May 9, 2016
Conference draft.

Women Working Longer: Facts and Some Explanations

Claudia Goldin
Harvard University and NBER

Lawrence F. Katz
Harvard University and NBER

For presentation at the *Women Working Longer* Conference, May 21-22, 2016, Royal Sonesta Hotel, Cambridge MA.

Acknowledgments: We thank the University of Michigan, David Wise and the staff at the NBER, especially Mohan Ramanujan, for enabling our use of the restricted access version of the HRS. We gratefully acknowledge the work of researchers at the RAND Corporation of Santa Monica CA for easing the lives of countless scholars through the creation of a harmonized version of the HRS, known as the RAND HRS data and the RAND HRS Family Files. We thank our research assistants, who labored over the CPS, HRS and the Social Security earnings files. Natalia Emanuel began the project by extending the RAND version of the HRS. Amira Abulafi worked tirelessly to make sense of the data and help create the first version of the paper. Jonathan Roth took over from Amira and created the lifecycle and career condition variables. We thanks Maria Fitzpatrick for providing her code to produce the "ever a teacher" variable. We gratefully acknowledge the financial support of the Alfred P. Sloan Foundation's Working Longer program under grant no. 2013-6-16, "Women Working Longer."

Women have been working longer for a long time in U.S. history. Their labor market participation increased decade after decade during the twentieth century, as more women were drawn into the labor force. But that is an old story. The new story is that a large fraction of women are working a lot longer, past their sixties and even into their seventies. Their increased participation at older ages started in the late 1980s before the turnaround in older men's labor force participation and before the economic downturns of the 2000s. [1]

Women's increased participation beyond their fifties is a change of real consequence. Rather than being an increase in marginal part-time workers, the higher labor force participation of older women consists disproportionately of those working at full-time jobs. Women are remaining on their jobs as they age rather than scaling down or leaving for positions with shorter hours and fewer days.[2]

Why have women as a group increased their participation at older ages? Increased labor force participation of women in their older ages, we will emphasize, is part of the general increase in cohort labor force participation rates. Successive cohorts, for various reasons, increased their participation at all ages, resulting in an upward shift of participation by birth cohort. As more women held jobs with greater advancement, were college graduates, were not currently married or were married to men who also extended employment into their senior years their participation did not wane as much in their fifties. More remained active in the labor force into their sixties and beyond.

Rising cohort effects in labor force participation across successive birth cohorts of U.S. women are clearly visible in the raw data from the Current Population Survey (CPS) and the Health and Retirement Survey (HRS). But these cohort effects are considerably dampened when education is considered. Higher participation at all ages has been due to greater levels of education, particularly college graduation. The increase in cohort effects in labor force participation for women in their late fifties and early sixties is also lessened

[1] According to OECD data, most nations have had increased labor force participation for women in their sixties. Among those 65-69 years old, the U.S and Japan lead in participation (not necessarily
[2] In future versions, we plan to use the HRS to see whether women once gradually left the labor force by transitioning to part-time work but do not do that in more recent years.

by including work experience at younger ages and by adding information on the main prior occupation. Although it would be useful to know more about the characteristics of women's prior jobs, the HRS data we are using are relatively mute on that score. We find some (negative) impact on the labor force participation at older ages of college graduate from having been a teacher and discuss what that is the case.

Since many of the cohorts we consider were those that experienced greater divorce, current marital status is also related to employment at older ages. Since couples often coordinate their work and leisure, current employment of the spouse is an additional correlate of whether a woman is working longer.

Most of the factors just mentioned, particularly educational attainment and earlier employment continuity, were determined prior to the employment decision under question. The addition of these factors almost nullifies the cohort effects, except in one important case. For the most recent cohorts of college graduate women we can study to their sixties (those born from 1949 to 1955), the predetermined, observable factors do not eliminate the cohort effect. Something else is keeping them in the labor force at older ages.

Labor force participation rates of women in their early sixties can be observed today for cohorts born up to the mid-1950s. Participation rates of forty and fifty year old women born in the late 1950s and early 1960s have not increased relative to those of prior cohorts and have even decreased for some. Lifecycle cohort labor force functions are no longer the inverted U-shaped or humped functions they once were. They have become flat lines, similar to those of men. These flat lines, moreover, have intersected (sort of pierced) the humped functions of prior cohorts, signaling a decrease in participation at mid-age for recent cohorts. What does this mean for the future employment of older women?[3]

For one thing, labor force participation rates of the more recent cohorts in their middle age will be less than for prior cohorts. In that sense, it would appear that

[3] Hurd and Rohwedder (2014) use questions in the Health and Retirement Study on subjective probabilities of employment to predict future labor force participation rates. See also Maestas and Zissimopoulos (2010) for participation forecasts at older ages to 2030 and for an excellent summary of the issues.

employment at older ages could stagnate or even decrease. But other factors are operating in a different direction. For one thing, college graduation rates are continuing to increase, especially for women.[4] Another point is that women are exiting for shorter periods.

In addition, our finding that college graduate women currently in their early sixties have positive cohort effects in labor force participation that remain substantial even after controlling for their earlier lifecycle participation rates suggests that today's younger women will likely retire later than one would have predicted based on their educational attainment and lifecycle participation rates. The finding is particularly noteworthy since college graduates are a much larger fraction of the total for cohorts born in the 1960s.

A. Labor Force Participation Rates

1. By Age, Sex and Education Level

The central facts concerning the labor force participation of women by age are shown in Figure 1, which uses the March CPS samples and gives rates for women by five-year age groups since 1962. Throughout almost the entire period shown, participation rates have increased for women in the 35 to 54 year group. But these rates were fairly flat for women 55 years and older until the 1980s, when an almost continuous increase ensued, even for the 70 to 74 year old group. Since schooling attainment will be of importance in understanding the change, it is useful to observe trends by education level.

The trends for college graduates are given in Figure 2. The series is restricted to the currently married group because a large fraction of the older cohorts of college graduate women, born from the 1890s to the 1910s, never married or married late in their lives (Goldin 1997). A very large fraction never had children. These women, in consequence, had higher labor force participation rates relative to others in their cohort.

[4] College graduation (meaning baccalaureate) rates for women by age 35 are expected to increase from 38.4 percent for those born in 1978 to around 43.7 percent for those born in 1987. The last number is extrapolated based on changes in prior cohorts from 25 to 35 years. Note that the data used here on college graduation rates is based on the responses of those around 60 years old. Responses are higher than for the same cohorts at 35 years old.

But because many in the 1890s to 1910s cohorts who married often had no children, their labor force participation rates are still high for the earliest cohorts shown. Participation rates for this group decline a bit over time as college graduate women began to marry and have children at rates, by age, that were more comparable to those of those without higher education.

If one ignores cohorts born before around 1920, the series for all women and that for college graduate women 55 or older are first relatively constant and then increase over time, particularly after the mid- to late-1980s.[5] The percentage point increase during the past 25 years, shown in Table 1, is not much different between the aggregate group of women and the college graduates. But because college graduate women have had considerably higher participation rates than lesser-educated women, the shift of women to college has increased participation rates for older women and increased the growth of women's employment at older ages.

Also clear in Table 1 is that the increased participation of older women exceeds that of older men both absolutely and relative to the base levels. For 60 to 64 year olds, for example, female participation increased by 17 percentage points on a base of 34 percent but for males the increase is just 6 percentage points on a base of 55 percent. The percentage point increase for 65 to 69 year olds males and females is similar in absolute magnitude, but the initial base for women is far lower (15 versus 26 percent). The relative increase for older women has meant that the gender gap in participation at older ages has greatly decreased, as can be seen in Figure 3.

The gender gap in participation has, of course, decreased more generally. But the absolute percentage point difference at some of the older ages is now smaller than at the younger age groups whereas it was once just as large. For 60 to 64 year olds, for example, the difference in participation rates between men and women was about 50 percentage points in 1962. In 2013 the difference was just 9 percentage points, when that for males and females in their thirties to mid-forties was around 16 percentage points.

[5] For the 60 to 64 year old group, participation rates after around 1980 refer to individuals born after about 1920.

Men and women are doing more of the same things throughout their lives and this is even truer at older ages. But is that also true within couples? The answer to that question is that far more couples between 59 and 63 years old are working together now rather than being retired together.[6] In addition, there are almost as many couples in that age group in which the wife is working and the husband is not than vice versa. We return to the issue of joint employment and leisure below.

2. Full-time versus Part-time Employment of Women at Older Ages

During the recent period of increased labor force participation among older women, the fraction working full-time and full-year has greatly increased. The increase is especially evident for the 65 years and older group. As seen in Figure 4, the fraction of 65 to 69 year old women in the labor force who worked full-time and full-year increased from around 30 to almost 50 percent, with much of the increase occurring after 2000. The fraction of the 70 to 74 year old labor force participants working full-time and full-year increased from 20 to almost 40 percent.

Although the timing could indicate the impact of changes in the Social Security earnings test, the increase begins before 2000 for both younger and older age groups of women.[7] Note, as well, that prior to the increased participation of older women, the fraction working full-time and full-year actually decreased.

3. Cohort Trends

Increased employment among older women would appear to be related to their increased participation earlier in their lives. This conclusion can be deduced from the fact that all cohorts in Figure 5.A that have increased participation in their sixties, relative to older cohorts, also had increased participation relative to older cohorts earlier in their lives. These cohorts had higher participation rates throughout their lifecycles than did previous cohorts. Although the graph begins with the cohort born in 1929, the pattern is

[6] This statement is true for couples in which the woman is between 59 and 63 years old using the Health and Retirement Study. For couples in which the woman is between 62 and 63 years, the statement is true beginning in 2008.
[7] **Note** changes in the law regarding the Social Security Earnings test.

evident as well for some of the earlier birth cohorts. Only with cohorts born in the early 1920s is there no discernible increase in participation among women in their sixties despite modest increases earlier in their lives.[8] The data for college graduates given in Figure 5.B reveal similar findings but participation levels are higher.

As will be emphasized later, regressions of the labor force rate at older ages on birth cohort dummies indicate that cohort effects are greatly muted by the addition of various predetermined factors such as education, earlier employment continuity, and women's past occupations. That is, cohort differences in labor force participation later in life are largely, but not entirely, a function of earlier changes in human capital accumulation. These human capital advances occurred to a great extent because women perceived that their investments would pay off in the labor market and that their employment would be higher and more continuous than it was for previous cohorts.

We noted before that the typical inverted U-shaped function tracing out a cohort's lifecycle labor force participation has become more of a flat line after the mid-1950s birth cohorts. Participation rates at around ages 25 to 35 greatly increased for the 1930s to 1950s birth cohorts, in large part because births have been delayed.

We also noted that the new flatter cohort lifecycle functions have begun to cross each other. This crossing creates an interest "twist" in participation for the most recent cohorts in Figures 5.A and more so for college graduate women in 5.B. The "twist" is the cohort representation of the oft-mentioned decrease in the participation of thirty-something women, the idea that women are somehow "opting out" of the labor force.

Participation rates for women in their thirties and forties have in recent years levelled off and have even decreased relative to slightly older cohorts. One clear way to see the change is to observe that slicing the cohort graphs at around ages 30 and 50 yields the usual cohort progression. Younger cohorts have higher participation rates than older cohorts. But slicing the cohort graphs between those ages, say at age 40, does not yield

[8] These general trends are also apparent in Figure 1. For example, the participation line for those 65 to 69 years begins to increase around 1987, therefore for women born in the early 1920s.

higher rates for the younger recent cohorts, such as those born from 1959 to 1973. These lines appear to have twisted.

Does this mean that participation rates for these women in their fifties, sixties and beyond will also be lower? Their increased education and labor force participation in their younger years would argue the opposite. Why they have decreased participation is still an on-going research question, although much concerns delayed births and an absence of mandated leaves of more than 12 weeks. The decrease in participation is not large, but the disruption of the increasing trend is clear and could argue for a break in the increase of women's working longer.[9]

The bottom line for cohort change is that increased participation at older ages has occurred for cohorts that increased their attachment to the labor force throughout their lives. The upshot is that the more one has been attached to the labor force, the longer one works. We now turn to using longitudinal information from the HRS matched to Social Security earnings records to understand the role of cohort effects.

B. Exploring the Role of Cohort Effects Using the HRS

Cohorts born later in the period have higher labor force participation rates at older ages than do those born earlier. We explore whether these cohort effects are primarily due to changes in factors determined largely prior to the retirement option. These variables can include educational attainment, number and ages of children and lifecycle labor force participation. We will also consider the degree to which the individual had relatively high earnings when employed, which we term the "career" condition. These largely predetermined characteristics will be measured in our empirical work prior to around age 54 whereas the retirement option is considered from ages 59 to 63.

[9] Hurd and Rohwedder (2014) note that subjective probabilities of future work at ages 62 and 65 are reliable predictors of actual employment and that current HRS respondents in their early fifties have subjective probabilities of future work that exceed the actual participation rates of individuals in their sixties. This finding indicates a continued increase in the employment of older individuals.

The retirement decision may instead be primarily determined by factors that are contemporaneous, such as a set of shocks or transitory factors. These factors may have served to increase participation at older years in the post-1980s period and may include marital status change, fluctuations in the value real estate or financial assets, pension losses, reductions in the value of Social Security payments, and deteriorating health status.

The evidence points to a large impact of changes in the predetermined factors. Educational differences across cohorts reduce cohort differences in labor force participation from ages 59 to 63 by about a half. Lifecycle labor force participation prior to age 54 produces an overshooting of the cohort effects, meaning that the cohort effects are over explained. Our measure of serious career earnings does not perform better than the simpler measure of life cycle participation, and generally performs worse.

Once these variables are considered, adding information on the number and birth years of children has no impact. Children serve to reduce participation in the 25 to 34 year range but have no separate effect in later life. The many contemporaneous factors mentioned are related to the variance within cohorts but do not do much to explain changes across cohorts.

The one interesting anomaly concerns the most recent of the cohorts of college graduate women that can be followed to their sixties. Those born from 1949 to 1951 have higher participation at ages 59 to 63 even given their measures of their lifecycle participation prior to age 54 and educational attainment. That is, the cohort effect for the 1949 to 1951 group remains significant even including the various predetermined factors, in particular lifecycle participation.

The finding that later cohorts have higher participation given their earlier lifecycle participation may be useful in forecasting what more recent cohorts will be doing when they reach their sixties. Recall that labor force participation rates across the lifecycle have become fairly flat from ages 25 to 45 and that the youngest cohorts of women do not always have higher participation relative to previous ones by age. In fact, the most recent data indicate a backtracking of younger cohorts of women in their forties. That is, for the

college graduate group, as well as for others, participation rates have not increased relative to prior cohorts and have even decreased at various ages.

The finding about those born between 1949 and 1951 may indicate that participation rates for even younger cohorts may be higher still in their sixties and seventies than prior generations at least for college educated women.

To explore the role of cohort and predetermined variables, data from the Health and Retirement Study (HRS) are used together with information on the earnings history of the respondents from Social Security earnings data and W-2 forms (after 1980).[10] Each of the respondents to the HRS, beginning with the first cohort in 1992, was asked whether their Social Security earnings history could be linked. If the individual agreed to the linkage, then all past records were linked.[11] If the individual did not, then the individual was asked again in the next biennial survey. Therefore, the older cohorts had more chances to agree to a linkage than the younger cohorts and the linkage rates will be higher in consequence.[12]

When we use information on lifecycle labor force participation, we must restrict the sample to individuals who gave permission to have their Social Security earnings (and W-2 forms) linked. Otherwise the full HRS sample is used, given age and other restrictions that may apply. Across all cohorts about 80 percent agreed to the linkage with Social Security (and W-2) records, but it is smaller for younger cohorts for the reasons just given.

---

[10] Appendix Figure A1 shows the close relationship between labor force participation rates in the HRS and in the CPS for women in their fifties and sixties. For women in their seventies the HRS has higher participation rates. The reasons are not yet clear since each survey is supposed to cover those in nursing homes and similar care units. Note that the W-2 data are also provided for 1977-79 but are incomplete for those years in the HRS-SSA linked data. We, and other researchers, begin the W-2 data with 1980 due to that issue.

[11] A curious aspect of the HRS is that until 2006 individuals were asked every year if they would continue the linkage to the Social Security earnings data. If at any point they decided not to, the prior data were allowed but the contemporaneous and future data were not. For most HRS respondents, the break in the linkage will not matter since the HRS itself collected information on labor supply and earnings. But the break will matter for a spouse who entered the HRS at a younger age and who was folded when the individual's birth cohort relevant HRS cohort was added. See Appendix, especially the section Social Security Earnings Record Linkage in the HRS, for details.

[12] See Appendix Table A1 for linkage rates by birth cohort.

In all cases we examine labor force participation rates of women 59 to 63 years old and include three-year birth cohort dummies. We begin in Table 2, cols. (1) to (5) by including characteristics largely determined prior to age 54, such as educational attainment and lifecycle participation during various intervals.[13] We add in col. (6) current marital status and a summary measure of current health status. Table 3 divides the group into two education levels, college graduates and those who did not graduate college.[14] Cols. (1) to (4) of Table 3 include the predetermined characteristics and cols. (5) and (6) add current marital status and health status. Table 4 includes only college graduate women. In addition to the previous variables, we add information on whether the individual was ever a teacher. A large fraction of college graduate women in the earlier cohorts were teachers for much of their working lifetimes and teachers often had defined benefit pensions.

The use of the HRS linked to the Social Security earnings records reduces the size of the sample, less so for the earlier than for the more recent cohorts. In Tables 2 and 4 we explore the sensitivity of the results to using the "full" sample and the "linked" sample. Because the HRS is a longitudinal sample, many of the respondents are in the sample more than once between the ages of 59 and 63 and thus we cluster the standard errors at the individual level. We also include dummy variables for the single-year ages.

In Table 2, col. (1) the baseline regression is provided for the full sample and in col. (2) when estimated for the smaller linked sample. The variables of interest are those giving the effect of cohort birth years, given in three-year bins from 1931 to 1951 (where 1931-33 is the omitted cohort group).[15] The effects are very similar between the two samples and both demonstrate the increase in participation at older ages for birth cohorts after 1943 and especially after 1949. The most recent of the cohorts that can be analyzed for the 59 to 63 year old group, born from 1949-51, has a participation rate that is around

[13] The addition of variables for children ever born adds no explanatory power for older women's labor force participation after including controls for earlier life cycle participation. Thus, we do not include specifications including controls for children.

[14] We provide similar regressions to those in Table 3 for women 56 to 58 years old in Appendix Table A5.

[15] The last year of the HRS is currently 2012. The 1952-54 cohort is incomplete and thus is omitted.

10 percentage points higher for the full sample and 9 percentage points higher for the linked sample than that for cohorts born in the 1930s. The only additional covariates included in the first two columns are single year of age and race dummies.

Educational attainment is added in col. (3) and lifecycle participation between ages 35 and 44 in included in col. (4). The lifecycle labor force variables give the fraction of years in the interval that the woman was in the labor force. These have been computed from restricted-access information on HRS participants from Social Security earnings data (since 1951) and W-2 forms (when available). Additional information was used from the HRS to add labor force data for those exempt from Social Security taxes, in particular government employees such as teachers. The HRS provides information concerning two periods prior to the start of HRS interviews in which the respondent was a government employee. When HRS survey responses are available regarding participation, they are used in place of Social Security earnings and W-2 data. (For more details, see Appendix: *Lifecycle Participation and Career Condition*.) Various life cycle employment variables were created for each of the three decades from age 25 to 54 and for the entire period.

The addition of educational attainment eliminates the economic and statistical significance of the cohort coefficients for all but the most recent of the birth cohorts. Although the linked sample coefficients are given, those for the full sample change in the same manner. The addition of the lifecycle participation variable in col. (4) further reduces the coefficient for the most recent of the birth cohorts to a slightly negative value. It also produces some modest reduction of the impact of educational attainment since the more educated have greater continuity in employment.

Instead of a variable that measures lifecycle participation, one that measures the degree to which a woman reaches some career level may be even more important in determining future participation. Since women with greater prior employment when first beginning their careers have greater attachment to the labor force later in the lives, those with higher earnings during their employment periods should have even greater attachment.

To test whether employment per se or years of better earning performance matter we create a variable giving the fraction of an age interval that a "career condition" was met. The condition used here is achieving an earnings level that is some fraction (50 percent in this case) of the median earnings of a full-time, year-round male worker for the ten-year age group considered during the relevant period.[16] That is, the career condition for a woman when she was in an age group is judged relative to the earnings of the median male in the same age group during the identical time period. Women who were never in the labor force in the age interval get zero, as do those who never earned more than the condition but were in the labor force. We find that the variable giving the career condition, in col. (5), is related to later employment but it less strongly related than is the simpler measure of the fraction of the interval a woman was employed.

Col. (6) augments the col. (4) specification by adding two contemporaneous variables: current marital status and current health status. Because the birth cohort coefficients were already extinguished with controls for education and earlier lifecycle participation, the added contemporaneous variables actually lead to overshooting. The addition of health status reduces the impact of education, and in most instances halves the coefficients in col. (4). The more highly educated are also the healthiest or, at least, they consider themselves to be healthy. The coefficient on earlier lifecycle participation remains substantial and is only slightly reduced.

It is useful to explore the impact of current marital status even if it does little to change the birth cohort coefficients. Being currently married decreases participation for older women but the effect is reduced if the woman's spouse is employed and the total impact is about equal to that of the omitted group (never married) and to widowed

[16] Earnings of the median male, in the same age group and year, are used. These data are available in published documents (U.S. Census Bureau *P-60 Reports*) prior to the micro-data for the CPS, which begins in 1962. The calculation of all the career conditions considered requires data from 1956 (1931 + 25 years). A fraction of the male median is used because the median is too high a bar for employed women during much of the period considered. Women in the exempt occupations are assumed to exceed the bar. See Appendix: *Health and Retirement Survey: Construction of Variables*.

women.[17] Divorced women have participation rates about 8 percentage points higher than the base group of never married women.

Disaggregating by education, as in Table 3, reveals substantial differences between the higher (college graduate) and lower (below a college graduate) educated groups in the correlates of their later employment. Note that within the college graduate group, dummy variables are added for degrees above the bachelors (MA and the various graduate and professional degrees) and, within the non-college group, dummy variables are added for high school diploma and having some college.

The regressions in cols. (1) and (2) of Table 3 include only cohort effects (plus age, race and education dummies). Cohort effects for the college graduates are substantial but the group that did not graduate from college does not reveal a gradient with respect to birth year. For college graduates, however, the latest cohort has participation rates about 11 percentage points higher than for those born in the 1930s.

In cols. (3) and (4) we add lifecycle participation variables including whether the woman was never in the labor force during the interval. The addition of the lifecycle measures has little impact on the cohort effect for the college graduate women born most recently. Earlier labor force participation matters more for the less educated group than for the college educated. For college graduates what matters greatly is whether the women did not work at all in the interval, even though that is a fairly small group. The much higher labor force participation for the 1949-51 cohort of college graduate women remains unexplained even with controls for current marital and health status, as seen in col. (6).

Lastly, Table 4 looks in more depth at college graduates in part because their participation rates are the highest at all ages, particularly among those in their sixties. In addition, the fraction of older women who are college graduates has greatly expanded and will continue to do so given the increase of college graduates at younger ages. Both the increase of college graduation for future cohorts and their higher participation at older ages would imply an increase in the future employment of older women.

[17] We discuss, below, changes in the joint employment and retirement of couples.

Table 4 includes the predetermined (lifecycle participation and education) and contemporaneous (marital and health status) variables. In addition, we include whether the woman was ever employed as a teacher.

Similar to the findings for college graduate women in Table 3, cohort effects are large for the latest of those shown. The coefficient remains large and statistically significant despite the inclusion of current marital status and lifecycle participation variables. Only in col. (5), with the inclusion of the fraction of years from 25 to 54 that the woman was in the labor force does the coefficient greatly decline.

Teaching was the single most important occupation for college graduate women for many of HRS cohorts. Around 45 percent of college graduate women in the cohorts born from 1931 to 1941 were teachers at some point in their working lives, as seen in Figure 6. A much smaller fraction of women (around 30 percent) for the later cohorts considered here, 1945 to 1951, were teachers. And an even smaller fraction (around 20 percent) were teachers in the late 1950s birth cohort, a group still too young to observe in their sixties.

Those who were ever a teacher had participation rates when they were in the 59 to 63 age range that were about 5 percentage points lower than other college graduate women. But controlling for whether a woman was ever a teacher does not knock out the cohort trend. In addition, the impact of ever being a teacher increases when controlling for lifecycle participation, implying that teachers work more than others earlier in their lives but are less likely to work than others later in their lives. Their earlier work would have made one think they would be more likely to work later but they are, in fact, less likely.[18]

The Table 4 analysis reinforces the findings from Table 3 that the cohort effect for the most recent birth group is not extinguished by the other covariates even when the lifecycle participation rate variable is included, as in cols. (3), (4) and (5). The coefficients are slightly smaller than in Table 3 because of the inclusion of the "ever a teacher" variable but they are still relatively large and significant.

[18] The paper by Maria Fitzpatrick in this volume explores the reason for the decrease in employment among teachers at older ages.

Another important finding for forecasting women's future participation at older ages is that employment in the 45 to 54 year old range is the best predictor of whether an individual will remain employed into her early sixties. Therefore, even though participation rates have "twisted," as noted in the discussion of Figure 5, the fact that participation is still higher for the most recent cohorts when they have reached their fifties suggests that these college graduate women will remain in the labor force to a greater extent through their sixties and seventies. The same is not true for the non-college graduate group.

We have also run the same regressions as in Tables 2, 3, and 4 where the outcome is working full-time, rather than being in the labor force. The results (given in Appendix Table A4 for the Table 3 comparison) for full-time work for college graduates and those below graduating from college reveal similar patterns to those for labor force participation for cohort trends.[19] Thus, the cross-cohort increases in labor force participation for older women are substantially driven by increases in full-time work.

The summary finding is that older women have had substantial increases in labor force participation. The inclusion of various covariates, such as education and lifecycle participation, reduces the cross-cohort pattern of rising cohort effects. But for the college graduate group, the labor force increase for the most recent cohorts in their sixties is not reduced by the inclusion of various covariates. The most recent cohorts with less than college completion, however, have had smaller increases and these do get extinguished with the expanded set of predetermined group of covariates (detailed education attainment and earlier labor force participation).

[19] We also run the same labor force regressions as in Tables 2, 3, and 4 for women 56 to 58 years old and give the Table 3 results in Appendix Table A5. Using women 56 to 58 years old allows us to include another birth cohort, 1952-54. Similar to the Table 3 results, college graduate women in cohorts born after 1948 show larger cohort effects than for earlier cohorts. Including the full set of pre-existing characteristics lowers the cohort effect estimates for the more recent cohorts at age 56 to 58 somewhat more than for the older group of women in Table 3.

C. Lifecycle Labor Force Participation

Given the importance of lifecycle labor force participation for later work, we now explore how lifetime employment changed across cohorts born from 1931 to 1954. We divide lifetime employment into five quintiles: from 0 to 20 percent of the years under consideration to 80 to 100 percent. Figure 7, part A, shows the percentage in the labor force in the five quintiles covering age from 25 to 54 years for all women. Part B gives these figures for the group in the youngest ten-year grouping, 25 to 34 years old. Figure 8 provides the same data for the earliest of the birth cohorts shown (born 1931 to 1938) and for the most recent of the birth cohorts (born 1948 to 1954).

The fraction of women in the labor force 80 to 100 percent of the time they were 25 to 54 years old expanded from 20 percent to more than 50 percent across these cohorts (see Figure 7, part A). The flip side of that increase is the sharp decrease of those in the labor force fewer than 20 percent of the thirty year period shown. The middle three groups change little.

More extreme changes occurred for women in the 25 to 34 year old group (see Figure 7, part B). Once again, the middle three quintiles show little change and all of the increase in the top quintile comes about because of a decrease in the lowest quintile.

One gets a clearer picture of cohort change in Figure 8. Cohorts born from 1931 to 1938 had lifecycle participation rates that were rather uniformly distributed across the quintiles. But by the 1947 to 1954 cohorts almost 50 percent were in the labor force for the entire period and very few were in the lowest quintile.

To make sense of these lifecycle trajectories, the concepts of heterogeneity and homogeneity will be useful.[20] When participation rates for a birth cohort increase with age, all women in the cohort could be working more weeks per year or more women could be entering the labor force. That is, change could be at the intensive or extensive margin (or a

[20] See Goldin (1989) and Heckman and Willis (1979) on the concepts of heterogeneity and homogeneity applied to labor force participation over the lifecycle.

combination). The group that exhibits the former is termed "homogeneous," since all women are increasing their work level, and the group that exhibits the latter is termed "heterogeneous," because only some women are increasing their participation.

The difference is far more than semantic.[21] The distinction helps us understand why the lifecycle labor force distributions in Figure 8 appear as they do given the cohort trajectories from Figure 5. A uniform distribution, as is the case for the 1930s cohort in Figure 8 part A, is fully consistent with heterogeneous participation when labor force participation rates trace out a clear hump or inverted U-shape. And a more skewed distribution, as in Figure 8 part B, for the late-1940s to 1950s cohorts is also fully consistent with heterogeneous participation when cohort rates are humped, but begin at a higher level and are flatter.

The distinction will be critical in making projections about whether more recent cohorts will attain higher participation than their older peers when they are in their sixties and seventies. Looking back at the synthetic cohort lines in Figure 5, the most recent cohorts are seen to display flat and even somewhat decreasing participation rates over their brief lifecycles. That is, participation rates are actually higher at the lower ages than at the middle.

But if most working women persist in the labor force, then the finding that early participation matters significantly implies that the reduction in participation, or the absence of an increase, for the most recent cohorts in their midyears will not matter much for their employment later in life. The key point is that in earlier cohorts, women who entered the labor force in their midyears were probably the least persistent.

D. Working Women, Working Couples

The regression results showed a commonly known relationship. Couples generally work together and play together. Currently married women are far more likely to be in the

[21] Olivetti (2006) models an underlying reason for greater persistence in the returns to experience and demonstrates the increased returns from the 1970s to the 1990s.

labor force in their older years if their husbands are also working. In the Table 2, col. (6) regression, married women with a spouse in the labor force are almost 20 percentage points more likely to be in the labor force than are other married women.

Figure 9 demonstrates two additional points beyond that in the regressions, which did not interact marital status with birth cohort. The data use three categories of women 59 to 63 years old: those currently married with a husband working, those currently married with a non-working husband and those not currently married. The first is that participation rates of all currently married women rose relative to the third group and the second is that the rates of currently married women with a working spouse increased most.

A greater fraction of couples are now working together rather than being retired together whereas twenty years ago a greater fraction of couples were retired together.[22] Even though husbands and wives generally have complementary leisure, the timing of the recent increase in participation rates differs between older men and women. Whereas the increase for women began around the late 1980s, that for men is not discernable until around 1998 or even later. That is, there is a decade difference between the increase for men and that for women. Neither corresponds with the dating of the Great Recession. The increase for women precedes some of the more important changes in Social Security, although that for men may be related to the change in the Social Security earnings test.[23]

E. Concluding Remarks

We have explored the increase in the labor force participation of older women that began around the late 1980s. Using CPS data, we have observed that the increase in participation was disproportionately in full-time, year-round jobs, rather than part-time positions.

---

[22] These findings are consistent with the complementarity of the leisure time of older husbands and wives. See Schirle (2008) for a summary of the literature on complementarity and Blau (1998) in particular. The Schirle (2008) article demonstrates, for three countries, that the increase in women's labor force participation at older ages has led to increased men's participation.

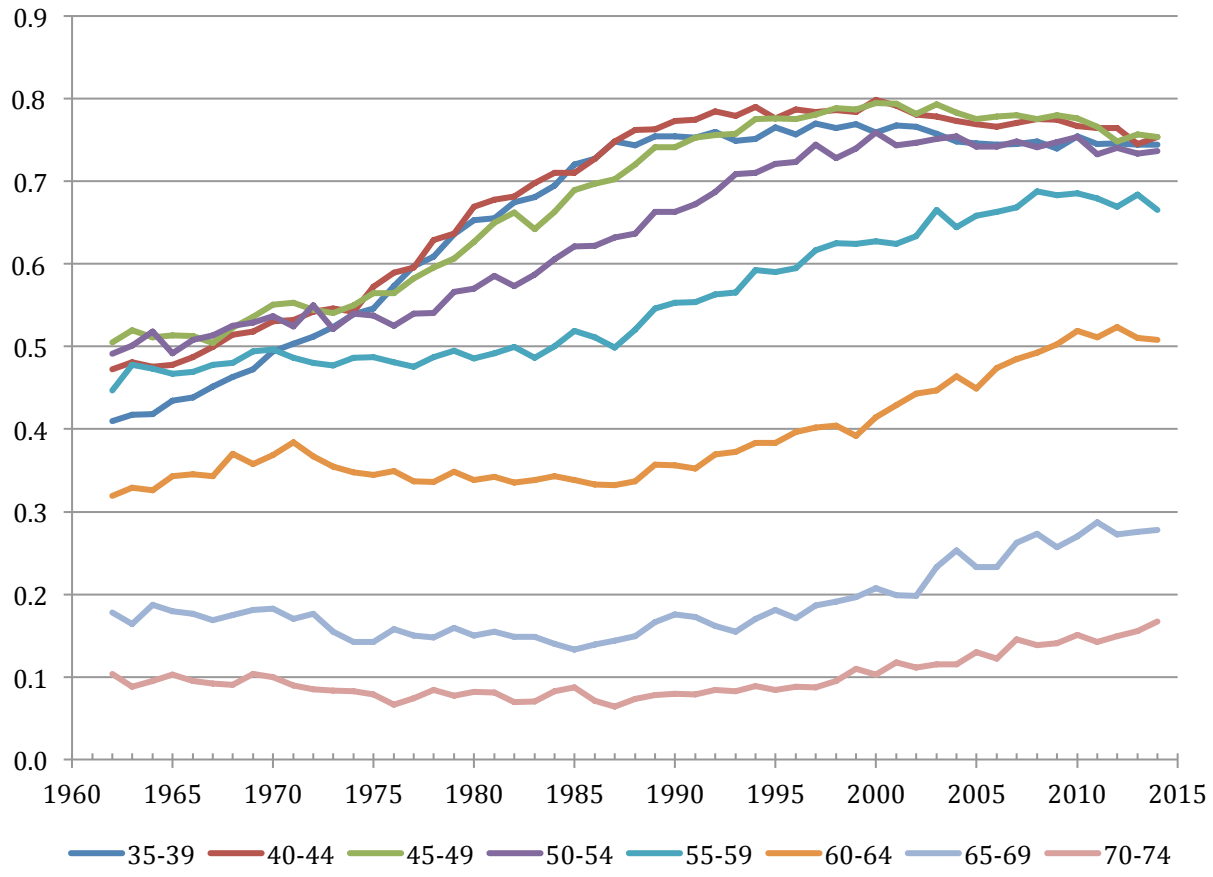
[23] The impact of recent Social Security changes on retirement trends is discussed in Gustman and Steinmeier (2009) and Mastrobuoni (2009), among others.

We have investigated why labor force participation increased by using the HRS and related Social Security earnings records to analyze the participation of women 59 to 63 years old across cohorts born from 1931 to 1951. We first analyze the effect of characteristics largely determined prior to labor market entry and then those determined before a woman is in her late fifties. For the aggregate sample, participation rates increase most for cohorts born after around 1943 and the increase in participation is greater for the more highly educated. Most important is that the cohort effects, noticed in the aggregate data, are greatly reduced when education is included in the regression.

Although the rising cohort effects for all women are virtually extinguished by the inclusion of education, the college graduate group displays an interesting and potentially important phenomenon. The cohort effect for the most recent birth cohort that can be explored, that from 1949 to 1951, remains large and statistically significant even after controlling for earlier labor force participation (from ages 35 to 44 or from ages 45 to 54).

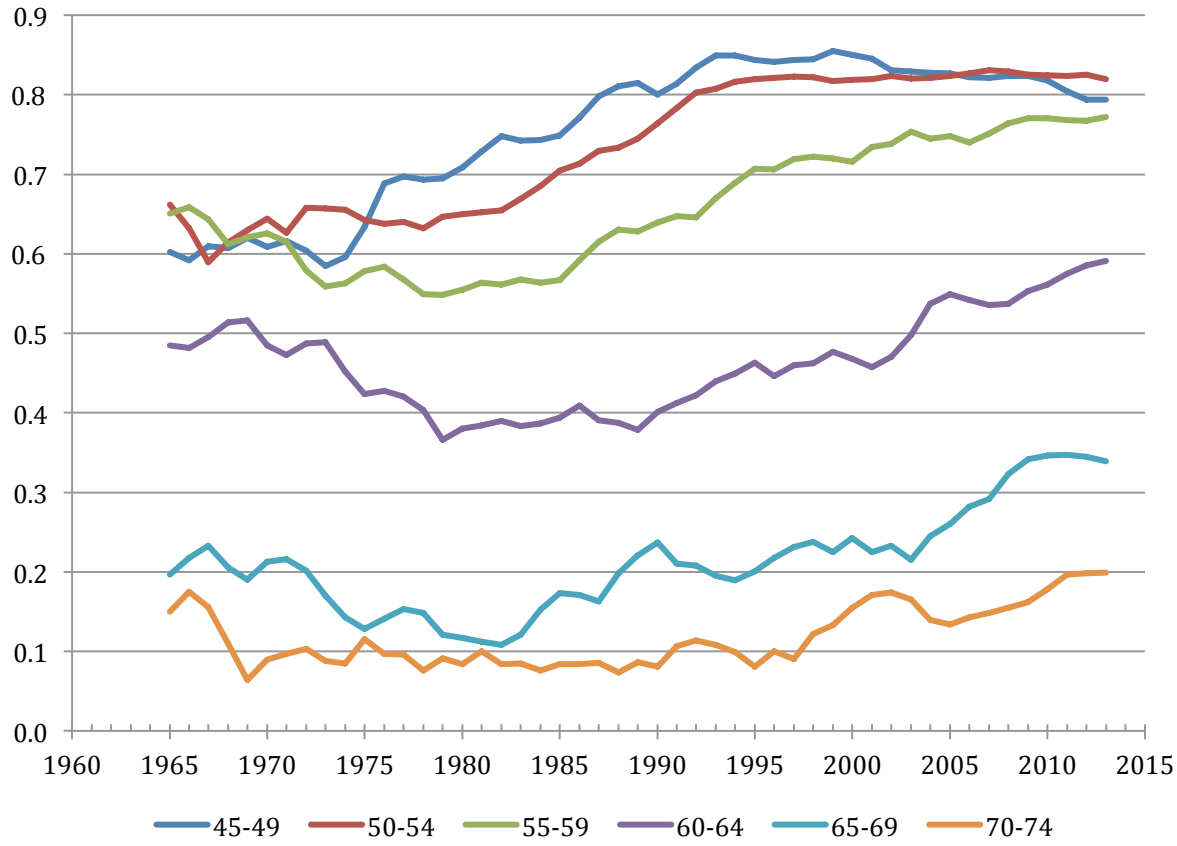
Whereas the increased participation rates of older women with birth cohorts in the early 1940s seem explicable due to their higher levels of education, that for the most recent cohort that can be explored appears to be due to other factors. Our current challenge is to understand what these factors are and how they are likely to impact the labor force participation at older ages of future cohorts.

Figure 1: Female Labor Force Participation by Five-Year Age Groups, 1962 to 2014



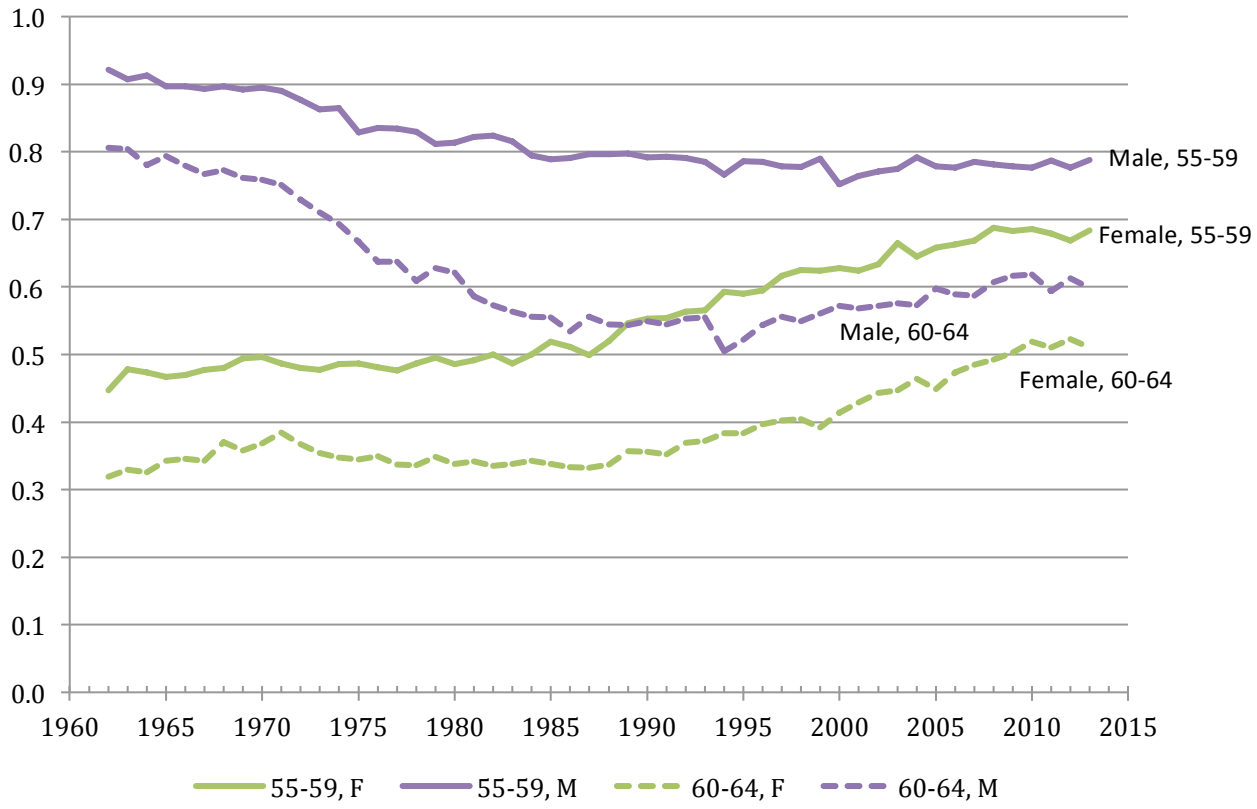
Source: CPS March.

Figure 2: Female Labor Force Participation by Five-Year Age Groups for Currently Married College Graduates, 1965 to 2013 (three-year centered moving averages)



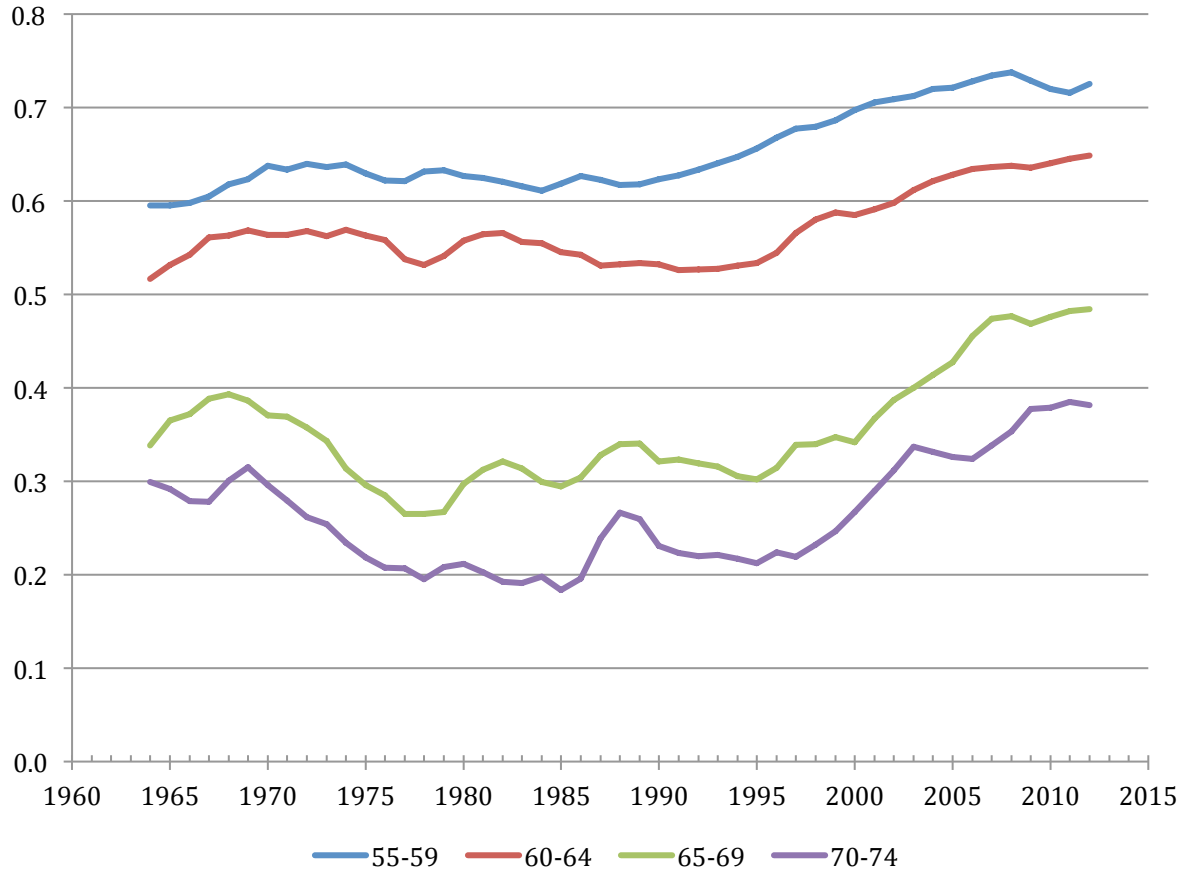
Source: CPS March.

Figure 3: Gender Gap in Labor Force Participation at Older Ages, 1962 to 2013: CPS



Source: CPS March.

Figure 4: Women Employed Full-Time, Full-Year among Labor Force Participants during the Year, 1963 to 2014: CPS

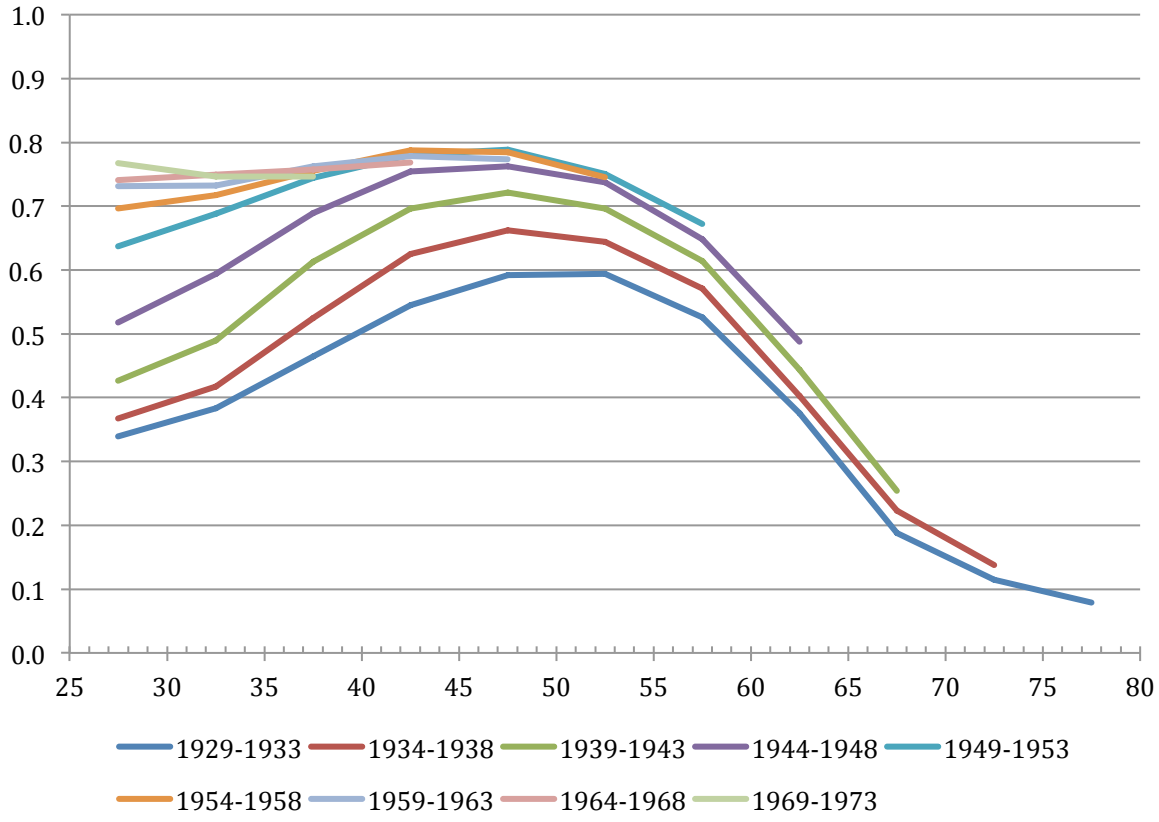


Source: CPS March.

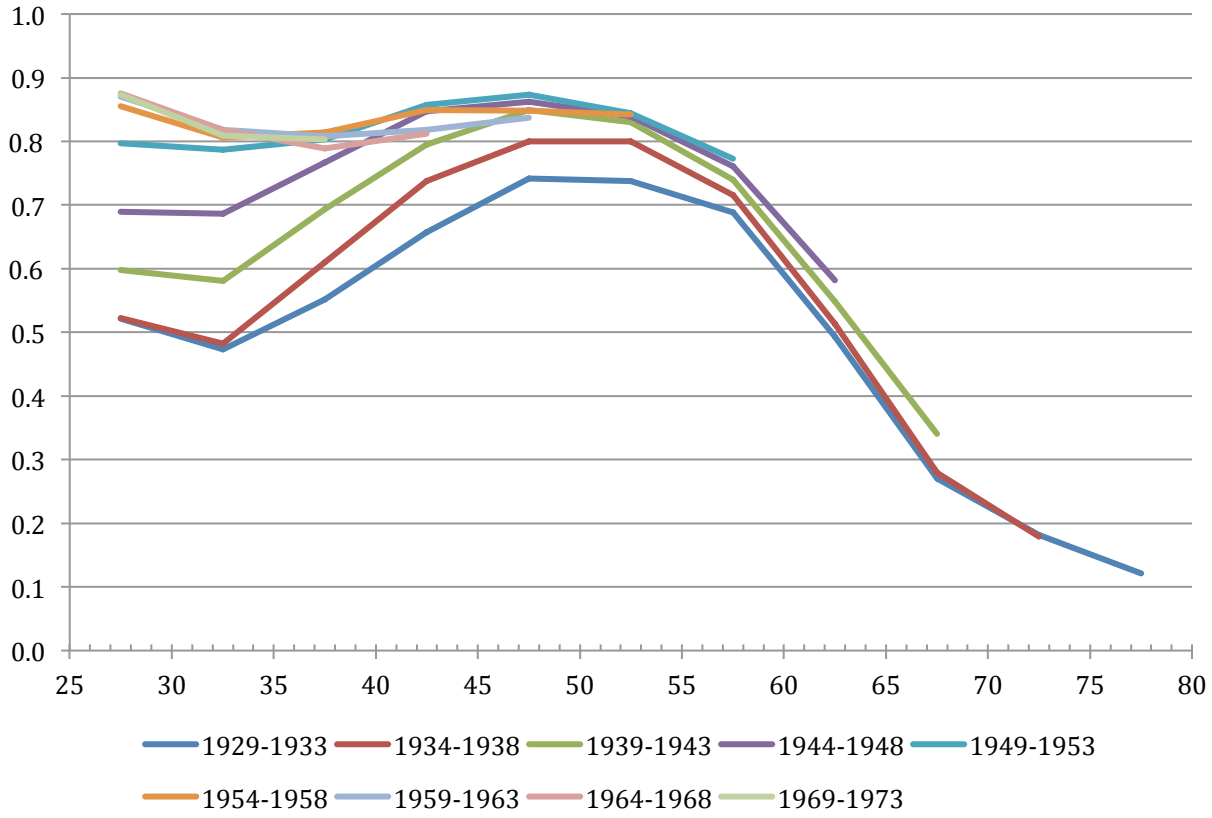
Notes: Both numerator and denominator refer to the calendar year. By a "labor force participant during the year" is meant anyone who worked during the year. Three-year centered moving averages are shown. Full-time, full-year workers are those who worked 40 or more weeks and 35 or more hours per week.

Figure 5: Labor Force Participation Rates for Women by Five-Year Birth Cohorts (1929-33 to 1969-73) and Five-Year Age Groups (25-29 to 74-79 years): CPS

A. All Education Groups



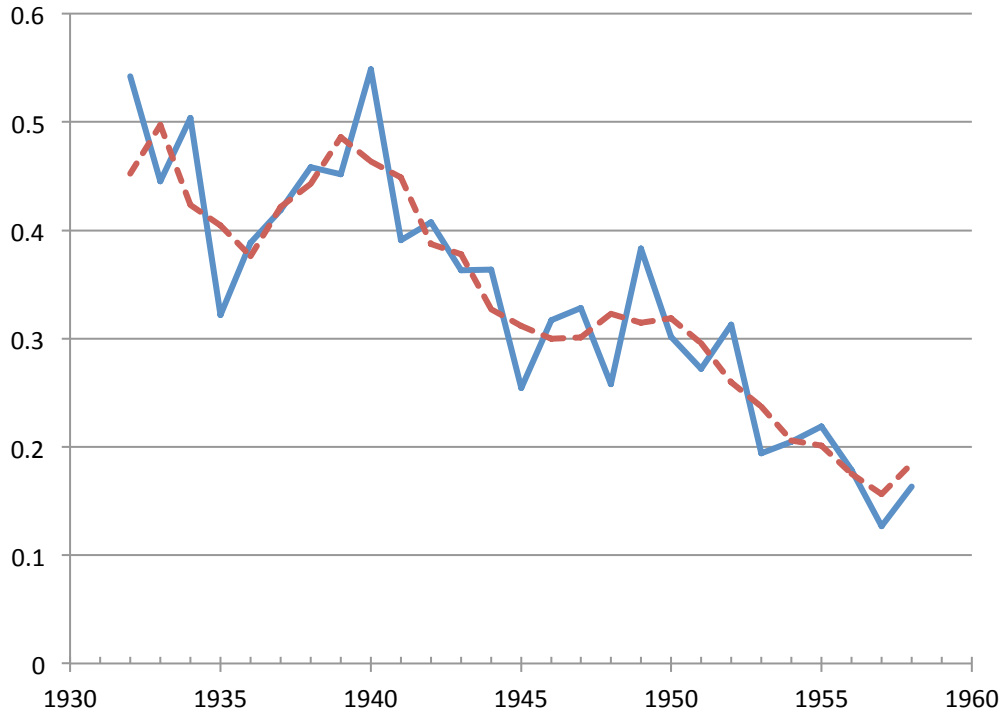
B. College Graduates



Source: CPS March.

Notes: Every data point in each graph contains 25 birth years and ages.

Figure 6: Fraction of College Graduate Women Ever Employed in Teaching, for 1931 to 1959 Birth Cohorts: HRS

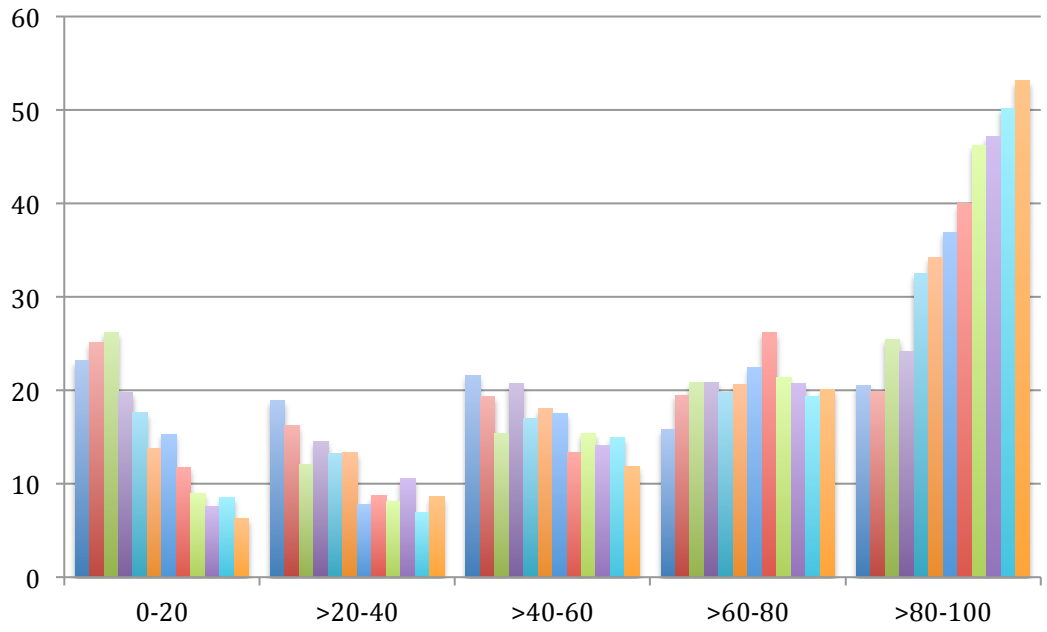


Source: HRS, Restricted Access Data.

Notes: "Ever employed in teaching" is calculated with code provided by Maria Fitzpatrick, which uses detailed occupations of respondents prior to their first HRS interview. The dashed line is the three-year centered moving average.

Figure 7: Lifecycle Labor Force Participation in the HRS and Social Security Earnings Data

A. All Women, 25 to 54 Years



B. All Women, 25 to 34 Years

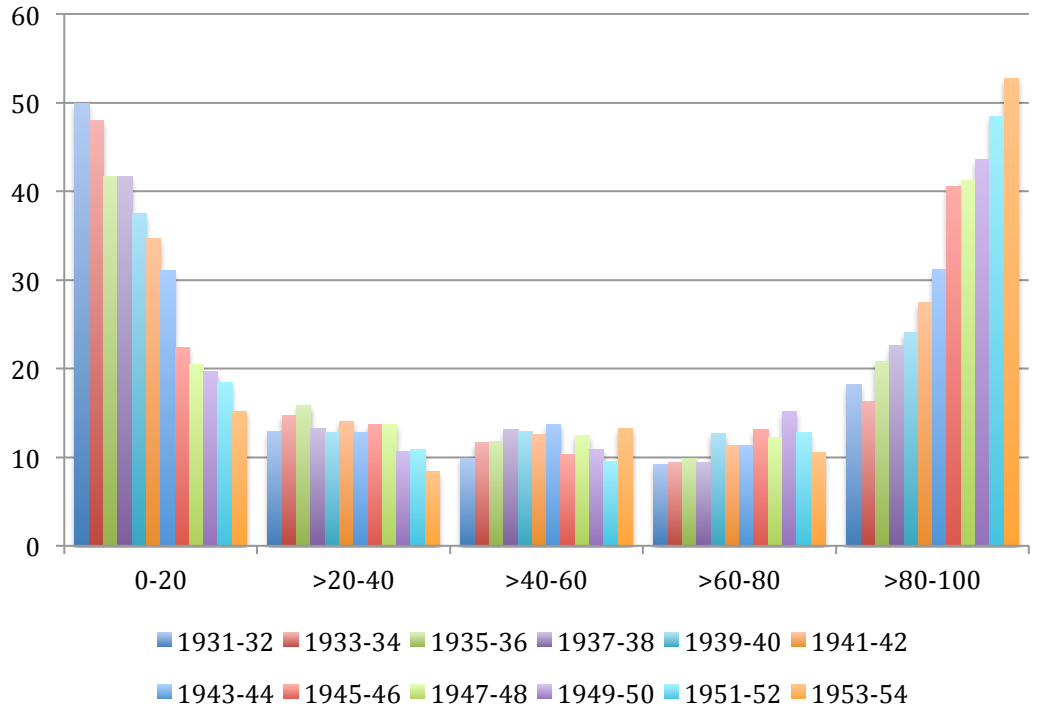
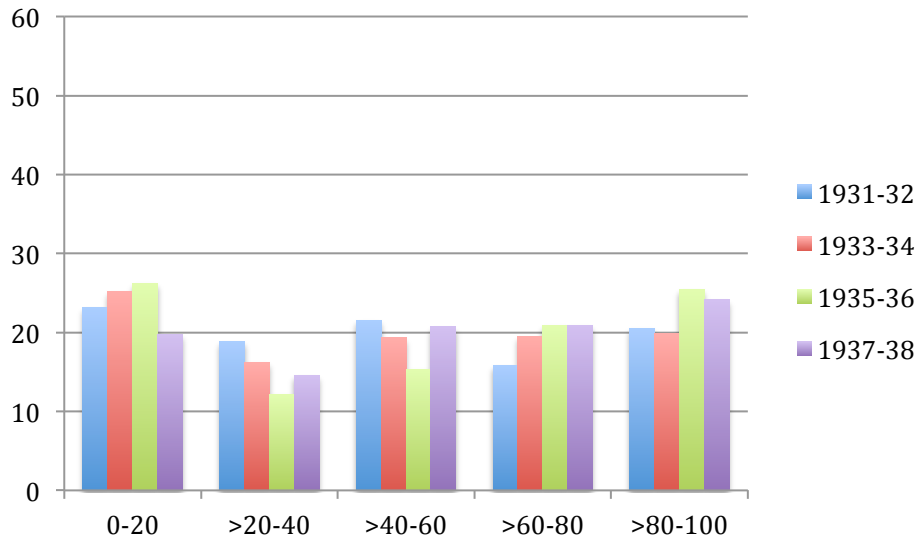


Figure 8: Lifecycle Labor Force Participation by Birth Cohorts, 25 to 54 Years Old: HRS

A. Birth Cohorts from 1931 to 1938



B. Birth Cohorts from 1947 to 1954

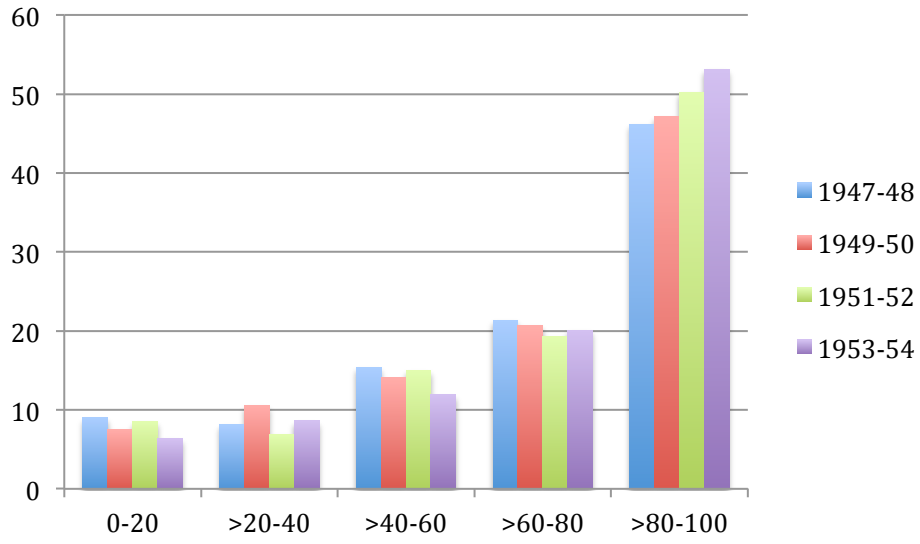
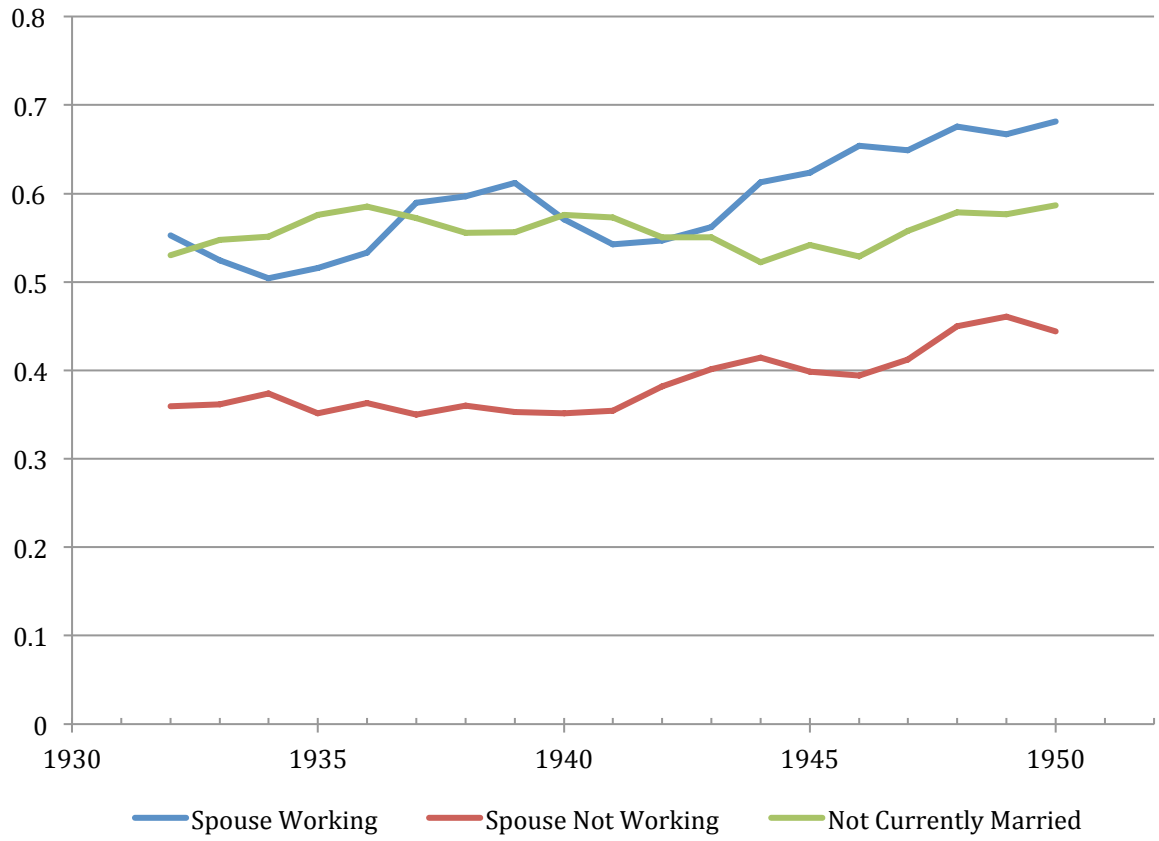


Figure 9: Labor Force Participation by Year of Birth, Current Marital Status and Husband's Employment for 59 to 63 Year Old Women



Source: HRS

Notes: All women 59 to 63 years old are included; person weights applied.

Table 1: Labor Force Participation Rates for Males and Females, 55 to 74 Years: CPS

Age Group	Educational Group	Labor Force Participation Rate in		Percentage Point Change c.1988 to c.2013
		1987-89	2012-14	
Women				
55-59	All	0.522	0.673	15.1
	College graduates	0.685	0.779	9.4
	Non-college grad.	0.499	0.627	12.8
60-64	All	0.341	0.514	17.3
	College graduates	0.454	0.612	15.8
	Non-college grad.	0.330	0.472	14.3
65-69	All	0.153	0.276	12.3
	College graduates	0.240	0.367	12.7
	Non-college grad.	0.145	0.244	9.9
70-74	All	0.072	0.157	8.6
	College graduates	0.130	0.214	8.3
	Non-college grad.	0.066	0.142	7.5
Men				
55-59	All	0.796	0.779	-1.8
	College graduates	0.886	0.896	1.0
	Non-college grad.	0.773	0.728	-4.6
60-64	All	0.548	0.607	5.9
	College graduates	0.682	0.727	4.5
	Non-college grad.	0.516	0.543	2.8
65-69	All	0.258	0.380	12.2
	College graduates	0.402	0.491	8.9
	Non-college grad.	0.231	0.321	9.0
70-74	All	0.155	0.232	7.7
	College graduates	0.254	0.324	7.0
	Non-college grad.	0.141	0.191	5.0

Sources: CPS March 1987, 1988, 1989, 2012, 2013, and 2014.

Table 2: Female Labor Force Participation at Age 59-63 Years, All Education Groups: HRS

	Labor Force Participation at Ages 59 to 63 Years					
	Full Sample	Linked Sample				
	(1)	(2)	(3)	(4)	(5)	(6)
Year of birth						
1934-36	-0.00810	-0.0117	-0.0144	-0.0165	-0.0134	-0.0151
	(0.0192)	(0.0207)	(0.0203)	(0.0197)	(0.0201)	(0.0186)
1937-39	0.0141	0.00418	0.000505	-0.0151	-0.00400	-0.00585
	(0.0191)	(0.0206)	(0.0202)	(0.0199)	(0.0202)	(0.0186)
1940-42	0.0163	0.0129	-0.00795	-0.0354	-0.0179	-0.0322
	(0.0203)	(0.0218)	(0.0211)	(0.0205)	(0.0209)	(0.0194)
1943-45	0.0464*	0.0330	-0.00191	-0.0390	-0.0164	-0.0326
	(0.0229)	(0.0247)	(0.0240)	(0.0235)	(0.0238)	(0.0223)
1946-48	0.0635**	0.0558*	0.00758	-0.0363	-0.0107	-0.0236
	(0.0217)	(0.0242)	(0.0240)	(0.0237)	(0.0240)	(0.0221)
1949-51	0.0973***	0.0897***	0.0320	-0.0117	0.0145	0.000859
	(0.0217)	(0.0250)	(0.0246)	(0.0243)	(0.0246)	(0.0231)
High school grad.			0.156***	0.117***	0.136***	0.0582***
			(0.0174)	(0.0170)	(0.0173)	(0.0166)
Some college			0.238***	0.189***	0.206***	0.102***
			(0.0198)	(0.0196)	(0.0200)	(0.0190)
College graduate			0.288***	0.231***	0.243***	0.114***
			(0.0247)	(0.0247)	(0.0253)	(0.0244)
MA			0.335***	0.270***	0.280***	0.154***
			(0.0275)	(0.0274)	(0.0285)	(0.0264)
PhD, MD, JD, etc.			0.465***	0.361***	0.395***	0.228***
			(0.0442)	(0.0441)	(0.0444)	(0.0440)
Lifecycle LFP 35-44				0.236***		0.219***
				(0.0168)		(0.0158)
Career cond. 35-44					0.131***	
					(0.0187)	
Currently married						-0.167***
						(0.0303)
Divorced						0.0785*
						(0.0313)
Widow						0.00653
						(0.0322)
Spouse in LF						0.189***
						(0.0141)
Health status	no	no	no	no	no	yes
Age dummies	Yes	yes	yes	yes	yes	yes

Race dummies	yes	yes	yes	yes	yes	yes
Constant	0.456***	0.428***	0.325***	0.257***	0.321***	0.167***
	(0.0331)	(0.0369)	(0.0344)	(0.0346)	(0.0347)	(0.0427)
N	18,383	15,431	15,431	15,431	15,431	15,431
R-squared	0.028	0.028	0.069	0.101	0.079	0.179

Sources: Health and Retirement Study (HRS) 1992 to 2012, RAND version with added variables from original HRS files. Social Security Administration earnings (and W-2) data are used to calculate life cycle labor force participation (Lifecycle LFP <ages>) and the career condition (Career cond. <ages>).

Notes: The dependent variable is 1 if the woman is in the labor force and 0 otherwise. The HRS asks respondents their labor force status and a woman is in the labor force if she reported being employed or unemployed and searching for work.
Health status is self-reported and is coded as 1 if "good" or better and 0 otherwise.
Marital status variables refer to current status.
"Lifecycle LFP <ages>" is the fraction of the interval the woman was in the labor force as determined by a combination of the data sources described in the Appendix.
The "linked sample" indicates that the individual gave permission for Social Security earnings data to be linked.
Omitted base group variables are: 1931-33 birth cohort; below high school graduate (overall or for the less than college graduate group); BA only for the college graduate group; never married; other race and age 59. Omitted from the table are dummy variables for missing variables regarding spouse in labor force, career condition 35-44, and health status.
HRS person weights have been used.

Standard errors in parentheses have been clustered at the individual level.
* p<0.05, ** p<0.01, *** p<0.001

Table 3: Female Labor Force Participation at Ages 59 to 63, by Education: HRS

	Labor Force Participation at Ages 59 to 63 Years					
	Not College Graduate	College Graduate	Not College Graduate	College Graduate	Not College Graduate	College Graduate
	(1)	(2)	(3)	(4)	(5)	(6)
Year of birth						
1934-36	-0.0238	0.0297	-0.0269	0.0327	-0.0254	0.0372
	(0.0219)	(0.0525)	(0.0212)	(0.0509)	(0.0200)	(0.0473)
1937-39	-0.00217	0.0170	-0.0225	0.0221	-0.0110	0.0165
	(0.0217)	(0.0545)	(0.0212)	(0.0548)	(0.0197)	(0.0522)
1940-42	-0.0139	0.0374	-0.0452*	0.0205	-0.0407	0.0169
	(0.0231)	(0.0515)	(0.0223)	(0.0506)	(0.0211)	(0.0473)
1943-45	-0.00774	0.0521	-0.0460	0.0213	-0.0380	0.0198
	(0.0268)	(0.0538)	(0.0261)	(0.0537)	(0.0246)	(0.0521)
1946-48	0.00184	0.0591	-0.0436	0.0226	-0.0309	0.0340
	(0.0271)	(0.0524)	(0.0266)	(0.0528)	(0.0250)	(0.0484)
1949-51	0.00297	0.135**	-0.0461	0.110*	-0.0270	0.104*
	(0.0287)	(0.0497)	(0.0282)	(0.0505)	(0.0268)	(0.0486)
High school grad.	0.154***		0.113***		0.0530**	
	(0.0174)		(0.0171)		(0.0167)	
Some college	0.237***		0.184***		0.0982***	
	(0.0198)		(0.0197)		(0.0192)	
MA		0.0471		0.0392		0.0348
		(0.0311)		(0.0307)		(0.0289)
PhD, MD, JD, etc.		0.172***		0.141**		0.114*
		(0.0472)		(0.0483)		(0.0481)
Lifecycle LFP 35-44			0.209***	0.0774	0.193***	0.0779
			(0.0267)	(0.0558)	(0.0249)	(0.0559)
Never in LF 35-44			-0.0592*	-0.151*	-0.0567*	-0.142*
			(0.0248)	(0.0700)	(0.0232)	(0.0679)
Currently married					-0.152***	-0.254***
					(0.0345)	(0.0625)
Divorced					0.0622	0.117
					(0.0359)	(0.0608)
Widow					0.00372	-0.0112
					(0.0363)	(0.0695)
Spouse in LF					0.180***	0.229***
					(0.0155)	(0.0345)
Health status	no	no	no	no	yes	yes
Age dummies	yes	yes	yes	yes	yes	yes
Race dummies	yes	yes	yes	yes	yes	yes

Constant	0.284***	0.850***	0.256***	0.830***	0.169***	0.628***
	(0.0365)	(0.0640)	(0.0415)	(0.0754)	(0.0509)	(0.0993)
N	12,789	2,642	12,789	2,642	12,789	2,642
R-squared	0.056	0.041	0.094	0.059	0.173	0.142

Sources: Health and Retirement Study (HRS) 1992 to 2012, RAND version with added variables from original HRS files. Social Security Administration earnings (and W-2) data are used to calculate life cycle labor force participation (Lifecycle LFP <ages>).

Notes: The dependent variable is 1 if the woman is in the labor force and 0 otherwise. The HRS asks respondents their labor force status and a woman is in the labor force if she reported being employed or unemployed and searching for work.
Health status is self-reported and is coded as 1 if "good" or better and 0 otherwise.
Marital status variables refer to current status.
"Lifecycle LFP <ages>" is the fraction of the interval the woman was in the labor force as determined by a combination of the data sources described in the Appendix. "Never in LF" is 1 if the individual was recorded as having no years in the labor force during those years.
All columns use the "linked sample."
Omitted base group variables are: 1931-33 birth cohort; below high school graduate (overall or for the less than college graduate group); BA only for the college graduate group; never married; other race and age 59. Omitted from the table are dummy variables for missing variables regarding spouse in labor force and health status.
Regressions are estimated separately for college graduates and those who did not graduate from college. For the college graduates degrees beyond a bachelors are added (MA, PhD, etc.), where MA includes all masters and PhD, MD, JD, etc. includes all graduate and professional degrees. For non-college graduates dummy variables are added for those with a high school diploma and some college.
HRS person weights have been used.

Standard errors in parentheses have been clustered at the individual level.
* p<0.05, ** p<0.01, *** p<0.001

Table 4: Labor Force Participation among College Graduate Women at Ages 59 to 63: HRS

	Labor Force Participation at Ages 59 to 63 Years				
	Full Sample	Linked Sample			
	(1)	(2)	(3)	(4)	(5)
Year of birth					
1934-36	0.00774	0.0261	0.0287	0.0438	0.0270
	(0.0497)	(0.0527)	(0.0510)	(0.0473)	(0.0511)
1937-39	0.00217	0.0155	0.0207	0.0193	0.00826
	(0.0503)	(0.0545)	(0.0547)	(0.0510)	(0.0540)
1940-42	0.0389	0.0342	0.0167	0.0195	0.00219
	(0.0484)	(0.0518)	(0.0508)	(0.0471)	(0.0508)
1943-45	0.0370	0.0412	0.00886	0.0115	-0.0214
	(0.0503)	(0.0542)	(0.0539)	(0.0520)	(0.0543)
1946-48	0.0465	0.0482	0.00996	0.00406	-0.0314
	(0.0475)	(0.0528)	(0.0529)	(0.0496)	(0.0539)
1949-51	0.0957*	0.127*	0.101*	0.113*	0.0593
	(0.0452)	(0.0500)	(0.0508)	(0.0465)	(0.0511)
Ever a teacher	-0.0477	-0.0494	-0.0561	-0.0920**	-0.0616*
	(0.0288)	(0.0311)	(0.0306)	(0.0289)	(0.0301)
MA	0.0538	0.0554	0.0486	0.0386	0.0402
	(0.0296)	(0.0321)	(0.0316)	(0.0295)	(0.0312)
PhD, MD, JD, etc.	0.160***	0.169***	0.137**	0.111**	0.112*
	(0.0447)	(0.0460)	(0.0470)	(0.0406)	(0.0456)
Lifecycle LFP 35-44			0.0714		
			(0.0559)		
Never in LF 35-44			-0.162*		
			(0.0703)		
Lifecycle LFP 45-54				0.384***	
				(0.0706)	
Never in LF 45-54				-0.214*	
				(0.0875)	
Lifecycle LFP 25-54					0.368***
					(0.0607)
Health status	no	no	no	no	no
Age dummies	yes	yes	yes	yes	yes
Race dummies	yes	yes	yes	yes	yes
Marital status dummies	yes	yes	yes	yes	yes
Job status of husband	yes	yes	yes	yes	yes
Constant	0.855***	0.868***	0.856***	0.572***	0.667***
	(0.0591)	(0.0645)	(0.0762)	(0.0861)	(0.0712)

N	3,137	2,642	2,642	2,642	2,642
R-squared	0.040	0.044	0.062	0.149	0.081

Sources: Health and Retirement Study (HRS) 1992 to 2012, RAND version with added variables from original HRS files. Social Security Administration earnings (and W-2) data are used to calculate life cycle labor force participation (Lifecycle LFP <ages>) and the career condition (Career cond. <ages>).

Notes: The dependent variable is 1 if the woman is in the labor force and 0 otherwise. The HRS asks respondents their labor force status and a woman is in the labor force if she reported being employed or unemployed and searching for work.
*Health status is self-reported and is coded as 1 if "good" or better and 0 otherwise.
Marital status variables refer to current status.
"Lifecycle LFP <ages>" is the fraction of the interval the woman was in the labor force as determined by a combination of the data sources described in the Appendix. "Never in LF" is 1 if the individual was recorded as having no years in the labor force during those years.
The "linked sample" indicates that the individual gave permission for Social Security earnings data to be linked.
Omitted base group variables are: 1931-33 birth cohort; BA only for the college graduate group; never married; and other race and age 59. Omitted from the table are dummy variables for missing variables regarding spouse in labor force and health status.
HRS person weights have been used.

Standard errors in parentheses have been clustered at the individual level.
* $p<0.05$, ** $p<0.01$, *** $p<0.001$

References

Blau, David M. 1998. "Labor Force Dynamics of Older Married Couples," *Journal of Labor Economics* 16(3): 595–629.

Goldin, Claudia. 1989. "Life-Cycle Labor-Force Participation of Married Women: Historical Evidence and Implications," *Journal of Labor Economics* 7(1) January: 20-47.

Goldin, Claudia. 1997. "Career and Family: College Women Look to the Past." In F. Blau and R. Ehrenberg, eds., *Gender and Family Issues in the Workplace*. New York: Russell Sage Press: 20-58.

Gustman, Alan and Thomas Steinmeier. 2009. "How Changes in Social Security Affect Recent Retirement Trends," *Research on Aging* (March 2009) 31 (2): 261-90.

Heckman, James J. and Robert J. Willis. 1977. "A Beta-logistic Model for the Analysis of Sequential Labor Force Participation by Married Women," *Journal of Political Economy* 85 (February): 27-58.

Hurd, Michael and Susann Rohwedder. 2014. "Predicting Labor Force Participation of the Older Population." http://siepr.stanford.edu/system/files/shared/events/wlc2014/Hurd-Rohwedder-paper.pdf

Maestas, Nicole and Julie Zissimopoulos. 2010. "How Longer Working Lives Ease the Crunch of Population Aging," *Journal of Economic Perspectives* 24(1) (Winter): 139-60

Mastrobuoni, Giovanni. 2009. "Labor Supply Effects of the Recent Social Security Benefit Cuts: Empirical Estimates Using Cohort Discontinuities," *Journal of Public Economics* 93: 1224-33.

Olivetti, Claudia. 2006. "Changes in Women's Hours of Market Work: The Role of Returns to Experience," *Review of Economic Dynamics* 9(4) October: 557-87.

Sandy Chien, Nancy Campbell, Orla Hayden, Michael Hurd, Regan Main, Josh Mallett, Craig Martin, Erik Meijer, Angela Mui, Michael Moldoff, Susann Rohwedder, and Patricia St. Clair. 2013. *HRS Data Documentation, Version M*. (August). RAND, Labor and Population Program. Santa Monica, CA.

Schirle, Tammy. 2008. "Why Have the Labor Force Participation Rates of Older Men Increased since the Mid-1990s?," *Journal of Labor Economics* 26(4) (October): 249-94.

Appendix[24]

Health and Retirement Study: General Comments

The Health and Retirement Study (known as the HRS and also as the University of Michigan Health and Retirement Study) is a widely used data set. More information can be found at: http://hrsonline.isr.umich.edu/. This brief Appendix will discuss certain details of importance to this paper.

The Health and Retirement Study, supported by the National Institute on Aging and the Social Security Administration, was begun in 1992 with a random sample of households in which one member was born between 1931 and 1941 and thus between 51 and 61 years old. This initial sample is known as the HRS cohort, also as the "Intermezzo" cohort. In households containing a married or partnered couple, the "spouse" and "respondent" categories were randomly assigned. "Spouses" were not given positive sample weights until 1998, if born from 1931 to 1941. If they were born from 1942 to 1947, they are not given positive sample weights until the "War Baby" (WB) cohort was added. The "War Baby" (WB) cohort was born 1942 to 1947. The "Early Baby Boomer" (EBB) cohort, born 1948 to 1953, was added in 2004. The Mid-Boomer (MBB) cohort, 1954 to 1959 was added in 2010. The WB, EBB, and MBB cohorts were between 51 and 56 years old.

The cohorts mentioned have all been surveyed every two years. Additional cohorts born before 1931 are also part of the HRS, but the HRS, WB and EBB are the primary ones we have used in our research. At the time of this writing, the HRS data are available to 2012.

Health and Retirement Study: Construction of Variables

Lifecycle Labor Force Participation

Lifecycle labor force participation is intended to measure the fraction of a period during which the individual was in the labor force. The time period we consider is from 25 to 54 years and we subdivide that into three decades. We primarily use the information from the Social Security earnings records (and the W-2 forms after 1977) to figure out whether an individual was employed during a year. We can do this only for individuals who gave permission to the HRS to link their survey to their Social Security earnings records. On average 80 percent of the sample agree to this linkage.

[24] Some of the material in this section will be included in an overall appendix for all volume papers that use the HRS.

In general, we define someone as a labor force participant if during a year their annual earnings were at least equal to the federal minimum wage in that year times 10 hours times 52 weeks. Complications arise because some individuals were exempt from the Social Security earnings tax. These exempt employees were generally government workers and for our sample of women, teachers would have been an important exempt category. During the initial interview the HRS asked whether the individual had been employed by the government (including municipal, state and federal government positions) and if that was the case, the person could list two periods of employment. We count the individual in the labor force if the person did not pay the Social Security earnings tax in some year but stated that their employment was in the government for that period. It should be noted that when the W-2 forms become available, there is no problem with exempt status since the forms include all W-2 income. In addition, some HRS respondents were surveyed when they were in their early fifties and we use the HRS survey data when it exists. Thus we determine labor force status on the basis of various pieces of information including the HRS survey, the Social Security earnings records and the W-2 forms.

Career Condition Variable

Similar to the construction of the lifecycle labor force variable, we create a "career condition" variable that assesses whether individuals who were in the labor force earned above some amount. The amount is given by some fraction (we have used both 0.5 and 0.75) of the median annual wage of a (full-time, full-year) male worker in the given year. Because the period we are considering predates the micro-data for the CPS, we use the published surveys to obtain the male median annual wage. In our empirical work we define the career condition between ages 35 and 44 years ("Career cond. 35-44") as the fraction of years in the age interval the individual exceeded 50 percent of the earnings of the median male full-time, full-year worker.

Earnings data for this calculation are obtained primarily from the Social Security earnings records, the W-2 forms when available, and the HRS earnings data when it exists for the individual. If the individual was in a tax-exempt employment (and did not have W-2 or HRS earnings data), we assume that the income was sufficient to exceed the given "career condition."

Social Security Earnings Record Linkage in the HRS

The fraction of female HRS respondents who agreed at some point to the linkage of their HRS study to their Social Security earnings record is fairly high. Just around 11 percent are not linked from the 1931 to 1942 birth cohorts. The fraction increases to 15 percent for 1943 to 1945 and then to 21 percent for 1946 to 1948. The high rate of non-

linkage for the 1950s cohorts is probably because they have had fewer years to agree to have their records linked since respondents are asked during each wave. The fraction not linked will probably fall during the next waves of the HRS as more respondents agree to the linkage.

Individuals who did not agree to the linkage do not differ on the basis of educational attainment and current marital status with those who did agree. The main determinant of linkage is the number of years the individual has been in the data set and, therefore, how many times the individual has been asked permission for the linkage.

Comparisons of the HRS and the CPS

Labor Force Participation Rates

The HRS closely tracks the labor force participation rates given by the CPS for the same age groups and birth cohorts. The only major differences occur for those 65 years and older. The HRS labor force data are always greater than the CPS data in every year from 1992 to 2010 for these age groups but are not for the younger groupings. The precise reason is unclear. One possibility is that the CPS does a better job interviewing individuals in group quarters.

Marital Status and Education

Both the HRS marital status and education variables track the CPS well for individuals 51 to 56 years old. Cohorts born from 1935 to 1952 are given in Appendix Tables A2 and A3. The HRS samples are fairly small and are subject to considerable sampling error. It should be noted that the education distributions for 51 to 56 year old women differ from those for younger women in the same birth cohorts due to a common phenomenon that individuals gain education, for real or fictitious reasons, as they age.

The HRS contains a potential complication because some people did not list a degree and their highest degree was inferred. According to the RAND HRS Codebook (Chien 2013, pp. 132-33): "The highest degree is assigned by looking at reports from Tracker and all waves of data. The first non-missing value is used." When the actual degree is missing, it is imputed and a bachelor's degree is given to those with RAEDYRS = 16 or 17. Interestingly, the HRS and the CPS data for the same age groups and birth cohorts is remarkably similar.

Children Ever Born

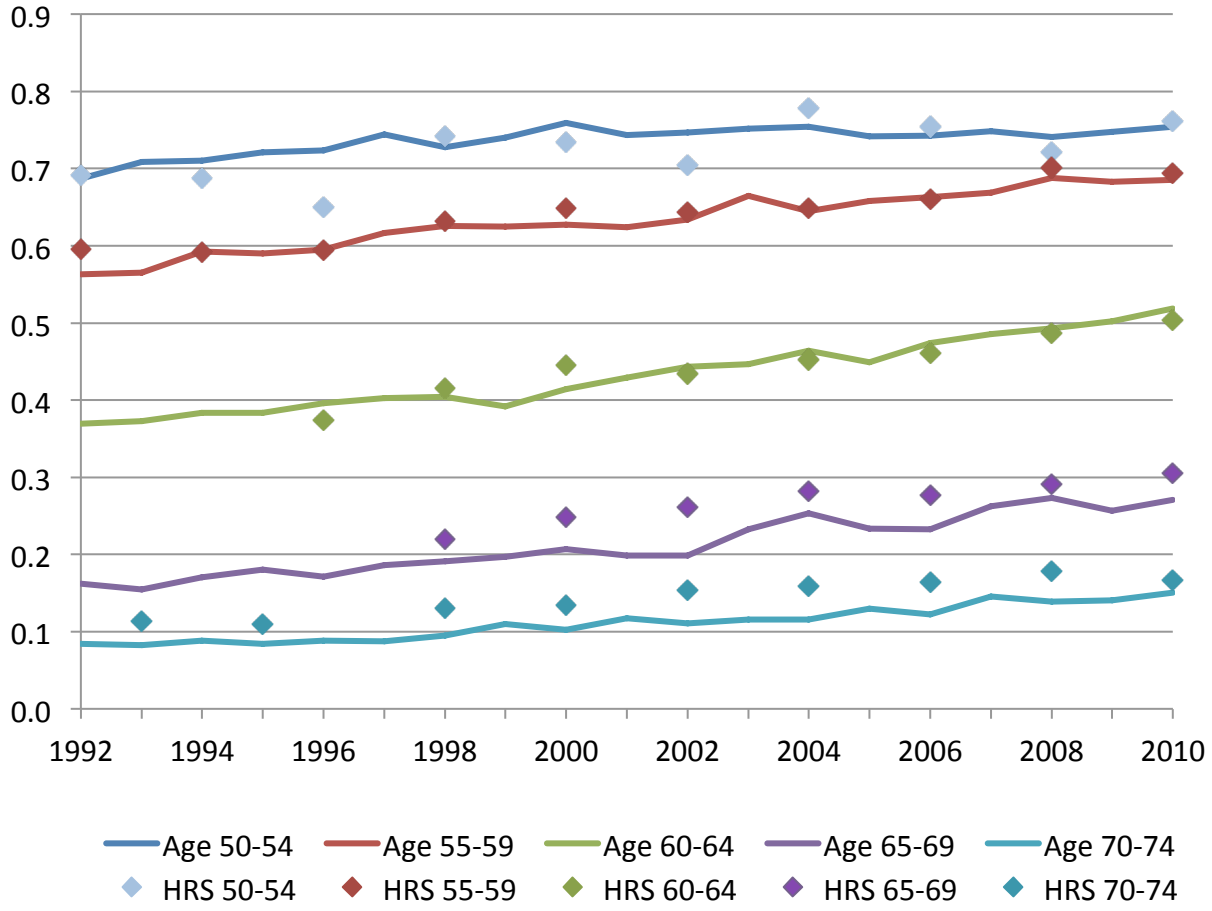
The information on children ever born in the HRS differs in various ways from that in the CPS June Fertility Supplements. But the mean number of children for the same

cohorts in each of the two sources is not much different. It appears that the main difference is that the fraction of women who report no births in the HRS is lower than reported in the CPS. For example, for women with a college degree born from 1947 to 1951 the fraction with zero births in the HRS (all of the respondents are older than 44 years) is 19 percent. But in the CPS the fraction with zero births at 40 to 44 years old is about 25 percent. For women with less than a college degree, the fraction with no births in the HRS for those born for 1947 to 1951 is 10 percent but is 13 percent in the CPS June Fertility Supplements.

Even though HRS respondents report a lower fraction with no birth, the mean number of children ever born, as given in Appendix Figure A2, is very similar to that given in the CPS June Fertility Supplements. The HRS number is almost always slightly higher, especially for cohorts from after 1945.

One possibility is that women in the HRS are also including adopted and step children. That possibility has been explored and does not appear to be the source of the difference.

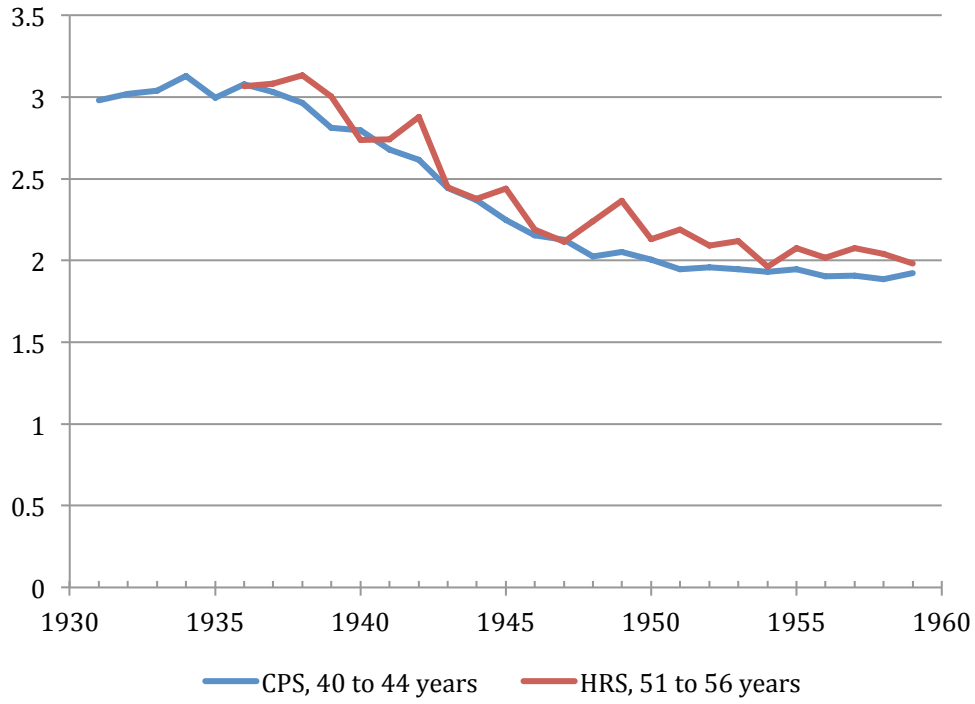
Appendix Figure A1: Comparing Labor Force Participation for the HRS and the CPS: Women 50-54 to 70-74 Years



Sources: March CPS; HRS.

Notes: The HRS is a biennial survey. Some age groups are not shown for the HRS because the group is incomplete and the participation rate would be biased since it would omit some of the older ages in the group.

Appendix Figure A2: Children Ever Born for Women 51 to 56 Years Old in HRS and 40 to 44 Years Old in CPS in Birth Cohorts 1936 to 1951



Sources: HRS; CPS June Fertility Supplements, micro-data (1973 to 2014).

Notes: HRS person weights are used; no weights are used for the CPS. Children ever born is truncated below 10 in both samples. In both data sets, the variable is supposed to give the number of children ever born to the respondent and not the number of live children or adopted or step children

Appendix Table A1: Fraction of Female HRS Respondents Linked to Social Security Earnings Records by Birth Cohort

Birth Years	Fraction Linked
1931-33	0.886
1934-36	0.888
1937-39	0.868
1940-42	0.893
1943-45	0.852
1946-48	0.790
1949-51	0.714
1952-54	0.682

Sources: HRS, Restricted Access Data.

Notes: Person weights used. Linkage uses HRS to 2012.

Appendix Table A2: Comparing Marital Status for the HRS and CPS: Women 51 to 56 Years

	Fraction Currently Married			Fraction Ever Married but not currently married			Fraction Never Married		
Year of Birth	HRS-SS	Full HRS	CPS	HRS-SS	Full HRS	CPS	HRS-SS	Full HRS	CPS
1931-32	.	.	0.730	.	.	0.230	.	.	0.040
1933-34	.	.	0.731	.	.	0.225	.	.	0.044
1935-36	0.737	0.730	0.716	0.236	0.246	0.237	0.026	0.024	0.047
1937-38	0.741	0.730	0.712	0.230	0.240	0.244	0.029	0.029	0.044
1939-40	0.739	0.734	0.704	0.214	0.222	0.244	0.047	0.043	0.051
1941-42	0.641	0.647	0.684	0.330	0.322	0.262	0.030	0.030	0.054
1943-44	0.552	0.537	0.678	0.378	0.387	0.273	0.070	0.075	0.049
1945-46	0.716	0.731	0.683	0.229	0.221	0.259	0.054	0.047	0.058
1947-48	0.647	0.643	0.669	0.302	0.305	0.269	0.052	0.052	0.063
1949-50	0.628	0.604	0.666	0.314	0.328	0.259	0.058	0.067	0.075
1951-52	0.676	0.684	0.661	0.266	0.266	0.261	0.058	0.049	0.078

Sources: HRS. Restricted Access Data for HRS-SS columns. CPS.

Notes: The HRS-SS columns refer to the sample linked to the Social Security Administration earnings data. The number of observations for the HRS sample is about 500 for the 1949-50 and 1951-52 birth cohorts and 1,000 for the 1937-38 and 1939-40 birth cohorts. Missing values (.) indicate lack of coverage using the particular HRS cohorts.