Retirement Lock and Prescription Drug Insurance: Evidence from Medicare Part D, Gal Wettstein, Online Appendix

# A Data Appendix

## Defining Treatment and Control Groups Based on employer-sponsored RHI

The RAND version M of the HRS contains information regarding whether or not individuals are offered RHI from their employer. Questions on whether a respondent has RHI of any sort (limited to age 65 or not) are asked from wave 3 (1996 and later). For waves 5 and on (interviews conducted in 2000 and later), this is taken from a question asking individuals under age 65 whether they have employer-sponsored plans that offer RHI (their own or their spouse's). If they reply that they do, a follow-up question asks if this coverage would extend past age 65.

The main variable I use to determine coverage is based on these questions from wave 5 and later, and provides a summary of the information regarding all employer-sponsored plans the individual reports (up to three different plans). These questions are asked only of individuals who have ESI while working and are under age 65. The possible values this variable takes are: "not covered in retirement"; "covered [in retirement] just to age 65"; "covered [in retirement] to age 65, don't know over"; "covered [in retirement] to and over age 65"; and a number of possible missing values: "age is 65 or older"; "don't know"; "source missing, question"; "missing"; "no respondent employer provided insurance"; "refused to answer"; "question not asked"; and "spouse is non response".

To be included in the sample for the main analysis an observation must be either in the treatment group (covered in retirement only until age 65) or in the control group (covered in retirement to and over age 65). If the observation cannot be definitively allocated into one of these groups it will not be included in the sample (e.g., if the respondent gave an answer of "covered to age 65, don't know over", but see strategies below for inferring insurance status in the

absence of clear answers).

The questions about RHI are only asked of respondents below age 65. To allocate observations with age above 65 to the treatment and control groups I employ two complementary strategies: 1) use of lagged values of the same individual from before age 65 to ascertain what manner of RHI, if any, she would have after age 65; 2) inference from current employment and insurance statuses (after age 65) what manner of RHI, if any, she has after age 65. I will now detail for each of these approaches what information is required and what assumptions are made.

**1) Use of Lagged Values from before Age 65.** This approach is straightforward. If a respondent is interviewed both under and over age 65 at different survey waves then from answers given regarding RHI offered by employers when asked at waves when she was younger than 65, it can be inferred what retiree insurance she will have when over 65. For example, a respondent replying at age 64 that she will be covered by her employer plan in retirement only until age 65 will be allocated to the treatment group (of individuals with retiree insurance only until age 65) in all waves, including later waves when she is not asked this question because she is 65 or older.

The assumption made in this approach is that employers did not change the terms of their retiree insurance plans for employees or retirees already covered by those plans, when they were over 65 years of age. This assumption is not completely innocuous: for example, employers who face financial distress such as bankruptcy may change the terms of their retiree health plans. It is assumed that such cases are relatively rare.

Misallocation of observations to the treatment and control groups due to this assumption should generally operate in the direction of allocation of a treated individual to the control group: an individual was promised RHI for life but at some point the employer decided not to honor that promise and the individual becomes de facto only covered until age 65 (if the abrogation of the promise occurs before age 65 then respondents' answers to the HRS question regarding retiree health coverage should reflect this and no error in allocation

would be made). In this case the identification concerns raised by violations of this assumption pertain to the assumption that the control group is, in fact, untreated. To the extent that no significant effect on the control group was found, this should not be of grave concern. Moreover, if any bias is implied by this regarding the effect of Part D eligibility on the treated group it is to bias that effect towards 0.

This concern is further allayed by use of the alternative control group in Appendix B.2. The finding of a significant effect on the treated group relative to this alternative control and the null effect of Part D eligibility on individuals with no ESI whatsoever provides further evidence of the mechanism of retirement lock irrespective of the assumption made here.

**2) Inference of Experimental Group from post Age 65 Employment and Insurance Statuses.** This approach is a little more complex, though the idea is simple: consider individuals who reported that they have retiree insurance but do not know if it is limited to age 65 or not, or who have missing values for the question on retiree insurance for any reason. If over time they retire it can be inferred whether or not their RHI extends past age 65 by observing whether they are covered by an employer plan when they retire and are over age 65.

This is especially useful for individuals who were 61-64 in 1996 or 63-64 in 1998: such individuals were asked if they had retiree insurance but were not asked if it was limited to age 65 or not. In future waves with more detailed questions they were not probed further because they were already over age 65, and thus not asked questions regarding retiree insurance. There were 684 such individuals in 1996, and 295 such individuals in 1998.

The main difficulty in putting this approach into practice is that it will not reveal the retiree insurance status of respondents over age 65 who still work. This difficulty can be partially circumvented by observing the same individual over time until she is retired. If her employer plan continues to cover her in retirement then it can be inferred that she was covered by a plan that would cover her in retirement even when she was employed. As she is over 65, this

places her unambiguously in the control group of individuals with RHI past age 65.

If, however, the respondent is observed retired and over age 65 without insurance, then it is not immediately clear if she would have had retiree insurance only until age 65 (and thus belong in the treatment group) or whether she had no retiree insurance at all (and thus should not be included in the sample). To deal with this ambiguity we must refer again to lagged responses of the same individual from before age 65. If at those ages the individual at some point replied she had RHI then she can be included in the treatment group. Otherwise she is assumed not to have had retiree insurance at all, and thus is excluded from the sample.

Concretely, the approach I take is to consider for each respondent the first period after age 65 in which she is retired and check whether or not she has retiree insurance at that point. If she does, I assign her to the control group in all previous periods as well. If she does not I check whether before age 65 she claimed she would have some form of retiree insurance should she retire. If she did she is assigned to the treatment group in all periods. If she did not she is excluded from the sample. In any case, the order of preference for assigning observations to treatment and control groups is: 1) explicit respondent answers when available; 2) lagged values from before age 65 when available, if no explicit response to RHI questions is available; and 3) inference from leading values if an observation is still unassigned based on (1) and (2).

This approach substantially increases the size of the sample, salvaging many observations with missing values or unknown age limits for retiree insurance. For example, in the baseline specification it increases the number of individuals observed from 4,934 using just strategy (1) to 6,516 using both (see column 3 of table 6 and column 2 of table 1). However, it implies some selection of workers out of the sample. Specifically, individuals who continue working throughout the period they are observed in the HRS cannot reveal their retiree insurance status in this way.

It is not clear that this selection should be different across the treatment and control groups and its overall magnitude is small as the vast majority of

individuals do, in fact, retire by the later ages considered (I check for retirement among individuals as old as 75-76 in 2010, covering to these ages even the youngest individuals in 1996-1998 who would not be asked about their post 65 retiree insurance status). Nevertheless, in order to be sure that this selection is not biasing the results, the next section replicates the main results of the paper using a sample where treatment and control groups are constructed only based on strategy (1). This leads to a smaller sample and thus larger standard errors but the qualitative results remain quite robust.

There are some inconsistencies in the resulting classification. Primarily, there are observations classified as in the control group who have no employer coverage after age 65 in retirement on the one hand; and observations classified as treatment who have employer coverage after age 65 despite being retired. The latter could be individuals who have RHI *only after* age 65. While rare, about 5 percent of employer retiree plans were structured in this way (Kaiser Family Foundation (2017)), but the data cannot identify them definitively. Overall in the sample used in this paper that is over age 65, 1.2 percent of person-year observations fall in the group with the first inconsistency, and 5.2 percent are in the second. I do not "correct" these inconsistencies, to let the directly asked survey responses have precedence (after all, individuals' actual conditions in retirement might differ from their expectations about retirement; for example, if a previous employer goes bankrupt and no longer offers insurance to retirees). However, in tests of the main analysis with different treatment of such inconsistencies (e.g., excluding those who display them) the results are not qualitatively different. On this point, it is also worth noting that randomly misclassifying individuals between the treatment and control should shrink the observed differences between the groups, and bias estimation results to 0.

A constraint that the HRS survey, even coupled with these two procedures, places on the sample is that in the early years of the sample (2000 and 2002) individuals of advanced age cannot have their retiree insurance status identified if they do not actually have retiree insurance after age 65: for example, a 69 year-old in 2000 would not have been asked regarding RHI in wave 5 (2000)

or waves 3 and 4 (1996-1998) because she would have been over age 65 in all those survey waves, and RHI was not inquired about in previous waves. If she is retired and insured by an employer past age 65 she can be placed in the control group- but if she is uninsured it is impossible to tell whether it is because her insurance was limited to age 65 or because she had no retiree insurance at all. It is for this reason that the sample for the entire analysis is based on ages 55-68: respondents over 68 who should properly belong in the treatment group would be unrepresented in most of the pre Part D period.

## Definition of Alternative Control Group for Appendix B.2 and Descriptive Statistics

Appendix B.2 examines the robustness of the central results to use of a different control group. The treatment group in all analyses is the same (individuals with RHI only until age 65); however while the main analysis is done with a control group of individuals who have RHI from their employers past age 65, Appendix B.2 uses a control group of individuals who have no ESI whatsoever. Construction of this latter group is straightforward: it includes only individuals who have no ESI. This includes insurance from a current or previous employer or union, of one's own or of one's spouse. All respondents are asked this question and there are few missing values (an average of 210 missing values out of about 20,000 observations each wave).

Table 5 gives descriptive statistics for the pre treatment sample (individuals aged 55-64, in the years 2000-2004) of the three experimental groups: the treatment group (individuals with retiree insurance only until age 65), the main control group (individuals with retiree insurance past age 65), and the alternative control group (individuals with no ESI).[41] Column 1 provides statistics for demographic variables, prescription drug insurance and utilization, and the main outcome variables of full- and part-time work and labor earnings for the treatment group, as well as the number of individuals included

[41] Except for statistics on age and number of unique individuals, which are not limited to observations of less than 64 years of age, before 2006 but rather encompass the entire sample.

in the group; columns 2 and 3 do the same for the main and alternative control groups, respectively.

There are about 4000 unique individuals in each of the treatment and main control groups, and the two groups are similar in their demographic characteristics. About 50 percent are women, 15 percent are African American, they have a mean age of 62 and between 13 and 14 years of education on average. None of these differences is statistically significant aside from education, and even there the magnitude of the difference is very small. Similarly, household assets, part-time work, and household income are not significantly different between the groups at conventional levels.

Likewise, the groups are similar in their coverage for prescription drugs, which is almost universal (as expected, since both groups before age 65 have RHI which almost invariably also includes drug coverage), in public coverage of prescription drugs, and in out-of-pocket drug spending. While some of the differences on these measures are significant (drug coverage is lower, and out of pocket spending higher in the treatment group), the magnitudes of the differences are small. It is worth noting here that to the extent that the treatment group includes individuals who, despite having ESI and RHI, did *not* have drug coverage before Part D eligibility, these individuals would not have been job-locked with respect to drugs and thus their inclusion in the treatment group would serve to bias the estimate of Part D's effect towards 0. In practice, however, drug coverage rates were nearly universal in both the treatment and control groups, and the differences between them (of 1.5 percentage points in coverage, or 0.1 standard deviations) are not likely to cause substantial bias. The treatment and control groups do differ significantly in their full-time work rates; however as discussed above it is parallel trends, rather than identical levels, which test the identifying assumption of the triple-differences estimation strategy.

The alternative control group is less similar to the treatment group in both demographic characteristics and in levels of the outcome variables than the main control group is. In particular, all the observed measures are significantly different between the groups besides out-of-pocket drug spending and part-

time work. As expected, the alternative control group is less educated, less wealthy, has lower income and has a much higher share of women than the treatment group. They are also less likely to have prescription drug insurance coverage, and more likely to have public prescription drug insurance coverage before introduction of Part D. Furthermore, their rate of full-time work is lower pre treatment, as are their average annual labor earnings. However, here too, the identifying assumption is one of parallel trends in the absence of treatment rather than identical levels. This assumption can be assessed by examining the pre trends in figure 6. The very different baseline characteristics of the alternative control group and the treatment group are a motivating factor in the choice of individuals with RHI past age 65 as the main control group for the analysis, rather than individuals with no ESI at all.

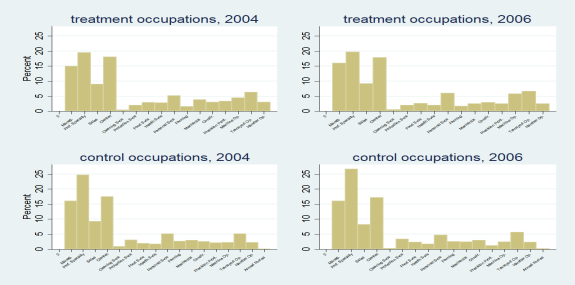
The treatment and main control groups are also similar in their occupations and industries. Further, the distribution across occupations and industries does not show substantial changes in reaction to the introduction of Part D. These distributions for each experimental group, in years 2004 and 2006, are presented in figure 4 (occupations) and figure 5 (industries). Both treatment and control groups are predominantly in managerial, clerical, and professional occupations (together accounting for over half of each group), with sales accounting for an additional 10 percent of each group. The remaining 30-40 percent are roughly uniformly distributed across a variety of occupations. With respect to industry, both treatment and control groups are most likely to work in professional services (between 30 percent and 40 percent), with public administration (between 5 percent and 12 percent), manufacturing (around 15 percent) and retail (about 15 percent) making up the bulk of the remainder.

Table 5: Descriptive Statistics by Experimental Group at Ages 55-64, Years 2000-2004

| | (1) | (2) | (3) |
|---|---|---|---|
| | RHI only up to age 65 | Main Control | Alternative Control |
| Share Women | 0.513 | 0.495 | 0.639 |
| | (0.5) | (0.5) | (0.48) |
| Share Black | 0.144 | 0.154 | 0.22 |
| | (0.352) | (0.361) | (0.414) |
| Age | 62.4 | 62.28 | 62.8 |
| | (3.86) | (3.85) | (3.79) |
| Years of Education | 13.07 | 13.51 | 11.11 |
| | (2.67) | (2.63) | (3.46) |
| non Housing Household Assets | 350,119 | 405,274 | 214,341 |
| | (1,293,232) | (2,039,182) | (874,601) |
| Share with Prescription Drug Insurance | 0.969 | 0.985 | 0.591 |
| | (0.173) | (0.122) | (0.492) |
| Share with Public Prescription Drug Insurance | 0.003 | 0.005 | 0.204 |
| | (0.051) | (0.071) | (0.403) |
| Out-of-Pocket Spending on Drugs/Month | 71.79 | 55.48 | 94.79 |
| | (258) | (200) | (1049) |
| Share Working Full-Time | 0.554 | 0.4 | 0.186 |
| | (0.5) | (0.49) | (0.389) |
| Share Working Part-Time | 0.143 | 0.158 | 0.161 |
| | (0.35) | (0.364) | (0.368) |
| Annual Labor Earnings | 32,930 | 28,104 | 6,374 |
| | (31,404) | (32,931) | (14,945) |
| Number of Individuals | 3,717 | 4,048 | 5,773 |

*Notes:* This table presents descriptive statistics for the three experimental groups in the analysis: column 1 shows the treatment group of individuals with retiree health insurance (RHI) only up to age 65; column 2 shows the main control group of individuals with retiree insurance past age 65; column 3 shows the alternative control group of individuals with no ESI. The sample is restricted to ages 55-64 (except for the statistics on age and number of individuals) and years 2000-2004. For the row of age the sample is ages 55-68, years 2000-2004. All monetary values are inflated to 2010 dollars using the consumer price index. Annual labor earnings are top-coded at $100,000. The number of individuals is the number of unique individuals included in the baseline specification of Equation (1) in the complete sample, within each experimental group; i.e., all individuals aged 55-68, in the years 2000-2010 in each of the experimental groups. Note that there are individuals who may appear in more than one group at different survey waves (e.g., if they move from a job which does not offer any ESI to one which offers retiree insurance). Each row besides the last presents the mean of the variable listed in that row for the three experimental groups, with standard deviations in parentheses.

Figure 4: Distribution of Occupations for Treatment and Control Groups, in 2004 and 2006




*Notes:* This figure represents the share of the relevant population in each of the occupations listed along the x-axis. The relevant population in each panel is: treatment group in 2004, treatment group in 2006, control group in 2004 and control group in 2006 for the upper left, upper right, lower left and lower right panels, respectively. Individuals who are no longer working are excluded.

Figure 5: Distribution of Industries for Treatment and Control Groups, in 2004 and 2006


*Notes:* This figure represents the share of the relevant population in each of the industries listed along the x-axis. The relevant population in each panel is: treatment group in 2004, treatment group in 2006, control group in 2004 and control group in 2006 for the upper left, upper right, lower left and lower right panels, respectively. Individuals who are no longer working are excluded.

## Main Results with Experimental Groups Defined only by Lagged Values from before Age 65, or only by Strictly pre 2004 Lagged Values

As discussed above, construction of the sample requires knowledge of the ESI status of respondents after retirement. If they are insured in retirement but only until age 65 they are in the treatment group; if they are insured in retirement past age 65 they are in the control group; if they are neither then they are not included in the sample (except for the sample in Appendix B.2).

While the HRS contains all the necessary information for construction of

these groups for individuals below age 65, at age 65 and over questions regarding RHI are not asked. It is therefore necessary to infer retiree insurance status for observations aged 65 or over. This is done by two strategies detailed above. The first uses answers given by individuals interviewed when they were younger than 65 to infer their retiree insurance status after age 65. The second fills in the gaps due to missing or ambiguous answers by inferring from the observed retiree insurance status after age 65 for a given individual what that individual's employer offered retirees.

This second method admits into the sample individuals who are observed retired and over age 65 at some point during the sample period. For them it is possible to see if they are insured in retirement past age 65, and thus infer that when they were not retired they were plausibly nevertheless offered RHI past age 65 should they retire. However, this method cannot admit into the sample individuals who are never observed retired, and thus selects out of the sample by construction some individuals who keep working throughout the sample period. This section aims to demonstrate robustness of the main results to using a sample constructed using only the first method, which does not involve possible selection on work status.

Table 6 replicates the main results of the paper using this smaller but less potentially selected sample. The table shows the effect of Part D eligibility on full-time and part-time work. The estimation method is the same as that described in Section III. Columns (1) and (5) for full-time work and part-time work, respectively, show differences-in-differences results using only the treatment group. Columns (2), (3), (6), and (7) estimate the effect of Part D eligibility using triple-differences as in Equation (1). Columns (2) and (6) do this without individual fixed effects (instead including richer demographic controls), for full and part-time work respectively. Columns (3) and (7) estimate the baseline specification.

An additional robustness check in this spirit is to identify treatment status based on RHI from years before Part D (and using only method 1, to avoid inferring treatment status from insurance status post Part D). In this exercise I use the values from 2002 or earlier. This has the advantage of not only tak-

ing treatment status as fixed before Part D's implementation in 2006, but also before passage of the Medicare Prescription Drug, Improvement and Modernization Act of 2003. Passage of the law was far from assured, since the law passed by a very narrow majority and through a contentious process; thus this determination of treatment status helps to ensure that individuals were not selecting their employer and retiree benefits with foresight of Part D's future existence. Columns (4) and (8) show estimates of the effect of Part D using treatment and control groups defined in this way, with results qualitatively similar to those in the main analysis. For full-time work, the estimates indicate Part D led to a reduction of 6 percentage points for the treatment group (although this is not significant due to the large associated standard errors).

Table 6: Main Results Using only Lagged Values to Determine Experimental Group

| Dependent Variable: | Full-Time Work | | | | Part-Time Work | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Post65*Post2006*RHI only up to age 65 | - | -0.0482 | -0.0893 | -0.0595 | - | 0.1202 | 0.0949 | 0.0724 |
| | - | (0.0428) | (0.0413) | (0.0498) | - | (0.0375) | (0.0384) | (0.046) |
| Post65*Post2006 | -0.0772 | -0.0491 | 0.014 | -0.0052 | 0.0918 | 0.0094 | -0.0106 | 0.0003 |
| | (0.0356) | (0.0252) | (0.0234) | (0.0266) | (0.0324) | (0.0251) | (0.0236) | (0.0268) |
| Age and Year Indicators*RHI only up to age 65 | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Demographic and Health Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Individual Fixed Effects | Yes | No | Yes | Yes | Yes | No | Yes | Yes |
| N | 4,784 | 11,920 | 11,982 | 8,637 | 4,784 | 11,920 | 11,982 | 8,637 |
| Number of Clusters | 2,646 | 4,906 | 4,934 | 3,321 | 2,646 | 4,906 | 4,934 | 3,321 |

*Notes:* This table presents the effects of Part D eligibility on the main outcomes of full-time and part-time work when treatment status is based only on lagged values. No inference of retiree insurance status is made from observed insurance status after age 65. The dependent variable of the first four columns is full-time work, and for the latter four columns part-time work. Columns (1) and (5) show the differences-in-differences estimates of the effect of Part D eligibility on the treatment group with no control group. The sample for these columns is restricted to the treatment group. Columns (2) and (6) show the same in a triple-differences design with standard control group, with no individual fixed effects. Columns (3) and (7) show the baseline specification. Columns (4) and (8) use only lagged values from 2002 or earlier to infer treatment status. The controls included in each specification are indicated in the table. Demographic controls include an indicator for being single, a set of indicators for each of the census divisions, and a fifth-order polynomial in non housing household wealth. Health controls include a set of indicators for self-reported health on a scale of 1-5; body-mass index; and a set of indicators for having any of the following physician-diagnosed conditions: cancer, lung disease, heart disease, stroke, arthritis, or psychiatric conditions. Columns (2) and (6) include additional demographic controls: gender, a full set of indicators for years of education, veteran status, and indicators for race (white, African American or other) and religion (Protestant, Catholic, Jewish, None or other). Robust standard errors clustered at the level of the individual are in parentheses.

The results in table 6 are similar to those in Table 1. In all specifications, with all samples, full-time work declines for the treated group by between 4.8 and 9 percentage points. The effect on part-time work is also of similar

magnitude using these smaller samples, between 7 and 12 percentage points. Furthermore, in all specifications in both samples the control group has no significant effect (except for a marginally significant effect in column (2)). The main difference between the results in the main analysis and here are the standard errors. Unsurprisingly, the standard errors in table 6 are somewhat larger, due to the smaller samples used here and the more severe measurement error in treatment status.

# B Robustness Checks

This section demonstrates that the results in Section IV are robust to a number of perturbations of the sample and design.

## A. Alternative Measurements of Labor Supply

The measures of labor force status in the main analysis are based on average hours of work per week and number of weeks worked per year. An interesting question in its own right, and a natural robustness check for previous results, is to consider the effect of Part D eligibility on the average of hours of work per week itself, as a measure of work intensity.

The results of using this variable as the outcome for the basic specification of equation (7) are in columns 1 and 2 of table 7. Column 1 shows the effect unconditional on working, with hours worked for individuals who do not work set to 0. Column 2 does the same, conditional on working. In both there is a negative effect of Part D eligibility on average hours of work a week, of between 2.7 and 4.9 hours a week less for the treated individuals upon eligibility. Column 3 constructs a new full-time work variable purely from reported average hours a week, with the variable equal to 1 if average hours a week are more than 35, and 0 otherwise. The estimated effect of Part D is similar to the main results, with a fall of 7.7 percentage points in full-time work for the treated.

Table 7: Alternative Definitions of Labor Supply

| | (1) | (2) | (3) |
|---|---|---|---|
| Dependent Variable: | Hours/Week | Hours/Week (If Working) | More Than 35 Hours/Week |
| Post65*Post2006*RHI only up to age 65 | -2.667 | -4.914 | -0.0770 |
| | (1.349) | (1.655) | (0.0319) |
| Post65*Post2006 | 0.553 | -0.922 | 0.0290 |
| | (0.947) | (1.290) | (0.0224) |
| Age and Year Indicators*RHI only up to age 65 | Yes | Yes | Yes |
| Demographics, Health, and Individual Fixed Effects | Yes | Yes | Yes |
| N | 15,076 | 7,511 | 15,076 |
| Number of Clusters | 6,465 | 4,038 | 6,465 |

*Notes:* This table presents estimates of the effect of Part D eligibility on various measures of labor supply. The dependent variable of each column appears in its heading. Individuals reporting more than 70 hours of work in a typical week are omitted. In the second column only individuals reporting strictly positive hours are included. The first row provides the triple-differences estimates of Part D eligibility on the dependent variable for individuals with employer-sponsored retiree health insurance (RHI) only until age 65. The third row provides the estimates of the effect of Part D eligibility on the dependent variable for the control group of individuals with RHI unlimited by age. Demographic controls include an indicator for being single, a set of indicators for each of the census divisions, and a fifth-order polynomial in non housing household wealth. Health controls include a set of indicators for self-reported health on a scale of 1-5; body-mass index; and a set of indicators for having any of the following physician-diagnosed conditions: cancer, lung disease, heart disease, stroke, arthritis, or psychiatric conditions. Robust standard errors clustered at the level of the individual are in parentheses.

## B. Alternative Control Group: No ESI

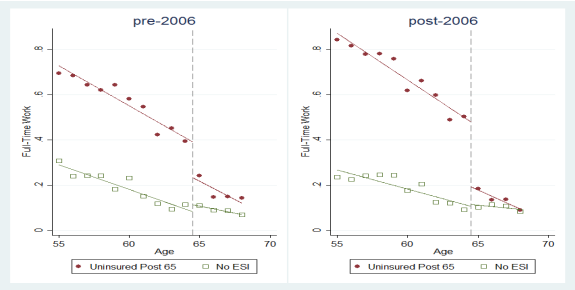
Thus far all the triple-differences regressions have used a control group of individuals who have RHI until any age. They are similar to the treatment group of individuals who have retiree insurance only until age 65, but differ in their prescription drug-induced retirement lock. A group less comparable to the treatment group, but equally unaffected by the relaxation of retirement lock, is workers who do not have any ESI.

Those without any ESI are less similar to the treatment group than those with RHI to any age on virtually every observable, from gender distribution to income (see columns 1 and 3 of table 5). This second control group nevertheless allows me to test the robustness of the main results by comparing the treated group to a different, yet still untreated (with respect to retirement lock), control group.

Figure 6 shows the pre trends of full-time work for the treatment group, who have RHI until age 65, in circles; and for this alternative control group of individuals with no ESI whatsoever, in squares. The gap between the treatment group's and this control group's full-time work rates before 2006 is larger

than when using the original control group, however the trends are roughly parallel.

Figure 6: Full-Time Work Rates in the Treatment and Alternative Control Groups by Age




*Notes:* This figure shows the triple differences of full-time work using an alternate control group off individuals who had no ESI. The sample is individuals aged 55-68, in the years 2000 until 2010. The squares depict full-time work by age for the control group of individuals who have no ESI whatsoever. The circles depict full-time work by age for the treatment group of individuals who have RHI only until age 65. The left-hand panel consists of observations in the years 2000-2004, before Part D; the right-hand panel consists of observations from the years 2006-2010, after the introduction of Part D. The dashed gray line differentiates between ages eligible for Part D, on the right, and those ineligible, on the left (in the post 2006 period).

Table 8 confirms that the qualitative results hold using this alternative control group. While the precise numbers are naturally different, they are of the same sign and order of magnitude. This estimation indicates a 6.7 percentage point decline in full-time work and a (statistically insignificant) 2.5 percentage point increase in part-time work for the treated in the baseline specification. As above, there are no statistically significant effects for the control group.

Table 8: Alternate Control Group with No ESI

| | (1) | (2) |
|---|---|---|
| Dependent Variable: | Full-Time Work | Part-Time Work |
| Post65*Post2006*Treated | -0.0668 | 0.0245 |
| | (0.0320) | (0.0327) |
| Post65*Post2006 | -0.00959 | 0.0189 |
| | (0.0160) | (0.0186) |
| Age and Year Indicators * Treated | Yes | Yes |
| Demographics, Health, and Individual Fixed Effects | Yes | Yes |
| N | 19,224 | 19,224 |
| Number of Clusters | 8,913 | 8,913 |

*Notes:* This table presents estimates of the effect of Part D eligibility on full-time and part-time work relative to a control group of individuals who had no ESI. The dependent variable of each column appears in its heading. The first row provides the triple-differences estimates of Part D eligibility on the dependent variable for individuals with retiree health insurance (RHI) only until age 65. The third row provides the estimates of the effect of Part D eligibility on the dependent variable for the control group of individuals with no ESI. Demographic controls include an indicator for being single, a set of indicators for each of the census divisions, and a fifth-order polynomial in non housing household wealth. Health controls include a set of indicators for self-reported health on a scale of 1-5; body-mass index; and a set of indicators for having any of the following physician-diagnosed conditions: cancer, lung disease, heart disease, stroke, arthritis, or psychiatric conditions. Robust standard errors clustered at the level of the individual are in parentheses.

## C. Different Specifications

**Differences-in-Differences: Older and Younger than 65/Before and After 2006** In addition to the above perturbations, the results are also robust to a number of other changes to the specification described in Section III. Table 9 shows the estimates for full-time (columns 1 and 2), part-time (columns 3 and 4), and any work (columns 5 and 6), replicating Table 1 with a differences-in-differences, rather than triple-differences, approach. In contrast to the triple-differences, this design keeps *only* the treatment group, and compares the change in their outcomes at age 65, before and after 2006. While this design is more straightforward, it cannot control for unobservable changes affecting individuals over age 65 after 2006 differently than individuals under age 65 (e.g., the increasing Social Security FRA or age-specific shocks from the Great Recession). Thus the point estimates vary somewhat from the main results; however their signs remain consistent with the main results. For example, the estimated effect on full-time work with controls is a reduction of 7 percentage points.

Table 9: Differences-in-Differences Estimates of the Effect of Medicare Part D Eligibility

| Dependent Variable: | Full-Time Work | | Part-Time Work | | Any Work | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Specification/Sub-Sample: | No Controls | Baseline | No Controls | Baseline | No Controls | Baseline |
| Post65*Post2006 | -0.146 | -0.0703 | 0.0556 | 0.0641 | -0.0901 | -0.0062 |
| | (0.0219) | (0.0305) | (0.0191) | (0.0289) | (0.0242) | (0.0326) |
| Age and Year Indicators*RHI only up to age 65 | Yes | Yes | Yes | Yes | Yes | Yes |
| Demographics, Health, and Individual Fixed Effects | No | Yes | No | Yes | No | Yes |
| N | 7,046 | 6,850 | 7,046 | 6,850 | 7,046 | 6,850 |
| Number of Clusters | 3,859 | 3,717 | 3,859 | 3,717 | 3,859 | 3,717 |

*Notes:* This table presents differences-in-differences estimates of the effect of Part D eligibility on full-time work (columns 1 and 2), part-time work (columns 3 and 4), and any work (columns 5 and 6). The controls in columns 1, 3, and 5 are only age and time fixed effects, and age and time fixed effected interacted with having retiree health insurance (RHI) only up to age 65. Columns 2, 4, and 6 also include an indicator for being single, a set of indicators for each of the census divisions, a fifth-order polynomial in non housing household wealth, a set of indicators for self-reported health on a scale of 1-5, body-mass index, and a set of indicators for having any of the following physician-diagnosed conditions: cancer, lung disease, heart disease, stroke, arthritis, or psychiatric conditions. The first row provides the differences-in-differences estimates of Part D eligibility on the dependent variable for individuals with RHI only until age 65. Robust standard errors clustered at the level of the individual are in parentheses.

**Differences-in-Differences: Treatment versus Control/Before and After 2006** Section III describes the identifying assumption that Part D had no effect on individuals below age 65 who were thus not eligible for its benefits. An alternative assumption could be made to identify the effect of Part D using a different design: if the control group provides a good counterfactual time trend to the treatment group over age 65 *on its own*, there is no need to control further for time trends using individuals in the treatment group below age 65. As stated in footnote (18) this alternative assumption is less consistent with the patterns in the data than the initial assumption: while a violation of the initial assumption would imply a decline in labor for the treated under age 65 in the years following Part D, in fact they display a substantial increase. In contrast, there is only a very mild increase in full-time work among the control individuals younger than 65 after 2006 relative to before. Thus it appears that some factor leading to increased work among the treated is missing among the controls, raising the need for both counterfactuals, and the triple-difference, to properly assess the effect of Part D.

Nevertheless, to test sensitivity of the results to relying more heavily on the similarity of the treatment and control groups, table 10 displays estimates of a

differences-in-differences design with only individuals over age 65. Individual fixed effects are not included here as the panels are too short, with only 1.6 observations per individual on average (recall that assignment to treatment and control groups is only possible up to age 68, so this approach restricts the sample to ages 65-68, with a bi-annual survey). In this approach results for full-time work are attenuated (because of the fact that the treatment group might have had higher work rates in the absence of Part D as indicated by the treated individuals below age 65 which are excluded here). The effect on full-time work is thus smaller in magnitude and statistically insignificant (although marginally significant without controls, p=0.064). However, the sign of the point estimates is consistent with the main results, and economically still large. Furthermore, in this approach a larger share of the reduction in full-time work is due to individuals leaving work completely, rather than shifting to part-time work, as evident from columns 3 and 4.

Table 10: Differences-in-Differences Estimates Based on Treatment and Control, Only After Age 65

| Dependent Variable: | Full-Time Work | | Part-Time Work | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Specification | No Controls | With Controls | No Controls | With Controls |
| Post2006*RHI only up to age 65 | -0.0624 | -0.0335 | 0.0086 | -0.0069 |
| | (0.0333) | (0.0297) | (0.0423) | (0.0329) |
| Year Indicators*RHI only up to age 65 | Yes | Yes | Yes | Yes |
| Age, Demographic, and Health Controls | No | Yes | No | Yes |
| N | 6,784 | 6,675 | 6,784 | 6,675 |
| Number of Clusters | 4,256 | 4,197 | 4,256 | 4,197 |

*Notes:* This table presents differences-in-differences estimates of the effect of Part D eligibility on full-time work, using only individuals over age 65 and comparing the treatment and control groups before and after 2006. The dependent variable is full-time work in columns 1 and 2, and part-time work in columns 3 and 4. The first row provides the estimates of the effect of Part D eligibility for individuals with retiree health insurance (RHI) only until age 65. Demographic controls include an indicator for being single, a set of indicators for each of the census divisions, and a fifth-order polynomial in non housing household wealth. Health controls include a set of indicators for self-reported health on a scale of 1-5; body-mass index; and a set of indicators for having any of the following physician-diagnosed conditions: cancer, lung disease, heart disease, stroke, arthritis, or psychiatric conditions. Robust standard errors clustered at the level of the individual are in parentheses.

**Hazard Models** Table 11 shows estimates of hazard models. This approach is an alternative to estimation with individual fixed effects which nevertheless accounts for unobserved heterogeneity across individuals. It assumes that those leaving full-time work do not return to full-time work, and that those

leaving work completely do not return to work. The qualitative results are in line with the main results of the paper, with Part D eligibility increasing the hazard of leaving full-time work by 72 percent for individuals in the treatment group (in the specification including controls), with no significant effect on the control group. While this estimate is very large, the standard errors are also large, and a hazard ratio as small as 22 percent cannot be rejected at 95 percent confidence. This would be in line with the semi-elasticity implied by the paper's main estimate, of 25 percent. Similarly, the hazard of leaving any work increases with Part D eligibility by 46 percent for the treated (again, with no effect for the control group). For this outcome I cannot reject an increase as small as 1.7 percent at 95 percent confidence.

Table 11: Hazard Models

| Outcome | Full-Time | | Any Work | |
|---|---|---|---|---|
| Specification: | No Controls | Controls | No Controls | Controls |
| | (1) | (2) | (3) | (4) |
| Post65*Post2006*RHI only up to age 65 | 1.759 | 1.718 | 1.484 | 1.461 |
| | (0.31) | (0.299) | (0.276) | (0.27) |
| Post65*Post2006 | 1.063 | 0.89 | 1.124 | 0.957 |
| | (0.125) | (0.103) | (0.143) | (0.12) |
| Age and RHI only up to age 65 Indicators | Yes | Yes | Yes | Yes |
| Demographic, and Health Controls | No | Yes | No | Yes |
| N | 5,913 | 5,663 | 8,329 | 7,989 |
| Number of Clusters | 3,321 | 3,122 | 4,331 | 4,082 |

*Notes:* This table presents estimates of the effect of Part D eligibility on full-time work (in columns 1 and 2) and on any work (columns 3 and 4) using a hazard model (failure is leaving full-time work in columns 1 and 2, and leaving work completely in columns 3 and 4). Columns 1 and 3 control only for being over age 65, being observed after 2006, having retiree health insurance (RHI) only up to age 65, and all the second and third order interactions of these variables, along with a full set of age indicators. Columns 2 and 4 also include indicators for gender, race, years of education, and veteran status. Further time-varying controls in columns 2 and 4 are: an indicator for being in the treatment group, an indicator for being single, a set of indicators for each of the census divisions, a fifth-order polynomial in non housing household wealth, a set of indicators for self-reported health on a scale of 1-5, body-mass index, and a set of indicators for having any of the following physician-diagnosed conditions: cancer, lung disease, heart disease, stroke, arthritis, or psychiatric conditions. The first row provides the estimated effect of Part D eligibility on the hazard ratio for individuals with RHI only until age 65. The third row does the same for individuals with RHI for life. Robust standard errors clustered at the level of the individual are in parentheses.

**Separate Year and Age Trends by Demographics** Table 12 estimates equation (7) with added controls: interactions of the demographic variables with the year effects (columns 1 and 3) and with the age and year effects (columns 2 and 4). The demographic groups allowed their own year and age

fixed effects are gender, a full set of indicators for years of education, veteran status, and indicators for race (white, African American, or other) and religion (Protestant, Catholic, Jewish, None, or other). The results are similar to those in the main analysis, with declines in full-time work around 8 percentage points, and increases in part-time work around 6.5 percentage points. As in the main results, there were no significant effects for the control group. Thus to the extent that the treatment and control groups are different on these observable characteristics, those differences do not seem to be driving the differential labor outcomes of the two groups over age 65, after 2006. This complements the results in columns 1 and 2 of Table 18 below, which similarly control for interactions of year and age with Census division.

Table 12: Triple Differences – Interactions of Demographics with Age and Time Fixed-Effects

| Dependent Variable: | Full-Time Work | | Part-Time Work | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Post65*Post2006*RHI only up to age 65 | -0.0831 | -0.0775 | 0.0646 | 0.0663 |
| | (0.0317) | (0.032) | (0.0311) | (0.0316) |
| Post65*Post2006 | 0.0146 | 0.0127 | 0.002 | 0.0025 |
| | (0.0219) | (0.0223) | (0.022) | (0.0225) |
| Age and Year Indicators*RHI only up to age 65 | Yes | Yes | Yes | Yes |
| Demographic, Health, and Individual Fixed Effects | Yes | Yes | Yes | Yes |
| Year Indicators * Demographics | Yes | Yes | Yes | Yes |
| Age Indicators * Demographics | No | Yes | No | Yes |
| N | 15,303 | 15,303 | 15,303 | 15,303 |
| Number of Clusters | 6,479 | 6,479 | 6,479 | 6,479 |

*Notes:* This table presents robustness checks for the triple-differences estimates of the effect of Part D eligibility on full-time work (columns 1 and 2) and part-time work (columns 3 and 4). The controls included in each specification are indicated in the table. The first row provides the triple-differences estimates of Part D eligibility for individuals with retiree heath insurance (RHI) only until age 65. The third row provides the estimates of the effect of Part D eligibility for the control group of individuals with RHI unlimited by age. Demographic controls include an indicator for being single, a set of indicators for each of the census divisions, and a fifth-order polynomial in non housing household wealth. Health controls include a set of indicators for self-reported health on a scale of 1-5; body-mass index; and a set of indicators for having any of the following physician-diagnosed conditions: cancer, lung disease, heart disease, stroke, arthritis, or psychiatric conditions. The demographic variables interacted with age and year are gender, a full set of indicators for years of education, veteran status, and indicators for race (white, African American, or other) and religion (Protestant, Catholic, Jewish, None, or other). Robust standard errors clustered at the level of the individual are in parentheses.

**Clustering by Household** The main results on full-time work are robust to clustering by household, rather than by individual:

Table 13: Triple Differences Estimates of Part D Eligibility's Effect on Labor

| Dependent Variable: | Full-Time Work | Part-Time Work | Any Work |
|---|---|---|---|
| | (1) | (2) | (3) |
| Post65*Post2006*RHI only up to age 65 | -0.0836 | 0.0589 | -0.0247 |
| | (0.0414) | (0.0409) | (0.0446) |
| Post65*Post2006 | 0.0199 | 0.00157 | 0.0215 |
| | (0.0289) | (0.0289) | (0.033) |
| Age and Year Indicators*RHI only up to age 65 | Yes | Yes | Yes |
| Demographics, Health, and Individual Fixed Effects | Yes | Yes | Yes |
| N | 15,382 | 15,382 | 15,382 |
| Number of Clusters | 5,017 | 5,017 | 5,017 |

*Notes:* This table presents triple differences estimates of the effect of Part D eligibility on full-time work (columns 1 and 2), part-time work (columns 3 and 4), and any work (columns 5 and 6). The controls in columns 1, 3, and 5 are only age, time, and having retiree health insurance (RHI) only up to age 65 fixed effects, and age and time fixed effects interacted with treatment group. Columns 2, 4, and 6 also include an indicator for being single, a set of indicators for each of the census divisions, a fifth-order polynomial in non housing household wealth, a set of indicators for self-reported health on a scale of 1-5, body-mass index, and a set of indicators for having any of the following physician-diagnosed conditions: cancer, lung disease, heart disease, stroke, arthritis, or psychiatric conditions. The first row provides the triple-differences estimates of Part D eligibility on the dependent variable for individuals with employer-sponsored RHI only until age 65. The third row provides the estimates of the effect of Part D eligibility on the dependent variable for the control group of individuals with RHI unlimited by age. Robust standard errors clustered at the household level are in parentheses.

**Other Specifications** Table 14 shows some other specifications with full-time work as the outcome. The estimated effect on the treated remains robust and uniformly insignificant effects persist for the control group. Column 1 excludes from estimation individuals younger than age 62, to verify that results are not driven by younger workers who may be less comparable to the treated group of over 65-year-olds; column 2 excludes individuals ages 63-64 who may be "treated" by the introduction of Part D due to its expected effect on them later in life when thye attain eligibility for the program; column 3 excludes individuals who are on Medicaid or Veteran Affairs, as these individuals would have had prescription drug insurance before Part D; and columns 4 and 5 take as the outcome variable the change in full-time work instead of including individual fixed effects (with and without demographic and health controls, in column 5 and 4, respectively). Results are broadly consistent with the main analysis (although only marginally significant in columns 1, 4, and 5).

Table 14: Triple Differences Estimates – Robustness Checks

| Dependent Variable: | Full-Time Work | | | First Difference of Full-Ti | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | |
| Specification/Sub-Sample: | Only Ages 62-68 | Excluding Ages 63-64 | Excluding Medicaid and VA | Years 2002-2010 | Years |
| Post65*Post2006*RHI only up to age 65 | -0.0802 | -0.0942 | -0.0938 | -0.0494 | -0 |
| | (0.0452) | (0.0377) | (0.0329) | (0.0257) | (0 |
| Post65*Post2006 | 0.00405 | 0.0189 | 0.0259 | 0.0033 | 0. |
| | (0.0286) | (0.0292) | (0.0228) | (0.0168) | (0. |
| Age and Year Indicators*RHI only up to age 65 | Yes | Yes | Yes | Yes | |
| Demographic and Health Controls | Yes | Yes | Yes | No | |
| Individual Fixed Effects | Yes | Yes | Yes | No | |
| N | 9,790 | 13,220 | 14,345 | 14,326 | 14 |
| Number of Clusters | 4,785 | 6,415 | 6,226 | 6,168 | 6 |

*Notes:* This table presents robustness checks for the triple differences estimates of the effect of Part D eligibility. The dependent variable is an indicator for full-time work (columns 1, 2, and 3) and the change in full-time work from the previous period (columns 4 and 5). Column 1 is estimated only on a sample of 62-68 year-olds, while column 2 excludes 63-64 year-olds. Column 3 is estimated on a sample excluding individuals on Medicaid or Veteran's Affairs. Columns 4 and 5 omit observations in year 2000. The first row provides the triple-differences estimates of Part D eligibility for individuals with retiree health insurance (RHI) only until age 65. The third row provides the estimates of the effect of Part D eligibility for the control group of individuals with RHI unlimited by age. Demographic controls include an indicator for being single, a set of indicators for each of the census divisions, and a fifth-order polynomial in non housing household wealth. Health controls include a set of indicators for self-reported health on a scale of 1-5; body-mass index; and a set of indicators for having any of the following physician-diagnosed conditions: cancer, lung disease, heart disease, stroke, arthritis, or psychiatric conditions. Robust standard errors clustered at the level of the individual are in parentheses.

An alternative to the regressions in section IV.G, showing heterogeneity in the effect by health status, is to estimate heterogeneity by regular use of prescription drugs. This is potentially an outcome of Part D coverage, and indeed, the literature shows utilization of drugs increased with Part D. Nevertheless, estimates relying on this measure have been helpful in other contexts such as the discussion of Krueger, Katz and Notowidigdo (2017). Table 15 replicates table 4 where instead of splitting the sample by physician-diagnosed health status, the sample is split based on self-reported regular drug use. The results are broadly similar to those split by health status: Part D is estimated to have had a significant negative effect on full-time work for those who regularly take prescription drugs (-11.5 percentage points), with a smaller (and statistically insignificant) effect on those who do not regularly take drugs.

Table 15: Heterogeneity by Regular Drug Use

| Dependent Variable: | Full-Time Work | |
|---|---|---|
| | (1) | (2) |
| Sub-Sample: | Regularly Takes Rx | Does Not Regularly Take Rx |
| Post65*Post2006*RHI only up to age 65 | -0.115 | -0.076 |
| | (0.032) | (0.07) |
| Post65*Post2006 | -0.042 | -0.002 |
| | (0.022) | (0.05) |
| Controls | Yes | Yes |
| N | 12,481 | 3,089 |
| Number of Clusters | 5,548 | 1,781 |

*Notes:* This table presents heterogeneity of the effect of Part D eligibility on full-time and part-time work by regular use of prescription drugs. The sub-sample of each column is detailed in the column's heading. Both columns control for age and time fixed effects, age and time fixed effects interacted with having retiree health insurance (RHI) only up to age 65, an indicator for being single, a set of indicators for each of the census divisions, and a fifth-order polynomial in non housing household wealth. The first row provides the triple-differences estimates of Part D eligibility for individuals with employer-sponsored RHI only until age 65. The third row provides the estimates of the effect of Part D eligibility for the control group of individuals with RHI unlimited by age.

## D. Different Samples

There are no other large-scale data sets, to my knowledge, that contain the information necessary to support the research design used in this paper (e.g., whether the individual will have RHI, and till what age). However, by giving up the focus on the population most affected by the introduction of Part D, it is possible to conduct a differences-in-differences analysis using the American Community Survey (ACS). Naturally, this leads to attenuated results relative to those found in the main analysis. To begin with, over half the population is in my "control groups" (having retiree insurance for life, or having no ESI whatsoever, the control group used in Appendix B2), and they should exhibit null effects which are included in the treatment effect averaged over the entire population. Furthermore, it is likely that individuals in jobs with RHI, even limited to age 65, value health insurance more highly than the general population, and are thus more responsive to the availability of Part D. Finally, this approach cannot control for changes that affect individuals over age 65 after 2006 differently than those younger than 65 besides Part D (such as the change in Social Security's FRA).

All these consideration notwithstanding, the qualitative effect can be found

in the larger ACS sample, as well, and the results are in table 16. On the extensive margin, eligibility for Part D led to a decline of 1 percentage point in any work; on the full-time work margin there was a decline of 0.37 percentage points. To compare apples to apples, columns 3 and 4 of the table estimate a similar specification on the HRS sample. This differences-in-differences does not focus on the treatment group, as in Section IV of the main analysis, but rather takes the entire HRS sample ages 55-68 in the years 2000-2010. The results are fairly similar across the two samples, particularly for the "any work" outcome. It is possible that hours of work are measured with greater error than any work, leading to greater attenuation bias in the full-time work estimates.

Table 16: Population-Wide Differences-in-Differences Results, in the ACS and the HRS

| | ACS | | HRS | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Dependent Variable: | Full-Time Work | Any Work | Full-Time Work | Any Work |
| Post65*Post2006 | -0.0037 | -0.0105 | -0.0168 | -0.0119 |
| | (0.0011) | (0.0012) | (0.01) | (0.0105) |
| Age and Year Indicators | Yes | Yes | Yes | Yes |
| Demographic Controls | Yes | Yes | Yes | Yes |
| Health Controls | No | No | Yes | Yes |
| Individual Fixed Effects | No | No | Yes | Yes |
| N | 2,268,986 | 2,268,986 | 42,025 | 42,025 |
| Number of Clusters | - | - | 14,247 | 14,247 |

*Notes:* This table presents estimates of the effect Part D eligibility on full-time work using a differences-in-differences design on the entire population, regardless of insurance status. The dependent variable is an indicator for full-time work, in columns 1 and 3, or for any work, in columns 2 and 4. Full-time work in the ACS is defined as reporting more than 35 hours of work in a usual week, and not being unemployed or out of the labor force. Being in any work in the ACS is defined as not being unemployed or out of the labor force. In the HRS these variables are defined as in the main analysis. The sample is 62-68 year-olds in the 2000-2010 ACS, and 55-68 in the HRS. The first row provides the differences-in-differences estimates of Part D eligibility's effect. Demographic controls in the ACS sample include marital status, education, and gender. Demographic controls in the HRS sample include an indicator for being single, a set of indicators for each of the census divisions and a fifth-order polynomial in non housing household wealth. Health controls in the HRS sample include a set of indicators for self-reported health on a scale of 1-5; body-mass index; and a set of indicators for having any of the following physician-diagnosed conditions: cancer, lung disease, heart disease, stroke, arthritis, or psychiatric conditions. Robust standard errors are in parentheses, and errors are clustered by individual in columns 3 and 4.

## E. Falsification Tests

As an additional falsification test, I check whether a placebo treatment defined as eligibility for Part D at ages other than age 65 yields no significant effect (excluding observations over age 65). Reassuringly, this is the case; table 17

shows estimates when placebo eligibility for Part D is assigned at ages 60 and 62 (columns 1-2 and 3-4, respectively). The point estimates indicate an *increase* in full-time work for the treated due to placebo Part D eligibility, and none are near significant. There are similar null effects for the control group.

Table 17: Falsification Tests

| | Placebo Treatment at age 60 | | Placebo Treatment at age 62 | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Dependent Variable: | Full-Time Work | Part-Time Work | Full-Time Work | Part-Time Work |
| Post65*Post2006*RHI only up to age 65 | 0.0143 | 0.0143 | 0.0408 | 0.0226 |
| | (0.0506) | (0.0459) | (0.0539) | (0.0467) |
| Post65*Post2006 | -0.0139 | -0.0075 | -0.0132 | -0.029 |
| | (0.0347) | (0.0316) | (0.0344) | (0.0322) |
| Age and Year Indicators*RHI only up to age 65 | Yes | Yes | Yes | Yes |
| Demographics, Health, and Individual Fixed Effects | Yes | Yes | Yes | Yes |
| N | 8,676 | 8,676 | 8,676 | 8,676 |
| Number of Clusters | 4,859 | 4,859 | 4,859 | 4,859 |

*Notes:* This table presents triple-differences estimates of the effect of "placebo" Part D eligibility on full-time work (columns 1 and 3) and part-time work (columns 2 and 4), and any work (columns 5 and 6). The sample is ages 55-64, in years 2000-2010. Placebo eligibility is assigned at age 60 in columns 1 and 2, and at age 62 in columns 3 and 4. Controls are age and time fixed effects, and age and time fixed effects interacted with having retiree health insurance (RHI) only up to age 65, as well as demographic and health controls: an indicator for being single, a set of indicators for each of the census divisions, a fifth-order polynomial in non housing household wealth, a set of indicators for self-reported health on a scale of 1-5, body-mass index, and a set of indicators for having any of the following physician-diagnosed conditions: cancer, lung disease, heart disease, stroke, arthritis, or psychiatric conditions. The first row provides the triple-differences estimates of the effect of placebo Part D eligibility for individuals with RHI only until age 65. The third row provides the estimates of the effect of placebo Part D eligibility for the control group of individuals with RHI unlimited by age. Robust standard errors clustered at the level of the individual are in parentheses.

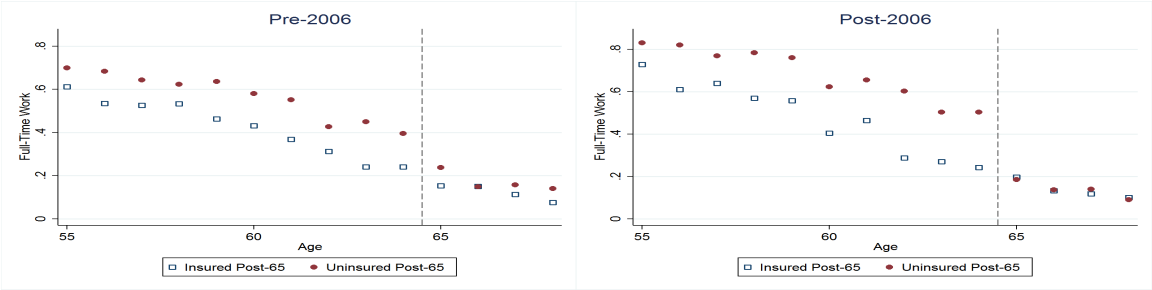
## F. Supporting Figures and Tables

This appendix displays figures and tables referenced in the main text but made available online for brevity.

Figure 7 shows shows the same data as Figure 1 but with the different dimensions of the three conditions (pre/post 2006, before/after age 65, treatment/control group) substituted so that the left-hand panel shows the full-time work rates for the treatment and control groups before 2006, while the right-hand panel shows them after 2006. This alternative presentation of the same data allows verification of the parallel pre trends between the two treatment groups. Again, while the control group displays no sharp drop in full-time work rates at age 65 either before or after Part D, the treatment group has a

substantially larger drop post 2006 relative to pre 2006. It is interesting to note that in the post 2006 period the treatment and control groups behave similarly after age 65, consistent with both groups at this point facing detachment of the labor and insurance decisions.

Figure 7: Full-Time Work Rates by Age and Treatment Status

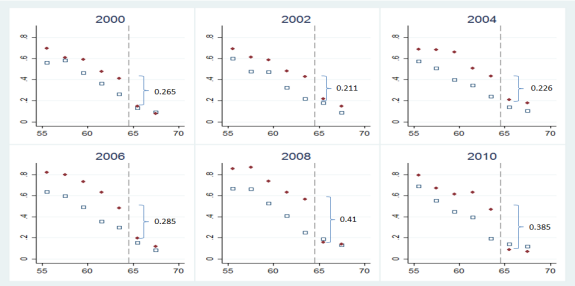



*Notes:* This figure shows the triple differences of full-time work. The sample is individuals aged 55-68, in the years 2000 until 2010, who have RHI through their employer. The squares depict the rates of full-time work by age for the control group of individuals who have retiree health insurance (RHI) through their employer unlimited by age. The circles depict full-time work rates by age for the treatment group of individuals who have RHI through their employer only until age 65. The panel on the left consists of observations in the years 2000-2004, before Part D; the panel on the right consists of observations from the years 2006-2010, after the introduction of Part D. The dashed gray line differentiates between ages eligible for Part D, on the right, and those ineligible, on the left (in the post 2006 period).

Figure 8 displays differences-in-differences figures separately for each year in the sample. Instead of pooling all three pre Part D survey years and all three post Part D survey years as in figure 7, figure 8 shows the same information on full-time work by age and by treatment group at an annual level. This illustrates two main points: the first is that the treatment and control groups have parallel pre trends every year, not just averaged out over the pre and post Part D years (consistent with the points estimates shown in figure 2). Second, it shows that the pivotal year in which the full-time work of the treatment group begins to decline more sharply at age 65 is in fact 2006. Whereas the decline in 2000-2004 is 23 percentage points (averaged over the three years), the fall at age 65 in 2006 is 28 percentage points.

Table 18 estimates the main specifications for full- and part-time work accounting for the Great Recession in various ways: by interacting age, year, and treatment group indicators with Census division (columns 1-2), and by excluding the Great Recession years from the sample. Columns 3-4 exclude all years after 2006, while columns 5-6 exclude only the actual years of the recession. In all cases the results are qualitatively similar to the main results.

Figure 8: Full-Time Work Rates by Age,Treatment Status, and Year




*Notes:* This figure shows the triple differences of full-time work, on a year-by-year level. The sample is individuals aged 55-68, in the years 2000 until 2010. Squares depict the rates of full-time work by every two consecutive ages for the control group of individuals who have retiree health insurance (RHI) through their employer unlimited by age. Circles depict full-time work rates by every two consecutive ages for the treatment group of individuals who have RHI through their employer only until age 65. The top row represents observations from the years 2000-2004, before Part D; the bottom row consists of observations from the years 2006-2010, after the introduction of Part D. The dashed gray line differentiates between ages eligible for Part D, on the right, and those ineligible, on the left (in the post 2006 period). The brackets indicate the difference in full-time work rates for the treated group between ages 63-64 and 65-66 in every survey wave.

Table 18: Accounting for the Great Recession

| | 2000-2010 | | 2000-2006 | | 2000-2006 and 2010 | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Dependent Variable: | Full-Time | Part-Time | Full-Time | Part-Time | Full-Time | Part-Time |
| Post65*Post2006*RHI only up to age 65 | -0.0725 | 0.0506 | -0.0560 | 0.0164 | -0.0928 | 0.0433 |
| | (0.0314) | (0.0308) | (0.0440) | (0.0427) | (0.0392) | (0.0370) |
| Post65*Post2006 | 0.0173 | 0.018 | 0.0109 | 0.0158 | 0.0237 | 0.00573 |
| | (0.0217) | (0.022) | (0.0289) | (0.0278) | (0.0262) | (0.0253) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Age, Year, "RHI only up to age 65" Indicators * Census Division | Yes | Yes | No | No | No | No |
| N | 15,382 | 15,382 | 11,646 | 11,646 | 13,039 | 13,039 |
| Number of Clusters | 6,515 | 6,515 | 5,741 | 5,741 | 6,201 | 6,201 |

*Notes:* This table presents estimates of the effect of Part D eligibility on full-time and part-time work, accounting for the Great Recession. The sub-sample of years after 2006 included in the estimation for each column is in the column's heading, with the entire 2000-2010 sample included in the first two columns, only 2006 in the post Part D period in the third and fourth columns, and 2006 and 2010 comprising the post Pat D period in the fifth and sixth columns. The dependent variable of each column appears in its heading. All columns control for age and time fixed effects, age and time fixed effects interacted with having retiree health insurance (RHI), an indicator for being single, individual fixed effects, a set of indicators for each of the census divisions, a fifth-order polynomial in non housing household wealth, a set of indicators for self-reported health on a scale of 1-5, body-mass index, and a set of indicators for having any of the following physician-diagnosed conditions: cancer, lung disease, heart disease, stroke, arthritis, or psychiatric conditions. Columns 1 and 2 also include the interaction of Census Division indicators with age and year indicators, as well as with an indicator for having RHI only until age 65. The first row provides the triple-differences estimates of Part D eligibility for individuals with employer-sponsored RHI only until age 65. The third row provides the estimates of the effect of Part D eligibility for the control group of individuals with RHI unlimited by age. Robust standard errors clustered at the level of the individual are in parentheses.

# C The Social Welfare Implications of Part D

This appendix develops a simple model for the effect of the Part D subsidy on the government budget which accounts for behavioral responses. Based on the elasticities estimated in Section IV, every dollar of drug insurance subsidy distributed by the government is estimated to have a fiscal cost of $1.68. I.e., a dollar spent subsidizing the prescription drug insurance of retirees costs the government an extra 68 cents due to behavioral responses. Coupled with the estimated willingness to pay of beneficiaires, this yields a marginal value of public funds in Part D of $2.

**The Social Cost of Medicare Part D** The estimate of $R$ in Section V yields the mean individual retiree's willingness to pay for the Part D subsidy. The social benefits of the subsidy must, however, also account for the social cost of providing it. If subsidization has no further deadweight loss then this

cost would be \$1 for every dollar of the subsidy distributed. However, the subsidy is likely to place further strain on the government budget through individuals' behavioral responses.

The two most obvious such behavioral responses are the immediate corollaries of the empirical findings in Section IV: that the subsidy leads to reduced labor supply, implying both higher takeup of the subsidy itself as individuals transition from ESI to Medicare, and generally increase their rate of drug coverage; as well as the reduction in income tax revenue resulting from decreased labor.[42] These three elements, the changes in labor, the changes in the amount of insurance bought, and changes in taxable income can be incorporated into a simple model of the government budget.[43]

Define the government budget per capita as:

$$B \equiv A - (1 - G(\overline{v}(s)))sx + \tau_a * I(s), \tag{8}$$

where $A$ signifies revenue per capita from sources other than income tax; $(1 - G(\overline{v}(s)))s$ is the average subsidy to the prescription drug insurance of those not working full-time per unit of insurance; $x$ is the average quantity of insurance they purchase; $\tau_a$ is the average income tax rate; and $I(s)$ is average income, so that $I(s) \equiv G(\overline{v}(s)) * I(1) + (1 - G(\overline{v}(s))) * I(0)$. The effect on the budget of offering another dollar of subsidy is therefore given by:

$$\frac{1}{x}\frac{dB}{ds} = -(1 - G(\overline{v}(s))) + s\frac{dG(\overline{v}(s))}{ds} - (1 - G(\overline{v}(s)))\frac{s}{x}\frac{dx}{ds} + \frac{\tau_a}{sx} * \frac{sdI(s)}{ds} \tag{9}$$

[42] I abstract from the cost of the subsidy to ESI included in the Medicare Modernization Act: this subsidy costs the government resources, of course, but also reduces the cost to the government implicit in individuals transitioning from ESI to other Part D plans. In this sense the additional policy of subsidizing ESI alongside individual plans interacts with the individual subsidies to reduce the fiscal cost of the latter.

[43] This model could also be expanded to account for potential other ways in which the Part D subsidies might impact the budget, such as changes in Social Security claiming or costs borne by other programs such as Veterans Affairs. A sufficient statistic for this could be the causal effect of Part D on the government budget (as described in Hendren 2016), however estimating this parameter is beyond the scope of the current paper.

The first term is the mechanical static cost of the subsidy, the additional dollar given to all those who were already retired; the second term states that the entire subsidy must now be given to individuals who choose to retire due to the change in subsidy; the third term indicates that the entire subsidy must be given to additional units of insurance that retirees are induced to purchase due to the lower price of insurance; the final term captures the reduction in income tax revenues due to individuals' behavioral responses to the subsidy, their lower rate of work. These last three terms together make up the fiscal externality.

All the terms in equation (9) were estimated in Section IV, with the exception of the elasticity of demand for insurance with respect to the subsidy, $\frac{s}{x}\frac{dx}{ds}$. This latter term is estimated in Appendix D using a differences-in-differences research design with only the treatment group, as in the first section of Appendix B3, with prescription drug insurance coverage as the outcome variable. The result of that estimation is that $\frac{s}{x}\frac{dx}{ds} = 0.15$.

The other quantities used in the calibration are, based on the results from Section IV:

$(1 - G(\overline{v}(s))) = 0.65$

$s\frac{dG(\overline{v}(s))}{ds} = 0.084$

$\frac{sdI(s)}{ds} = 1,477$

an average income tax rate of $\tau_a = 0.28$ (using 2006 rates for federal and average state income taxes, Tax Policy Center, 2014)

and $sx = 1,588$.[44] Plugging these numbers into equation (9) and normalizing by the share of the population receiving the subsidy gives:

$$\frac{1}{x}\frac{dB}{ds}/(1 - G(\overline{v}(s))) = 1.68.$$

Following Hendren (2016) we can get the marginal value of public funds spent on the subsidy to prescription drugs of retirees by integrating the willing-

[44]This is different than the number used in Section V because it is the subsidy for one year, rather than discounted over the lifetime, to keep it in the same units as the change in annual labor income due to Part D estimated in Section IV. The value $sx = 1588$ is the net subsidy per capita in 2010 (Medicare Board of Trustees (2014)).

ness to pay for a dollar of subsidy over the population, and accounting for the fiscal cost of providing that subsidy. The willingness to pay estimated above is the average among retirees. The willingness to pay of full-time workers for a subsidy they do not benefit from is 0.[45] Therefore the average willingness to pay in the population is $willingness - to - pay * (1 - \%full\_time)$. Dividing this by the fiscal cost of a dollar of subsidy estimated above gives a marginal value of public funds of \$2, in terms of dollars of income to retirees.[46]

# D Estimation of Elasticity of Insurance Demand

In this appendix I estimate the response of insurance coverage to introduction of Part D for use in the calibration of Part D's costs in Appendix C. Estimation is based on the differences-in-differences design described in the first section of Appendix B3. The dependent variable is an indicator for prescription drug insurance coverage. Results are in table 19.

Table 19: Semi-Elasticity of Demand for Insurance with Respect to Medicare Part D

| Dependent Variable: | Full-Time Work |
|---|---|
| Post65*Post2006 | 0.13 |
| | (0.022) |
| Age and Year Indicators | Yes |
| Demographic, Health, and Individual Fixed Effects | Yes |
| N | 6,557 |
| Number of Clusters | 3,628 |

*Notes:* This table presents the effects of Part D eligibility on prescription drug insurance coverage. The sample is restricted to the treatment group: individuals with retiree health insurance only until age 65. Demographic controls include an indicator for being single, a set of indicators for each of the census divisions, and a fifth-order polynomial in non housing household wealth. Health controls include a set of indicators for self-reported health on a scale of 1-5; body-mass index; and a set of indicators for having any of the following physician-diagnosed conditions: cancer, lung disease, heart disease, stroke, arthritis, or psychiatric conditions. Robust standard errors clustered at the level of the individual are in parentheses.

The prescription drug insurance coverage rate increases by 13 percentage points upon Part D eligibility for the treatment group. The baseline insur-

[45] In the static model in Section I an individual with low disutility of labor is assumed to have a lifelong low disutility of labor. In a richer dynamic model individuals would have a time-varying willingness to pay.

[46] This calculation does not account for the cost of *raising* public funds, only the value of spending funds already raised by the government.

ance coverage rate for this group is 0.887. Thus the elasticity of coverage is $0.13/0.887 = 0.15$. This is a proxy for the parameter required in equation (9), assuming that everyone who buys insurance buys the average quantity of insurance.

# E Extensive Policy Change Model

The policy change considered in Section I is a marginal increase in subsidy for prescription drug insurance for retirees. However, the introduction of Part D was not an incremental increase in a subsidy but large change, from no subsidy at all to around $1,600 worth of subsidy per capita a year. In addition, Part D is more than just a subsidy; for example, it involved the creation of online "markets" to compare and select different plans. Individuals may value these miscellaneous changes apart from their valuation of dollars of subsidy.

The model in Section I also assumes some structure on insurance markets, and how individuals interact with them. This provides intuition regarding what might drive a valuation of the Part D subsidy above and beyond valuation of simple income. However, the fact of such excess valuation is not dependent on the specifics of the modeling assumptions made. In this appendix I present a simple variation of the model in Section I which allows for a discrete policy change, which is not necessarily denoted in dollars. Furthermore, I impose no structure on the insurance markets mechanisms underlying individuals' valuations of the subsidy.

This analysis finds the equivalent variation of the Part D policy change, with the only difference from the typical equivalent variation analysis being that it is measured in labor responses, which are then put into dollar terms. Such an approach is closely related to that described in Hendren (2016).

**Setup** Individuals have preferences over two goods, consumption, $c$, and some policy, $s \in \{0, 1\}$, as well as a disutility from labor, $v_i$.

$$U_i = u_i(c(l_i), s(l_i)) - v_i * l_i \tag{10}$$

Labor is once again modeled as an extensive margin decision, where $l_i = 1$ if individual $i$ works full-time, and $l_i = 0$ if not. Consumption is assumed to be equal to income in this static model and so is larger when working full-time than when not, $c(1) \equiv c_1 > c_0 \equiv c(0)$. Labor disutility is distributed according to a cumulative distribution function $G(v)$, with a probability density function of $g(v)$.

Furthermore, the policy is dependent on labor. Before the policy change $s(l_i) = 0$ for all $i$, and for any $l$. After the policy change retirees enjoy the policy while full-time workers do not: $s(1) = 0, s(0) = 1$.

**Optimal Labor Choice** Before the policy change individual $i$ works full-time if and only if:

$u_i(c_1, 0) - u_i(c_0, 0) \equiv \overline{v}_0 > v_i$

In other words, $i$ works full-time only if the utility from the added consumption of full-time work minus her labor disutility is larger than the utility of consumption from less than full-time work. This defines a labor disutility cutoff below which individuals work full-time and above which they do not.

Similarly, after the policy change individual $i$ works full-time if and only if:

$u_i(c_1, 0) - u_i(c_0, 1) \equiv \overline{v}_1 > v_i$

Here, too, there is a labor disutility cutoff below which individuals work full-time and above which they do not. $\overline{v}_1 < \overline{v}_0$ because the utility in the non working state is higher with the policy change.

**Analysis of the Policy Change** Define the change in utility when not working full-time due to $s$ as:

$$\overline{v}_0 - \overline{v}_1 = u_i(c_0, 1) - u_i(c_0, 0) \equiv \Delta u \tag{11}$$

This change in utility is precisely equal to the change in the labor disutility cutoff. Therefore the policy change will lead to a decline in labor supply associated with a decline in the cutoff labor disutility of full-time work. The change in labor supply associated with $s$ is therefore:

$$\Delta G(v) \equiv G(\overline{v}_0) - G(\overline{v}_1) = \int_{\overline{v}_0 - \Delta u}^{\overline{v}_0} G(v)dv \tag{12}$$

**Equivalent Variation Calculation** Consider a different policy change, which increases retirement consumption, $c_0$, to $\tilde{c}_0 = c_0 + \Delta c$:

As before, prior to the policy change individual $i$ works full-time if and only if:

$u_i(c_1, 0) - u_i(c_0, 0) \equiv \overline{v}_0 > v_i$

After the policy change individual $i$ works full-time if and only if:

$u_i(c_1, 0) - u_i(\tilde{c}_0, 0) \equiv \tilde{v} > v_i$

Where $\tilde{v}$ is the labor disutility cutoff when retirement consumption has been increased by $\Delta c$. As before, $\tilde{v} < \overline{v}_0$, this time due to the added utility of additional consumption in retirement.

Define the change in utility when not working full-time due to $\Delta c$ as:

$$\overline{v}_0 - \tilde{v} = u_i(\tilde{c}_0, 0) - u_i(c_0, 0) \equiv \widetilde{\Delta u} \tag{13}$$

As above, this decline in the labor disutility cutoff leads to a decline in the share of the population working full-time:

$$\Delta \tilde{G}(v) \equiv G(\overline{v}_0) - G(\tilde{v}) = \int_{\overline{v}_0 - \widetilde{\Delta u}}^{\overline{v}_0} G(v)dv \tag{14}$$

**Claim.** If $\Delta c$ is such that $\widetilde{\Delta u} = \Delta u$ then: 1) Individuals value $\Delta c$ precisely as much as they value the policy $s$; 2) $\widetilde{\Delta u} = \Delta u$ if and only if $\Delta \tilde{G}(v) = \Delta G(v)$.

**Proof.** (1) follows immediately from the definitions in equation (11) and equation (13). (2) follows immediately from the definitions of equation (12) and equation (14).

This shows that if we choose $\Delta c$ such that $\Delta\tilde{G}(v) = \Delta G(v)$ then we will have found the equivalent variation of $s$ such that individuals value $s$ as much as they value $\Delta c$.

**Calibration** Section IV estimated that $\Delta G(v) = 0.0836$. As described in Section V, Gelber, Isen and Song (2016) found that \$6,126 increase in Social Security leads to a decline in participation of 0.004. Thus the equivalent variation of Part D is \$128,000; i.e., Part D is valued as another \$128,000 of lifetime discounted (annually at 3 percent) Social Security wealth. If we further assume the sum of the policy change implicit in Part D is the additional subsidy to prescription drug insurance, the same calibration used in Section V can get us that the monetary value of Part D is \$25,000. Therefore willingness to pay for one dollar of the subsidy can be calibrated as $\frac{128{,}000}{25{,}000} \approx 5.12$, as in Section V.