The impact of health on wages: Evidence from Europe before and during the Great Recession

Manuel Flores[a], Melchor Fernández[a], Yolanda Pena-Boquete[b]

[a]University of Santiago de Compostela; Galician Studies and Development Institute (IDEGA)

[b]University of Vigo; Galician Studies and Development Institute (IDEGA)

October 31, 2014

ABSTRACT

This paper adds to the empirical literature of health as a potential endogenous explanatory variable in wage equations by addressing problems such as unobserved heterogeneity, sample selection and measurement error (in the health variable) in one comprehensive framework. Moreover, by using European individual-level panel data from before and during the Great Recession (GR)—which started in Europe in 2008—we gain insights into whether and how the current crisis has altered the relationship between health and wages. Our results provide empirical evidence of measurement error in the self-reported health variable when estimating its impact on wages for men, and of selectivity bias in wages for both men and women. We also show that in the period prior to the GR, working-age men (20–64 years old) who are in relatively better health (measured by a one-unit increase in a health index) have, on average, a 9 percent higher hourly wage rate. This effect is concentrated (and largest) among older workers (50–64 years old). Instead, during the GR the positive impact of health on wages disappears. One possible explanation for these findings is that *presenteeism* (i.e. attending work even though being sick)—which has become more common during the GR—may have reduced the impact of (poor) health on wages. With regard to working-age women (20–59 years old), we do not find evidence of an effect of health on wages, both before and during the GR.

JEL Classification: D00, I10, J20, J24, J30

Keywords: Health, Wages, Great Recession, EU-SILC

1. Introduction

According to economic theoretical models, health, as a component of human capital, affects wages through productivity. An individual in bad health is assumed to have a lower productivity, and thus a lower wage rate (Becker, 1962; Mushkin, 1962). Hence, if workers are paid according to their marginal productivity, wages will be determined by firm-level supply and demand factors and by Mincerian-type employee-level human capital (i.e. education and experience) and health effects (cf. Jäckle and Himmler, 2010).[1] However, while there is a wide support in the empirical literature of the positive impact of experience and education on wages (see Card (1999) for a survey), the relationship between wages and health is less clear-cut.

Yet health, and in particular its possible endogeneity in wage equations and labor market equations in general, has received a great deal of attention in the literature (see Currie and Madrian (1999) for a survey). Common reasons for this endogeneity are nonrandom sample selection, because we only observe wages for those who choose to work and because participation is possibly correlated with idiosyncratic changes in wages, and unobserved heterogeneity, as there may be unobserved characteristics such as parental socioeconomic status (Flores *et al.*, 2014) or even prenatal circumstances (Almond and Currie, 2011) that affect both health and wages. Another reason is the potential measurement error that arises in many studies, which, like ours, use self-reported health (SRH) status as a measure for actual health. This potential measurement error may stem from three sources: *pure* measurement error (see Bound *et al.*, 2001; Crossley and Kennedy, 2002), the justification bias (see Bound, 1991; Stern, 1989), or the basing of SRH on subjective judgment, which may hinder comparison across individuals (Kapteyn *et al.*, 2007; Meijer *et al.*, 2011).[2]

[1] Alternatively, Contoyannis and Rice (2001) argue that the (positive) relation between poor health and low wages may stem from employer beliefs that poor health correlates with unobserved characteristics that are negatively associated with productivity or from discrimination against individuals perceived to be in poor health (see also Currie and Madrian (1999, pp. 3332–3) and references therein).

[2] There is some consensus in the literature that the pure measurement error and the reporting differences are likely to attenuate the impact of SRH on labor market outcomes (including wages), whereas the justification bias will most probably exaggerate its impact (Bound, 1991; Brown *et al.*, 2010). The direction of the other biases due to unobserved heterogeneity and sample selection is, *a priori*, unclear and will depend, respectively, on the correlation between the unobserved characteristics and health, and on whether selection is positive or not, as well as on the correlation between participation (i.e. selection) and health.

Most previous studies that analyze the impact of health on wages have considered the aforementioned problems to different extents. For instance, Haveman *et al.* (1994) estimate a simultaneous equations model for work, wages and health. They use panel data for men with "histories of significant labor force attachment" and report that (lagged) poor health does affect both wages and work-time negatively. However, they do not consider the aforementioned endogeneity problems of unobserved heterogeneity, sample selection and measurement error in their subjective health variable (which is constructed from two questions on whether the individual is work limited by health, and by how much). Contoyannis and Rice (2001) use data from employed British men and women and find that the significant effect of (excellent) SRH on wages in ordinary least squares (OLS) models becomes insignificant for men and remains only marginally significant for women when individual fixed effects (FE) are controlled for. However, reduced psychological health is found to reduce hourly wages for men. Using further Hausman-Taylor type instrumental variables estimators leaves these results virtually unchanged. Brown *et al.* (2010) use a Health Index approach in a similar vein to Bound *et al.* (1999) to correct for measurement error in SRH, and additionally control for unobserved heterogeneity and sample selection bias. They do not find an effect of health on men's (reservation) wages in Britain.[3] Jäckle and Himmler (2010), on their side, use German data and compare different estimators that account for selectivity bias, unobserved heterogeneity and measurement error to assess the effect of health on wages. Their results suggest that health has in general a positive effect on wages for men but not for women.

Some studies have also explored the reverse causality of wages on health.[4] The empirical evidence on this issue is rather mixed and the models used often require additional assumptions for identification (see also Currie and Madrian, 1999, pp. 3320, 3331). For instance, Lee (1982) and Cai (2009), report a positive and no effect of wages on health for U.S. and Australian men, respectively. Also Haveman *et al.* (1994) model the reverse impacts of work-time and wages on health (although the estimation results that include the reverse effect of wages on health are not reported). Finally, it is also

[3] However, and as will be made clear in Section 3, one possible drawback within their approach is that they plug into the wage equation the fitted values of a health stock variable obtained from a first-stage health equation that is estimated to correct for measurement error *before* correcting for selectivity bias (Semykina and Wooldridge, 2008).

[4] Grossman (2001) discusses a theoretical model on this issue.

worth noting that all previous studies focus on prime-age individuals, except Lee (1982) who uses data from older workers.

Our contribution to the literature regarding the relationship between health and wages is twofold. First, we expand the findings of Jäckle and Himmler (2010) for Germany and use individual-level panel data from the European Union Statistics on Income and Living Conditions (EU-SILC) to assess the potential importance of unobserved heterogeneity, selection and measurement error when estimating the impact of health on wages for men and women in Europe, and by age groups. Second, we use data from before and during the Great Recession (GR), which started in Europe in 2008 (Arpaia and Curci, 2010), to gain insights into whether and how the current crisis has altered the relationship between health and wages, an issue not yet addressed in the literature. For instance, budget cuts during the GR have restricted (overall) access to health care (Karanikolos *et al.*, 2013), which we would expect to result in an increase in the impact of health on wages, on average. On the other hand, the rise in uncertainty and unemployment (Leduc and Zheng, 2012) combined with decreases in (un) employment protection during the GR (OECD, 2013) may be pushing some workers to go to work when they are sick (CareerBuilder, 2011). This phenomenon is known as *presenteeism*, and may, at least in the short run, reduce the impact of (poor) health on wages. Also, during recent years, there is some evidence of a reduction in the variable component of wages (Vandekerckhove *et al.*, 2012), the one that is likely to be more responsive to productivity-related components such as health, which might further reduce the impact of health on wages.

Our primary empirical findings show that in the period prior to the GR, and similar to Jäckle and Himmler's (2010) study for Germany, European working-age men (20–64 years old) who are in relatively better health (measured by a one-unit increase in a health index) have, on average, a 9 percent higher hourly wage rate. But in addition to their study, we show that this effect is concentrated (and largest) among older workers (50–64 years old). However, during the GR the positive impact of health on wages disappears. Two possible explanations for these findings are the abovementioned increase in *presenteeism* (i.e. attending work even though being sick) and the reduction of the (more) productivity-related component of wages during the current crisis. With regard to working-age women (20–59 years old), we do not find evidence of an effect of health on wages, neither before nor during the GR. Finally, and mainly for men, our results provide further empirical evidence of measurement error in the SRH variable when estimating its impact on wages, and of selectivity bias in wages.

The remainder of the paper is organized as follows. Section 2 describes the data and the main analytical variables. Section 3 outlines the empirical model and discusses a number of related econometric issues. Section 4 reports the estimation results and Section 5 analyzes their robustness. Section 6 summarizes the main findings and concludes.

2. Data and descriptive statistics

We use individual-level panel data from the European Union Statistics on Income and Living Conditions (EU-SILC), a harmonized and representative cross-national panel of the European population aged 16 and over. EU-SILC contains household and individual-level information on various components, e.g., income, work, health, housing and other social indicators about living conditions. It is a four-year rotational panel except for France and Luxemburg, where it is an eight-year and a pure panel, correspondingly. Therefore, we use two different panels for analyzing the period before (2005–2007) and during the GR (2008–2011), which correspond to release 2008 and 2011, respectively.[5]

EU-SILC covers the 27 countries of the European Union (EU) plus Croatia, Iceland, Turkey, Norway and Switzerland. However, it was not implemented in all countries at the same time. On one hand, Bulgaria, Romania, Croatia, Turkey and Switzerland took part in the EU-SILC project after 2005. On the other hand, in release 2011 data for this year were not available for Greece, France, Ireland, Sweden and Slovakia. In our analysis, we consider 15 countries that participate in the whole sample period (2005–2011) covering Northern (Denmark, Finland, and Norway), Central (Austria, Belgium, Luxembourg, and the Netherlands), Southern (Cyprus and Spain), and Eastern Europe (the Czech Republic, Estonia, Hungary, Lithuania, Poland, and Slovenia).

[5] The year 2008 is taken as the start of the GR in Europe (Arpaia and Curci, 2010) and, hence, is dropped from the release 2008.

Our empirical analysis is based on data for male (female) respondents aged 20–64 (20–59).[6] This selection yields 125,985 observations for 41,995 respondents for the period 2005–2007, and 131,200 observations for 32,800 respondents for the period 2008–2011. Panel attrition in EU-SILC is relatively low; of about 14 (11) percent by wave in the period 2005–2007 (2008–2011). In our analysis, we exclude students, and in general individuals below age 20, as they may not have established work patterns (Haveman *et al.*, 1994). Moreover, we leave out self-employed workers since their motives with respect to labor market participation are specific and their reported earnings and hours are a poor proxy for their hourly wage rate. Additionally, we do not consider individuals in compulsory military service or those in (early) retirement. We also exclude individuals that are permanently disabled (handicapped) because they are not likely to be paid according to their marginal productivity for reasons such as discrimination and due to the fact that most of them work at sheltered workshops with subsidized wages (Jäckle and Himmler, 2010). Missing values forces a 32 (34) percent reduction in sample size for the period 2005–2008 (2008–2011). The result is a balanced panel comprising 61,071 observations for 9,914 male and 10,443 female European respondents for the period 2005–2007, and 60,528 observations for 7,128 male and 8,004 female European respondents for the period 2008–2011.[7]

Details on the definitions of all variables used in the empirical analysis are given in appendix Table A.1. The (log) hourly gross wage rate—measured in PPP's adjusted 2005 €—is defined for employed individuals and is obtained from dividing the amount of gross wage earnings by the number of hours (usually) worked. Countries such as Greece, Italy, Latvia and Portugal report only net wage earnings in the first waves and were, therefore, excluded from the analysis.[8] Self-reported health (SRH) status is rated on a five-point scale (from 1 to 5: very bad, bad, fair, good, and very good), and the variable determining selection, i.e. participation in the labor market, is defined as 1 if the individual works and 0 otherwise. Among our explanatory variables, experience

---

[6] For women, mandatory retirement age in many of our sample countries (e.g., in the Eastern European countries) is age 60 (www.oecd.org/els/social/pensions/PAG).

[7] Our choice of using a balanced panel is dictated by the correction procedures that we use, which are obtained under the assumption that there is no attrition. In particular, the variables in vector z_i are always observed (see Section 3 and Semykina and Wooldridge (2008, pp. 33) for more details).

[8] The reason for not considering net hourly wages is that they are likely to be affected by family conditions and tax legislation not related to an individual's labor market productivity.

refers to actual experience, hence avoiding additional measurement error in this variable.[9] In order to capture differences in labor market institutions across the countries within our sample, we also include country- and gender-specific employment and unemployment rates in our empirical model in Section 3. This information is taken from the Labour Force Survey available from Eurostat (http://epp.eurostat.ec.europa.eu/). Table 1 provides summary statistics on these and other health, demographic and socioeconomic characteristics used in the empirical analysis for both periods, before and during the GR, and for working-age employed and nonemployed men and women. Similar descriptive statistics by age groups are given in appendix Tables A.2 to A.4.

As shown in Table 1, both before (2005–2007) and during the GR (2008–2011), men earn higher (real) hourly wages than women, and employed men and women report better health than their nonemployed peers. The most notable change, however, when comparing the two time periods is the increase in the probability of working in women, which increases from 0.73 to 0.79 (this is known as the “added-worker effect” in the literature). This increase, moreover, is largest among older women (see Tables A.2–A.4 in the appendix). Instead, for men, the probability of working in our sample remains about the same in both periods (between 0.92–0.93). Furthermore, we do not find large changes in real hourly wages and SRH across time, with the possible exception of employed and nonemployed men for which real hourly wages and SRH increase slightly, respectively.

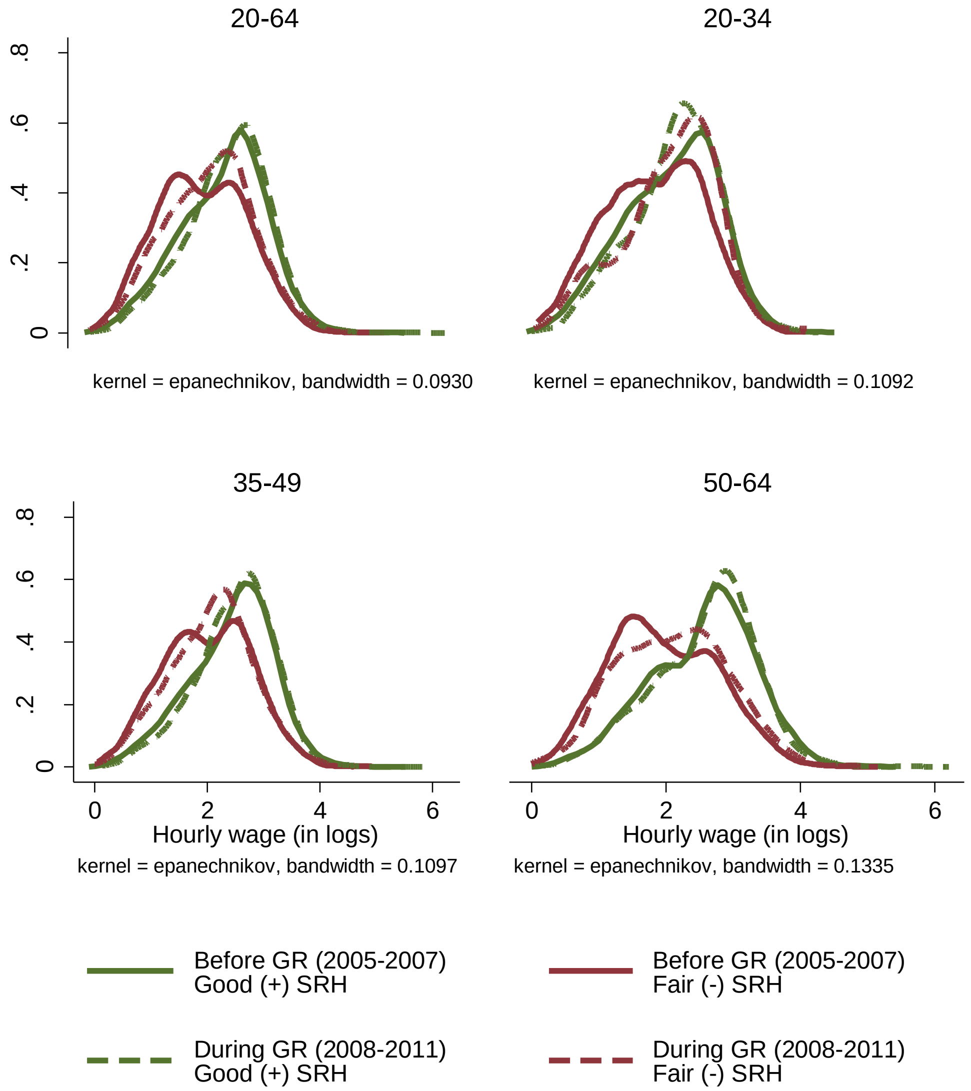
Figures 1 and 2 show kernel density estimates for the logarithm of (real) hourly wages before and during the GR for those who report good or very good health (Good (+) SRH) and fair or worse health (Fair (-) SRH), by age groups and for men and women, respectively. As the figures show, across age groups and gender, the main changes in the density of wages occur for those who report worse health, whose wage density shifts slightly to the right, mainly because of a reduction in the lower tail of the density. Thus, the difference in wage densities between those reporting good or very good and fair or worse health becomes smaller during the GR, which suggests a diminishing role for health in wages during this period.

[9] We also had to exclude the UK from our analysis, as it reports no data on experience.

Table 1: Descriptive statistics

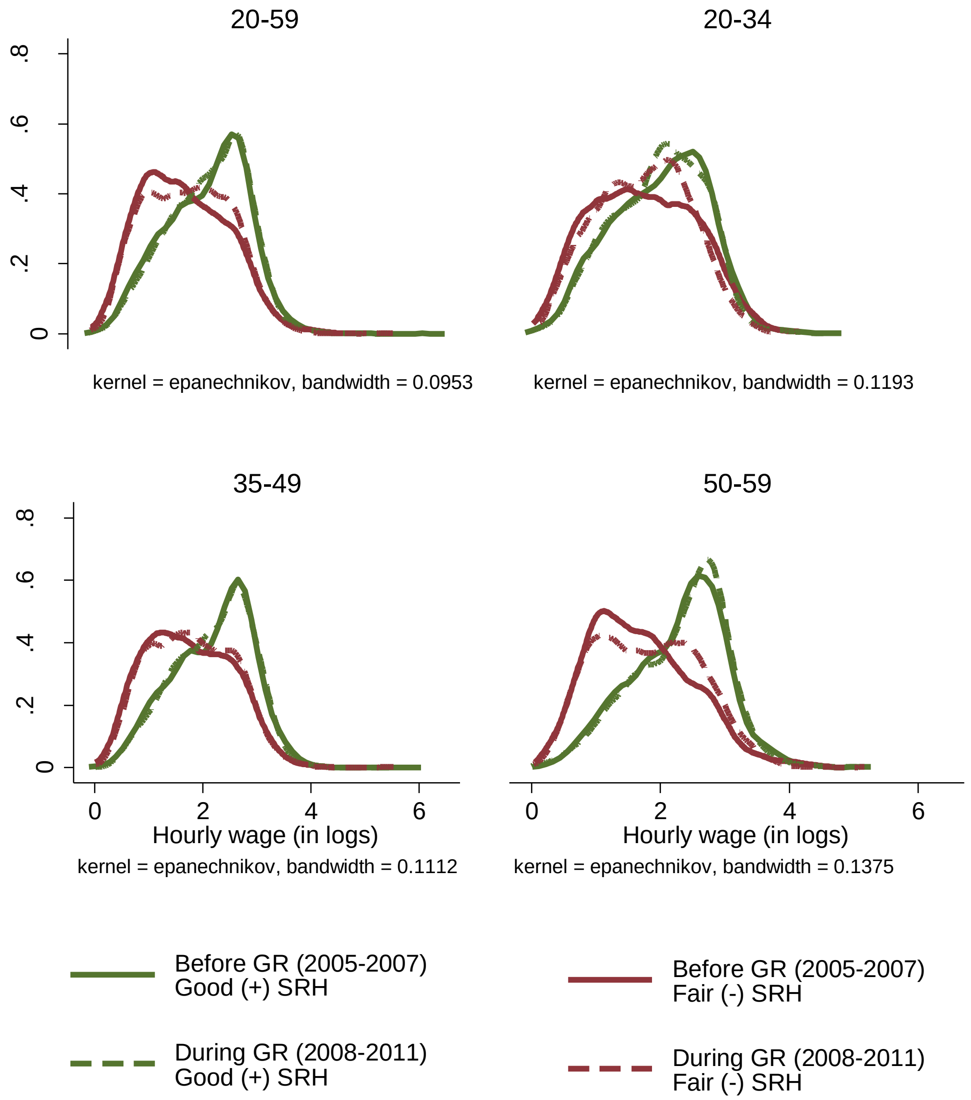
	Employed				Nonemployed			
	2005–2007		2008–2011		2005–2007		2008–2011	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
Men (20–64)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Hourly gross wage rate	12.66	10.55	13.52	11.73				
Self-reported health (1-5)	4.05	0.76	4.07	0.73	3.54	0.99	3.67	0.9
1+ chronic diseases	0.17	0.38	0.18	0.39	0.28	0.45	0.27	0.45
Limited with GALI (1-3)	1.13	0.4	1.13	0.39	1.33	0.62	1.28	0.55
Experience	20.7	11.35	20.18	11.5	17.13	12.58	17.72	12.6
ISCED 0–2	0.18	0.38	0.2	0.4	0.36	0.48	0.38	0.48
ISCED 3-4	0.56	0.5	0.49	0.5	0.56	0.5	0.52	0.5
ISCED 5-6	0.26	0.44	0.31	0.46	0.08	0.28	0.11	0.31
Annual income from nonemployment	18,530	20,415	19,779	20,604	15,283	16,346	15,649	15,800
Household size	3.3	1.39	3.26	1.41	3.2	1.6	3.28	1.73
Single	0.24	0.43	0.24	0.43	0.44	0.5	0.49	0.5
Age	40.97	10.36	41.98	10.03	42.25	11.88	42.55	11.75
Unemployment rate (%)	5.93	3.01	7.82	5.04	7.91	4.08	10.9	5.73
Employment rate (%)	72.68	6.07	71.07	6.06	68.76	6.6	67.64	5.61
Prob. work (entire sample)	0.92	0.27	0.93	0.26				
Observations (N)	27,462		26,386		2,280		2,126	
Women (20–59)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Hourly gross wage rate	10.31	10.03	10.34	8.6				
Self-reported health (1-5)	4.01	0.77	4.01	0.74	3.8	0.89	3.8	0.83
1+ chronic diseases	0.18	0.39	0.2	0.4	0.24	0.43	0.26	0.44
Limited with GALI (1-3)	1.14	0.4	1.15	0.41	1.24	0.54	1.23	0.5
Experience	18.39	10.25	17.32	10.41	9.43	9.04	9.79	9.24
ISCED 0–2	0.15	0.36	0.16	0.37	0.42	0.49	0.39	0.49
ISCED 3-4	0.52	0.5	0.47	0.5	0.46	0.5	0.46	0.5
ISCED 5-6	0.33	0.47	0.37	0.48	0.12	0.32	0.14	0.35
Annual income from nonemployment	25,790	27,946	27,300	29,630	32,700	33,229	27,459	38,773
Household size	3.18	1.29	3.19	1.32	3.76	1.49	3.79	1.54
Single	0.28	0.45	0.27	0.44	0.15	0.35	0.17	0.38
Age	40.72	9.44	41.81	9.14	41.38	10.29	41.54	10.25
Unemployment rate (%)	7.82	3.66	8.39	4.69	8.77	3.87	10.61	5.55
Employment rate (%)	58.66	7.48	59.7	7.3	55.77	5.79	56.6	5.99
Prob. work (entire sample)	0.73	0.44	0.79	0.41				
Observations (N)	22,954		25,153		8,375		6,863	

Figure 1: Kernel density estimates for the logarithm of hourly wages before and during the Great Recession by levels of SRH and age groups (Men)



Source: Author calculations based on EU-SILC (2005-2007, 2008-2011). The figures show kernel density estimates for the logarithm of real gross hourly wages before and during the Great Recession for men and by age groups for those who report good or very good (Good (+) SRH) and fair or worse health (Fair (-) SRH).

Figure 2: Kernel density estimates for the logarithm of hourly wages before and during the Great Recession by levels of SRH and age groups (Women)



Source: Author calculations based on EU-SILC (2005-2007, 2008-2011). The figures show kernel density estimates for the logarithm of real gross hourly wages before and during the Great Recession for women and by age groups for those who report good or very good (Good (+) SRH) and fair or worse health (Fair (-) SRH)..

3. The empirical model

Our aim in this paper is to estimate the impact of health on wages for the entire population (with panel data), but we only observe the wage rates for individuals who work. As already mentioned in Section 1, this creates a selection problem as the decision to participate in the labor market is likely to be nonrandomly determined and this is unlikely to be fully covered by observable factors (Heckman, 1979). Moreover, exploiting the panel nature of our data using, e.g., a fixed effects approach will not solve this problem unless the selection process is time constant (Dustmann and Rochina-Barrachina, 2007). If unobserved time-varying, e.g., lifestyle-related factors such as health determinants or motivation affect selection, this kind of selection will influence wages through the error term and lead to inconsistent estimation (Brown *et al.*, 2010; Jäckle and Himmler, 2010). Therefore, to tackle this selection problem, we estimate the following system of equations that model the relationship between health and wages:

$$w_{it} = \beta_0 + x_{it}\beta_1 + c_{i1} + u_{it1}, \qquad (1)$$

$$S^*_{it} = \delta_0 + z_{it}\delta_1 + c_{i2} + u_{it2}, \qquad S_{it} = 1\left[S^*_{it} > 0\right]. \qquad (2)$$

Equation (1) is a wage equation, and Equation (2) is a reduced-form equation that describes an individual's decision to participate in the labor market. The subscripts i and t index individuals and time periods, respectively. w_{it} is the logarithm of the hourly (gross) market wage and we only observe it when S^*_{it}, the latent propensity to work, is positive. S_{it} denotes actual labor market participation, and 1[.] is an indicator function which equals one if its argument is true.

The system of equations comprises two sets of explanatory variables, x_{it} and z_{it}. In the wage equation, x_{it} is a 1 x K vector of explanatory variables, with β_1 being the corresponding parameter (column) vector. The variables in x_{it} include SRH (which, as discussed below is not necessarily exogenous due to measurement error problems), as well as variables such as experience and macro-level labor market variables that affect both wages and labor market participation (Montuenga *et al.*, 2003). In the participation equation, z_{it} is a 1 x G vector of explanatory variables including all exogenous variables in x_{it} among others, with δ_1 being the corresponding parameter (column) vector.

Both the wage and participation equations include unobserved heterogeneity in the form of time-constant individual effects denoted by c_{i1} and c_{i2} which are possibly correlated with x_{it} and z_{it}, and two error terms u_{it1} and u_{it2}, with u_{it1} likely to have a non-zero conditional expectation due to selective sampling, i.e., wages being observed for labor market participants only. Hence, the model combines problems of (time-constant) unobserved heterogeneity, sample selectivity and possibly additional endogeneity in one explanatory variable due to measurement error.

To deal with these issues, in this paper we use Semykina and Wooldridge's (2010) framework, which basically extends Wooldridge's (1995) method for correcting for sample selection in fixed effects models by allowing some variable(s) in the main equation (wage equation) to be correlated with the (idiosyncratic) error term. Therefore, z_{it} needs to include at least one variable that is correlated with the possibly endogenous explanatory variable in x_{it} (in our case, SRH) but is at the same time not correlated with the error term in the wage equation u_{it1}. More specifically, z_{it} is assumed to be strictly exogenous in Equation (1) conditional on c_{i1}. Here, we include two health limitation variables (chronic conditions and a Global Activity Limitations Indicator, GALI) that are assumed to be exogenous instruments for SRH. Although not strictly necessary, we also add exclusion restrictions to z_{it} such as nonlabor income and other household composition variables (household size and marital status) that drive selection but can be omitted from the wage equation. We also define z_{it1} as a subset of z_{it} including all the exogenous variables in x_{it} (i.e., all variables except SRH) plus the two health limitations variables.

To deal with the issue of unobserved heterogeneity, Semykina and Wooldridge (2010)—in a similar way to Wooldridge (1995)—use Mundlak's (1978) device and write the unobserved individual effects as a linear projection onto the (individual) time averages of z_{it}, denoted $\bar{z}_{i.}$, and an error term. This is particularly important for the participation equation (2), as it allows to circumvent the incidental parameters problem in nonlinear models (Wooldridge, 2002). After introducing some self-explanatory notation, we can rewrite Equations (1) and (2) in the following way:

$$w_{it}^{*} = \beta_0 + x_{it}\beta_1 + \bar{z}_{i.}\xi_1 + v_{it1}\,, \qquad (3)$$

$$S_{it} = 1\left[d_0 + z_{it}d_1 + \bar{z}_{i.}x_2 + n_{it2} \square 0\right], \qquad (4)$$

where the composite error terms $v_{itj} = \varepsilon_{ij} + u_{itj}, j = 1, 2$ are likely to be correlated with each other due to sample selectivity. Semykina and Wooldridge (2010) (and Wooldridge, 1995) basically extend Heckman's (1979) procedure to an unobserved effects framework, where v_{it1} is a linear function of v_{it2} and mean independent of z_i conditional on v_{it2}:[10]

$$E\left(n_{it1} \mid n_{it2}, z_{i1}, \ldots, z_{iT}\right) = gn_{it2}. \qquad (5)$$

While this is a formal assumption which keeps the model manageable, it still provides flexibility by allowing residual correlation to subsist even after the introduction of the Mundlak terms controlling for persistent individual features in Equations (3) and (4). If we substitute Equation (5) into (3), we get:

$$w_{it}^{*} = \beta_0 + x_{it}\beta_1 + \bar{z}_{i.}\xi_1 + \gamma\lambda_{it} + e_{it1}\,, \qquad (6)$$

where e_{it1} is an error term.

Finally, the probit selection model requires the composite error term $v_{it2} = \varepsilon_{i2} + u_{it2}$ to be standard normally distributed. The participation equation (4) then forms a sequence of T standard probit models that are estimated to calculate the inverse Mills ratios (IMR), $\hat{\lambda}_{it}$. The final estimating equation is obtained by substituting $\hat{\lambda}_{it}$ for λ_{it} in Equation (6).

As summarized in Semykina and Wooldridge (2010), a consistent way of estimating Equation (6) is to use, for the selected sample ($S_{it} = 1$), pooled 2SLS using z_{it1}, $\bar{z}_{i.}$, and

[10] This corresponds to part (iv) of Assumption 4.1.1 in Semykina and Wooldridge (2010).

$\hat{\lambda}_{it}$ as instruments. $\hat{\lambda}_{it}$ can be interacted with time dummies to allow γ to be different across time periods. Standard errors robust to serial correlation and heteroskedasticity are calculated as suggested in Semykina and Wooldridge (2010), and are adjusted for the additional variation introduced by the estimation of T probit models in the first step.[11]

4. Estimation results

As explained in Section 1, one of the aims of this study is to gain insights into whether the current crisis has altered the relationship between health and wages. We, therefore, use data from before (2005–2007) and during (2008–2010) the GR and discuss the estimation results in Sections 4.1 and 4.2, respectively. Moreover, because the effect of health on labor market outcomes is expected to increase with age (Currie and Madrian, 1999), we also explore if the impact of health differs across age groups (in addition to gender). Furthermore, in order to investigate the potential role of unobserved heterogeneity, selection and measurement error when estimating the impact of health on wages, we compare Semykina and Wooldridge's (2010) estimator (SemWool10) to the one of Wooldridge (1995) (Wool95) and computationally undemanding pooled OLS, pooled two-stage least squares (2SLS), fixed effects (FE), and fixed effects two-stage least squares (FE-2SLS) estimators.[12] This also enables us to compare our results with those obtained in some of the studies discussed in Section 1. For the sake of completeness, it is worth noting that we include dummies for every age year (in the OLS and 2SLS models) and survey year (in all models), and thereby control for (birth) cohort effects. In addition, in the OLS and 2SLS models we add dummies for medium and high educational attainment, and for country.

The first step in the estimation procedures of Semykina and Wooldridge (2010) and Wooldridge (1995) is to estimate the participation equation as a sequence of T standard cross-sectional probit models to calculate the selectivity correction terms, i.e., the inverse Mills ratios (IMR). These results are reported in appendix tables A.5 to A.8 for

[11] We, therefore, use the Stata do-files that are available in Semykina's webpage (http://mailer.fsu.edu/~asemykina/). Instead of using the analytical formulae for the asymptotic variance one can also apply "panel bootstrap". The bootstrap estimator will be consistent for $N \to \infty$ and T fixed (Semykina and Wooldridge, 2010).

[12] See Semykina and Wooldridge (2008) and Jäckle and Himmler (2010) for a similar analysis using U.S. and German data, respectively.

men and women. In the wage equations, we tested for (contemporaneous) selection bias as proposed in Semykina and Wooldridge's (2010) Procedure 3.1 which requires estimating for the selected sample ($S_{it} = 1$) by FE-2SLS the equation $w_{it} = \beta_0 + x_{it}\beta_1 + c_{i1} + \gamma_t \hat{\lambda}_{it} + e_{it1}$, using z_{it1} and $\hat{\lambda}_{it}$ as instruments. The resulting p-values from Wald tests to test $H_0 : \gamma_1 = ... = \gamma_T = 0$ are reported in Table 2 for both men and women, and by age groups. These indicate the presence of selection bias in the (main) wage equations for men although not for women when using the FE-2SLS estimator, which justifies (in part) the use of the SemWool10 estimator (see Columns 2).[13]

Table 2: Tests for selection bias by gender and age groups. P-values from Wald tests on the joint significance of 3 (period 2005–2007) and 4 (period 2008–2011) IMR are provided.[a]

	2005–2007		2008–2011	
	FE[b]	FE-2SLS[c]	FE[b]	FE-2SLS[c]
Men	(1)	(2)	(1)	(2)
20–64	0.00	0.01	0.00	0.01
20–34	0.10	0.18	0.61	0.66
35–49	0.02	0.04	0.03	0.01
50–64	0.09	0.09	0.21	0.16
Women	(1)	(2)	(1)	(2)
20–59	0.75	0.71	0.61	0.75
20–34	0.11	0.19	0.95	0.85
35–49	0.43	0.59	0.76	0.79
50–59	0.31	0.19	0.80	0.48

[a] FE and FE-2SLS estimation. Robust p-values are reported.
[b] Under the null hypothesis the FE estimators are consistent.
[c] Under the null hypothesis the FE-2SLS estimators are consistent.

Finally, in the 2SLS and FE-2SLS models, we provide two additional tests on the health limitations variables (chronic conditions and GALI), which are shown in the bottom parts of Tables 3 to 8. The first is an F-test that the coefficients on the health limitations in the first stage (population) health reduced-form regressions are all zero. The test statistics are always high and well over the rule-of-thumb of ten (Bound *et al.*, 1995). The second is an overidentification test where the null hypothesis is that the health limitations variables are orthogonal to the error term in the wage equation. As shown by the p-values, in the FE-2SLS models we do not reject the null hypothesis at any sensible level for men and do so only for older female workers and in the 2005–2007 period, where we reject it at a five percent significance level. This, however, is to be expected,

[13] We also tested for selection bias in the FE model using Wooldridge's (1995) Procedure 3.2. As expected, the results from this test also indicate the presence of selection bias in the (main) wage equations for men but not for women (see Columns 1 in Table 2). However, it is worth mentioning that another condition required for consistency of, e.g., the FE-2SLS estimator (apart from no contemporaneous selection bias) is that selection in one time period is not correlated with the idiosyncratic errors in other time periods (Semykina and Wooldridge 2008, pp. 20).

as overidentification tests that ignore selection (and/or unobserved heterogeneity) will tend to reject the validity of instruments too often (Semykina, 2012).

4.1. Before the Great Recession (2005–2007)

Wage equations

The top and bottom panels in Table 3 contain the estimation results of the wage equation for the period prior to the GR for men and women, respectively. As shown in the top part, men in better health earn higher (hourly) wages, although the magnitude differs across models. For instance, in Column 1 the parameter of the health variable using the pooled OLS estimator (0.07) is higher than the coefficient in the FE model (0.012), which suggests a positive correlation between the individual FE and health, with those individuals that are more productive (and, hence, earn higher wages) having unobserved characteristics which lead to better health. Controlling furthermore for selection in the Wooldridge (1995) estimator leaves the coefficient (0.015) and significance level virtually unchanged. Turning to the 2SLS models, a comparison of the parameters shows that the coefficients of the health variable in columns 1, 2 and 3 are smaller than their 2SLS counterparts in columns 4, 5 and 6, respectively, which is to be expected if SRH is error-ridden (see also Cai, 2009; Jäckle and Himmler, 2010).[14] Using the 2SLS estimator yields the highest parameter estimate of 0.106, whereas accounting for unobserved heterogeneity (on top of measurement error) in the FE-2SLS estimator again scales the health coefficient down to 0.055. Controlling furthermore for nonrandom selection into the workforce increases the parameter to 0.085.

The same six econometric models using the female sample are presented in the bottom panel of Table 3. The results, however, are much less intuitive than in the male sample. Throughout the specifications, only the pooled OLS and 2SLS estimators yield a significant and rather similar effect of health on wages, which is substantially smaller when using the Wooldridge (1995) estimator, and insignificant in the other specifications.

[14] As explained in Flores and Kalwij (2013), such an attenuation bias in the impact of health on wages can most likely be attributed to a standard errors-in-variables downward bias in the effect of SRH on wages because of a dominating (pure and reporting) measurement error in SRH (see Bound, 1991, p. 111; Bound *et al.*, 1999).

Table 3: Wage equations, 2005–2007[a]

Men (20–64)	OLS (1)	FE (2)	Wool95 (3)	2SLS (4)	FE-2SLS (5)	SemWool10 (6)
Health (+1)	0.070*** (0.005)	0.012*** (0.004)	0.015*** (0.005)	0.106*** (0.013)	0.055*** (0.017)	0.085*** (0.020)
Experience	0.014*** (0.002)	0.005*** (0.002)	0.010*** (0.002)	0.014*** (0.002)	0.005*** (0.002)	0.011*** (0.002)
Experience^2	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)
ISCED 3-4	0.215*** (0.012)			0.211*** (0.012)		
ISCED 5-6	0.614*** (0.014)			0.605*** (0.014)		
Observations	27462	27462	27462	27462	27462	27462
F-test[b]				1920.97 (2)	435.64 (2)	
Overid. test[c]				11.46 (0.00)	0.21 (0.65)	
Unobserved effects[d]			194.10 (0.00)			119.80 (0.00)
Women (20–59)	(1)	(2)	(3)	(4)	(5)	(6)
Health (+1)	0.072*** (0.006)	0.009 (0.005)	0.013** (0.007)	0.059*** (0.012)	-0.006 (0.020)	0.001 (0.023)
Experience	0.012*** (0.002)	0.003 (0.002)	0.009*** (0.003)	0.012*** (0.002)	0.003 (0.002)	0.009*** (0.003)
Experience^2	-0.000** (0.000)	-0.000 (0.000)	-0.000*** (0.000)	-0.000** (0.000)	-0.000 (0.000)	-0.000*** (0.000)
ISCED 3-4	0.230*** (0.013)			0.231*** (0.013)		
ISCED 5-6	0.659*** (0.014)			0.663*** (0.014)		
Observations	22954	22954	22954	22954	22954	22954
F-test[b]				1673.08 (2)	387.73 (2)	
Overid. test[c]				8.43 (0.00)	0.82 (0.36)	
Unobserved effects[d]			194.06 (0.00)			119.88 (0.00)

[a] The dependent variable is the PPP-adjusted hourly gross wage rate (in logs). All estimates include country- and gender-specific employment and unemployment rates and dummy variables for survey year. Columns (1) and (4) also include dummy variables for age year and country. Columns (3) and (6) also include interactions of the IMR with the survey year dummy variables. Standard errors robust to serial correlation and heteroskedasticity are in parentheses. The standard errors in Columns (3) and (6) are also corrected for the probit first-stage estimation. Significance levels: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

[b] F-test statistic with degrees of freedom in parentheses. The null hypothesis is that the health limitations (Chronic and GALI) have no joint impact on SRH.

[c] Overidentification test statistic with corresponding p-value in parentheses. The null hypothesis is that the health limitations variables (Chronic and GALI) are orthogonal to the error term in the wage equation (2.6).

[d] χ^2 test statistics with corresponding p-value in parentheses for the joint significance of 8 and 9 variables (vector $\overline{z}_{i.}$) are reported.

Heterogeneous effects across age groups

The impact of health on labor market outcomes is likely to increase with age (Currie and Madrian, 1999). To explore this possibility we re-estimate the six previous econometric models for young (20–34), middle-aged (35–49), and older male (50–64) and female (50–59) workers. Because our main interest is concerning the impact of health on wages, we again discuss only the results from the wage equations, which are shown in Table 4 for men and in Table 5 for women.

As shown in Columns 1 and 3 in Table 4 for men, the parameter estimates obtained when using the pooled OLS and 2SLS estimators increase (slightly) with age, but remain significant when accounting for individual unobserved heterogeneity and sample selection for older workers only. For the latter, the health coefficient ranges from 0.021 (FE model) to 0.164 (SemWool10). Table 5 contains the estimation results for women. Only the pooled OLS estimator yields a significant effect of health on wages for all age groups, which, again, slightly increases with age. Moreover, this effect is only robust to individual unobserved heterogeneity and sample selection for middle-aged women. However, the FE-2SLS and SemWool10 estimators yield an insignificant effect of health on wages for all female age groups.

In sum, in accordance with Jäckle and Himmler (2010) we find that once we take account of the potential problems of unobserved heterogeneity, sample selection and measurement error in SRH, health has a significant impact on wages for men but not for women in Europe. In addition to their study, our results show that the positive impact of health on wages for men is driven by older workers, which is in line with the expectation of health having an increasing impact on labor market outcomes with age (Currie and Madrian, 1999).

Table 4: Wage equations by age groups, 2005–2007 (Men)[a]

	OLS	FE	Wool95	2SLS	FE-2SLS	SemWool10
Age 20–34	(1)	(2)	(3)	(4)	(5)	(6)
Health (+1)	0.051***	0.004	0.006	0.097***	0.036	0.070
	(0.010)	(0.009)	(0.010)	(0.033)	(0.041)	(0.049)
Experience	0.034***	0.014**	0.021***	0.034***	0.014**	0.020***
	(0.006)	(0.007)	(0.008)	(0.006)	(0.007)	(0.008)
Experience^2	-0.002***	-0.001***	-0.001***	-0.002***	-0.001***	-0.001**
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)
ISCED 3-4	0.200***			0.195***		
	(0.023)			(0.022)		
ISCED 5-6	0.499***			0.492***		
	(0.030)			(0.029)		
Observations	7545	7545	7545	7545	7545	7545
F-test[b]				296.18 (2)	82.36 (2)	
Overid. test[c]				3.52 (0.06)	0.07 (0.80)	
Unobs. effects[d]			61.01 (0.00)			49.12 (0.00)
Age 35–49	(1)	(2)	(3)	(4)	(5)	(6)
Health (+1)	0.067***	0.013*	0.010	0.108***	0.045*	0.037
	(0.009)	(0.007)	(0.007)	(0.020)	(0.026)	(0.027)
Experience	0.007*	0.000	0.004	0.006*	-0.000	0.004
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Experience^2	-0.000	0.000	-0.000	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ISCED 3-4	0.246***			0.242***		
	(0.019)			(0.019)		
ISCED 5-6	0.664***			0.655***		
	(0.024)			(0.024)		
Observations	10821	10821	10821	10821	10821	10821
F-test[b]				802.99 (2)	206.22 (2)	
Overid. test[c]				3.91 (0.05)	0.06 (0.80)	
Unobs. effects[d]			68.6 (0.00)			41.24 (0.00)
Age 50–64	(1)	(2)	(3)	(4)	(5)	(6)
Health (+1)	0.079***	0.021***	0.036***	0.110***	0.101***	0.164***
	(0.011)	(0.008)	(0.009)	(0.021)	(0.033)	(0.039)
Experience	0.006*	0.004	0.007	0.006*	0.003	0.006
	(0.003)	(0.003)	(0.005)	(0.003)	(0.003)	(0.005)
Experience^2	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ISCED 3-4	0.200***			0.197***		
	(0.026)			(0.026)		
ISCED 5-6	0.657***			0.647***		
	(0.031)			(0.032)		
Observations	5896	5896	5896	5896	5896	5896
F-test[b]				606.68 (2)	105.09 (2)	
Overid. test[c]				2.97 (0.08)	0.32 (0.57)	
Unobs. effects[d]			45.27(0.00)			17.84 (0.00)

[a] The dependent variable is the PPP-adjusted hourly gross wage rate (in logs). All estimates include country- and gender-specific employment and unemployment rates and dummy variables for survey year. Columns (1) and (4) also include dummy variables for age year and country. Columns (3) and (6) also include interactions of the IMR with the survey year dummy variables. Standard errors robust to serial correlation and heteroskedasticity are in parentheses. The standard errors in Columns (3) and (6) are also corrected for the probit first-stage estimation. Significance levels: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

[b] F-test statistic with degrees of freedom in parentheses. The null hypothesis is that the health limitations (Chronic and GALI) have no joint impact on SRH.

[c] Overidentification test statistic with corresponding p-value in parentheses. The null hypothesis is that the health limitations variables (Chronic and GALI) are orthogonal to the error term in the wage equation (2.6).

[d] χ^2 test statistics with corresponding p-value in parentheses for the joint significance of 8 and 9 variables (vector $\overline{z}_{i.}$) are reported.

Table 5: Wage equations by age groups, 2005–2007 (Women)[a]

	OLS	FE	Wool95	2SLS	FE-2SLS	SemWool10
Age 20–34	(1)	(2)	(3)	(4)	(5)	(6)
Health (+1)	0.054***	-0.004	-0.011	-0.005	-0.018	-0.079*
	(0.012)	(0.013)	(0.014)	(0.030)	(0.040)	(0.046)
Experience	0.020***	0.007	0.020**	0.020***	0.007	0.020**
	(0.007)	(0.008)	(0.009)	(0.007)	(0.008)	(0.009)
Experience^2	-0.001*	-0.001	-0.001**	-0.001*	-0.001	-0.001**
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)
ISCED 3-4	0.182***			0.188***		
	(0.027)			(0.027)		
ISCED 5-6	0.571***			0.585***		
	(0.030)			(0.031)		
Observations	5809	5809	5809	5809	5809	5809
F-test[b]				249.11(2)	75.33 (2)	
Overid. test[c]				1.16 (0.28)	0.02 (0.88)	
Unobs effects[d]			40.15 (0.00)			27.98 (0.00)
Age 35–49	(1)	(2)	(3)	(4)	(5)	(6)
Health (+1)	0.065***	0.016**	0.024***	0.055***	-0.000	0.015
	(0.008)	(0.007)	(0.009)	(0.017)	(0.028)	(0.032)
Experience	0.012***	0.005	0.009**	0.012***	0.005	0.010**
	(0.004)	(0.003)	(0.005)	(0.004)	(0.003)	(0.005)
Experience^2	-0.000	-0.000*	-0.000*	-0.000	-0.000	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ISCED 3-4	0.228***			0.230***		
	(0.020)			(0.020)		
ISCED 5-6	0.686***			0.689***		
	(0.022)			(0.022)		
Observations	10166	10166	10166	10166	10166	10166
F-test[b]				743.71 (2)	196.31 (2)	
Overid. test[c]				0.75 (0.39)	0.22 (0.64)	
Unobs effects[d]			93.98 (0.00)			51.13 (0.00)
Age 50–59	(1)	(2)	(3)	(4)	(5)	(6)
Health (+1)	0.097***	0.021**	0.018	0.120***	0.014	0.039
	(0.014)	(0.011)	(0.012)	(0.024)	(0.035)	(0.037)
Experience	0.011***	0.008**	0.003	0.011***	0.008**	0.003
	(0.004)	(0.003)	(0.005)	(0.004)	(0.003)	(0.004)
Experience^2	-0.000	-0.000**	-0.000	-0.000	-0.000**	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ISCED 3-4	0.248***			0.246***		
	(0.026)			(0.026)		
ISCED 5-6	0.713***			0.707***		
	(0.032)			(0.032)		
Observations	4078	4078	4078	4078	4078	4078
F-test[b]				485.56 (2)	79.65 (2)	
Overid. test[c]				8.64 (0.00)	6.60 (0.01)	
Unobs effects[d]			45.91 (0.00)			19.54 (0.00)

[a] The dependent variable is the PPP-adjusted hourly gross wage rate (in logs). All estimates include country- and gender-specific employment and unemployment rates and dummy variables for survey year. Columns (1) and (4) also include dummy variables for age year and country. Columns (3) and (6) also include interactions of the IMR with the survey year dummy variables. Standard errors robust to serial correlation and heteroskedasticity are in parentheses. The standard errors in Columns (3) and (6) are also corrected for the probit first-stage estimation. Significance levels: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

[b] F-test statistic with degrees of freedom in parentheses. The null hypothesis is that the health limitations (Chronic and GALI) have no joint impact on SRH.

[c] Overidentification test statistic with corresponding p-value in parentheses. The null hypothesis is that the health limitations variables (Chronic and GALI) are orthogonal to the error term in the wage equation (2.6).

[d] χ^2 test statistics with corresponding p-value in parentheses for the joint significance of 8 and 9 variables are reported.

4.2. During the Great Recession (2008–2011)

Wage equations

The top and bottom panels in Table 6 contain the estimation results of the wage equation during the GR period for men and women, respectively. As shown in the top panel for men, and similar to the period prior to the GR, the parameter of the health variable using the pooled OLS and 2SLS estimators (0.081 and 0.092, respectively) are larger than the coefficients in the FE and FE-2SLS estimators (0.008 and 0.033, respectively); all these estimated effects are significantly different from zero at the five percent level. However, contrary to the period prior to the GR, accounting for sample selection in the SemWool10 and Wool95 estimators reduces the health effects further, which become insignificant at any sensible level.

The estimation results using the female sample are presented in the bottom panel of the table and are very similar to the ones using the sample period prior to the GR. Only the pooled OLS and 2SLS estimators yield a significant and rather similar effect of health on wages (0.068 and 0.061, respectively), which becomes insignificant once individual unobserved heterogeneity is accounted for (FE and FE-2SLS), and remains insignificant when further controlling for sample selection (Wool95 and SemWool10).

Heterogeneous effects across age groups

As done previously, we re-estimate the six econometric models for young (20–34), middle-aged (35–49), and older male (50–64) and female (50–59) workers and discuss the results obtained from the wage equations, which are given in Tables 7 and 8 for men and women, respectively.

Table 6: Wage equations, 2008–2011[a]

Men (20–64)	OLS (1)	FE (2)	Wool95 (3)	2SLS (4)	FE-2SLS (5)	SemWool10 (6)
Health (+1)	0.081***	0.008**	-0.001	0.092***	0.033**	0.002
	(0.006)	(0.004)	(0.005)	(0.013)	(0.015)	(0.021)
Experience	0.015***	0.104***	0.029	0.015***	0.103***	-0.039
	(0.003)	(0.025)	(0.050)	(0.003)	(0.026)	(0.051)
Experience^2	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ISCED 3-4	0.234***			0.232***		
	(0.013)			(0.013)		
ISCED 5-6	0.618***			0.615***		
	(0.015)			(0.015)		
Observations	26040	26040	26040	26040	26040	26040
F-test[b]				1561.36 (2)	377.05 (2)	
Overid. test[c]				15.3 (0.00)	1.71 (0.19)	
Unobserved effects[d]			283.33 (0.00)			209.2 (0.00)
Women (20–59)	(1)	(2)	(3)	(4)	(5)	(6)
Health (+1)	0.068***	0.007	0.006	0.061***	0.018	-0.002
	(0.006)	(0.005)	(0.005)	(0.013)	(0.017)	(0.020)
Experience	0.014***	0.043***	0.206***	0.014***	0.043***	0.194***
	(0.003)	(0.012)	(0.023)	(0.003)	(0.012)	(0.024)
Experience^2	-0.000***	-0.000***	-0.000	-0.000***	-0.000***	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ISCED 3-4	0.230***			0.231***		
	(0.014)			(0.014)		
ISCED 5-6	0.670***			0.672***		
	(0.015)			(0.015)		
Observations	23859	23859	23859	23859	23859	23859
F-test[b]				1488.42 (2)	447.7 (2)	
Overid. test[c]				6.92 (0.01)	2.98 (0.08)	
Unobserved effects[d]			140.34 (0.00)			76.35 (0.00)

[a] The dependent variable is the PPP-adjusted hourly gross wage rate (in logs). All estimates include country- and gender-specific employment and unemployment rates and dummy variables for survey year. Columns (1) and (4) also include dummy variables for age year and country. Columns (3) and (6) also include interactions of the IMR with the survey year dummy variables. Standard errors robust to serial correlation and heteroskedasticity are in parentheses. The standard errors in Columns (3) and (6) are also corrected for the probit first-stage estimation. Significance levels: *** p < 0.01 ** p < 0.05 * p < 0.10.
[b] F-test statistic with degrees of freedom in parentheses. The null hypothesis is that the health limitations (Chronic and GALI) have no joint impact on SRH.
[c] Overidentification test statistic with corresponding p-value in parentheses. The null hypothesis is that the health limitations variables (Chronic and GALI) are orthogonal to the error term in the wage equation (2.6).
[d] χ^2 test statistics with corresponding p-value in parentheses for the joint significance of 8 and 9 variables (vector $\overline{z}_{i.}$) are reported

As Columns 1 and 3 in Table 7 show, for men, only the pooled OLS and to some extent the 2SLS estimators yield a significant effect of health on wages across age groups. Once individual unobserved heterogeneity is accounted for (FE or FE-2SLS estimators) the parameter estimate of the health variables becomes insignificant (at the five percent level). Notably, the positive effect of health on wages for older workers found in the period prior to the GR disappears when controlling for individual unobserved heterogeneity and sample selection (Wool95 and SemWool10 estimators)[15] and, in general, we no longer find evidence of an increasing effect of health on wages with age. The estimation results by age groups for women are shown in Table 8. The pooled OLS and 2SLS estimators in Columns 1 and 3 show a positive association between health and wages. In the youngest age group, the health effect becomes insignificant once individual unobserved heterogeneity is accounted for (FE and FE-2SLS estimators). On the contrary, in middle-aged women, Columns 1 to 5 yield a significant positive impact of health on wages, which becomes insignificant when using only the SemWool10 estimator. Nevertheless, since we do not reject the null hypothesis of no (contemporaneous) selection bias (see Table 2), our preferred estimation results for these women are those from the FE-2SLS estimator. In older female workers the significant positive effect of health on wages in Columns 1 and 3 (OLS and 2SLS estimators) reverses its sign from positive to negative when controlling for individual fixed effects and sample selection, and moreover increases in significance in the 2SLS models (FE-2SLS and SemWool10 estimators). Here we cannot reject the null hypothesis of no (contemporaneous) selection bias either and, therefore, we take the FE-2SLS estimator as our preferred specification. This negative effect of health on hourly wages, we find, is due to a dominating effect on hours of work.[16]

To summarize, during the GR the effect of health on wages in the working-age population becomes insignificant for men and remains insignificant for women. However, across age groups we find some evidence of a positive effect of health on wages (FE-2SLS estimator) for middle-aged female workers, which is only marginally significant for older male workers. We also find a significant negative effect of health on wages for older female workers that is due to a dominating effect of health on hours of work.

[15] However, health yields a marginally significant, positive impact on wages in the FE-2SLS estimator, which is the one we should consider, as we find no evidence of (contemporaneous) selection bias for older male workers (see Table 2).

[16] More specifically, the effect of health on wage earnings is positive and statistically significant in the OLS and 2SLS models only. Instead, the effect of health on hours of work is always positive and statistically significant (results not shown).

Table 7: Wage equations by age groups, 2008–2011 (Men)[a]

	OLS	FE	Wool95	2SLS	FE-2SLS	SemWool10
20–34	(1)	(2)	(3)	(4)	(5)	(6)
Health	0.066***	0.006	0.003	0.048	0.000	-0.005
	(0.012)	(0.009)	(0.010)	(0.030)	(0.037)	(0.048)
Experience	0.051***	0.093	0.190*	0.051***	0.093	0.175*
	(0.009)	(0.069)	(0.108)	(0.009)	(0.069)	(0.101)
Experience^2	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ISCED 3-4	0.159***			0.161***		
	(0.023)			(0.024)		
ISCED 5-6	0.452***			0.456***		
	(0.032)			(0.033)		
Observations	5580	5580	5580	5580	5580	5580
F-test[b]				152.39 (2)	69.94 (2)	
Overid. test[c]				1.76 (0.18)	0.54 (0.46)	
Unobs. effects[d]			75.5(0.00)			63.14 (0.00)
35–49	(1)	(2)	(3)	(4)	(5)	(6)
Health	0.092***	0.007	-0.002	0.100***	0.042	-0.019
	(0.009)	(0.006)	(0.007)	(0.021)	(0.027)	(0.035)
Experience	0.014**	0.112***	0.081	0.014**	0.112***	-0.041
	(0.007)	(0.040)	(0.084)	(0.007)	(0.040)	(0.083)
Experience^2	-0.000*	0.000	-0.000	-0.000*	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ISCED 3-4	0.275***			0.274***		
	(0.021)			(0.022)		
ISCED 5-6	0.659***			0.657***		
	(0.025)			(0.026)		
Observations	9881	9881	9881	9881	9881	9881
F-test[b]				573.76 (2)	114.44 (2)	
Overid. test[c]				6.56 (0.01)	0.59 (0.44)	
Unobs. effects[d]			118.5 (0.00)			77.32 (0.00)
50–64	(1)	(2)	(3)	(4)	(5)	(6)
Health	0.084***	0.005	-0.004	0.102***	0.051*	0.036
	(0.013)	(0.008)	(0.011)	(0.026)	(0.029)	(0.037)
Experience	0.005	0.111**	0.003	0.005	0.109**	-0.103
	(0.010)	(0.047)	(0.103)	(0.009)	(0.047)	(0.102)
Experience^2	-0.000	0.000	0.000	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ISCED 3-4	0.205***			0.201***		
	(0.029)			(0.029)		
ISCED 5-6	0.632***			0.626***		
	(0.034)			(0.035)		
Observations	5696	5696	5696	5696	5696	5696
F-test[b]				523.23 (2)	131.97 (2)	
Overid. test[c]				6.28 (0.01)	0.00 (0.97)	
Unobs. effects[d]			58.85 ((0.00)			40.86 (0.00)

[a] The dependent variable is the PPP-adjusted hourly gross wage rate (in logs). All estimates include country- and gender-specific employment and unemployment rates and dummy variables for survey year. Columns (1) and (4) also include dummy variables for age year and country. Columns (3) and (6) also include interactions of the IMR with the survey year dummy variables. Standard errors robust to serial correlation and heteroskedasticity are in parentheses. The standard errors in Columns (3) and (6) are also corrected for the probit first-stage estimation. Significance levels: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

[b] F-test statistic with degrees of freedom in parentheses. The null hypothesis is that the health limitations (Chronic and GALI) have no joint impact on SRH.

[c] Overidentification test statistic with corresponding p-value in parentheses. The null hypothesis is that the health limitations variables (Chronic and GALI) are orthogonal to the error term in the wage equation (2.6).

[d] χ^2 test statistics with corresponding p-value in parentheses for the joint significance of 8 and 9 variables (vector $\overline{z}_{i.}$) are reported.

Table 8: Wage equations by age groups, 2008–2011 (Women)[a]

	OLS	FE	Wool95	2SLS	FE-2SLS	SemWool10
20–34	(1)	(2)	(3)	(4)	(5)	(6)
Health	0.054***	0.012	-0.001	0.090**	0.055	0.030
	(0.013)	(0.013)	(0.015)	(0.042)	(0.059)	(0.062)
Experience	0.046***	0.091***	0.327***	0.046***	0.091***	0.283***
	(0.009)	(0.034)	(0.070)	(0.009)	(0.034)	(0.069)
Experience^2	-0.003***	-0.002**	-0.002*	-0.003***	-0.002**	-0.002*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
ISCED 3-4	0.195***			0.192***		
	(0.034)			(0.033)		
ISCED 5-6	0.570***			0.563***		
	(0.035)			(0.035)		
Obser	4840	4840	4840	4840	4840	4840
F-test[b]				137.05 (2)	63.66 (2)	
Overid. test[c]				0.49 (0.49)	0.44 (0.51)	
Unobs. effects[d]			28.32 (0.00)			20.37 (0.00)
35–49	(1)	(2)	(3)	(4)	(5)	(6)
Health	0.077***	0.017***	0.023***	0.064***	0.049**	0.025
	(0.009)	(0.006)	(0.007)	(0.020)	(0.022)	(0.026)
Experience	0.017***	0.027	0.183***	0.017***	0.028	0.174***
	(0.005)	(0.017)	(0.033)	(0.005)	(0.017)	(0.034)
Experience^2	-0.000	0.000	0.000	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ISCED 3-4	0.238***			0.241***		
	(0.022)			(0.023)		
ISCED 5-6	0.680***			0.685***		
	(0.024)			(0.025)		
Obser	9996	9996	9996	9996	9996	9996
F-test[b]				632.84 (2)	209.17 (2)	
Overid. test[c]				8.18 (0.00)	2.7 (0.10)	
Unobs. effects[d]			70.78 (0.00)			41.14 (0.00)
50–64	(1)	(2)	(3)	(4)	(5)	(6)
Health	0.064***	-0.021*	-0.020*	0.045*	-0.069**	-0.101**
	(0.015)	(0.011)	(0.012)	(0.025)	(0.034)	(0.042)
Experience	0.023**	0.011	0.082**	0.022**	0.011	0.052
	(0.010)	(0.028)	(0.040)	(0.010)	(0.028)	(0.044)
Experience^2	-0.000*	0.000	0.000	-0.000*	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ISCED 3-4	0.214***			0.217***		
	(0.029)			(0.029)		
ISCED 5-6	0.679***			0.685***		
	(0.033)			(0.033)		
Obser	4484	4484	4484	4484	4484	4484
F-test[b]				456.45 (2)	81.66 (2)	
Overid. test[c]				2.42 (0.12)	1.54 (0.21)	
Unobs. effects[d]			32.37 (0.00)			16.01 (0.00)

[a] The dependent variable is the PPP-adjusted hourly gross wage rate (in logs). All estimates include country- and gender-specific employment and unemployment rates and dummy variables for survey year. Columns (1) and (4) also include dummy variables for age year and country. Columns (3) and (6) also include interactions of the IMR with the survey year dummy variables. Standard errors robust to serial correlation and heteroskedasticity are in parentheses. The standard errors in Columns (3) and (6) are also corrected for the probit first-stage estimation. Significance levels: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

[b] F-test statistic with degrees of freedom in parentheses. The null hypothesis is that the health limitations (Chronic and GALI) have no joint impact on SRH.

[c] Overidentification test statistic with corresponding p-value in parentheses. The null hypothesis is that the health limitations variables (Chronic and GALI) are orthogonal to the error term in the wage equation (2.6).

[d] χ^2 test statistics with corresponding p-value in parentheses for the joint significance of 8 and 9 variables (vector $\overline{z}_{i.}$) are reported.

5. Sensitivity analyses

We performed several sensitivity analyses to test the robustness of our results. For instance, one possible concern in our findings is that the significant effect of health on the wages of (older) male workers disappears during the crisis because (unhealthy) workers (early-) retire from the labor market and, hence, are dropped from our sample (see Section 2). Therefore, we re-estimated all models keeping the (early-) retirees in our sample, and it does not change our main results for men and women, neither before (2005–2007) nor during (2008–2011) the GR (these results are available upon request).

We also investigated our (implicit) model assumption of no reverse impacts of wages on health limitations (i.e., the assumption of strict exogeneity of the health limitation variables conditional on the individual fixed effect) by re-estimating the two-stage models with health limitations restricted to only chronic conditions. These limitations are those unlikely to be directly affected by current labor market outcomes (employment). Recent work by Westerlund *et al.* (2010) provides indirect support for these methodological choices. First, it shows that in France, retirement does not change the risk of major chronic diseases, which supports the inclusion of chronic diseases in the (first-stage) health equation. Secondly, it demonstrates that retirement is associated with a reduction in mental and physical fatigue and depression symptoms, which justifies the omission of GALI from the (first-stage) health equation. The estimation results using only chronic illnesses as an instrument for SRH are given for both men and women in Table 9, respectively. These results show that using only chronic conditions to correct for measurement error in SRH does not change our main empirical findings in either period, lending support to the assumption of no reverse impact of current wages on health limitations.[17]

[17] The estimation results by age groups (and in both periods before and during the GR) also remain much unchanged when using only chronic conditions as an instrument for SRH (these results are available upon on request).

Table 9: Wage equations. Sensitivity analysis using only chronic conditions to instrument for SRH.[a]

	2005–2007			2008–2011		
	2SLS	FE-2SLS	SemWool10	2SLS	FE-2SLS	SemWool10
Men (20–64)	(4)	(5)	(6)	(4)	(5)	(6)
Health	0.086***	0.049**	0.061**	0.067***	0.020	-0.006
	(0.014)	(0.020)	(0.025)	(0.015)	(0.018)	(0.025)
Experience	0.014***	0.005***	0.011***	0.015***	0.104***	-0.053
	(0.002)	(0.002)	(0.002)	(0.003)	(0.025)	(0.052)
Experience^2	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ISCED 3-4	0.213***			0.237***		
	(0.012)			(0.013)		
ISCED 5-6	0.610***			0.623***		
	(0.014)			(0.016)		
Observations	27462	27462	27462	26040	26040	26040
F-test[b]	2653.65 (1)	530.60 (1)		2088.86 (1)	547.19 (1)	
Unobserved effects[c]			125.83 (0.00)			228.1 (0.00)
Women (20–59)	(4)	(5)	(6)	(4)	(5)	(6)
Health	0.042***	-0.018	0.003	0.045***	-0.001	-0.025
	(0.014)	(0.023)	(0.027)	(0.014)	(0.019)	(0.023)
Experience	0.012***	0.003	0.009***	0.015***	0.043***	0.200***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.012)	(0.024)
Experience^2	-0.000**	-0.000	-0.000***	-0.000***	-0.000***	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ISCED 3-4	0.233***			0.234***		
	(0.013)			(0.014)		
ISCED 5-6	0.667***			0.677***		
	(0.015)			(0.015)		
Observations	22954	22954	22954	23859	23859	23859
F-test[b]	2450.56 (1)	604.93 (1)		2130.59 (1)	636.57 (1)	
Unobserved effects[c]			121.85 (0.00)			85.16 (0.00)

[a] The dependent variable is the PPP-adjusted hourly gross wage rate (in logs). All estimates include country- and gender-specific employment and unemployment rates and dummy variables for survey year. Columns (1) and (4) also include dummy variables for age year and country. Columns (3) and (6) also include interactions of the IMR with the survey year dummy variables. Standard errors robust to serial correlation and heteroskedasticity are in parentheses. The standard errors in Columns (3) and (6) are also corrected for the probit first-stage estimation. Significance levels: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

[b] F-test statistic with degrees of freedom in parentheses. The null hypothesis is that Chronic conditions has no impact on SRH.

[c] χ^2 test statistics with corresponding p-value in parentheses for the joint significance of 8 variables (vector $\overline{z}_{i.}$) are reported.

6. Summary and discussion

Theoretical economic models predict, based on productivity arguments, that an individual’s health affects his or her wage rate. However, econometric problems such as unobserved heterogeneity, sample selection and measurement error are likely to bias the estimate of health in a wage equation. In this paper, we add to this empirical literature by expanding on the findings of Jäckle and Himmler (2010) for Germany to Europe and across age groups, based on the idea of health having an increased effect on wages with age. Moreover, by using data from before and during the Great Recession (GR) we gain insights into whether, and how, the current crisis has altered the relationship between health and wages.

Our results provide empirical evidence of measurement error in the SRH variable when estimating its impact on wages and of selectivity bias in wages for men mainly. Our findings for Europe also show that the positive impact of health on wages for (older) male workers in the period prior to the GR largely disappears during the GR. Moreover, although overall for women we do not find evidence of an impact of health on wages both before and during the GR, for middle-aged (35–49) and older female workers (50–59) we find a (significant) positive and negative impact of health on (hourly) wages during the GR, respectively.

These latter results might be related to the "added-worker effect" shown in Table 1 and appendix Tables A.2–A.4. There is some evidence that suggests that such women (who manage to get a job in a tight labor market) have rather higher levels of education and, hence, also relatively higher wages (see Landivar, 2012, and Tables A.3–A.4). The negative effect of health on hourly wages that we find for older female workers is due to a dominating effect on working hours, and might be related to the fact that for these women the increase in employment probability during the GR (i.e. the "added-worker effect") is largest. Still, further research should also consider information on their partners in order to better understand personal circumstances that motivate such women to re-enter into the labor market.

With regard to our findings in the working-age population, although during the GR overall welfare generosity (e.g., access to health care) across countries in Europe has declined, which we would expect to result in an increase in the impact of health on wages, *presenteeism* (i.e. attending work even though being sick) has become more common among workers. For instance, a recent survey by CareerBuilder (2011) in the U.S finds that more than 70 percent of workers typically go to work when they are sick. There is also some evidence during the GR of a reduction in the variable component of wages, which is likely to be more responsive to productivity-related components such as health, for instance, through cuts in bonuses and other rewards (Vandekerckhove *et al.*, 2012). Both the increase in *presenteeism* and the reduction in the variable component of wages could explain why health has become less responsive to wages in the working-age (male) population during the GR.

From a policy point of view, if the reduction in the impact of (poor) health on wages is attribuitable to *presenteeism*, while being a short-run effect it actually raises concern on the potentially negative long-run effects for these workers and on the negative impacts on public health, e.g., by affecting co-workers' health. For example, U.S. employees who attended work while infected with H1N1 are estimated to have caused the infection of as many as seven million co-workers (15 percent of the 44 million infected with H1N1) over just three months during the height of the H1N1 pandemic in 2009 (Drago and Miller, 2010).

Acknowledgments

We wish to thank Rob Alessie, Pilar García-Gómez, Sergi Jiménez, Adriaan Kalwij, Ronald Oaxaca, and the seminar participants at Erasmus University School of Economics, the 33th Jornadas de la Asociación de Economía de la Salud, and the 10th Jornadas de la Asociación Española de Economía del Trabajo for valuable comments and discussions. Manuel Flores gratefully acknowledges the financial support from Xunta de Galicia through its María Barbeito fellowship program.

References

Almond D, Currie J, 2011. Killing me softly: The Fetal Origins Hypothesis. *Journal of Economic Perspectives* **25**:153–172.

Arpaia A, Curci N. 2010. EU labour market behaviour during the Great Recession. European Economy Economic Papers No. 405.

Becker GS. 1962. Investment in human capital: A theoretical analysis. *Journal of Political Economy* **70**: 9–49.

Bound J. 1991. Self-reported versus objective measures of health in retirement models. *Journal of Human Resources* **26**: 106–138.

Bound J, Brown C, Mathiowetz N. 2001. Measurement error in survey data, In: Heckman JJ, Leamer E (Eds.), *Handbook of Econometrics*, vol. 5, Elsevier Science: New York, pp. 3705–3843.

Bound J, Jaeger DA, Baker RM. 1995. Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association* **90**: 443–450.

Bound J, Schoenbaum M, Stinebrickner T, Waidmann T. 1999. The dynamic effects of health on the labor force transitions of older workers. *Labour Economics* **6**: 179–202.

Brown S, Roberts J, Taylor K. 2010. Reservation wages, labour market participation and health. *Journal of the Royal Statistical Society: Series A* **173**: 501–529.

Cai L. 2009. Effects of health on [the] wages of Australian men. *Economic Record* **85**: 290–306.

Card D. 1999. The causal effect of education on earnings. In: Ashenfelter O, Card D (Eds.), *Handbook of Labor Economics*, vol. 3. Elsevier Science B.V.: Amsterdam, pp. 1801–1863.

CareerBuilder. 2011. *Nearly three-in-four workers go to work when they are sick* [WWW Document]. URL http://www.careerbuilder.com/share/aboutus/pressreleasesdetail.aspx?sd=1%2f19%2f2011&siteid=cbpr&sc_cmp1=cb_pr616_&id=pr616&ed=12%2f31%2f2011. [1 May 2014]

Contoyannis P, Rice N. 2001. The impact of health on wages: Evidence from the British Household Panel Survey. *Empirical Economics* **26**: 599–622.

Crossley TF, Kennedy S. 2002. The reliability of self-assessed health status. *Journal of Health Economics* **21**: 643–658.

Currie J, Madrian BC. 1999. Health, health insurance and the labor market. In: Ashenfelter O, Card D (Eds.), *Handbook of Labor Economics*, vol. 3. Elsevier Science B.V.: Amsterdam, pp. 3309–3415.

Drago R, Miller K. 2010. Sick at work: Infected employees in the workplace during the H1N1 pandemic. Institute for Women's Policy Research Working Paper No. B264.

Dustman C, Rochina-Barrachina MR. 2007. Selection correction in panel data models: An application to labour supply and wages. *Econometrics Journal* **10**: 263–293.

Flores M, García-Gómez P, Kalwij A. 2014. Early life circumstances and life-cycle labor market outcomes. Mimeo.

Flores M, Kalwij A. 2013. What do wages add to the health-employment nexus? Evidence from older European workers. Netspar Discussion Paper No. 03/2013-005.

Grossman M. 2001. The Human Capital Model. In: Culyer AJ, Newhouse JP (Eds.), *Handbook of Health Economics*, vol. 1A. Elsevier Science B.V.: Amsterdam, pp. 347–408.

Haveman R, Wolfe B, Kreider B, Stone M. 1994. Market work, wages, and men's health. *Journal of Health Economics* **13**: 163–82.

Heckman JJ. 1979. Sample selection bias as a specification error. *Econometrica* **47**: 153–161.

Jäckle R, Himmler O. 2010. Health and wages: Panel data estimates considering selection and endogeneity. *Journal of Human Resources* **45**: 364–406.

Kapteyn A, Smith JP, Van Soest A. 2007. Vignettes and self-reports of work disability in the United States and the Netherlands. *American Economic Review* **97**: 461–473.

Karanikolos M, Mladovsky P, Cylus J, Thomson S, Basu S, Stuckler D, Mackenbach JP, McKee M. 2013. Financial crisis, austerity, and health in Europe. *Lancet* **381**: 1323–31.

Landivar LC. 2012. The impact of the Great Recession on mothers' employment, In: Blair SL (Ed.), *Economic Stress and the Family*, Emerald Group Publishing Limited, pp.163–185.

Leduc S, Zheng L. 2012. Uncertainty, unemployment, and inflation. FRBSF Economic Letter No. 2012-28.

Lee L-F. 1982. Health and wage: A simultaneous equation model with multiple discrete indicators. *International Economic Review* **23**: 199–221.

Meijer E, Kapteyn A, Andreyeva T. 2011. Internationally comparable health indices. *Health Economics* **20**: 600–619.

Mincer J. 1974. *Schooling, Experience, and Earnings*. Columbia University Press for the National Bureau of Economic Research: New York.

Montuenga V, García I, Fernández M. 2003. Wage flexibility: Evidence from five EU countries based on the wage curve. *Economics letters* **78**: 169–74.

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

Mushkin SJ. 1962. Health as an investment. *Journal of Political Economy* **70**: 129–157.

OECD. 2013. *OECD Employment Outlook 2013* [WWW Document]. OECD Publishing. URL http://dx.doi.org/10.1787/empl_outlook-2013-en. [1 May 2014]

Semykina A. 2012. Specification tests and tests for overidentifying restrictions in panel data models with selection. *Economics Letters* **115:** 53–55.

Semykina A, Wooldridge JW. 2008. Estimating panel data models in the presence of endogeneity and selection: Theory and application. Michigan State University. Unpublished.

Semykina A, Wooldridge JW. 2010. Estimating panel data models in the presence of endogeneity and selection. *Journal of Econometrics* **157**: 375–380.

Stern S. 1989. Measuring the effect of disability on labor force participation. *Journal of Human Resources* **24**: 361–95.

Vandekerckhove S, Van Peteghem J, Van Gyes G. 2012. Wages and working conditions in the crisis. European Foundation for the Improvement of Living and Working Conditions. Working Paper No. EF/12/44/EN.

Westerlund H, Vahtera J, Ferrie JE, Singh-Manoux A, Pentti J, Melchior M, Leineweber C, Jokela M, Siegrist J, Goldberg M, Zins M, Kivimäki M. 2010. Effect of retirement on major chronic conditions and fatigue: French GAZEL occupational cohort study. *British Medical Journal* **341**: c6149.

Wooldridge JW. 1995. Selection corrections for panel data models under conditional mean independence assumptions. *Journal of Econometrics* **68:** 115–132.

Wooldridge JW. 2002. *Econometric Analysis of Cross Section and Panel Data*. The MIT Press: Cambridge, MA.

Appendix

Table A.1: Variable definitions

Variable	Definition
Dependent variable	
Log hourly (gross) wage	Hourly gross wages are measured in PPP-adjusted 2005 €. They are defined for paid job workers and obtained from dividing the amount of gross wage-earnings by the number of hours (usually) worked. Both variables are available for the main job. We treat extreme values in hourly gross wages such as those below 1 and above 300 PPP-€ as missing. It does *not* include reimbursements made by an employer for work-related expenses (e.g. business travel), severance and termination pay to compensate employees for employment ending before the employee has reached the normal retirement age for that job and redundancy payments, allowances for purely work-related expenses such as those for travel and subsistence or for protective clothes, lump sum payments at the normal retirement date, and union strike pay.
Participation	Participation (i.e. selection) is equal to 1 if a respondent reports working a positive number of hours per week in his/her main job, 0 otherwise.
Respondent's Health	
Self-reported health (SRH)	Includes five SRH categories, from 1 to 5: very bad, bad, fair, good, and very good.
Chronic	Chronic refers to chronic diseases; it is equal to 1 if a respondent has one or more conditions, 0 if none. Conditions are defined as hypertension, high blood cholesterol, diabetes, asthma, arthritis, osteoporosis, stomach condition, cataracts, and other conditions.
GALI	GALI refers to the global activity limitation indicator. The question for this index is the following: "For the past six months at least, to what extent have you been limited because of a health problem in activities people usually do. It is equal to 1 if the respondent is "not limited", 2 if "limited but not severely," and 3 if "severely limited."
Respondent's Socioeconomic Characteristics	
Experience	Includes the actual number of years worked by the respondent.
Education	Includes three levels of education defined from the ISCED Code 1997: no education, primary education, or lower secondary education (ISCED 0–2), upper secondary and postsecondary nontertiary education (ISCED 3–4), and tertiary education (ISCED 5–6).
Annual nonlabor income	Annual nonlabor income is measured in PPP-adjusted 2005 €. It is defined as total income received by all household members in the previous year minus individual income from employment in the previous year.
Log household size	Includes the logarithm of the number of household members.
Single	Single is equal to 1 if single, 0 otherwise (married or cohabiting).
Age	Includes dummy variables for each age year.
Time	Survey year dummies.

Table A.2: Descriptive statistics. Men and women aged 20–34.

	Employed				Nonemployed			
	2005–2007		2008–2011		2005–2007		2008–2011	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
Men (20–34)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Hourly gross wage rate	9.68	6.84	10.13	6.7				
Self-reported health (1-5)	4.3	0.68	4.33	0.64	4.04	0.89	4.08	0.82
1+ chronic diseases	0.11	0.31	0.1	0.3	0.14	0.35	0.19	0.39
Limited with GALI (1-3)	1.09	0.33	1.08	0.3	1.17	0.45	1.17	0.47
Experience	8.09	4.52	7.59	4.48	4.08	4.5	4.81	4.2
ISCED 0–2	0.17	0.37	0.21	0.41	0.35	0.48	0.38	0.49
ISCED 3-4	0.59	0.49	0.5	0.5	0.56	0.5	0.52	0.5
ISCED 5-6	0.24	0.43	0.28	0.45	0.09	0.28	0.09	0.29
Annual income from nonemployment	20,099	21,674	20,925	21,407	15,841	15,503	19,667	16,941
Household size	3.23	1.44	3.21	1.5	3.69	1.57	3.82	1.69
Single	0.47	0.5	0.46	0.5	0.7	0.46	0.73	0.44
Age	28.08	3.59	28.22	3.41	26.71	3.9	26.54	3.74
Unemployment rate (%)	6.05	3.1	7.72	4.77	7.96	4.08	11.06	5.74
Employment rate (%)	72.16	6.07	70.51	5.71	68.29	6.54	67.17	5.29
Prob. work (entire sample)	0.92	0.27	0.91	0.28				
Observations (N)	7,545		5,676		630		528	
Women (20–34)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Hourly gross wage rate	9.19	7.19	8.65	6.56				
Self-reported health (1-5)	4.27	0.68	4.26	0.66	4.09	0.76	4.15	0.71
1+ chronic diseases	0.11	0.32	0.11	0.31	0.13	0.34	0.11	0.31
Limited with GALI (1-3)	1.09	0.33	1.09	0.31	1.13	0.39	1.11	0.37
Experience	7.26	4.17	6.89	4.04	4.47	4.18	4.27	4.22
ISCED 0–2	0.11	0.31	0.14	0.35	0.28	0.45	0.28	0.45
ISCED 3-4	0.51	0.5	0.42	0.49	0.58	0.49	0.52	0.5
ISCED 5-6	0.38	0.49	0.43	0.5	0.14	0.35	0.2	0.4
Annual income from nonemployment	24,875	22,622	24,814	21,111	23,561	22,000	22,466	18,036
Household size	3.07	1.33	3.09	1.35	4.07	1.44	4.05	1.43
Single	0.39	0.49	0.37	0.48	0.2	0.4	0.24	0.43
Age	28.16	3.51	28.66	3.15	28.36	3.48	28.06	3.38
Unemployment rate (%)	7.97	3.67	8.57	4.77	9.05	3.98	9.56	4.76
Employment rate (%)	57.86	7.17	58.54	6.89	55.41	5.81	56.8	6.06
Prob. work (entire sample)	0.72	0.45	0.75	0.43				
Observations (N)	5,809		5,145		2,279		1,703	

Table A.3: Descriptive statistics. Men and women aged 35–49.

	Employed				Nonemployed			
	2005–2007		2008–2011		2005–2007		2008–2011	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
Men (35–49)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Hourly gross wage rate	13.54	10.73	14	10.91				
Self-reported health (1-5)	4.04	0.73	4.08	0.71	3.38	0.99	3.6	0.86
1+ chronic diseases	0.17	0.37	0.17	0.38	0.31	0.46	0.28	0.45
Limited with GALI (1-3)	1.13	0.39	1.13	0.39	1.4	0.67	1.31	0.59
Experience	21.68	6.34	20.01	7.39	16.14	7.62	16.15	8.5
ISCED 0–2	0.17	0.37	0.18	0.39	0.35	0.48	0.34	0.47
ISCED 3-4	0.56	0.5	0.51	0.5	0.59	0.49	0.54	0.5
ISCED 5-6	0.27	0.44	0.31	0.46	0.06	0.23	0.12	0.32
Annual income from nonemployment	17,167	19,267	18,873	17,857	13,430	15,892	13,579	15,679
Household size	3.58	1.35	3.55	1.39	3.29	1.75	3.47	1.81
Single	0.15	0.36	0.17	0.38	0.43	0.5	0.4	0.49
Age	41.99	3.81	41.92	3.58	42.45	3.86	42.26	3.68
Unemployment rate (%)	5.96	3.06	7.71	5.05	7.77	4.06	11.36	5.8
Employment rate (%)	72.75	6.16	71.27	6.06	69.11	6.69	67.27	5.34
Prob. work (entire sample)	0.94	0.24	0.94	0.24				
Observations (N)	10,821		10,002		738		662	
Women (35–49)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Hourly gross wage rate	10.67	9.01	10.67	9.14				
Self-reported health (1-5)	4	0.75	4	0.73	3.8	0.9	3.8	0.8
1+ chronic diseases	0.18	0.39	0.19	0.39	0.24	0.43	0.25	0.43
Limited with GALI (1-3)	1.14	0.4	1.15	0.41	1.24	0.54	1.23	0.5
Experience	19.66	6.82	17.81	7.62	10.18	7.71	10.22	7.95
ISCED 0–2	0.15	0.36	0.14	0.35	0.42	0.49	0.37	0.48
ISCED 3-4	0.53	0.5	0.49	0.5	0.46	0.5	0.48	0.5
ISCED 5-6	0.32	0.47	0.37	0.48	0.12	0.32	0.15	0.35
Annual income from nonemployment	26,389	29,261	28,794	32,964	34,385	35,156	27,723	33,653
Household size	3.48	1.24	3.56	1.28	4.14	1.44	4.15	1.55
Single	0.23	0.42	0.23	0.42	0.12	0.32	0.15	0.35
Age	42.07	3.8	42.16	3.62	41.8	3.78	41.97	3.61
Unemployment rate (%)	7.88	3.78	8.34	4.65	8.89	3.92	11.28	5.84
Employment rate (%)	58.74	7.61	59.88	7.25	55.87	5.95	56.31	5.91
Prob. work (entire sample)	0.77	0.42	0.82	0.38				
Observations (N)	10,166		10,434		2,980		2,246	

Table A.4: Descriptive statistics. Men aged 50–64 and women aged 50–59.

	Employed				Nonemployed			
	2005–2007		2008–2011		2005–2007		2008–2011	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
Men (50–64)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Hourly gross wage rate	14.42	12.94	15.55	16.01				
Self-reported health (1-5)	3.77	0.81	3.82	0.74	3.32	0.91	3.39	0.91
1+ chronic diseases	0.26	0.44	0.28	0.45	0.34	0.47	0.36	0.48
Limited with GALI (1-3)	1.19	0.46	1.19	0.45	1.38	0.63	1.36	0.59
Experience	34.6	7.17	32.58	9.83	30.14	9.09	30.17	9.88
ISCED 0–2	0.21	0.41	0.21	0.41	0.38	0.49	0.42	0.49
ISCED 3-4	0.52	0.5	0.46	0.5	0.52	0.5	0.5	0.5
ISCED 5-6	0.27	0.44	0.33	0.47	0.1	0.3	0.08	0.27
Annual income from nonemployment	19,326	21,265	20,722	25,417	17,398	18,151	15,064	14,878
Household size	2.82	1.23	2.77	1.25	2.74	1.27	2.73	1.48
Single	0.12	0.33	0.16	0.37	0.2	0.4	0.38	0.49
Age	55.14	3.21	55.41	3.15	55.88	3.22	56.13	3.39
Unemployment rate (%)	5.76	2.84	8.06	5.22	8.14	4.1	10.35	5.47
Employment rate (%)	73.3	5.91	71.38	6.4	68.63	6.63	68.17	5.86
Prob. work (entire sample)	0.89	0.31	0.90	0.30				
Observations (N)	5,896		5,773		698		635	
Women (50–59)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Hourly gross wage rate	10.73	10.08	11.33	9.66				
Self-reported health (1-5)	3.7	0.8	3.8	0.77	3.51	0.91	3.49	0.84
1+ chronic diseases	0.28	0.45	0.3	0.46	0.36	0.48	0.41	0.49
Limited with GALI (1-3)	1.22	0.48	1.22	0.48	1.38	0.64	1.34	0.58
Experience	29.68	8.76	26.47	11.26	13.46	11.81	14.42	11.67
ISCED 0–2	0.23	0.42	0.21	0.41	0.58	0.49	0.55	0.5
ISCED 3-4	0.51	0.5	0.48	0.5	0.33	0.47	0.38	0.49
ISCED 5-6	0.26	0.44	0.31	0.46	0.09	0.28	0.07	0.25
Annual income from nonemployment	24,746	30,137	27,835	33,567	39,303	38,924	31,988	61,555
Household size	2.51	1.13	2.5	1.13	2.85	1.26	2.9	1.26
Single	0.3	0.46	0.29	0.46	0.14	0.35	0.18	0.38
Age	53.78	2.28	54.08	2.22	54.43	2.42	54.56	2.27
Unemployment rate (%)	7.5	3.31	8.02	4.46	8.32	3.64	10.99	5.9
Employment rate (%)	59.44	7.44	60.78	7.76	56.13	5.62	56.91	5.94
Prob. work (entire sample)	0.66	0.47	0.74	0.44				
Observations (N)	4,078		4,799		2,105		1,665	

Table A.5: Participation equations, Men aged 20-64 (Wool95)

	2005	2006	2007	2008	2009	2010	2011
Experience	-0.109***	0.120***	0.141***	-2.498***	-4.310***	3.653***	2.055***
	(0.018)	(0.024)	(0.022)	(0.142)	(0.269)	(0.240)	(0.118)
Experience^2	0.001	-0.002***	0.000	-0.006*	-0.015**	-0.007	0.003
	(0.001)	(0.001)	(0.001)	(0.003)	(0.006)	(0.005)	(0.003)
SRH	0.059	-0.070	0.143***	0.195**	-0.076	0.116*	0.097
	(0.044)	(0.052)	(0.047)	(0.076)	(0.074)	(0.062)	(0.068)
Single	0.096	-0.135	-0.470**	-0.626**	0.162	-0.098	-0.266
	(0.178)	(0.244)	(0.190)	(0.260)	(0.329)	(0.295)	(0.250)
Household size	-0.219***	0.244***	0.111*	-0.035	-0.015	0.041	0.120*
	(0.056)	(0.092)	(0.065)	(0.080)	(0.104)	(0.083)	(0.072)
ln nonlaborincome	0.006	-0.060**	-0.087***	-0.019	-0.067**	-0.104***	-0.036
	(0.019)	(0.026)	(0.022)	(0.030)	(0.029)	(0.027)	(0.023)
Employment rate	0.161***	-0.056	-0.046	0.054	0.942***	0.399***	-0.141*
	(0.039)	(0.121)	(0.047)	(0.069)	(0.118)	(0.131)	(0.073)
Unemployment rate	0.259***	-0.052	-0.184***	0.029	0.785***	0.322**	-0.352***
	(0.039)	(0.141)	(0.042)	(0.068)	(0.115)	(0.125)	(0.086)
mexperience	0.168***	-0.080***	-0.103***	2.462***	4.309***	-3.645***	-2.048***
	(0.020)	(0.025)	(0.023)	(0.149)	(0.273)	(0.237)	(0.115)
mexperience^2	-0.001***	0.002**	-0.001	0.006*	0.015**	0.007	-0.004
	(0.001)	(0.001)	(0.001)	(0.003)	(0.006)	(0.005)	(0.003)
msrh	0.500***	0.665***	0.434***	-0.124	0.276***	0.188**	0.052
	(0.054)	(0.062)	(0.055)	(0.096)	(0.090)	(0.077)	(0.084)
msingle	-0.451**	-0.308	0.040	0.116	-0.550	-0.403	0.028
	(0.183)	(0.253)	(0.199)	(0.272)	(0.342)	(0.306)	(0.259)
mhousehold_size	0.266***	-0.146	-0.045	-0.016	-0.024	-0.070	-0.120
	(0.059)	(0.093)	(0.067)	(0.085)	(0.108)	(0.086)	(0.074)
mln_nonlaborincome	-0.077***	-0.053*	0.015	-0.011	0.053	0.065**	0.044
	(0.023)	(0.029)	(0.022)	(0.036)	(0.035)	(0.029)	(0.028)
memployment_rate	-0.110***	0.071	0.071	-0.015	-0.942***	-0.340***	0.161**
	(0.037)	(0.123)	(0.050)	(0.081)	(0.121)	(0.121)	(0.067)
munemployment_rate	-0.329***	-0.014	0.089**	-0.020	-0.808***	-0.337***	0.358***
	(0.050)	(0.142)	(0.039)	(0.068)	(0.117)	(0.122)	(0.094)
Constant	-4.349***	-1.071*	-2.103***	-3.739***	-0.470	-4.656***	-2.779***
	(0.608)	(0.591)	(0.581)	(1.201)	(0.641)	(0.947)	(0.872)
Observations	9914	9914	9914	7128	7128	7128	7128

[a] Probit coefficient estimates. Robust standard errors are in parentheses. Significance levels: *** p<0.01 ** p<0.05 * p<0.10.

Table A.6: Participation equations, Men aged 20-64 (SemWool10)

	2005	2006	2007	2008	2009	2010	2011
Experience	-0.099***	0.121***	0.126***	-2.505***	-4.394***	3.770***	2.090***
	(0.018)	(0.024)	(0.023)	(0.141)	(0.268)	(0.237)	(0.118)
Experience^2	0.000	-0.003***	0.000	-0.006*	-0.015**	-0.008	0.003
	(0.001)	(0.001)	(0.001)	(0.003)	(0.006)	(0.005)	(0.003)
Chronic	0.117	0.009	-0.119	0.019	0.031	-0.003	-0.026
	(0.089)	(0.101)	(0.094)	(0.153)	(0.138)	(0.121)	(0.130)
GALI	0.039	0.057	0.129*	0.349***	0.154	0.028	0.072
	(0.069)	(0.072)	(0.066)	(0.120)	(0.116)	(0.096)	(0.107)
Single	0.114	-0.106	-0.482***	-0.612**	0.159	-0.088	-0.280
	(0.174)	(0.237)	(0.185)	(0.260)	(0.330)	(0.291)	(0.250)
Household size	-0.215***	0.209**	0.109*	-0.027	-0.029	0.047	0.115
	(0.056)	(0.090)	(0.064)	(0.080)	(0.104)	(0.083)	(0.072)
ln nonlaborincome	0.008	-0.055**	-0.079***	-0.018	-0.066**	-0.100***	-0.035
	(0.019)	(0.025)	(0.021)	(0.030)	(0.029)	(0.027)	(0.023)
Employment rate	0.185***	-0.058	-0.086*	0.053	0.933***	0.400***	-0.150**
	(0.038)	(0.121)	(0.046)	(0.070)	(0.118)	(0.131)	(0.073)
Unemployment rate	0.226***	0.094	-0.183***	0.030	0.767***	0.304**	-0.354***
	(0.038)	(0.143)	(0.042)	(0.068)	(0.114)	(0.125)	(0.085)
mexperience	0.146***	-0.093***	-0.098***	2.470***	4.392***	-3.763***	-2.084***
	(0.019)	(0.026)	(0.023)	(0.147)	(0.272)	(0.235)	(0.114)
mexperience^2	-0.001**	0.002***	-0.001	0.006*	0.015**	0.008	-0.003
	(0.001)	(0.001)	(0.001)	(0.003)	(0.006)	(0.005)	(0.003)
mchronic	-0.042	0.027	0.178	0.086	0.014	-0.056	0.137
	(0.115)	(0.131)	(0.121)	(0.208)	(0.194)	(0.166)	(0.169)
mgali	0.645***	0.747***	0.580***	-0.279	0.093	0.389***	0.004
	(0.094)	(0.101)	(0.093)	(0.188)	(0.171)	(0.141)	(0.156)
msingle	-0.476***	-0.334	0.051	0.103	-0.532	-0.391	0.054
	(0.179)	(0.245)	(0.194)	(0.272)	(0.343)	(0.303)	(0.259)
mhousehold_size	0.261***	-0.112	-0.048	-0.020	-0.012	-0.075	-0.116
	(0.059)	(0.092)	(0.065)	(0.085)	(0.109)	(0.086)	(0.075)
mln_nonlaborincome	-0.072***	-0.047*	0.016	-0.012	0.055	0.065**	0.044
	(0.023)	(0.028)	(0.022)	(0.036)	(0.035)	(0.028)	(0.028)
memployment_rate	-0.117***	0.092	0.126***	-0.011	-0.926***	-0.332***	0.174***
	(0.036)	(0.123)	(0.048)	(0.082)	(0.121)	(0.120)	(0.067)
munemployment_rate	-0.279***	-0.159	0.097**	-0.021	-0.785***	-0.311**	0.362***
	(0.049)	(0.144)	(0.038)	(0.068)	(0.117)	(0.121)	(0.093)
Constant	-5.272***	-2.340***	-2.914***	-4.047***	-1.067	-5.268***	-2.930***
	(0.619)	(0.602)	(0.593)	(1.220)	(0.673)	(0.953)	(0.896)
Observations	9914	9914	9914	7128	7128	7128	7128

[a] Probit coefficient estimates. Robust standard errors are in parentheses. Significance levels: *** p<0.01 ** p<0.05 * p<0.10.

Table A.7: Participation equations, Women aged 20-59 (Wool95)

	2005	2006	2007	2008	2009	2010	2011
Experience	-0.151***	0.151***	0.210***	-3.175***	-5.134***	4.287***	2.003***
	(0.013)	(0.016)	(0.016)	(0.103)	(0.188)	(0.152)	(0.073)
$Experience^2$	0.002***	-0.003***	-0.003***	0.004	-0.013**	-0.009**	0.005**
	(0.000)	(0.000)	(0.000)	(0.003)	(0.006)	(0.005)	(0.002)
SRH	0.098***	0.004	-0.057	-0.005	-0.047	0.012	0.013
	(0.035)	(0.038)	(0.036)	(0.062)	(0.056)	(0.050)	(0.050)
Single	0.652***	0.286	-0.421***	0.216	-0.158	0.243	0.253
	(0.144)	(0.197)	(0.142)	(0.230)	(0.273)	(0.249)	(0.189)
Household size	-0.039	-0.129*	-0.117**	-0.165**	-0.085	-0.372***	-0.207***
	(0.049)	(0.068)	(0.049)	(0.071)	(0.079)	(0.071)	(0.057)
ln nonlaborincome	0.000	0.054**	-0.120***	-0.063*	-0.038	0.023	-0.064**
	(0.021)	(0.026)	(0.024)	(0.034)	(0.029)	(0.026)	(0.026)
Employment rate	0.110***	-0.065	-0.155***	-0.368***	1.370***	-0.087	0.362***
	(0.030)	(0.068)	(0.031)	(0.067)	(0.083)	(0.091)	(0.079)
Unemployment rate	0.079***	0.211**	-0.105***	0.145***	1.213***	-0.220***	-0.006
	(0.018)	(0.088)	(0.018)	(0.054)	(0.093)	(0.062)	(0.062)
mexperience	0.283***	-0.021	-0.074***	3.198***	5.173***	-4.219***	-1.959***
	(0.014)	(0.017)	(0.017)	(0.106)	(0.190)	(0.151)	(0.073)
$mexperience^2$	-0.004***	0.001	0.000	-0.004	0.012**	0.008*	-0.006**
	(0.000)	(0.001)	(0.000)	(0.003)	(0.006)	(0.005)	(0.002)
msrh	0.379***	0.478***	0.540***	0.039	0.243***	0.278***	0.158**
	(0.042)	(0.045)	(0.043)	(0.076)	(0.068)	(0.061)	(0.062)
msingle	-0.239	0.188	0.913***	-0.420*	0.248	-0.188	-0.240
	(0.150)	(0.204)	(0.148)	(0.241)	(0.286)	(0.258)	(0.197)
mhousehold_size	-0.034	0.093	0.067	0.112	-0.010	0.311***	0.106*
	(0.051)	(0.070)	(0.051)	(0.075)	(0.082)	(0.074)	(0.061)
mln_nonlaborincome	-0.097***	-0.209***	0.011	0.023	0.080**	-0.004	0.178***
	(0.026)	(0.031)	(0.024)	(0.042)	(0.033)	(0.031)	(0.029)
memployment_rate	-0.065**	0.110	0.210***	0.474***	-1.430***	0.068	-0.370***
	(0.030)	(0.068)	(0.032)	(0.073)	(0.085)	(0.085)	(0.075)
munemployment_rate	-0.051**	-0.188**	0.135***	0.020	-1.265***	0.164**	-0.054
	(0.025)	(0.088)	(0.015)	(0.043)	(0.093)	(0.065)	(0.074)
Constant	-4.384***	-3.792***	-4.882***	-8.891***	2.338***	-0.146	-1.370***
	(0.327)	(0.324)	(0.333)	(0.685)	(0.381)	(0.481)	(0.471)
Observations	10443	10443	10443	8004	8004	8004	8004

[a] Probit coefficient estimates. Robust standard errors are in parentheses. Significance levels: *** p<0.01 ** p<0.05 * p<0.10.

Table A.8: Participation equations, Women aged 20-59 (SemWool10)

	2005	2006	2007	2008	2009	2010	2011
Experience	-0.147***	0.149***	0.202***	-3.175***	-5.154***	4.334***	2.024***
	(0.013)	(0.016)	(0.016)	(0.103)	(0.188)	(0.151)	(0.073)
Experience^2	0.002***	-0.003***	-0.003***	0.004	-0.014**	-0.009**	0.005**
	(0.000)	(0.000)	(0.000)	(0.003)	(0.006)	(0.005)	(0.002)
Chronic	0.139**	-0.011	0.030	0.058	0.124	0.021	-0.018
	(0.068)	(0.075)	(0.072)	(0.116)	(0.103)	(0.090)	(0.096)
GALI	-0.012	0.067	-0.077	-0.051	0.016	0.096	0.135*
	(0.055)	(0.058)	(0.055)	(0.095)	(0.089)	(0.073)	(0.081)
Single	0.729***	0.222	-0.482***	0.223	-0.151	0.197	0.235
	(0.142)	(0.193)	(0.140)	(0.229)	(0.273)	(0.248)	(0.188)
Household size	-0.054	-0.128*	-0.095*	-0.164**	-0.090	-0.365***	-0.204***
	(0.048)	(0.067)	(0.049)	(0.071)	(0.079)	(0.071)	(0.058)
ln nonlaborincome	0.004	0.049**	-0.112***	-0.064*	-0.034	0.020	-0.063**
	(0.020)	(0.025)	(0.023)	(0.034)	(0.028)	(0.026)	(0.025)
Employment rate	0.088***	-0.097	-0.111***	-0.368***	1.345***	-0.086	0.377***
	(0.030)	(0.068)	(0.031)	(0.067)	(0.083)	(0.090)	(0.080)
Unemployment rate	0.052***	0.285***	-0.074***	0.150***	1.233***	-0.262***	-0.004
	(0.017)	(0.090)	(0.018)	(0.054)	(0.093)	(0.062)	(0.061)
mexperience	0.277***	-0.021	-0.068***	3.197***	5.193***	-4.267***	-1.981***
	(0.014)	(0.017)	(0.017)	(0.106)	(0.190)	(0.151)	(0.073)
mexperience^2	-0.004***	0.001	0.000	-0.004	0.013**	0.008*	-0.006***
	(0.000)	(0.001)	(0.000)	(0.003)	(0.006)	(0.005)	(0.002)
mchronic	0.207**	0.303***	0.257***	-0.130	0.107	0.301**	0.303**
	(0.088)	(0.096)	(0.095)	(0.161)	(0.142)	(0.125)	(0.132)
mgali	0.394***	0.449***	0.559***	0.227	0.042	0.040	-0.269**
	(0.075)	(0.079)	(0.079)	(0.141)	(0.129)	(0.109)	(0.120)
msingle	-0.283*	0.291	1.009***	-0.426*	0.252	-0.139	-0.227
	(0.147)	(0.200)	(0.145)	(0.241)	(0.285)	(0.258)	(0.197)
mhousehold_size	-0.028	0.078	0.036	0.109	-0.007	0.303***	0.101*
	(0.050)	(0.069)	(0.050)	(0.075)	(0.082)	(0.074)	(0.061)
mln_nonlaborincome	-0.076***	-0.167***	0.029	0.025	0.083**	0.002	0.180***
	(0.025)	(0.029)	(0.024)	(0.041)	(0.033)	(0.030)	(0.029)
memployment_rate	-0.037	0.148**	0.171***	0.475***	-1.402***	0.076	-0.381***
	(0.030)	(0.068)	(0.032)	(0.073)	(0.085)	(0.085)	(0.076)
munemployment_rate	-0.027	-0.271***	0.096***	0.016	-1.284***	0.216***	-0.053
	(0.025)	(0.089)	(0.015)	(0.043)	(0.093)	(0.064)	(0.073)
Constant	-4.613***	-4.440***	-5.204***	-9.185***	2.211***	-0.506	-1.132**
	(0.331)	(0.331)	(0.337)	(0.715)	(0.411)	(0.506)	(0.491)
Observations	10443	10443	10443	8004	8004	8004	8004

[a] Probit coefficient estimates. Robust standard errors are in parentheses. Significance levels: *** p<0.01 ** p<0.05 * p<0.10.