

Supplemental Appendix

Social Security and Retirement of Older Blacks and Hispanics

By Emma Aguila and Zeewan Lee

Appendix A: Sample Description

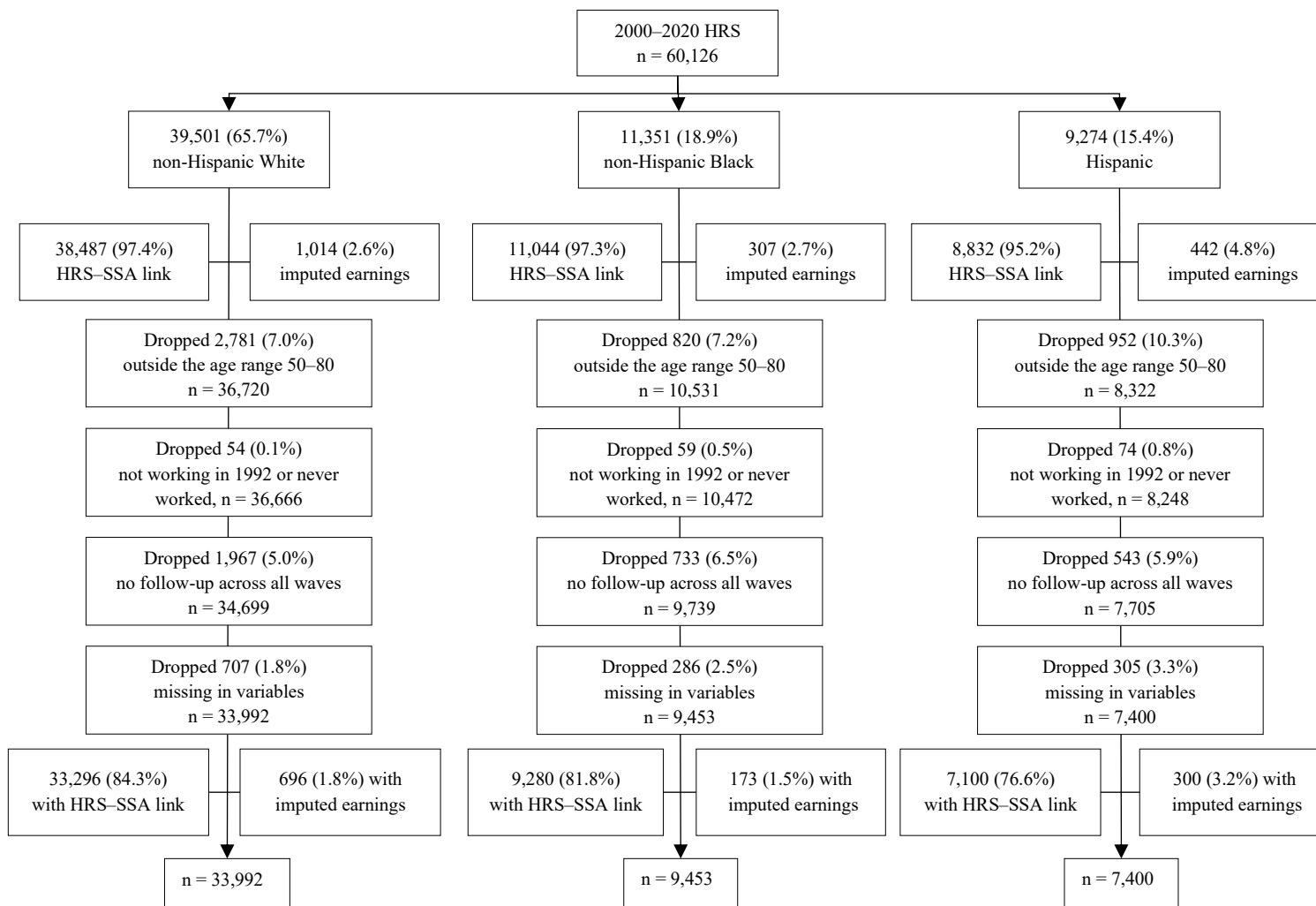


Figure A1. Sample Description using HRS 2000-2020

Source: author's calculations using HRS 2000-2020 (RAND HRS, 2025).

The 2000-2020 Health and Retirement Study (HRS) includes 39,501 non-Hispanic Whites (hereafter ‘Whites’), 11,351 non-Hispanic Blacks (hereafter ‘Blacks’), and 9,274 Hispanics, as shown in Figure A1. The HRS has earnings history through linkages to U.S. Social Security Administration (SSA) data (summary earnings data for HRS respondents) for 38,487 Whites, 11,044 Blacks, and 8,832 Hispanics (Health and Retirement Study (HRS)). For the remainder of each group (1,014 Whites, 307 Blacks, and 442 Hispanics), we imputed past earnings. We excluded from our analysis 2,781 (7.0%) Whites, 820 (7.2%) Blacks, and 952 (10.3%) Hispanics for being outside our target age range of 50 to 80. We then dropped 54 of the Whites (0.1% of 39,501), 59 Blacks (0.5% of 11,351), and 74 Hispanics (0.8% of 9,274) who were in the 2000 wave (the first year of our analysis) but who did not work that year, or who had never worked. We removed 1,967 (5.0% of 39,501) Whites, 733 (6.5% of 11,351) Blacks, and 543 (5.9% of 9,274) Hispanics without follow-up interviews after their first survey. Finally, we dropped 707 (1.8% of 39,501) Whites, 286 (2.5% of 11,351) Blacks, and 305 Hispanics (3.3% of 9,274) because they had missing covariates. That left us with a sample of 33,992 Whites, 9,453 Blacks, and 7,400 Hispanics.

Appendix B: Social Security Benefits, Wealth, and Peak Value Calculation

We calculated each respondent’s Social Security benefits as follows. First, we obtained the *past annual earnings* for sample respondents from age 25 onward. For respondents with HRS-SSA linkages, we used SSA records of past earnings. For individuals with missing HRS-SSA linkages, as well as for HRS respondents whose labor statuses were shown as working but whose earnings were incorrectly recorded as zero or negative, we imputed past annual earnings. We used a multiple imputation technique that involved iterative stochastic imputation. This method allowed for imputation of zero as a possible value as well as imputation of earnings brackets used in the surveys to recover nonresponse (Aguila et al. 2023). We used age, sex, cohort, race, and education covariates for imputations conducted separately by race/ethnicity, i.e. Whites, Blacks, and Hispanics. We ensured that the distribution of imputed values matched that of observed incomes in the HRS-SSA data (the SSA summary earnings data for HRS respondents), thereby accounting for the uncertainty around the true earnings value that was missing (Johnson and Young 2011; White et al. 2011). We generated five imputations of earnings for each person-year observation, and set the mean of the five imputations as the value to be used for estimating individuals’ Social Security wealth (Ni et al. 2005; Lokupitiya R., Lokupitiya E. and Paustian 2006; White et al. 2018). If the imputed value of earnings was (a) equal to or below zero or (b) above the 99th percentile of the unimputed earnings distribution in the SSA-HRS linked records for each year, we re-imputed the earnings. Our final earnings distribution—including both the imputed values and observed values—was similar to the distribution of observed earnings in the HRS-SSA linked data.

Next, we imputed *future annual earnings* (i.e., projected earnings past 2020) of respondents by assuming their earnings would increase 1% every year until they reached the maximum age (T) of 120. Previous studies assumed all respondents work continuously in their lifetime when estimating future annual earnings (e.g., Coile and Gruber 2001, 2007). In this study, we departed from the strict assumption and estimated respondents’ employment gaps. We examined respondents’ working histories and applied similar patterns of employment gaps to future imputed earnings. Specifically, we counted the actual number of months the respondents worked between 1992 and 2020 and applied the same number of months worked

randomly to every 29-year timeframe for the future years until individuals reached age 120.^{1,2} Finally, we applied the imputed future employment patterns to estimate *future annual earnings* more accurately. If a respondent was expected to work only 6 months out of 12 in 2030, the projected annual earnings for 2030 was halved to reflect his work trajectories.

Following the approximation of the past and projected annual earnings, we adjusted respondents' lifetime earnings by the yearly National Average Wage Indices, which adjusted for inflation across years (U.S. Social Security Administration 2019a). Then, we applied 'earnings cap' to respondents' annual income, which ensured that their income is counted only up to the indexed cap in our Social Security benefits calculation (U.S. Social Security Administration 2019a). The cap is the first key redistributive feature of the Social Security formula. Under the year-specific earnings cap (set at \$137,700 in 2020), individuals do not pay Social Security taxes and hence have no additional Social Security benefits for earnings above the cap (U.S. Social Security Administration n.d.). This ensures that high-income earners do not claim extremely high Social Security benefits relative to those earning below the cap. For simplicity in our calculations, we also excluded any cost-of-living adjustments to which recipients were entitled.

From the capped and indexed annual salaries that respondents earned across their entire lifetime, we selected the 35 highest-earning years. We then calculated each respondent's *Average Indexed Monthly Earnings* (AIME) by dividing the sum of his highest 35 annual earnings by 420 (the number of months in 35 years). Any employment gaps (for individuals who did not work all 12 months in a year) in individuals' working trajectories led to lower Social Security benefits, all else equal.

Once we computed the AIME for each respondent, we used a nonlinear function to estimate Social Security benefits an individual would receive should this person retire and begin claiming at the full retirement age (FRA)—also known as the *Primary Insurance Amount* (PIA). The nonlinear function summed three separate percentages of portions of the AIME predetermined by the Social Security Administration to estimate an individual's PIA. The PIA is the sum of (1) 90 percent of the earnings up to the first threshold or bend-point, (2) 32 percent of the remaining earnings up to the second threshold, and (3) 15 percent of the remainder. The full list of thresholds or bend-points that vary by year can be found at the Social Security Administration webpage (U.S. Social Security Administration n.d.). These separate percentages serve as the second key redistributive feature of the Social Security wealth formula. We computed PIAs for each retirement age from current year t to 120.

Next, using the PIAs, we computed Social Security wealth (SSW_t) as the expected net present value of a worker's Social Security benefits received until death if retiring at age t , using the following equation:

¹ We did not consider individuals' mortality in the imputation of future earnings or working trajectories. However, individuals' gender and race-specific year-to-year survival probabilities are reflected in the Social Security wealth calculation.

² Respondents' earnings prior to 1992 (the year of the first HRS survey) were based on actual SSA records. These records already reflected any truncation in respondents' work histories. For instance, if an individual worked only five months in one year, the total income reported for the year was a sum of the earnings for those five months. Because the work histories are accurately reflected in the past earnings for the respondents, we did not apply our newly identified work trajectories in such cases.

$$(1) \quad SSW_t = \sum_{s=r}^S \frac{pr_{s|t} * B_s}{(1+d)^{s-t}},$$

This formula allowed us to account for the difference between each respondent’s retirement age and the FRA as well as survival probabilities and discount rates. In the above equation, $pr_{s|t}$ was the probability the individual was alive at time s conditional on being alive at time t . B_s was the primary insurance amount (PIA) to be collected on a monthly basis if the individual chose to retire at time s (i.e., $s=r$, where r denoted the timing of retirement). S was the year in which the individual died. We set d , the real discount rate, to 3 percent,³ and S , the maximum possible age reachable by an individual, equal to age 120 according to U.S. life-expectancy tables (Arias et al. 2016).

Next, we computed the survival probabilities ($pr_{s|t}$) as $pr_{s|t} = \prod_t^{s-1} (1 - \lambda_t)$. Here, λ_t was a hazard function where $\lambda_t = \frac{d_t}{S_t}$, with d_t denoting the number of people dying in period t , and S_t denoting the number of survivors at time t . We obtained information on survival and mortality prospects for the computation of $pr_{s|t}$ from Life Tables 2000 to 2020, published by the U.S. Centers for Disease Control and Prevention (National Center for Health Statistics 2025). Different demographic groups vary significantly in survival probabilities, with Hispanics and females outliving their counterparts (Case and Deaton 2015). Accounting for these differences, we applied different survival probabilities by gender and race/ethnicity, generating six distinct probabilities for male and female among Whites, Blacks and Hispanics.

Figure A2 shows our calculations of the median survival probabilities by race/ethnicity. Specifically, Figure A2 shows the survival probabilities from age x to $x+1$, $x+2$, $x+3$... starting at age 50, separately by race/ethnicity. Survival probabilities were highest for Hispanics, followed by Whites and Blacks. This confirmed research showing mortality increasing for Whites in recent decades while remaining lowest for Hispanics.

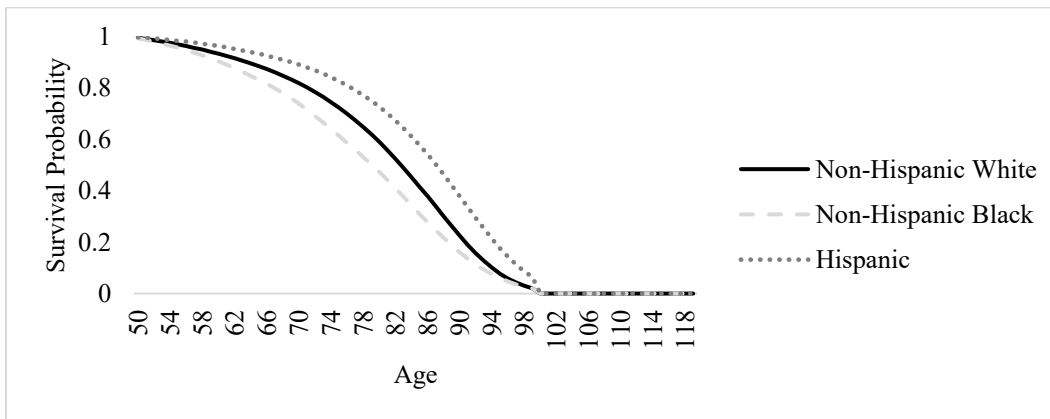


Figure A2. Average Survival Probability by Race/Ethnicity
Source: author’s calculations based on the CDC Life Table 2012 (Arias, Heron and Xu 2016).

³ This follows the practice of previous literature (e.g. Börsch-Supan 1992; Coile and Gruber 2001, 2000).

We computed Social Security wealth to be nonzero for individuals starting at age 60 and increasing afterward. Although the early retirement age (ERA) is 62, this strategy was consistent with that of previous research that accounted for cases where individuals retired before the ERA (e.g., at age 60) but delayed claiming benefits until they reached the ERA (Coile and Gruber 2007).⁴ While both Social Security benefits and tax rates influence retirement decisions (Fields and Mitchell 1984; Coile and Gruber 2007), we did not consider the effect of tax rates due to the unavailability of data. We do not believe this poses a limitation to our work because previous research has demonstrated that Social Security tax rates have only a small effect on the primary earner's labor supply decisions (Knapp 2014).

Next, we modeled the financial retirement incentive generated by the Social Security wealth accumulation as a *peak value* (PV) (Coile and Gruber 2001). As noted above, we computed the lifetime Social Security wealth for individuals based on their monthly benefits to be received until death with an immediate retirement (today at age t , in year s) as well as the expected lifetime Social Security wealth this individual could accrue by retiring at age $t+1$, $t+2$, $t+3$, and so on. The peak value measured the difference in the expected Social Security wealth if an individual retired at a future optimal age (and received the maximum expected value of Social Security wealth) rather than retiring immediately at age t , appropriately discounted, with

$$(2) \quad PV_t(r^*) = \sum_{s=r}^S \frac{pr_{s|t} E_t(B_s(r^*))}{(1+d)^{s-t}} - \sum_{s=t}^S \frac{pr_{s|t} E_t(B_s(t))}{(1+d)^{s-t}},$$

where $pr_{s|t}$ was the probability of living to age s conditional on being alive at age t , $B_s(r)$ was the Social Security benefits in year s if the individual started claiming in year r (r^* was the future optimal year that maximized the expected value of Social Security wealth), and d referred to the subjective discount factor. The first term in the equation referred to one's optimal lifetime Social Security wealth, calculated by summing the monthly benefits to be collected from the optimal initial claiming time of r^* until the time of death S , subject to discount rates and survival probabilities. The second term referred to one's actual lifetime Social Security wealth attainable by retiring at time t . In the end, the peak value measured the arithmetic difference between (1) the expected Social Security wealth if one retired at a future optimal age, and (2) the expected Social Security wealth if one were to retire at age t . Although the peak value is comparable to the option value (Ausink and Wise 1996; Lumsdaine et al. 1992)⁵ as a metric of retirement incentives, it requires fewer assumptions (Stock and Wise 1990). The peak value does not include salary income and hence provides more accurate estimates of the association between Social Security retirement incentives and the probability of retirement (Coile and Gruber 2007).

⁴ Following Coile and Gruber (2007), we assumed that, (1) individuals retiring before age 62 delayed claiming until age 62 but not later, and (2) individuals retiring at age 62 or later claimed Social Security benefits immediately.

⁵ In an option value model (Stock and Wise 1990), an employee would compare the expected present value of retiring at the current age with the value of retiring at each subsequent age. This process assumes workers' utility maximization tendencies whereby an employee weights the indirect utility of future income to that of future retirement benefits. The option value of postponing retirement is equal to the maximum of the present value of retiring at each future age minus the expected present value of immediate retirement. Previous studies have shown that the option value model approximates workers' retirement decisions as well as a more complex dynamic programming model (e.g. Ausink and Wise 1996; Lumsdaine et al. 1992).

Appendix C: Identifying Past Employment Gaps and Approximating Future Working Trajectories

The working trajectories file, publicly available as a contributed file from Christian Dudel for the HRS, contains respondents' household ID number, personal ID number, and labor force status by month and year (indicators for inactive, disabled, unemployed, retired, employed) (Dudel 2018). The file provided HRS respondents' *monthly self-reported work status* from January 1992 to April 2015, *with no further updates thereafter*.

For social security wealth calculation, we needed to impute the work trajectories of HRS respondents from May 2015 forward so that we could identify patterns of employment gaps for all respondents from age 20 to age 120. We used Dudel's methodology to obtain a complete file of HRS respondents' work trajectories between 1992 and 2020. We identified individuals' patterns of past labor force participation observed in our newly created Working Trajectories data (1992-2020) and applied the same patterns to every subsequent 29-year period to impute future work trajectories.

We employed the following methodology to construct the Working Trajectories data from 1992 to 2020. First, we imputed missing monthly labor force statuses for HRS survey years between January 1992 and April 2015. In addition to the original HRS labor force status variable categories for employed, unemployed, disabled, retired, and out of the labor force (inactive), we generated three additional categories for missing, dead, and censored. The status 'Missing' was used when there was no information on the status of a respondent whatsoever, the information in the HRS did not suffice to assign a status, or no interview date was recorded. The status 'Dead' was used when the respondent had passed away. The status 'Censored' was used when an individual participated in the last wave of data available and had not died or dropped out of the survey before this wave. We imputed missing monthly work status in two cases. One of these involved missing monthly work status in the transition between jobs for the respondent, for whom we imputed the missing months as 'unemployed.' The other involved respondents working multiple jobs and reporting more than one monthly job status, for whom we imputed labor force status as 'employed' as long as a respondent was working in any of the jobs. In the HRS surveys, self-reported changes in labor status included the month and year in which respondents became unemployed, disabled, or retired. We applied this information to identify the beginning of respondents' 'non-working' periods in the work trajectories. Lastly, using the exit interviews, we identified the exact month and year of a respondent's death (and exit from the survey).

Second, because the HRS is a biennial survey conducted only in even-numbered years, we had to identify the monthly work statuses of the respondents for the odd years between 1992 and 2015. Information from the (even) year $t-1$ was used to fill the work statuses in year t —but not thereafter. If a respondent reported being 'unemployed' in July 2012 (HRS survey year) and 'employed' in June 2014 (next HRS survey year), the status 'unemployed' was assigned to months from July 2012 to May 2014. In some cases, not all months between two HRS surveys could be assigned a status—leaving gaps in working trajectories. Respondents' monthly work statuses in these gaps were imputed using the following rules:

- **“LEFT”**: initial Status A was assigned for all missing months.
- **“RIGHT”**: subsequent Status B was assigned for all missing months.
- **“MID”**: A transition between A and B was assumed to occur in the *middle* of the interval of missing information. The first half of the missing months was assigned Status A, and the next half was assigned Status B.

- “OTHER1”: The missing months were assigned as ‘out of the labor force,’ a work status different from Status A or B.
- “OTHER2”: The missing months were assigned as ‘unemployed,’ a work status different from Status A or B.
- “OTHER3”: A transition between Status A and a different status was assumed to occur in the *middle* of the interval of missing information. The first half of the missing months were assigned Status A, and the next half was assigned ‘unemployed’.

Table A1 below summarizes the imputation rules. The initial observed work statuses (Status A) are given in the row, and the subsequent observed statuses (Status B) after the missing monthly statuses are given in the column.

Table A1. Imputation Rules for Missing Monthly Labor Force Statuses

		Subsequent (Status B)					
		<i>death</i>	<i>out of labor force</i>	<i>disabled</i>	<i>unemployed</i>	<i>retired</i>	<i>employed</i>
Initial (Status A)	<i>out of labor force</i>	left	left	left	left	left	other3
	<i>disabled</i>	left	mid	left	left	left	other3
	<i>unemployed</i>	left	mid	left	left	left	left
	<i>retired</i>	left	mid	left	left	left	left
	<i>employed</i>	left	right	other1	other1	other1	other2

Once we identified monthly labor force statuses from January 1992 to April 2015, we imputed the respondents’ work trajectories beyond April 2015. The monthly labor statuses were simplified into a binary variable, indicating whether one worked (1) or not (0). We first counted the number of months the respondents worked between January 1992 and April 2015, and updated the Working Trajectories data file from 2016 to 2020 following the structure of the 1992 to 2015 file. Then, we imputed a respondent’s future monthly work status beyond 2020 in a way that the same number of working months observed in 1992-2020 was assigned (randomly) to every 29-year timeframe for years beyond 2020. We imputed the work trajectories for everyone until the maximum age of 120 (T). For example, if a respondent worked for 240 months between 1992 and 2020, we imputed his future working trajectory by randomly assigning 240 working months for every 29-year period beyond 2020. Based on these imputations, we calculated the *number of working months per year for each year beyond 2020 until individuals reached age 120*. We used the imputed future working months, which closely reflected the observed work trajectories of the respondents, to calculate the *annual projected earnings of the respondents*. The annual salary is one of the central components of the social security wealth calculation.

We did not consider individuals’ mortality in the imputation of future earnings or working trajectories. However, individuals’ gender and race/ethnicity-specific year-to-year survival probabilities are reflected in the Social Security wealth calculation. Moreover, respondents’ earnings prior to 1992 (i.e., the year of the first HRS survey) were based on SSA records. These records already reflected any past

truncations in respondents' work histories. For instance, if an individual worked for only five months in one year, the total income reported for the year was a sum of the earnings for those five months. Because the work histories are accurately reflected in the past earnings for respondents with HRS-SSA linkages, and because the past earnings of respondents with missing HRS-SSA linkages were imputed based on HRS-SSA linked records of other respondents, we did not apply our imputation rules to respondents' past work trajectories before 1992.

Appendix D: Descriptive Statistics

Table A2 summarizes characteristics for 33,992 White, 9,453 Black, and 7,400 Hispanic person-wave observations. It reveals several differences across these groups. Whites had the highest average level of Social Security wealth, followed by Blacks and Hispanics. The peak value incentive was greatest for Hispanics (\$3,090.68), followed by Whites (\$3,087.61) and Blacks (\$1,810.36). Put another way, Hispanics could increase their Social Security wealth by nearly \$3,000 by delaying retirement until a future optimal time, while Blacks would increase their Social Security wealth approximately \$1,800 by doing so. Based on their Social Security wealth accumulation trajectories, Blacks had the lowest incentives to delay retirement.

As can be seen from the distribution of birth years shown in Figure A3, Hispanics and Blacks were younger than Whites. This helps explain why Blacks (14.70%) and Hispanics (13.96%) were less likely than Whites (15.56%) to retire in the next year. Nearly half of White and Hispanic respondents were male. Hispanics had fewer years of education than the respondents in the other two groups. Blacks were less likely than others to be married or partnered. Hispanics had higher average household sizes. White respondents were most likely to be U.S. veterans (> 95.00%). More than 95 percent of Whites and more than 90 percent of Blacks were born in the United States, but only about 34 percent of Hispanic respondents were.

Whites reported working more total years worked than did Blacks or Hispanics, suggesting they had the longest labor force engagement (see Table A2). Figure A4 shows the percentage of each group that worked 12 months, 6 to 11 months, one to five months, or zero months on an annual basis. All else equal, Blacks and Hispanics worked fewer months per year than Whites did. This, in turn, would lead to *lower annual earned incomes* and *lower* accumulation of Social Security benefits for them.

As observed in Table A2, all three groups differed in financial status and by health conditions. White and Black respondents earned significantly more than Hispanics did, and the spouses of White and Black respondents earned more than the spouses of Hispanic respondents did. Real net wealth of Whites was 3.6 times that of Blacks and 3.3 times that of Hispanics. About 70 percent of Whites and Blacks have contributed to employer-provided pension but fewer than 50 percent of Hispanic respondents have done so. Blacks were more likely than Hispanics or Whites to report chronic health conditions.⁶ Hispanics had the highest average CES-D scores, indicating more depressive symptoms.

⁶ Lower prevalence of chronic conditions for Hispanics may be due to underdiagnosis rather than better health (Kim et al. 2018).

Table A2. Summary Statistics

	Non-Hispanic White (W)	Non-Hispanic Black (B)	Hispanic (H)	Significance of differences		
	% or Mean (SD)	% or Mean (SD)	% or Mean (SD)	W-B	B-H	W-H
<i>Outcome variable</i>						
Retire in t+1, conditional on working in t	15.56	14.70	13.96	**	-	***
<i>Explanatory variables</i>						
Social Security Wealth (in \$10,000)	31,786.20 (48,231.52)	19,427.33 (31,389.03)	19,118.15 (33,077.98)	***	-	***
Peak Value (in \$10,000)	3,087.61 (8,703.16)	1,810.36 (6,175.55)	3,090.69 (7,859.45)	***	***	-
<i>Covariates</i>						
Age	59.64 (6.17)	58.71 (5.57)	58.29 (5.22)	***	***	***
Age difference with spouses (1=yes, 0=no)						
Respondent is older	32.87	22.62	33.53	***	***	-
Same age as spouse	7.74	3.96	6.58	***	***	***
Respondent is younger	31.35	21.81	28.20	***	***	***
Missing	28.04	51.61	31.69	***	***	***
Gender (1=male, 0=female)	48.28	40.18	50.72	***	***	***
Years of education	13.91 (2.36)	13.15 (2.55)	10.51 (68.36)	***	***	***
Couple (1=yes, 0=no)	71.98	48.39	68.36	***	***	***
No. of household members	2.35 (1.09)	2.58 (1.44)	3.24 (1.74)	***	***	***
U.S.-born (1=yes, 0=no)	95.87	90.44	34.26	***	***	***
Veteran status (1=yes, 0=no)	17.74	14.43	7.69	***	***	***
Years of tenure at current job (1=yes, 0=no)						
less than 1 year	30.80	33.77	38.59	***	***	***
1 to < 5 years	17.18	17.45	17.23	***	***	-
5 to < 10 years	50.14	46.05	40.97	***	***	***
10+ years	1.88	2.73	3.20	***	*	***
Total years worked	41.66	35.25	30.61	***	***	***

	(12.38)	(13.98)	(14.91)			
Real annual salary income (\$)	50,526.38 (82,730.14)	33,186.14 (39,629.64)	25,605.08 (29,023.96)	***	***	***
Real net wealth (\$)	687,627.80 (1,843,860.81)	189,359.50 (615,171.11)	207,690.80 (524,200.20)	***	**	***
Real annual salary income of spouse (\$)	32,616.12 (67,223.42)	19,388.56 (41,291.77)	16,880.63 (34,951.64)	***	***	***
Real annual out-of-pocket exp (\$)	3,422.96 (7,810.73)	2,417.89 (5,114.53)	2,470.79 (7,054.09)	***	***	***
Contributed to employer-provided pension (1=yes, 0=no)						
Defined benefit plans (DB)	27.68	28.02	17.45	***	***	***
Defined contribution plans (DC)	40.25	35.63	24.34	***	***	***
Both	2.79	4.13	2.74	***	***	-
None	27.72	30.35	53.24	***	***	***
Chronic Conditions (1=one or more, 0=none)	55.98	69.95	56.57	***	***	-
CES-D score (0-8)	0.99 (1.54)	1.34 (1.58)	1.59 (1.95)	***	***	***
Health insurance, employer-sponsored (1=yes, 0=no)	34.57	28.55	27.07	***	**	***
Medicare coverage (1=yes, 0=no)	17.68	13.45	9.84	***	***	***
No. observations	33,992	9,453	7,400			

Notes: Standard deviation (SD) in parenthesis.

Source: author's calculations using HRS 2000-2020.

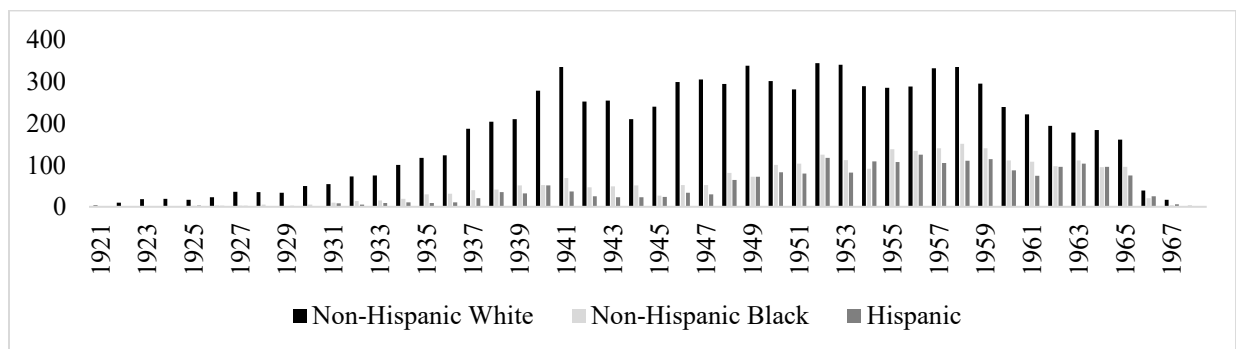


Figure A3. Number of HRS Respondents by Year of Birth and Race/Ethnicity

Source: author's calculations.

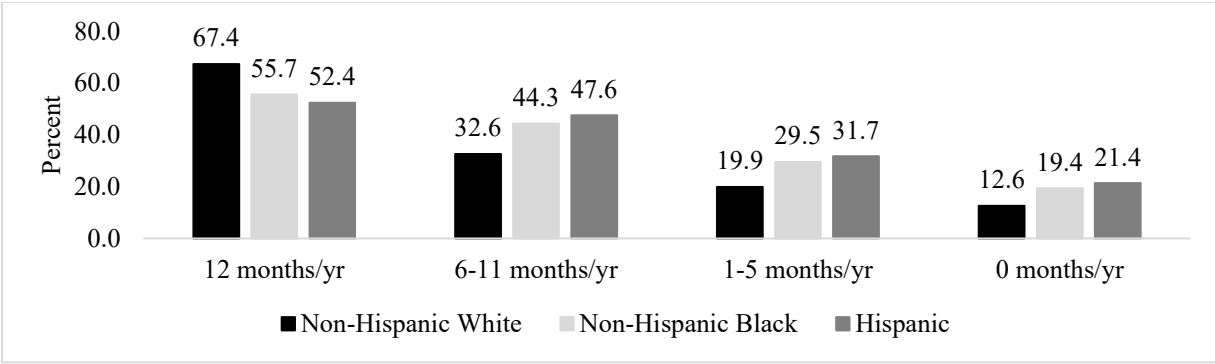


Figure A4. Percent of Respondents Working 12 Months, 6-11 Months, 1-5 Months, and 0 Months per Year by Race/Ethnicity

Source: author's calculations.

Appendix E: Regression Results Including All Covariates

Tables A3 and A4 show the marginal effects of the probit regressions, estimated separately for Whites, Blacks, and Hispanics. For all models, the outcome variable was a binary indicator for retirement, which took a value of 1 if working in wave t and retiring in $t + 1$, and 0 if working in t and $t + 1$. We include covariates sequentially in Specifications I-IV. Table A3 shows the effects of Social Security wealth and the peak value incentive on retirement decisions absent any other controls (columns 1-3) and when controlling for demographic and work variables as well as for age fixed-effects and year fixed-effects (columns 4-6). Table A4 adds further controls, specifically those for financial status (columns 1-3) and those for health and insurance status (columns 4-6).

Table A3. Retirement Likelihood: Marginal Effects

	(1)	(2)	(3)	(4)	(5)	(6)
	Non-Hispanic White	Non-Hispanic Black	Hispanic	Non-Hispanic White	Non-Hispanic Black	Hispanic
Covariate Specifications:	Spec. I: No controls			Spec. II: Demographics and work		
Social Security Wealth (in \$10,000)	0.014*** (0.001)	0.020*** (0.001)	0.016*** (0.002)	0.002*** (0.001)	0.003 (0.002)	0.003 (0.002)
Peak Value (in \$10,000)	-0.020*** (0.003)	-0.024*** (0.006)	-0.017*** (0.006)	-0.010*** (0.004)	-0.010 (0.007)	-0.027*** (0.008)
Age difference with spouses (1=yes, 0=no)						
Same age as spouse	-	-	-	0.008 (0.008)	0.040* (0.023)	-0.011 (0.017)
Respondent is younger	-	-	-	0.010* (0.005)	0.016 (0.011)	0.015 (0.011)
Missing	-	-	-	0.220* (0.114)	0.001 (0.000)	-0.367*** (0.004)
Gender (1=male, 0=female)	-	-	-	-0.021***	-0.004	-0.056***

	-	-	-	(0.005)	(0.008)	(0.009)
Years of education	-	-	-	-0.005***	-0.002	-0.001
	-	-	-	(0.001)	(0.001)	(0.001)
Couple (1=yes, 0=no)	-	-	-	0.172***	0.001	-0.631***
	-	-	-	(0.060)	(0.000)	(0.004)
No. of household members	-	-	-	-0.009***	0.005*	0.004
	-	-	-	(0.002)	(0.003)	(0.002)
U.S.-born (1=yes, 0=no)	-	-	-	0.012	0.061***	0.031***
	-	-	-	(0.009)	(0.010)	(0.009)
Veteran status (1=yes, 0=no)	-	-	-	0.020***	-0.010	-0.006
	-	-	-	(0.006)	(0.011)	(0.016)
Years of tenure at current job (1=yes, 0=no)						
Less than 1 year	-	-	-	0.018***	0.046***	0.022*
	-	-	-	(0.006)	(0.011)	(0.011)
5 to < 10 years	-	-	-	0.015***	0.030***	0.003
	-	-	-	(0.005)	(0.010)	(0.011)
10+ years	-	-	-	-0.132***	-0.110***	-0.096***
	-	-	-	(0.008)	(0.013)	(0.015)
Total years worked	-	-	-	-0.003***	-0.001***	-0.002***
	-	-	-	(0.001)	(0.001)	(0.001)
Age fixed-effects	no	no	no	yes	yes	yes
Year fixed-effects	no	no	no	yes	yes	yes
Dependent Variable Mean	0.156	0.147	0.140	0.156	0.147	0.140
No. Observations	33,992	9,453	7,400	33,992	9,453	7,400

Notes: Social Security wealth, peak value, and real net wealth are in U.S. dollars. Reference categories for spouse age-differences is “the respondent is older than the spouse”, and for years of tenure at current job is “1 to < 5 years”. Standard errors (SE) in parentheses. * p<0.1 ** p<0.05 *** p<0.01.

Source: author’s calculations using HRS 2000-2020.

Table A4. Retirement Likelihood: Marginal Effects (Cont.)

	(1)	(2)	(3)	(4)	(5)	(6)
	Non-Hispanic White	Non-Hispanic Black	Hispanic	Non-Hispanic White	Non-Hispanic Black	Hispanic
	Spec. III. Demographics, work, and financial conditions			Spec. IV. Demographics, work, financial conditions, health and health insurance		
Social Security Wealth (in \$10,000)	0.002***	0.004**	0.003	0.002***	0.005***	0.003
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
Peak Value (in \$10,000)	-0.009**	-0.009	-0.026***	-0.007**	-0.007	-0.024***

	(0.004)	(0.007)	(0.008)	(0.004)	(0.007)	(0.008)
Age difference with spouses (1=yes, 0=no)						
Same age as spouse	0.008 (0.008)	0.032 (0.022)	-0.014 (0.017)	0.009 (0.008)	0.032 (0.022)	-0.017 (0.017)
Respondent is younger	0.007 (0.005)	0.013 (0.015)	0.013 (0.011)	0.008 (0.005)	0.014 (0.012)	0.011 (0.011)
Missing	0.222** (0.113)	- -	-0.368*** (0.004)	0.221** (0.107)	- -	-0.367*** (0.004)
Gender (1=male, 0=female)	-0.022*** (0.005)	-0.004 (0.008)	-0.056*** (0.009)	-0.022*** (0.005)	-0.003 (0.008)	-0.053*** (0.009)
Years of education	-0.005*** (0.001)	-0.001 (0.001)	-0.001 (0.001)	-0.004*** (0.001)	0.001 (0.001)	-0.001 (0.001)
Couple (1=yes, 0=no)	0.174*** (0.059)	0.001 (0.000)	-0.631*** (0.004)	0.174*** (0.057)	0.001 (0.000)	-0.632*** (0.004)
No. of household members	-0.008*** (0.002)	0.005* (0.003)	0.004 (0.002)	-0.009*** (0.002)	0.005* (0.003)	0.004 (0.002)
US-born (1=yes, 0=no)	0.013 (0.009)	0.063*** (0.010)	0.035*** (0.010)	0.011 (0.009)	0.059*** (0.010)	0.031*** (0.009)
Veteran status (1=yes, 0=no)	0.019*** (0.006)	-0.008 (0.011)	-0.005 (0.016)	0.018** (0.008)	-0.009 (0.011)	-0.007 (0.015)
Years of tenure at current job (1=yes, 0=no)						
Less than 1 year	0.016*** (0.006)	0.039*** (0.010)	0.019* (0.011)	0.014** (0.007)	0.036*** (0.010)	0.017 (0.011)
5 to < 10 years	0.012** (0.006)	0.034*** (0.010)	0.004 (0.011)	0.012** (0.005)	0.034*** (0.010)	0.007 (0.011)
10+ years	-0.132*** (0.007)	-0.105*** (0.014)	-0.090*** (0.016)	-0.131*** (0.007)	-0.106*** (0.014)	-0.090*** (0.016)
Total years worked	-0.003*** (0.000)	-0.001*** (0.000)	-0.002*** (0.000)	-0.003*** (0.000)	-0.001*** (0.000)	-0.002*** (0.000)
Real annual income (ln)	0.002*** (0.000)	-0.001 (0.000)	0.001 (0.000)	0.002*** (0.000)	-0.001 (0.000)	0.001 (0.000)
Real net wealth (in \$100,000)	0.001 (0.000)	0.001 (0.001)	0.001* (0.001)	0.001 (0.000)	0.001 (0.000)	0.001* (0.001)
Spouse real annual income (ln)	-0.001 (0.000)	-0.001** (0.000)	-0.001 (0.000)	-0.001* (0.000)	-0.001** (0.000)	-0.001 (0.000)
Annual out-of-pocket exp (ln)	0.001	0.001	0.001	-0.001	0.001	0.001

	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
Contributed to employer-provided pension (1=yes, 0=no)						
DC plans	-0.042*** (0.005)	-0.015* (0.009)	-0.008 (0.012)	-0.040*** (0.004)	-0.014* (0.009)	-0.008 (0.012)
Both DB and DC	-0.021** (0.011)	-0.030* (0.017)	0.015 (0.028)	-0.022* (0.012)	-0.030* (0.017)	0.015 (0.028)
None	0.007 (0.005)	0.046*** (0.010)	0.026** (0.012)	-0.001 (0.005)	0.040*** (0.010)	0.014 (0.012)
Chronic Conditions (1=one or more, 0=none)	-	-	-	0.021*** (0.004)	0.012 (0.008)	0.024*** (0.008)
CES-D score (0-8)	-	-	-	0.006*** (0.001)	0.006*** (0.002)	0.004** (0.002)
Health insurance, employer-sponsored (1=yes, 0=no)	-	-	-	-0.042*** (0.010)	-0.031*** (0.007)	-0.036*** (0.008)
Medicare coverage (1=yes, 0=no)	-	-	-	-0.001 (0.010)	0.023 (0.018)	0.039* (0.021)
Age fixed-effects	yes	yes	yes	yes	yes	yes
Year fixed-effects	yes	yes	yes	yes	yes	yes
Dependent Variable Mean	0.156	0.147	0.140	0.156	0.147	0.140
No. observations	33,992	9,453	7,400	33,992	9,453	7,400

Notes: Social Security wealth, peak value, and real net wealth are in U.S. dollars. Reference categories for spouse age-differences is “the respondent is older than the spouse”, for years of tenure at current job is “1 to < 5 years,” and for contributed to employer-provided pension is “only to DB pension plans.” DC refers to Defined Contribution and DB to Defined Benefit. Standard errors (SE) in parentheses.

* p<0.1 ** p<0.05 *** p<0.01.

Source: author’s calculations using HRS 2000-2020.

References

- Aguila, Emma, Zeewan Lee, and Rebeca Wong. 2023. “Migration, Work, and Retirement: The Case of Mexican-Origin Populations.” *Journal of Pension Economics & Finance* 22 (2): 167–87. <https://doi.org/10.1017/s1474747221000342>.
- Arias, Elizabeth, Melonie Heron, and Jiaquan Xu. 2016. *United States Life Tables, 2012*. 8. https://www.cdc.gov/nchs/data/nvsr/nvsr65/nvsr65_08.pdf.
- Ausink, John A., and David A. Wise. 1996. “The Military Pension, Compensation and Retirement of U.S. Air Force Pilots.” In *Advances in the Economics of Aging*, edited by David A. Wise. National Bureau of Economic Research. <http://www.nber.org/chapters/c7319>.

- Börsch-Supan, Axel. 1992. "Population Aging, Social Security Design, and Early Retirement." *Journal of Institutional and Theoretical Economics (JITE) / Zeitschrift Für Die Gesamte Staatswissenschaft* 148 (4): 533–57.
- Case, Anne, and Angus Deaton. 2015. "Rising Morbidity and Mortality in Midlife among White Non-Hispanic Americans in the 21st Century." *Proceedings of the National Academy of Sciences of the United States of America* 112 (49): 15078–83. <https://doi.org/10.1073/pnas.1518393112>.
- Coile, Courtney, and Jonathan Gruber. 2000. *Social Security and Retirement*. Working Paper W7830. With D. A. Wise. National Bureau of Economic Research. <https://www.nber.org/system/files/chapters/c10332/c10332.pdf>.
- Coile, Courtney, and Jonathan Gruber. 2001. "Social Security Incentives for Retirement." In *Themes in the Economics of Aging*, Wise David A., edited by D. A. Wise. National Bureau of Economic Research. <https://www.nber.org/system/files/chapters/c10332/c10332.pdf>.
- Coile, Courtney, and Jonathan Gruber. 2007. "Future Social Security Entitlements and the Retirement Decision." *The Review of Economics and Statistics* 89 (2): 234–46. <https://doi.org/10.1162/rest.89.2.234>.
- Dudel, Christian. 2018. "Working Trajectories in the HRS." Distributed by the University of Michigan. <https://hrsdata.isr.umich.edu/data-products/working-trajectories-hrs>.
- Fields, Gary S., and Olivia S. Mitchell. 1984. "The Effects of Social Security Reforms on Retirement Ages and Retirement Incomes." *Journal of Public Economics* 25 (1): 143–59. [https://doi.org/10.1016/0047-2727\(84\)90048-3](https://doi.org/10.1016/0047-2727(84)90048-3).
- Johnson, David R., and Rebekah Young. 2011. "Towards Best Practices in Analyzing Datasets with Missing Data: Comparisons and Recommendations." *Journal of Marriage and Family* 73 (5): 926–45. <https://doi.org/10.2307/41329640>.
- Kim, Eun Ji, Taekyu Kim, Joseph Conigliaro, Jane M. Liebschutz, Michael K. Paasche-Orlow, and Amresh D. Hanchate. 2018. "Racial and Ethnic Disparities in Diagnosis of Chronic Medical Conditions in the USA." *Journal of General Internal Medicine* 33 (7): 1116–23. <https://doi.org/10.1007/s11606-018-4471-1>.
- Knapp, David. 2014. *The Effect of Social Security Auxiliary Spouse and Survivor's Benefits on the Household Retirement Decision*. Working paper. University of Michigan, the 16th Annual Joint Meeting of the Retirement Research Consortium. http://crr.bc.edu/wp-content/uploads/2014/06/Panel-3_3-Knapp.pdf.
- Lokupitiya, Ravindra, Erandathie Lokupitiya, and Keith Paustian. 2006. "Comparison of Missing Value Imputation Methods for Crop Yield Data." *Environmetrics* 17 (4): 339–49. <https://doi.org/10.1002/env.773>.
- Lumsdaine, Robin L., James H. Stock, and David A. Wise. 1992. "Three Models of Retirement: Computational Complexity versus Predictive Validity." In *Topics in the Economics of Aging*, edited by David A. Wise. The Economics of Aging. National Bureau of Economic Research. <http://www.nber.org/chapters/c7097>.

- National Center for Health Statistics. 2025. “National Vital Statistic System - Life Expectancy.” <https://www.cdc.gov/nchs/nvss/life-expectancy.htm>.
- Ni, Daiheng, John D. Leonard II, Angshuman Guin, and Chunxia Feng. 2005. “Multiple Imputation Scheme for Overcoming the Missing Values and Variability Issues in ITS Data.” *Journal of Transportation Engineering* 131 (12): 931–38. [https://doi.org/10.1061/\(ASCE\)0733-947X\(2005\)131:12\(931\)](https://doi.org/10.1061/(ASCE)0733-947X(2005)131:12(931)).
- Stock, James H., and David A. Wise. 1990. “The Pension Inducement to Retire: An Option Value Analysis.” *Issues in the Economics of Aging*, 205–30. <https://www.nber.org/system/files/chapters/c7118/c7118.pdf>.
- U.S. Social Security Administration. 2019. “National Average Wage Index.” <https://www.ssa.gov/OACT/COLA/AWI.html>.
- U.S. Social Security Administration. n.d. “Benefits Planner | Social Security Tax Limits on Your Earnings | SSA.” Table values. Maximum Taxable Earnings. <https://www.ssa.gov/benefits/retirement/planner/maxtax.html>.
- U.S. Social Security Administration. n.d. “Benefit Formula Bend Points.” <https://www.ssa.gov/OACT/COLA/bendpoints.html>.
- White, Ian R., Patrick Royston, and Angela M. Wood. 2011. “Multiple Imputation Using Chained Equations: Issues and Guidance for Practice.” *Statistics in Medicine* 30 (4): 377–99. <https://doi.org/10.1002/sim.4067>.
- White, Kirk, Jerome P. Reiter, and Amil Petrin. 2018. “Imputation in US Manufacturing Data and Its Implications for Productivity Dispersion.” *The Review of Economics and Statistics* 100 (3): 502–9. https://doi.org/10.1162/REST_a_00678.