Chaire de recherche
sur les enjeux économiques
intergénérationnels

21-03

Stress and Retirement

Raquel Fonseca, Hugo Morin and Ana I. Moro-Egido

Cahier de recherche
Working paper
Mars / March 2021

HEC MONTRÉAL ESG UQÀM

Chaire de recherche
sur les enjeux économiques
intergénérationnels

est une chaire multi-institutionnelle qui s'appuie
sur un partenariat avec les organisations suivantes :

CIRANO — Allier savoir et décision — Centre interuniversitaire de recherche en analyse des organisations

Retraite Québec

Dépôt légal : Bibliothèque et Archives nationales du Québec et Bibliothèque et Archives Canada, 2021. ISSN 2563-0008

Stress and Retirement

Raquel Fonseca* Hugo Morin† Ana I. Moro-Egido‡§

March 8, 2021

Abstract

This paper investigates the bi-directional causal relationship between retirement and stress. We use PSID data for the period 2007-2015. Using a simultaneous equations approach, we find that a rise in stress increases the probability to retire by roughly 15.4 percentage points, while retirement decreases stress by 34.5 percentage points. We find the same results when we disaggregate by individuals' characteristics, but the former effect is stronger for males, for people working in typical blue-collar jobs, and for people whose wealth is below the mean; while the latter is stronger for males, for white-collar workers, for people whose wealth is above the mean, and for white individuals. We show that official retirement ages are a strong instrument for actual retirement age, and that lagged physical activity levels are a non-linear instrument for perceived stress. We also confirm that objective measures of mental health are a strong instrument for perceived stress.

JEL Classification: C30, I10-12, J26

Keywords: Stress, retirement, physical activity, simultaneous equation models.

*Corresponding author. ESG-Université du Québec à Montréal & CIRANO. Email:fonseca.raquel@uqam.ca

†ESG-Université du Québec à Montréal and ENAP

‡University of Granada

§We thanks Kathleen Mullen for her useful comments. This research is also part of the program of the Research Chair in Intergenerational Economics. Errors are our own.

1 Introduction

Many factors may be associated with one's retirement decision. While financial incentives, such as public or private pension eligibility, come to mind as the most obvious determinants, research has been pointing to the importance of non-financial factors as well. Among others, physical health is found to be an important predictor of one's retirement decision: ill physical health or a negative health shock, for instance, increase the probability of a worker leaving the labor force (Blundell et al., 2017, Disney et al., 2006, Dwyer & Mitchell, 1999, Macken, 2019, McGarry, 2004, Quinn, 1977). Some other factors may be associated with one's decision to retire such as working conditions and job strains (Blekesaune & Solem, 2005, Sutinen et al., 2005, Wahrendorf et al., 2012), leisure interests (Shultz et al., 1998), work-life balance (Angrisani et al., 2017) or working effort-reward imbalance (Hoven et al., 2015, Wahrendorf et al., 2012).

However, some other studies have focused on the opposite question, that is, whether retirement has an effect on either physical or mental health. They have shown that transition from the labor force to retirement will sometimes be associated with a change in one's mental health and well-being, and that the effect might become stronger or weaker with time spent out of the labor force (Binh Tran & Zikos, 2019, Butterworth et al., 2006, Charles, 2004, Coe & Zamarro, 2011, Dave et al., 2008, Eibich, 2015). All in all, the literature has yielded important findings on the relationship between health and retirement, but the results remain mixed as to the direction of this relationship.

The aim of this paper is to disentangle the possible direction of causality between mental health and retirement. When analyzing this question, we emphasize one aspect of mental health, stress[1]. Using PSID data for the period 2007-2015, and focusing on respondents who are close to the normative retirement age, we attempt to determine whether stress induces workers to retire earlier, or if it is rather that early retirement provokes stress in new retirees. In contrast with the previous literature, we perform a simultaneous equation analysis where both outcomes (retirement and stress) are instrumented. This allows us to determine which effect prevails after controlling for the other. Our approach follows that of Charles (2004) and Fonseca et al. (2017), who use simultaneous equations to analyze the relationship between retirement and a general measure of well-being. Another contribution is the choice of the instrument for stress, which is physical activity. It has been proven to be a mediator between retirement status and physical health (Binh Tran & Zikos, 2019, Eibich, 2015).

[1]Michaud et al. (2016) deemed the effect of stress caused by income shock and job loss on mental health and well-being.

Our results are unequivocal. There is a bidirectional effect between retirement and stress, but it is stronger in one direction. In particular, we find that a rise in stress increases the probability to retire by roughly 15.4 percentage points, while retirement decreases stress by 34.5 percentage points. We find the same results when we disaggregate by individuals' characteristics, but the former effect is stronger for males, for people working in typical blue-collar jobs, and for people whose income is below the mean; while the latter is stronger for males, for white-collar workers, for people whose income is above the mean, and for white individuals.

Additionally, we also find that our instruments are strong. Being above full retirement age affects positively the probability of retirement, and the magnitude of this effect (around 20 p.p.) is larger than the one we found for being merely above early retirement age (around 11 p.p.). Physical activity is a non-linear predictor of the level of stress, but of a lower magnitude than mental health which exerts a stronger effect (3.2 p.p. at the its highest effect compared to 26.6 p.p.).

The paper will be structured as followed. Relevant studies are going to be presented in section 2. Section 3 will present data. The econometric model will be addressed in section 4 and the estimation results will be found in section 5, with robustness checks in section 6. Finally, section 7 will discuss the results and conclude.

2 Selected Studies

In this section, we will detail several studies onto which the analysis is based. Globally, the focus is on the literature investigating the effect of health on retirement decisions, and the relationship of retirement and mental health.

We first review key studies analyzing the effect of health on retirement. Health is clearly one of the most important non-financial determinants of retirement plans or decision. One's health condition may directly affect one's working capacity and push someone out of the labor force even if that person would rather have stayed at work for financial reasons or by simple preference. An early paper investigating this question is that written by Quinn (1977). Using the Retirement History Study and focusing on white married men between 58 and 63, the author finds that health and pension eligibility are clear determinants of early retirement, among a large set of other covariates. More specifically, health was found to have the greatest impact; even greater than pension eligibility, and sharply greater than wage and income. Similarly, Dwyer & Mitchell (1999) insisted on the unimportance of economic variables in comparison with health variables. The authors used men from 51 to 61 in the 1992 Health and Retirement Study (HRS) and estimated a model of retirement explained by retirement income, health and other covariates. It was estimated that men living with a low level of overall health will retire earlier. More recently, Disney et al. (2006) found that a higher level of health stock impacted positively the probability to stay in the labor force and the results were robust to the tackling of the endogeneity problem. McGarry (2004) reached similar results, focusing on the retirement at the age of 62. Using the first and second waves of the HRS data and the expected probability to continue working at the age of 62 and 65, the results indicated that a fair or poor subjective health is negatively correlated with labor participation. Moreover, the plan to quit labor force is more impacted by changes in health than changes in financial variables. More recently, Blundell et al. (2017), using the HRS and the English Longitudinal Study of Ageing (ELSA), found that, overall, declines in health explain up to 15% of the decline in employment between ages 50 and 70. The effects drop with education and are larger in the US than in England. Macken (2019), using the longitudinal Survey of Health, Aging and Retirement in Europe (SHARE) for Germany, shows that three health outcomes, like self-rated health, depression, and cardiovascular disease, can be caused by work stress and thus may lead to a lower retirement age.

As revised in French (2005), Blundell et al. (2016), and French & Jones (2017), the main mechanisms through which health affects labor supply are of different nature. First, illness can make work extremely unpleasant. Second, bad health can also reduce worker productivity. People in sufficiently bad health may receive disability benefits, and people receiving disability benefits are usually prohibited from working. Third, with shorter expected lifespans, individuals in bad health may not need to work as long in order to accumulate financial and pension wealth for retirement. Fourth, health may also affect labor supply through medical expenses. Because many US workers only receive health insurance while they continue to work, expensive medical conditions may lead them to delay retirement. Alternatively, uninsured workers may leave their jobs in order to qualify for health care provided through disability or means-tested social insurance.

In this literature the distinction on both objective and subjective health variables is important. The objective and subjective measures do not contain the same information and will affect the retirement decision differently. For example, we could take two people with the same objective ill health, but one not being limited by this condition and, therefore, not perceiving it as a problem, while the other one feeling the effects of that condition. The effect on labor supply choices will be different between these two. On the opposite, someone without any objective (or labeled by doctors) health problems, but having a subjective feeling that his health is problematic, will probably retire earlier than someone in the same condition but without this felt limitation. Moreover, subjective measures have a stronger impact on retirement (Dwyer & Mitchell, 1999).

If the literature has shown that health may impact retirement decisions, previous research has also shown that retiring may, by itself, have an impact on health, either physical and mental health. Recent literature has shown mixed findings regarding the effect of retirement on both subjective (e.g., physical and mental health) and objective (e.g., limitations in activities of daily living, mortality) health. Focusing the review on mental health, it is common to think of retirement as a period of relative stress-free enjoyment following a stressful professional life. But in reality, retirement appears not always to relieve stress. For many, retirement is a period of mental health issues. Surely, aging comes with its challenges, and more physical and mental incapacity may emerge as years pass. The question that arises is whether retirement itself is causing any degradation of mental conditions, apart from the effect of aging, or not.

Several studies report a positive effect of retirement on health (Coe & Zamarro, 2011, Eibich, 2015, Insler, 2014, Neuman, 2008, Zhu, 2016) but other studies find negative effects (Behncke, 2012, Dave et al., 2008). There is also evidence that retirement may have a negative effect (Bonsang & Perelman, 2012, Rohwedder & Willis, 2010) or little effect (Coe & Zamarro, 2011) on cognitive abilities. Concerning the size of the effect, Dave et al. (2008) estimated that within a period of six years after retirement, respondents of HRS saw a decline of 6 to 9 percent in their mental health condition. Similarly, Binh Tran & Zikos (2019) observe that retirement has a positive effect on self-reported health. The estimated coefficient appears to be sizeable, as retirement leads to an approximately 28 percentage-point increase.

The mechanism behind the negative effect rely on the fact that the decline in their mental health condition may be caused by a drop in physical activity and social interactions. In other words, life slows down and this has an effect on one's physical and mental capacities. As found in Dave et al. (2008), these results were robust to the control for job conditions and several lifestyle variables. An individual fixed effect modelization was added and the relation between full retirement and mental health incapacity is argued to be a causal mechanism. But the effect is not that clear: early retirees could be more likely to suffer from mental health problems than the workers in the same ages, but this relationship does not seem hold when using older subsamples (Butterworth et al., 2006).

There is also a literature in economics inspired by biological theories that focus on the potential negative health effects of the cumulative adverse mental effects (allostatic load) from stress (Seeman et al., 1997). In particular, Michaud et al. (2016) study how different labour market shocks, such as income variation and job loss, among other factors, affect health through the stress they cause. Yet, the effect of stress, as a source of bad health, on retirement decisions, has been poorly treated in the literature.

3 Data and Main Variables

The main dataset we use is the Panel Study of Income Dynamics (PSID) for the selected period 2007-2015. We use the Main Family Data, the PSID Individual-Level Data taken from the raw data, and PSID data prepared by the Cross-National Equivalent File project at the Ohio State University.[2,3] We focus on older ages by restricting the sample to respondents aged between 50 and 80. The average age is roughly 60 years old.[4] In addition, we discard people who are self-employed (totally or partly) since they are known to have different behaviors regarding retirement (Blau, 1987).

3.1 Main Variables

We present here the main definitions of our two variables of interest, stress and retirement. We begin by defining the variable *Stress*.[5] The PSID data has a variable that would be most suitable to characterize the stress felt by an individual: the question asked is *"How often do you feel rushed or pressed for time"*. This variable captures one's feeling of time pressure and, as older workers may feel that working is taking too much time in their daily schedule, they could choose to retire to get more time and lower this pressure. Unfortunately, this variable is only available for the year 2003. In order to use the panel structure of the data, we look for a variable that is available for the whole studied period.

There is a set of variables evaluating the feelings about some mental conditions of the respondents. These variables represent one's feeling of restlessness, nervosity, sadness, hopelessness, worthlessness and that everything is an effort. For example, for each feeling,

[2] For details, https://psidonline.isr.umich.edu/Guide/default.aspx, and https://cnef.ehe.osu.edu/data/

[3] The PSID is a longitudinal dataset with a rotating panel design, and representative household panel survey for the US population. It is collected in a biennially from 1997, and low attrition, around 10000 families and 24000 individuals nowadays.

[4] We cut off between 50 and 80 for the availability of number of observations. In particular only 1.29% of those younger than 50 are retired, similar figure for those beyond 80. We have also considered the interval 45-85 in the robustness section 6.

[5] Some alternative definitions have been included into the robustness analysis in section 6.

it is constructed using respondents' answer to the questions *"In the past 30 days about how often did you feel restless or fidgety"*. The answers are self reported on a five values scale: none of the time; a little of the time; some of the time; most of the time; all of the time. We check the correlation of all these variables with the stress variable, presented in previous paragraph for the 2003 sub-sample when all of them are available (Table 1). The *nervous* and *restless* variables are strongly and positively related with the *rushed for time* variable, particularly the nervous feeling. Both *nervous* and *restless* are statistically suitable proxies for the variable.

Table 1: Regression of the stress variable on several mental health conditions, 2003 sub-sample

	Feeling rushed for time
Sadness	-0.111**
	(0.003)
Nervous	0.259***
	(0.000)
Restless	0.121***
	(0.000)
Hopeless	0.107*
	(0.037)
Everything's an effort	0.097**
	(0.001)
Worthless	-0.152**
	(0.002)
Constant	1.511***
	(0.000)
Observations	2495

p-values in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Authors' calculations using the PSID dataset (2003).

Giving these results, we build an index with both *nervous* and *restless* as the stress proxy along the period 2007-2015.[6] The indicator is a dichotomous variable taking the value of one if, for at least one of these two variables, the respondent reports any feeling of nervosity and/or restlessness (28,86% of respondents); and zero otherwise.

[6]These two variables are unavailable in 2005. Given that some key variables are built with lag considerations, we do opt for keeping the length of the panel constant, Therefore we only consider the period 2007-2015. We have included waves 2001 and 2003 in the robustness analysis in section 6.

Our other key variable is the retirement status (*Retirement*). We follow Rohwedder & Willis (2010) in defining retirement as a state that captures two conditions: the person is "out of the labor force"; the person has reached the age of retirement (50 to 80 years old). To build our variable, we use the self-reported retirement variable available in the PSID data. The question asked in the survey is the following: *"We would like to know about what you do. Are you working now, looking for work, retired, keeping house, a student,.. or what?"*. The dummy variable takes a value of one if the respondents said they were retired or keeping house, and they did not work any hours. Conditional on working in the previous period, they were considered to be retired. The variable takes a value of zero if individuals were working. This retirement definition will be the main one used in the analysis. On average, in our sample, 23% of respondents are retired.

3.1.1 Stress, Retirement and Age

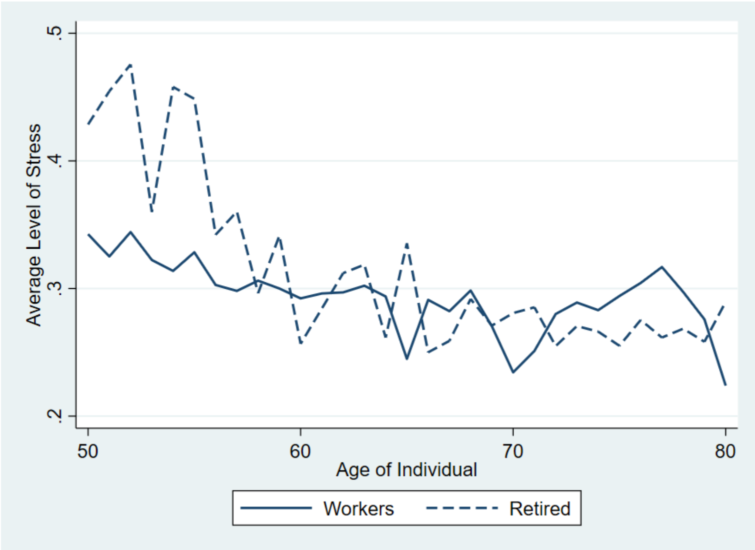
The relationship between stress and age in our sample is reported in Figure 1. It shows the relationship for two groups: retirees and workers.[7] Taking both groups into consideration, we can see a clear negative relationship. As they get older, respondents report lower levels of stress.[8] An inevitable point to be made here is the difference between workers and retirees who are younger than 60. In this age range, retirees appear to have a higher level of stress than working people. After age 60, the evolution of stress is very similar between workers and retirees until age 70, where the stress felt by workers is slightly higher. As no control is used in this visual representation, it is impossible to argue a causal relationship. Yet, if retirement is sometimes said to be a period free of worries, our respondents' answers do not reflect this idea.

Moreover, an earlier retirement is correlated with a higher level of stress. For example, a 55-year-old worker has an average level of stress of 0.32, when a retiree of the same age reports a stress of 0.45. Early retirees seem the ones experiencing the higher stress. The question still remains: are these people stressed because they retired too early (and, for instance, found themselves without an environment in which to socialize, or encountered unexpected financial problems) or has stress caused them to retire earlier? The modeling and estimations presented further will attempt to provide an answer by using the simultaneity of those two outcomes.

[7]We only consider, at each age, the ones who report working or being retired according to our definition of retirement.

[8]Indeed, if we include in the analysis everyone from 20 years old to 80 years old, there is a strong negative relationship with age. There seems to be a peak of stress at around the age of 25 to 30 and, thereafter, a decline of stress.

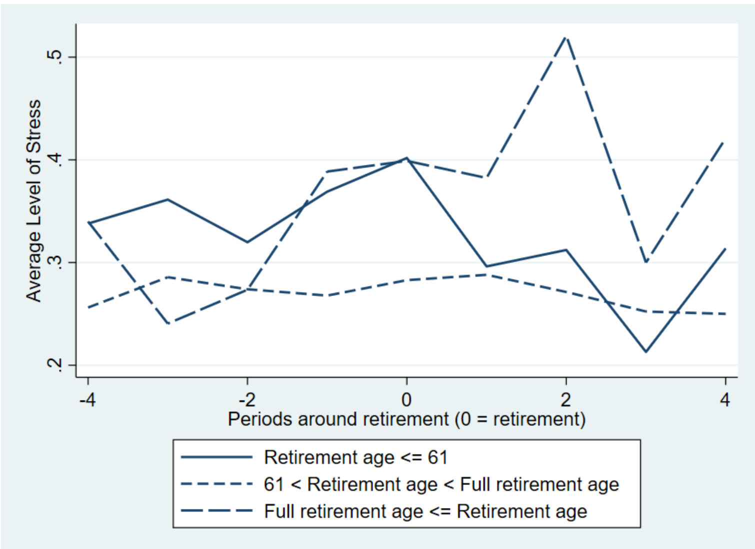
Figure 1: Relationship between stress and age, population between 50 and 80, by work status



Own calculations: PSID dataset (2007-2015).

Before getting into this analysis, we present in Figure 2 the average stress level reported at retirement as well as during the four periods preceding and following retirement for those who retired between 50 and 80 years old. Periods are defined as two years, the length of PSID panel data, and period 0 is the retirement year. The sample has been

Figure 2: Relationship between stress and period around retirement, by retirement age subgroups



Own calculations: PSID dataset (2007-2015).

divided into three different groups, depending on the age of retirement. The first group (solid line) is early retirees, i.e. respondents who retired before 61 years old. For this group, stress seems to be increasing as retirement approaches, with a peak of stress felt at the retirement period. Once respondents retire, the level of stress begins to decrease, roughly at the same rate as it was growing before. We could understand that retirement is particularly important for this group, as it seems to be highly correlated with a shift in the stress trend over the periods. The second group (dash line) retired between 61 and the normal retirement age (which depends on the birth year in the USA). These retirees quit the labor force earlier than the full retirement age, but still later than the early retirement threshold. Their level of stress does not seem to be correlated with retirement, being relatively stable over the periods around retirement. The third group (long dash line) shows yet another pattern of stress around retirement. These are the ones who retired after the full retirement age. They seem to experience a drastic rise of stress before retirement and, after a last bump on the period right after retirement, stress seems to be capped around a level of 0.5. Thereby, retirement and stress are correlated in this subgroup, as respondents experienced an increasing stress that is leading retirement decision, but retirement does not lower the stress level.

3.2 Heterogeneity

We find heterogeneity in the stress-retirement relationship, as it differs by gender, ethnicity, occupation and wealth (Figure 3). We define gender as a boolean variable that equals one if the respondent reports being a woman. We also consider a distinction by race (White vs. Non-white individuals). We include some information on the type of occupation by using a boolean variable that takes the value of one if the respondent is in a white collar occupation.[9] Finally, to the extent that wealth provides more accurate information on long-term living conditions, it is well recognized as a determinant of retirement but also as a determinant of self-reported health. We thus create a variable to identify individuals whose level of wealth is above (or equal), or below the median.

The top left panel informs us that male and female follow a similar trend of stress around retirement, with female revealing a slightly higher level of stress. There is nothing here that could lead us to conclude that a different level of stress in men and women leads to differentiated patterns in retirement. In the top right panel, the sample is divided by ethnicity, i.e. white and non-white people. It is often considered as another division upon which discrimination is based. Here, we see that non-white people report a higher level

[9] Following Mazzonna & Peracchi (2012), the blue collar occupations include the primary or the secondary sectors, or working in fields that consist of manual tasks, and the white collar occupations from the service and administration sectors. Military personnel was taken out of the sample.

of stress. Moreover, this level of stress does not decrease at all after retirement. Rather, it appears to be the opposite. In the bottom left panel, the sample is divided between blue and white collar workers. Here again, we see that the group that is considered less advantaged reports a higher level of stress. The bottom right panel shows a large difference in stress by wealth. Essentially, those living below the median report a sharply higher stress level overall. Those at or above the median have a low, stable level of stress. For the purpose of graphical simplicity, we kept only these two wealth subgroups. But dividing the sample in smaller slices (not shown) exacerbates the results, and shows that poorer people live with a higher stress level.

Figure 3: Relationship between stress and period around retirement, by subgroups

Authors' calculations: PSID dataset (2007-2015).

4 Empirical Strategy

As shown in Section 2, many authors have found that one's health condition is one of the most important non-financial factors which determine retirement decisions, but on the other hand, retirement has also been found to be a period of mental health issues. Hence, when studying the relationship between stress and retirement, it appears important to take into account the simultaneity of causality. This is what we attempt. Our decision is motivated by the evidence of simultaneous causality shown in the literature review, but also by the insights gained from the descriptive analysis presented in Section 3. To model the relationship between our two variables of interest properly, we perform a simultaneous equation analysis, where both outcomes (retirement and stress) are instrumented. The model is written as such:

$$\begin{aligned}
Retirement_{i,t} &= \beta_0 + \beta_1 Stress_{i,t} + \beta_2 I_{i,t} + \beta_3 X_{i,t} + \delta_t + \xi_{i,t} && (1)\\
Stress_{i,t} &= \gamma_0 + \gamma_1 Retirement_{i,t} + \gamma_2 Physical\ activity_{i,t-1} \\
&\quad + \gamma_3 Physical\ activity^2_{i,t-1} + \gamma_4 O_{i,t} + \gamma_5 X_{i,t} + \delta_t + \mu_{i,t} && (2)
\end{aligned}$$

where

- *Retirement* is individual i's retirement probability in period t, conditional on not being retired in the previous period ($Pr(Retirement_{i,t} = 1 \mid Retirement_{i,t-1} = 0)$). It takes the value of 0 when a respondent is not retired and 1 from the moment the respondent is retired;
- $Stress_{i,t}$ is respondent i's self-reported level of stress in period t;
- $I_{i,t}$ is a matrix of institutional variables regarding retirement age;
- $O_{i,t}$ is an objective measure of mental health condition;
- $Physical\ activity_{i,t-1}$ is the weekly average number of heavy physical activities during the previous period;
- $X_{i,t}$ is a control vector;
- δ_t is a period fixed effects;
- $\xi_{i,t}$ and $\mu_{i,t}$ are error terms.

If we had considered only equation (1), as has been often done in the literature, one might have easily thought that unobserved variables could have affected the effect of stress onto one's retirement decision, causing an endogeneity problem (*i.e.* $corr(Stress_{it}, \xi_{i,t}) \neq 0$). On the other hand, looking at equation (2), we could also argue that retirement can affect stress. Thus, to solve this potential problem of endogeneity and reverse causality, we have opted for simultaneous equations, instrumenting both outcomes (retirement and stress). The specific instruments are described below in subsection 4.2.

Additionally, to make use of the panel structure of the data set, the estimation as shown above includes dummies for time. With this in mind, we allow individual correlation over time by splitting the error term in equation (1) as $\xi_{i,t} = \epsilon_i + \varepsilon_{it}$, where ϵ_i is the individual time-invariant effect and ε_{it} is an independent error term with $\epsilon_i \sim N(0, \sigma_\epsilon^2)$, $\varepsilon_{it} \sim N(0, 1)$ and $Cov(\epsilon_i, \varepsilon_{it}) = 0$.[10] In this setting it is assumed that the error terms are random and not correlated with the observable explanatory variables. It may not be plausible given the potential correlation between individual unobserved characteristics and the explanatory variables. The relevant and widely-used solution to address this issue would be that proposed by Mundlak (1978), as justified in Ferrer-i-Carbonell & Frijters (2004) or Ferrer-i-Carbonell (2005). With this method, the individual random effect ϵ_i is decomposed into two terms: (i) the pure error term, u_i, which is normally distributed with zero mean and independent of the idiosyncratic error ε_{it}; and (ii) a part that is correlated with a subset of observable time-varying regressors, ω_{it}, with correlation $\overline{\omega_i}\eta$, where $\overline{\omega_i}$ is the average of ω_{it} across time. The subset of variables ω_{it} includes variables that vary across time such as family income, years of education, and members of the household.

The panel structure of our data also allows us to control for a period fixed effect (δ_t). Labor force participation is surely related with macroeconomic idiosyncratic shocks that could constraint people to stay in the labor force longer than they would have preferred. This control captures this macroeconomic context and permits us to focus on individuals' characteristics. In a similar vein, those variables also capture how such characteristics affect stress in a different manner than individual characteristics do.

It is important to underline a computational limitation. The estimations did not include a nonlinear relation between our dependent variable and the explanatory ones. As the retirement variable is dichotomous, we should have used a *probit* or a *logit* estimation to bound the estimated coefficient between 0 and 1. This was not possible with the *Three Stages Least Square* command that we used, (*reg3* in *Stata*). We tested our results by transforming the stress variable in a dichotomous one too (i.e., a new variable that equals to one if the actual variable takes the value of two or three and zero otherwise). As the dependent and the explanatory variables are dichotomous, we avoid this problem. To confirm the soundness of our approach, we have also carried on a robustness analysis with an alternative measure of stress as a 3-level category variable in Section 6.

[10] Equivalently in the other equation.

4.1 Explanatory variables

Concerning the specific variables for our analysis, we start describing the standard control variables included in vector X_{it}. We relegate the detailed description of the instrumental variables ($I_{i,t}$ in equation (1) and *Physical activity*$_{i,t-1}$ and $O_{i,t}$ in equation (2), respectively) to the following subsection.

Beside the variables included in Section 3 for the heterogeneity analysis, we also include in vector X_{it} the following standard controls. Marital status is a dichotomous variable that equals to one if the respondent is married/in couple and zero if not (whether he is single, separated or divorced). The education variable has 5 categories: lower than High School; High School; Some College, no degree; B.A.; and higher than B.A. Health is captured by two dichotomous variables. The first one equals to one if the respondent has any major health condition (cancer, stroke, heart attack or lung disease) and zero otherwise. The second one equals to one if the respondent has any minor health condition (arthritis, hypertension or diabetes). This strategy was previously used by Fonseca et al. (2017) and seems appropriate, as one's retirement decision is based on major health condition (Bound, 1991, Bound et al., 1999, Disney et al., 2006, Dwyer & Mitchell, 1999, McGarry, 2004). Moreover, minor conditions could have an effect on its own that could be related to the respondent's level of stress. Finally, we include the size of the household the individual lives in. For main descriptive statistics, see Table A.1 in the Appendix.

4.2 Instrumental Variables

Regarding the retirement decision (equation (1)), we include two exogenous institutional variables concerning retirement age. In particular, in matrix $I_{i,t}$ we include two dichotomous variables indicating whether the individual is eligible for full or early retirement public pensions using cohort and gender-specific pension eligibility ages. The official retirement ages refer to the law that was in place when respondents in PSID were facing retirement decisions. In the US, the normal retirement age depends on the birth year (for example it is 67 for persons born before 1959), and the early retirement age is 62. In our sample, 13.8% and 23.5% of respondents are older than the early and full retirement age respectively. This strategy has been used often in the literature, and these instruments have been shown to be very strong predictors of retirement (see e.g. Charles, 2004, Coe & Zamarro, 2011, Fonseca et al., 2017, Rohwedder & Willis, 2010). For detailed information on the early and normal statutory retirement ages by cohort and gender[11].

[11] See https://www.ssa.gov/oact/progdata/nra.html for normal retirement age regulation.

When we regress a linear probability model for retirement, the dummies indicating that an individual's age is above or below a statutory retirement age are positive and significant.

Regarding equation (2), we consider two different variables. Our perceived stress variable will be instrumented by respondents' level of physical activity on the previous period ($Physical\ activity_{i,t-1}$). The idea is that heavy physical activity on a regular basis helps to temper depression and anxiety (Carek et al., 2011, Dunn et al., 2001, Ströhle, 2009) and helps mental well-being in general (Fox, 1999, Windle et al., 2010). In addition, we used the lagged form of this variable to make sure there is no correlation with contemporaneous retirement decision. We cannot know when the person retired during the period, as PSID data are structured in waves of two years. It is possible that a respondent has retired the day after the previous wave questionnaire (period $t-1+1$ day), or the day just before the present period questionnaire. Between both of these extreme cases, the available time for physical activity could be sharply different. This could lead to an erroneous endogeneity correction, as physical activity and retirement could be related. Using the lag prevents this problem from arising. As far as we know, physical activity has never been used as an instrument to capture the endogeneous problem of stress, or any mental or physical health condition. This is an intuitively and, as we will see in the next section, statistically suitable instrument. Binh Tran & Zikos (2019) and Eibich (2015) found that the relationship between retirement status and physical health is mediated positively, among other things, by physical activity.

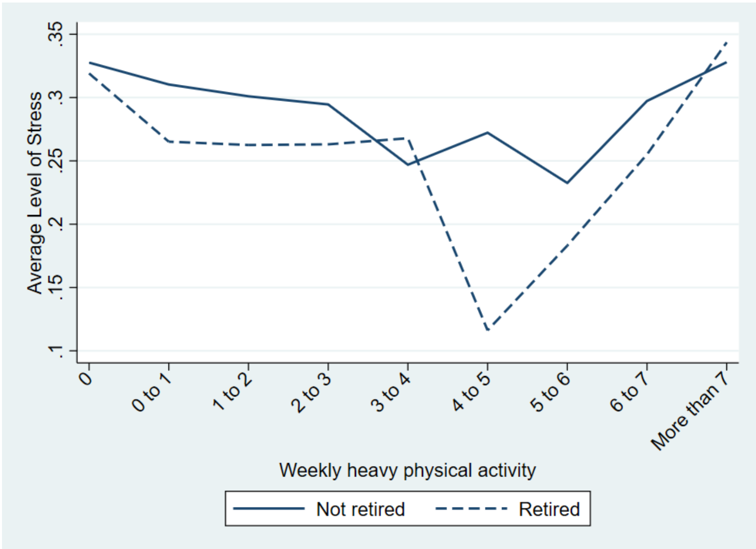
PSID contains information on how often respondents do *vigorous* physical activities for at least 10 minutes that cause heavy sweating or large increases in breathing or heart rate measured in number of times and in time unit (day, week, month, year). We build a 8-scale variable that ranges from 0 heavy physical activities per week to 7 or more activities per week. In our sample the average number of activities per week is 2.47.

We present this relationship between heavy physical activity and stress in Figure 4 for workers and retirees. First of all, taking both groups, a convex relation can be stated between those two variables. We see that a higher level of heavy physical activity is linked with a lower level of stress, with a floor at around 4 and 6 times a week, and then a rise back for those doing more exercise. This non-linear form will be used is the subsequent analysis. Windle et al. (2010) found that training programs help to raise the level of well-being of sedentary people. We find similar results here. Starting from the sedentary point (i.e. 0 exercise per week), a rise in the number of times one practices a physical activity is linked with a lower level of stress. Given that, we choose to introduce the variable $Physical\ activity_{i,t-1}$ in a non-linear way (i.e., we add the squared variable)

But the effect seems to depend on how many times the individual actually exercises. An incentive to raise the frequency of exercise in people who are already really active could have a negative effect.

Focusing on retirees, we confirm what has been observed before: on average, retirees report a higher level of stress, but with a slightly lower floor. The relationship between those two variables seems to be more important for them than for workers. This is intuitive: workers may see their stress level affected by other factors, such as work environment and engagement, onto which physical activity could have a lighter effect. But, still, the convex form of the relation is seen.

Figure 4: Relationship between stress and physical activity, by working status



Own calculations: PSID dataset (2007-2015).

In addition, a contemporaneous objective measure of mental health ($O_{i,t}$), following Benítez-Silva et al. (2004), Bound (1991), Disney et al. (2006), Dwyer & Mitchell (1999) will be used as an instrument. It reports whether the respondent "*has or had psychiatric problems*" (anxiety, depression or nervosity) and is part of the questionnaire's section on medical conditions. It will be used as an objective condition affecting one's self-reported, or perceived, level of stress. A lagged objective mental health problem could affect today's capacity to work, not necessarily by passing through today's subjective condition, which is consistent with the literature. In our sample, 7.9% reports such problems.

5 Results

This section presents the estimation results of our modeling. First of all, the straightforward model is going to be presented. We show what is the average relationship between our variables of interest in the whole sample. Afterwards, we divide the sample in selected groups to analyze whether the effect found is stronger for some specific individuals, whereas it would be less significant for others. The relationship between stress and retirement is surely not homogeneous and depend on one's characteristics. Finally, the model has been run with several similar dependent and explanatory variables to test the robustness of our estimated coefficients, which we relegate to next section.

The estimations of our structural equation modeling is to be found in Table 2. The table is horizontally divided in two sections. The upper one presents the results of the equation (1), i.e. the determinants of the retirement probability. The lower part reports the estimates of the equation (2), estimating the determinants of the stress level. As both outcomes explain one another in a simultaneous equation, we find the stress level's effect estimated in the retirement probability equation and vice versa.

The columns (1) to (4) introduce the control variables gradually. As we see, the inclusion of every control variable does not change substantially the estimated coefficients. Stress is found to have a strong effect on the retirement decision. A marginal increase in the stress variable raises the probability to retire by 31.4 percentage points (p.p.) in the baseline model (1), and by 15.4 p.p. in the model including all controls (4). This effect is similar in magnitude to what has been previously found for physical health (Disney et al., 2006, Dwyer & Mitchell, 1999, McGarry, 2004, Quinn, 1977). Regarding the effect of retirement on stress, we find that the choice to leave the labor force does reduce likelihood of stress by 34.5 p.p. These two findings are consistent with patterns shown in Figure 2. The level of stress is high in periods before retirement and it did seem to decrease after retirement, even after several years.

Taking a look at the instruments used for the retirement equation, we see that being older than the early retirement age (62 years old) raises significantly the probability to retire (about 19.3 p.p.). Meanwhile, being older than the full retirement age (which depends on birth year in the USA) has a weaker but still significant effect (11.3 p.p).[12]

On the stress equation side, there is a strong nonlinear relation between the stress and

[12] Test on differential effect has been made. The difference is 0.079 p.p. and is significantly different from zero. It is important to note that each of the legal retirement ages' effect are not significantly different from stress' effect on retirement.

physical activity, in line with Figure 4. If this relation is intuitive (heavy physical activity lowering stress levels), it is a statistically strong instrument as well. We find that physical activity decreases by 1.6 p.p. the probability of having stress, and the effect is increasing with the level of physical activity (estimated parameter of the squared variable is positive). In particular, and confirming Figure 4, the highest effect arises with a frequency of physical activity 4 times a week (around 3.2 p.p.), and then a rise back for those doing more exercise.[13]

Objective mental condition is also found, once more, to be a statistically strong instrument, consistently with what has been found in the literature (Benítez-Silva et al., 2004, Bound, 1991, Disney et al., 2006, Dwyer & Mitchell, 1999). The corresponding effect is an increase of 26.6 p.p. of the probability of being stressed. The effect on stress is significantly different among the three variables (retirement, physical activity and mental health).[14] As reported in Table 2, the null hypothesis that our instruments are valid cannot be rejected. Or put another way, the test gives no reason to doubt the validity of these instrumental variables.

Considering now the other explanatory variables, we find that gender, working or not in a white-collar occupation and marital status affect similarly retirement and stress.[15] However, we find that race exerts a stronger effect (of 5.6 p.p.) on the retirement decision than on the stress level. The same pattern is found regarding wealth, and such differences increased with wealth (from 4.9 p.p. at the bottom quintile to 12.9 p.p. at the top quintile). Major and minor physical health condition are more relevant for stress determination than for retirement, as well as household size.

In sum, we have found that retirement affects negatively the level of stress, but also that stress levels affect positively the decision to retire. The former is larger in absolute value (19.1 p.p.) than the latter. The variables we chose to instrument retirement are strong, that is, being above the full or early retirement age affects positively the probability of retirement. As expected, the effect of being above full retirement age is larger. We also find a strong effect of instruments for stress, mainly through objective mental health and in a smaller, but non-linear, way, through lagged physical activity.

[13]The quantitative effect of physical activity does not change dramatically when we exclude objective mental health from stress equation, displaying a maximum effect of 6.6 p.p.

[14]All differences are tested and are significantly different from zero.

[15]The test is not significant at the 5% level

Table 2: Estimation of the effect of stress on retirement decision, structural equations modeling

	(1)	(2)	(3)	(4)
Retirement				
Stress	0.314**	0.334*	0.145***	0.154***
	(0.003)	(0.015)	(0.000)	(0.000)
Age Above Full Retirement Age	0.210***	0.209***	0.200***	0.193***
	(0.000)	(0.000)	(0.000)	(0.000)
Age Above Early Retirement Age	0.125***	0.124***	0.118***	0.113***
	(0.000)	(0.000)	(0.000)	(0.000)
Stress				
Retirement	-0.294***	-0.288***	-0.280***	-0.345***
	(0.000)	(0.000)	(0.000)	(0.000)
Physical activity, t-1	-0.032***	-0.026***	-0.021***	-0.016**
	(0.000)	(0.000)	(0.000)	(0.004)
(Physical activity, t-1)^2	0.004***	0.003***	0.003***	0.002**
	(0.000)	(0.000)	(0.001)	(0.006)
Anxiety/depression/ bad nerves			0.289***	0.266***
			(0.000)	(0.000)
Observations	12587	12587	12587	12587
Mundalk	NO	YES	YES	YES
Demographic	NO	NO	NO	YES

p-values in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Column 4, Hansen-Sargan overidentification statistic is 4.179 distributed as Chi-sq(3) with a p-value of 0.243.
Calculations made by the authors with the PSID dataset (2007-2015).

5.1 Retirement-Stress by Individual Characteristics

The main results presented were population averages. An interesting twist in the methodology is to analyze the possible heterogeneity of these estimates, focusing on different subgroups as we have shown in Figure 3. These results are reported in the Table 3.

The first two columns report a sharp difference between men and women. The stress seems to affect men strongly, with an effect as high as 38.8 p.p., six folds of the women's coefficients. Many mechanisms could be operating here. A first one is that men may be dealing with stress less efficiently and could be more affected by this mental condition. Another could be that they may, on average, have more stressful jobs, whether it is manual or otherwise dangerous jobs. Retirement affects stress levels almost equally for both genders. The hypothesis of more stressful jobs could be confirmed by the columns (3) and (4), focusing on people working as blue or white collar. In this case, one can see that stress has a sharply higher effect on retirement for respondents with blue-collar jobs (occupied by 60 percent of men in our dataset), than it does for those with white-collar jobs. We also find that retirement has a higher impact on stress in white-collar occupation, consistently with Figure 3.

By level of wealth, columns (5) and (6), we observe that stress only affects the retirement decision for those individuals under the median wealth. There is an equity question that arises here. The people in the lower wealth groups are pushed out of labor force earlier by their stress condition, while the upper ones may continue to work longer. If they could have worked longer, they may have accumulated more wealth to enjoy retirement with less financial constraints. This mechanism accentuates the inequality over the retirees. Note that it is only among those with wealth levels higher than the median that the effect of retirement on stress is significantly different from zero. By ethnicity, we only find differences in how retirement affects stress, with no effect among the non-whites.

Concerning the instruments, we observe that being older than the full or early retirement age exhibit no differences in retirement decisions by gender, occupation, wealth and ethnicity. The effect of mental health by gender, occupation, wealth and ethnicity do no present significant differences neither. However, physical activity only affects reducing the males', white collar workers', above mean wealth individuals', and white individuals' level of stress.

Table 3: Estimation of the effect of stress on retirement decision, by subgroups

	Gender		Occupation		Wealth		Ethnicity	
	Male	Female	White	Blue	Over mean	Under mean	White	Non-white
Retirement								
Stress	0.388***	0.063+	0.056	0.304***	0.094	0.204***	0.148**	0.184***
	(0.000)	(0.084)	(0.206)	(0.000)	(0.108)	(0.000)	(0.002)	(0.000)
Age Above Full Retirement Age	0.201***	0.195***	0.188***	0.197***	0.188***	0.190***	0.184***	0.214***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age Above Early Retirement Age	0.135***	0.096***	0.104***	0.129***	0.105***	0.128***	0.104***	0.134***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Stress								
Respondent is retired	-0.406***	-0.296**	-0.380***	-0.252*	-0.375***	-0.231+	-0.444***	-0.128
	(0.000)	(0.003)	(0.000)	(0.024)	(0.000)	(0.056)	(0.000)	(0.231)
Physical activity, t-1	-0.024**	-0.010	-0.016*	-0.014	-0.016*	-0.016+	-0.019**	-0.012
	(0.002)	(0.218)	(0.022)	(0.111)	(0.019)	(0.081)	(0.008)	(0.178)
(Physical activity, t-1)^2	0.003**	0.002	0.002*	0.002	0.002*	0.002	0.002*	0.002
	(0.008)	(0.129)	(0.027)	(0.127)	(0.014)	(0.130)	(0.018)	(0.125)
Anxiety/depression/bad nerves	0.212***	0.297***	0.242***	0.302***	0.218***	0.316***	0.229***	0.382***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	6239	6348	8062	4525	7954	4633	8133	4454

p-values in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Controls for demographic characteristics and Mundalk correction are included in all specifications.
Calculations made by the authors with the PSID dataset (2007-2015).

6 Robustness

In this section, we test our results by modifying our specifications in order to determine if the relationships we found are robust. The results may be found in the Table 4.

In the first column, we used an alternative and more restrictive retirement definition. We constructed a retirement variable based on the respondent's answer to a general labor force participation question. For this variable, we used their answer to the question *"In what year did you retire?"*. Starting from the year given by the respondent – again conditional on not working any hour – the respondent was considered to be retired (but not before). As retirement may be a vague term (people may retire from their main job, but still work on sideline jobs), we prefer the main definition. However, the results are not significantly modified by the use of this definition: stress nonetheless raises the probability to retire by 15.6 p.p.

In columns (2)-(4) we use alternative stress definitions. In column (2) we have coded the variable to take value one if at most the individual reports any level of nervous or/and restless feeling. In column (3) we use a three level variable built as follows. The lowest value represents an absence of reported stress; the middle value corresponds to the respondent reporting one of these feelings (nervous or restless); and the highest level is attained when nervous or/and restless feelings are both reported at the same time, which is a good indication of one's level of stress and lack of peace. On this 1 to 3 scale, respondents in our sample has an average level of stress of 1.4. In the column (4), instead of the whole stress variable, we only used the "restless" variable. Although the effect are qualitative similar to the main specification, we find some small differences in the quantitative effect for specification (3) and (4). The effect of stress on retirement is lower but the effect of retirement on stress goes up. Instruments for retirement, the fact of being older than full or early retirement age, do exert the same quantitative effect. We observe that instruments for stress, physical activity and mental health, produces a higher effect on stress.

In columns (5) to (7), we substitute the stress variable with a variable for mental health conditions, as has been done in the previous literature. The estimation reported in the column (5) used a general mental health condition variable, which used information about whether or not the respondent suffered emotional, nervous or psychological problems, while the fact of suffering depression was used in column (6) and a 5-scale life satisfaction variable was used in column (7).[16] These results are all qualitatively similar to those

[16]Life satisfaction might exhibit opposite sign, as a high level of life satisfaction is correlated with a

found in our main models, although we found some quantitative differences. The most similar specification is the one using the fact of suffering depression (column 6). In terms of the other alternative mental health conditions, we find that the effect of stress on retirement is lower, but the instruments are quite robust to any change in the definition. Retirement does not affect stress measured as the general mental health condition (column 5). Physical activity do not affect neither the latter measure nor life satisfaction.

In column (8), we substitute the health condition included into the vector X_{it} by a self-rated health level. Self-declared health may suffer from different forms of biases, possibly leading to some biases in the estimation.[17] As true health may not be observed, a proxy would need to be used. The problem is that the estimates computed in a case where the correlation between true health and the proxy is low may be biased too, and the weaker the correlation, the greater the bias. Unfortunately, this bias could only go in one direction and could not be canceled by another one. At any level of objective health measure, if someone does not consider himself able to work any longer, he will retire. On the whole, using self-reported health in our case results in a slightly lower coefficient for retirement and stress, but the significance is not affected. Instruments exert pretty much the same effect as in the main specification.

In column (9) we check the validity of our findings by estimating our main specification with an alternative data set, the Health and Retirement Study (HRS). This data set is often used to estimate models for the determinant of retirement and other behaviors in later professional life. We chose to use the PSID because the "nervous" and "restless" variables were available, which we used to construct the stress variable. Both are important aspect representing one's stress level. However, the HRS dataset only includes the "restless" variable. As the latter has the strongest relation with the stress variable, as shown in Table 1, the information is more complete using the PSID dataset. In addition, the physical activity variable in the HRS is a dichotomous one that takes the value of one if the respondent made vigorous physical activity three times a week and zero otherwise. Being not continuous as the one we have in PSID, we cannot exploit the convex form of the relation between stress and physical activity. See descriptive statistics for HRS in Table A.2 in the Appendix.

Facing this choice, one might legitimately ask if our results would be robust to the use of

lower level of stress and a drop in the probability to leave the labor force.

[17]If bias comes from the difficulty to compare respondents' subjective perception of true health condition, this could lead to an underestimated effect. If bias would results from endogeneity (the justification hypothesis), the effect would be overestimated. But, as it was argued by Bound (1991), these opposite possible biases could be preferable to the use of an objective measure.

another dataset. These findings are directly comparable to column (4). We see that the coefficient is higher, raising the probability to retire to 21.6 p.p. Moreover, the effect of being retired on stress is larger than with PSID (63.9 p.p.). Without any doubt, having the same variables in both surveys would have been ideal, giving us us the possibility to construct similarly the stress variable. It would also have given us the possibility to exploit the nonlinear relationship between stress and heavy physical activity in the instrumental variables. Nonetheless, considering this imperfect set up, the results found in this section give us confidence that the results estimated with the PSID dataset are credible.

Finally, as in the main sample, we dropped years 2001 and 2003 to avoid the extra length in some of the data gaps when building variables, we know included them in column (10) to check that this choice would not bias the result. Main findings do slightly change in magnitude, but the main results hold. Additionally, we also check that our choice of age range, from 50 to 80, does not affect results. In column (11) we consider those individuals whose age is between 45 and 85 years old. Again, main results do not change qualitatively, only in magnitude. In both cases, the effect of stress on retirement goes down, while the effect of retirement on stress increases.

Table 4: Estimation of the effect of stress on retirement (more restricted definition), structural equations modeling

					Estimation with different...						
	(1) Retirement	(2) Stress	(3) Stress	(4) Stress	(5) Mental Health	(6) Mental Health	(7) Mental Health	(8) Health	(9) HRS	(10) Waves	(11) Age range
Retirement											
Stress	0.156***	0.154***	0.083***	0.060***	0.053***	0.161***	-0.104**	0.127***	0.216***	0.132***	0.106***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)	(0.001)	(0.001)	(0.000)	(0.000)
Age Above Full Retirement Age	0.193***	0.193***	0.192***	0.191***	0.184***	0.187***	0.189***	0.193***	0.245***	0.178***	0.178***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age Above Early Retirement Age	0.121***	0.113***	0.113***	0.112***	0.108***	0.109***	0.116***	0.113***	0.132***	0.109***	0.115***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Stress											
Resp. retired	-0.367***	-0.345***	-0.564***	-0.676***	-0.015	-0.117*	0.352**	-0.288***	-0.639***	-0.428***	-0.538***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.509)	(0.011)	(0.006)	(0.000)	(0.000)	(0.000)	(0.000)
Physical activity, t-1	-0.018**	-0.016**	-0.026**	-0.030*	0.000	-0.016***	0.012	-0.010^{+}	-0.021***	-0.013**	-0.010**
	(0.002)	(0.004)	(0.003)	(0.013)	(0.893)	(0.000)	(0.259)	(0.072)	(0.000)	(0.009)	(0.009)
(Physical activity, t-1)^2	0.002**	0.002**	0.004**	0.004*	-0.000	0.002***	-0.000	0.001^{+}		0.002*	0.001*
	(0.003)	(0.006)	(0.002)	(0.015)	(0.979)	(0.000)	(0.752)	(0.066)		(0.020)	(0.045)
Anxiety/depression /bad nerves	0.267***	0.266***	0.484***	0.666***	0.705***	0.246***	-0.309***	0.247***	0.086***	0.276***	0.296***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	12372	12587	12587	12587	12583	12587	10236	12583	55119	15995	30810

p-values in parentheses. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Column 2-4 correspond to different measures of stress. Column 5-7 consider mental health as dependent variable instead of stress. Column 8 includes in vector X_{it} an alternative measure of health. Columns 9-11 correspond to different samples. See specific details in section 6.

Calculations made by the authors with the PSID dataset (2007-2015) and the HRS dataset (1992-2014)

7 Discussion and Conclusion

This paper aimed at investigating the effect of mental health, and more specifically, *stress*, on retirement decision, and conversely, the effect of retirement on stress. The literature has shown that lower levels of health (physical and mental), increases the probability to retire, while retiring may, or may not, contribute to improved health. The existing literature is not so conclusive about whether leaving the workforce had a positive, of negative, influence on health. One of the reasons for which the literature has not yet given clear answers as to the way in which health and retirement affect each other is that those variables are jointly determined. This paper addresses this endogeneity issue by using a simultaneous equations approach to study the way in which retirement and stress affect each other.

The approach we take in this paper consists in modelling the relationship between our two variables of interest, retirement and stress, by performing a simultaneous equation analysis in which both of these outcomes are instrumented. To be more specific, we instrument retirement with individual eligibility for full or early retirement public pensions using cohort and gender-specific pension eligibility age. Stress is then instrumented using two different variables: physical activity levels during the previous period, which is an instrument that we first propose in this paper; and an objective measure of mental health, which has been used frequently in the literature. All instruments for both, retirement and stress, are found to be appropriated.

The main results show that a rise in stress increases the probability to retire, consistently with the relevant literature. Our analyses further show that this effect is stronger for males, blue-collar workers, and respondents with below-average wealth, implying an equity issue. On the other hand, taking into account simultaneity of causality, we find that retirement reduces stress levels. This result is a novelty in the literature, since previous contributions did not address the endogeneity problem. We find a higher effect for males, white-collar workers, respondents with above-average wealth, and white respondents. We find that this latter effect of retirement on stress is of larger incidence than the effect of stress in retirement. We know that sometimes the endogeneity correction may bias the estimation upward, but given the inconclusive evidence in the existing literature, it was crucial to endogeneize. Physical activity reduces stress in an increasing rate and mental health problems raise up the level of stress. This latter result is especially interesting in terms of policy, and further research should be done to analyse the effects of physical activity on economic and well-being variables at older ages.

References

Angrisani, M., Casanova, M. & Meijer, E. (2017), 'Work-life balance and labor force attachment at older ages'.

Behncke, S. (2012), 'Does retirement trigger ill health?', *Health Economics* **21**, 282â300.

Benítez-Silva, H., Buchinsky, M., Man Chan, H., Cheidvasser, S. & Rust, J. (2004), 'How large is the bias in self-reported disability?', *Journal of Applied Econometrics* **19**(6), 649–670.

Binh Tran, D. & Zikos, V. (2019), 'The causal effect of retirement on health: Understanding the mechanisms', *Australian Economic Review* **52**(4), 427–446.

Blau, D. M. (1987), 'A time-series analysis of self-employment in the united states', *Journal of Political Economy* **95**(3), 445–467.

Blekesaune, M. & Solem, P. E. (2005), 'Working conditions and early retirement: A prospective study of retirement behavior', *Research on Aging* **27**(1), 3–30.

Blundell, R., French, E. & Tetlow, G. (2016), Chapter 8 - retirement incentives and labor supply, Vol. 1 of *Handbook of the Economics of Population Aging*, North-Holland, pp. 457 – 566.

Blundell, R. W., Britton, J., Costa Dias, M. & French, E. (2017), 'The impact of health on labor supply near retirement'.

Bonsang, E., A. S. & Perelman, S. (2012), 'Does retirement affect cognitive functioning?', *Journal of Health Economics* **31**, 490â501.

Bound, J. (1991), 'Self-reported versus objective measures of health in retirement models.', *Journal of Human Resources* **26**(1).

Bound, J., Schoenbaum, M., Stinebrickner, T. R. & Waidmann, T. (1999), 'The dynamic effects of health on the labor force transitions of older workers', *Labour Economics* **6**(2), 179–202.

Butterworth, P., Gill, S. C., Rodgers, B., Anstey, K. J., Villamil, E. & Melzer, D. (2006), 'Retirement and mental health: analysis of the australian national survey of mental health and well-being', *Social science & medicine* **62**(5), 1179–1191.

Carek, P. J., Laibstain, S. E. & Carek, S. M. (2011), 'Exercise for the treatment of depression and anxiety', *The International Journal of Psychiatry in Medicine* **41**(1), 15–28.

Charles, K. K. (2004), Is retirement depressing?: Labor force inactivity and psychological well-being in later life, Technical report.

Coe, N. B. & Zamarro, G. (2011), 'Retirement effects on health in europe', *Journal of Health Economics* .

Dave, D., Rashad, I. & Spasojevic, J. (2008), 'The effects of retirement on physical and mental health outcomes', *Southern Economic Journal* **75**, 497–523.

Disney, R., Emmerson, C. & Wakefield, M. (2006), 'Ill health and retirement in britain: A panel data-based analysis', *Journal of health economics* **25**(4), 621–649.

Dunn, A. L., Trivedi, M. H. & O'Neal, H. A. (2001), 'Physical activity dose-response effects on outcomes of depression and anxiety'.

Dwyer, D. S. & Mitchell, O. S. (1999), 'Health problems as determinants of retirement: Are self-rated measures endogenous?', *Journal of health economics* **18**(2), 173–193.

Eibich, P. (2015), 'Understanding the effect of retirement on health: Mechanisms and heterogeneity', *Journal of Health Economics* **43**, 1â12.

Ferrer-i-Carbonell, A. (2005), 'Income and well-being: an empirical analysis of the comparison income effect', *Journal of Public Economics* **89**(5), 997–1019.

Ferrer-i-Carbonell, A. & Frijters, P. (2004), 'How important is methodology for the estimates of the determinants of happiness?*', *The Economic Journal* **114**(497), 641–659.

Fonseca, R., Kapteyn, A., Lee, J. & Zamarro, G. (2017), 'Does retirement make you happy? a simultaneous equations approach'.

Fox, K. R. (1999), 'The influence of physical activity on mental well-being', *Public health nutrition* **2**(3a), 411–418.

French, E. (2005), 'The effects of health, wealth, and wages on labour supply and retirement behaviour', *The Review of Economic Studies* **72**(2), 395–427.

French, E. & Jones, J. B. (2017), 'Health, health insurance, and retirement: A survey', *Annual Review of Economics* **9**, 383–409.

Hoven, H., Wahrendorf, M. & Siegrist, J. (2015), 'Occupational position, work stress and depressive symptoms: a pathway analysis of longitudinal share data', *J Epidemiol Community Health* **69**(5), 447–452.

Insler, M. (2014), 'The health consequences of retirement', *Journal of Human Resources* **49**, 195â233.

Macken, J. (2019), 'Work stress among older employees in germany: Effects on health and retirement age', *PlosOne* **14**(2), e0211487.

Mazzonna, F. & Peracchi, F. (2012), 'Ageing, cognitive abilities and retirement', *European Economic Review* **56**, 691–710.

McGarry, K. (2004), 'Health and retirement do changes in health affect retirement expectations?', *Journal of Human Resources* **39**(3), 624–648.

Michaud, P.-C., Crimmins, E. M. & Hurd, M. D. (2016), 'The effect of job loss on health: Evidence from biomarkers', *Labour economics* **41**, 194–203.

Mundlak, Y. (1978), 'On the pooling of time series and cross section data', *Econometrica* **46**, 69–85.

Neuman, K. (2008), 'Quit your job and get healthier? the effect of retirement on health', *Journal of Labor Research* **29**, 177â201.

Quinn, J. F. (1977), 'Microeconomic determinants of early retirement: A cross-sectional view of white married men', *Journal of Human Resources* pp. 329–346.

Rohwedder, S. & Willis, R. (2010), 'Mental retirement.', *Journal of Economic Perspectives* **24**, 119–138.

Seeman, T. E., Singer, B. H., Rowe, J. W., Horwitz, R. I. & McEwen, B. S. (1997), 'Price of adaptation allostatic load and its health consequences.', *Archives of Internal Medicine* **157**(19), 2259–2268.

Shultz, K. S., Morton, K. R. & Weckerle, J. R. (1998), 'The influence of push and pull factors on voluntary and involuntary early retirees' retirement decision and adjustment', *Journal of vocational behavior* **53**(1), 45–57.

Ströhle, A. (2009), 'Physical activity, exercise, depression and anxiety disorders', *Journal of neural transmission* **116**(6), 777.

Sutinen, R., Kivimäki, M., Elovainio, M. & Forma, P. (2005), 'Associations between stress at work and attitudes towards retirement in hospital physicians', *Work & Stress* **19**(2), 177–185.

Wahrendorf, M., Dragano, N. & Siegrist, J. (2012), 'Social position, work stress, and retirement intentions: a study with older employees from 11 european countries', *European sociological review* **29**(4), 792–802.

Windle, G., Hughes, D., Linck, P., Russell, I. & Woods, B. (2010), 'Is exercise effective in promoting mental well-being in older age? a systematic review', *Aging & mental health* **14**(6), 652–669.

Zhu, R. (2016), 'Retirement and its consequences for womens health in australia', *Social Science Medicine* **163**, 117â25.

Appendix

Table A1: Descriptive statistics, PSID

Variable	Mean	Std. Dev.	Min.	Max.	N
Gender (1 = Female)	0.54	0.5	0	1	25076
Marital status (1 = married)	0.67	0.47	0	1	25076
Ethnicity (1 = Non-White)	0.36	0.48	0	1	24251
Education	2.83	1.31	1	5	25076
Wealth (log)	14.77	0.22	13.42	18.26	25076
Major Health Condition	0.25	0.43	0	1	25076
Minor Health Condition	0.64	0.48	0	1	25076
Occupation (1 = White collar)	0.64	0.48	0	1	20283
Number of Persons in HH	2.44	1.24	1	12	25076

Table A2: Descriptive statistics, HRS

Variable	Mean	Std. Dev.	Min.	Max.	N
Stress	0.55	0.5	0	1	32436
Retirement	0.61	0.49	0	1	73078
Gender	1.57	0.49	1	2	80128
Marital status (1 = married)	0.63	0.48	0	1	80119
Ethnicity (1 = Non-White)	1.34	0.62	1	3	79977
Education	3.26	1.37	1	5	80118
Wealth (log)	11.3	3.1	0	17.73	74497
Major Condition	0.39	0.49	0	1	80135
Minor Condition	0.8	0.4	0	1	80135
Number of Persons in HH	2.31	1.25	1	15	80135
Age Above Full Retirement Age	0.48	0.5	0	1	80135
Age Above Early Retirement Age	0.61	0.49	0	1	80135
Physical Activity	0.32	0.47	0	1	80135
Mental Health	0.18	0.39	0	1	80002