Economic Modelling 131 (2024) 106617

Contents lists available at ScienceDirect

Economic Modelling

journal homepage: www.journals.elsevier.com/economic-modelling

Stress and retirement☆

Raquel Fonseca [a,b,*], Ana I. Moro-Egido [c], Hugo Morin [a,d]

[a] ESG-Université du Québec à Montréal, Canada
[b] CIRANO, Canada
[c] University of Granada, Spain
[d] HRSDC, Canada[1]

ARTICLE INFO

JEL classification:
C30
I10-12
J26

Keywords:
Stress
Retirement
Physical activity
Simultaneous equation models

ABSTRACT

This paper examines the bidirectional causal relationship between retirement and stress. We use data from the PSID for the period 2007–2015. By employing a simultaneous equations approach, we find that experiencing stress increases the likelihood of retirement by approximately 34.8 percentage points, whereas retirement lowers the probability of feeling stress by 19.6 percentage points. We obtain consistent results when we disaggregate data according to various individual characteristics, including gender, occupation, wealth, and ethnicity. To determine actual retirement age, we use official retirement ages and a proxy of health stock as instruments, while lagged physical activity levels serve as a non-linear instrument for perceived stress. Our findings are particularly relevant in terms of policy, and further research is needed to assess the impact of physical activity on economic and well-being outcomes for older individuals.

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 (Quinn, 1977; McGarry, 2004; Dwyer and Mitchell, 1999; Disney et al., 2006; Jones et al., 2010; Macken, 2019; Blundell et al., 2023). Some other factors may be associated with one's decision to retire such as working conditions and job strains (Blekesaune and Solem, 2005; Sutinen et al., 2005; Messe, 2011; Wahrendorf et al., 2012; Mao et al., 2014), leisure interests (Shultz et al., 1998), work-life balance (Angrisani et al., 2020) 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 physical and mental health and well-being, and that the effect might become stronger or weaker with time spent out of the labor force (Charles, 2004; Butterworth et al., 2006; Dave et al., 2008; Coe and Zamarro, 2011; Eibich, 2015; Binh Tran and Zikos, 2019). 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.

Our contribution is to disentangle the possible direction of causality between mental health and retirement. When analyzing this question, we emphasize on one aspect of mental health: stress. 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

☆ We thanks the editor, the associate editor and two anonymous referees for their useful comments. We also thank Kathleen Mullen, Mayssun El-Attar for their comments and Simon Lord for his research assistance. This research is also part of the program of the Research Chair in Intergenerational Economics. Ana I. Moro-Egido thanks financial support through Grant P18-RT-4115 funded by Consejería de Universidad, Investigación e Innovación and FEDER. Errors are our own. The data that support the findings of this study are available from https://psidonline.isr.umich.edu and https://hrs.isr.umich.edu/about. Restrictions apply to the availability of these public data. However, the files generating our variables and findings are available from the journal, which will make them public upon request to the authors.

* Corresponding author at: ESG-Université du Québec à Montréal, Canada.
E-mail address: fonseca.raquel@uqam.ca (R. Fonseca).

[1] This paper should not be considered as representing the views of the HRSDC.

https://doi.org/10.1016/j.econmod.2023.106617
Received 1 February 2023; Received in revised form 28 November 2023; Accepted 4 December 2023
Available online 6 December 2023

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, physical activity, which has been proven to be a covariate between retirement status and physical health (Eibich, 2015; Binh Tran and Zikos, 2019).

The novelty of our study lies in the integration of essential components from two distinct branches of the existing literature. One branch examines how health affects retirement decisions, while the other investigates the impact of retirement on health. By employing a simultaneous equation analysis, we combine these two perspectives and instrument both outcomes, retirement and stress. To the best of our knowledge, there is no existing theoretical study that explores the bidirectional relationship between stress and retirement. However, we believe that this question is relevant and merits empirical investigation.

Health is clearly one of the most important non-financial determinants of retirement plans or decisions. A person's health condition may directly affect their working capacity and push them out of the labor force even if they would rather have stayed at work for a number of reasons, including financial considerations. An early paper by Quinn (1977) found that health was found to have the greatest impact on early retirement; even greater than pension eligibility, and sharply greater than wage and income. Similarly, Dwyer and Mitchell (1999) found that men living with a low level of overall health will retire earlier. More recently, Disney et al. (2006) found that, after correcting for endogeneity, a higher level of health stock impacted positively the probability to stay in the labor force. McGarry (2004) reached similarly that a fair or poor subjective health is negatively correlated with labor participation. Moreover, the plan to quit the labor force is more impacted by changes in health than changes in financial variables. Macken (2019) showed that three health outcomes, like self-rated health, depression, and cardiovascular disease, potentially caused by work stress, may lead to a lower retirement age. More recently, Blundell et al. (2016) found that, overall, declines in health explain up to 15% of the decline in employment between ages 50 and 70.

The main mechanisms through which health affects labor supply are of different nature as shown in French (2005), Blundell et al. (2016), and French and Jones (2017). 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.

There is also a key distinction regarding health measurement. 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 and Mitchell, 1999).

Recent literature has shown mixed findings regarding the effect of retirement on both subjective health (e.g., physical and mental health) and objective health (e.g., limitations in activities of daily living, mortality). 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 (Neuman, 2008; Coe and Zamarro, 2011; Insler, 2014; Eibich, 2015; Zhu, 2016; Grotting and Lillebo, 2020) but other studies find negative effects (Dave et al., 2008; Behncke, 2012).[2]

The mechanism behind the negative effect may be that a drop in physical activity and social interactions induces a decline in mental health. 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. 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 to hold when using older sub-samples (Butterworth et al., 2006).

Finally, there is also new evidence 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 labor market shocks, such as income variation and job loss, among other factors, affect health and well-being 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.

Recall that one of our contributions relies on the inclusion of all these mechanisms to try to disentangle the casual relationship between stress and retirement. Our results show that there is a bidirectional effect between retirement and stress, but it is stronger in one direction. In particular, we find that feeling stress increases the probability to retire by roughly 34.8 percentage points (p.p.), while retirement decreases the probability of feeling stress by 19.6 p.p.

Being above full retirement age affects positively the probability of retirement, and the magnitude of this effect (around 29.7 p.p.) is larger than the one we found for being merely above early retirement age (around 16.6 p.p.). The stock of health has an effect of 16.5 p.p. Physical activity is a non-linear predictor of the level of stress. An increase in the frequency of physical activity decreases the probability of feeling stress by 2.0 p.p at a decreasing rate. We find the same results when we disaggregate by individuals' characteristics, such as gender, white-collar and blue-collar occupations, people whose income is below or above the mean and for non-white and white individuals. Main results apply for all the groups analyzed.

The paper will be structured as followed. Section 2 will describe the data and main variables. The empirical strategy will be addressed in Section 3 and the estimation results will be found in Section 4, with robustness checks in Section 5. Finally, Section 6 will discuss the results and conclude.

2. 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.[3] It is an unbalanced panel, so we keep individuals that are present in at least two periods.

[2] For instance, Dave et al. (2008) estimated that within a period of six years after retirement, there is a decline of 6 to 9 percent in mental health condition. However, Binh Tran and Zikos (2019) observe that retirement leads to an approximately 28 percentage-point increase on self-reported health.

[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 survey from 1997, and low attrition, around 10,000 families and 24,000 individuals nowadays. For details, https://psidonline.isr.umich.edu/Guide/default.aspx, and https://cnef.ehe.osu.edu/data/.

We focus on older ages by restricting the sample to respondents aged between 50 and 80, with an average age of roughly 60 years old.[4] We also drop individuals who were retired for the whole observed period so that we can capture the transition from work to retirement. We disregard those individuals who, eventually, came back to the labor market after a period of inactivity. In addition, we discard people who are self-employed (totally or partly) since they are known to have different behaviors regarding retirement (Blau, 1987). We end up in our main sample with 4426 individuals which make a total number of 10,677 observations.

To be sure that our results are not driven by our choices in the sample selection, in the robustness Section 5, we perform our analysis considering a balanced version of the panel, a wider range of ages and the consideration of self-employment together with the employees and by themselves. We also repeat our analyses using another data set, the Heath and Retirement Study (HRS). As we will discuss later on, our main results hold.

2.1. Main variables

We present here the main definitions of our two variables of interest, stress and retirement. We begin by defining the variable *Stress*. 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 proxy that is available for the whole studied period.

There is a set of variables evaluating the feelings about some mental conditions of the respondents repeated along the panel. These variables represent one's feeling of restlessness, nervousness, sadness, hopelessness, worthlessness and that everything is an effort. For each feeling, the variable is constructed using respondents' answer to a question taking this form (this is the example for restlessness): *"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.

Given these results, we build an index with both *nervous* and *restless* as the stress proxy for the period 2007–2015. The indicator is a dichotomous variable taking the value of one if the respondent reports any feeling of nervousness and/or restlessness (28.9% of respondents); and zero otherwise.[5] To show that this choice does not shape our results, we have considered a variety of alternative measures for stress and our results hold (see robustness Section 5).

Our other key variable is the retirement status (*Retirement*). We follow Rohwedder and 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?"* We also use the information regarding the number of hours worked. With these two variables, we construct a dummy variable that takes a value of one if the respondents said they were retired or keeping house, conditional on working in the previous period, and they did not work any hours in the current period. The variable takes a value of zero if individuals were still working. This retirement definition will be the main one used in the analysis. On average, in our sample, 22.9% of respondents are retired. To check the robustness of our results we perform the analysis with a broad definition of retirement and using only the information of the perceived retirement status (see robustness Section 5) and our results remain the same.[6]

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

	(1) 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	2,495
R-squared	0.075

p-values in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

2.1.1. Stress, retirement and age

The relationship between stress and age in our sample is reported in Fig. 1. It shows the relationship for two groups: retirees and workers, defined as described above. Taking both groups into consideration, we can see a clear negative relationship. As they get older, the incidence of stress is lower.[7]

An inevitable fact to notice here is the difference between workers and retirees who are younger than 60. In this age range, the incidence of stress among retirees is higher than among working people. After age 60, the incidence of stress is very similar between workers and retirees until age 70, where the incidence of stress among 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, the respondents' answers do not reflect this idea.

Moreover, an earlier retirement is correlated with a higher incidence of stress. For example, among 55-year-old workers 32.8% report to feel stress, when among retirees of the same age the incidence is 44.9%.

[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. The proportion of individuals beyond 80 years old is similar.

[5] 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. Note that the variable takes one if the actual variable takes the value from three to five and zero otherwise.

[6] Notice that we do not consider individuals who came back to the labor market after retirement. We also discard self-employed workers given their different labor market and retirement behavior. However, we include this occupational choice in our robustness analysis in order to check if our results could have a potential selection bias.

[7] Indeed, if we have also included 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.

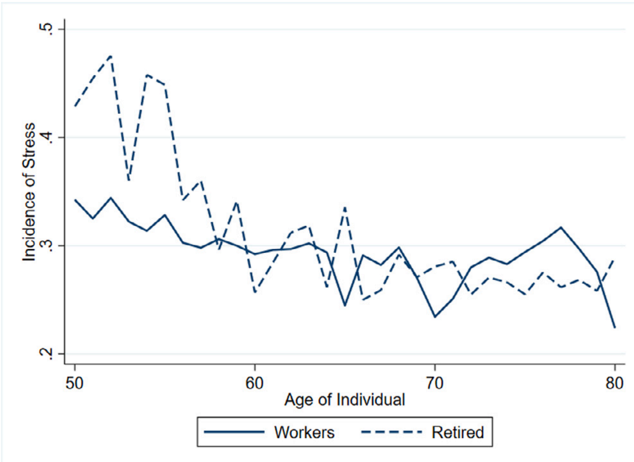

Fig. 1. Relationship between stress and age, population between 50 and 80, by work status.
Own calculations: PSID dataset (2007–2015).

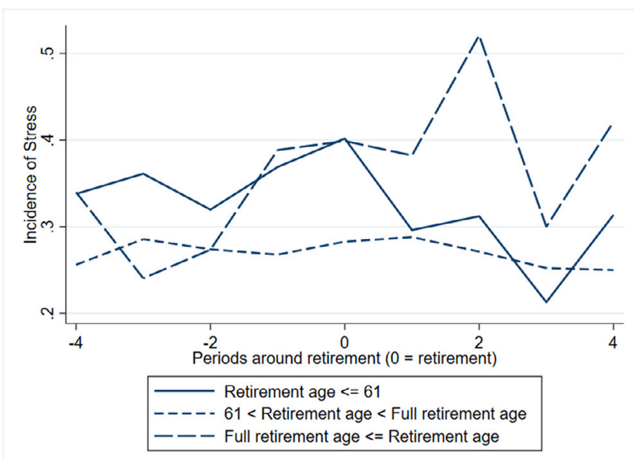

Fig. 2. Relationship between stress and period around retirement, by retirement age subgroups.
Own calculations: PSID dataset (2007–2015).

Early retirees seem to be the ones experiencing the higher levels of 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.

Before getting into this analysis, we present in Fig. 2 the average incidence of reported stress at retirement as well as during the four periods preceding and following retirement for those who retired between 50 and 80 years old. This figure could help us to understand better the time path of incidence of stress feelings, whether or not stress feelings increase when approaching retirement, and how these feelings evolve after retirement. Periods are defined as two years, the length of PSID panel data, and period 0 is the retirement year.

The sample has been divided into three different groups, depending on the age of retirement. The first group (solid line) consists of early retirees, i.e. respondents who retired before 61 years old. For this group, the incidence of stress seems to be increasing as retirement approaches, with a peak at the retirement period. Once respondents retire, the incidence 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) includes individuals 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. The incidence 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. The incidence of stress seems to experience a drastic rise before retirement and, after a last bump on the period right after retirement. Thereby, retirement and stress are correlated in this subgroup, as incidence of stress experienced an increase that is leading retirement decision, but retirement does not lower the incidence of stress.

2.2. *Heterogeneity*

We explore the existence of differences in retirement behaviors across groups. Following French et al. (2016), Nardi et al. (2021), we analyze the relationship between stress and retirement by gender, ethnicity, occupation, and wealth (Fig. 3).[8]

We define gender as a dichotomous variable that equals one if the respondent reports being a woman. We consider a distinction by race (white and non-white individuals). We include some information on the type of occupation by using a dummy 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 the incidence of stress follows a similar trend for male and female, with female revealing a slightly higher incidence of stress. There is nothing here that could lead us to conclude that 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 present a higher incidence of stress that increases until retirement, and decreases afterwards. In the bottom left panel, the sample is divided between blue and white collar workers. We find that the former report a higher incidence of stress until retirement. The bottom right panel shows a large difference in stress by wealth. Essentially, those living below the median report a sharply higher stress incidence overall. 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 among poorer people the incidence of stress is higher.

3. Empirical strategy

As discussed in the introduction, many authors have highlighted the significant role of one's health condition as a non-financial factor in retirement decisions. However, retirement has also been associated with mental health issues. Hence, when studying the relationship between stress and retirement, it appears important to take into account the simultaneity of causality between these variables. This is precisely what we aim to do in our study.

8 Table A.1. in the Appendix shows descriptive statistics for the main variables.

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

Fig. 3. Relationship between stress and period around retirement, by subgroups. Authors' calculations: PSID dataset (2007–2015).

Our decision to adopt a simultaneous equation analysis is motivated by the existing evidence of simultaneous causality, but also by the insights gained from the descriptive analysis presented in Section 2. To effectively model the relationship between our two variables of interest, retirement and stress, we employ a simultaneous equation analysis in which both outcomes are instrumented. The model is written as such:

$$Retirement_{i,t} = \beta_0 + \beta_1 Stress_{i,t} + \beta_2 I_{i,t} + \beta_3 h_{i,t} + \beta_4 X_{i,t} + \delta_t + \xi_{i,t} \tag{3.1}$$

$$Stress_{i,t} = \gamma_0 + \gamma_1 Retirement_{i,t} + \gamma_2 Physicalactivity_{i,t-1} + \gamma_3 Physicalactivity^2_{i,t-1} + \gamma_4 X_{i,t} + \delta_t + \mu_{i,t} \tag{3.2}$$

where

- *Retirement* is individual i's retirement probability in period t, conditional on not being retired in the previous period ($Pr(Retirement_{i,t} = 1 | Retirement_{i,t-1} = 0)$). 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 a dichotomy variable where the value is 1 if the respondent i's self-reported any level of feeling stress in period t. It takes the value of 0 when a respondent does not feel any stress;
- $I_{i,t}$ is a matrix of institutional variables regarding retirement age;
- $h_{i,t}$ is a measure of the stock of health;
- $Physicalactivity_{i,t-1}$ 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 Eq. (3.1), 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 Eq. (3.2), we could also argue that retirement can affect stress (*i.e.* $corr(Retirement_{it}, \mu_{i,t}) \neq 0$). 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 Section 3.1.

Additionally, to make use of the panel structure of the data set, the estimation includes 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 individuals' characteristics do. With this in mind, we allow individual correlation over time by splitting the error term in Eq. (3.1) 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 $\epsilon_i \sim N(0, \sigma^2_\epsilon)$, $\varepsilon_{it} \sim N(0,1)$ 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 and Frijters (2004) or Ferrer-i-Carbonell (2005). With this method, the individual random effect ϵ_i is modeled as $\epsilon_i = u_i + \overline{\omega_i}\eta$. The first term is a pure error term, u_i, which is normally distributed with zero mean and independent of the idiosyncratic error ε_{it}. The second term

[10] We present here the composition of the error term in Eq. (3.1), but is similar in Eq. (3.2).

represents 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.

It is important to underline a computational limitation. The estimations are performed assuming a linear probability model for both dependent variables.[11] As the dependent and the explanatory variables are dichotomous, we minimize the potential drawbacks of this assumption. A reason for this choice is that empirical implementation of a simultaneous equations with binary dependent variables is not empirical well-developed.[12]

3.1. Instrumental variables

We start describing the instrumental variables. We relegate the detailed description of the other control variables to the following subsection.

Regarding the retirement decision, Eq. (3.1), we will make use of two types of instruments already used in the related literature. First, in matrix $I_{i,t}$ we include two exogenous institutional variables concerning retirement age. In particular, we built 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; Rohwedder and Willis, 2010; Fonseca et al., 2017; Coe and Zamarro, 2011.[13] 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. Secondly, following Disney et al. (2006), we build a *health stock* taking into consideration individual's characteristics as well as health indicators.[14]

It has been shown in related literature (Dwyer and Mitchell, 1999; Disney et al., 2006; Bound, 1991; Bound et al., 1999; McGarry, 2004; Fonseca et al., 2017) that considering both mental health and physical health directly in the explanation of retirement decisions could lead to some bias.

In terms of Eq. (3.2), perceived stress variable will be instrumented by respondents' level of physical activity on the previous period ($Physicalactivity_{i,t-1}$). The idea is that heavy or vigorous physical activity on a regular basis helps to temper depression and anxiety (Dunn et al., 2001; Ströhle, 2009; Carek et al., 2011) and helps mental well-being in general (Fox, 1999; Windle et al., 2010). Furthermore, we incorporated the lagged form of this variable for two reasons: (i) to prevent from collinearity between stress and physical activity. It is highly improbable that an individual's stress level today gives the same information that the number of days they engaged in exercise two years ago. And (ii) to make sure there is no correlation with contemporaneous retirement decision. By using the lag, we avoid the possibility of an erroneous endogeneity correction, as there could be a potential relationship between physical activity and retirement. As far as we know, physical activity has rarely been used in this literature.

In order to uncover the suitability and the validity of this instrument, we present additional arguments and some evidence.[15] There is no reason to think that the number of times you practiced physical activity or going the gym in the previous years may related to retirement decisions today, if it is not through health status as we claim. Finally, one can think that physical activity (even when lagged) is likely to be correlated to some other unobservable characteristics ($\mu_{i,t}$) that influence stress. The potential correlation with unobservables could be mitigated by controlling for individual fixed effects. We have controlled for different individual and household's characteristics. We also include the Mundlak correction, as it incorporates potential confounders that are fixed over time within individuals. See Table 2 for details.

The PSID contains information on how often respondents do *vigorous* physical activities for at least 10 min 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 also analyze the alternative intensity of physical activity (light or moderate) to check the robustness of our results (Section 5). The pair-wise correlation among retirement and physical activity is about −0.06.

We present this relationship between heavy physical activity and stress in Fig. 4 for workers and retirees. First of all, taking both groups, a non-linear relation can be stated between those two variables. We see that a higher level of heavy physical activity is linked with a lower incidence of stress, with a floor at around 4 and 6 times a week, and then a rise back for those doing more exercise. 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 incidence of stress.

Given these findings, we choose to introduce the variable $Physicalactivity_{i,t-1}$ 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. As shown in Fig. 1, on average, there is a higher incidence of stress among non-retirees. 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 non-linear form of the relation is seen.

[11] As articulated by Wooldridge (2002) and Angrist and Pischke (2008), when the primary objective is to assess the partial impact of an independent variable on the probability of a response, averaged over the distribution of that variable, the occurrence of some predicted values outside the unit interval may not carry significant consequences. In light of this perspective, using Linear Probability Models (LPM) in our study is suitable for investigating the potential bi-directional relationship between retirement and health.

[12] Theoretical developments by Chesher and Rosen (2012) consider theoretically identification in a simultaneous equations model with discrete outcomes, as demonstrated in their Example 1. We would like to thank the anonymous referee for bringing this to our attention. To further validate our approach, we conducted a robustness analysis in Section 5 using an alternative measure of stress as a 3-level categorical variable. Additionally, we estimated a simultaneous equation method using the "cdsimeq" command in STATA, which provide a simultaneous estimation with a dummy variable and a continuous variable as dependent variables. Yet, this methodology does not to allow us to estimate simultaneous equations with two dummies as dependent variables. Main results remain unchanged and are available upon request.

[13] For detailed information on the early and normal statutory retirement ages by cohort and gender, see https://www.ssa.gov/oact/progdata/nra.html.

[14] We use the same type of variables as in Disney et al. (2006). We include socio-economic variables: age, squared age, sex, race, household size, marital status, education; and health related variables: major health condition (cancer, stroke, heart attack or lung disease), minor health condition (arthritis, hypertension or diabetes) and having psychiatric problems (anxiety, depression or nervosity). As a robustness test we have also consider the index in lagged value to capture how the evolution of health could influence retirement decisions. Results do not change and are available upon request from authors.

[15] We added some information in the result section regarding empirical tests.

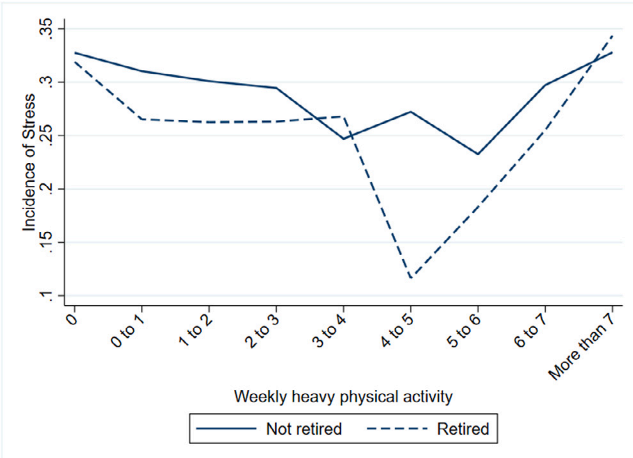

Fig. 4. Relationship between stress and physical activity, by working status.
Own calculations: PSID dataset (2007–2015).

3.2. Explanatory variables

Concerning the specific variables for our analysis, we start describing the standard control variables included in vector X_{it}. Beside the variables included in Section 2 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. Finally, we include the size of the household the individual lives in.

Finally, we consider that health conditions could influence the level of self-perceived stress in Eq. (3.2). While major health condition is well-known recognized determinant of retirement decisions, minor conditions could have an effect on its own that could be related to the respondent's level of stress. Moreover, controlling for health measures implies that physical activity could help to temper the stress, and well-being in general. We will also consider the inclusion of measures of objective mental health.[16] The intuition to include these more objectives measures of health is that the self-reported level of stress is heterogeneous in the sense that individuals with different objective health could report different levels of stress. For main descriptive statistics, see Table A.1 in Appendix.

4. Results

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 Eq. (3.1), i.e., the determinants of the retirement probability. The lower part reports the estimates of Eq. (3.2), estimating the determinants of the probability of feeling stress. We report here only the estimated parameters corresponding to the variables of interest, relegating the whole set of estimated parameters to Table A.2 in Appendix.

The columns (1) to (3) introduce the control variables gradually. As we see, the inclusion of every control variable does not change substantially the estimated coefficients. Feeling stress raises the probability to retire by 42.6 p.p. in the baseline model (1), and by 34.8 p.p. in the model including all controls (3). This effect is similar in magnitude to what has been previously found for physical health (Quinn, 1977; McGarry, 2004; Disney et al., 2006; Dwyer and Mitchell, 1999). The potential mechanisms, as reviewed in the Introduction, could be the unpleasant effect of suffering from stress and still be working, or just that the productivity decreases.

Regarding the effect of retirement on the probability of feeling stress, we find that the choice to leave the labor force does reduce likelihood of stress by 19.6 p.p. These two findings are consistent with patterns shown in Fig. 2. The incidence of stress increases in periods before retirement and it did seem to decrease after retirement, even after several years. The intuition behind this results may relay on the consideration that retirement can be considered as a stress-free enjoyment following a stressful professional life.

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 16.6 p.p.). Meanwhile, being older than the full retirement age (which depends on birth year in the USA) has a stronger significant effect (around 29.7 p.p) on retirement.[17] Stock of health is also found to be a statistically good instrument, consistently with what has been found in the literature (Bound, 1991; Dwyer and Mitchell, 1999; Benítez-Silva et al., 2004; Disney et al., 2006). The corresponding effect is an increase of 16.5 p.p. of the probability of being retired. So, as in Disney et al. (2006), the positive effect of health stock on retirement would confirm

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

	(1)	(2)	(3)
Retirement			
Stress	0.426 [0.005]	0.492 [0.009]	0.348 [0.000]
Age Above Full Retirement age	0.313 [0.000]	0.315 [0.000]	0.297 [0.000]
Age Above Early Retirement age	0.182 [0.000]	0.182 [0.000]	0.166 [0.000]
Health Stock			0.165 [0.006]
Stress			
Retirement	−0.200 [0.000]	−0.203 [0.000]	−0.196 [0.000]
Physical activity, t-1	−0.032 [0.000]	−0.028 [0.000]	−0.020 [0.000]
(Physical activity, t-1)^2	0.004 [0.000]	0.003 [0.000]	0.003 [0.000]
Observations	10,677	10,677	10,677
Ind. Charact.[a]	NO	NO	YES
Mundalk[b]	NO	YES	YES
Time Fixed effects	NO	NO	YES

p-values in parentheses.
Hansen–Sargan overidentification statistic is 4.179, p-value of 0.243.
Own calculations: PSID dataset (2007–2015).
[a] Here we consider individual fixed effects (gender, race, education); individual/household characteristics (marital status, household size, white-collar occupation, wealth quintiles); objective health charact. (mental health, major and minor health conditions).
[b] Here we consider family income, years of education, and size of the household.

[16] The correlation among self-reported stress and objective mental health is low, about 0.16. In the case of objective measures of physical health, the correlation is even lower, of about 0.05. So one cannot think that self-reported stress is fully determined by either mental or physical health.

[17] The difference is 13 p.p. is significantly different from zero. It is important to note that each of the legal retirement ages' effect is also significantly different from stress' effect on retirement.

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

	All	Gender		Occupation		Wealth		Ethnicity	
		Male	Female	White	Blue	Under mean	Over mean	White	Non-white
Retirement									
Stress	0.348	0.245	0.485	0.541	0.214	0.441	0.292	0.329	0.636
	[0.000]	[0.075]	[0.000]	[0.005]	[0.024]	[0.002]	[0.004]	[0.010]	[0.000]
Above Full Ret. Age	0.297	0.272	0.344	0.246	0.316	0.308	0.288	0.283	0.354
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Above Early Ret. Age	0.166	0.141	0.205	0.137	0.178	0.182	0.138	0.167	0.166
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Health Stock	0.165	0.320	−0.011	0.227	0.133	0.044	0.240	0.081	0.107
	[0.006]	[0.000]	[0.920]	[0.061]	[0.063]	[0.646]	[0.001]	[0.371]	[0.207]
Stress									
Retirement	−0.196	−0.293	−0.118	−0.094	−0.225	−0.253	−0.098	−0.283	−0.05
	[0.000]	[0.000]	[0.065]	[0.312]	[0.000]	[0.000]	[0.298]	[0.000]	[0.528]
Physical activity, *t-1*	−0.020	−0.025	−0.015	−0.017	−0.022	−0.023	−0.017	−0.026	−0.012
	[0.000]	[0.002]	[0.041]	[0.036]	[0.002]	[0.001]	[0.092]	[0.001]	[0.120]
(Physical activity, *t-1*)^2	0.003	0.003	0.002	0.002	0.003	0.003	0.002	0.003	0.002
	[0.000]	[0.005]	[0.028]	[0.047]	[0.004]	[0.001]	[0.136]	[0.002]	[0.064]
N	10,677	5,658	5,019	3,788	6,889	6,851	3,826	6,999	3,678

P-values in parenthesis. Calculations made by the authors with the PSID dataset (2007–2015).

that for there is a link between changing health stocks and changes in labor market state rather than simply an underlying association between poor health and inactivity.[18]

On the stress side, there is a strong non-linear relation between stress and physical activity, in line with Fig. 4. If this relation is intuitive (heavy physical activity lowering stress levels), it is a statistically good instrument as well. We find that physical activity decreases the probability of feeling stressed by 2 p.p., and the effect is increasing with the level of physical activity (estimated parameter of the squared variable is positive). In particular, and confirming Fig. 4, the strongest effect arises with a frequency of physical activity of 4 times a week.[19]

Considering now the other explanatory variables, we find that working or not in a white-collar occupation affects similarly retirement and stress (see Table A.2 into Appendix). While we find that gender and marital status only affects retirement, race exerts similar effects (of 3.1 p.p. and 3.2 p.p., respectively) on the retirement decision and on the stress level. Regarding wealth, we find that it only affects stress (from 7.4 p.p. at the bottom quintile to 10.4 p.p. at the top quintile) and only the top quintile displays an effect on retirement. Finally, major and minor physical health condition are less relevant for stress determination than mental health.

In sum, we have found that retirement affects negatively the incidence of stress, but also that the probability of feeling stress levels affect positively the decision to retire. The latter is larger in absolute value (15.2 p.p.) than the former. The variables we chose to instrument retirement are good, 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, and the stock of health is positively related to retirement (in the same line as in Disney et al., 2006). We also find that lagged physical activity displays a non-linear effect on stress.

[18] We have estimated the model considering health stock as another determinant not an instrument and also dropping from the model. Main results hold and are available upon request from the authors. We have also controlled by the lagged stock of health with no significant effect.

[19] The test for the validity of instruments has been developed for the non-structural case, where a single outcome depends on a variable that is instrumented. As mentioned in Andrews et al. (2019), we assess instrument strength using the F-statistic along with identification-robust Anderson-Rubin confidence intervals. We find that in the first step the F-statistic is F(2, 10654) = 4.20. So, there is evidence of the explanatory power of the instruments. Additionally, Sargan statistic is 0.243 with p-value of 0.623. Therefore, we cannot reject the null hypothesis that they are not correlated with error terms. Thus, instruments are valid.

Table 4
Retirement-stress relationship (different definitions of retirement), structural equations modeling.

	(0)	(1)
Retirement		
Stress	0.348	0.341
	[0.000]	[0.000]
Above Full Retirement Age	0.297	0.271
	[0.000]	[0.000]
Above Early Retirement Age	0.166	0.152
	[0.000]	[0.000]
Health Stock	0.165	0.131
	[0.006]	[0.020]
Stress		
Retirement	−0.196	−0.241
	[0.000]	[0.000]
Physical activity, *t-1*,	−0.020	−0.019
	[0.000]	[0.001]
(Physical activity, *t-1*)^2	0.003	0.003
	[0.000]	[0.001]
Observations	10,677	10,534

P-values in parenthesis. Calculations made by the authors with the PSID dataset (2007–2015).

4.1. Retirement-stress: Heterogeneity

Following the intuition presented in Fig. 3, we estimate our model by different groups to check the existence of heterogeneous effects. These results are reported in Table 3. The first column represents the baseline results in Table 2.

The subsequent two columns report results for men and women. The same mechanisms could be operating here for both of them. Retirement affects the probability of feeling stress, and the effect is slightly larger for males; feeling stress exerts a positive effect on retirement, and the effect is slightly larger in females. The hypothesis of more stressful jobs could be confirmed by the columns (4) and (5), focusing on people working as blue or white collar, consistently with Fig. 3.

By level of wealth, columns (6) and (7), we observe that feeling stress affects the retirement decision, largely for individuals under the mean 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

Table 5
Retirement-stress relationship (different definitions of stress), structural equations modeling.

	(0)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Retirement											
Stress	0.348	0.339	0.192	0.142	0.082	0.054	0.200	0.129	0.177	0.134	0.176
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Above Full Ret. Age	0.297	0.290	0.295	0.293	0.289	0.287	0.276	0.293	0.285	0.286	0.277
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Above Early Ret. Age	0.166	0.162	0.165	0.164	0.161	0.160	0.154	0.164	0.158	0.159	0.154
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Health Stock	0.165	0.161	0.161	0.157	0.164	0.170	0.192	0.165	0.209	0.184	0.223
	[0.006]	[0.006]	[0.005]	[0.009]	[0.002]	[0.001]	[0.000]	[0.002]	[0.000]	[0.000]	[0.000]
Stress											
Retirement	−0.196	−0.129	−0.314	−0.387	−0.427	−0.502	−0.001	−0.394	−0.134	−0.209	0.02
	[0.000]	[0.003]	[0.000]	[0.000]	[0.002]	[0.005]	[0.986]	[0.000]	[0.029]	[0.043]	[0.716]
Physical activity, *t-1*	−0.020	−0.022	−0.032	−0.039	−0.084	−0.122	−0.052	−0.031	−0.043	−0.054	−0.024
	[0.000]	[0.000]	[0.000]	[0.002]	[0.000]	[0.000]	[0.000]	[0.005]	[0.000]	[0.000]	[0.000]
(Physical activity, *t-1*)^2	0.003	0.003	0.004	0.005	0.011	0.015	0.006	0.005	0.005	0.007	0.003
	[0.000]	[0.000]	[0.000]	[0.002]	[0.000]	[0.000]	[0.000]	[0.002]	[0.000]	[0.000]	[0.003]
Observations	10,677	10,677	10,677	10,677	10,677	10,677	10,677	10,677	10,677	10,677	10,677

P-values in parenthesis. Calculations made by the authors with the PSID dataset (2007–2015).

the effect of retirement on stress is significantly different from zero. By ethnicity, columns (8) and (9), we only find differences in how retirement affects the probability of feeling stress, with no effect among the non-whites. Therefore, we do not appreciate changes on over whole results for the different subgroups.

5. Robustness

In this section, we test our results by modifying our specifications in order to determine if the relationships we found are robust. We consider different definitions of our two dependent variables: retirement and stress. We then modify the definition of our instrument for stress, that is, physical activity, and finally we change some definitions in the independent variables, data set and years of the study. In all tables we report our baseline model in column (0), corresponding to column (3) in Table 2.

5.1. Robustness in the retirement measure

In this section, we test our results by modifying our specifications in order to determine if the relationships we found are robust. We consider different definitions of our two dependent variables: retirement and stress. We then modify the definition of our instrument for stress, that is, physical activity, and finally we change some definitions in the independent variables, data set and years of the study. In all tables we report our baseline model in column (0), corresponding to column (3) in Table 2.

5.2. Robustness in the retirement measure

We first used an alternative and more restrictive retirement definition (Table 4). In column (1), 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: feeling stress nonetheless raises the probability to retire by 34.1 p.p. and stress is reduced more the probability of retirement by 24.1 p.p related to the baseline model.

5.3. Robustness in the stress measure

In Table 5 we present alternative measures of stress. In column (1) 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 (2) 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 feeling 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 (3), instead of the whole stress variable, we only used the "restless" variable. Although the effect are qualitative similar to the main specification in column (0), we find some small differences in the quantitative effect for specification (3). 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, produces a slightly higher effect on stress.

We have also consider a set of seven other measures of stress going beyond feelings of restlessness and nervousness. In column (4) we use a continuous variable to account for the accumulation of feelings of sadness, nervousness, restlessness, hopelessness, worthlessness, and that everything is an effort, that is, all the variables considered in Table 1 for estimating our proxy of stress. In column (5), we use an index with all the previous variables constructed using the Principal Component Analysis. In column (6) to (10) we consider a dummy variable taking value 1 if there is a feeling of sadness, nervousness, hopelessness, worthlessness, and that everything is an effort, respectively. The results are qualitatively robust to different measures.[20]

5.4. Robustness on physical activities and independent variables

In Table 6 we present the results for different measures of physical activity, health condition, age range and labor status. In column (1) we define physical activity as in the main specification with the difference

[20] We have performed alternative analyses considering the stress variable with a variable for mental health conditions, as has been done in the previous literature: suffered emotional, nervous or psychological problems; suffering depression; life satisfaction. Main results hold qualitatively.

Table 6
Retirement-stress relationship (different definitions of independent variables), structural equations modeling.

	(0)	(1)	(2)	(3)	(4)	(5)
Retirement						
Stress	0.348	0.338	0.391	0.290	0.308	0.177
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.323]
Above Full Ret. Age	0.297	0.296	0.278	0.297	0.251	0.236
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Above Early Ret. Age	0.166	0.165	0.154	0.169	0.144	0.106
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Health Stock	0.165	0.170	0.128	0.180	0.465	0.466
	[0.006]	[0.006]	[0.000]	[0.000]	[0.000]	[0.000]
Stress						
Retirement	−0.196	−0.189	−0.011	−0.225	−0.187	−0.472
	[0.000]	[0.000]	[0.804]	[0.000]	[0.000]	[0.000]
Physical activity, *t-1*	−0.020	−0.013	−0.013	−0.011	−0.020	−0.025
	[0.000]	[0.021]	[0.010]	[0.025]	[0.000]	[0.075]
(Physical activity,*t-1*)^2	0.003	0.002	0.002	0.001	0.002	0.002
	[0.000]	[0.018]	[0.008]	[0.036]	[0.000]	[0.180]
Observations	10 677	10 677	10 673	14 574	14 280	2176

P-values in parenthesis. Calculations made by the authors with the PSID dataset (2007–2015).

that now we consider the intensity of the activity to be light or moderate instead of vigorous. Notice that results are almost the same.

In column (2) we substitute the health condition included into the vector X_{it} by a self-rated health level. As pointed out in Disney et al. (2006) and previously in this manuscript, self-declared health may suffer from different forms of biases, possibly leading to some biases in the estimation.[21] For this reason, we substitute it with an estimated stock of health in our main specification, but here we present the estimation results with the self-reported health. Again, there is a positive effect of stress on retirement, while in this case the effect of retirement on stress vanishes.

In column (3) we consider those individuals whose age is between 45 and 85 years old, extending then the age range with respect to the main specification (50–80 years old). 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.

We are aware of the different labor market behavior of self-employed as pointed out in Section 2, however in order to check whether the exclusion of such type of workers could have biased our results, we run the estimation with all workers. In column (4) we consider employees and the self-employed in the sample and in column (5) we only consider the self-employed. When taking employees and the self-employed together, results do not change, except for the magnitude of the effect. When considering the self-employed separately, we observe that physical activity does not affect their level of stress. The rest of our results hold for this group of individuals. We do not want to comment much on this group as the pattern for labor supply, and therefore for retirement, might be determined by other factors. Even, the self-selection on self-employment should be taken into account, which we have not done; we present here only the results for them in order to establish that including them in the whole sample does not significantly change our main results as shown in column (4).

5.5. Robustness on time period and data set

As in the main sample, we dropped the years 2001 and 2003 to avoid data gaps when building variables. We know include them in column (1) of Table 7, to check that this choice would not bias the result. Main findings do slightly change in magnitude, but the main results hold. Although by construction the PSID is an unbalanced panel data, we run the estimation restricting the analysis to a balanced version of the dataset (individuals observed all periods, in particular for more than 7 periods). In column (2) we report the results for the balanced panel. Again, the main findings do slightly change in magnitude, but the main results hold.

In column (3) 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 are interested in a general stress variable, therefore we use the PSID. However, the HRS dataset includes a proxy about how much the current job involves stress with four categories from strongly disagree to strongly agree. We build a dichotomy variable with one, of responses of agree or strongly agree and zero otherwise. Nervousness and restlessness represent general stress variables and not only the

Table 7
Retirement-stress relationship (different time period and data set), structural equations modeling.

	(0)	(1)	(2)	(3)
Retirement				
Stress	0.348***	0.285***	0.397***	0.198*
	[0.000]	[0.000]	[0.000]	[0.013]
Above Full Ret. Age	0.297***	0.259***	0.298***	0.189***
	[0.000]	[0.000]	[0.000]	[0.000]
Above Early Ret. Age	0.166***	0.155***	0.174***	0.102***
	[0.000]	[0.000]	[0.000]	[0.000]
Health Stock	0.165**	0.126**	0.079	−0.193***
	[0.006]	[0.009]	[0.272]	[0.000]
Stress				
Retirement	−0.196***	−0.281***	−0.173**	−0.632***
	[0.000]	[0.000]	[0.002]	[0.000]
Physical activity, *t-1*	−0.020***	−0.016**	−0.015*	−0.022***
	[0.000]	[0.003]	[0.013]	[0.000]
(Physical activity,, *t-1*)^2	0.003***	0.002**	0.002**	
	[0.000]	[0.004]	[0.007]	
	10,677	13,509	8,065	55,119

P-values in parenthesis. Calculations made by the authors with the PSID dataset (2007–2015)(columns 1 and 2) and the HRS dataset (1992–2014) (column 3).

[21] 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.

stress caused by the job. We have also checked with other mental health variables in HRS. Available upon requests.

The closer definition to the PSID variable of the physical activity in the HRS is a dichotomous one, that takes the value of one if the respondent made vigorous physical activity more than three times a week, and zero otherwise. Being not continuous, we cannot exploit the non-linear form of the relation between stress and physical activity. The rest of the variables are built in similar way as possible to the PSID variables. See descriptive statistics for HRS in Table A.3 in Appendix.

We then test the robustness of our analyses to the use of another data set altogether. In this case, the effect of stress on the probability to retire raises to 19.8 p.p. Moreover, the effect of being retired on stress is larger than with the PSID. Without any doubt, having the same variables in both surveys would have been ideal, giving us the possibility to construct the stress variable similarly. 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.

6. Conclusion and discussion

This paper aimed at investigating the effect of stress on retirement decision, and conversely, the effect of retirement on stress. The literature has shown that lower levels of health, physical and mental, increase the probability to retire, while retiring may, or may not, improve health. One of the reasons for which the literature might have 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. Specifically, retirement is instrumented using individual eligibility for full or early retirement public pensions, taking into account cohort and gender-specific pension eligibility age, and a measure of health stock, which has been used frequently in the literature. Stress, on the other hand, is instrumented using physical activity levels from the previous period, which is a novel instrument proposed in this paper. Instruments for both, retirement and stress, are found to be appropriate and valid.

The main findings show that, accounting for the simultaneity of causality, feeling stress increases the likelihood of retirement, which is consistent with the existing literature, and that retirement reduces stress. This result is a novelty in the literature, since previous contributions did not address the issue of endogeneity in both sides. We find that these patterns remain by gender and occupation, with minor variations observed in stress among different wealth groups and ethnicity. The effect of retirement on stress is found to be less substantial than the impact of stress on retirement, both in the overall case and across all demographic groups. 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 endogenize.

While our study sheds light on the relationship between stress and retirement decisions, there are some limitations regarding data availability to acknowledge. It would have been ideal, for instance, to have longitudinal data on biomarkers as a stress-related medical condition or legal reforms regarding stress to provide stronger evidence of the impact of stress on retirement decisions. Despite this limitation, our study employs robust empirical methods and a two dataset analysis to examine the association between stress and retirement, providing valuable insights. In the same reasoning, there is no perfect measure of stress in our dataset, although the proxies for stress used in this study were found to be robust across different specifications.

Our study highlights the need for further research and policy interventions to mitigate the adverse effects of stress on individuals' health and retirement outcomes. As our results show, feeling stress has an important impact on the decision to retire. This is especially important given the ongoing debate on the need to work longer to address labor shortages and pension finances. It is crucial to manage stress effectively to prevent it from becoming a deciding factor in retirement. Managing stress for older workers is necessary to ensure that retirement decisions are made based on personal preferences and financial stability, rather than being driven by feelings of stress.

The reduction of stress after retirement also raises important questions about mental health in society. Employers could consider implementing stress-management resources for their employees, such as mindfulness training or mental health counseling. Beyond the workplace, policymakers should consider enhancing retirement benefits and providing support for individuals as they transition into retirement, including wider access to mental healthcare and social support networks. Our study also suggests that policies at societal and business levels could encourage physical activity in the workplace or outside work hours as a means of managing stress among those approaching retirement age. Further research is needed to analyze the effects of physical activity on economic and well-being variables at older ages.

Declaration of competing interest

We state that there is no financial/personal interest or belief that could affect their objectivity, or if there is, stating the source and nature of that potential conflict.

Data availability

The authors do not have permission to share data.

Acknowledgments

This paper should not be considered as representing the views of the HRSDC. This research is also part of the program of the Research Chair in Intergenerational Economics. Ana I. Moro-Egido thanks financial support from R&D&R programmes of the Spanish Government through project PID2019-111765GB-I00 and from the Regional Government of Andalusia through project PY18-4115 is gratefully acknowledged.

Appendix

See Table A.1, A.2 and A.3.

Table A.1
Descriptive statistics, PSID.

Variable	Mean	Std. Dev.	Min.	Max.	N
Gender (1 = Female)	0.54	0.5	0	1	25,076
Ethnicity (1 = Non-White)	0.36	0.48	0	1	24,251
Occupation (1 = White collar)	0.64	0.48	0	1	20,283
Wealth (log)	14.77	0.22	13.42	18.26	25,076
Marital status (1 = married)	0.67	0.47	0	1	25,076
Education	2.83	1.31	1	5	25,076
Major Health Condition	0.25	0.43	0	1	25,076
Minor Health Condition	0.64	0.48	0	1	25,076
Number of Persons in HH	2.44	1.24	1	12	25,076

Table A.2
Estructural equations modeling (full set of covariates).

	Retirement	Stress
Stress	0.348	
	[0.000]	
Above Full Ret. Age	0.297	
	[0.000]	
Above Early Ret. Age	0.166	
	[0.000]	
Retirement		−0.196
		[0.000]
Physical activity$_{t-1}$		−0.020
		[0.000]
Physical activity$^2_{t-1}$		0.003
		[0.000]
Health Stock	0.165	
	[0.039]	
Major condition		0.047
		[0.000]
Minor condition		0.028
		[0.007]
Mental health		0.251
		[0.000]
Gender	0.102	0.016
	[0.000]	[0.198]
Household size	−0.010	0.004
	[0.139]	[0.604]
Race	0.032	−0.031
	[0.023]	[0.012]
Marital Status	0.094	−0.001
	[0.000]	[0.964]
White collar	−0.024	0.010
	[0.019]	[0.431]
Wealth (Q2)	0.027	−0.074
	[0.228]	[0.002]
Wealth (Q3)	0.006	−0.079
	[0.782]	[0.000]
Wealth (Q4)	0.035	−0.100
	[0.096]	[0.000]
Wealth (Q5)	0.063	−0.104
	[0.005]	[0.000]
Const.	−0.697	0.003
	[0.187]	[0.996]

P-value in brackets.

Table A.3
Descriptive statistics, HRS.

Variable	Mean	Std. Dev.	Min.	Max.	N
Stress	0.55	0.5	0	1	32 436
Retirement	0.61	0.49	0	1	73 078
Gender	1.57	0.49	1	2	80 128
Marital status (1 = married)	0.63	0.48	0	1	80 119
Ethnicity (1 = Non-White)	1.34	0.62	1	3	79 977
Education	3.26	1.37	1	5	80 118
Wealth (log)	11.3	3.1	0	17.73	74 497
Major Condition	0.39	0.49	0	1	80 135
Minor Condition	0.8	0.4	0	1	80 135
Number of Persons in HH	2.31	1.25	1	15	80 135
Age Above Full Retirement Age	0.48	0.5	0	1	80 135
Age Above Early Retirement Age	0.61	0.49	0	1	80 135
Physical Activity	0.32	0.47	0	1	80 135

References

Andrews, I., Stock, J., Sun, L., 2019. Weak instruments in IV regression: Theory and practice. Annu. Rev. Econ. (4).

Angrisani, M., Casanova, M., Meijer, E., 2020. Work-life balance and labor force attachment at older ages. J. Labor Res. 41, 34–68.

Angrist, J., Pischke, J., 2008. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press.

Behncke, S., 2012. Does retirement trigger ill health? Health Econ. 21, 282–300.

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

Binh Tran, D., Zikos, V., 2019. The causal effect of retirement on health: Understanding the mechanisms. Australian Econ. Rev. 52 (4), 427–446.

Blau, D.M., 1987. A time-series analysis of self-employment in the United States. J. Polit. Econ. 95 (3), 445–467.

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

Blundell, R., Britton, J., Dias, M.C., French, E., 2023. The impact of health on labor supply near retirement. J. Hum. Resour. 58 (1), 282–334.

Blundell, R., French, E., Tetlow, G., 2016. Chapter 8 - retirement incentives and labor supply. In: Piggott, J., Woodland, A. (Eds.), In: Handbook of the Economics of Population Aging, vol. 1, North-Holland, pp. 457–566.

Bound, J., 1991. Self-reported versus objective measures of health in retirement models. J. Hum. Resour. 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 Econ. 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. Soc. Sci. Med. 62 (5), 1179–1191.

Carek, P.J., Laibstain, S.E., Carek, S.M., 2011. Exercise for the treatment of depression and anxiety. Int. J. Psychiatry Med. 41 (1), 15–28.

Charles, K.K., 2004. Is retirement depressing? Labor force inactivity and psychological well-being in later life. In: Polachek, S.W. (Ed.), Accounting for Worker Well-Being. Elsevier.

Coe, N.B., Zamarro, G., 2011. Retirement effects on health in Europe. J. Health Econ. 30 (1), 77–86.

Dave, D., Rashad, I., Spasojevic, J., 2008. The effects of retirement on physical and mental health outcomes. Southern Econ. J. 75, 497–523.

Disney, R., Emmerson, C., Wakefield, M., 2006. Ill health and retirement in Britain: A panel data-based analysis. J. Health Econ. 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. Med. Sci. Sports Exerc. 33 (6 Suppl), S587–610.

Dwyer, D.S., Mitchell, O.S., 1999. Health problems as determinants of retirement: Are self-rated measures endogenous? J. Health Econ. 18 (2), 173–193.

Eibich, P., 2015. Understanding the effect of retirement on health: Mechanisms and heterogeneity. J. Health Econ. 43, 1–12.

Ferrer-i-Carbonell, A., 2005. Income and well-being: an empirical analysis of the comparison income effect. J. Public Econ. 89 (5), 997–1019. http://dx.doi.org/10.1016/j.jpubeco.2004.06.003.

Ferrer-i-Carbonell, A., Frijters, P., 2004. How Important is Methodology for the estimates of the determinants of Happiness?*. Econ. J. 114 (497), 641–659. http://dx.doi.org/10.1111/j.1468-0297.2004.00235.x.

Fonseca, R., Kapteyn, A., Lee, J., Zamarro, G., 2017. Does retirement make you happy? A simultaneous equations approach. In: Insights in the Economics of Aging. University of Chicago Press, National Bureau of Economic Research, pp. 339–372. http://dx.doi.org/10.7208/chicago/9780226426709.001.0001.

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

French, E., 2005. The effects of health, wealth, and wages on labour supply and retirement behaviour. Rev. Econom. Stud. 72 (2), 395–427.

French, E., von Gaudecker, H.-M., Jones, J.B., 2016. The Effect of the Affordable Care Act on the Labor Supply, Savings, and Social Security of Older Americans. Working Papers 2016-353, (wp354), Ann Arbor, MI. University of Michigan Retirement Research Center (MRRC).

French, E., Jones, J.B., 2017. Health, health insurance, and retirement: A survey. Annu. Rev. Econ. 9, 383–409.

Grotting, M.W., Lillebo, O.S., 2020. Health effects of retirement: evidence from survey and register data. J. Popul. Econ. 33 (2), 671–704.

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. J. Hum. Resour. 49, 195–233.

Jones, A.M., Rice, N., Roberts, J., 2010. Sick of work or too sick to work? Evidence on self-reported health shocks and early retirement from the BHPS. Econ. Model. 27 (4), 866–880.

Macken, J., 2019. Work stress among older employees in Germany: Effects on health and retirement age. PLoS One 14 (2), e0211487.

Mao, H., Ostaszewski, K.M., Wang, Y., 2014. Optimal retirement age, leisure and consumption. Econ. Model. 43, 458–464.

Mazzonna, F., Peracchi, F., 2012. Ageing, cognitive abilities and retirement. Eur. Econ. Rev. 56, 691–710.

McGarry, K., 2004. Health and retirement do changes in health affect retirement expectations? J. Hum. Resour. 39 (3), 624–648.

Messe, P., 2011. Taxation of early retirement windows and delaying retirement: The French experience. Econ. Model. 28 (5), 2319–2341.

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

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

Nardi, M.D., French, E., Jones, J.B., McGee, R., 2021. Why Do Couples and Singles Save During Retirement?. NBER Working Papers 28828, (28828), National Bureau of Economic Research, Inc.

Neuman, K., 2008. Quit your job and get healthier? The effect of retirement on health. J. Labor Res. 29, 177–201.

Quinn, J.F., 1977. Microeconomic determinants of early retirement: A cross-sectional view of white married men. J. Hum. Resour. 12 (3), 329–346.

Rohwedder, S., Willis, R., 2010. Mental retirement. J. Econ. Perspect. 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. Arch. Intern. Med. 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. J. Vocat. Behav. 53 (1), 45–57.

Ströhle, A., 2009. Physical activity, exercise, depression and anxiety disorders. J. Neural Transm. 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. Eur. Sociol. Rev. 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 Ment. Health 14 (6), 652–669.

Wooldridge, J., 2002. Econometric Analysis of Cross Section and Panel Data. The MIT Press.

Zhu, R., 2016. Retirement and its consequences for womens health in Australia. Soc. Sci. Med. 163, 117–125.