ONLINE APPENDIX

for

Social Security, Labor Supply and Health of Older Workers: Quasi-Experimental Evidence from a Large Reform

Itay Saporta-Eksten (Tel Aviv University)
Ity Shurtz (Ben-Gurion University of the Negev)
Sarit Weisburd (Tel Aviv University)

May 2020

Appendix 1: Creating the Database

Our database was created by combining the following administrative datasets for women in Israel born between 1929 and 1932 and their husbands. The sample includes women who were married in 1996, and their husbands as of 1996, excluding new immigrants (who immigrated to Israel after 1989). In practice, when we condition on employment in 1994 (our main sample), we further constrain the sample to individuals that are not kibbutz members or self-employed, due to availability of their employment data. All datasets below are administrative NII data that are linked via an individual identifier.

Marital History (Up to 2015)

This dataset follows the marriages and divorces of each individual in our sample. It provides information on dates where changes in marital status were made for each individual in the data. This allows us to create a 1-1 match between each wife and her spouse which can then be linked to all other administrative datasets.

Demographics (Up to 2015)

This dataset provides information by id (for both women and their spouses) on country of birth, ethnicity, year and month of birth, year and month of immigration, locality of residence, and year and month of death.[1] The latter variables are used to create our outcome variables for the age-by-age longevity analysis. Additionally, we use these data to construct age of immigration and age difference between spouses. This database also allows us to determine wife's age at first marriage when combined with the marital history data.

We use the following characteristics as household level controls in our analysis of husband employment: religion, immigrant status of both spouses, and age difference between spouses.

We use the following characteristics in the machine learning process to determine housewife status: wife's ethnicity, wife's age at first marriage, age difference between spouses, country of birth of both spouses, age of immigration for both spouses, and locality of residence (See Appendix 2).

Employment and Earnings Data (1984-2015)

This dataset provides annual data for each individual in our sample (wives and husbands) on: earnings, number of months worked and industry employed for each year between 1984 and 2015. We calculate monthly earnings as total earnings divided by months worked and convert to 1996 values using the Consumer Price Index (CPI) provided by the Bank of Israel. Our data include all employee salaries, excluding self-employed and kibbutz members. Because this is administrative data, we treat a missing year of data as evidence that an individual did not work as an employee that year. We merge this information with annual data on both minimum wage and the earnings test.

[1] When month of birth is missing, it is coded as April (see Appendix Figure A5).

We use these data to construct a few key variables in our analysis: (1) year of retirement and cumulative employment post-Reform (see Section 4); (2) pre-Reform earnings (which are used as controls in some of the regressions (see for example Table 5)) ; (3) number of years a wife worked in the 10 years leading up to the Reform (1996) (a key factor in determining housewife status for women for whom we lack the administrative flag).

Benefit Records (2003-2015)

These data provide monthly information on public pension benefits received each year between 2003 and 2015 for each individual in our data (wives and husbands). These data also provide information on whether or not the wife is flagged as a housewife.

Private Pension Records (2001-2015)

These data provide information on the number of months over which the individual withdrew private pension funds and the total amount of funds withdrawn for each year between 2001 and 2015 for each individual in our data (wives and husbands).

Births (Up to 2015)

This dataset can be linked by wife-id and provides information on the date of each birth. This database is used to determine both number of children and age at first birth which are both factors applied in the machine learning process to determine housewife status (See Appendix 2).

Appendix 2: The Machine Learning Process

As discussed in Section 3 of the paper, we use a machine learning approach to recover housewife classification for about 18% of women in our sample who lack sufficient earning history or a social security indicator to determine housewife status.

To assign a housewife flag for this sample, we train a random forest algorithm using the sample of women for which we observe true housewife status from the social security housewife indicator.[2] After an exploratory analysis, removing variables that had no predictive power, we conduct the training using the following set of variables: wife's birth cohort; wife's sum of years worked in the last 10 years before the reform year (1996); wife's number of months worked at age 55; wife's income in 1992; wife's ethnicity; number of children; wife's marriage age; wife's age at first birth; age difference between spouses; country of birth of both spouses; age of immigration for both spouses; locality of residence.

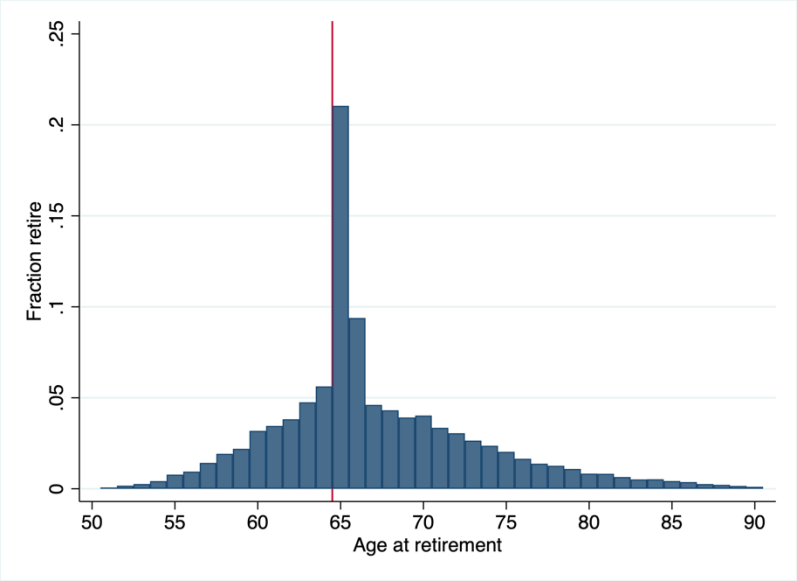
Given that our housewives group is the treated group that makes-up roughly 26% of the population, misclassifying non-housewives as housewives is costlier than the opposite. We therefore assign a housewife indicator to be 1 only if 70% of the trees predict housewife. The implied error rate in our prediction is 26% with only 3% of non-housewives incorrectly identified as housewives. Recall that this algorithm is only applied to 18% of the sample, which implies that we correctly classify housewives in over 95% of the cases within our full sample.

[2] We use the "ranger" package in R, growing 1,000 trees.

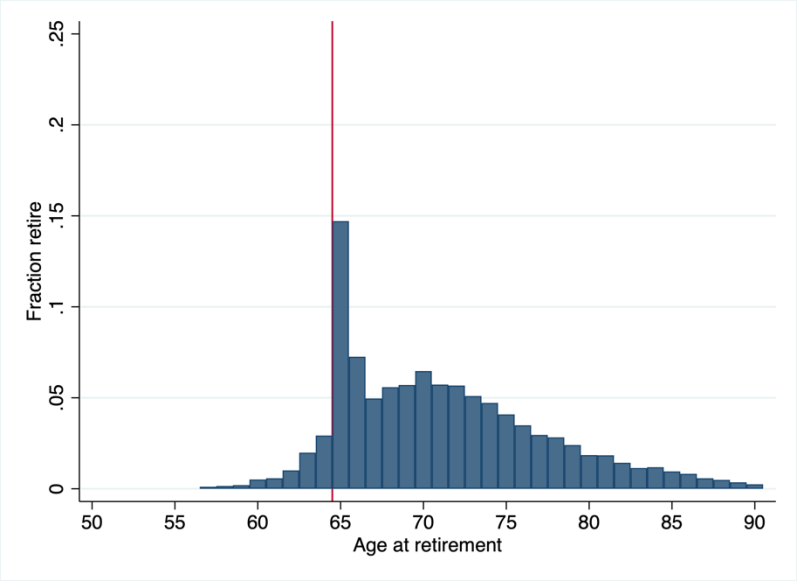
Appendix Figures

Figure A1. The Distribution of Retirement age

A. Population

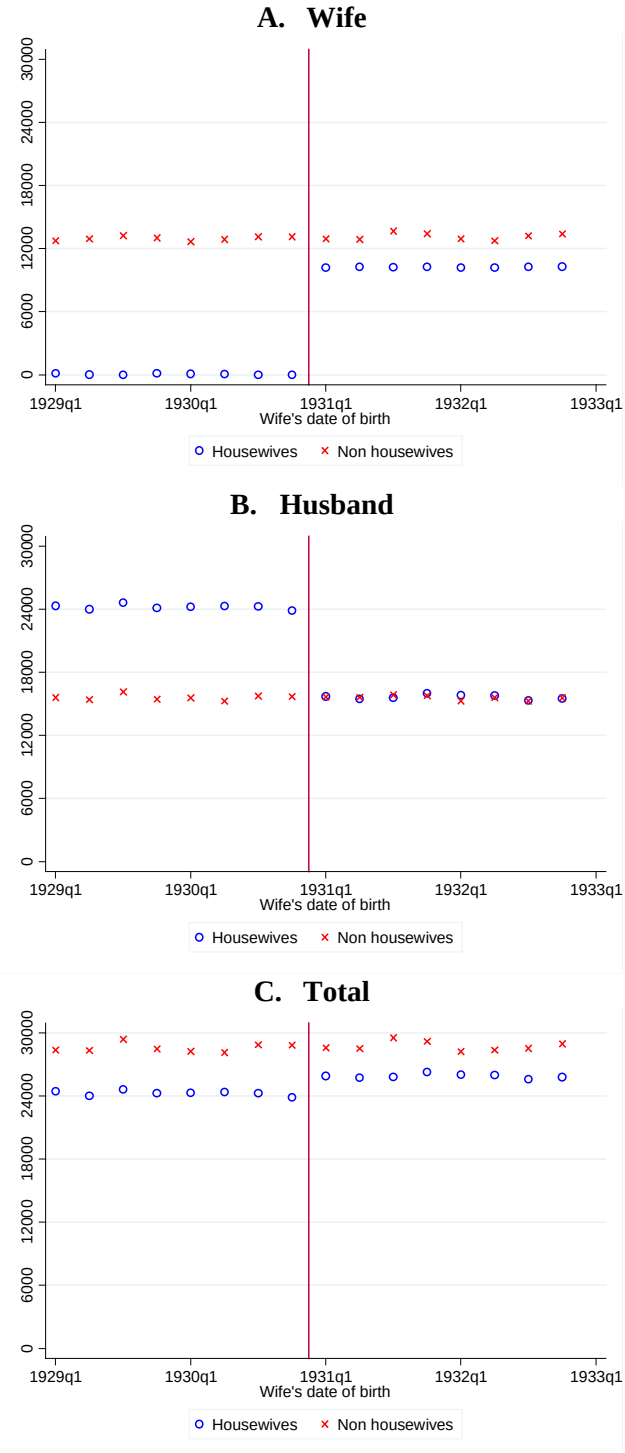


B. Conditioning on work in 1994



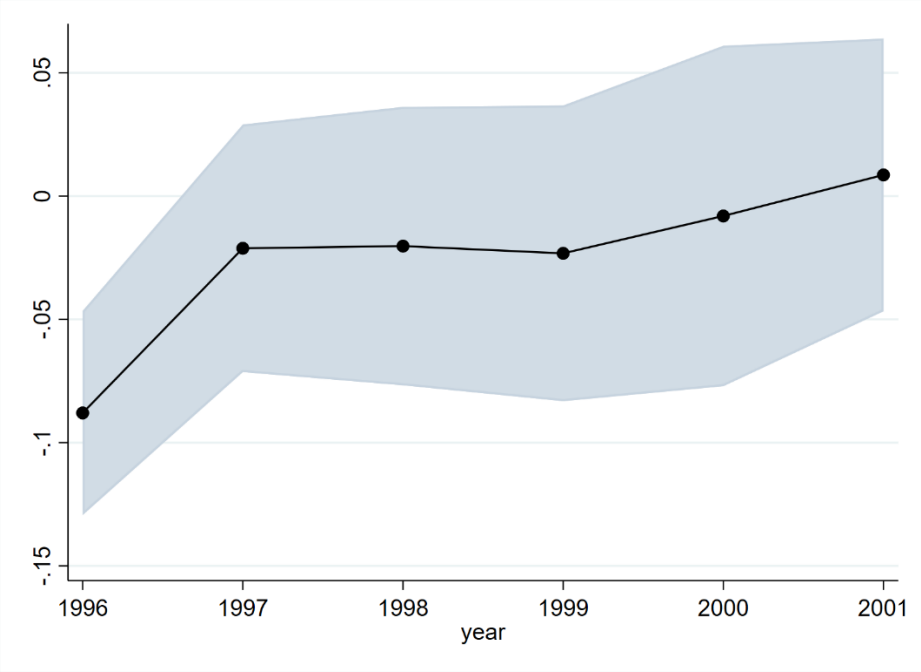
Note. This figure displays the distribution of husbands' retirement ages. Panel A displays the distribution for the entire population of husbands married to wives born between the years 1930-1932, while panel B displays the distribution conditioning on work in 1994. Note that retirement ages with less than 10 observations were suppressed due to NII disclosure policies.

Figure A2. Change in benefits around the date of birth cutoff



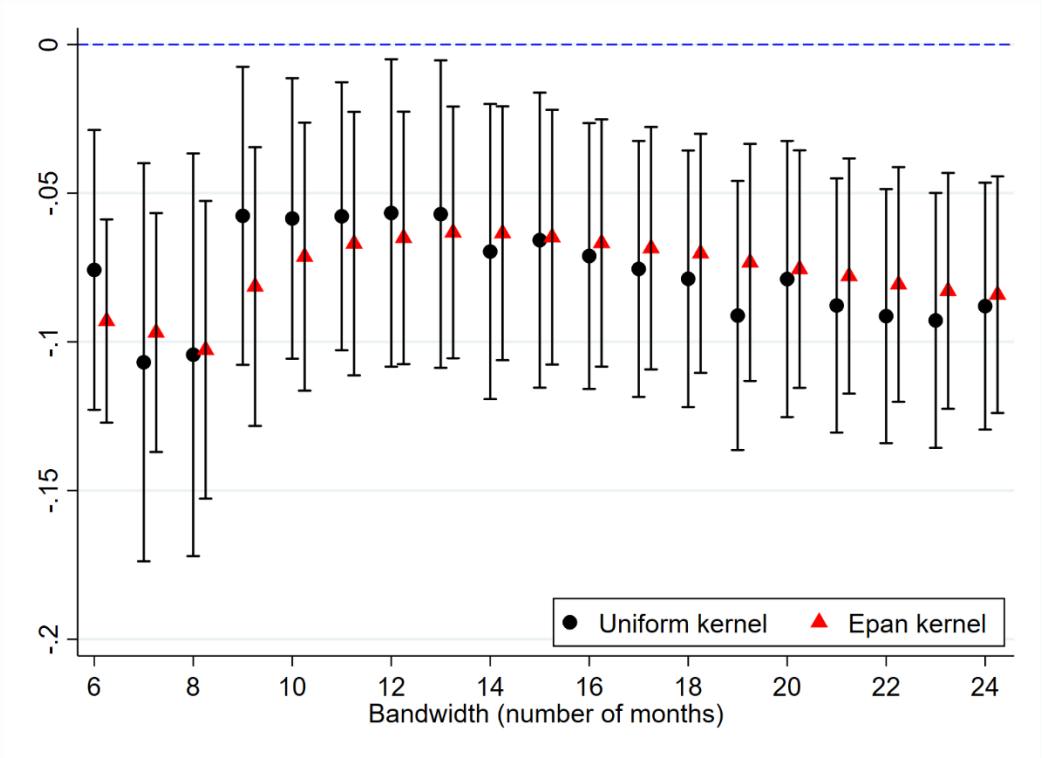
Note. Calculated for the sample of married couples with wives born 1929 to 1932, conditioning on husband's employment in 1994 and survival until 2007. Each symbol represents average retirement benefits in 2003-2007 by quarter of birth of the wife. All amounts are in NIS and deflated to 1996.

Figure A3. RDD Estimates for the Effect on Retirement Year-by-Year



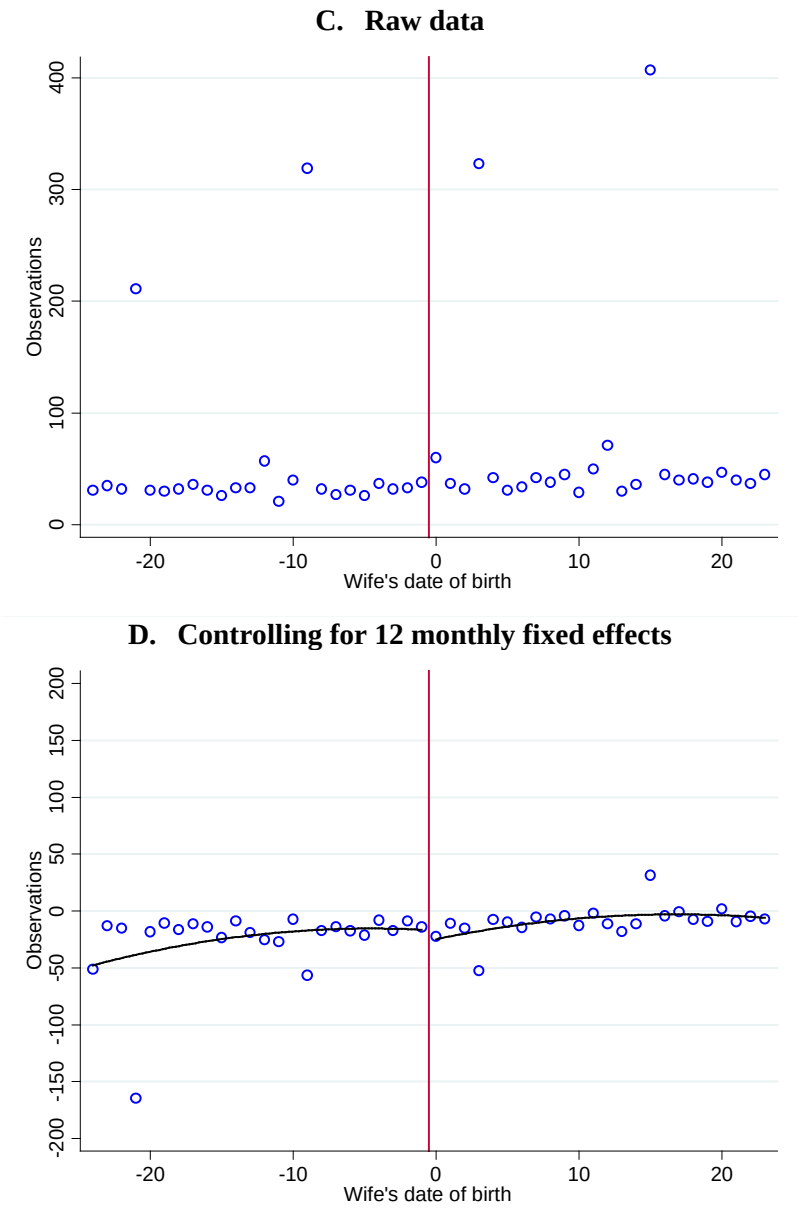
Note: This figure summarizes the results of the RDD analyses for retirement in the years 1996-2001. The points report the RDD coefficient (β in Equation (1)) and the shaded area is the 90% confidence intervals of the estimates. The specification corresponds to column (1) of Table 6.

Figure A4. RDD Estimates for the Effect on Retirement by 1996 (Alternative Bandwidths and Kernels)



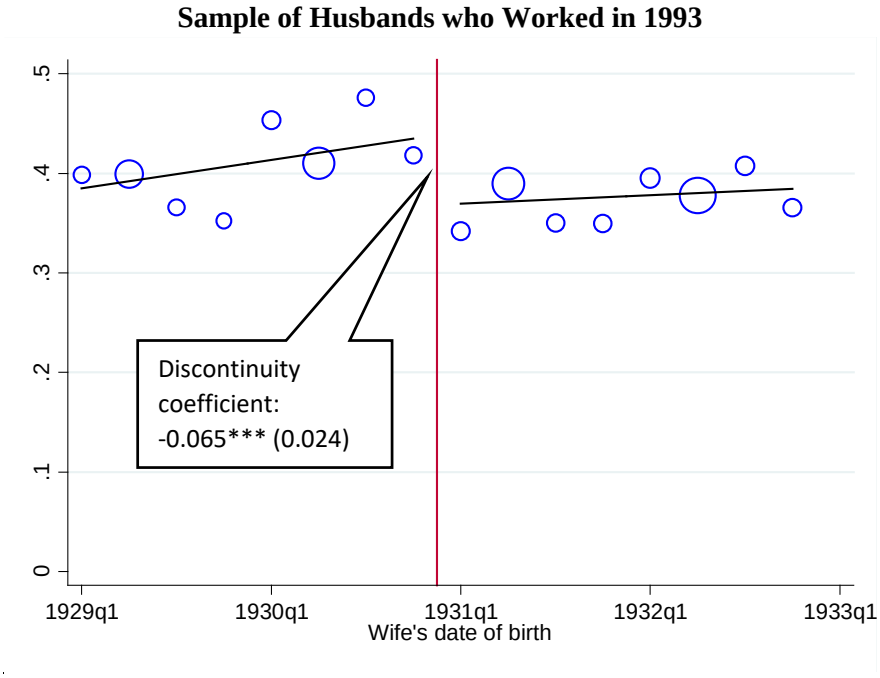
Note: This figure summarizes the results of the RDD analyses for retirement corresponding to column 1 of Table 6, for all bandwidths 6-24 months, and for Uniform and Epan kernels. The spikes correspond to the 90% confidence intervals of the estimates.

Figure A5. Density of observations around the RDD cutoff



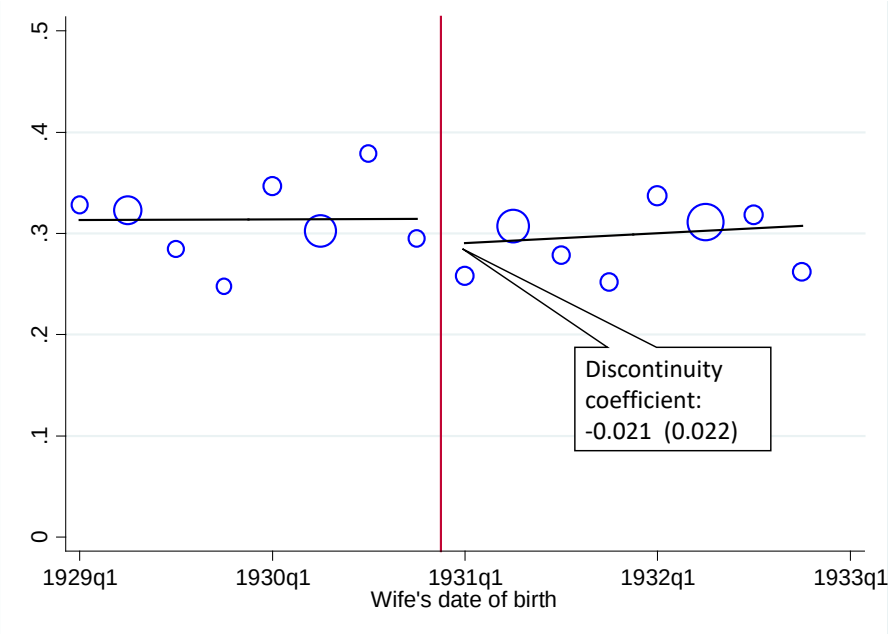
Note. This figure displays the count of observations by monthly bins. Panel A displays the raw data, while panel B displays the residuals from a regression that controls for 12 month fixed effects. Sample is husbands married to housewives born 1929 to 1932.

Figure A6. RDD, Retirement Rate by 1996, by Wife's Birth Quarter



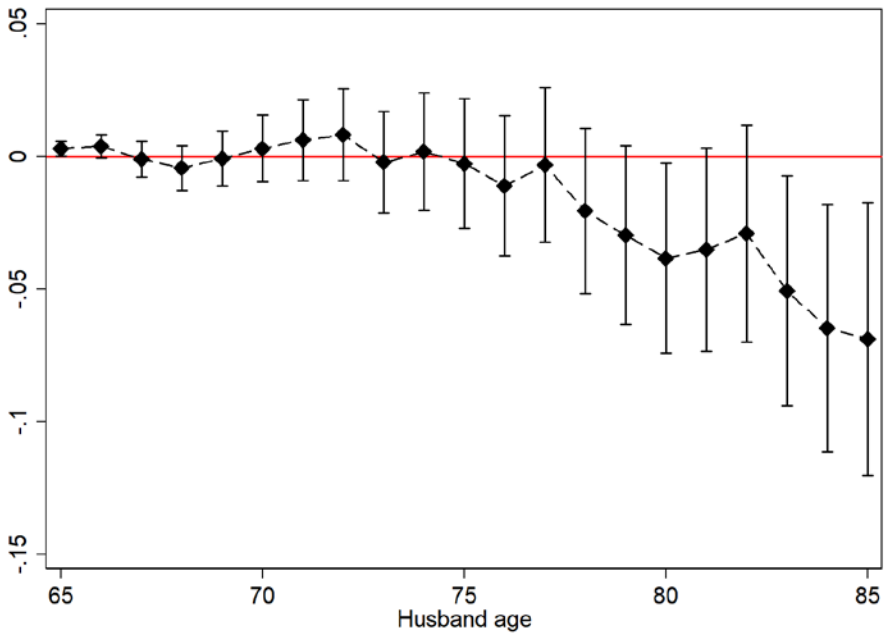
Note: This figure shows retirement rates by 1996 of husbands married to wives born 1929 to 1932, conditional on working in 1993. Circle size is proportional to the number of observations in the cell. Straight lines represent best linear fit on each side of the cutoff.

Figure A7. RDD, Retirement Rate in 1995, by Wife's Birth Quarter, placebo



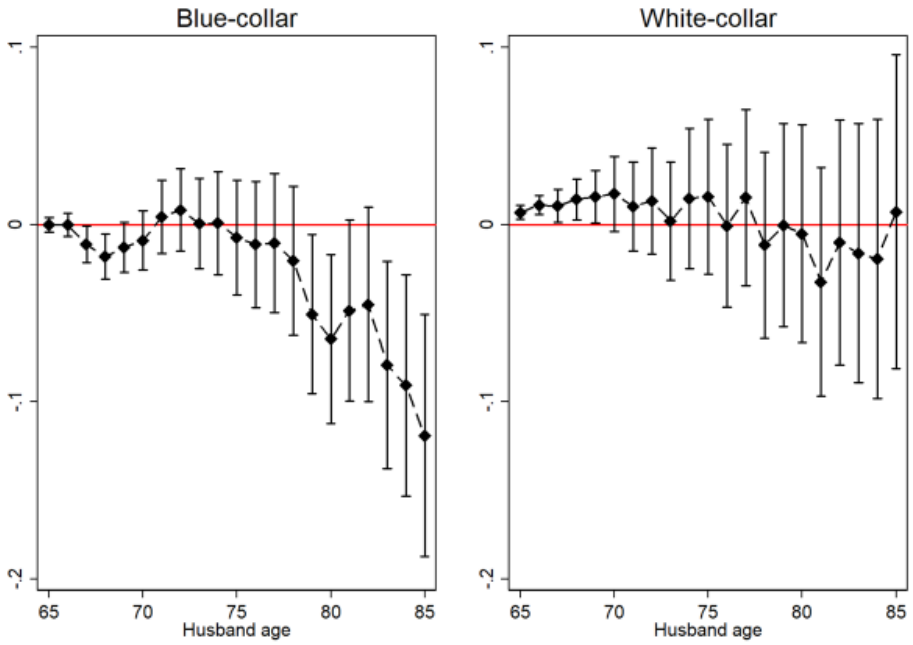
Note. This figure shows retirement rates by 1995 of husbands of housewives born 1929 to 1932, conditional on working in 1993. Circle size is proportional to the number of observations in the cell. Straight lines represent best linear fit on each side of the cutoff.

Figure A8. The Effect of the Housewives Reform on Survival (Excluding Young Husbands)



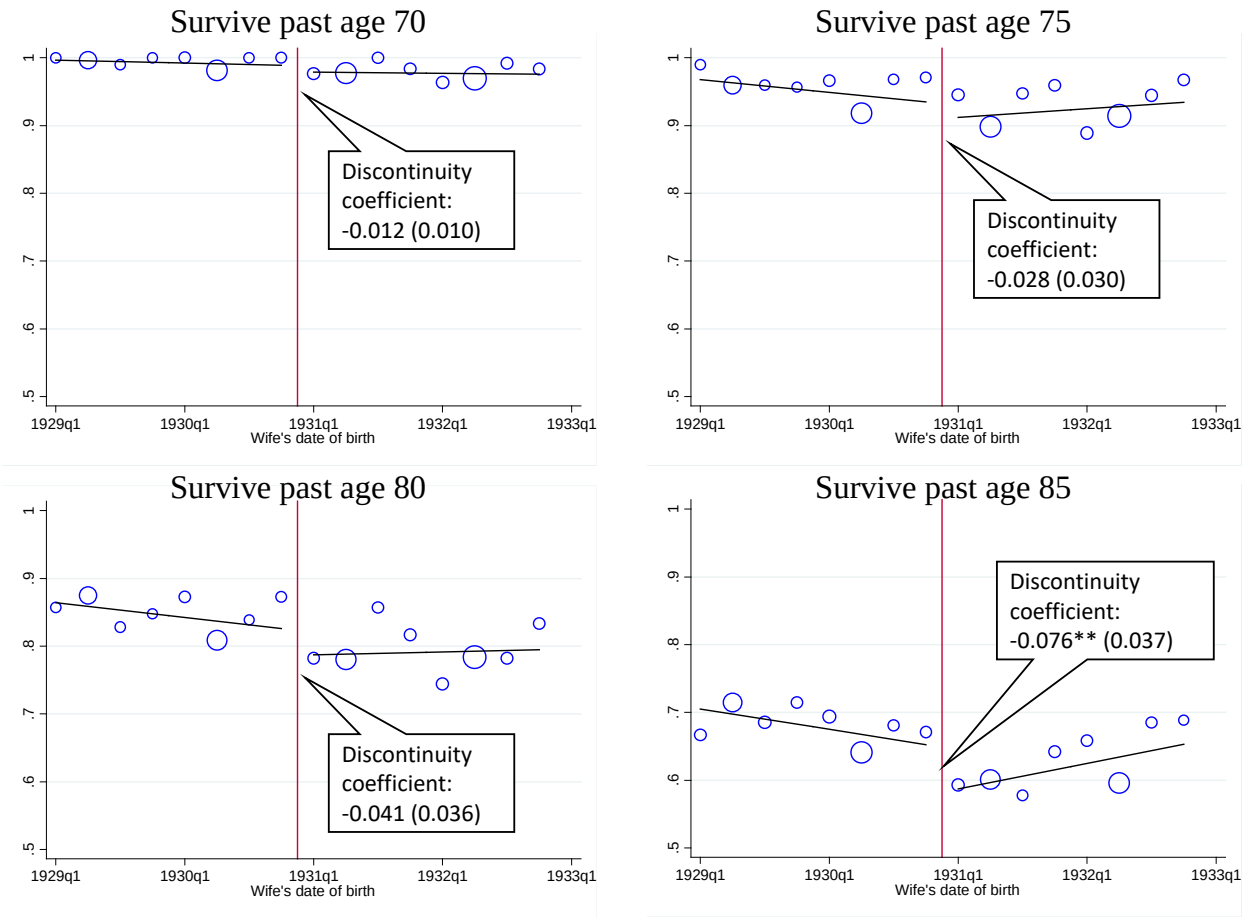
Note: This figure displays the reduced form results for survival by age when excluding younger husbands (born after 1935). The figure displays the coefficient of the instrument – the interaction term HW× Wife born Jan. 1st 1931 or later and its 90% bootstrap confidence interval, for each of the 21 Survival indicators (survival past 65 - survival past 85) as the outcome variables.

Figure A9. The Effect of the Housewives Reform on Survival by Occupation (Excluding Young Husbands)



Note: This figure displays the age-by-age reduced form results when excluding younger husbands (born after 1935), by industry type. The figure displays the coefficient of the instrument – the interaction term HW× Wife born Jan. 1st 1931 or later and its 90% confidence interval, for each of the 21 Survival indicators (survival past 65 - survival past 85) as the outcome variables.

Figure A10. RDD Estimates for the Effect of the Housewives Reform on Longevity



Note. This figure shows survival rates past different age cutoffs of husbands of housewives born 1929 to 1932, conditional on working in 1994. Circle size is proportional to the number of observations in the cell. Straight lines represent best linear fit on each side of the cutoff.

Appendix Tables

Table A1. Timing of the Reform

Date	Progress
June 1994	The parliament State Control Committee requests the National Insurance Institute of Israel (NII) to evaluate the discrimination
February 1995	The parliament Labor and Welfare Committee initiates the discussion about a new law (involves legislators and women rights activists)
August 1995	The Ministry of Labor forms the initial draft for the law
October 1995 to January 1996	Final Law is drafted, and signed. Most press coverage

Table A2. Comparing the main sample to the population

	1930 Cohort		1931-1932 Cohort	
	(1)	(2)	(3)	(4)
	Main Sample	Population	Main Sample	Population
Wife's characteristics				
Immigrant flag	0.807	0.834	0.802	0.824
	(0.395)	(0.373)	(0.398)	(0.381)
Jewish	0.964	0.914	0.959	0.915
	(0.186)	(0.281)	(0.198)	(0.279)
Immigration year	1952.5	1953.6	1953	1953.4
	(11.1)	(10.9)	(11.4)	(10.7)
husband's characteristics				
Husband's age	65.6	67.7	64.1	66.2
	(4.5)	(4.9)	(4)	(4.7)
Immigrant flag	0.84	0.868	0.826	0.852
	(0.367)	(0.338)	(0.379)	(0.355)
Jewish	0.963	0.891	0.959	0.895
	(0.19)	(0.311)	(0.198)	(0.307)
Immigration year	1951.7	1952.8	1952.2	1952.8
	(11.9)	(11.9)	(11.6)	(11.5)
Average Income when in 1993	77,523	21,178	83,011	26,974
	(102785.3)	(62408.4)	(103286.1)	(69081.9)
Average Income when wife is 1993|income>0	88,169	63,333	93,073	70,574
	(105248.4)	(94759.1)	(104998.5)	(97003)
Observations	2,642	10,181	6,438	20,460
(% in sample)	(26)		(31.5)	

Notes: Descriptive statistics for the sample of couples where wife was born between January 1930 and December 1932. Columns (1) and (3) show statistics for sample conditioning on husband's employment in 1994 (the main sample in the paper), while columns (2) and (4) show statistics for the entire population. All amounts are in NIS and deflated to 1996.

Table A3 Private Pensions (calculated for the years 2003-2007)

	1930 Cohort		1931-1932 Cohort		Diff in Diff
	(1)	(2)	(3)	(4)	(5)
	HW	Non-Hw	HW	Non-Hw	1931-32 vs 1930
Wife's average pension	27.9	15,531.5	4.9	15,680.9	-172.5
	(646.2)	(22,613.5)	(159.3)	(23,199.6)	(1,186)
Husband's average pension	33,085.8	40,284.4	32,050.2	39,440.2	-191.4
	(43,879.9)	(51,830.6)	(44,527.6)	(51,822.2)	(3,059.5)
Total household average pension	33,113.7	55,815.9	32,055	55,121.1	-363.9
	(43,980.1)	(61,656.9)	(44,528.3)	(61,948)	(3,519.7)
Observations	521	1,104	1,265	2,948	

Notes: Calculated for the sample of households with married wives born in 1930 or 1931-32, conditioning on husband's employment in 1994 and survival until 2007. Columns (1) and (2) for the 1930 cohort, and columns (3) and (4) for the 1931-32 cohorts. Column (5) shows the differences-in-differences for each characteristic (first taking the difference between HW and non-HW within cohort, and then taking the difference of the difference between them). All amounts are in NIS and deflated to 1996. Average pensions are calculated for the years 2003-2007.

Table A4

	Cumulative number of extra years worked							
Sample	Baseline		Spouse is immigrant		Spouse is Jewish		Spouse is 1-5 years older than wife	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HW× Wife born Jan. 1st	0.327**	0.285**	0.314**	0.291**	0.312**	0.272**	0.537***	0.473**
1931 or later	(0.131)	(0.127)	(0.141)	(0.138)	(0.137)	(0.134)	(0.206)	(0.2)
HH level controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	9,080	9,080	7,539	7,539	8,718	8,718	3,804	3,804

Note: Columns (1) and (2) replicate Table 5. Columns (3) and (4) restrict the sample to immigrant spouses, (5)-(6) restrict the sample to Jewish spouses, and (7)-(8) restrict the sample to spouses who are 1-5 years older than their wife. All regressions include a constant, HW dummy, and a dummy for 1931-32 cohort. HH level controls include a 3rd degree polynomial of husband log monthly earnings in 1993 (replacing log no-earnings with 0), and a 3rd degree polynomial of the husband-wife age difference as well as dummies for Jewish, and for immigrant status of both husband and wife, when applicable. Standard errors are calculated using Huber-White heteroscedasticity correction.

Table A5: The Effect of Employment on Life-Expectancy: RDD Estimates

		Reduced form	IV
		(1)	(2)
Outcome variable	Number of years survived between age 65 and 85	-0.385 [-0.89,0.059]	-1.059 [-6.147,0.502]
	Number of years survived between age 65 and 74	-0.075 [-0.19,0.02]	-0.198 [-1.097,0.076]
	Number of years survived between age 75 and 85	-0.31 [-0.741,0.084]	-0.861 [-4.551,0.562]
	First stage		0.386*** (0.108)
	Observations	2894	2894

Note: All specifications refer to housewife households and condition on husband's employment in 1994. Polynomials are allowed to differ on two sides of the 1931 quarter one cutoff. Household controls includes dummies for Jewish, and for immigrant status of both husband and wife, as well as a 3rd degree polynomial of husband log monthly earnings in 1993 (replacing log no-earnings with 0), and a 3rd degree polynomial of the husband-wife age difference. For each statistic, we report in square brackets its 90% bootstrap confidence interval (with 500 replications).