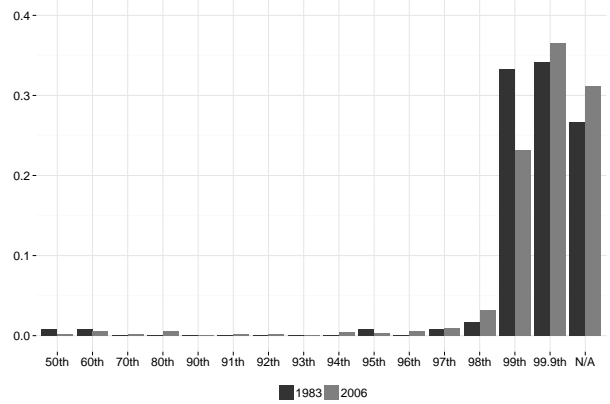
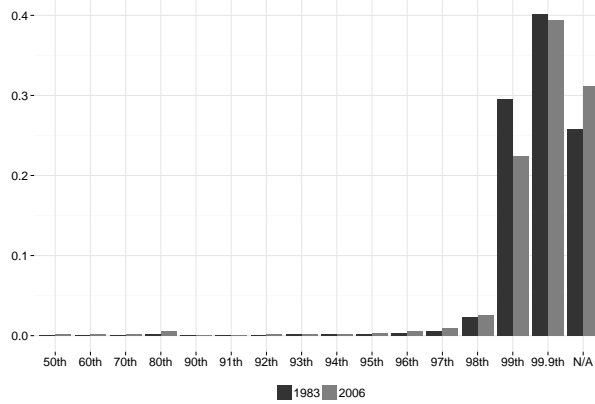
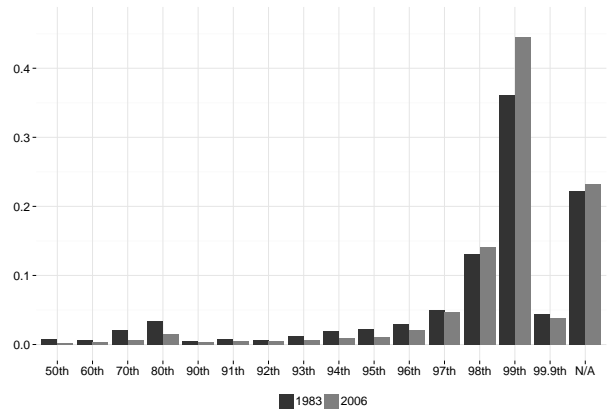


FIGURE B.9 – Changes in Five-Year Transition Probabilities out of Top Percentiles into Finer Percentile Groups, by Gender

(A) Men, Transition out of Second 0.9pct



(B) Women, Transition out of Second 0.9pct



(C) Men, Five-Year Transition out of Top 0.1pct

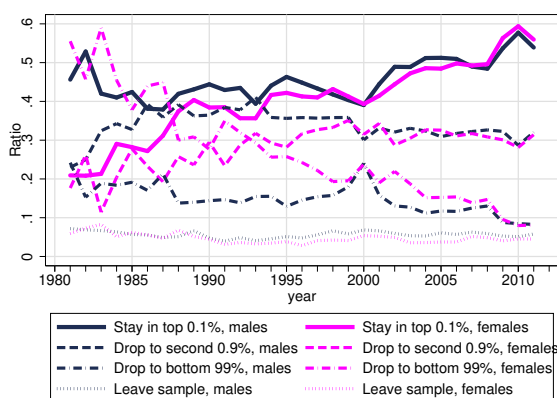
(D) Women, Five-Year Transition out of Top 0.1pct

## C Mobility within gender-specific distributions

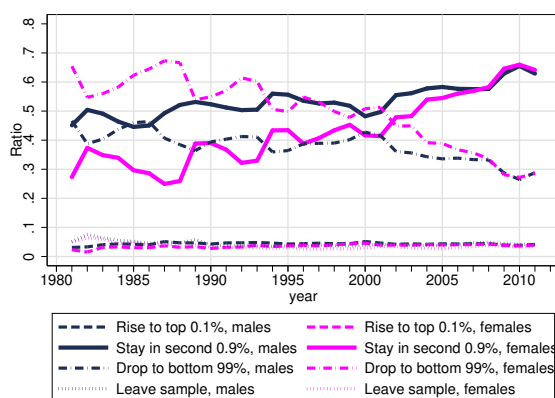
This appendix reports figures that are analogous to those in Section 4, but in which individuals are defined as top earners based on their position in their gender-specific earnings distribution, rather than the overall earnings distribution.

FIGURE C.1 – Transition probabilities in and out of top percentiles of earnings distribution, by gender

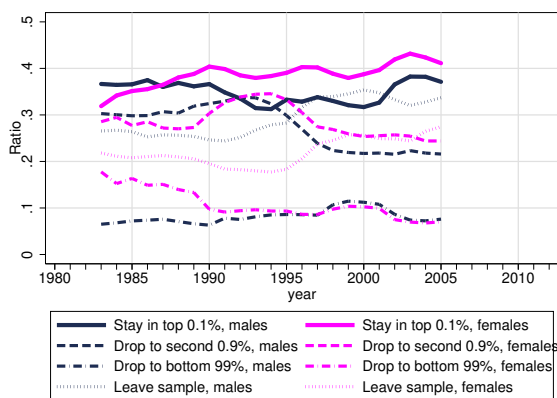
(A) One-year transition probabilities for annual earnings, top 0.1pct



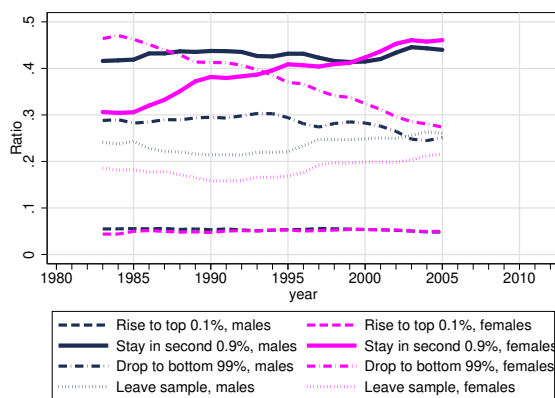
(B) One-year transition probabilities for annual earnings, second 0.9pct



(C) Five-year transition probabilities for five-year earnings, top 0.1pct



(D) Five-year transition probabilities for five-year earnings, second 0.9pct



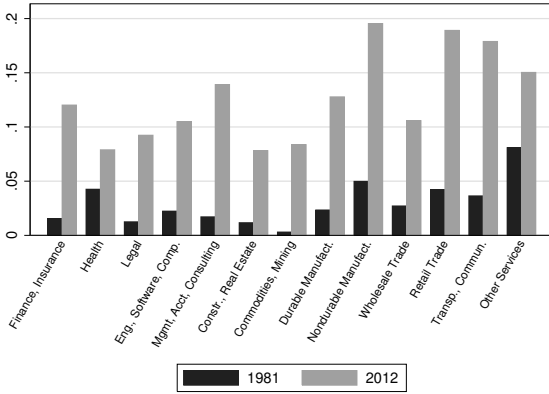
Notes: These figures show the probability that a top earner based on average earnings over the period  $t - 2, \dots, t + 2$  is a top earner based on average earnings over the period  $t + 3, \dots, t + 7$ , separately for male top earners (blue) and female top earners (pink). Individuals are classified as top earners based on gender-specific earnings distributions.

# D Industry analysis further figures

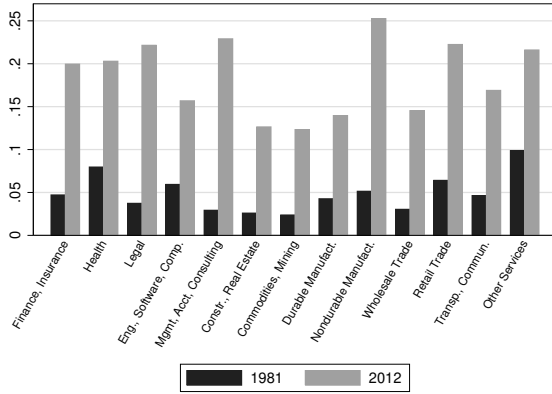
This appendix contains figures that are analogous to those in Section 5, but which are constructed using annual earnings rather than five-year average earnings.

FIGURE D.1 – Top earners by industry and gender, annual earnings

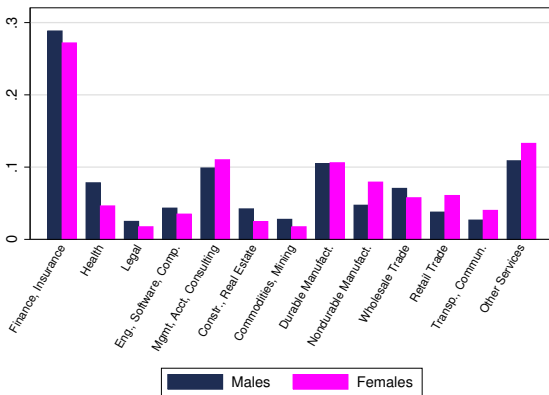
(A) Share of women by industry within top 0.1pct



(B) Share of women by industry within top 0.9 percent



(C) Industry shares by gender within top 0.1pct, 2008–12



(D) Industry shares by gender within second 0.9pct, 2008–12

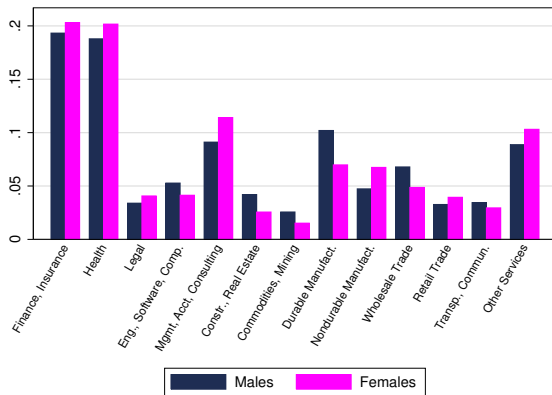


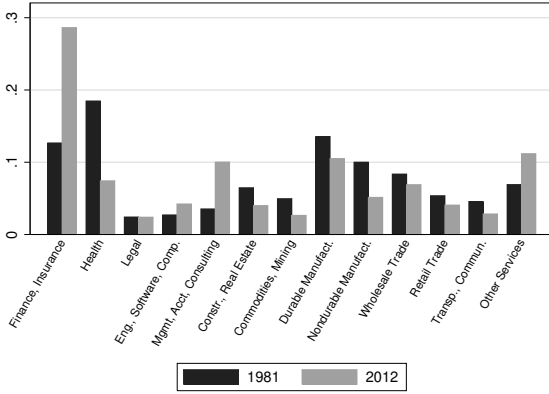
TABLE D.1 – Selected US Companies and Associated (Primary) SIC Codes

Company Name:	Primary SIC Code	Descriptions
Google	7370	Computer Programming, Data Processing, And Computer Services
Apple,Dell	3571	Electronic computers
HP	3570	Computer and office equipment
Microsoft	7372	Prepackaged software
IBM	7371	Computer programming services
Intel	3674	Semiconductors and related services
Oracle	7372	Prepackaged software
Cisco	5045	Wholesale-Computers and Peripheral equipment and Software
Qualcomm	3663	Radio and TV broadcasting and communication equipment
Boeing	3721	Aircraft and parts
Amazon.com	5961	Retail-Catalog and Mail Order Houses
3M	3291	Abrasive products
Walmart	5331	Retail-Variety stores
Exon, Chevron, BP	2911	Petroleum refining
Total SA	1211	Crude petroleum and natural gas
Ford, GM, Tesla	3711	Motor vehicles and passenger car bodies
Benkshire-Hathaway, State Farm	6331	Fire, Marine and Casualty Insurance
General Electric:	3600	Electronic and other electrical equipment except computers
Cargill Inc	5153	Grain and field beans; Domestic Transportation of Freight
Bank of America, JP Morgan	6021	Banks
Goldman Sachs	6022	Investment bank
Morgan Stanley	6199	Investment bank
Mettle	6311	Life insurance

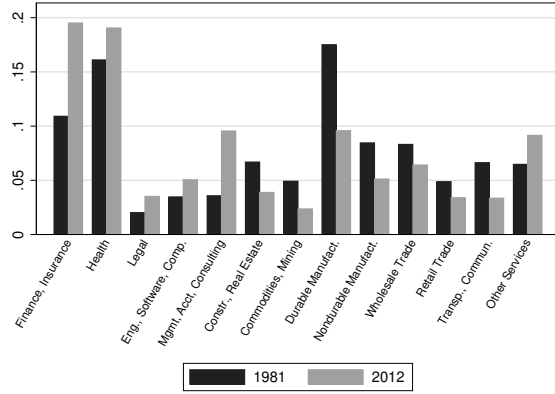
Notes: Some companies listed here have further SIC codes associated with them. For example, Microsoft: 7371, 7372, 7379 (Prepackaged software, primary), and 3944 (electronic games) and 3861 (photographic equipment). And similarly, Cargill Inc: 5153 (Grain & Field Beans); 4424 (Deep Sea Domestic Transportation of Freight); 6221 (Commodity Contracts Brokers & Dealers); 2041 (Flour & Other Grain Mill Products.)

FIGURE D.2 – Industry composition of top earners, annual earnings

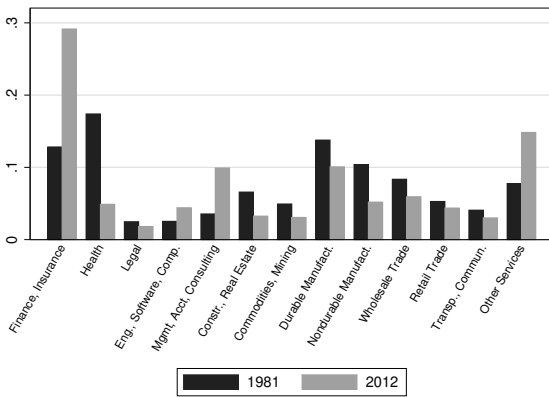
(A) Population shares, top 0.1pct



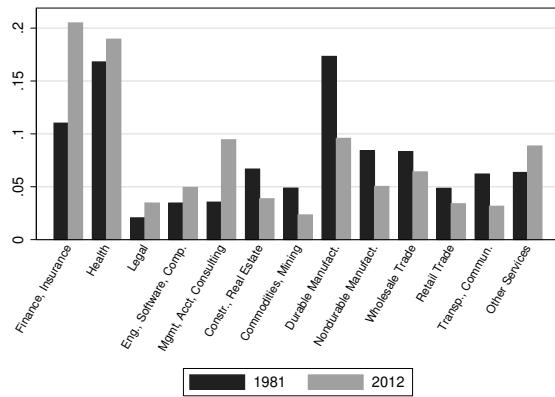
(B) Population shares, second 0.9pct



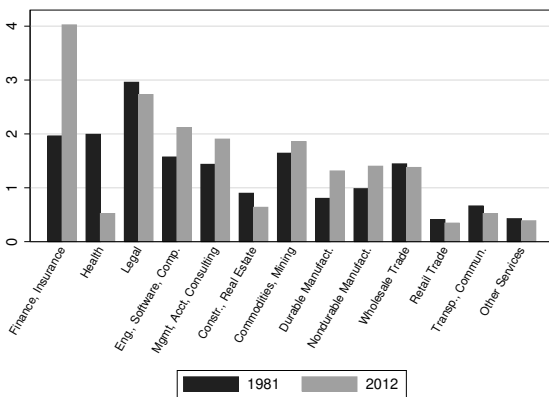
(C) Earnings shares, top 0.1pct



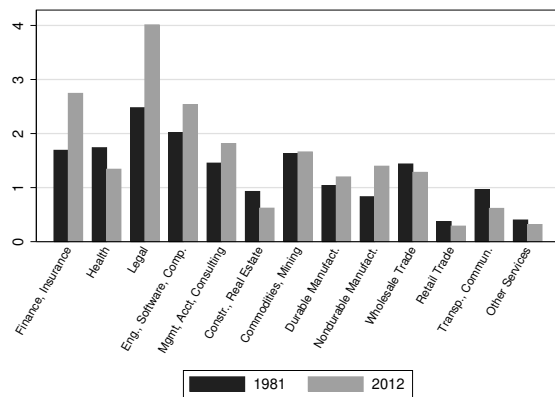
(D) Earnings shares, second 0.9pct



(E) Population shares, top 0.1pct relative to bottom 99 percent



(F) Population shares, second 0.9pct relative to bottom 99 percent



## E Age analysis further figures

This appendix contains figures that are analogous to those in Section 6, but which are constructed using annual earnings rather than five-year average earnings, and additional figures that are references in Section 6.

FIGURE E.1 – Age distribution of workers, annual earnings

(A) Age distribution of individuals in top 0.1pct (B) Age distribution of individuals in second 0.9pct

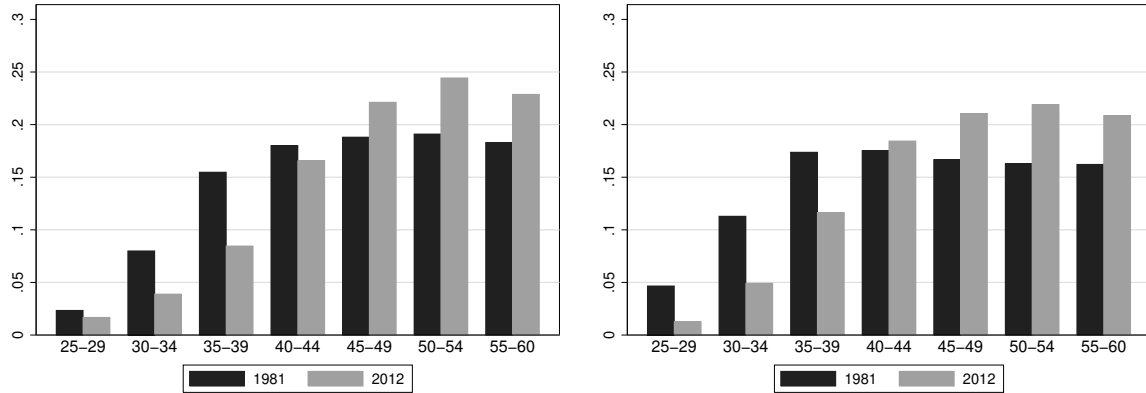


FIGURE E.2 – Age distribution of workers by gender, overall distribution, five-year average earnings

(A) 1981-85

(B) 2008-12

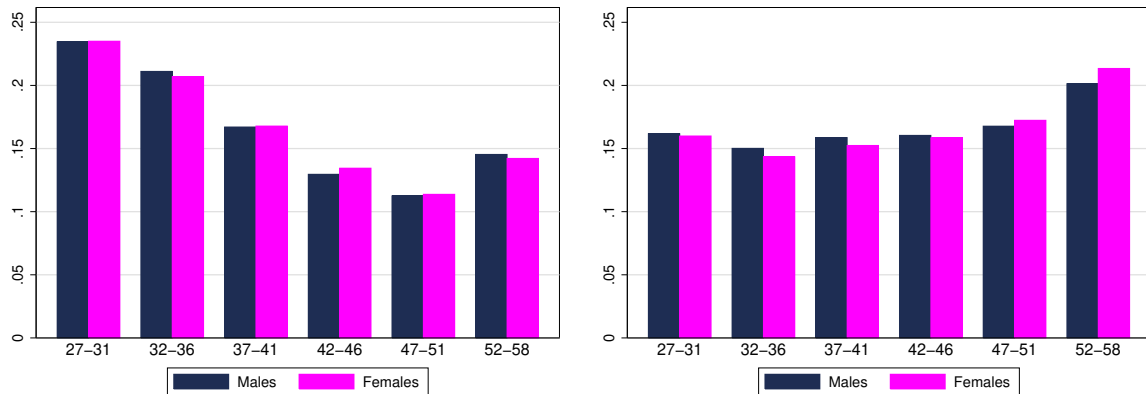
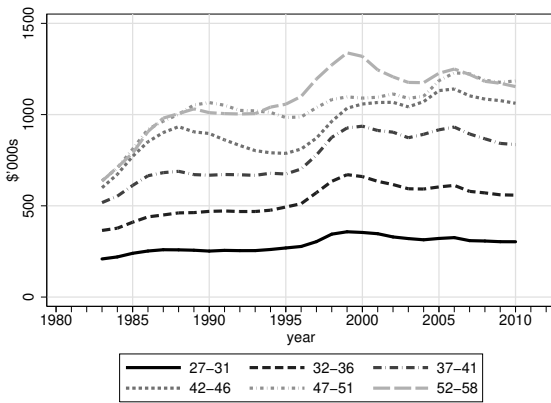
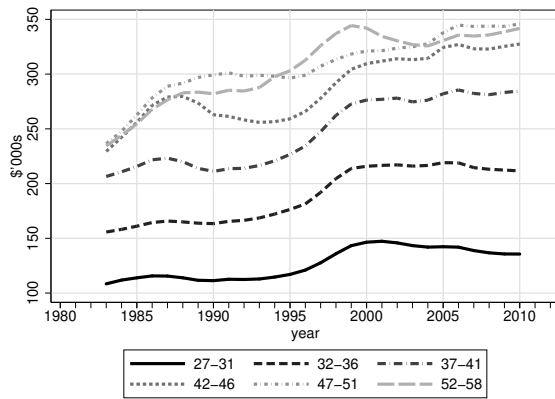


FIGURE E.3 – Top-earning thresholds within age groups, five-year average earnings



(A) Thresholds for top 0.1pct, by age group



(B) Thresholds for top 1pct, by age group

## F Lifetime earnings analysis for 30-59 year age range

This appendix reports analogous tables and figures to those in Section 7, but where the 30 year age range is taken to be the ages 30 to 59, rather than 25 to 54.

TABLE F.1 – Lifetime earnings top earnings statistics

	<u>Top 0.1%</u>	<u>Second 0.9%</u>	<u>Bottom 99%</u>
<u>30-year earnings thresholds:</u>			
- 99.9th percentile (\$'000s)	20,704		
- 99th percentile (\$'000s)		7,043	
Mean 30-year earnings (\$'000s)	38,092	10,545	1,276
Median 30-year earnings (\$'000s)	29,467	9,443	1,043
Mean no. working years	27.9	28.3	25.6
<u>Mean fraction of working years in age-specific:</u>			
- top 0.1pct	35%	5%	0%
- next 0.9pct	40%	42%	0%
- bottom 99 percent	25%	53%	100%

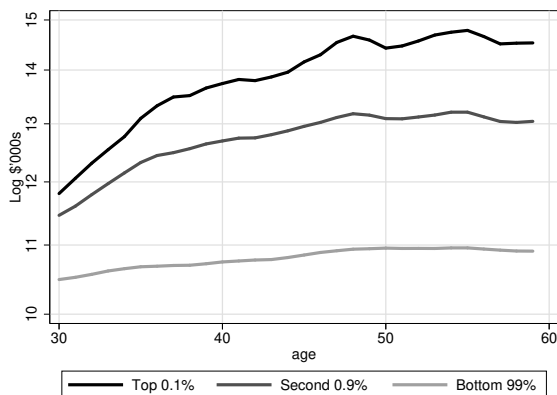
TABLE F.2 – Gender differences among lifetime top earners

	<u>Top 0.1%</u>	<u>Second 0.9%</u>	<u>Bottom 99%</u>
<u>Panel A: Overall top earners</u>			
Female worker share	9%	11%	49%
Female earnings share	9%	10%	38%
Log mean gender gap	-0.01	0.06	0.46
Log p50 gender gap	-0.05	0.05	0.48
No. working years gender gap	0.40	0.20	0.90
<u>Panel B: Gender-specific top earners</u>			
Male threshold (\$'000)	27,512	9,320	
Female threshold (\$'000)	9,487	3,828	
Log mean gender gap	1.18	0.97	0.52
Log p50 gender gap	1.16	0.96	0.49
No. working years gender gap	-0.19	-0.01	0.94

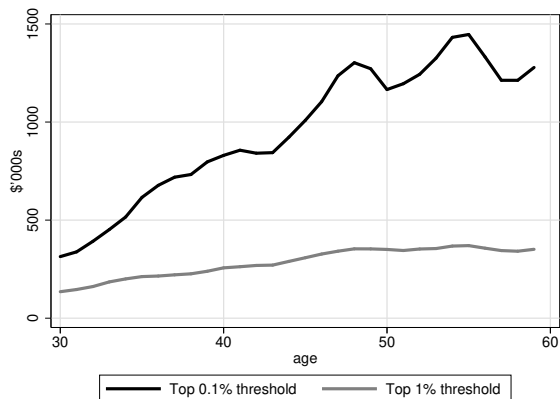


FIGURE F.1 – Age profiles by 30-year top earning groups

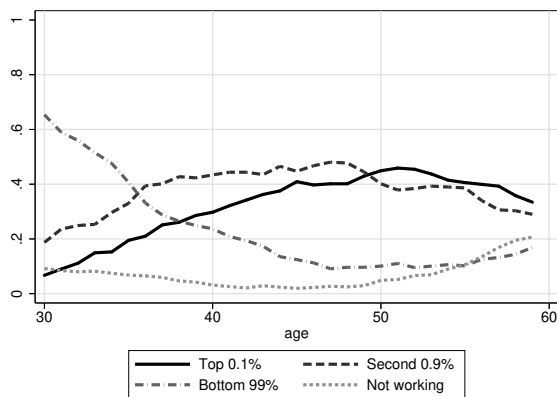
(A) Mean earnings by age



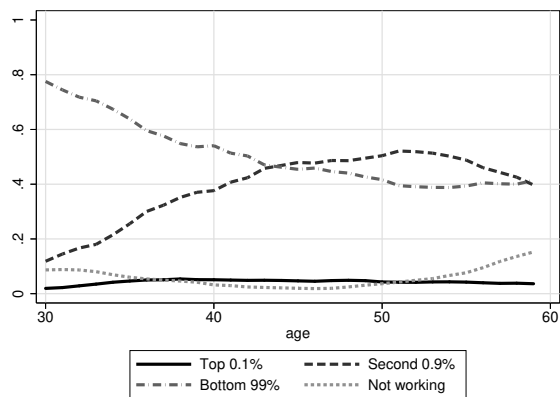
(B) Age-specific top-earning thresholds



(C) Location of lifetime top 0.1pct in age-specific distributions



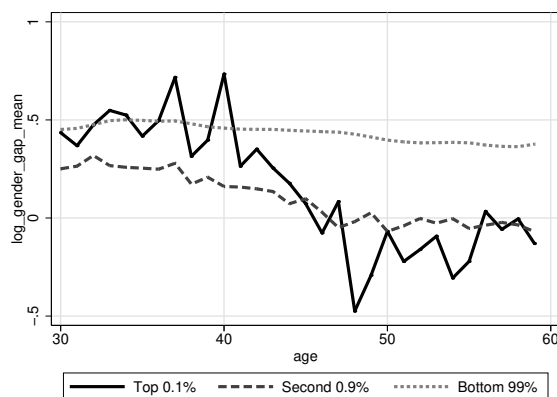
(D) Location of lifetime top 1pct in age-specific distributions



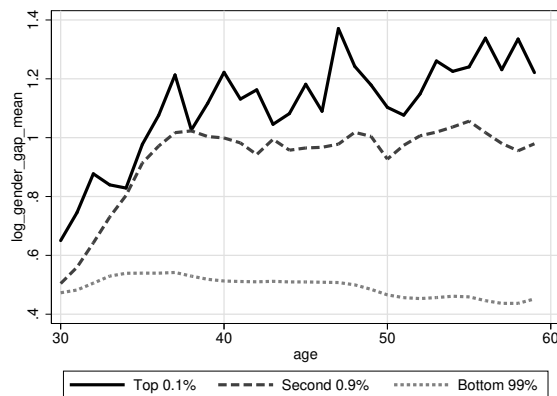
Notes: Figures refer to individuals from the 1951, 1952, and 1953 birth cohorts. Age-specific top-earning thresholds and groups are computed using only these three cohorts.

FIGURE F.2 – Gender gap among 30-year top earners by age

(A) Overall lifetime top earners

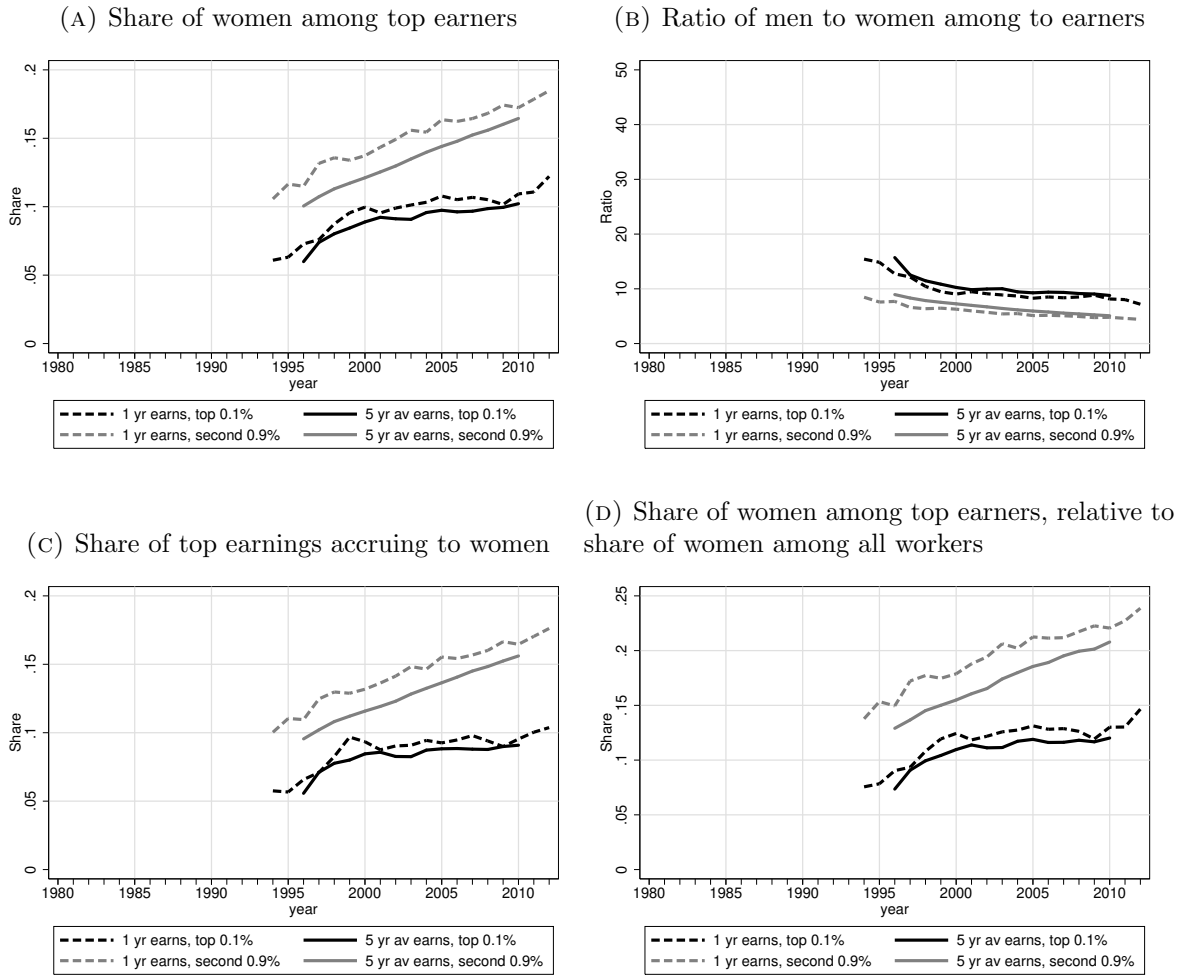


(B) Gender-specific lifetime top earners



Notes: Figures refer to individuals from the 1951, 1952, and 1953 birth cohorts. Age-specific top-earning thresholds and groups are computed using only these three cohorts. Figures show mean gender gap in each part of the earnings distribution.

FIGURE G.1 – Gender composition of top earners

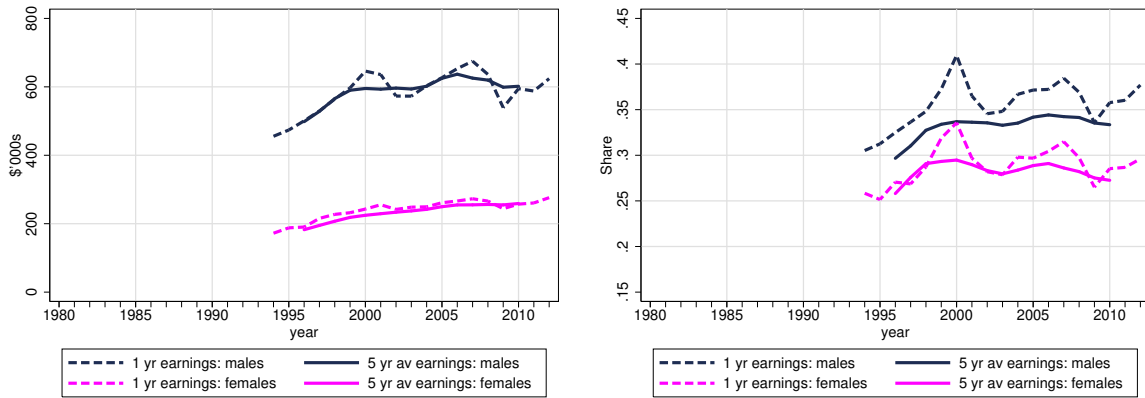
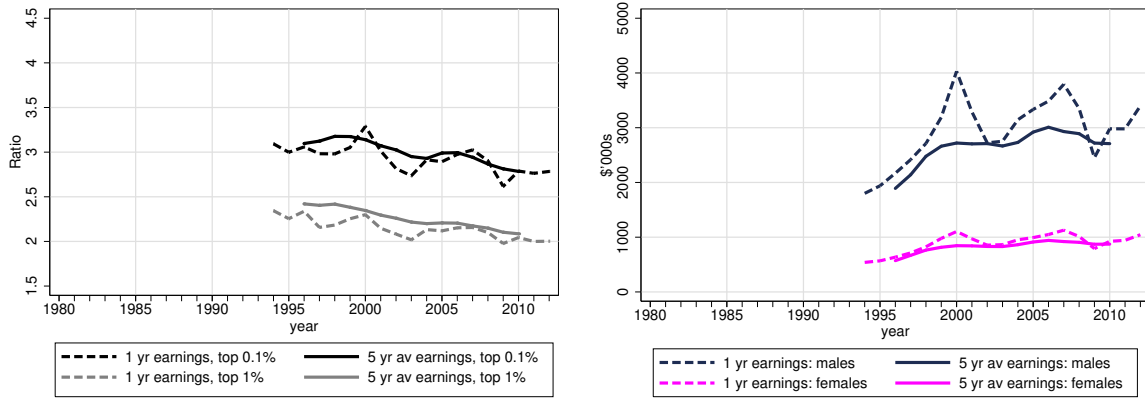


## G Including self-employment income

This appendix contains deleted figures from the main text, constructed using a definition of income that includes both wage and salary earnings, and earnings from self-employment income.

FIGURE G.2 – Male top earners versus female top earners

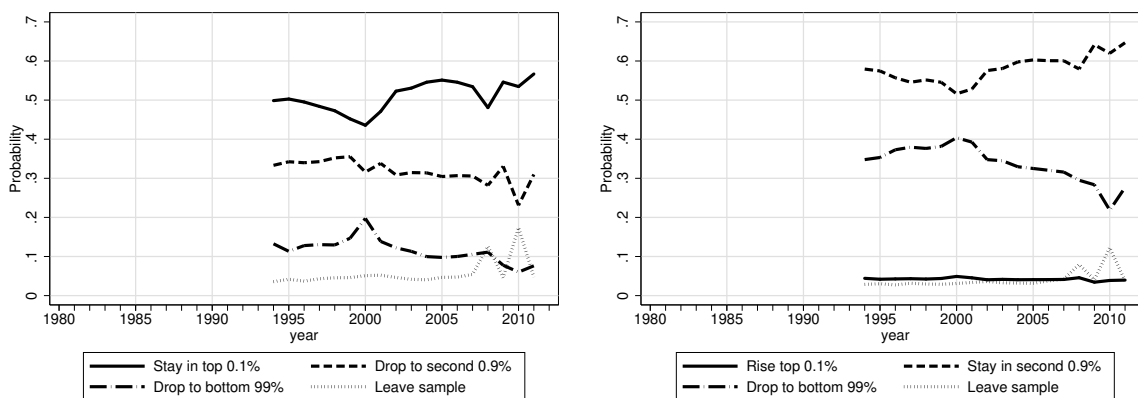
(A) Ratio of male to female top earning threshold top 0.1pct of women  
 (B) Average earnings among top 0.1pct of men and



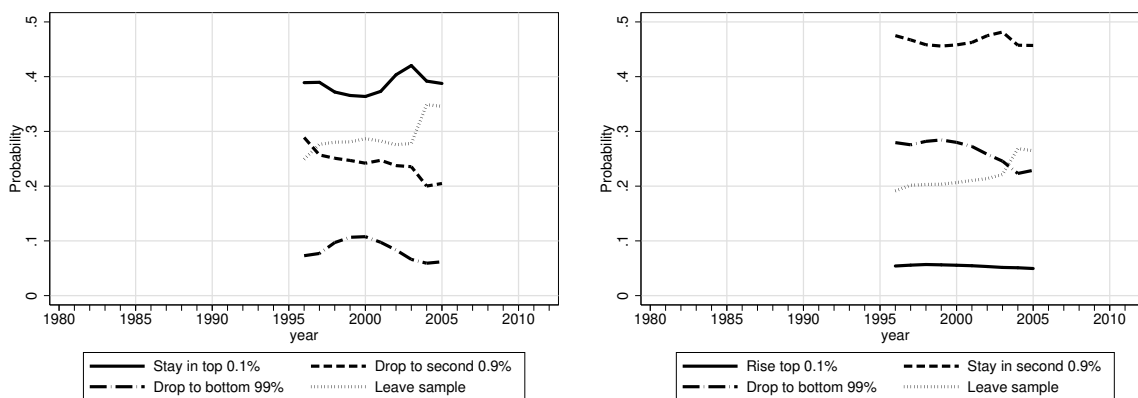
(C) Average earnings among second 0.9pct of men and second 0.9pct of women  
 (D) Share of top 0.1pct earnings in top 1pct earnings for men and women

FIGURE G.3 – Transition probabilities in and out of top percentiles of earnings distribution

(A) 1-year transition prob. for annual earnings, top 0.1pct (B) 1-year transition prob. for annual earnings, second 0.9pct



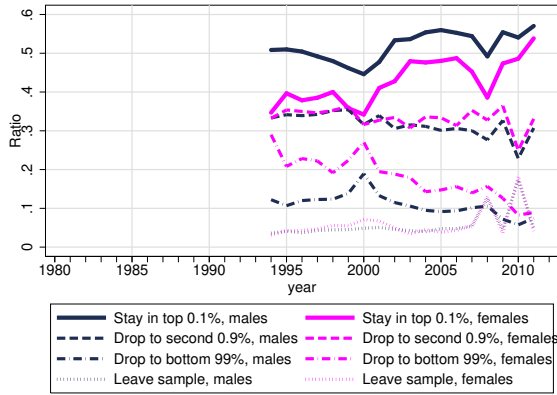
(C) 5-year transition prob. for 5-year earnings, top 0.1pct (D) 5-year transition prob. for 5-year earnings, second 0.9pct



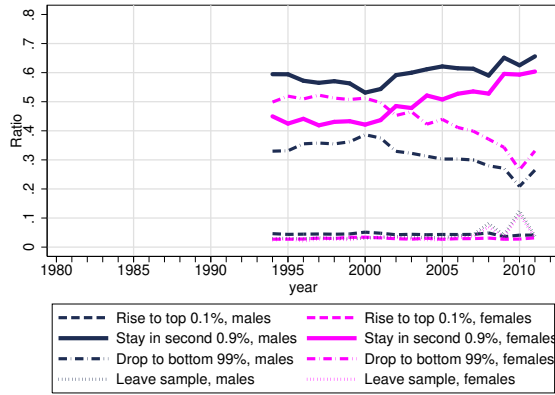
Notes: These figures show the probability that a top earner based on average earnings over the period  $t - 2, \dots, t + 2$  is a top earner based on average earnings over the period  $t + 3, \dots, t + 7$ .

FIGURE G.4 – Transition probabilities in and out of top percentiles of earnings distribution, by gender

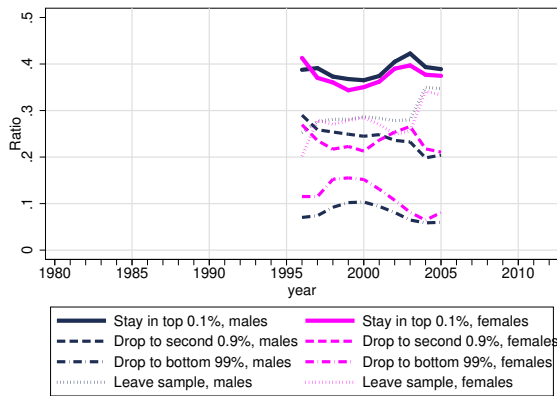
(A) 1 year transition probabilities for annual earnings, top 0.1pct



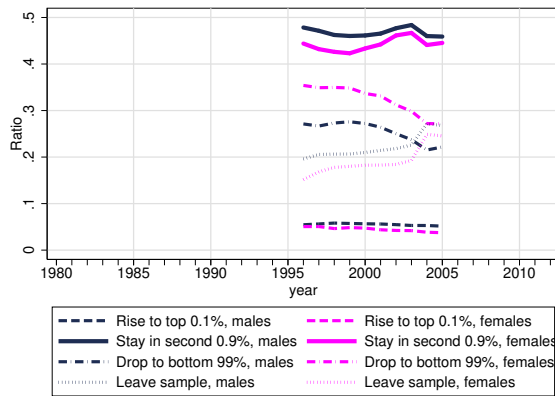
(B) 1 year transition probabilities for annual earnings, second 0.9pct



(C) 5 year transition probabilities for 5-year earnings, top 0.1pct



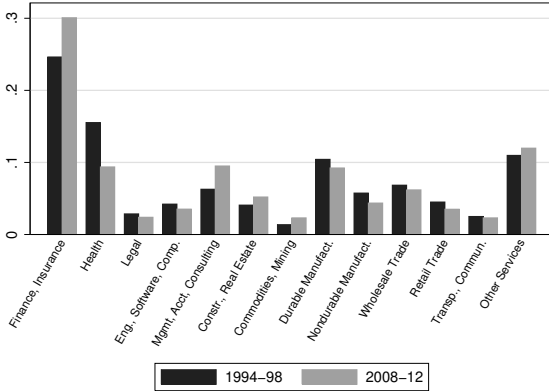
(D) 5 year transition probabilities for 5-year earnings, second 0.9pct



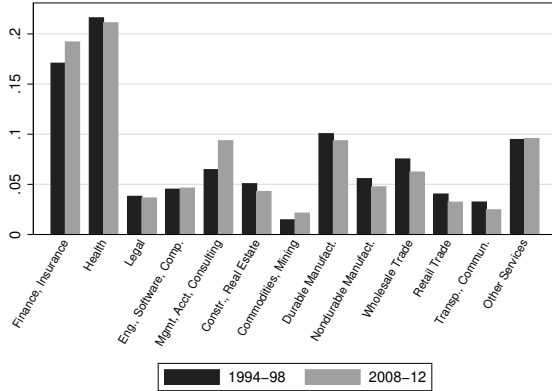
Notes: These figures show the probability that a top earner based on average earnings over the period  $t - 2, \dots, t + 2$  is a top earner based on average earnings over the period  $t + 3, \dots, t + 7$ , separately for male top earners (blue) and female top earners (pink).

FIGURE G.5 – Industry composition of top earners, 5-year average earnings

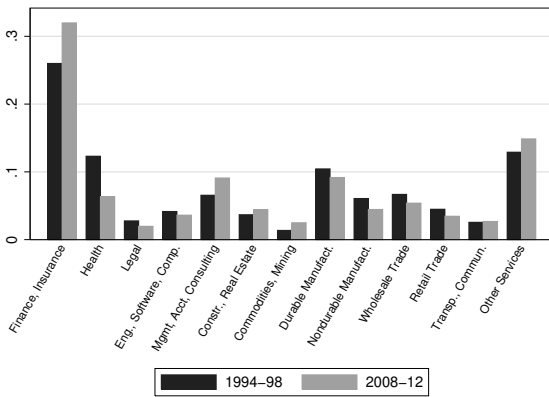
(A) Population shares, top 0.1pct



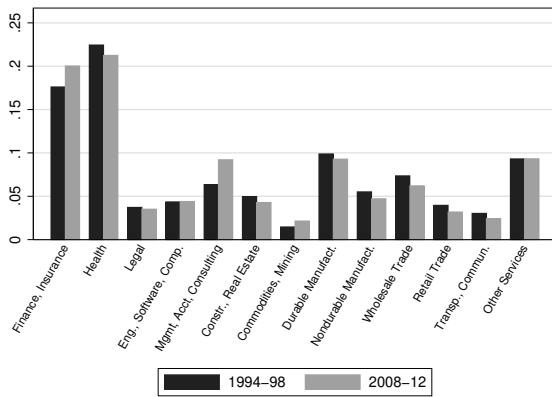
(B) Population shares, second 0.9pct



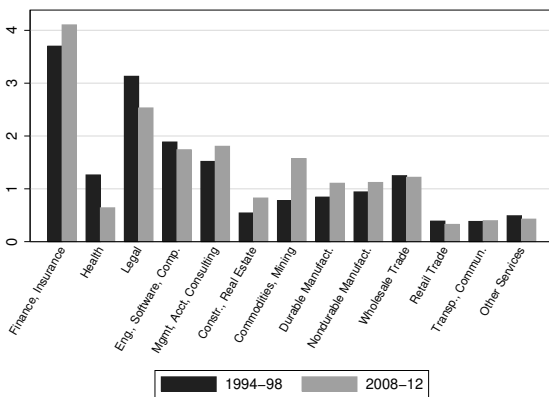
(C) Earnings shares, top 0.1pct



(D) Earnings shares, second 0.9pct



(E) Population shares, top 0.1pct relative to bottom 99 percent



(F) Population shares, second 0.9pct relative to bottom 99 percent

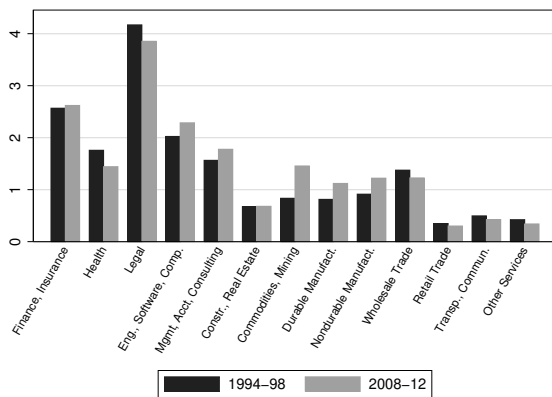
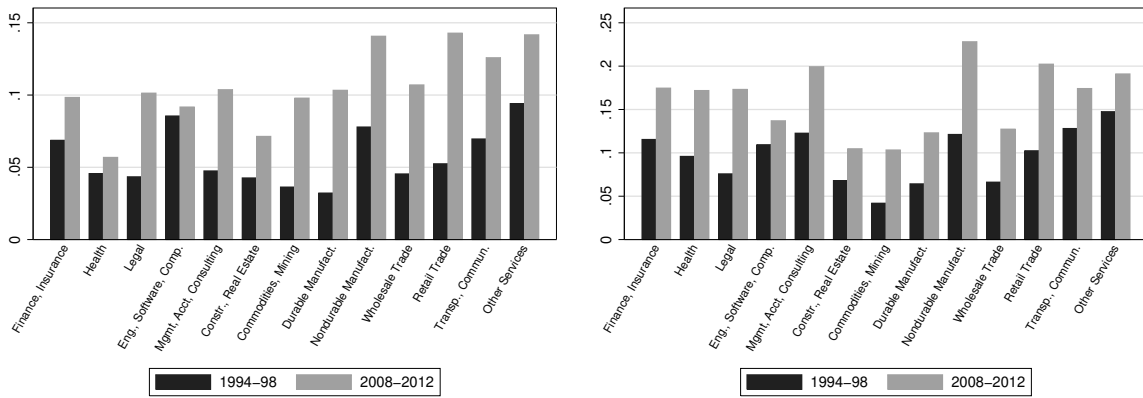


FIGURE G.6 – Top earners by industry and gender, 5-year average earnings

(A) Share of women by industry within top 0.1pct (B) Share of women by industry within top 0.9pct



(C) Industry shares by gender within top 0.1pct, 2008-12 (D) Industry shares by gender within second 0.9pct, 2008-12

