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

Itay Saporta-Eksten (Tel Aviv University and UCL)
Ity Shurtz (Hebrew University)
Sarit Weisburd (Tel Aviv University)

August 2018

Abstract We study the effects of public pension systems on the retirement timing of older workers and, in turn, the health consequences of delaying retirement by those workers. Causal inference relies on a social security reform in Israel that shifted payments from husbands to their (non-working) wives, thereby substantially reducing the implied tax on the husband's employment while keeping overall household wealth constant. Using administrative social security data, we estimate extensive-margin labor supply elasticities w.r.t. the average net-of-tax rate of about 0.6 for men over 65. Using the reform to instrument for employment, we find that working an additional full year at old age decreases survival probability past age 80 by 12%. This effect is driven by workers holding blue-collar jobs. Finally, we evaluate the effect of the reform on earnings. The results imply a small value for an additional year of life, suggesting that workers underestimate the health cost of employment at older ages.

Keywords: labor supply, social security, tax Reform, health, mortality

Acknowledgements: We are very grateful to Daniel Gottlieb, Miriam Shmelzer, Gabriella Heilbronn, and members of the National Insurance Institute of Israel staff for their help in conducting this study. Research funding from the Israel Science Foundation is gratefully acknowledged.

1. Introduction

Facing aging populations, almost half of all OECD countries implemented, in recent years, reforms to their public pension systems that encourage work among older workers.[1] Such policies, however, can potentially take a toll on older workers' health and longevity. Evidence on the effects of public pension systems on the employment of older worker and on the consequences of work at older ages on health is therefore key to informing a better design of such systems. In this paper, we use administrative data to draw causal inference on these two issues by leveraging a social security reform in Israel that changed the implied tax on delayed retirement while holding benefit generosity constant for a well-defined segment of the population—housewife households (hereinafter the "Housewives Reform" or the "Reform").

Concretely, we first study how the implied tax on delayed retirement, often induced by the existence of an earnings test in public pension systems, affects employment. Namely, we measure the extensive margin labor supply elasticity with respect to the average net-of-tax rate of older workers. While this is an important statistic in the context of public pension systems policy, the evidence pertaining to its size is mixed and inconclusive. The Reform provides a particularly favorable setting to estimate this statistic due to the existence of both a within-cohort unaffected group (non-housewife households) and a sharp eligibility age cutoff. Thus, we are able to implement a triple difference framework as well as a regression discontinuity framework. We then use the Reform to estimate the effect of employment at older ages on longevity. Despite the importance of the relationship between work and health and its increasing relevance for older workers, existing evidence on this issue is scant. Because the Reform created a change in employment incentives without affecting wealth, it provides an opportunity to estimate this relationship. We therefore use eligibility to the Reform as an instrument for the additional employment of elderly workers to study this question.

Our analysis uses the Housewives Reform that was implemented in 1996. The Reform shifted public pension payments from husbands to their (non-working) wives, thereby reducing the implied tax on the husband's employment while keeping overall benefit generosity roughly constant. The Reform implied a 10,000 NIS[2] (about $2,600) reduction in the implied annual tax on employment for husbands married to an eligible housewife—a housewife born after January 1st, 1931. This reduction was large, roughly amounting to a 16% increase in the average net-of-tax rate.[3]

[1] According to an OECD report (OECD, 2015), between 2013 and 2015 almost all OECD countries implemented some reforms to their public pension systems, with almost half of the countries conducting reforms that target work incentives.

[2] All NIS figures are reported in real 1996 terms.

[3] Because the reform shifted payments from husbands to their wives, a change in household decisions may arise due to a shift in spouses' respective bargaining power (see Chiappori & Mazzocco (2017) for a recent survey). However,

We apply a triple difference (DDD) specification to compare cumulative retirement rates of the full population of husbands in Israel married to housewives born in 1931 (treatment) and those married to housewives born in 1930 (comparison) pre- and post-Reform implementation in 1996. The identifying assumption underlying this analysis is that a husband married to a housewife born in 1931 would have followed a similar retirement trend to a husband married to a housewife born sometime in the previous calendar year absent the introduction of the Reform. To account for potential differences across these cohorts, we apply a third difference using husbands married to non-housewives that were born in 1930 versus 1931. We find that the Reform reduced retirement rates of affected husbands by about 6.4 percentage points at the year of its implementation (1996). Once husbands turn 70 years old there is no longer a tax disincentive to work, and indeed as husbands age the differences in cumulative retirement rates between treated and control husbands disappears.

Using the same triple differencing strategy we find that husbands married to housewives born in 1932 also show a significant reduction in retirement in 1996 immediately when the Reform is passed. Thus, these men decrease their retirement rates despite the fact that the reduction in the work tax will only appear one year later when their wives' turn 65. Our estimated effect of information regarding a future change in tax on delayed retirement is very similar in size to that found for older workers who incurred an actual tax change (the main treatment group). This result suggests that the Reform was well publicized and understood by the relevant population and is consistent with retirement being an irreversible (or close to irreversible) decision.

Because the Housewives Reform targeted a very specific segment of the population (housewife households where the wife was born after 1931) with two similar comparison groups, it naturally lent itself to a DDD specification. However, it is also true that the closest comparison group is households with housewives born just a little bit too early to be included in the Reform. We therefore complement the DDD analysis with a regression discontinuity design (RDD), where we use the sharp cutoff in birth date of the housewives (January 1st 1931). Focusing on the effect of the Reform on retirement rates on impact, we find RDD estimates, which are very close in magnitude to the DDD estimates. The RDD estimates corroborate the DDD results by showing that the response we measure arises sharply around the eligibility threshold of the Reform. Altogether, these estimates imply that the elderly exhibit moderate to high extensive-margin labor supply elasticities of about 0.6. We run two types of placebo tests for this analysis. First, we show that this decrease in retirement rates around the birthdate cutoff (January 1st 1931) does not hold for non-housewife households. Second, we show that the retirement rates of

because our setting involves elderly households, wives with no labor market attachment, and with almost zero divorce rates, it is likely that the effect of the reform on household behavior through this channel was small.

husbands in housewife households do not show any shift in retirement behavior around the birthday cutoff in 1994 and 1995, i.e. prior to the 1996 change in legislation.

Next, we use this exogenous shift in employment at older ages to examine the impact of work on longevity. While we find a positive association between employment and survival probability past the age of 80 in our sample, when we instrument for employment using the Reform, we find that the effect of employment on survival past age 80 is *negative*. Specifically, working an additional year decreases the probability of survival past age 80 by about 12%. Interestingly, this effect appears to be concentrated among blue-collar workers, suggesting that such jobs may be less suited for work at older ages.

We check the robustness of our health results using an RDD strategy where we examine the probability of surviving past 80 around the sharp cutoff in birth date of the housewives (January 1st 1931). In line with our instrumental variable analysis, in this reduced form analysis we continue to find a negative effect of employment on health. Namely, a decrease in survival rates after the birthday cutoff. Reassuringly, this effect disappears when examining non-housewife households.

Finally, we estimate the effect of the Reform on earnings, finding that the Reform increased earnings (net of benefits lost) by about 46,100 NIS. Combining this result with the effect of the Reform on survival we recover the value that treated individuals attribute to an additional year of life. Such calculation implicitly assumes that individuals internalize the health effect. The results indicate a very small value for an additional year of life, in the order of 47,500 NIS ($14,900). The small implied value calls into question the notion that when making employment decisions around retirement, workers fully understand the cost of employment in terms of their health and longevity and support the view that workers tend to underestimate these costs.

This paper is related to the large literature on employment incentives created by the social security system. The pioneering work by Krueger and Pischke (1992) studies the "notch generation's" employment response to the large reduction in social security benefits. While they find a limited effect, recently, Gelber, Isen, and Song (2016), using administrative data and an RDD designs, find that benefit reduction led to a substantial increase in labor supply that reflects a large income effect.[4]

Another body of literature studies the implications of the earnings test, which creates an implied tax on delayed retirement, to extensive-margin labor supply.[5] Taking a structural approach, French (2005) shows that eliminating the earnings test in the U.S. system (canceling the implied tax on delayed retirement) has a larger impact on the timing of retirement than reducing benefits or delaying the benefits

[4] Additional body of work studies the defined early retirement age (ERA) and full retirement age (FRA). See works by Mastrobuoni (2009), Blau and Goodstein (2010), Behaghel and Blau (2012) and Manoli and Weber (2016). See also the volume edited by Gruber and Wise (2004) and Coile and Gruber (2007) for international micro evidence using the option value approach outlined in Stock and Wise (1990).

[5] The earnings test implies that earnings above the earnings threshold means forgoing at least some retirment benefits.

eligibility age. Song & Manchester (2007) study the implications of the U.S. earnings test removal of 2000 and find inconclusive evidence on labor force participation while Friedberg & Webb (2009) report an increase in employment following the elimination of the earnings test in 2000. Baker and Benjamin (1999) and Disney and Smith (2002) report small or no changes in participation for elimination of the tests in Canada and the UK, respectively.

In a recent study, Gelber et al. (2017) develop a method to estimate the elasticity of participation with respect to the average net-of-tax rate using the kink in the budget set induced by the earnings test. They find larger elasticities than those previously reported in this literature. Leveraging the richness of the Housewives Reform, we estimate extensive margin labor supply elasticities as well. However, we rely on a very different identification approach, taking advantage of the sharp discontinuity in tax rates implied by date of birth within the treated population (housewife households) as well as the existence of a non-treated population within birth cohort (non-housewife households). Remarkably, the magnitudes of the elasticities we find in our quasi-experimental setting are very similar to those found by Gelber et al. (2017).[6]

A smaller strand of literature studies how employment affects health. Recent work is this area shows that job-loss is harmful to health (see e.g. Sullivan & Von Wachter (2009)) and that higher unemployment rates are associated with mortality (Gerdtham & Ruhm, 2006). However, these studies do not separate the role of employment from other aspects of job loss. Particularly, they do not distinguish between exogenous separations (layoffs), and endogenous employment decisions, which are more relevant in the context of retirement choices.

Snyder & Evans (2006) point out that while the notch generation in the U.S. received lower social security payments they had lower mortality rates than similar individuals who were born one quarter earlier and received higher benefits. They propose that this outcome may have been driven by the fact that the notch cohort was 5 percent more likely to work between the ages of 68-70 than the slightly older cohort. Thus, suggesting that the impact of the decrease in wealth on mortality was more than offset by the increase in employment. Our findings provide direct evidence against this hypothesis, especially when it comes to blue-collar jobs. We show that, in blue-collar jobs, additional employment has an adverse effect on mortality and health.

Studying the impact of early retirement on health, Kuhn, Wuellrich, & Zweimüller (2010) find that early retirement increases premature deaths. They show, however, that the adverse effects are likely to be focused on involuntary rather than voluntary job losses. Several recent studies find that early

[6] Another related strand of research examines the implications of the earnings test for intensive margin labor supply decision. See works by Burtless and Moffitt (1985), Baker and Benjamin (1999), Friedberg (2000), Disney and Smith (2002), Gruber and Orszag (2003), Haider and Loughran (2008), Gelber, Jones, and Sacks (2013) and Engelhardt and Kumar (2014).

retirement causes a reduction in mortality (see e.g. Hallberg et al. (2014) and Bloemen et al. (2017)).[7] Our work provides support for these results using a clean quasi-experimental source of variation in employment, which is quite rare in this literature.

The paper proceeds as follows. In section 2 we provide a brief review of the social security system in Israel and the Housewives Reform. Section 3 presents the data. Section 4 provides analysis of the impact of the implied tax of delayed retirement on labor supply. Section 5 reviews our results with respect to the effect of employment on longevity. Section 6 concludes.

2. The Social Security System in Israel and the Housewives Reform

Israel has a universal pay-as-you-go social security pension system, where contributions are withheld from the worker's salary up to a contribution cap. Each worker then receives his/her retirement benefits from the system at the end of his/her working life. Eligibility for retirement benefits depends on an individual's age and the *eligibility period*. There are two different age requirements. The *retirement age*—individuals who reach this age qualify for retirement benefits subject to an earnings test; the *eligibility age*—an age above which individuals are eligible to collect their retirement benefits regardless of their earnings. The retirement age in the relevant time period for the Housewives Reform was 65 and 60 for men and women respectively, and the eligibility age was 70 and 65 for men and women respectively.[8] The *eligibility period* is the total periods of social security coverage that an individual must accrue in order to qualify for retirement benefits. Unlike social security systems in many other developed countries that require accrual of employment periods in order to receive retirement benefits, to be covered by the Israeli social security system one only needs to be an Israeli resident.[9]

A *housewife* is defined in social security law as a married woman whose spouse is insured by the social security retirement benefits program and who does not have a working history. Prior to 1996 a housewife was ineligible for social security retirement benefits. Instead, her spouse would receive a supplemental payment to his social security retirement benefits to account for his wife as a "dependent". This means that while a man who did not work and was married to a woman who was eligible for social security retirement benefits would receive a social security pension at eligibility age, a woman in an otherwise similar situation would not. The discrimination between married men and married women

[7] Eibich (2015) and Insler (2014) show that retirement improves health through improved health behavior (see Eibich (2015) for a review of recent literature on these issues.).

[8] The retirement age is 65 for men who were born before March 1939 and 60 for women who were born before June 1944 - the relevant time for the Housewives Reform. For younger men this age gradually increased to 67 and for younger women it was gradually increased to 62. The eligibility age for men has not changed over the years. For women however it was gradually increased to age 67.

[9] While baseline benefits do not depend on employment, benefits are increasing linearly in number of employment years up to a cap (but are not a function of earnings).

received substantial public criticism and eventually the Israeli parliament changed the social security law with the aim of eliminating this discrimination against women. The change was applied, however, only to women born on January 1st 1931 or later (the Housewives Reform).[10] Thus, the Reform created a sharp difference in the benefits schedule of married couples with housewives born pre- vs. post- January 1st 1931. Under the old regime, the couple would have lost all benefits if the husband is employed (and earns a salary above the breakeven point in which the benefits are taxed), while under the new regime, a significant portion of the benefits is paid unconditional on husband employment, once his wife reaches 65. For husbands, this Reform is equivalent to a large decrease in the implied tax on delayed retirement.

In order to analyze the impact of the Reform, it is important to understand the timing and publicity of the announcement. To shed light on this issue, we searched news articles in two out of the three large newspapers in Israel (in terms of circulation), for the term "housewives" or "social security" from January 1st 1994 (two years before the Reform) through December 31st 1996. Table A1 summarizes the time-line of the legislation as recovered from these news articles. While the idea for the law was first raised in 1994, the first draft only appeared in the summer of 1995, and the final law was drafted and signed in the last quarter of 1995. Throughout this later period there were multiple mentions of the law in the press, including detailed articles explaining the change. This suggests that the public did have access to information about the nature of the Reform, but that this information was made available late in 1995, hence the behavioral responses are expected primarily in 1996.

3. Data and Sampling

Our analysis draws on administrative data from the National Insurance Institute of Israel ("NII")—Israel's Social Security Administration. These data are collected by the NII from various sources (including the IRS and the Ministry of Interior affairs) for internal use. The data contain information on employment history, earnings and social security benefits. They also contain demographic information such as country of origin, nationality, gender, date of birth, marital status and the birth date of each child. Importantly, these data can link spouses, allowing us to differentiate between husbands who are married to housewives and non-housewives and to determine whether the wife's precise birth-date results in the household being impacted by the legislative change. Furthermore, the data provide the date of death of each individual in our sample up to 2015, thus we can create an indicator for survival and use it to measure the impact of employment on health in the empirical analysis. Finally, the data includes information about eligibility for the NII long-term-care benefits. Eligibility requires undergoing a test that establishes the degree of inability to perform routine everyday tasks that is performed by an NII expert at the individual's home.

[10] In 2013, the law was changed again and applied the new rule to women who were born before January 1931.

Hence, data on long-term-care benefits eligibility may be used as a measure for an individual's general health. We create an indicator variable for eligibility to long-term-care benefits as an additional measure of health.[11]

In order to execute the analysis of the Housewives Reform, we created a dataset of all women who were born between January 1st, 1929 and December 31st, 1932, who were married in 1996, excluding self-employed, kibbutz members and new immigrants. For each woman, we indicate whether she is a housewife. We trace each woman's spouse and obtain information about both partners' age, employment history, earnings, and social security benefits. As we are interested in job exit behavior, we restrict the sample to households in which husbands are (still) employed when their wives are 62 years old unless stated otherwise. Finally, we restrict our sample to couples in which husbands and wives both survived until the wife turned 74.[12]

Table 1 provides some descriptive statistics of the Housewives Reform sample. The DDD identification strategy that we employ in Section 4.1 compares retirement decisions of husbands married to housewives belonging to the 1930 and 1931 cohorts after differencing out birth cohort effects using husbands married to non-housewives within each cohort. Therefore, columns (1) - (4) summarize the characteristics of housewife (or "HW") households and non-housewife (or "non-HW") households where the wife was born in either 1930 or 1931. Additionally, for each of the observable characteristics, we compare the differences-in-differences between HW and non-HW households over the two cohorts, in column (5) of the table. The 1930 and 1931 cohorts look well balanced on observables. Other than slightly younger husbands in HW households (compared to non-HW households) belonging to the 1931 cohort, there are no significant differences between the two age cohorts after the removal of cohort effects using non-Housewives.

4. The Impact of (Implied) Income Tax on Retirement

4.1. Conceptual framework

We introduce a static labor supply model to illustrate how the Housewives Reform is related to a broader set of typical social security reforms and to tie it directly to the impact of tax on labor supply on the extensive margin at old age. We then use the model to derive some predictions about the impact of the Housewives Reform on labor force participation.

[11] We consider an individual who passes away as eligible for long-term-care.

[12] We identify housewives using data on the receipt of social security benefits that is only available from 2003. Women that were born in 1929 (the oldest cohort in our sample) turned 74 in 2003. Thus, this is the minimal age for which we observe housewife status.

Consider a worker who is working from age 0 until age R and lives for another $T - R$ years after retirement. Suppose that there is no choice of labor supply on the intensive margin, and that the wage rate for each year of work is w. The worker is eligible for social security retirement benefits b per year starting at age R^0. The benefits, however, are subject to an earnings test, implying that for each additional year that the worker stays employed (and earns above the test cutoff) after R^0, the worker loses a portion τ of her annual retirement benefits. This setup captures the fact that delayed retirement schemes are not actuarially fair (or at least not perceived as such by some workers). For illustrative purposes, suppose that the earnings test threshold is zero. The worker's lifetime earnings as a function of retirement age are illustrated in Panel A of Figure 1.[13] If $\tau = 0$ (dashed blue line), the system is actuarially fair, i.e., for each forgone dollar due to delayed retirement, the worker receives an extra dollar after retirement.[14] In many social security systems (including the Israeli system), however, a forgone dollar in delayed retirement is compensated by less than an extra dollar post-retirement, introducing the kink in the budget constraint, namely, an implied tax on delayed retirement (τ^A , black solid line).

The effect of the Housewives Reform on the life-time budget constraint is very similar to a *decrease* in the implied tax on delayed retirement. Pre-Reform, delayed retirement is associated with forgoing the husband *and* the dependent's benefits. However, post-Reform, delayed retirement is associated with forgoing only the husband's benefits while the dependent's benefits are paid regardless of husband's employment status. In Figure 1, this corresponds to a reduction in τ (τ^B, dotted red line). The main takeaway from this budget set analysis is that *the Reform can be interpreted as reduction in the implied tax on delayed retirement* when the worker is older than R^0, *holding the generosity of benefits constant*. This feature of the Reform—creating a pure tax change—is quite unique among social security reforms analyzed in the literature, where it is usually the case that generosity of the benefits is changed (either directly or through change in eligibility age) at the same time that the tax on employment is changing.[15] Panel B of Figure 1 demonstrates the effect of another reform typically analyzed in the literature—a reduction in social security benefits (red dotted line). Such a reform indeed reduces the implied tax rate on delayed retirement, but at the same time decreases total benefits distributed. The figure demonstrates that analyzing such a reform captures both the wealth effect and effect of the tax change. Thus, the Housewives Reform provides an opportunity to study the implications of taxes on the employment decisions of older workers.

---

[13] We consider the benefits change to have the first order effect when forming predictions since our empirical specification is focused on agents who are affected by the policy after paying most or all of their life-time taxes. This allows us to abstract away from effects that could arise due to the impact of the changes in benefits on tax collection prior to retirement age.
[14] In this simple model this is also equivalent to a system without an earnings test.
[15] One exception is a reform that eliminates the earnings test.

Panel A of Figure 2 also provides a framework for interpreting the elasticities we recover. The change in policy is equivalent to a change in the after tax earnings when employed. To the extent that intensive margin responses are small (i.e. that w is not affected by the Reform), the change in slope captures a change in the average net-of-tax rate paid to an employed worker.[16]

To form a prediction about the effect of this policy change on the timing of retirement, consider a worker who derives utility from total life-time goods consumption and from life-time leisure. The worker maximizes utility under the life-time budget constraint described above (Panel A of Figure 1). Note that this formulation requires assuming perfect capital markets. If workers are heterogeneous in their disutility from work (or in wages) then pre-Reform some workers are working less than R^0 years, some bunch at R^0, and some work more then R^0. The first order and unambiguous prediction of the model is an (average) increase in labor supply for workers who pre-Reform choose to work exactly R^0 years (this is illustrated using indifference curves in Panel A of Figure 2). Workers who pre-Reform retired before age R^0 are expected to be unaffected by the policy change. The effect of the policy on workers who pre-Reform have chosen to work more than R^0, assuming that the labor supply curve is upward sloping, is delaying retirement. However, a tax reduction may induce earlier retirement if the income effect dominates the substitution effect in the labor supply response to wage changes.[17]

A final point, which is important to highlight using this framework, is that from the point of view of the household, the Reform has little effect on life-time earnings *other than* through changes in employment. This last insight is important because it motivates the use of the Reform as an instrument for employment when studying the impact of employment on health outcomes for these workers.[18]

4.2.The impact of the Housewives Reform on benefits

In this section, we show the effect of the Reform on the household's social security retirement benefits using data from 2003-2007 when almost all workers in our sample were eligible for old age benefits regardless of their employment status. The first four columns of Table 2 show how benefits are allocated between husband and wife for the 1930 and 1931 housewife and non-housewife households. Column (5) reports the differences-in-differences between HW and non-HW households in the two cohorts. The first row of the table shows the wives' retirement benefits. Housewives that were born in 1930 (column (1))

---

[16] In this simple model, total tax paid is $b\tau$, hence the average net-of-tax rate is given by $\left(1-\frac{\tau b}{w}\right)$.

[17] This framework does not capture credit constraints. Some individuals may be retiring at R^0 simply because they cannot borrow against their future retirement benefits. For these individuals, whose wives gained access to the housewife's benefit before they retired, the reform might actually incentivize early retirement.

[18] The only effect the reform had on household earnings, not through employment, is through the reduction in tax rate for those working during the years they work. We discuss this further in the context of the assumptions for the validity of the instrument in section 5.

received essentially zero retirement benefits while housewives that were born in 1931 (column (3)) received on average over 10,000 NIS, illustrating how, following the Reform, housewives became recipients of retirement benefits.

In the second row, the table displays the effect of the Reform on Housewives' spouses. Spouses in the 1930 HW households (column (1)) received on average about 24,000 NIS while those in the 1931 HW households (column (3)) received about 15,000 NIS on average. This difference arises because the Reform canceled the supplemental "dependent" payment for housewives' husbands. These numbers illustrate that the Reform caused a sharp change in the incentives to retire, substantially reducing the penalty on employment for the 1931 HW households. Notably, while wives in 1931 HW households receive a much larger share of household benefits than 1930 HW households, the total benefits collected at the household level are, on average, only slightly higher for the 1931 cohort. In other words, the Reform held the overall benefit level almost constant, shifting payments from husbands to wives. Overall, Table 2 establishes that the Reform corresponds to a pure change in the implied tax on delayed retirement. Appendix Figure A1 provides a graphical representation of the Reform. While the message of the figure is similar to that of Table 2, the figure illustrates the sharp cutoff in benefits collection for households with a housewife born pre- vs. post-January 1st 1931.

Note that while the average difference in benefits of wives in HW and non-HW households in the 1931 cohort is much smaller than the same difference in 1930 (due to the Reform), benefits of the non-HW group are somewhat larger. The source of this gap is the difference in employment histories between wives in the two types of households. Social security retirement benefits do not depend on earnings histories; they do depend, however, on employment histories. By construction, housewives do not receive credit for employment history, explaining the differences between wives' benefits in columns (3) and (4).

4.3. The response to the Housewives Reform – a DDD approach

The ideal experiment to study the effects of the Housewives Reform would involve a random assignment of housewife-households to a "treatment group" that is subject to the Reform's new rule and a control group that remains under the old, pre-Reform, rule. The environment we study lends itself to a standard triple difference (DDD) design that closely approximates such a thought experiment. Concretely, treated households are those with a housewife born on January 1st 1931 or later and the comparison group includes households with housewives that were born before this date. The identifying assumption is then that absent the Reform, post-Reform trends would have been similar for those two groups. However, pooling different birth cohorts for treatment and control to conduct the analysis, one might be skeptical about the common trend assumption. To address that, we invoke the DDD approach, much in the spirit of Gruber (1994), where we use within cohort non-HW households, to correct for potentially different trends across cohorts. We also show, in Section 4.4, that an RDD approach, which addresses this issue by comparing households around the cutoff date of January 1st, delivers very similar results.

Our first step is to run a differences-in-differences analysis to examine how the Reform affected retirement rates of the "treatment" households—those with housewives born in 1931—relative to non-HW households of the same birth cohort. We then run a similar DID analysis with our "comparison" households—those with housewives that were born in 1930. In the second step, we combine these two analyses to one DDD framework. Thus, the DDD framework provides estimates of the effect of the Reform on retirement rates in the treatment group relative to the comparison group while differencing out any birth cohort effects using non-HW households.

Figure 3 illustrates graphically the results from our DDD approach outlined above.[19] The figure shows, side by side, the retirement trends of husbands married to 1931-born wives (panel A) and husbands married to 1930-born wives (panel B). In each panel, cumulative retirement rates of husbands (conditional on working when their wife is 62) is graphed against wife's age for both husbands married to housewives, and to non-housewives.[20] The first thing to note in this figure, is that pre-Reform, conditional on working when their wife is 62, retirement trends are essentially identical within each cohort between husbands married to housewives and husbands married to non-housewives. These identical trends within cohort pre-Reform are reassuring for the validity of our DDD identification strategy which relies on changes in retirement trends for the 1931 housewives post-Reform after differencing out birth cohort effects using non-HW households.

Second, starting at the first year of the Reform, there is an apparent divergence between husbands married to housewives and husbands married to non-housewives in the 1931 cohort, with no similar divergence in the 1930 cohort. This divergence is due to a lower retirement rate of 1931 husbands married to housewives (relative to those married to non-housewives). This slower retirement is consistent with an increase in labor supply caused by a reduction in the implied tax on employment for this cohort.

Columns (1) and (3) of Table 3 report the regressions results that map to the lines in Figure 3. Each row in column (1) reports the difference between housewife- and non-housewife households in 1930 by wife's age (this corresponds to the difference between the two lines in the right panel of Figure 3). Column (3) reports the same numbers for the 1931 cohort. Column (5) reports the DDD estimates, i.e. the difference between columns (3) and (1). Columns (2), (4) and (6) pool together the pre-Reform years and report the same estimates as is in columns (1), (3) and (5), respectively. As the table shows, the effect of

[19] In our baseline specification, we define the last year before retirement to be the last year for which we see the individual working for at least 6 months and earning on average at least the monthly minimum wage. We show below that our results are not sensitive to the retirement definition.

[20] Note that the figures would be identical with calendar years on the x-axis instead of wife's age. In the regression analysis, one needs to choose between wife's age dummies and calendar year dummies. While we conduct the analysis using wife's age, the results are essentially identical using year dummies.

the Reform on retirement is statistically significant and economically large in the first year of the Reform. Using the estimates from column (6), conditional on working when their wife is 62, husbands married to 1931 cohort housewives are 6.4 percentage points (s.e. 1.9) less likely to retire when their wife turns 65 (i.e. the year they are affected by the Reform) compared to husbands married to wives that were born in 1930. As the wives age, the effect vanishes. This is not surprising, given that husbands are on average 2.5 years older than their wives, as the wives grow older, more spouses reach the eligibility age of 70, where they are not subject to an earnings test anymore (and therefore see no tax on delayed retirement).

These estimates can be used to recover an extensive margin elasticity of labor supply with respect to the average net-of-tax rate.[21] In recovering this elasticity, we must make some assumptions about the way that workers perceive the delayed retirement credit. We assume that workers are myopic or alternatively that they do not fully understand the delayed retirement credit system, whereby they do not realize that forgone benefits are replaced by delayed retirement credit. Hence, our assumptions place an upper bound on the percent change in average net-of-tax rate, and thereby a lower bound on the elasticity. In this case, the 6.4 percentage point decline in the retirement rate on impact implies an elasticity of 0.6. This magnitude is on the high side of estimates of extensive margin elasticities (See e.g. Table 1 in Chetty et al (2012)), however, it is comparable with the recent estimates reported by Gelber et al. (2017).[22]

Robustness tests and alternative designs.

We check the robustness of our results to sample and outcome definitions, as well as to alternative empirical designs. Columns (1) and (2) of Table 4 report the findings for the baseline DDD specification when restricting the control sample of non-HW households to include only households where the wife has less attachment to the labor force. To do so, we require that the wife does not work at age 62 or later, making the non-HW households more similar to the HW households. Reassuringly, the results are almost identical to the ones reported for the full sample. In columns (3) - (6) we evaluate the sensitivity of the results to different definitions of the retirement year. In columns (3) and (4) we report the results with an employment definition that requires only 3 months of employment in a given year,

[21] An alternative approach would have been to estimate directly the elasticity of employment w.r.t to the implied tax, instrumenting the tax change using the reform (see for example Gruber and Saez (2002)). However, we do not observe the benefits collected at the household level around the reform (only for later years), hence instead of imputing these, we estimate directly the effect of the reform on employment, and recover average elasticities using average social security benefits.

[22] In a recent paper, Manoli and Weber (2016b) estimate participation semi-elasticities w.r.t. financial incentives of between 0.1 and 0.3 applying bunching methods to Austrian data in the context employer provided severance payments. They find that elasticities are most significant for financial incentives that have a time horizon of 6 to 9 months.

while in columns (5) and (6) we require a monthly income above the earnings test threshold in order to consider someone employed.[23] The results are again very similar to our baseline results.

Since we define retirement based on the last year of employment, our empirical analysis is implicitly based on a "traditional" notion of uninterrupted employment that ends upon retirement. While this may be quite plausible for the population we study, we also examine whether our results are sensitive to this issue. To do so, we analyze actual employment, rather than retirement, in a given year by defining the independent variable as an employment dummy that takes the value of 1 if the husband is employed in a given year and zero otherwise. We then use this variable to analyze whether the Reform induced employment. The results of this robustness exercise, reported in columns (7) and (8) of Table 4, are almost identical to the retirement estimation results, with the opposite sign, showing that the results are not sensitive to this issue.

For the 1931 cohort analyzed above, the timing of learning about the Reform and actual receipt of payments overlapped at the end of 1995. However, for the 1932 cohort, while information about the Reform came in the end of 1995, actual payments for housewives started only in the beginning of 1997 (upon wives reaching the age of 65). Therefore, examining the response of the 1932 cohort provide an opportunity to understand how the knowledge regarding a future change in benefit taxation affected behavior. Figure 4, as well as column (9) of Table 4, report the results for the 1932 cohort analysis. As Figure 4 clearly illustrates, the response of the 1932 cohort began upon the arrival of the information about the Reform. Namely, spouses of housewives in the 1932 cohort begin to delay retirement in 1996, the first year of the Reform. This kind of response is consistent with irreversibility of the retirement decision – even though payments start only in 1997, retiring in 1996 implies that the husband does not work in 1997 as well, hence will not enjoy the implied tax reduction. The results are similar to the results of the 1931 cohort analysis in terms of their magnitude, with a somewhat more persistent effect of the Reform on retirement. This more persistent response is consistent with the fact that this group was initially exposed to the Reform at a younger age and thus, had longer to work before reaching age 70.

Cumulative effect on employment.

So far, we have focused on the year-by-year response of retirement to the tax decline induced by the Reform. Table 5 reports the effect of the Reform on cumulative husband's employment.[24] This is useful for two purposes. First, the cumulative effect summarizes the total effect of the Reform on employment. Second, in Section 5, we explore the effect of delayed retirement on long run health outcomes and survival. Naturally, survival is affected by the entire history of employment, hence

[23] Note that these definitions affect sample size as we condition on non-retirement by the year the wife turns 62.
[24] We calculate the husband's cumulative employment during the 5 years after his wife turns 65. We experimented with extending this period to as long as 10 years and found no qualitative difference in the results.

neglecting the cumulative effect (for example by associating the entire health effect with the 1996 employment effect) would result in an over-estimate of the effect of employment on health. Our approach compares the difference in cumulative employment of husbands married to housewives and non-housewives born in 1931 to the equivalent difference for husbands married to housewives born in 1930.[25] Columns (1) and (2) of Table 5 report the effect for the sample used in our DDD analysis, showing that, overall, the Reform increased cumulative employment (or delayed retirement) by 0.29 to 0.36 years of work. In columns (3) and (4) we analyze a sample of husbands who were still working in 1993 and 1994.[26] These husbands, who were more consistently employed prior to the Reform, are likely to be most affected by the 1996 change. Indeed, when constraining ourselves to this sample, we find a larger effect of the Reform on cumulative employment of 0.47 years of work. We will use this sample, and take advantage of this larger employment effect, to maximize statistical power when we study the effect of employment on health. In other words, these estimates are the first stage results for the analysis of the effect of employment on health that we perform in Section 5.

4.4. The Response to the Housewives Reform – an RDD Approach

In this section, we complement the DDD analysis with evidence from a regression discontinuity design. This approach exploits the sharp age-based rule within a regression discontinuity design framework. To illustrate how this would work, consider two housewives: one that was born on January 1st 1931 and another that was born on December 31st 1930. Assuming that the wives' exact date of birth is uncorrelated with their other characteristics and particularly their husbands' retirement decision, comparing the retirement patterns of their husbands resembles the ideal experiment that examines the effect of the Reform on husbands' retirement, that we described in the previous section.

More formally, let τ indicate the wife's date of birth in terms of days elapsed since January 1st 1931. For example, if a wife was born on December 20th 1930, $\tau = -12$; if she was born on January 10th 1931, $\tau = 9$. Let the treatment indicator, D, equal 1 if the wife was born in January 1st 1931 or later, and 0 otherwise. Consider the following model relating the husband's timing of retirement (y) with the wife's date of birth in terms of τ and the treatment indicator:

(1) $$y = \alpha + \beta D + f(\tau) + \epsilon.$$

$f(\tau)$, is a completely flexible control function, and it is continuous at $\tau = 0$. The parameter of interest in this model is β that measures the causal effect of the Reform on y. Intuitively, given that $f(\tau)$ absorbs

[25] This approach is consistent with the results in Tables 3 and 4 showing that there are no pre-trends in those differences.
[26] Because we have data on date of death through 2015, we further require that husbands' year of birth is before 1935 in order to have full information about their survival at age 80.

any continuous relationship between a wife's date of birth and her husband's retirement decision, the coefficient β estimates the discontinuous relations between the Reform and the husband's retirement decision. We estimate such a model using standard regression discontinuity design methods (see Lee and Lemieux (2010) for a survey).

Motivated by the DDD results, the RDD analysis aims to examine how likely a husband that was employed in 1993 is to retire by 1996—the first year of the Reform—as a function of his wife's date of birth.[27] Panel A of Figure 5 displays the results of the RDD analysis for HW households. The figure shows the retirement probabilities by quarter of birth of the wife, illustrating that there is a sharp drop in retirement probabilities for affected husbands. The corresponding estimates are reported in columns (1) - (4) of Table 6. Column (1) shows a statistically significant drop of over 7 percentage points in retirement rates of husbands of housewives using a specification with a linear polynomial and no household level controls. The result is unaffected by the inclusion of household level controls, as column (2) demonstrates. The results are also very similar when we repeat the analysis using a quadratic polynomial, as columns (3) and (4) of the table show.

Our setting provides an opportunity to examine the validity of these results by taking advantage of the existence of the non-HW group. Panel B of Figure 5 illustrates the results of a placebo exercise using the non-HW households. To focus on households that resemble HW households, we restrict the sample to include households where the wife has less attachment to the labor force, using the same criterion we applied in the previous section, namely, non-HW households in which the wife was not employed as of age 62. As the figure illustrates, this group does not exhibit a similar pattern of retirement around the January 1931 threshold as they were unaffected by the Reform. Columns (5) - (8) of Table 6 provide the corresponding estimates. Overall, consistent with the visual impression of Figure 5, the estimates show no indication of an effect in the non-HW group. These results corroborate the interpretation of the results as stemming from the Housewife Reform. They clearly show that the delay in retirement that we documented in the DDD analysis arises sharply around the January 1931 threshold and only in the case of the HW group.

Next, we look for any indication that observable characteristics (that are determined pre-Reform) change sharply around the January 1931 threshold. If this were the case, it could raise the concern that selection could be affecting our results. In panels A - C of Figure 6 we examine the behavior of the age-gap between husbands and wives, the log of husband's earnings when the wife is 63 (i.e. pre-Reform), and the share of immigrants among husbands, respectively. All three variables appear to trend quite

[27] Conditioning on employment in 1993 is useful because it leaves room to validate the results by using retirement in 1994 as a placebo test. In principle, retirement in 1995 is another placebo, but it practice it may be "contaminated" by the effect of the reform. Below we run both tests and further discuss this point.

smoothly around the January 1931 threshold. In panel D we report the husband's predicted probability to retire by 1996 using a model that includes 3rd order polynomials in the first two variables (age-gap and husband's log monthly earnings), and the husband's immigration status. As Panel D of Figure 6 illustrates, the predicted values generated by this model also appear to trend smoothly around the threshold. Table 7 provides the corresponding estimates. As the table indicates, there are no statistically significant discontinuities in the three observables we analyze, as well as in the predicted values of the probability to retire.[28] Overall, this analysis shows no indication that our results are an artifact of sample selection.

As we described above, information about the Reform only became available towards the end of 1995. Therefore, while there may have been some response towards the end of 1995, the Reform should not have affected the retirement decisions of households in 1994. However, if the drop in retirement rates arises because spouses of the 1931 housewives cohort tend to retire later regardless of the Reform, we would expect this to also manifest in the 1994 retirement decision. Thus, we conduct an additional validity test in which we replicate the RDD analysis using retirement in 1994 as the outcome variable. Panel A of Figure 7, and Table 8 display the results of this validation exercise for the HW households. As one might expect, retirement rates are lower in 1994, yet they are still substantial. Retirement rates trend smoothly around the treatment threshold, with no indication that husbands belonging to the 1931 housewives cohort have a particular tendency to retire less before the Reform. For completeness, Panel B of Figure 7 displays the corresponding analysis for the non-HW group, also showing no effect.[29]

5. The Effect of Employment on Health

So far, we have established a causal link between the decline in taxation and delaying retirement. We turn now to the second question that we have posed in this paper – what is the effect of extended employment on health? To address this question, given our setting, we analyze a model of the form

$$(2) \qquad Survival = \alpha + \beta_1 \cdot Employment + \beta_2 \cdot HW + \beta_3 \cdot \text{born\_1931} + X \cdot \gamma + \epsilon$$

[28] Appendix Figure A2 complements this analysis, by showing that there is no discontinuity in the density around the cutoff. Our birth date data is discrete at the monthly level, hence we cannot conduct a formal McCrary test (McCrary, 2008). Panel A of the figure shows the distribution of birth months without any controls. As missing birth months are recorded as April, there are noticeable spikes in number of records in each April, as well as slightly more records in January. Panel B, shows that controlling for 12 monthly dummies, the density is very smooth across the cutoff.

[29] We repeat the placebo exercise using retirement in 1995 as the outcome variable. This placebo exercise is not as clean because we can expect some effect on retirement towards the end of 1995 (see Appendix Table A2 for the timing of the reform). The results are reported in Figure A2 and Table A2, showing a negative, small and statistically insignificant result.

$Survival$ is our main health outcome. It is defined as a dummy variable that equals 1 if a husband survives past a given age threshold (we use 80 in the main specification). $Employment$ is defined as in Table 5–the number of years a husband worked after his wife has turned 65. HW is a dummy variable that equals 1 if the husband belongs to a HW household, $\mathrm{born\_1931}$ is an indicator for households whereby the wife was born in 1931, and X is a vector of household characteristics. Naïvely analyzing this model using OLS for example, is likely to provide biased estimates of β_1, the effect of employment on survival, because of underlying unobserved factors that affect both the employment decision and the survival of the individual. For example, Individuals with a health condition, unobserved by the econometrician, may tend to work less and have a lower likelihood to survive past 80, generating a positive association between employment and survival. Here, we aim to study this relation using the Housewives Reform as an exogenous source of variation in employment. Namely, we estimate the model in Equation (2) using the Housewives Reform and specifically the interaction term between HW and $\mathrm{born\_1931}$ as an instrument for $employment$.

Before proceeding to the results of this analysis, a discussion about the validity and interpretation of this instrument is warranted. Is it reasonable to assume that the exclusion restriction holds? Namely, that the Reform affected health only through its effect on employment. One obvious alternative channel is that the Reform affected household *resources* not through employment. However, Table 2 shows that overall household social security benefits remained on average almost unchanged for households with retired husbands affected by the Reform compared to households with unaffected husbands. Therefore it is unlikely that differences in health are driven by differences in life-time benefits post-retirement.[30] It is important to note that while there was not a substantial change in the household benefits for the 1931 cohort post-retirement, the Reform could have increased income through two other channels: First, delaying retirement may increase income as workers accumulate more years of labor earnings. Second, the Reform affected net income of the employed that *did not* change their employment behavior by increasing their average net-of-tax rate. This implies that on average those treated by the Reform earned higher incomes until they reached age 70. We argue that these income increases are not a source of concern to our identification strategy for two reasons. First, they are likely to have only a small impact on lifetime earnings. Second, even if the Reform did increase household resources, we would expect the increased resources to increase survival of the treated husbands (for example by providing additional funds for heating or medical expenses). Yet, the results we report below indicate that the Reform caused a decrease in survival of the treated husbands. Therefore, to the extent that increased resources affect the results, they attenuate them towards zero.

[30] For the years 2003-2007, using non-HW households to control for benefit differences across cohorts, column (5) of Table 2 shows that annual social security benefits are only 5.8% (1,379/23,971) higher for treated households.

The second condition for the validity of the Reform as an instrument is the existence of a first stage. We reported the first stage results, indicating a statistically significant increase of close to six months in employment, as part of our extensive discussion in Section 4 about the effect of the Reform on employment.[31] Notably, the F-stat for the effect of the Reform on employment is 7.8, which is slightly below the standard threshold often discussed in the literature.[32] Given that we have one endogenous variable with one instrument our IV estimates are median-unbiased, however inference could be problematic. To address this issue we calculate confidence intervals following Chernozhukov & Hansen (2008), and correcting for heteroscedasticity.

5.1. The Causal Link between Employment and Health

To implement the analysis, we only include husbands in our sample who were employed in 1993 and 1994. We analyze the effect of extended employment on their health, taking advantage of the Reform as an exogenous source of variation in employment. As we reported in columns (3) and (4) of Table 5, for this sample, the Housewives Reform induced on average close to six months of additional employment in the 5 years after their wives turned 65. In columns (1) and (2) of Table 9 we report the reduced form regression results. As the table indicates, the Housewives Reform was associated with a decrease of about 6.5 percentage points in the likelihood of the husbands in the sample surviving past age 80.

Before turning to our IV specification, we show results from the OLS estimation of equation (2) in columns (3) and (4) of Table 9. The results are small, but positive and significant. Even though they merely capture the correlation between employment and longevity, we note that the positive sign of these estimate is consistent with the typical simultaneity problem of the retirement-health link (see for example Insler (2014) for a literature review).

In the IV estimation results in Table 9, column (5), we find a *negative* and statistically significant (at the 10% level) 14 percentage point decrease in the likelihood to survive past 80 for an additional year of employment. The Chernozhukov-Hansen ("C-H") 90% confidence interval also contains only negative values. When adding household level controls, the estimates decrease to 11 percentage points (column (6)). These estimates reflect a decline of about 12% in survival past 80 for an extra full year of husband's employment at older age. We find very similar results to the ones reported in Table 9 when examining the eligibility for long-term-care benefits (See Appendix Table A3).

Similar to our approach in Section 4.4, we take advantage of the January 1931 threshold and we analyze the effect of employment on health outcomes applying an RDD approach. Figure 8 illustrates the

[31] The first stage results, those corresponding exactly to the sample we analyze in this section, namely, husbands who worked in the years 1993 and 1994 are reported in columns (3) and (4) of Table 5.
[32] E.g. in chapter 4 of Angrist & Pischke (2008) an F-stat of 10, based on Stock, Wright, & Yogo (2002) is regarded as the safe zone.

first stage. As the figure shows, our measure of employment, the number of years a husband worked after his wife has turned 65, is downward sloping as a function of his wife's date of birth. This is induced mechanically because we condition on work in 1993 and 1994, creating a negative correlation between having an older wife and more years of employment after the wife reaches age 65. Around the January 1931 threshold, there is a discernable jump of about 0.3 years in the employment measure. Columns (1) and (2) of Table 10 display the corresponding first stage results, reflecting an increase of 0.34-0.5 years in employment, similar to the effect we find in the differences-in-differences results.[33]

Figure 9 shows the probability of husbands surviving past the age of 80 as a function of their wife's quarter of birth. As panel A of the figure shows, there is a discontinuity at the first quarter of 1931 in the survival probability for husbands married to housewives. The magnitudes are reported in columns (3) and (4) of Panel A of Table 10. These magnitudes are consistent with the reduced form magnitudes reported in Table 9. To recover the effect of employment on health we can run an IV-RDD specification, whereby we use the discontinuity as an instrument for employment. The effect on cumulative employment as measured using this RDD approach is very similar to the effect reported in Table 5. Columns (5) and (6) of Panel A of Table 10 report the IV estimates using the discontinuity as the instrument for employment. The point estimates are very similar to the ones we obtain using the differences-in-differences approach in Table 9, and the C-H 90% confidence intervals indicate that the results are significant at the 10% level. Appendix Table A4 shows that our results are again similar when using an alternative health measure of Long-Term Care eligibility.

Columns (1) and (2) of Panel B of Table 10 examine non-HW households as a placebo test and illustrates that there is no discontinuity around the first quarter of 1931 of cumulative employment for husbands married to non-housewives (see also Panel B of Figure 8). Similarly, columns (3) and (4) of Panel B of Table 10 show no discontinuity in survival past age 80 around this date (see also Panel B of Figure 9).

5.2. Heterogeneity

Next, we examine whether the effect of employment on health varies across worker types. Ideally, we would like to compare the effect of work at older ages on health across different occupations. While we do not have direct information about the individuals' occupation, we use the industry in which the individual was employed in 1993 as a proxy for the individual's likelihood of working in blue- versus white-collar occupations. Specifically, we characterize each industry as blue- or white-collar using the

[33] Note that the sample size here is different from that of the sample used in the RDD analysis in Table 6. This difference arises because here we condition on working in both 1993 and 1994 (as we do throughout the health analysis) while in the previous RDD analysis we only conditioned on work in 1993 in order to allow the placebo analysis using retirement in 1994 as an outcome (Figure 7 and Table 9).

Israeli labor force surveys from the periods 1995-2000. These surveys contain information about the composition of employee occupations in each industry. We define an industry as blue-collar if at least 60% of the industry's employees have blue collar occupations.[34] With blue- and white-collar defined, we run the same analysis as above separately for each of the two groups.

Table 11 reveals a stark difference between the health effect of employment at older ages on blue and white collar workers. While the IV estimate for the blue-collar group is negative and statistically significant at the 10% level with a point estimate of -0.334 (s.e. 0.196), the corresponding estimate for the white-collar group is positive, close to zero, and statistically insignificant (0.012 (s.e. 0.069)). These results indicate that the effect we find in the full sample is concentrated among blue-collar workers. Namely, while it is true that on average employment adversely affects health, this effect arises primarily in blue-collar jobs.

It is interesting to try to gain a better understanding of the drivers of this response. The first-stage results in the blue- and white-collar groups are 0.479 (s.e. 0.251) and 0.485 (s.e. 0.225) years of work, respectively. Turning to the reduced form regression (column (1) and (4)), the estimates are -0.16 (s.e. 0.039) for the blue-collar group and 0.006 (s.e. 0.033) for the white-collar group. Putting together these results, it appears that while the response to the Reform, in terms of employment, is very similar for the blue- and white-collar groups, the effect on health seems to occur only in the blue-collar group.[35] That is to say, we find that the health of blue-collar workers is affected more adversely by the Reform, not because they delay retirement more than white-collar workers do, but because their jobs take a higher toll on their health.

5.3. The Health Effect of the Reform and the Value of Life

The results we report in the previous sections indicate that employment has a large negative effect on health. However, they also raise another question – are workers aware of the significant cost they pay for prolonged employment in terms of their health? Providing an answer to this question is difficult. Here, we offer a two-step approach to gain some insight about this question by using the Reform to recover the value of an additional year of life as it is reflected in our results. First, we estimate the effect of the Reform on household earnings. Second, we recover the effect of the Reform on life expectancy using our estimates for the effect of the Reform on survival and some assumptions about the functional form of the

[34] White-collar occupations are defined as the following major categories (equivalent to major categories 1 through 4 in the ISCO-88 occupation classification system): Legislators, senior officials and managers, Professionals, Technicians and associate professionals and Clerks.
[35] The C-H confidence intervals confirm that these findings are unlikely to be driven by differences in the predictive power at the first stage.

survival distribution. Together, this allows us to elicit the value that workers attach to an extra year of life, under the assumption that workers are aware of the health consequences of employment at older ages.

Column (7) of Table 9 reports the reduced form effect of the Reform on cumulative labor earnings. Labor earnings of husbands to 1931-born housewives, who are affected by the Reform, are about 53,400 NIS higher compared to the control group. While we do not observe benefits directly for that period, we can calculate that for the affected husbands, the average benefit loss from working an extra few months was about 7,300 NIS. Thus, overall net labor earnings for the treated group increased by about 46,100 NIS due to the Reform. Column (8) repeats the exercise of using the Reform as an instrument for employment in order to recover the effect of an extra year of employment on earnings, showing an increase in earnings of about 114,300 NIS.

To recover the change in life expectancy, we must make some functional form assumptions about the distribution of survival. First, we note that due to our sample restrictions, almost all husbands in our sample survived to age 75. Second, we assume that we can reasonably characterize the survival function using a Weibull distribution, and we assume the survival to age of a 100 is approaching zero for both groups. We separately fit a Weibull distribution to affected and non-affected husbands, and find that under these assumptions, the difference in the probability of survival to age 80 implies that husbands affected by the Reform lose about 12 months of their life due to working more.[36] Combining this increase in mortality with our measured increase in net labor earnings implies a value of about 47,500 NIS ($14,900) for an extra year of life. This estimate is well below the "Value of a Life-Year" typically found in the literature.[37]

Returning to the question we posed in the beginning of this section, the results indicate that if workers are aware of the health costs of prolonged employment then the value they assign to an additional life-year is extremely low. Since there is no reason to suspect that this is the case, the results suggest that elderly workers may be jeopardizing their health because they are unaware of the longevity costs of employment.

6. Conclusions

In this paper, we study the effects of public pension systems on employment of older worker and, in turn, the health consequences of older workers' employment. We leverage a social security Reform in Israel that shifted payments from husbands to their (non-working) wives, thereby reducing the implied tax on the husband's employment while keeping overall benefit generosity roughly constant. We estimate

[36] If we assume a 1% survival rate to the age of a 100 (rather than survival approaching zero), the implied loss increases slightly to 13 months.
[37] Murphy & Topel (2006), for example, calibrate that the value of a life-year at age 70 is over $200,000 in 2004 dollars (see also Hall & Jones (2007) for further discussion).

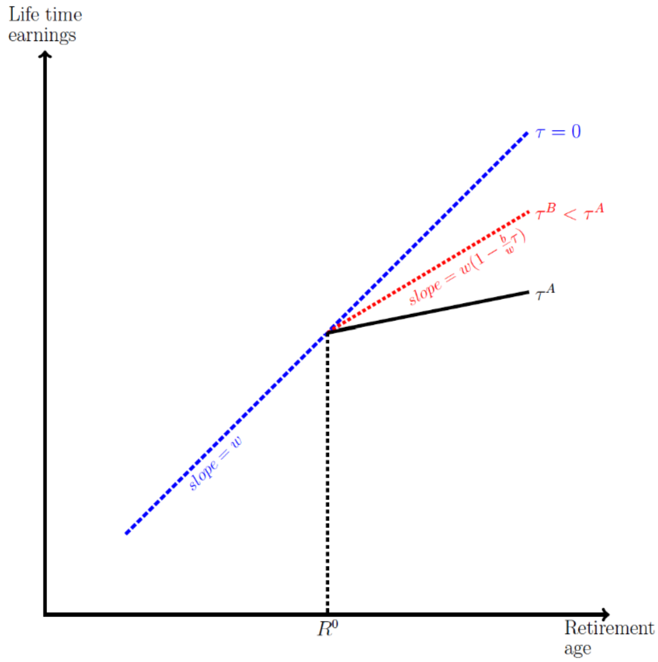
extensive-margin labor supply elasticities of about 0.6 for men over 65 w.r.t. the average net-of-tax rate. These numbers, which are consistent with other recent evidence on this issue, support the view that the existence of an implied tax on employment substantially affects the retirement timing of older workers.

We then estimate the effect of employment on longevity of older workers using the Reform to instrument for employment. While we find a cross-sectional positive correlation between employment at old age and longevity, IV estimates indicate that working an additional full year at old age decreases survival probability past age 80 by about 12%. Importantly, this effect appears to be driven by workers in blue-collar occupations. Finally, we estimate the effect of the Reform on earnings, finding that the Reform increased earnings (net of benefits lost) by about 46,100 NIS. Combining this result with the effect of the Reform on survival, we find that treated individuals attribute a very small value to an additional year of life. This result calls into question the notion that when making employment decisions around retirement, workers fully understand the cost of employment in terms of their health and longevity and supports the view that workers tend to underestimate these costs.

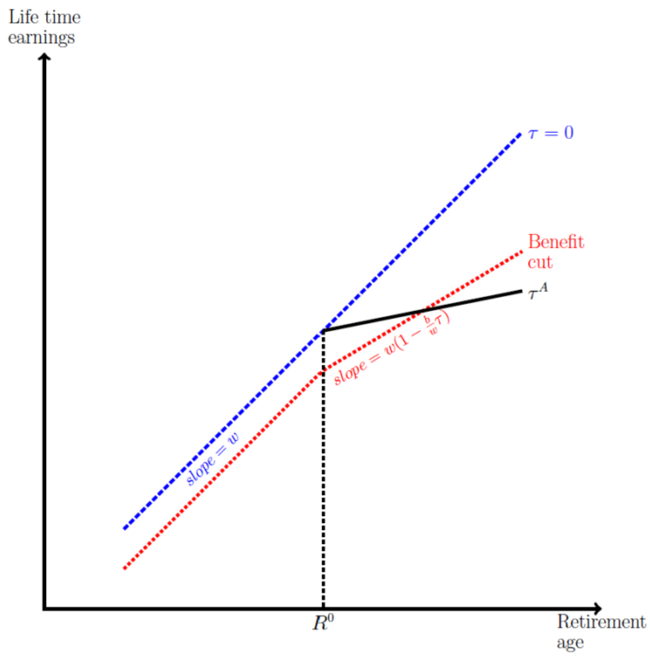
Granted that recent years have seen many developed countries implement changes to their public pension systems that encourage work among older workers, the results in this paper suggest that a better understanding of factors underlying employment decision-making, as well as of the broader impact of employment at old ages, are crucial for policymaking.

Figure 1. Life Time Budget Constraint Effects of Social Security Reforms

A. The Housewives Reform



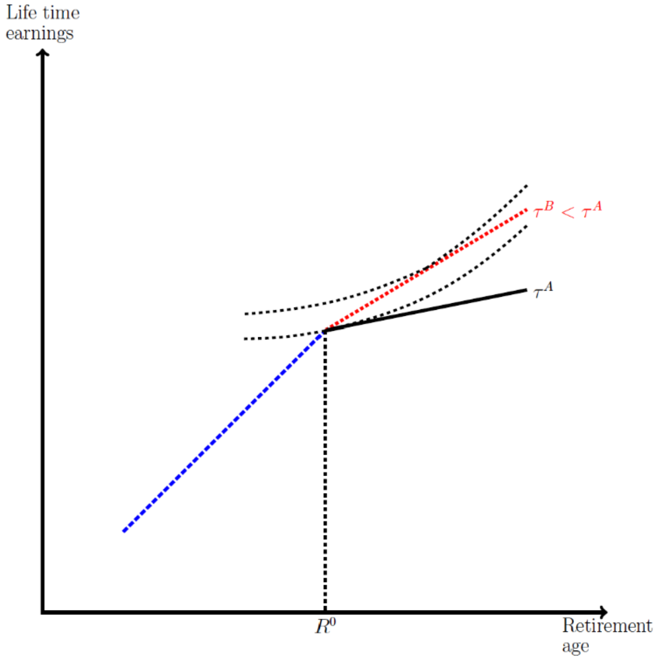
B. A benefit reducing social security reforms



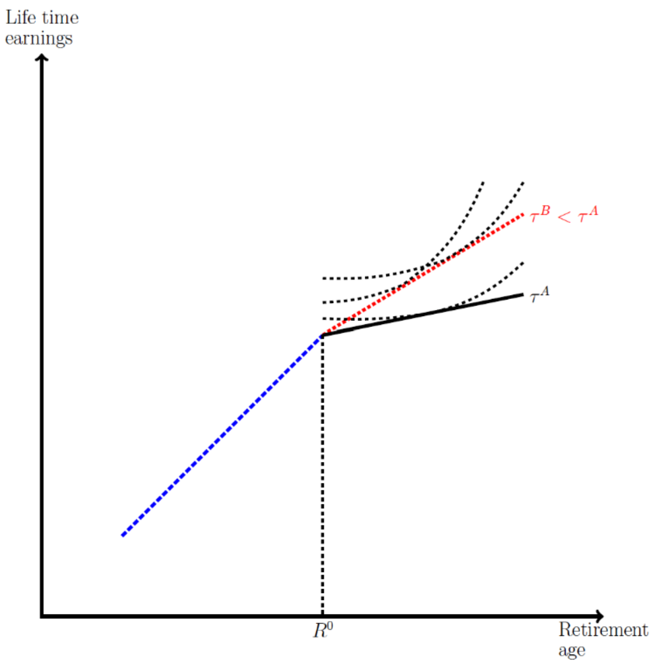
Note: Authors illustration. In panel A, the (blue) dashed line, denoted $\tau = 0$, represents actuarially fair social security system. The (black) solid line, denoted τ^A, represents an earnings test—a tax on delayed retirement. The (red) dotted line, denoted τ^B, illustrates the effect of the Housewives Reform on the budget line, a decrease in the implied tax on delayed retirement. In panel B, the (red) dotted line, denoted Benefit cut, shows the impact of a reduction in social security benefits on life-time earnings, illustrating that such a reform causes a decrease in the implied tax on delayed retirement as well as a decrease in overall life-time earnings.

Figure 2. **Expected Behavioral Responses to the Housewives Reform**

A. Retire at R^0

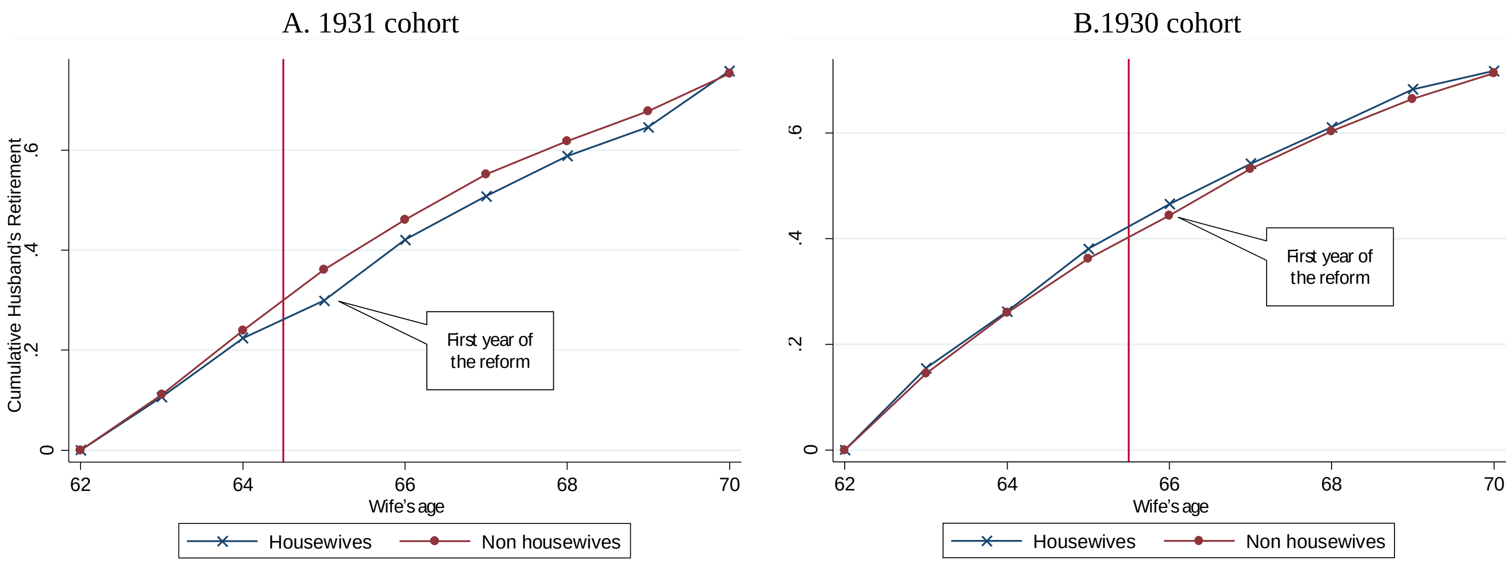


B. Retire after R^0



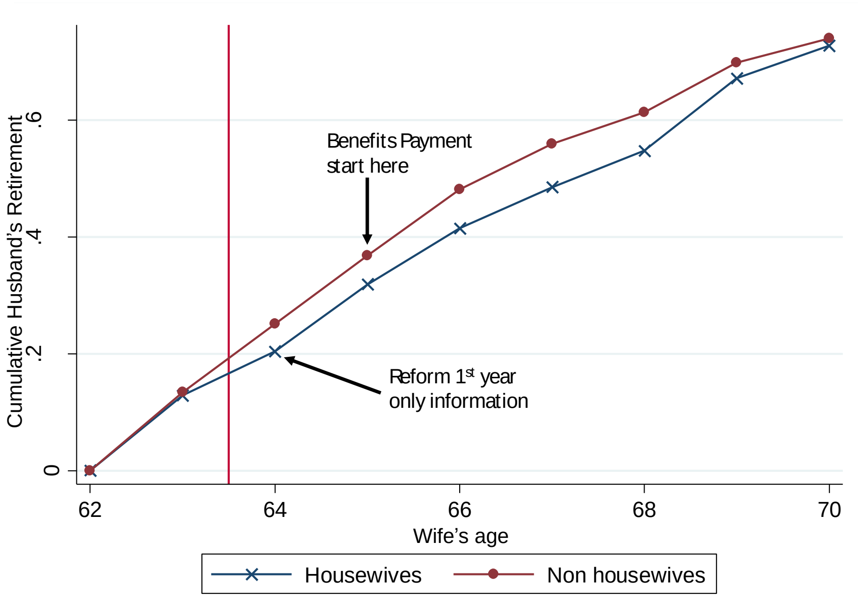
Note: Authors illustration. Panel A of the figure illustrates the response to the Reform by an individual whose intended timing of retirement pre-Reform was R^0. Panel B illustrates the response of an individual whose intended date of retirement was later than R^0.

Figure 3. DDD Results: Husbands Affected by the Reform Retire Later



Note: Results from differences-in-differences regressions using the sample of couples that were married in 1996, conditioning on husband's employment when wife is 62. The differences-in-differences regressions were conducted separately for the 1930 and the 1931 cohorts to illustrate the patterns for all 4 groups involved in the analysis. The results reported in the text are from the DDD regression which takes the difference between the differences between the two graphs.

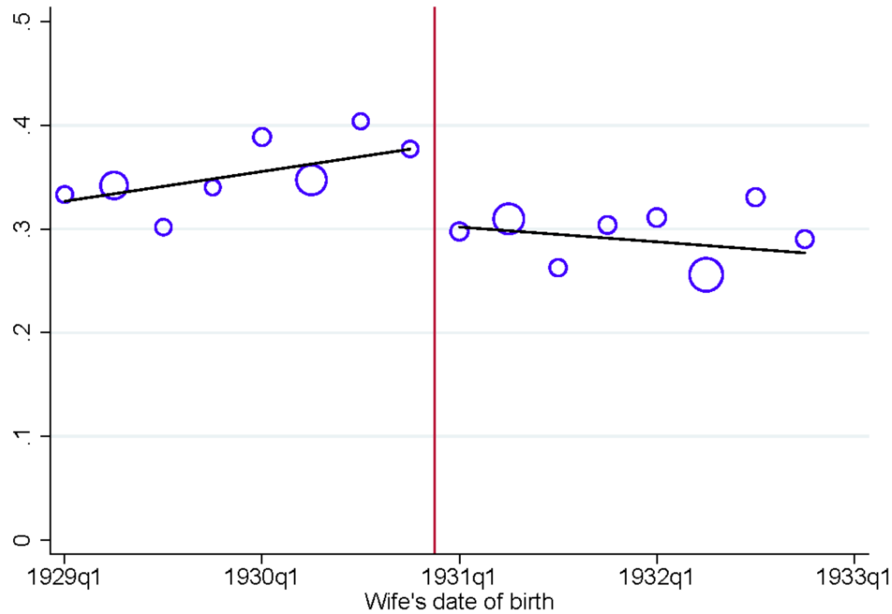
Figure 4. Effect Persists Longer for Husbands Married to 1932 Housewives



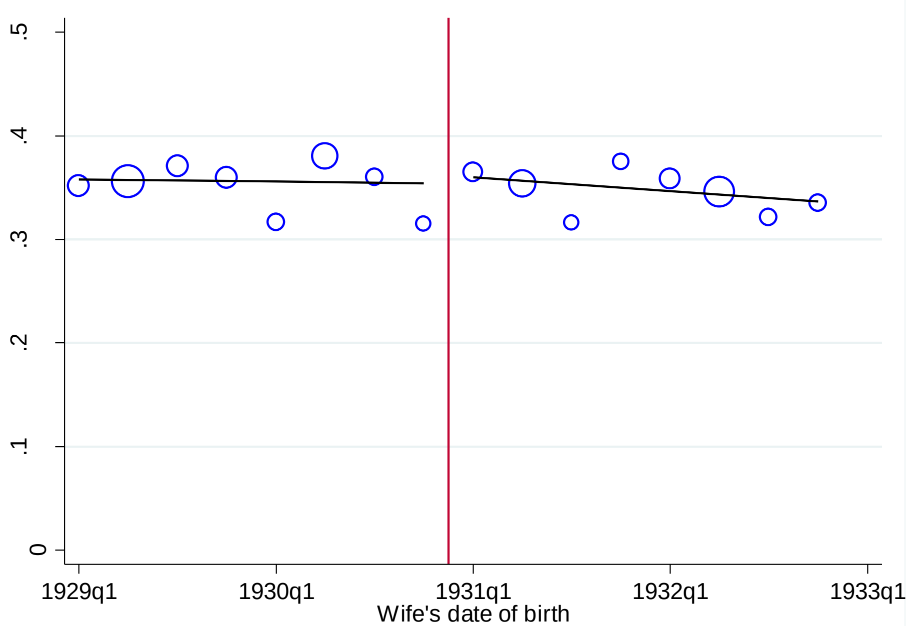
Note: Results from differences-in-differences regressions using the sample of couples that were married in 1996, conditioning on husband's employment when wife is 62. The differences-in-differences regressions were conducted for the 1932 cohort. The results reported in the text are from the DDD regression which takes the difference between the 1932 cohort and the 1930 cohort.

Figure 5. RDD, Retirement Rate by 1996, by Wife's Birth Quarter

A. Housewives

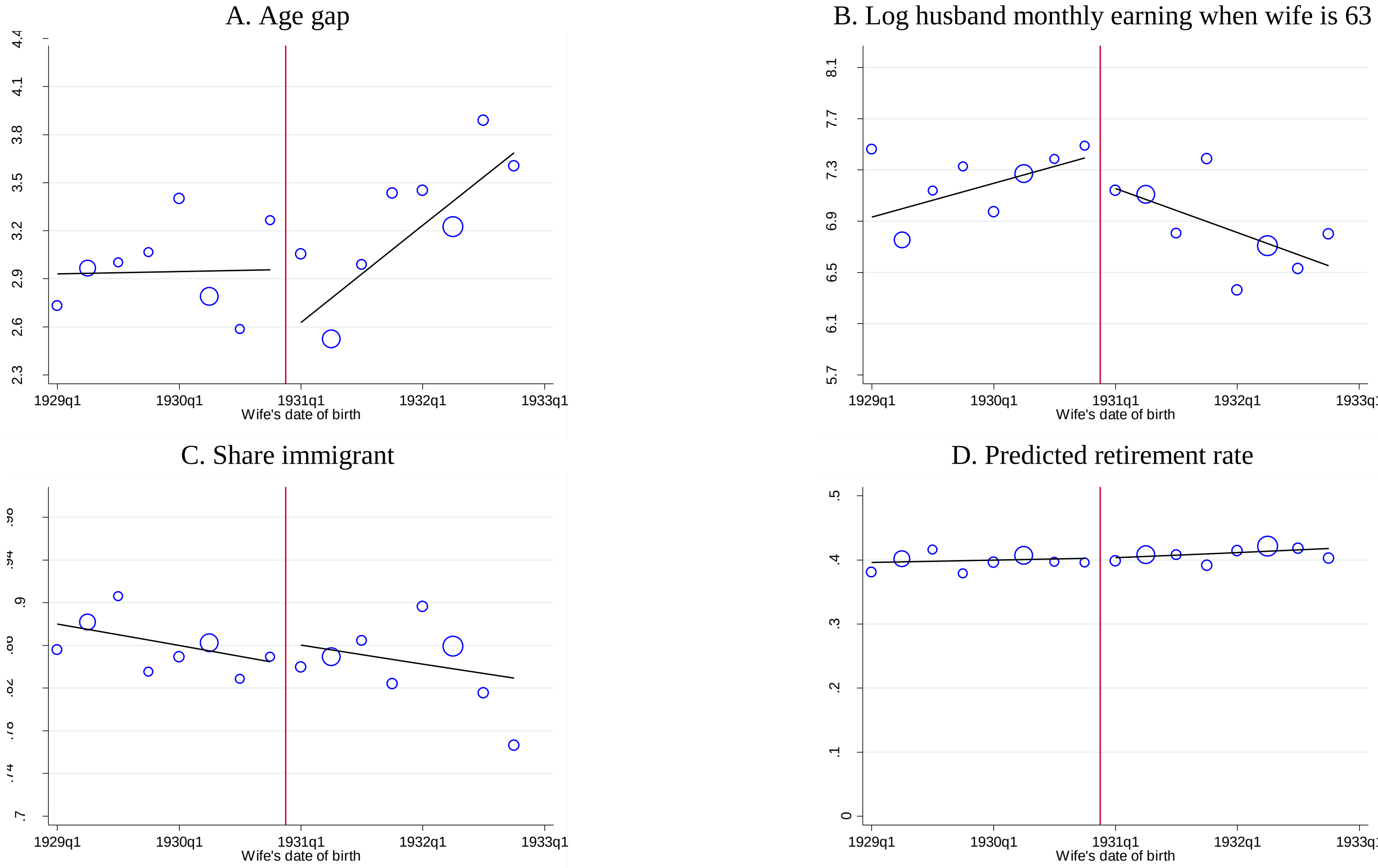


B. Non-Housewives



Note: This figure shows retirement rates by 1996 of husbands married to wives born 1929 to 1932, conditional on working in 1993. Sample in Panel A includes husbands of housewives and in Panel B, the sample includes husbands to non-Housewives who were not employed as of age 62. Circle size is proportional to the number of observations in the cell. Straight lines represent best linear fit on each side of the cutoff.

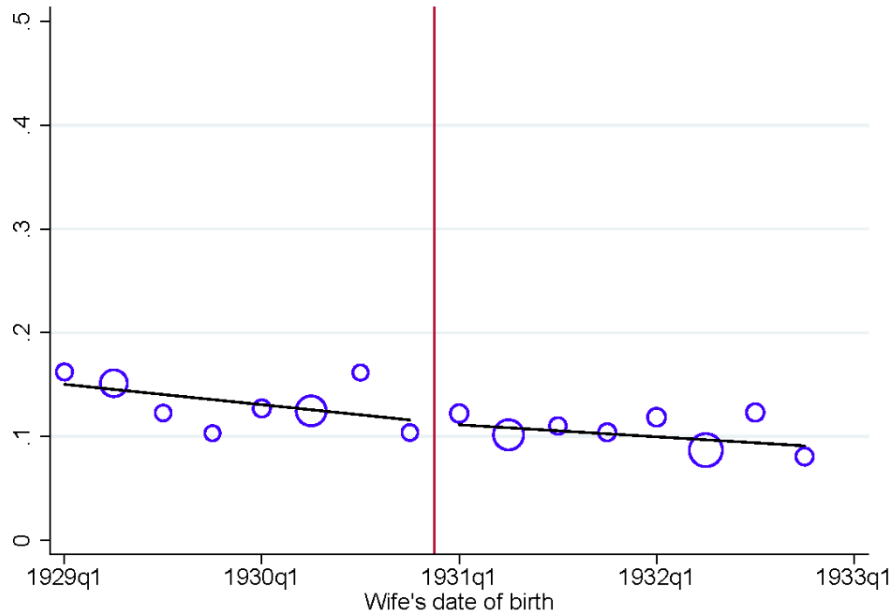
Figure 6. RDD, Selection on Observables



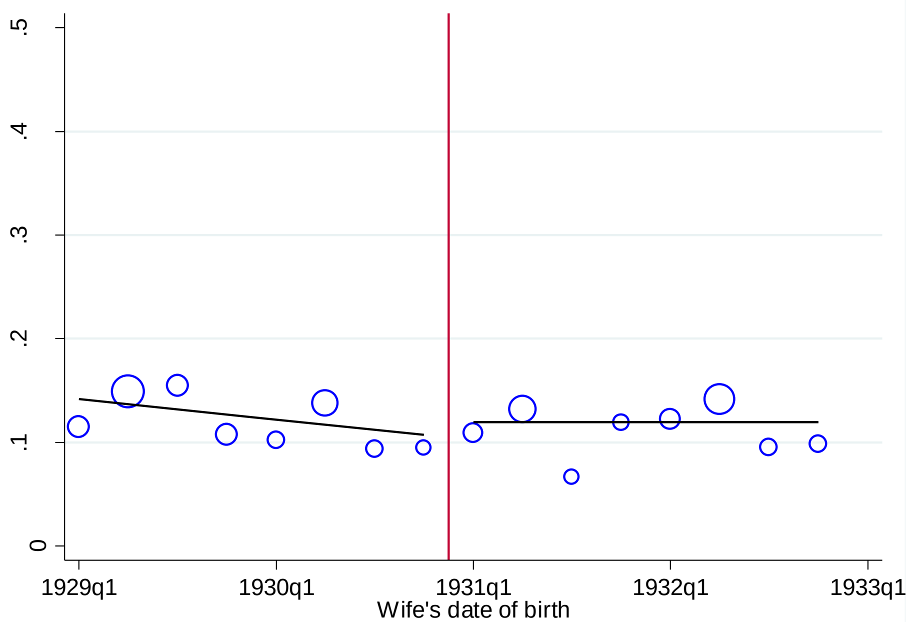
Note: Panels (A)-(C) of this figure display RDD analysis for the covariates age-gap between husband and wife, husband's log monthly earnings when wife is 63, and whether the husband is an immigrant. Panel (D) shows the predicted probabilities from a model that includes all three covariates, as well as 3[rd] degree polynomials of the first two.

Figure 7. RDD, Retirement Rate by 1994, by Wife's Birth Quarter (Placebo)

A. Housewives



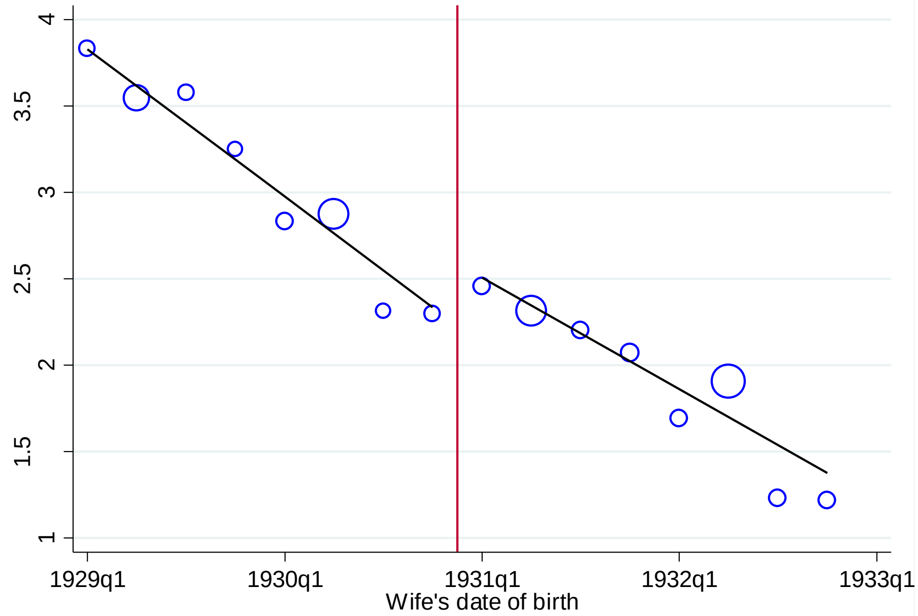
B. Non-housewives



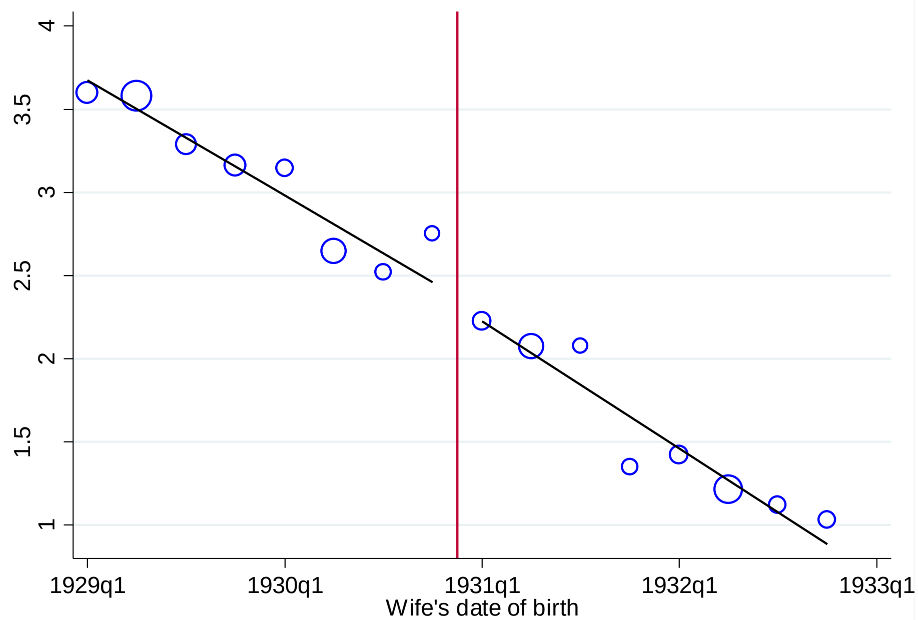
Note: This figure shows retirement rates by 1994 of husbands married to wives born 1929 to 1932, conditional on working in 1993. Sample in Panel A includes husbands of housewives and in Panel B, the sample includes husbands to non-Housewives who were not employed as of age 62. Circle size is proportional to the number of observations in the cell. Straight lines represent best linear fit on each side of the cutoff.

Figure 8. RDD, Cumulative Employment, by Wife's Birth Quarter

A. Housewives



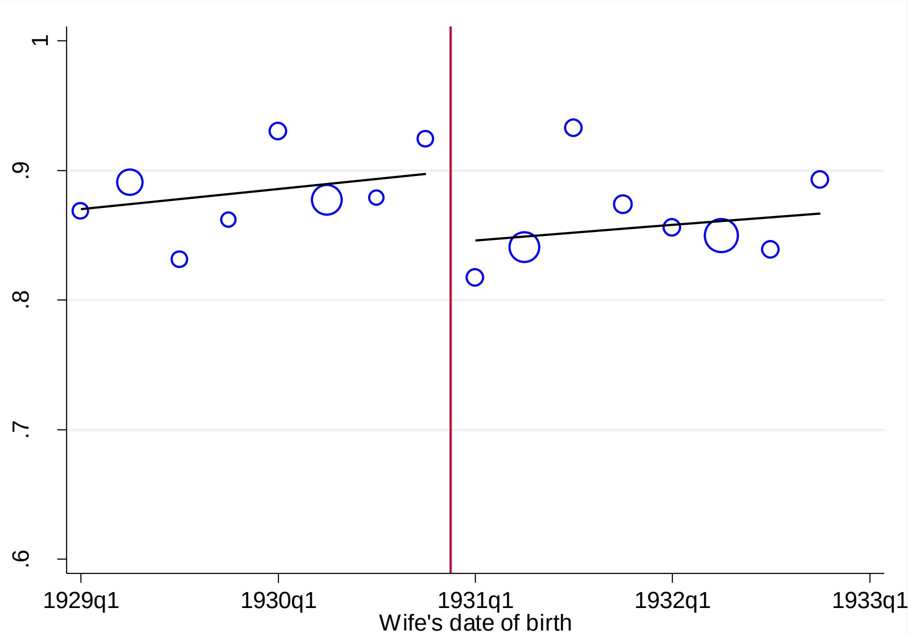
B. Non-Housewives



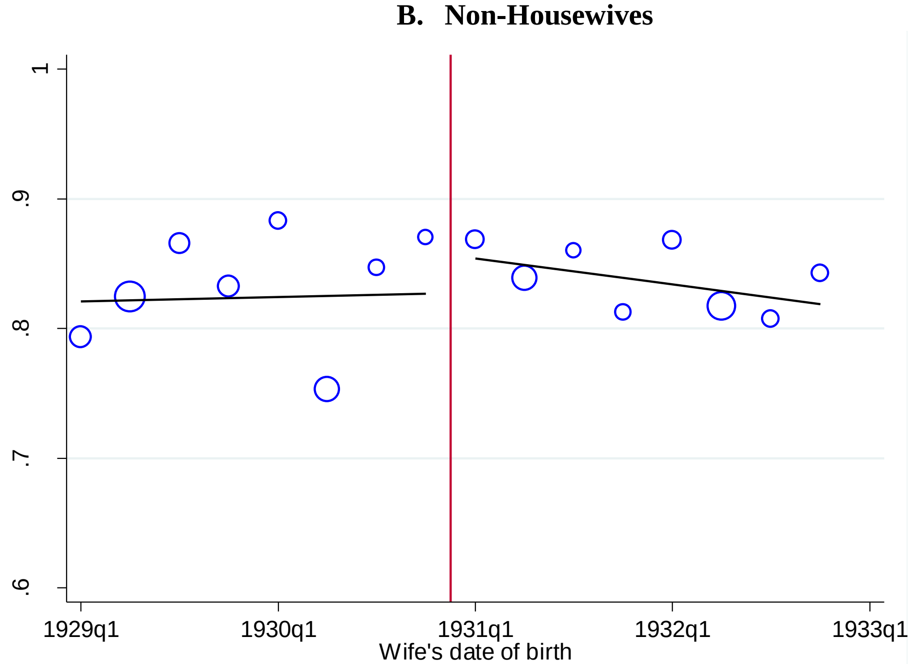
Note: RDD estimates for cumulative employment of husbands, married to wives born 1929 to 1932, after wife turns 65, conditional on working in 1993 and 1994. Sample in Panel A includes husbands of housewives and in Panel B, the sample includes husbands to non-Housewives who were not employed as of age 62. Circle size is proportional to the number of observations in the cell. Straight lines represent best linear fit on each size of the cutoff.

Figure 9. RDD, Survival Past 80, by Wife's Birth Quarter

A. Housewives



B. Non-Housewives



Note: RDD estimates for survival past age 80, of husbands married to wives born 1929 to 1932, conditional on working in 1993 and 1994. Sample in Panel A includes husbands of housewives, and in Panel B the sample includes husbands to non-Housewives who were not employed as of age 62. Circle size is proportional to the number of observations in the cell. Straight lines represent best linear fit on each size of the cutoff.

Table 1. Descriptive Statistics

	1930 Cohort		1931 Cohort		Diff in Diff
	(1)	(2)	(3)	(4)	(5)
	HW	Non-HW	HW	Non-HW	1931 vs 1930
Wife's characteristics					
Immigrant flag	0.832	0.804	0.803	0.796	-0.020
	(0.374)	(0.397)	(0.398)	(0.403)	(0.024)
Jewish	0.911	0.991	0.905	0.987	-0.001
	(0.286)	(0.095)	(0.293)	(0.115)	(0.011)
Immigration year	1951.8	1953.6	1952.1	1953.2	0.577
	(9.6)	(12)	(9.8)	(12.4)	(0.766)
husband's characteristics					
Husband's age (in 1993)	66	65.3	64.6	64.4	-0.5**
	(3.9)	(4.3)	(4.2)	(3.7)	(0.243)
Immigrant flag	0.865	0.839	0.846	0.814	0.006
	(0.342)	(0.367)	(0.362)	(0.39)	(0.022)
Jewish	0.909	0.99	0.905	0.987	-0.001
	(0.287)	(0.1)	(0.293)	(0.115)	(0.011)
Immigration year	1951.1	1952.6	1951.1	1952.7	-0.069
	(10.6)	(12.6)	(10)	(12.6)	(0.782)
Average Income when wife is 64	59,750	63,459.3	59,299.1	73,179.6	-10,171.1
	(109,435.6)	(94,132.1)	(79357.1)	(112,181.6)	(60,86.3)
Average Income when wife is 64|income>0	81,681	81,794.9	74,914.5	91,116.3	-16,087.9
	(120,767.8)	(99,609.3)	(82,382.6)	(118,473.8)	(7,376.5)
Observations	838	1,887	758	1,717	
(% HW within cohort)	(30.8)		(30.6)		

Note: Descriptive statistics for the sample of couples where wife was born between January 1930 and December 1931, conditioning on husband's employment when wife is 62. Columns (1) and (2) for wife's birth cohort of 1930, and columns (3) and (4) for birth cohort 1931. 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 cohorts). All amounts are in NIS and deflated to 1996.

Table 2. Social Security Benefits by Cohort

	1930 Cohort		1931 Cohort		Diff in Diff
	(1)	(2)	(3)	(4)	(5)
	HW	Non-HW	HW	Non-HW	1931 vs 1930
Average retirement benefits 2003-2007					
Wife	54.4	12,947.5	10,252.2	13,066.9	10,078.3***
	(710.3)	(3,046.5)	(833.4)	(3,386.9)	(173.9)
Husband	23,916.9	15,231.8	15,616.6	15,631.1	-8,699.6***
	(2,751.4)	(3,535.8)	(3,043.9)	(3,469)	(213.5)
Total	23,971.3	28,179.3	25,868.7	28,698	1,378.7***
	(2,730.3)	(5,149)	(3,033.7)	(4,542.7)	(278.7)
Observations	838	1,887	758	1,717	

Note: Calculated for the sample of couples with married wives born 1930 or 1931, conditioning on husband's employment when wife is 62. Columns (1) and (2) for wife's birth cohort of 1930, and columns (3) and (4) for wife's birth cohort 1931. 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 cohorts). All amounts are in NIS and deflated to 1996.

Table 3. DDD Results for the Effect of the Reform on Retirement by Wife's Age

	Difference in Differences Estimates				DDD 1930, 1931	
	1930		1931			
Coefficient	(1)	(2)	(3)	(4)	(5)	(6)
Wife age=63 × HW	0.01		-0.004		-0.015	
	(0.015)		(0.014)		(0.02)	
Wife age=64 × HW	0.002		-0.015		-0.017	
	(0.018)		(0.018)		(0.026)	
Wife age=65 × HW	0.019	0.012	-0.062***	-0.052***	-0.08***	-0.064***
	(0.02)	(0.014)	(0.02)	(0.014)	(0.029)	(0.019)
Wife age=66 × HW	0.022	0.015	-0.04*	-0.03*	-0.062**	-0.046*
	(0.021)	(0.017)	(0.022)	(0.018)	(0.03)	(0.024)
Wife age=67 × HW	0.01	0.003	-0.044**	-0.034*	-0.053*	-0.037
	(0.021)	(0.018)	(0.022)	(0.019)	(0.03)	(0.027)
Wife age=68 × HW	0.007	0.001	-0.03	-0.02	-0.037	-0.021
	(0.02)	(0.019)	(0.021)	(0.02)	(0.029)	(0.028)
Observations	21,800	21,800	19,800	19,800	41,600	41,600

Note: Sample comprised of married couples with wives born 1930 or 1931, conditioning on husband's employment when wife is 62. Columns (1) and (2) show differences-in-differences results for the year 1930. Columns (3) and (4) show differences-in-differences results for the year 1931. Columns (5) and (6) show DDD for the 1931 vs. 1930 cohorts. Standard errors are clustered at the individual level.

Table 4. Robustness Tests for the DDD Estimation

	DDD 1930, 1931								DDD 1932
	Nearly housewives		Less restrictive retirement definition		More restrictive retirement definition		Actual work		
Coefficient	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Wife age=63 × HW	-0.011		-0.025		-0.010		0.000		-0.016
	(0.023)		(0.019)		(0.022)		(0.028)		(0.021)
Wife age=64 × HW	-0.014		-0.023		-0.031		0.040		-0.049*
	(0.029)		(0.025)		(0.028)		(0.032)		(0.025)
Wife age=65 × HW	-0.070**	-0.058***	-0.077***	-0.053***	-0.087***	-0.066***	0.089***	0.069**	-0.068**
	(0.033)	(0.022)	(0.028)	(0.019)	(0.031)	(0.021)	(0.034)	(0.027)	(0.029)
Wife age=66 × HW	-0.043	-0.031	-0.075***	-0.051**	-0.080***	-0.060**	0.054	0.034	-0.088***
	(0.034)	(0.028)	(0.029)	(0.024)	(0.031)	(0.025)	(0.034)	(0.031)	(0.030)
Wife age=67 × HW	-0.035	-0.023	-0.058**	-0.034	-0.057*	-0.037	0.027	0.006	-0.084***
	(0.034)	(0.03)	(0.029)	(0.026)	(0.031)	(0.028)	(0.033)	(0.032)	(0.030)
Wife age=68 × HW	-0.025	-0.013	-0.031	-0.007	-0.044	-0.024	0.011	-0.010	-0.073**
	(0.033)	(0.032)	(0.029)	(0.027)	(0.030)	(0.029)	(0.031)	(0.033)	(0.029)
Observations	27,824	27,824	43,384	43,384	38,648	38,648	33,688	33,688	42,568

Note: Columns (1) and (2) show results for a control sample that only includes wives who were not employed as of age 62. Columns (3) - (8) show results for alternative definitions of employment. "Less restrictive" (columns (3) and (4)) requires only 3 months of employment per year while our main specification requires a minimum of 6 months of work to be considered employed in that year. "More restrictive" (columns (5) and (6)) requires monthly income to be above the earnings test threshold while our main specification only requires earning at least minimum wage. Columns (7) and (8) use actual work at the particular year (rather than a retirement definition). Column (9) provides the DDD estimates for the 1932 cohort. Standard errors are clustered at the individual level.

Table 5. Cumulative Effect of the Reform on Employment

Independent variable	Cumulative number of extra years worked			
Sample	DDD sample		Worked in 1994	
	(1)	(2)	(3)	(4)
HW × wife born 1931	0.355**	0.287*	0.467***	0.467***
	(0.166)	(0.165)	(0.168)	(0.167)
HH level controls	No	Yes	No	No
Observations	5,200	5,200	3,477	3,477

Note: Analysis of the impact of the Reform on husbands' cumulative years of work after wife turns 65. In columns (1) and (2), the sample includes couples with wives born in 1930 and 1931, conditioning on husband's employment when wife is 62. In columns (3) and (4) we further condition on employment in 1993 and 1994. All regressions include a constant, HW, and 1931 cohort dummies. The controls in columns (2) and (4) include dummies for Jewish, and for immigrant status of both husband and wife, as well as a 3rd degree polynomial of husband log monthly earnings when wife is 63, and a 3rd degree polynomial of the husband-wife age difference. Standard errors are calculated using Huber-White heteroscedasticity correction.

Table 6. RDD Results for the Effect of the Reform on Retirement by 1996

Polynomial degree	Panel A: Housewives				Panel B: Non-Housewives			
	One		Two		One		Two	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Wife born Jan. 1st 1931 or later	-0.074***	-0.077***	-0.072**	-0.076**	0.013	0.001	0.011	-0.001
	(0.023)	(0.022)	(0.036)	(0.038)	(0.031)	(0.032)	(0.051)	(0.057)
Household Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	2,934	2,934	2,934	2,934	4,285	4,285	4,285	4,285

Note: Analysis of retirement of husbands in 1996, conditioning on husband's employment in 1993. Polynomials are allowed to differ on two sides of the 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 when wife is 63, and a 3rd degree polynomial of the husband-wife age difference. Following Lee & Card (2008), standard errors are clustered by wife's month of birth (the running variable for wife date of birth is discrete at the monthly level).

Table 7. RDD Selection on Observables Tests

Polynomial degree	Age gap One	Age gap Two	Log husband's earnings when wife is 63 One	Log husband's earnings when wife is 63 Two	Husband immigrant One	Husband immigrant Two	Predicted values One	Predicted values Two
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Wife born Jan. 1st 1931 or later	-0.461	-0.437	-0.283	-0.338	0.020	-0.017	0.002	0.001
	(0.331)	(0.449)	(0.184)	(0.378)	(0.023)	(0.045)	(0.008)	(0.011)
Observations	2,934	2,934	2,934	2,934	2,934	2,934	2,934	2,934

Note: RDD analysis for the covariates age-gap between husband and wife, log husband's monthly earnings when wife is 63, and whether the husband is an immigrant. Columns (7) and (8) show the predicted probabilities from a model that flexibly includes all three covariates. Following Lee & Card (2008), standard errors are clustered by wife's month of birth (the running variable for wife date of birth is discrete at the monthly level).

Table 8. RDD Placebo, using Retirement in 1994

	Panel A: Housewives				Panel B: Non-Housewives			
Polynomial degree	One		Two		One		Two	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Wife born Jan. 1st 1931 or later	0.000	-0.001	-0.021	-0.024	0.022	0.010	0.015	0.003
	(0.017)	(0.017)	(0.031)	(0.031)	(0.024)	(0.022)	(0.044)	(0.044)
Household Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	2,934	2,934	2,934	2,934	4,285	4,285	4,285	4,285

Note: Analysis of retirement of husbands by 1994, conditioning on husband's employment in 1993. Polynomials are allowed to differ on two sides of the cutoff. 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 when wife is 63, and a 3rd degree polynomial of the husband-wife age difference. Following Lee & Card (2008), standard errors are clustered by wife's month of birth (the running variable for wife date of birth is discrete at the monthly level).

Table 9. The Effect of Employment on Life-Expectancy: Difference-in-Differences Estimates

Independent variable	Survival Past 80						Cumulative Income	
Specification	Reduced form		OLS		IV		Reduced form	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Employment	-0.065***	-0.051**	0.007**	0.004**	-0.139*	-0.110*	53,363**	114,278**
	(0.025)	(0.024)	(0.003)	(0.002)	(0.075)	(0.066)	(21,735)	(44,549)
Weak IV robust 90% C.I.					[-0.387,-0.046]	[-0.313,-0.024]		[41,002,187,555]
HH level controls	No	Yes	No	Yes	No	Yes	Yes	Yes
Observations	3,477	3,477	3,477	3,477	3,477	3,477	3,477	3,477

Note: Analysis of employment impact on husbands' life expectancy, conditioning on husband's employment in 1993 and 1994. Columns (1) and (2) show the effect of the Reform on survival past age 80 (without controls and with controls). Columns (3) and (4) show the OLS relation between employment and survival past 80. Columns (5), and (6) show the IV estimates for the effect of employment years on the same outcome. Columns (7) and (8) report the effect of the Reform on cumulative income in the 5 years subsequent to the Reform. 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 when wife is 63, and a 3rd degree polynomial of the husband-wife age difference. Standard errors are calculated using Huber-White heteroscedasticity correction. In columns (5) and (6) we also report in squared brackets the weak IV robust 90% confidence intervals following Chernozhukov & Hansen (2008).

Table 10. The Effect of Employment on Life-Expectancy: RDD Estimates

Specification	First Stage		Reduced form		IV	
Independent variable	Employment		Survival Past 80			
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Sample of husbands married to a housewife						
Wife born Jan. 1st	0.339**	0.504***	-0.057**	-0.039*	-0.167	-0.078
1931 or later	(0.133)	(0.136)	(0.027)	(0.023)	(0.107)	(0.048)
Weak IV robust 90% C.I.					[-0.643,-0.022]	[-0.192,-0.002]
Observations	2,051	2,051	2,051	2,051	2,051	2,051
Panel B: Sample of husbands married to a non-housewife (Placebo)						
Wife born Jan. 1st	-0.090	-0.011	0.026	0.038	N/A	N/A
1931 or later	(0.175)	(0.15)	(0.044)	(0.032)		
Household Controls	No	Yes	No	Yes		
Observations	3,040	3,040	3,040	3,040		

Note: Analysis of the effect of employment on husbands' life expectancy, conditioning on husband's employment in 1993 and 1994. Panel A includes only husbands who are married to housewives, and panel B includes only husbands married to non-housewives who were not employed as of age 62. Linear trends are allowed to differ on two sides of the cutoff. In columns (1) and (2) the dependent variable is cumulative years of work after wife turns 65. In columns (3) to (6) it is survival past age 80. 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 when wife is 63, and a 3rd degree polynomial of the husband-wife age difference. Following Lee & Card (2008), standard errors are clustered by wife's month of birth (the running variable for wife date of birth is discrete at the monthly level). In columns (5) and (6) we also report in squared brackets the weak IV robust 90% confidence intervals following Chernozhukov & Hansen (2008).

Table 11. The Effect of Employment on Life-Expectancy: Blue vs. White Collar Occupations

	Blue Collar			White Collar		
Specification	Reduced form	First stage	IV	Reduced form	First stage	IV
Independent variable	Survival	employment	Survival	Survival	employment	Survival
	(1)	(2)	(3)	(4)	(5)	(6)
HW × wife born 1931	-0.160***	0.479*		0.006	0.485*	
	(0.039)	(0.251)		(0.033)	(0.225)	
Employment			-0.334*			0.012
			(0.196)			(0.069)
Weak IV robust 90% C.I.			[-2.511,-0.146]			[-0.15,0.198]
Observations	1,404	1,404	1,404	2,073	2,073	2,073

Note: Analysis of the impact of the Reform on husbands' Life-Expectancy, conditioning on husband's employment in 1993 and 1994. Columns (1)-(3) report the analysis for blue collar workers, and columns (4)-(6) report the analysis for white collar workers (see text for blue- and white-collar definitions). Columns (1) and (4) show the effect of the Reform on survival past age 80. Columns (2) and (5) show the first stage. Columns (3) and (6) show the IV estimates for the effect of employment years on the same outcomes. Standard errors are calculated using Huber-White heteroscedasticity correction. In columns (3) and (6) we also report in squared brackets the weak IV robust 90% confidence intervals following Chernozhukov & Hansen (2008).

Bibliography

Angrist, J. D., & Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Baker, M., & Benjamin, D. (1999). How do retirement tests affect the labour supply of older men? *Journal of Public Economics*, *71*(1), 27–51.

Behaghel, L., & Blau, D. M. (2012). Framing social security reform: Behavioral responses to changes in the full retirement age. *American Economic Journal: Economic Policy*, *4*(4), 41–67.

Blau, D. M., & Goodstein, R. M. (2010). Can Social Security Explain Trends in Labor Force Participation of Older Men in the United States? *Journal of Human Resources*, *45*(2), 328–363. http://doi.org/10.3368/jhr.45.2.328

Bloemen, H., Hochguertel, S., & Zweerink, J. (2017). The causal effect of retirement on mortality: Evidence from targeted incentives to retire early. *Health Economics*.

Burtless, G., & Moffitt, R. A. (1985). The Joint Choice of Retirement Age and Postretirement Hours of Work. *Journal of Labor Economics*, *3*(2), 209–236. Retrieved from http://www.jstor.org/stable/2534983

Chernozhukov, V., & Hansen, C. (2008). The reduced form: A simple approach to inference with weak instruments. *Economics Letters*, *100*(1), 68–71.

Chetty, R., Guren, A., Manoli, D. S., & Weber, A. (2012). Does Indivisible Labor Explain the Difference between Micro and Macro Elasticities ? A Meta- Analysis of Extensive Margin Elasticities. In *NBER Macroeconomics Annual*.

Chiappori, P.-A., & Mazzocco, M. (2017). Static and intertemporal household decisions. *Journal of Economic Literature*, *55*(3), 985–1045.

Coile, C., & Gruber, J. (2007). Future social security entitlements and the retirement decision. *The Review of Economics and Statistics*, *89*(2), 234–246.

Disney, R., & Smith, S. (2002). The labour supply effect of the abolition of the earnings rule for older workers in the United Kingdom. *The Economic Journal*, *112*(478), C136--C152.

Eibich, P. (2015). Understanding the effect of retirement on health: mechanisms and heterogeneity. *Journal of Health Economics*, *43*, 1–12.

Engelhardt, G., & Kumar, A. (2014). Taxes and the labor supply of older Americans: recent evidence from the Social Security earnings test. *National Tax Journal*, *67*(2), 443–458.

Friedberg, L. (2000). The labor supply effects of the social security earnings test. *Review of Economics and Statistics*, *82*(1), 48–63.

Friedberg, L., & Webb, A. (2009). 1 New Evidence on the Labor Supply Effects of the Social Security Earnings Test. *Tax Policy and the Economy*, *23*(1), 1–36.

Gelber, A. M., Isen, A., & Song, J. (2016). *The Effect of Pension Income on Elderly Earnings: Evidence from Social Security and Full Population Data*.

Gelber, A. M., Jones, D., & Sacks, D. W. (2013). *Earnings adjustment frictions: Evidence from the Social Security earnings test*.

Gelber, A. M., Jones, D., Sacks, D. W., & Song, J. (2017). *The Employment Effects of the Social Security Earnings Test*.

Gelber, A. M., Jones, D., Sacks, D. W., Song, J., Chetty, R., Cole, J., … Slemrod, J. (2017). *Using Kinked Budget Sets to Estimate Extensive Margin Responses: Evidence from the Social Security Earnings Test* (Working Paper No. 23362).

Gerdtham, U.-G., & Ruhm, C. J. (2006). Deaths rise in good economic times: evidence from the OECD. *Economics & Human Biology*, *4*(3), 298–316.

Gruber, J. (1994). State-mandated benefits and employer-provided health insurance. *Journal of Public Economics*, *55*(3), 433–464.

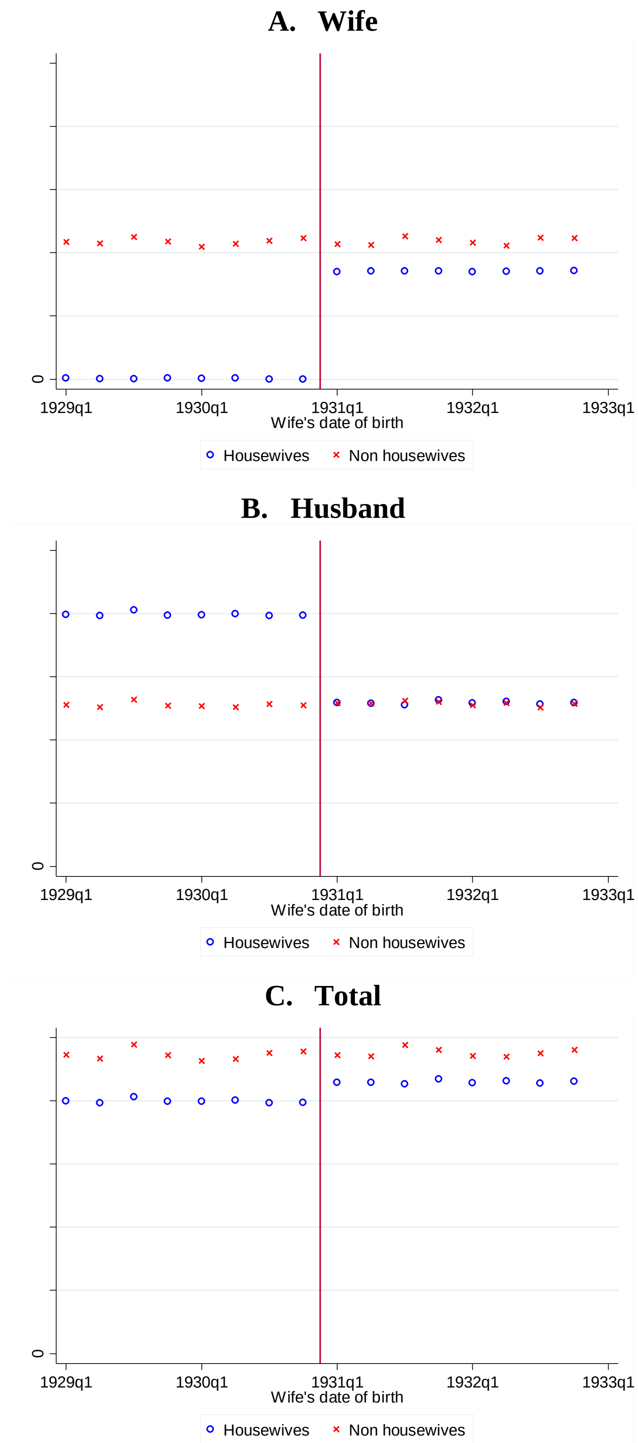
Gruber, J., & Orszag, P. (2003). Does the Social Security Earnings Test Affect Labor Supply and Benefits Receipt? *National Tax Journal*, *56*(4), 755–773. http://doi.org/10.3386/w7923

Gruber, J., & Saez, E. (2002). The elasticity of taxable income: evidence and implications. *Journal of Public Economics*, *84*(1), 1–32.
Gruber, J., & Wise, D. A. (Eds.). (2004). *Social Security Programs and Retirement around the World: Micro Estimation, vol. 2*. Chicago: The University of Chicago Press.
Haider, S. J., & Loughran, D. S. (2008). The Effect of the Social Security Earnings Test on Male Labor Supply: New Evidence from Survey and Administrative Data. *Journal of Human Resources*, *43*(1), 57–87. http://doi.org/10.1353/jhr.2008.0031
Hall, R. E., & Jones, C. I. (2007). The value of life and the rise in health spending. *The Quarterly Journal of Economics*, *122*(1), 39–72.
Hallberg, D., Johansson, P., & Josephson, M. (2014). *Early retirement and post retirement health*.
Insler, M. (2014). The health consequences of retirement. *Journal of Human Resources*, *49*(1), 195–233.
Krueger, A. B., & Pischke, J.-S. (1992). The Effect of Social Security on Labor Supply: A Cohort Analysis of the Notch Generation. *Journal of Labor Economics*, *10*(4), 412–437.
Kuhn, A., Wuellrich, J.-P., & Zweimüller, J. (2010). Fatal attraction? Access to early retirement and mortality.
Lee, D. S., & Card, D. (2008). Regression discontinuity inference with specification error. *Journal of Econometrics*, *142*(2), 655–674.
Lee, D. S., & Lemieuxa, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, *48*(2), 281–355.
Manoli, D. S., & Weber, A. (2016). *The Effects of the Early Retirement Age on Retirement Decisions* (Working Paper Series). Retrieved from http://www.nber.org/papers/w22561
Manoli, D., & Weber, A. (n.d.). Nonparametric Evidence on the Effects of Financial Incentives on Retirement Decisions. *American Economic Journal: Economic Policy*.
Mastrobuoni, G. (2009). Labor supply effects of the recent social security benefit cuts: Empirical estimates using cohort discontinuities. *Journal of Public Economics*, *93*(11–12), 1224–1233. http://doi.org/http://dx.doi.org/10.1016/j.jpubeco.2009.07.009
McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, *142*(2), 698–714.
Murphy, K. M., & Topel, R. H. (2006). The value of health and longevity. *Journal of Political Economy*, *114*(5), 871–904.
OECD. (2015). *Pensions at a Glance*. *Pensions* (Vol. 15). http://doi.org/10.1787/9789264107007-en
Snyder, S. E., & Evans, W. N. (2006). The effect of income on mortality: evidence from the social security notch. *The Review of Economics and Statistics*, *88*(3), 482–495.
Song, J. G., & Manchester, J. (2007). New evidence on earnings and benefit claims following changes in the retirement earnings test in 2000. *Journal of Public Economics*, *91*(3), 669–700.
Stock, J. H., & Wise, D. A. (1990). Pensions, the Option Value of Work, and Retirement. *Econometrica*, *58*(5), 1151–1180. Retrieved from http://www.jstor.org/stable/2938304
Stock, J. H., Wright, J. H., & Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics*, *20*(4), 518–529.
Sullivan, D., & Von Wachter, T. (2009). Job displacement and mortality: An analysis using administrative data. *The Quarterly Journal of Economics*, *124*(3), 1265–1306.

Hebrew Publications

National Insurance Institute of Israel (NII), August 2016. “Monthly Statistical Report.

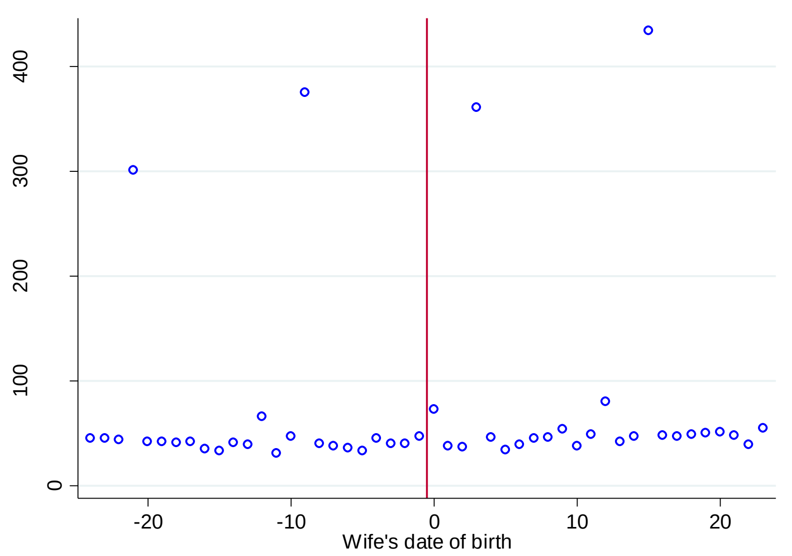
Figure A1. Change in benefits around the date of birth cutoff



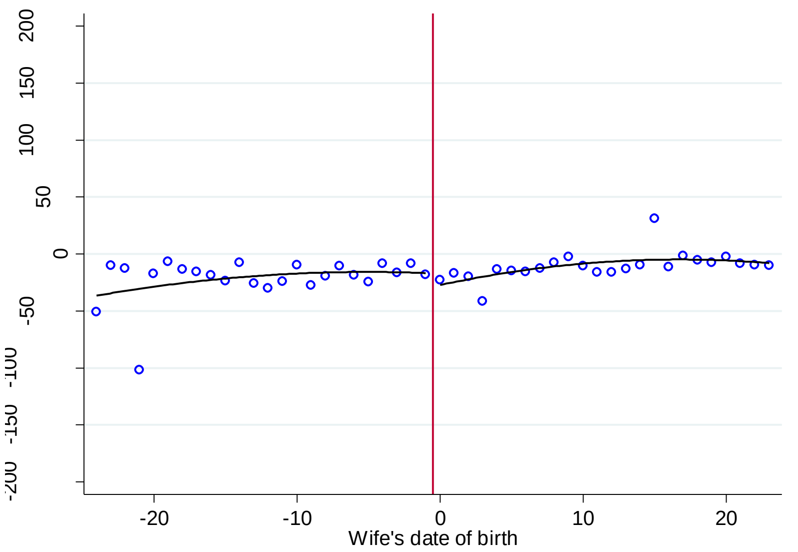
Note. Calculated for the sample of married couples with wives born 1929 to 1932, conditioning on husband's employment in 1992. Each symbol represents average retirement benefits in 2003-2007 by quarter of birth of the wife. All amounts are in NIS and deflated to 1996.

Figure A2. Density of observations around the RDD cutoff

A. Raw data



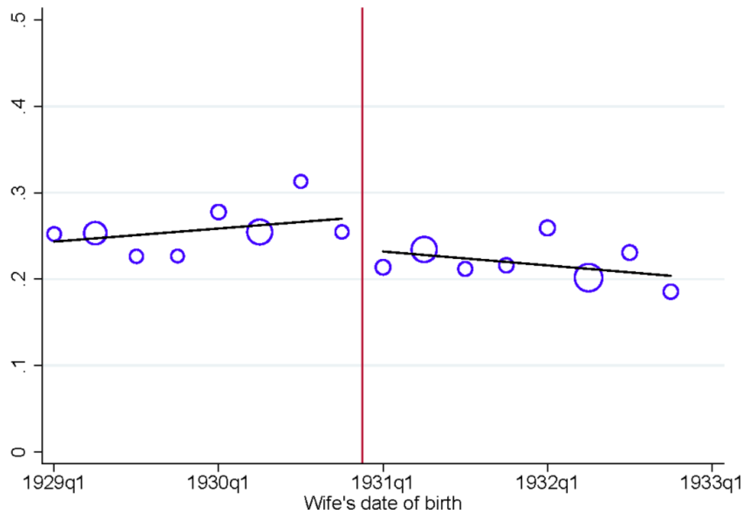
B. Controlling for 12 monthly fixed effects



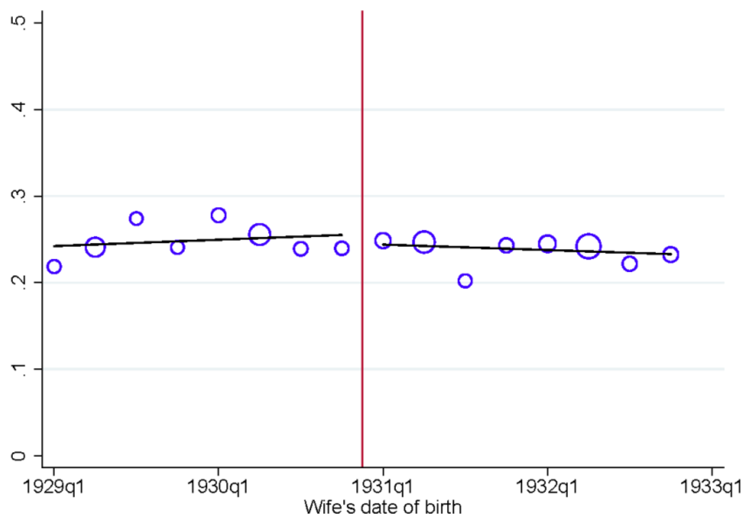
Note. This figure displays the count of observations by monthly bins. Panel A displays the raw data, while panel B displays the residuals from a regression that controls for 12 month fixed effects. Sample is husbands married to housewives born 1929 to 1932.

Figure A3. RDD, Retirement Rate in 1995, by Wife's Birth Quarter, placebo

A. Housewives



B. Non-Housewives



Note. This figure shows retirement rates by 1995 of husbands married to wives born 1929 to 1932, conditional on working in 1993. Sample in Panel A includes husbands of housewives and in Panel B, the sample includes husbands to non-Housewives who were not employed as of age 62. 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. RDD placebo, using retirement in 1995

	Panel A: Housewives				Panel B: Non-Housewives			
Polynomial degree	One		Two		One		Two	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Wife born Jan. 1st 1931 or later	-0.037	-0.034	-0.059	-0.056	-0.008	-0.022	0.012	0.003
	(0.023)	(0.024)	(0.044)	(0.049)	(0.020)	(0.020)	(0.033)	(0.031)
Household Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	2,934	2,934	2,934	2,934	6,801	6,801	6,801	6,801

Note: Analysis of retirement of husbands in 1996, conditioning on husband's employment in 1993. Polynomials are allowed to differ on two sides of the cutoff. Household controls includes dummies for Jewish, and for immigrant status of both husband and wife, as well as a 3rd degree polynomial in husband log monthly earnings when wife is 63, and a 3rd degree polynomial of the husband-wife age difference. Following Lee & Card (2008), standard errors are clustered by wife's month of birth (the running variable for wife date of birth is discrete at the monthly level).

Table A3: The Effect of Employment on Long-Term Care Eligibility: Difference-in-Differences Estimates

Independent variable	No LTC Eligibility					
Specification	Reduced form		OLS		IV	
	(1)	(2)	(3)	(4)	(5)	(6)
Employment	-0.074**	-0.051*	0.016***	0.011***	-0.159*	-0.109
	(0.032)	(0.031)	(0.003)	(0.003)	(0.093)	(0.078)
Weak IV robust 90% C.I.					[-0.463,-0.04]	[-0.341,-0.002]
HH level controls	No	Yes	No	Yes	No	Yes
Observations	3,477	3,477	3,477	3,477	3,477	3,477

Note: Analysis of the effect of employment on husbands' health, conditioning on husband's employment in 1993 and 1994. Columns (1) and (2) show the effect of the Reform on Long-Term-Care (LTC) eligibility (without controls and with controls). Columns (3), and (4) show the OLS relation between employment and LTC eligibility. Columns (5), and (6) show the IV estimates for the effect of employment years on the same outcome. Controls includes dummies for Jewish, and for immigrant status for both husband and wife, as well as a 3rd degree polynomial of husband log monthly earnings when wife is 63, and a 3rd degree polynomial of the husband-wife age difference. Standard errors are calculated using Huber-White heteroscedasticity correction. In columns (5) and (6) we also report in squared brackets the weak IV robust 90% confidence intervals following Chernozhukov & Hansen (2008)

Table A4: The Effect of Employment on Long-Term Care Eligibility: RDD Estimates

Specification	First Stage		Reduced form		IV	
Independent variable	Employment		No LTC Eligibility			
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Sample of husbands married to a housewife						
Wife born Jan. 1st	0.339**	0.504***	-0.109*	-0.070*	-0.321	-0.139*
1931 or later	(0.133)	(0.136)	(0.058)	(0.04)	(0.232)	(0.08)
Weak IV robust 90% C.I.					[-1.305,0.018]	[-0.293,0.028]
Observations	2,051	2,051	2,051	2,051	2,051	2,051
Panel B: Sample of husbands married to a non-housewife (Placebo)						
Wife born Jan. 1st	-0.090	-0.011	-0.008	0.011	N/A	N/A
1931 or later	(0.175)	(0.15)	(0.072)	(0.05)		
Household Controls	No	Yes	No	Yes		
Observations	3,040	3,040	3,040	3,040		

Note: Analysis of the effect of employment on husbands' health, conditioning on husband's employment in 1993 and 1994. Panel A includes only husbands who are married to housewives, and panel B only husbands married to non-housewives. Linear trends are allowed to differ on two sides of the cutoff. In columns (1) and (2) the dependent variable is cumulative years of work after wife turns 65. In columns (3) to (6) it is no Long-Term-Care (LTC) eligibility. 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 when wife is 63, and a 3rd degree polynomial of the husband-wife age difference. Following Lee and Card (2008), standard errors are clustered by wife's month of birth (the running variable for wife date of birth is discrete at the monthly level). In columns (5) and (6) we also report in squared brackets the weak IV robust 90% confidence intervals following Chernozhukov and Hansen (2008).