



## Research and Statistics Note

No. 2008-04

September 2008

---

# Estimating the First Instance of Substantive-Covered Earnings in the Labor Market

by Michael Compson\*

---

## Introduction

---

It is well documented that the average age at which individuals begin to receive retirement benefits has decreased substantially over the life of the Social Security (or, OASDI—Old-Age, Survivors, and Disability Insurance) program. However, very little information is available about the “front end” of an individual’s work history: The average age at which individuals first generate substantive earnings in the labor force, which might be measured as full-time employment or as a predetermined level of earnings. If the age at which individuals first generate substantive earnings and thus, make substantive contributions to the program has risen over time, then this could have an impact on the program.

Social Security’s administrative earnings data provide a unique source of information about an individual’s work history in covered employment.<sup>1</sup> Specifically, the data identify the point at which an individual first contributes to the Social Security program, and for the most part, contain detailed annual earnings information throughout an individual’s covered work history. Although the administrative data do provide annual total earnings information, they do not reveal the time period in which an individual first works full-time for a complete or substantial portion of a calendar year. As a result, predetermined annual earning levels are used as proxy measures for the first instance of substantive earnings.

The purpose of this note is to estimate the age at which individuals first generate a substantive level of earnings in the labor market and determine whether this age has changed over time. The operational question is what level of covered earnings is the best proxy for measuring substantive earnings? The analysis compares an individual’s reported covered earnings with three proxy measures of substantive activity in the labor market. The estimates are then aggregated by birth cohorts (persons born within the same time period) to determine if the age when first meeting the various proxy thresholds has changed over time.

---

<sup>1</sup> Covered earnings and earnings are used interchangeably throughout the analysis and refer to earnings from employment covered under the Social Security program.

---

\*Division of Earnings, Statistics, and Analysis, Office of Research, Evaluation, and Statistics, Office of Retirement and Disability Policy, Social Security Administration.

*Acknowledgments:* I would like to thank John Hennessey for all of his help in retrieving and processing data within SAS Enterprise Guide. In addition, I would like to thank Russell Hudson, Scott Mueller, Jeffrey Hemmeter, Jae Song, Paul Davies, Alexi Strand, Fred Galeas, Greg Diez, and other reviewers for their comments and suggestions on previous drafts of this note.

---

This note should be viewed as a preliminary investigation into the ability to use the Social Security Administration's (SSA's) earnings data to determine an individual's first instance of substantive-covered activity in the labor market and the first instance of substantive contributions to the program. There are a number of economic, technical, and programmatic issues that have a direct impact on the earnings data and hence, the results of this analysis. Additional detailed analysis of some of these issues is clearly warranted, but beyond the limited scope of this analysis.

### ***Proxy Measures for Substantive Activity in the Labor Market***

---

By definition, the amount an individual contributes to the Social Security program is directly related to his or her covered earnings. The question is what level of covered earnings constitutes substantive earnings in the labor market and the corresponding level of substantive contributions to the Social Security program? This analysis compares an individual's covered earnings in a given year to three proxy measures that estimate substantive activity in the labor market: (1) the national average wage index (AWI), (2) low-wage earners, and (3) very low-wage earners (see Chart 1).<sup>2</sup>

#### ***National Average Wage Index***

The national average wage index is a wage time series generated by SSA's Office of the Chief Actuary (OCACT). The AWI is used for a number of programmatic purposes, such as indexing a person's benefits and computing the OASDI contribution and benefit base. The AWI for a given year is generated by comparing the average wages subject to federal income tax reported on Form W-2 over a 2-year period. Specifically, the average wage for a given year is generated by dividing net compensation (wages and tips *plus* contributions to deferred compensation plans *less* distributions already included in reported wages) by the total number of workers. The resulting average wage is compared with the previous year's average wage to determine if there has been any increase over the year. The final step in the process is to multiply the previous year's AWI by the growth rate in average wages.

#### ***Low-Wage Earners and Very Low-Wage Earners***

For research and policy evaluation purposes, OCACT has also created several hypothetical workers with their level of covered earnings based on the AWI (Clingman and Nichols 2004). These wage earners are assumed to have "steady" annual earnings equal to a fixed percentage of the AWI throughout their work history. This analysis adopts the annual values of these hypothetical wage earners for proxy measures of substantive activity in the labor market. The annual value of substantive earnings for the low-wage earners' proxy is equal to 45 percent of the AWI. The annual value of the proxy for the very low-wage earners is equal to 25 percent of the AWI.<sup>3</sup>

### ***Earnings Data***

---

In general, Social Security's administrative earnings data contain the historical annual values of covered earnings over an individual's work history. As previously noted, the data do not contain any direct measures of an individual's participation in the workforce, and therefore it is not possible to determine

---

<sup>2</sup> Two additional proxy measures were considered but rejected for presentation in this note: (1) the full-time, full-year value of the minimum wage and (2) per capita income. The ad hoc manner in which the minimum wage has been changed over time makes the interpretation of the age estimates somewhat problematic. The results for the per capita income proxy were very similar to the results for the AWI so they were not included.

<sup>3</sup> The analysis also includes the average age at which individuals first report any amount of covered earnings as a benchmark for the three proxy measures of substantive activity in the labor market.

for any given earnings year whether the reported annual earnings were generated over a 3-, 6-, 9-, or 12-month period. Further, the data do not contain information regarding the individual's wage rate, hours worked, weeks worked, full-time or part-time working status, or full-year or part-year participation. These data limitations are important to remember throughout the analysis and explain why full-time work status or duration of the work period over the calendar year cannot be used as measures of substantive earnings in the workforce.

The earnings data used in this analysis are from the Continuous Work History Sample (CWHS)—a series of annual files that were specifically created to support research and statistical analysis. The annual files in the CWHS are based on a 1 percent sample of all Social Security numbers ever issued and include data from the Master Earnings File (MEF) and the Master Beneficiary Record (MBR). The CWHS provides a longitudinal history of earnings for the sample population, and the sample is updated annually to incorporate newly issued Social Security numbers.

The data used in the analysis are from the CWHS active file for tax year 2003 (earnings reported for tax year 2003). In general, the active file includes anyone with earnings from any employment, independent of whether the earnings are covered under the Social Security program. The active file includes individuals who reported covered earnings at some point during the 1951–2003 period and noncovered earnings reported sometime during the 1978–2003 period. Given the limited time frame of reported noncovered earnings and the limited scope of the current study, this analysis is restricted to workers in the active file who reported covered earnings at some point during their work history. The 2003 active file contains historical earnings records for 3,188,426 individuals who reported covered earnings at some point in time. The key variables used in the analysis are as follows:

1. The individual's year of birth;
2. The individual's sex;
3. The year when an individual's first report of employment with Social Security taxable earnings occurs. The values for this variable range from 1937 through 2003. This field includes all covered wages and salary and self-employed income; and
4. The nominal value of annual total Social Security–covered earnings from 1951 through 2003. This includes wages from all employers and earnings from self-employment. It is possible to have negative values for this variable (if self-employment earnings are negative).

A two-step process determines the individual's age when he or she first meets the proxy thresholds. The first step begins with the individual's first year of covered earnings and compares it with the corresponding proxy value for that year. This process continues for subsequent years until the individual's covered earnings exceed the threshold for a given proxy measure or there are no remaining years of earnings data to compare. The value of that year is assigned to a variable called "first-year threshold" for a proxy measure. If the earnings never exceed the proxy threshold, a value of zero is assigned to the first-year-threshold variable. The second step determines the individual's age when first meeting the threshold, by subtracting the individual's year of birth from the first-year threshold. This method for estimating an individual's age is not exact because we do not have access to one's date of birth, and the

earnings data are reported on an annual basis. Therefore, the estimates for an individual's age when first meeting the threshold could be plus or minus one year.<sup>4</sup>

Several groups of individuals were dropped from the analysis. The largest group removed from the 2003 active file comprises individuals who were born before 1946 and those born after 1986. There are two reasons why the 1946 birth cohort is the first one included in this analysis: (1) 1946 is typically considered the first year of the baby boom generation, and (2) there were significant expansions in the types of employment covered under the Social Security program in 1950 and 1955, and this could have a significant impact on the estimates for the first instance of substantive contributions to the program during those years. Individuals in the 1946 birth cohort were age 16 (typically considered the first year when individuals can enter the labor force) in 1962, several years beyond the transitions that one would expect because of the change in Social Security coverage in 1955. The last birth cohort included in the analysis is 1986—the last birth cohort where everyone would be at least age 16 in 2003. Limiting the sample to individuals who were born after 1945 and before 1987 removed 1,443,911 individuals, or approximately 45 percent of the individuals in the 2003 active file who reported covered earnings.

As previously noted, the active file contains an indicator of the first report of earnings covered under the Social Security program. Any of the individuals whose first report of covered earnings occurred between 1937 and 1950 had to be removed from the sample because there are no values for annual income during those years. This makes it impossible to determine if and when their annual earnings before 1951 exceeded the proxy measures for substantive activity in the labor market. This adjustment removed an additional 12,069 records of individuals with covered earnings, less than 1 percent of the subset of the 2003 active file.

A third group of individuals were removed from the analysis because of problems incurred when estimating their age when they first reported covered earnings. The analysis is limited to those whose age when first reporting covered earnings was between 11 and 61. This was done to remove anomalous age results among those appearing too young and those old enough to retire. This adjustment removed an additional 7,515 individuals from the analysis.

In total, these adjustments removed 1,463,495 records or nearly 46 percent of the records in the 2003 active file of workers reporting covered earnings. The final sample used in the analysis is still very large, consisting of 1,724,931 workers with covered earnings.

### ***Caveats to Consider***

---

There are a number of programmatic and data collection issues that have a major impact on the data used in the current analysis. The following discussion highlights three programmatic and data issues critical to the process of estimating the first instance of substantive-covered earnings in the labor market. In addition, a number of socioeconomic issues are identified that are likely to have an impact on the estimates.

First, it is critical to note that the analysis relies solely on covered earnings data to estimate an individual's first instance of substantive activity in the labor market. As such, the results do not reflect the experience of all workers because individuals who have earnings in noncovered employment will have

---

<sup>4</sup>For example, consider someone who was born on January 1, 1946. If the first year was 1966 when his or her earnings were above the very low-wage threshold, the individual was clearly age 20 when meeting the threshold. However, consider the case of an individual who was born on December 26, 1946. In the current analysis, this individual would be assigned to age 20 when he or she met the threshold. However, one could reasonably argue that this individual was really only age 19 when meeting the threshold.

an incomplete time series of earnings.<sup>5</sup> Although this problem is somewhat mitigated as coverage under the program has become nearly universal, it is very real for the early years of the estimates. Nevertheless, relying solely on covered earnings is appropriate for this analysis because covered earnings provide the revenue stream for the Social Security program.

Second, there has been an expansion over time in the types of employment covered under the program.<sup>6</sup> The impact of an expansion in coverage in estimating the age when first meeting a threshold will depend on the number of workers affected by the change in coverage. The larger the number of workers who are converted from noncovered to covered employment in a given year, the larger the bias in the estimates. Consider the case of a 40-year-old individual who has earnings from noncovered employment above the low-wage earner threshold for the past 15 years. If coverage was extended to his or her type of employment when reaching age 40, the first instance of covered earnings above the low-wage earner threshold would be at age 40, not 25 years old.

Third, a change in 1978 from quarterly reporting to annual reporting of earnings data represents a significant loss in SSA's ability to track an individual's participation in the labor force over time. At a minimum, quarterly data provided some insight into the individual's activities in the labor market. For example, if an individual reported earnings in all four quarters, one could reasonably argue that he or she was active in the labor market year round. Under annual reporting it is not possible to determine the pattern or duration of time during which the earnings were generated, for example, over a 6-week or 3-, 6-, or 12-month period.

There are also a number of socioeconomic issues that may have an impact on an individual's age when he or she first reports substantive activity in the labor market. For example, the annual inflow of immigrants (legal and illegal) into the United States has increased substantially since the early 1960s. In addition, the number of immigrants as a percentage of the population has also increased fairly dramatically over the same time period. The impact of this growing influx of immigrants on the estimates of the age at which individuals first make substantive contributions to the Social Security program is uncertain.

To the extent that immigrants are more likely to enter the United States at young ages and work the majority of their careers in the country, their effect on the results may be relatively small. But immigrants who enter the United States later in life (for example, under family reunification provisions), could skew the results substantially. Those immigrants would have their first instance of Social Security-covered labor market activity at older ages, even though their actual (but unobserved) labor market entry occurred at a younger age in their home country. These effects may change over time and across birth cohorts because of changes in immigration patterns.

Other socioeconomic issues include, but are not limited to, more individuals pursuing higher education (particularly those seeking deferment from the draft during the Vietnam War), the business cycle, the increase in female labor force participation over time, and the demographic changes that altered the age distribution (for example, the baby boom population). Although all of these are important issues, a detailed analysis is beyond the limited scope of the current analysis.

---

<sup>5</sup> As previously mentioned, SSA's administrative data contain annual measures of noncovered earnings from 1978 to 2003, but they were not used in this analysis.

<sup>6</sup> See *Annual Statistical Supplement to the Social Security Bulletin*, 2005, Table 2.A1, for a listing of provisions that have affected covered employment and self-employment. The largest increases in coverage occurred in 1950 and 1955.



## ***Results of Measuring Trends in Initial Substantive Contributions***

---

One way of measuring trends in the first instance of substantive contributions is to estimate the average age at which individuals in a given birth cohort first meet the proxy thresholds. These estimates reveal a consistent downward trend across cohorts (Chart 2). The average age at which individuals born in 1946 first report earnings that exceed the average wage index threshold is age 29 compared with age 22 for those born in 1980. The decline for the remaining thresholds is not as dramatic, but it does exhibit a clear downward trend. Even the first report of covered earnings decreases by about 3 years across birth cohorts.

At first glance, this result seems to contradict the hypothesis that the pursuit of higher education or technical training would lead to an increase in the average age when individuals begin making substantive contributions to the Social Security program. However, the average age at which individuals first meet the thresholds decreases across subsequent birth cohorts by definition because: (1) the maximum age that individuals can first meet a given threshold decreases across subsequent cohorts, and (2) many individuals in the later cohorts have not been in the labor market long enough to meet the various thresholds, essentially resulting in unobserved behavior or censored data.

The maximum age that someone born in 1950, 1960, 1970, and 1980 can attain by 2003 (the final year of data) is 53, 43, 33, and 23, respectively. This explains the relatively steep decline in the average age when first reporting covered earnings or first meeting the thresholds for those born after 1970. By definition, the average age when individuals born in 1980 can first meet the threshold is 23 or younger, a classic censored data problem. As a result, the average age for this cohort has to be 23 or younger.

Another way of measuring trends in the first instance of substantive contributions is to estimate the percentage of individuals in a given birth cohort who meet the thresholds. These estimates decrease fairly steeply after the 1964 birth cohort (Chart 3). For example, the percentage of individuals meeting the low-wage threshold decreases from 85 percent for the 1965 cohort to just below 50 percent for the 1980 cohort. This indicates that many of the individuals in the later cohorts have not had enough time in the labor force to meet the thresholds. Estimating the average age when individuals meet the thresholds by birth cohort results in unobserved behavior or a right censoring of the data.

To eliminate the censoring problem, the cumulative percentage of individuals meeting the various thresholds for six specific age categories is calculated for individuals who are younger than age 18, 20, 22, 24, 26, and 28. The first cohort for each of these age categories begins with 1946 and ends in the year when the maximum age for the category is attained. For example, the category aged 20 or younger includes those who are either 16, 17, 18, or 19, and the birth cohorts run from 1946 through 1984. This eliminates the possibility of including individuals who have not reached age 20 as of 2003 and allows valid comparisons across birth cohorts.

### ***Very Low-Wage Earner Threshold***

---

The first results considered are those generated using the very low-wage earner proxy for substantive earnings. The very low-wage earner is the lowest earnings level of all three measures for substantive earnings (equals 25 percent of the average wage index).

#### ***Aggregate Estimates***

The cumulative percentage of individuals meeting the very low-wage thresholds across birth cohorts is presented in Chart 4. The percentage of all individuals who were younger than age 18 when they first met the very low-wage earner threshold begins at nearly 5 percent for those born in 1946 and peaks at nearly 11 percent for the 1962 birth cohort. This percentage drops significantly to 4 percent for the 1966

birth cohort and decreases even lower to approximately 3 percent for the 1975 and subsequent birth cohorts. The decrease in the percentage of individuals who were younger than age 20 when first meeting the threshold is even larger, from 46 percent for the 1960 cohort to just 26 percent for the 1984 birth cohort. The percentage of all individuals who were younger than age 22 when first meeting the threshold peaks at nearly 63 percent for those born in 1959 compared with just 53 percent for the 1982 cohort. It is very clear that the percentage of individuals meeting the very low-wage earner threshold for the three lowest age categories decreases across subsequent birth cohorts.

Interestingly, the cumulative percentage of individuals who were younger than age 24 when first meeting the threshold has remained fairly steady across birth cohorts. The percentage drops slightly below 70 percent for the 1960–1974 cohorts, but then returns to previous levels after 1974. However, the cumulative percentages presented in Chart 4 mask some underlying differences occurring for different age categories. For example, the percentage of individuals who were younger than age 18 when they first met the very low-wage earner thresholds decreases slightly from 4.5 percent for the 1946 cohort to just over 3 percent for the 1980 cohort. But in tabulations not presented here, the percentage of individuals who were age 18 or 19 when they first met the threshold decreases from 34 percent for the 1946 cohort to 27 percent for the 1980 cohort. The bulk of the 8.5 percentage-point decrease for individuals younger than age 20 occurs for those who were either age 18 or 19.

The 8.5 percentage-point decrease for those first meeting the threshold when they were younger than age 20 is offset by an increase for those aged 20–23. Specifically, the percentage of individuals who were age 20 or 21 when first meeting the threshold increases by slightly more than 6 percentage points. The percentage of individuals who were age 22 or 23 when they first met the threshold increased by 2 percentage points. These results reveal a clear decrease in the cumulative percentage of individuals who were younger than age 22 when meeting the very low-wage thresholds across birth cohorts. The bulk of this decrease occurs for individuals who were aged 16–19.

### ***Estimates by Sex***

The results presented above reflect changes for birth cohorts as a whole; they do not distinguish by sex. Given the changes in labor force participation among female and male workers over the past several decades, an important question is whether there are any discernable differences by sex or changes in the age when female and male workers first meet the very low-wage earner threshold. One way to answer this question is to present the number of female and male workers who first met the very low-wage earner thresholds at a given age, as a percentage of the total number of sample members in the birth cohort. Chart 5 presents these results for female workers (top panel) and male workers (bottom panel).

The top panel reveals clear trends in the percentage of female workers first meeting the very low-wage earner thresholds across subsequent birth cohorts. Approximately 2 percent of the female sample born in 1946 first met the very low-wage earner threshold when they were younger than age 18. This increased to a peak of 7.5 percent for those born in 1962 and then decreased to 2.5 percent for the 1974 and subsequent birth cohorts. The percentage of female workers who were younger than age 20 when first meeting the very low-wage earner threshold was 29 percent for those born in 1946. This increased to 38 percent for those born in 1961 and has subsequently decreased to approximately 23 percent for those born in 1984—a fairly substantial decrease over the years. Although there is considerable fluctuation in the percentage of the female sample meeting the threshold before age 22, the percentage is virtually the same for those born in 1946 and 1982. Finally, there is a clear upward trend in the percentage of female workers meeting the threshold for the top three age categories (under 28, 26, and 24 years of age), bringing them very close to the levels achieved by their male counterparts for the same age categories.

The pattern for the lowest two age categories (under 18 and 20 years of age) for male workers is very similar to the results for female workers except that the magnitude of the decline is much larger for the male sample. Unlike the case for their female counterparts, the percentage of male workers who were younger than age 22 when first meeting the threshold exhibits a clear and substantial decrease across subsequent birth cohorts. There are also differences between male and female workers for the highest three age categories. The change in the percentage of male workers first meeting the threshold before they were ages 28 and 29 is relatively small across birth cohorts.

One can reasonably argue that the aggregate decrease in the percentage of individuals meeting the thresholds when they were younger than age 22 is primarily driven by the decrease in the male sample meeting the threshold. It is also clear that the upward trend in the aggregate percentage of individuals meeting the threshold when they were younger than ages 28 and 26 is the result of the increase for female workers. The relatively flat line for the cumulative percentage of individuals who were younger than age 24 is the result of the decrease for male workers meeting the threshold across subsequent birth cohorts, offset by the increase for their female counterparts.

### ***Low-Wage Earner Thresholds***

---

The second results considered are those generated using the low-wage earner proxy for substantive earnings. The low-wage earner is the middle earnings level of the three measures for substantive earnings (equals 45 percent of the average wage index).

#### ***Aggregate Estimates***

As expected, the cumulative percentage of individuals meeting the low-wage earner thresholds (Chart 6) is lower across the board relative to the very low-wage earner thresholds. Consider the extreme age categories—those younger than age 18 and those younger than age 28. The percentage of individuals who were younger than age 18 when they first met the low-wage earner threshold is always less than 2 percent and does not exceed 0.4 percent from 1966 forward. This is much less than the 3 percent to 4 percent meeting the very low-wage earner threshold. The cumulative percentage of those younger than age 28 when first meeting the low-wage earner threshold ranges from 70 percent to 73 percent across subsequent birth cohorts. This too is much lower than the 80 percent to 88 percent range for the very low-wage earners.

The percentage of individuals who were younger than age 20 when they first met the low-wage threshold peaks at just over 26 percent for those born in 1955 and decreases substantially to just 8 percent for those born in 1980. There is also a sharp decline in the percentage of individuals younger than age 22 when first meeting the low-wage earner threshold, dropping from 46 percent for the 1959 cohort to only 29 percent for those born in 1980. Unlike the case for the very low-wage earner thresholds, the percentage of individuals first meeting the threshold when they were younger than age 24 drops from 61 percent for the 1946 cohort to just under 48 percent for the 1980 cohort.

A closer look at the underlying ages reveals minimal change in the percentage of individuals who were ages 16 and 17 when first meeting the thresholds across cohorts. In tabulations not shown, there is substantial fluctuation for the percentages of individuals who were either aged 20–21 or aged 22–23 when first meeting the low-wage earner threshold. Both of these age categories have essentially the same values for the 1946 and 1980 birth cohorts. However, there is a substantial drop in the percentage of individuals who were age 18 or 19 when first meeting the threshold. The percentage peaks at 25 percent for the 1955 birth cohort and drops to just less than 8 percent for those born in 1980.



It is interesting to note that there is only a small difference in the cumulative percentage of individuals younger than age 26 when first meeting the low-wage earner threshold across birth cohorts. Notice that this relative stability in the percentage of individuals meeting the threshold occurs at a higher age and lower percentage compared with the results for the very low-wage earners.

### ***Estimates by Sex***

As expected, the cumulative percentage of female sample members meeting the low-wage earner threshold is lower for each of the age categories (Chart 7, top panel) relative to the very low-wage earner threshold (Chart 5, top panel). For example, the percentage of female workers who were younger than age 18 when first meeting the low-wage earner threshold is essentially zero throughout all birth cohorts, much lower than the percentage for the very low-wage earner threshold. In addition, there is a clear downward trend in the percentage of female workers meeting the low-wage earner thresholds when they were younger than ages 20, 22, and 24. The percentage of the female sample meeting the threshold when they were younger than age 24 decreases from 50 percent for those born in 1946 to 42 percent for the 1980 birth cohort. In contrast, the trend for this age category for the very low-wage earner was relatively unchanged. The percentage of female workers meeting the threshold when they were younger than ages 26 and 28 increases across birth cohorts but at a much slower rate relative to the very low-wage earner threshold.

For the male sample, the decrease is much more pronounced. The percentage of male workers meeting the threshold when they were younger than age 20 decreases from a peak of 36 percent for the 1955 birth cohort to less than 10 percent for those born in 1984 (Chart 7, bottom panel). The percentage of male workers meeting the threshold when they were younger than age 22 peaks at nearly 57 percent for those born in 1953 and drops to just 30 percent for those born in 1982. The percentage of male workers who first met the threshold when they were younger than age 24 decreases by nearly 20 percentage points across birth cohorts. Although not as dramatic, there is a clear downward trend in the percentage of the male sample meeting the low-wage thresholds for those who were younger than ages 28 and 26.

As was the case with the very low-wage earner threshold, one can reasonably argue that the bulk of the decrease in the aggregate percentages of individuals meeting the low-wage threshold is driven mostly by the decreases in the male sample meeting the thresholds.

### ***Average Wage Index Threshold***

---

The final results considered are those generated using the average wage index proxy for substantive earnings. The average wage index is the highest earnings level of all three measures for substantive earnings.

### ***Aggregate Estimates***

Chart 8 reveals a dramatic decline in the cumulative percentage of individuals first meeting the AWI threshold for all age categories across subsequent cohorts. The compression of the distribution by age for the average wage index relative to the very low-wage and low-wage earners is quite striking. For example, less than 40 percent of individuals born in 1946 had earnings that exceeded the AWI threshold by the time they were age 28, and this decreases to less than 30 percent for the 1978 cohort. At the other end of the spectrum, virtually none of the individuals younger than age 18 had met the threshold across all birth cohorts.

In addition, there are only a few years in which a very small percentage of individuals younger than age 20 met the AWI threshold. The percentage of individuals who first met this threshold when they

were younger than age 22 decreases from less than 10 percent for those born in 1946 to 1 percent for the 1972 and subsequent birth cohorts. The decline is just as dramatic for each of the other age categories. It is very clear that the percentage of individuals meeting the AWI threshold has decreased substantially for all birth cohorts. This evidence suggests that it is taking subsequent birth cohorts much longer to meet the AWI threshold. For example, approximately 20 percent of the individuals born in 1946 reported earnings that exceeded the AWI threshold before they reached age 24. This percentage declines dramatically to less than 10 percent for the 1968 and subsequent birth cohorts.

### ***Estimates by Sex***

Chart 9, top panel, reveals that virtually none of the female sample members met the AWI thresholds when they were younger than ages 18, 20, and 22. The percentage of female workers meeting the thresholds when they were younger than age 24 decreases from approximately 10 percent for the 1946 birth cohort to 4 percent for those born in 1982. There is relatively little change in the percentage of female workers who were younger than age 26 or 28 when first meeting the AWI earner threshold. As a result, the aggregate decreases in the percentage of individuals meeting the threshold when they were younger than ages 22, 24, 26, and 28 are driven primarily by the steep decreases in the percentage of male workers meeting the threshold shown in the bottom panel of the chart.

### ***Summary and Conclusion***

---

The increasing life expectancy in the United States and the decline in the average age at which individuals begin receiving retirement benefits have implications for the Social Security program. However, very little is known about the “front end” of an individual’s work history. If individuals are delaying the age at which they begin to generate substantive-covered earnings in the labor market, then such changes may delay or reduce patterns and levels of payroll tax contributions to the Social Security trust funds.

This note is a preliminary investigation into the front end of an individual’s covered work history, defined here as when he or she first generates substantive earnings in the labor market. SSA’s administrative earnings data used in the analysis are a unique historical time series of an individual’s covered earnings.

In general, the findings support the proposition that the first instance of substantive-covered earnings is occurring later in life. As such, the associated larger contributions to the Social Security program appear to be occurring later now than for previous generations. This is evidenced by the decreasing percentage of individuals in the “younger” age categories meeting the very low- and low-wage earner thresholds. The percentage of individuals meeting the average wage index threshold decreases for each of the age categories considered. Overall it is clearly taking subsequent cohorts longer to meet the various thresholds used to approximate substantive earnings in the labor market. The results by sex seem to indicate that for the most part, the reduction in the percentage of male workers meeting the thresholds for the lower age categories is driving this result.

These findings are somewhat tempered by the unknown impact of a number of critical data issues that are identified in the analysis. In addition, the preliminary results raise a number of policy issues relating to total contributions to the program over time. Additional research on the potential impact of the limitations of the data used to estimate the first instance of substantive contributions to the Social Security program and other policy issues is warranted.

Future research might consider the following four issues:

1. Using statistical techniques to address the issue of censored data,
2. Using multivariate techniques to separate out underlying trends in the data,
3. Evaluating the impact of not including noncovered earnings in the analysis, and
4. Examining differences between immigrants and nonimmigrants in the average age when achieving the first instance of covered substantive activity in the labor market.

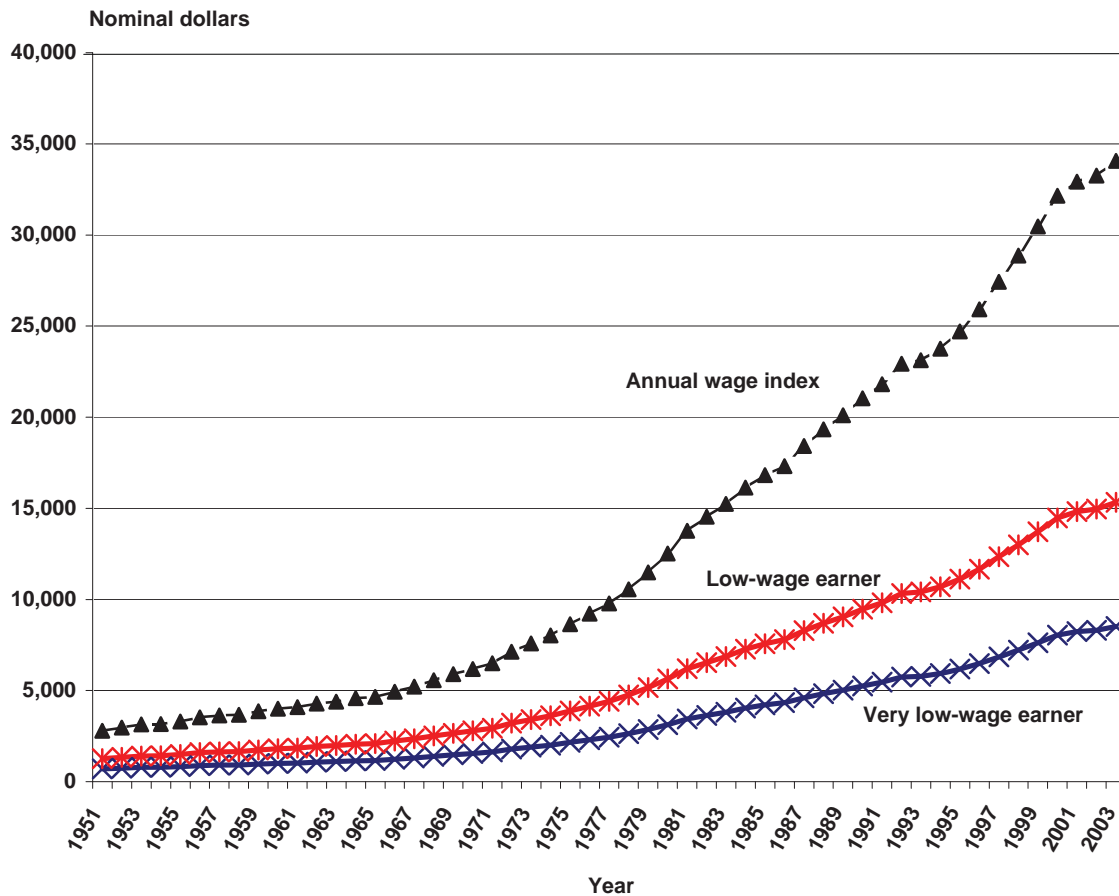
Finally, although this note estimates the individual's first instance of making substantive contributions to the Social Security program, it says nothing about the individual's contributions beyond this initial event. Further research might consider whether or not the decrease in the percentage of "younger" individuals meeting the various thresholds across birth cohorts has any implications for lifetime contributions to the Social Security program. Specifically—how do differences in the age at which individuals first meet a threshold affect lifetime-covered earnings and hence, contributions to the program over an individual's work history?

## **References**

---

- Clingman, Michael, and Orlo Nichols. 2004. Scaled factors for hypothetical earnings examples under the 2004 trustee report assumptions. Actuarial Note, No. 2004.3. Social Security Administration (December).
- Smith, Creston, M. 1989. The Social Security Administration's Continuous Work History Sample. *Social Security Bulletin* 52(10): 20–28.
- Social Security Administration. 2005. *Annual Statistical Supplement, 2004 to the Social Security Bulletin*. Washington, DC: Office of Policy, Office of Research, Evaluation, and Statistics.
- U.S. House of Representatives, Committee on Ways and Means, 2004. *2004 Green Book: Background material and data on the programs with the jurisdiction of the Committee on Ways and Means*. Washington, DC: Government Printing Office (March).

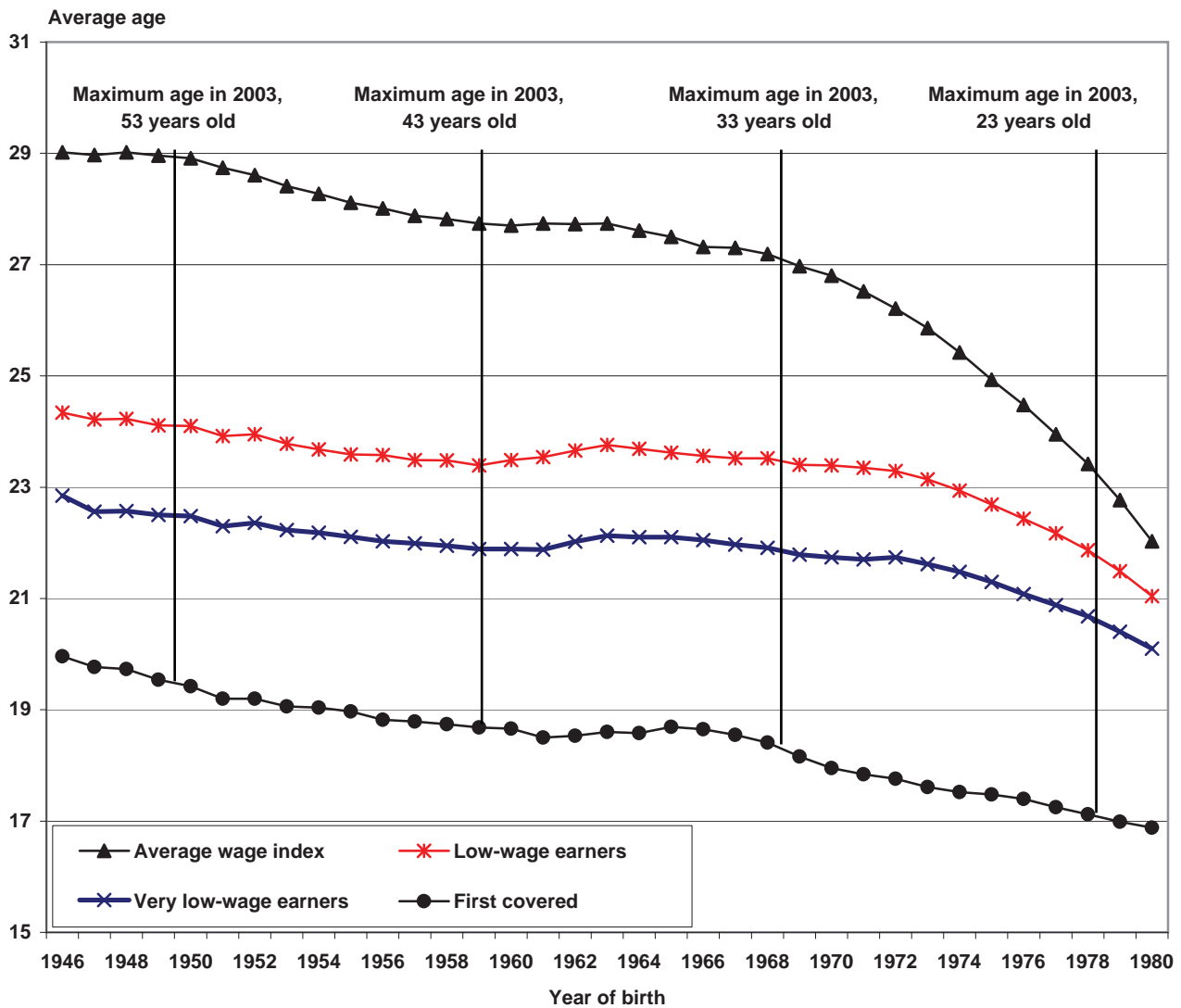
**Chart 1.**  
**Nominal values of the proxy measures for substantive activity in the labor market, by year**



SOURCE: Social Security Administration, Office of the Chief Actuary. Available at <http://www.ssa.gov/OACT/COLA/AWI.htm1#Series>.

NOTE: Low-wage earner equals 45 percent of average wage index (AWI) and very low-wage earner equals 25 percent of AWI. Both earner types are assumed to be "steady" workers earning a constant percentage of AWI throughout their careers.

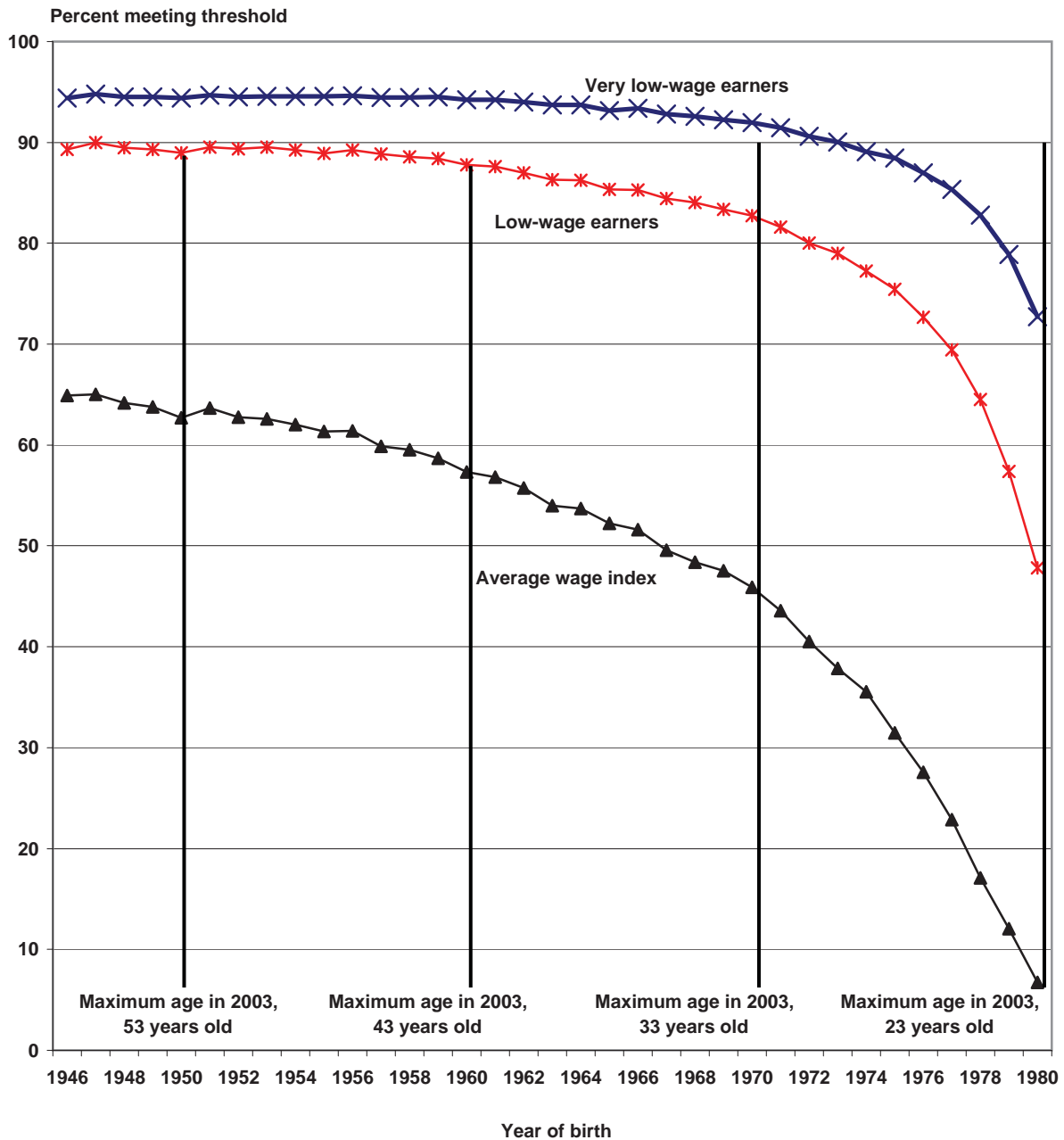
**Chart 2.**  
**Average age when first reporting covered earnings or first meeting the thresholds among those reporting covered earnings, by year of birth**



SOURCE: Author's calculations using the Continuous Work History Sample active file for tax year 2003.

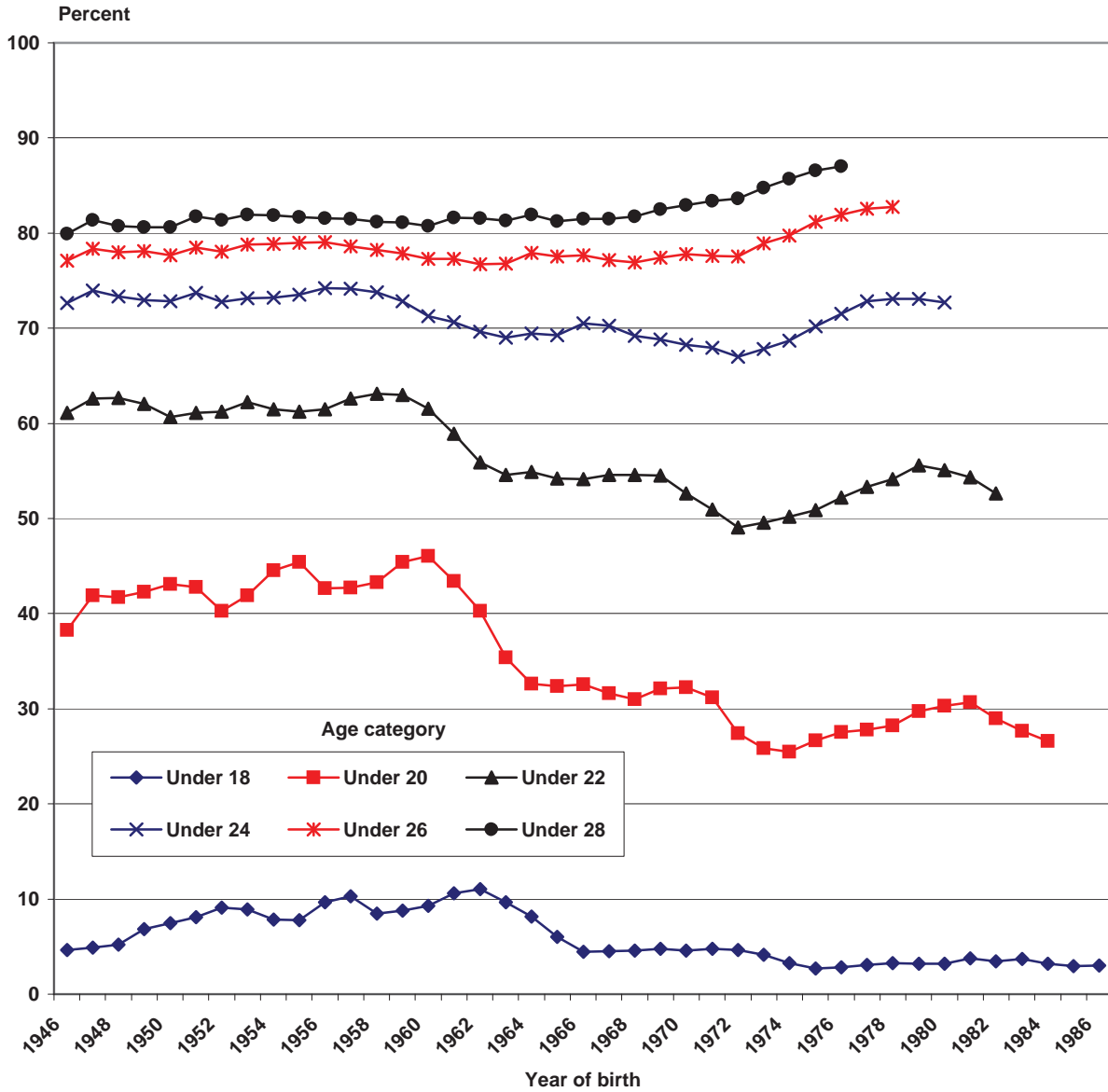


**Chart 3.**  
**Percentage of individuals meeting the thresholds among those reporting covered earnings, by year of birth**



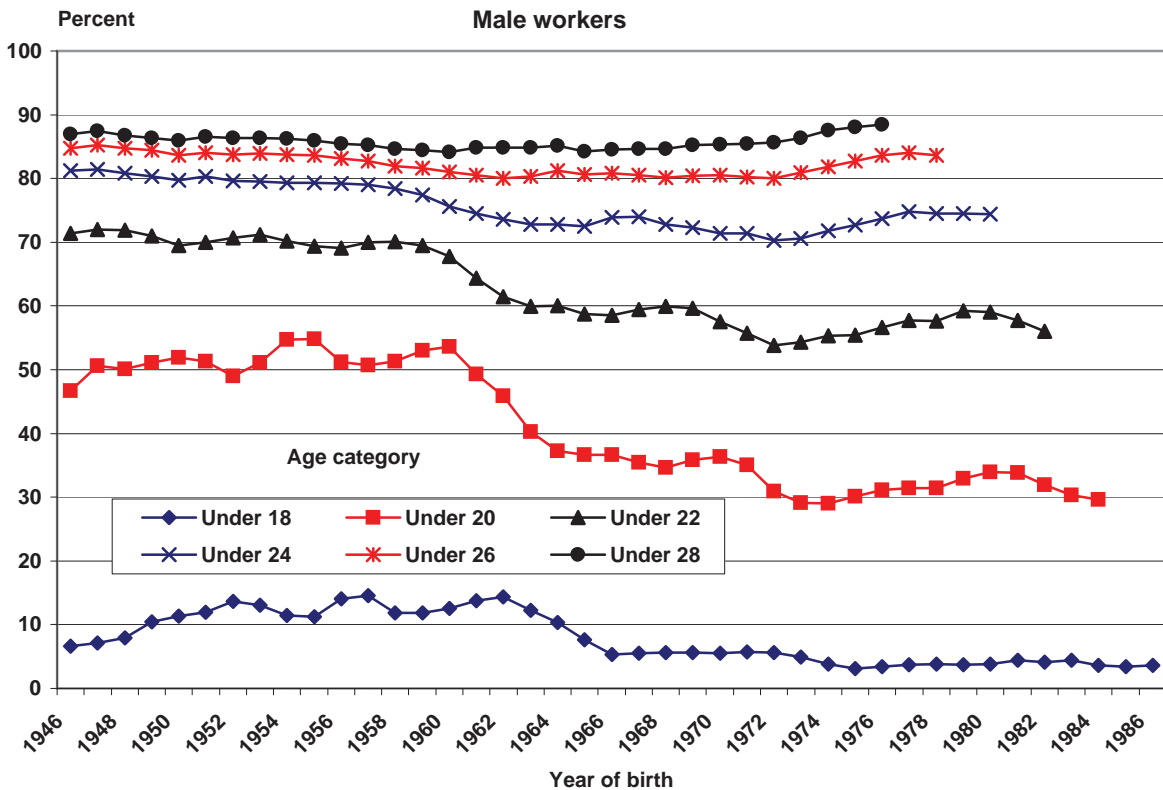
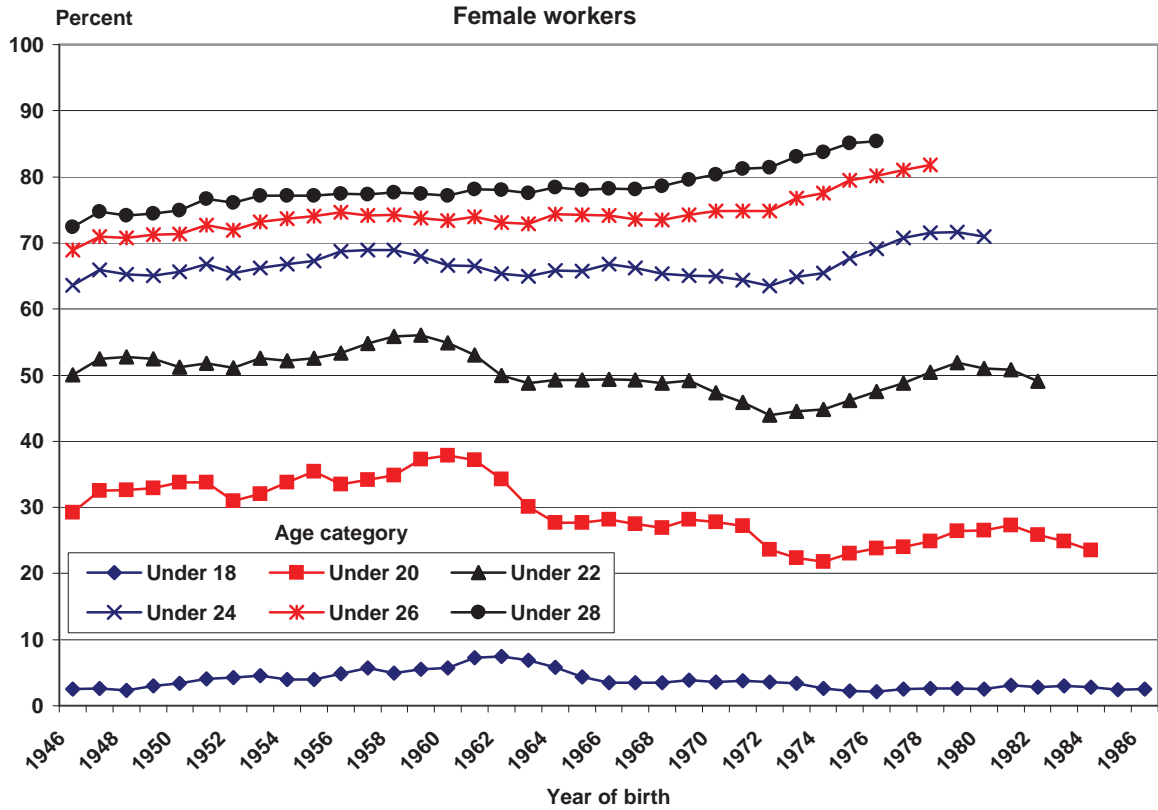
SOURCE: Author's calculations using the Continuous Work History Sample active file for tax year 2003.

**Chart 4.**  
**Cumulative percentage of individuals meeting the very low-wage earner thresholds among those reporting covered earnings, by age category and year of birth**

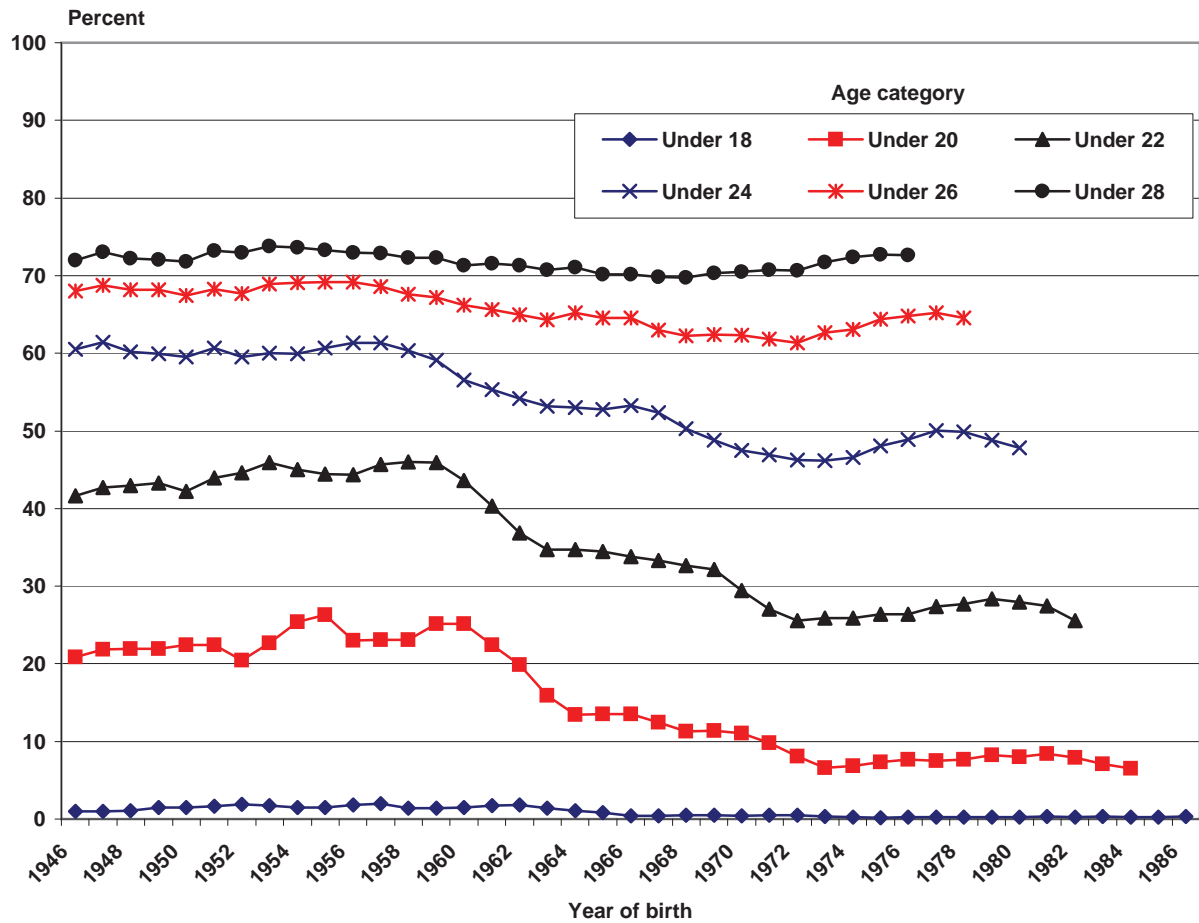


SOURCE: Author's calculations using the Continuous Work History Sample active file for tax year 2003.

**Chart 5.**  
**Cumulative percentage of female and male workers first meeting the very low-wage earner thresholds among women reporting covered earnings, by age category and year of birth**

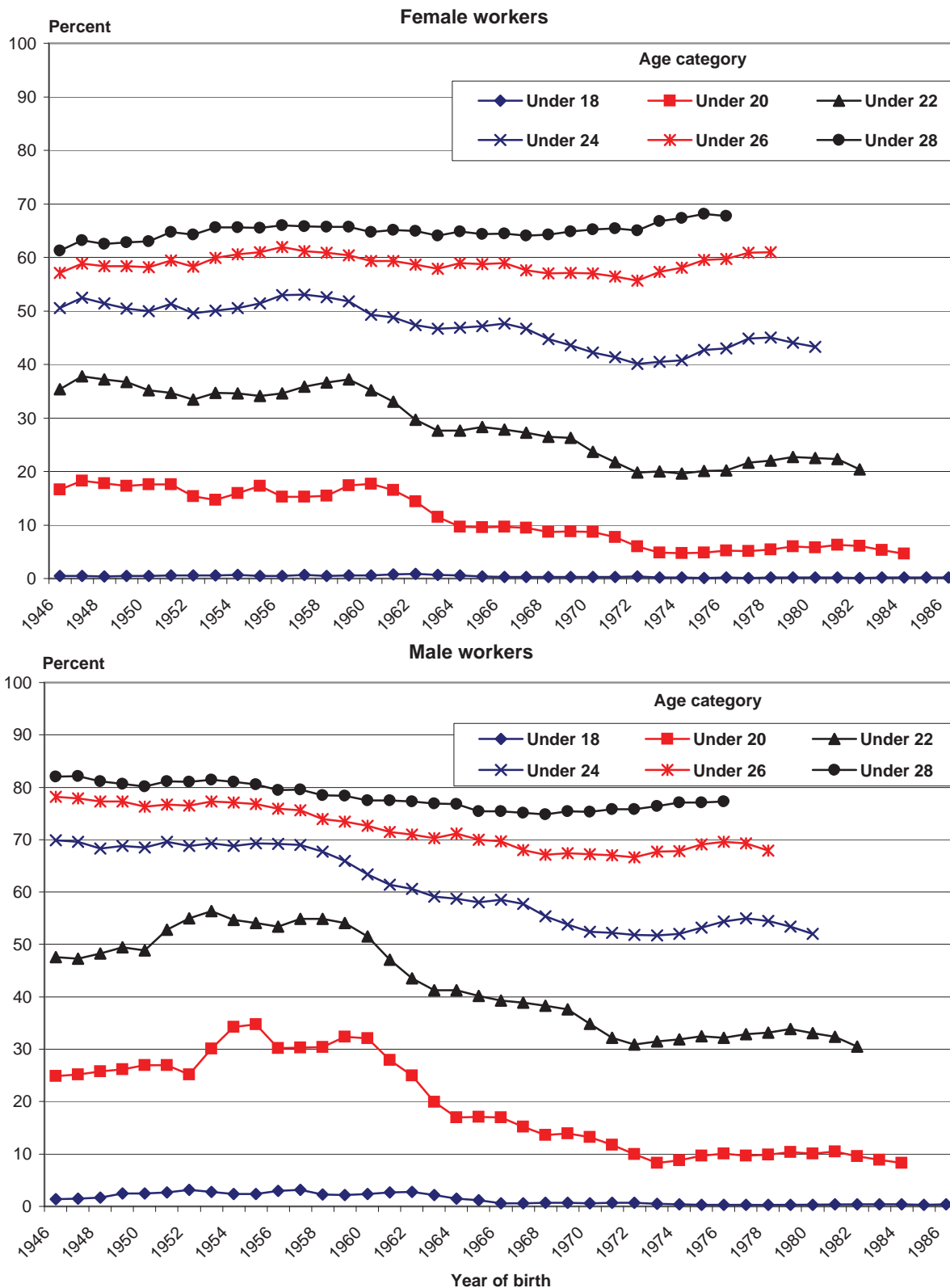


**Chart 6.**  
**Cumulative percentage of individuals meeting the low-wage earner thresholds among those reporting covered earnings, by age category and year of birth**



SOURCE: Author's calculations using the Continuous Work History Sample active file for tax year 2003.

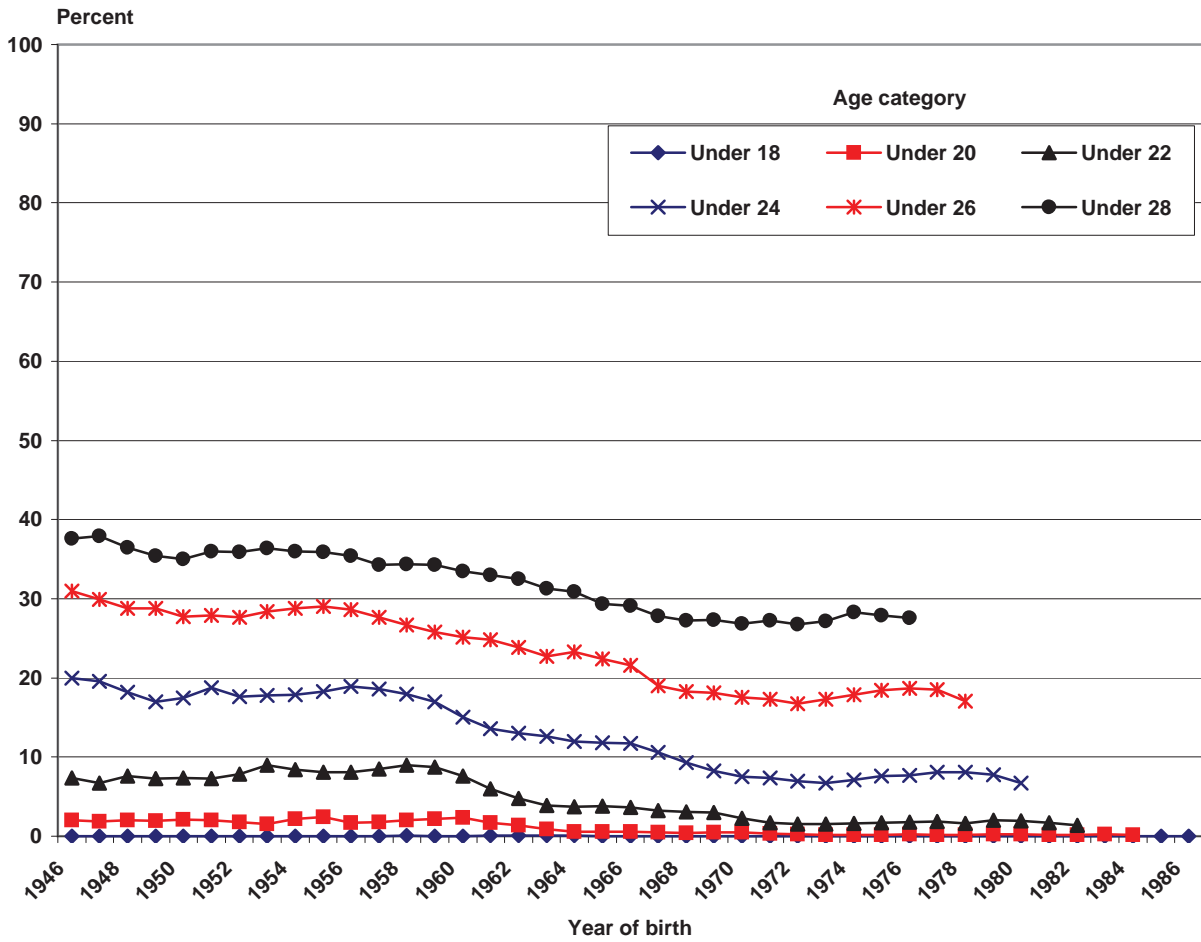
**Chart 7.**  
**Cumulative percentage of female and male workers first meeting the low-wage earner thresholds among those reporting covered earnings, by age category and year of birth**



SOURCE: Author's calculations using the Continuous Work History Sample active file for tax year 2003.

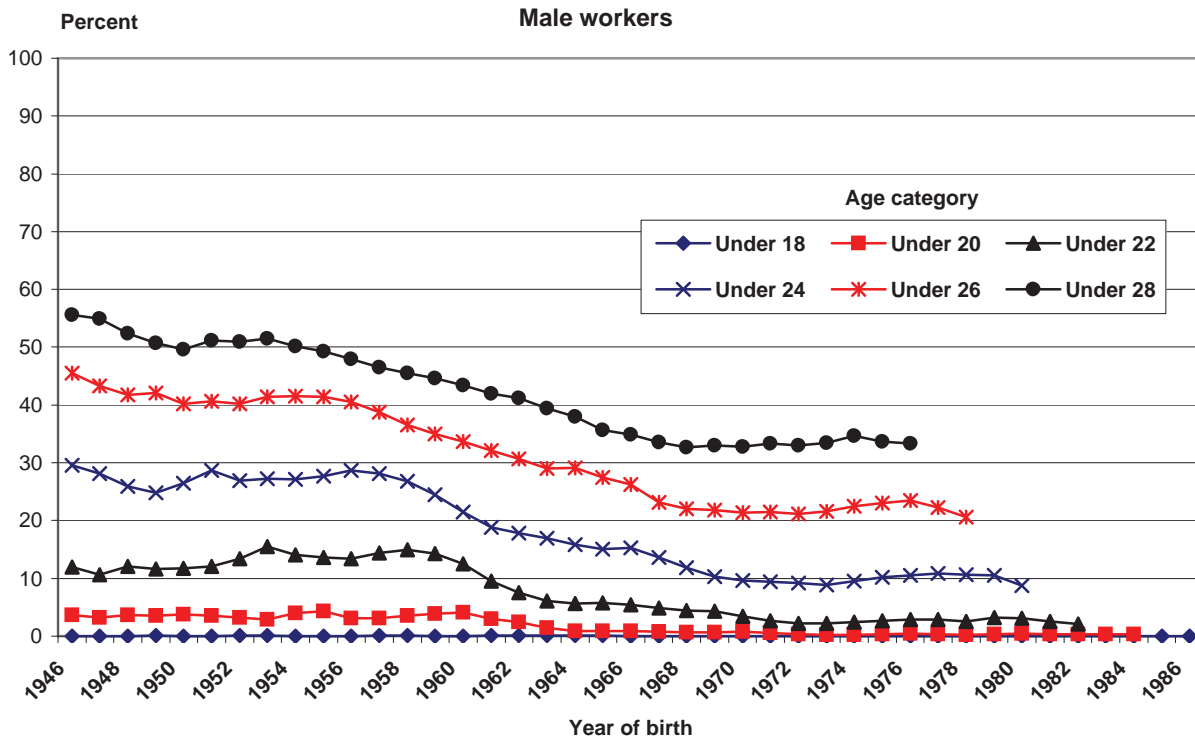
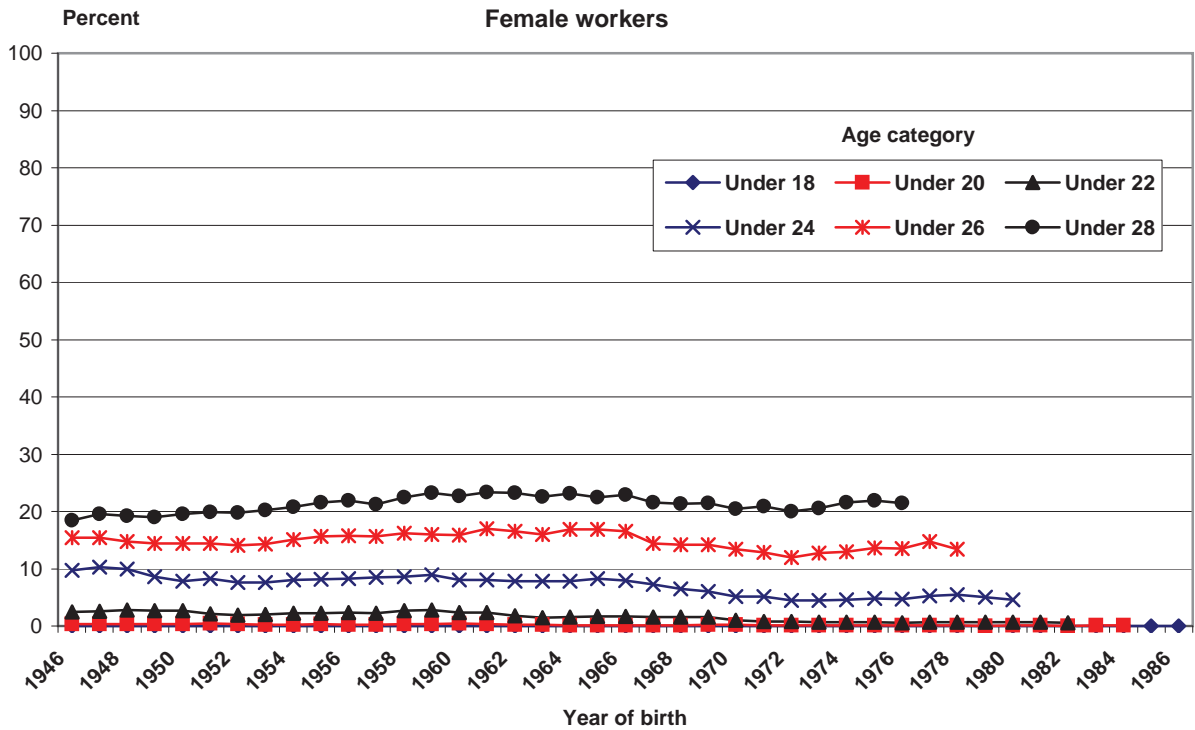


**Chart 8.**  
**Cumulative percentage of individuals meeting the average wage index earner thresholds among those reporting covered earnings, by age category and year of birth**



SOURCE: Author's calculations using the Continuous Work History Sample active file for tax year 2003.

**Chart 9.**  
**Cumulative percentage of female and male workers first meeting the average wage index earner thresholds among those reporting covered earnings, by age category and year of birth**



SOURCE: Author's calculations using the Continuous Work History Sample active file for tax year 2003.