HEALTH ECONOMICS
Health Econ. **24**: 1302–1313 (2015)
Published online 22 July 2014 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/hec.3087

SELF-EMPLOYMENT AND HEALTH: BARRIERS OR BENEFITS?

CORNELIUS A. RIETVELD[a,*], HANS VAN KIPPERSLUIS[a,b] and A. ROY THURIK[a,b,c,d]

[a]*Department of Applied Economics, Erasmus School of Economics, Erasmus University Rotterdam, DR, Rotterdam, the Netherlands*
[b]*Tinbergen Institute Rotterdam, PA, Rotterdam, the Netherlands*
[c]*Panteia, Zoetermeer, AA, the Netherlands*
[d]*GSCM-Montpellier Business School, Montpellier, France*

ABSTRACT

The self-employed are often reported to be healthier than wageworkers; however, the cause of this health difference is largely unknown. The longitudinal nature of the US Health and Retirement Study allows us to gauge the plausibility of two competing explanations for this difference: a contextual effect of self-employment on health (benefit effect), or a health-related selection of individuals into self-employment (barrier effect). Our main finding is that the selection of comparatively healthier individuals into self-employment accounts for the positive cross-sectional difference. The results rule out a positive contextual effect of self-employment on health, and we present tentative evidence that, if anything, engaging in self-employment is bad for one's health. Given the importance of the self-employed in the economy, these findings contribute to our understanding of the vitality of the labor force. Copyright © 2014 John Wiley & Sons, Ltd.

Received 08 October 2013; Revised 27 May 2014; Accepted 20 June 2014

KEY WORDS: health; HRS; self-employment

1. INTRODUCTION

Many governments stimulate self-employment (European Commission, 2004; Gilbert *et al.*, 2004) because of the assumed positive link with economic vitality (Audretsch and Keilbach, 2004; Carree and Thurik, 2010; Koellinger and Thurik, 2012). Recently, self-employment was also suggested to have a positive link with individuals' physical vitality (Tetrick *et al.*, 2000; Bradley and Roberts, 2004; Stephan and Roesler, 2010). If such a link exists, governments may also want to encourage self-employment as an alternative to early retirement to relieve the economic pressures generated by aging populations. The effectiveness of these measures depends on the extent to which self-employment indeed positively affects the health of the 50+ population. Existing evidence on this topic is, however, scarce, conflicting, and, partly because of the cross-sectional nature of existing analyses, poorly understood (Torres, 2012). This is surprising given the importance of the self-employed in the current economic system (Audretsch and Thurik, 2000; 2001).

While some of the earlier-mentioned papers show that self-employment has health benefits, others show that the self-employed are at higher risk for certain diseases than wageworkers (Lewin-Epstein and Yuchtman-Yaar, 1991; Buttner, 1992; Jamal, 1997; Parslow *et al.*, 2004; Dahl *et al.*, 2010, Yoon and Bernell, 2013). The cited studies emphasize structural differences between self-employment and wage work to explain the difference in health between the self-employed and wageworkers. The self-employed operate their business independently, without the control of a supervisor, while wageworkers are not fully responsible for the survival of the business (Bjuggren *et al.*, 2012). The associated differences in the amount and intensity of work and freedom versus controllability may result in different health outcomes (Stephan and Roesler, 2010).

*Correspondence to: Department of Applied Economics, Erasmus School of Economics, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR, Rotterdam, the Netherlands. E-mail: nrietveld@ese.eur.nl

Another explanation for health differences is almost entirely overlooked: namely, the selection of comparatively healthier individuals into self-employment. Only Jamal (1997) and Stephan and Roesler (2010) mention this possibility in their study limitations as an alternative explanation for their findings.[1] Such a selection mechanism is difficult to reveal because instrumental variables or longitudinal data are required.

In this paper, we use the Health and Retirement Study (HRS, Juster and Suzman, 1995), a population-wide panel dataset with information about employment status and several health outcomes, to study the association between self-employment and health. The longitudinal nature of the HRS allows us to gauge the plausibility of a contextual (benefit) effect versus a selection (barrier) effect, which is essential to fully understand the association between self-employment and health. Because it remains notoriously difficult—even with longitudinal data—to discriminate between the two effects, we use several methods to investigate which effect prevails.

We show that the self-employed are generally healthier than wageworkers for all three available measures of health: the number of health conditions ever had, self-reported health, and mental health. This correlation does not disappear when controlling for health history, suggesting that contemporaneous reverse causality from health to self-employment cannot entirely explain the correlation. However, the longitudinal fixed-effects analyses rule out a positive contextual effect of self-employment on health. These results suggest that the selection of comparatively healthier individuals into self-employment accounts for the cross-sectional association. We present tentative evidence that the contextual effect of self-employment on health could even be negative if the selection into self-employment based on unobservables is as large as the selection based on observables.

This paper is organized as follows. Section 2 discusses related literature, after which section 3 describes the data. In section 4, the empirical methods and results are presented. Section 5 presents robustness checks, and section 6 concludes.

2. RELATED LITERATURE

Health can be influenced by the characteristics of a given occupation (Ravesteijn *et al.*, 2013), which may result in a 'contextual effect' of self-employment on health. In contrast, self-employment can attract individuals with a different health profile and prospect than wageworkers. This could occur when individuals decide to quit or enter self-employment for health reasons and when predetermined individual cognitive and noncognitive skills simultaneously affect health and self-employment decisions. We will denote the latter as the 'selection effect' and discuss each of the two effects in the succeeding texts.

2.1. The contextual effect

A useful theoretical framework for understanding the contextual effect of self-employment on health is the so-called job-demand-control model (Karasek, 1979; Karasek and Theorell, 1990; Theorell and Karasek, 1996) that is rooted in sociology and epidemiology. The model emphasizes two aspects of the work environment, job control and job demand, that relate occupational characteristics to health. Job control refers to how much decision-making authority an individual has over when and how to perform the necessary work. Job demand refers to the experienced work intensity and workload. The mismatch between job demands and job control determines the level of occupational stress, which can influence disease incidence and longevity (Cooper and Marshall, 1976; Karasek, 1979; Cooper and Smith, 1985).

Compared with wageworkers, the self-employed have higher levels of job control. As owners of their business, the self-employed have more control over the organization of different tasks and the allocation of resources (Hébert and Link 1989; Prottas and Thompson, 2006). These positive features of self-employment

[1]Yoon and Bernell (2013) use an instrumental variable approach to overcome the selection problem, yet their instruments include the number of self-employed family members, immigrant status, years of labor market experience, the number of children, and having uninsured children, all of which are likely to have a direct relation with health, and thus violate the exclusion restriction.

DOI: 10.1002/hec

also have a downside, which has been called 'a double-edged sword' (Lewin-Epstein and Yuchtman-Yaar, 1991). The self-employed experience higher levels of job demands and workload as opposed to wageworkers (Buttner, 1992; Stephan and Roesler, 2010). Self-employment can turn into 'self-exploitation' as income, job, property, and assets are at stake (Lewin-Epstein and Yuchtman-Yaar, 1991).

An additional mechanism through which self-employment potentially affects health is the lack of health insurance (Zissimopoulos and Karoly, 2007). This may have an effect on health through the use of too little or inappropriate medical care and even directly decrease mental well-being and increase anxiety about financial matters (Finkelstein *et al.* 2012).

The empirical evidence regarding structurally different influences on health of self-employment and wage work is both limited and mixed, and the relative strengths of the positive (job control) and negative (job demand and lack of health insurance) health stimuli of self-employment have not been assessed.

2.2. The selection effect

The entrance into self-employment may be associated with an individual's health status for several reasons. First, ill health decreases the ability to focus on business opportunities (Gielnik *et al.*, 2012). Second, ill health may limit the access to crucial start-up financing (Beck and Demirguc-Kunt, 2006; Klapper *et al.*, 2007), as investors consider health as an important factor determining the investment risk. Third, self-employment is a financially less attractive option compared with wage work for less healthy individuals because income in self-employment hinges much more on the individual ability to work. Fourth, less healthy individuals may stay in wage work because it would be more expensive or impossible to be insured while self-employed.

The aforementioned arguments all would suggest a positive selection of healthier individuals into self-employment. Nonetheless, those with health problems may also have strong difficulties in finding suitable wage work. Employers may discriminate against them in the job-selection procedure, which could push them into the so-called necessity self-employment (Verheul *et al.*, 2010). The empirical evidence about health as an explanatory variable for self-employment is, however, mixed and limited. Using the HRS, Zissimopoulos and Karoly (2007) show that having a health limitation is a pull factor into self-employment. However, Fuchs (1982), Evans and Leighton (1989), and Van Praag and Van Ophem (1995) all show that having a health limitation is not associated with the transition from wage work to self-employment.

Self-employment is also associated with sociodemographic characteristics that independently affect health and health behavior (Lewin-Epstein and Yuchtman-Yaar, 1991), such as age (Zissimopoulos and Karoly, 2007, Parker, 2009), education (Blanchflower, 2000; Lleras-Muney, 2005), perseverance (Beugelsdijk and Noorderhaven, 2005), and risk aversion (Ekelund *et al.*, 2005). This implies that there are several reasons to expect that individuals entering self-employment have a different health profile than wageworkers. However, in which direction the joint effect of this selection mechanism points remains unclear. Most existing papers use cross-sectional data from which, in the absence of exclusion restrictions, it is difficult to disentangle the contextual effect of self-employment on health from a selection effect.[2] In our study, we focus on the aggregate 'net' contextual effect and the aggregate 'net' selection effect. The longitudinal nature of our data allows us to study which of these two effects prevails.

3. DATA

Our study uses the HRS, a longitudinal panel study that surveys a representative sample of Americans over the age of 50 and their spouses every 2 years. The dataset has three advantages: First, the HRS is a population-wide

[2]An exception is Dolinsky and Caputo (2003), who show in a longitudinal sample of women that self-employment has no effect on health, whereas working for wages has a positive effect on health. The significance of the effect difference is, however, not provided, and selection into self-employment is assumed to occur only on the basis of observables.

DOI: 10.1002/hec

study and thus includes both the self-employed and wageworkers. Second, the sample of relatively older individuals represents a phase of life in which many health issues become relevant and apparent and in which there is much policy scope to increase labor-force participation rates. Third, the dataset includes information on several health measures. Despite these advantages, the generalizability of our findings to other age groups may be limited.

We use the HRS Research and Development (RAND) v.L dataset, which consists of 10 biennial waves of data collection (1992–2010). The HRS RAND v.L dataset includes 30,671 individuals, coming from five subsamples. The initial HRS cohort ($N = 13{,}635$, born between 1931 and 1941) and Asset and Health Dynamics among the Oldest Old (AHEAD) cohort ($N = 8334$, born before 1924) participate from wave one onward. The Children of the Depression Age (CODA) cohort ($N = 2420$, born between 1924 and 1930) and War Baby (WB) cohort ($N = 2760$, born between 1942 and 1947) participate from wave four onward. The Early Baby Boomer (EBB) cohort ($N = 3522$, born between 1948 and 1953) joined in wave seven.

We study three health indicators as dependent variables: the number of health conditions, self-reported health, and mental health. We dichotomize these measures to ensure compatibility for all of our empirical methods but made sure that the results are not driven by the dichotomization (refer to section 5 for robustness checks).

The number of health conditions is measured using a nine-point scale, indicating for a set of eight common chronic diseases (arthritis, cancer, diabetes, heart problems, high blood pressure, lung disease, psychiatric problems, and stroke) how many of these that a doctor has *ever* told the respondent that he or she has. Our binary variable *no health conditions* takes the value of 1 if a person has none of the mentioned diseases and 0 if otherwise. Self-reported health is measured on a five-point Likert scale, with categories excellent, very good, good, fair, and poor. Our binary variable *self-reported health* takes the value of 1 if self-reported health is excellent or very good and 0 if otherwise. Mental health is measured on a nine-point Center for Epidemiological Studies Depression Scale (CESD) scale, ranging from 0 (the absence of depression symptoms) to 8 (the presence of all measured depression symptoms). CESD is consistently measured in wave 2–10. Our variable *mental health* takes 1 if CESD equals 0 and 0 if otherwise.

Our main explanatory variable is the binary variable *self-employment*. In each wave, those who were identified as self-employed or to be running their own business are coded as 1, and those who were identified as working for someone else are coded as 0.[3] In addition, we have the following demographic control variables: *gender* (0: female and 1: male), *age* (in years at time of interview), *age-squared*, *race* (0: White, 1: non-White), *years of education* (0–17+ years), *years of education of the father* (0–17+ years), and *years of education of the mother* (0–17+ years). These are well-known factors influencing health and self-employment that are, in general, determined before labor-force entrance. The variables *industry* (1: primary sector, 2: secondary sector, and 3: tertiary sector),[4] *job type* (1: white collar, 2: blue collar, and 3: other),[5] and *working hours* (1: 0–10, 2: 11–30, 3: 31–50, and 4: 51+) are constructed to control for heterogeneity within *self-employment*. We refer to these three variables as the employment controls.

All person-year observations with complete information on health, demographics, and employment are included in the analysis, except those with an age higher than 65 years (the normal retirement age and the age at which Medicare becomes available). This results in a sample size of 44,930 (12,247 individuals) for *no health conditions* and *self-reported health*. A subsample of 35,649 (10,723 individuals) observations is available for *mental health* because this variable has no observations in wave 1. Descriptive statistics of the sample are presented in Table I. In total, there are 36,461 person-year observations for wageworkers and 8469 for the self-employed. Differences in health between the self-employed and wageworkers are small but apparent. Differences in the mean values of the control variables indicate the necessity to control for these observables.

[3]Because our interest is the comparison between the self-employed and wage workers, we do not construct a separate group of retired or unemployed individuals. Our study sample thus reflects the working population.
[4]The primary sector includes agriculture, forestry, fishing, and mining (and construction). The secondary sector includes manufacturing, utilities, and construction. The tertiary sector includes all other job industries.
[5]We followed Forman-Hoffman *et al*. (2008) in the construction of this categorical variable.

DOI: 10.1002/hec

Table I. Descriptive statistics of the analysis sample

	Wageworkers	Self-employed
Health measures		
No health conditions (0/1)	0.36 (0.48)	0.39 (0.49)
Self-reported health (0/1)	0.57 (0.49)	0.61 (0.49)
Mental health (0/1)	0.54 (0.50)	0.56 (0.50)
Demographic controls		
Gender (0: female, 1: male)	0.42 (0.49)	0.57 (0.49)
Age (years)	55.88 (5.15)	56.74 (5.01)
Race (0: White, 1: non-White)	0.18 (0.39)	0.11 (0.32)
Years of education (0–17+ years)	13.21 (2.81)	13.46 (2.80)
Years of education of the father (0–17+ years)	9.80 (3.94)	10.25 (3.88)
Years of education of themother (0–17+ years)	10.12 (3.53)	10.60 (3.52)
Employment controls		
Industry		
Primary sector	4.87%	16.67%
Secondary sector	17.71%	8.76%
Tertiary sector	77.43%	74.57%
Job type		
White collar	65.96%	62.94%
Blue collar	32.77%	30.88%
Other	1.27%	6.19%
Working hours		
0–10	2.72%	8.89%
11–30	14.10%	26.89%
31–50	73.80%	40.71%
51+	9.38%	23.51%
N, person-year observations	36,461	8469
N, individuals	10,399	3050

Note. Mean values are reported, and standard deviations are given in parentheses. For the categorical employment controls, percentages are given per category.
The sum of individuals in wage work and self-employment is larger than the total sample size of 12,247 individuals because of switchers between wage-work and self-employment.

4. METHODS AND RESULTS

4.1. Pooled regressions controlling for observables

First, we compare the average health status of the self-employed with that of wageworkers. Using pooled logit regressions, we explain *no health conditions*, *self-reported health*, and *mental health*. In these models, a significant positive coefficient for *self-employment* means that the self-employed are healthier than wageworkers. Wave dummies are included in each regression, and the standard errors are clustered at the individual level. We run three model specifications for each dependent variable. In the first specification, we only include *self-employment*, which produces the simple association between self-employment and health. In the second specification, we add the demographic control variables to investigate whether observed characteristics such as education and age are responsible for the association between self-employment and health. In the third specification, we add the employment controls to verify that the association not simply reflects differences in the industry sector, occupational level, or working hours across the self-employed and wageworkers. The results are presented in Table II. The regression coefficients for the demographic controls in the second specification are not reported because they are very similar to those in the model including both demographic and employment controls.

The coefficients for *self-employment* are all positive and statistically significant in the univariate regressions for *no health conditions*, *self-reported health*, and *mental health*. Odds ratios are respectively 1.15 for *no health conditions*, 1.17 for *self-reported health*, and 1.11 for *mental health*. The inclusion of demographic and employment controls lowers the value of the coefficient in the *self-reported health* regression. The coefficient remains,

DOI: 10.1002/hec

Table II. Regression coefficients for the pooled logit models explaining *no health conditions*, *self-reported health*, and *mental health*

	No health conditions	Self-reported health	Mental health
Pooled logit			
Self-employment	0.15** (0.04)	0.16*** (0.04)	0.10** (0.04)
+ Demographic controls	No	No	No
+ Employment controls	No	No	No
Pooled Logit			
Self-employment	0.17*** (0.04)	0.10** (0.04)	0.01 (0.04)
+ Demographic controls	Yes	Yes	Yes
+ Employment controls	No	No	No
Pooled logit			
Self-employment	0.17*** (0.05)	0.11** (0.04)	0.03 (0.04)
Gender	0.18*** (0.04)	−0.08* (0.04)	0.27*** (0.03)
Age	−0.03 (0.04)	−0.03 (0.03)	−0.01 (0.03)
Age-squared	−0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Race	−0.15** (0.05)	−0.49*** (0.04)	−0.31*** (0.04)
Years of education	0.03** (0.09)	0.13*** (0.01)	0.06*** (0.01)
Years of education of the father	0.01 (0.01)	0.02** (0.01)	0.02*** (0.01)
Years of education of the mother	−0.01 (0.01)	0.03*** (0.01)	0.02** (0.01)
Industry (base: primary sector)			
Secondary sector	0.00 (0.09)	−0.03 (0.07)	−0.02 (0.07)
Tertiary sector	−0.05 (0.08)	−0.05 (0.06)	−0.01 (0.07)
Job type (base: white collar)			
Blue collar	−0.12** (0.04)	−0.27*** (0.04)	−0.28*** (0.04)
Other	0.07 (0.13)	−0.22* (0.11)	−0.32** (0.11)
Working hours (base: 0–10 h)			
11–30	−0.04 (0.07)	0.05 (0.07)	−0.06 (0.06)
31–50	0.00 (0.07)	0.10 (0.06)	−0.04 (0.06)
51+	0.06 (0.08)	0.18* (0.07)	−0.15* (0.07)
N, person-year observations	44,930	44,930	35,649
N, individuals	12,247	12,247	10,723

Note. Standard errors are in parentheses and clustered per individual. *$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.

however, significant. For *mental health*, the coefficient also becomes smaller but becomes insignificant. Interestingly, the adjustment for demographics and employment characteristics increases the *self-employment* coefficient in the *no health conditions* regression. Altogether, we find that the self-employed are in better health than wageworkers, although the difference in mental health is not statistically significant once demographic variables are controlled for.

Among the diseases that make up the *no health conditions* variable, we find that heart problems ($p = 0.008$) and high blood pressure ($p = 0.001$) are significantly negatively associated in the models including demographic and employment controls; the self-employed have these conditions less often than wageworkers. The other health conditions (arthritis, cancer, diabetes, lung disease, psychiatric problems, and stroke) are not significantly associated at the 5% level.

4.2. Regressions controlling for lagged health and time-invariant unobservables

Next, we perform longitudinal analyses to investigate reverse causality from health to self-employment, if health is a pull or push factor into or out of self-employment. Inspired by Granger (1969), Adams *et al*. (2003), and Stowasser *et al*. (2011), we investigate whether the lagged self-employment status has explanatory power for current health, while controlling for lagged health. A coefficient for self-employment that is qualitatively similar to the coefficient for self-employment in the pooled logit regression would strongly suggest that the association between self-employment and health is not completely the result of reverse causality. Again, we use a pooled logit regression with wave dummies and standard errors clustered per individual. As dependent variables, we take only *self-reported health* and *mental health*. The way in which *no health conditions* is measured in the HRS makes it

DOI: 10.1002/hec

Table III. Regression coefficients for *self-employment* in the models explaining *self-reported health* and *mental health*

	Self-reported health	Mental health
Pooled logit (lag self-employment and lag health)		
Univariate regression	0.11** (0.04)	0.07 (0.04)
+ Demographic controls	0.08* (0.04)	−0.00 (0.04)
+ Employment controls	0.08* (0.04)	0.00 (0.04)
N, person-year observations	30,918	23,600
N, individuals	9970	8315
Fixed-effects logit		
Univariate regression	0.01 (0.09)	−0.06 (0.09)
+ Demographic controls	0.01 (0.09)	−0.06 (0.09)
+ Employment controls	0.02 (0.09)	−0.06 (0.09)
N, person-year observations	20,909	20,323
N, individuals	4502	4718
Bivariate probit		
$\rho = 0.00$	0.06*** (0.01)	0.02 (0.02)
$\rho = 0.10$	−0.10*** (0.01)	−0.14*** (0.02)
$\rho = 0.20$	−0.26*** (0.01)	−0.30*** (0.02)
Equal selection	−0.14*** (0.01)	−0.20*** (0.02)
N, person-year observations	44,930	35,649
N, individuals	12,247	10,723

Note. Standard errors are in parentheses and clustered per individual. Standard errors in the bivariate probit model are based on 250 bootstrap replications. *$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.

unsuitable for inclusion in longitudinal analyses.[6] Again, we implement three model specifications; the only difference is that we include a lag of the dependent variable and the lag of *self-employment* instead of current *self-employment*.

Panel 1 of Table III reports the results. We find qualitatively the same results as those presented in Table II.[7] The coefficient for *self-employment* is significant in the regressions for *self-reported health*. For *mental health*, only the univariate model shows this association. We conclude that for these two health measures, the association from Table II cannot be entirely because of reverse causality.

As argued by Granger (1969), the explanatory power of self-employment for future realizations of health implies a form of causality. However, this type of causality cannot distinguish between a contextual effect of self-employment on health and third factors influencing both self-employment and health. We use a fixed-effects logit regression to control for unobserved heterogeneity deriving from possible time-invariant third factors influencing both self-employment and health. Examples of such variables could be risk aversion, time preferences, and genetic factors. A significant association between self-employment and health that remains after controlling for fixed unobserved determinants of self-employment and health would be consistent with a contextual effect of self-employment on health. Because time-invariant variables are accounted for in the fixed effect, we only control for *self-employment* and time-varying control variables in our three model specifications.

The results of the fixed-effects panel regressions are in panel 2 of Table III. Note that the sample size is somewhat smaller here because in the fixed-effects logit regression, the individuals without a change in the dependent

[6]*No health conditions* is measured in such a way that it only increases with age because the question is asked whether the doctor has *ever* told the respondent to have a certain chronic condition. The only possible change is from 0 (no health condition ever had) to 1 (at least one health condition ever had). This approach makes the correlation between measures in two consecutive waves almost 1. Moreover, this measure does not necessarily precisely reflect the change in the health status of an individual. For example, someone who completely recovered from a heart attack will always be seen in the data as having at least one health condition.

[7]We repeated the pooled logit analyses without the lagged dependent variable using only the person-year observations from the analyses with the lagged dependent variable. The regression results are qualitatively the same as those presented in Table II.

DOI: 10.1002/hec

variable are dropped.[8] The associations for *self-reported health* and *mental health* with *self-employment* are not significant. Hence, changes in *self-reported health* and *mental health* do not appear to be related to changes in *self-employment*. We consider this result as evidence against a contextual effect of self-employment on health. It also suggests that unobserved time-invariant individual characteristics influence both self-employment and health and that the positive association between self-employment and health mainly reflects a 'selection effect' in which intrinsically healthier individuals select into self-employment.

An additional piece of evidence for a selection effect comes from the inclusion of higher-order lags of self-employment into the univariate pooled logit regressions. The coefficients for higher-order lags of self-employment remain similar in size as those presented in Table II and statistically significant at the 5% level up to the fourth (*self-reported health*) and second lag (*mental health*), which would be counterintuitive if self-employment were to cause good health. Rather, these results suggest that the coefficient for *self-employment* picks up the effect of unobserved time-invariant individual characteristics that are associated with better health.

4.3. Pooled regressions controlling for unobservables

The fixed-effects logit model has two limitations. First, the model only controls for *time-invariant* third factors, while *time-varying* factors influencing both health and self-employment could also play a role. Second, the coefficients are only identified based on individuals who switch between self-employment and wage work. Such a switch is relatively rare (less than 5% between every two waves), resulting in a small and possibly nonrandom sample if switching is induced by time-varying factors that are not controlled for. To reduce this concern, we also implement a method proposed by Altonji *et al*. (2005) that uses the selection on observable variables as an indication for the potential selection on unobservable variables. Essentially, their idea is that it is unlikely that by controlling for the observed individual characteristics available in the dataset, all factors influencing both self-employment and health are controlled for. There will always be unobserved factors affecting decisions with respect to health and self-employment. However, the authors argue that the observed characteristics available in the dataset are typically carefully chosen, such that the selection of observable characteristics can be seen as an upper bound to the selection based on unobservable characteristics.

Specifically, Altonji *et al*. (2005) suggest upping a bivariate probit model to quantify how large the selection on the basis of unobservable variables into self-employment would have to be to fully account for the association between self-employment and health. Their suggested model depends on an assumption about the correlation ρ between the error components in the equations for self-employment and health.[9] Altonji *et al*. (2005) additionally suggest estimating a 'worst-case' scenario where it is assumed that the selection on observable variables is equal to the selection on unobservable variables, which places a particular constraint on the value of ρ in the estimation of the bivariate probit model (refer to Altonji *et al*., 2005 for details). This scenario creates an alternative way to gauge the plausibility of a contextual effect of self-employment on health without the need to rely solely on individuals switching jobs.

The bivariate probit results are given in the bottom panel of Table III. These regressions include both the demographic and employment controls. Obviously, in the models where we impose $\rho = 0$, we obtain qualitatively the same results as those presented in Table II because both models correspond to running separate probit/logit regressions for health and self-employment. When we constrain and increase the correlation between the error components in the health and self-employment regressions ($\rho = 0.10$, $\rho = 0.20$), the coefficient for *self-employment* becomes strongly significant in the opposite direction. This result suggests that a relatively small correlation

[8] A fixed-effect ordinary least squares regression that includes all available person-year observations gives qualitatively the same results. Moreover, the pooled logit results from Table II remain qualitatively the same if the regression analyses are restricted to the person-year observations that are included in the fixed-effect logit regressions.

[9] The equations take the form $E = \mu + \gamma X + \eta$ and $H = \alpha + \beta E + \tau X + \varepsilon$, where individual subscripts are omitted; E is self-employment, and X are the observed characteristics. H represents health, and η and ε are the error terms for self-employment and health, respectively. The correlation between these error terms is typically denoted by ρ in the bivariate probit model.

DOI: 10.1002/hec

between the error components (unobserved factors) of self-employment and health already accounts for the entire positive association and, in fact, even turns it negative.

However, in practice, we do not know the value of ρ. We therefore also present the 'worst-case scenario', where the selection on observable variables is assumed to be similar to the selection on unobservable variables. Under this scenario, the parameter ρ is estimated to be 0.12 for *self-reported health* and 0.16 for *mental health*. These positive correlations indicate that the unobserved factors influencing self-employment and health are positively correlated, which implies that healthier individuals are more likely to become self-employed. Imposing this restriction on ρ, the coefficients for *self-employment* are in both models negative and significant (Table III, bottom panel). The results thus confirm that, as in the fixed-effects regressions, if selection on unobserved variables is considered, the positive cross-sectional association between self-employment and health disappears.

In fact, the association even becomes negative, which would suggest a negative effect of self-employment on health. Because these latter results depend on a subjective judgment on the importance of unobserved explanatory variables in the regressions, we see these results as complimentary to the fixed-effects panel regression outcomes, which showed no contextual effect of self-employment on health. Altogether, in our view, the results provide compelling evidence that the contextual effect of self-employment on health is nonpositive, possibly zero, and, if anything, negative.

5. ROBUSTNESS CHECKS

In this section, we discuss a number of analyses performed to gauge the robustness of our findings. All results are available upon request. First, we investigated whether health-related attrition out of work was different between the self-employed and wageworkers. We estimated a pooled logit model explaining the probability of not working in the next wave by self-employment, current health, and the interaction between self-employment and current health. We find that this interaction term is not significant at the 5% level for *no health conditions*, *self-reported health*, and *mental health*. This result suggests that health-related attrition out of work is not different across the self-employed and wageworkers.

For consistency across methods, we dichotomized our three dependent variables, which unavoidably requires introducing an arbitrary threshold. We tested the sensitivity to this dichotomization in two different ways. First, when estimating ordered logit regressions rather than binary logit regressions, the initial association between self-employment and our three health outcomes exists and is similar in magnitude to the one presented in Table II. Second, we varied the thresholds in the dichotomization. For *no health conditions*, if we place the threshold on the nine-point scale between 1 and 2 and between 2 and 3, we get a significant coefficient of 0.16 and 0.14, respectively, very much in line with the base result. Imposing the threshold higher on the scale is difficult because less than 3% of the sample has more than three diseases. For *self-reported health*, if we place the threshold between excellent and very good, we get a significant coefficient of 0.29. Interestingly, the coefficient for *self-employment* is not significant if the threshold lies between good and fair. While this may partly be because of lack of power (the bad health group in this analysis is only 13%), it suggests that the association mainly reflects a disproportionate fraction of the self-employed reporting their health to be 'very good' and 'excellent'. Placing the threshold for dichotomization on the nine-point CESD scale for mental health between 1 and 2, between 2 and 3, and between 3 and 4 gives a coefficient of 0.07, 0.09, and 0.12 for *self-employment*, respectively (only the latter is significant). Less than 7% of the person-year observations have more than four depression symptoms.

We also looked at other, more indirect, health measures available in the dataset, to investigate the consistency of the reported findings and to get a better impression of the mechanisms involved. As alternative, more physical health measures, we selected *overweight* (BMI > 25), *obese* (BMI > 30), and *back problems* (1/0). We find that *self-employment* is negatively associated with *overweight* ($p = 0.04$) and *obese* ($p = 0.05$) but not associated with having back problems ($p = 0.56$). The HRS also asks respondents whether *health limits work* (1/0), which we find to be positively associated with self-employment ($p < 0.01$). This corroborates our main finding that self-employment does not strongly affect one's health but rather that health is a strong

DOI: 10.1002/hec

determinant of self-employment decisions. Rather than exiting the labor force completely, less healthy individuals may decide to become self-employed if they cannot find wage work.

6. CONCLUSION

It is notoriously difficult to discriminate between a contextual effect of self-employment on health and health-related selection of individuals into self-employment. However, this discrimination is a prerequisite for health policy development concerning this quantitatively and qualitatively important part of the labor force. Therefore, we use several methods to distinguish between these two effects. We find the self-employed to be generally healthier than wageworkers, both in terms of subjective health outcomes as well as in more objective outcomes such as the absence of chronic conditions. While it is tempting to attribute these results to the high level of job control and to even consider self-employment as a viable alternative to health-induced early retirement, our results suggest that the health differences are explained by a selection effect, in which healthier individuals self-select into self-employment.

This main conclusion is supported by the absence of a statistically significant effect of self-employment on health in fixed-effects regressions, which suggests that time-invariant individual characteristics influence both self-employment and health. Additionally, applying methods proposed by Altonji *et al*. (2005) suggests that it only takes a relatively small amount of selection based on unobserved characteristics into self-employment and health to fully account for the positive association between the two. These results are in line with the two-time-period, female subject only, study on the relation between self-employment and health by Dolinsky and Caputo (2003).

Our results not only emphasize the importance of a selection of comparatively healthier individuals into self-employment but also provide suggestive evidence that the contextual effect of self-employment on health, if anything, is negative. This conclusion is, however, tentative and based upon relatively strong assumptions on the amount of selection on the basis of unobserved individual characteristics. Nonetheless, the results do show that health does not seem to be a nonmonetary benefit of self-employment as was proposed by Stephan and Roesler (2010). In fact, the influence of self-employment is potentially even negative.

Further research is needed to identify the factors influencing both self-employment and health. Apart from traditional and more obvious variables such as risk aversion and perseverance, a recent line of inquiry has stressed the role of genes. Self-employment is, to a certain extent, influenced by genetic factors (Nicolaou *et al.*, 2008; Van der Loos *et al.*, 2013). It is perceivable that the same genetic factors influence both self-employment and health (such a mechanism is called pleiotropy in genetics). Although it falls outside the scope of this paper to reveal these and other joint causal factors, the possible finding of a shared causal factor for self-employment and health would be a major breakthrough.

Awareness of the presence of the selection mechanism is important for both policy makers and individuals who consider becoming self-employed. Stimulating self-employment is a key objective in many countries because of its assumed contribution to economic growth. The existence of entrance barriers may prevent such a policy to be successful. Our results indicate that health status may be such a barrier. Because we cannot distinguish between health itself and correlates of health (such as expected health or health of a spouse), future research should further disentangle the selection mechanism to establish whether health status is a perceived barrier (the less healthy do not even try to become self-employed) or an actual barrier (the less healthy are faced with more obstacles, such as in the process of securing loans, when they want to start a business).

ACKNOWLEDGEMENTS

The HRS is sponsored by the National Institute on Aging (NIA; grant number NIA U01AG009740) and is conducted by the University of Michigan. Van Kippersluis gratefully acknowledges funding from the NIA under grant R01AG037398 and the Netherlands Organization of Scientific Research (NWO Veni grant 016.145.082).

DOI: 10.1002/hec

The authors are grateful for the valuable comments of Martin Carree and André van Stel on earlier versions of this paper. All errors that remain are of the authors.

REFERENCES

Adams P, Hurd MD, McFadden D, Merrill A, Ribeiro T. 2003. Healthy, wealthy, and wise? Tests for direct causal paths between health and socioeconomic status. *Journal of Econometrics* **112**: 3–56.

Altonji JG, Elder TE, Taber CR. 2005. Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools. *Journal of Political Economy* **113**: 151–184.

Audretsch DB, Keilbach M. 2004. Entrepreneurship capital and economic performance. *Regional Studies* **38**: 949–959.

Audretsch DB, Thurik AR. 2000. Capitalism and democracy in the 21st century: From the managed to the entrepreneurial economy. *Journal of Evolutionary Economics* **10**: 17–34.

Audretsch DB, Thurik AR. 2001. What is new about the new economy: sources of growth in the managed and entrepreneurial economies. *Industrial and Corporate Change* **10**: 267–315.

Beck T, Demirguc-Kunt A. 2006. Small and medium-size enterprises: Access to finance as a growth constraint. *Journal of Banking & Finance* **30**: 2931–2943.

Beugelsdijk S, Noorderhaven N. 2005. Personality characteristics of self-employed: an empirical study. *Small Business Economics* **24**: 159–167.

Bjuggren CM, Johansson D, Stenkula M. 2012. Using self-employment as proxy for entrepreneurship: some empirical caveats. *International Journal of Entrepreneurship and Small Business* **17**: 290–303.

Blanchflower, D.G. 2000. Self-employment in OECD countries. *Labour Economics* **7**: 471–505.

Bradley DE, Roberts JA. 2004. Self-employment and job satisfactions: Investigating the role of self-efficacy, depression, and seniority. *Journal for Small Business Management* **42**: 37–58.

Buttner EH. 1992. Entrepreneurial stress: is it hazardous to your health? *Journal of Managerial Issues* **4**: 223–240.

Carree MA, Thurik AR. 2010. The impact of entrepreneurship on economic growth. In Audretsch DB, Acs ZJ (Eds), *Handbook of entrepreneurship research*. Springer Verlag: Berlin, Heidelberg. 557–594.

Cooper C, Marshall J. 1976. Occupational sources of stress: a review of the literature relating to coronary heart disease and mental ill health. *Journal of Occupational Psychology* **49**: 1–28.

Cooper C, Smith M. 1985. Job Stress and Blue Collar Work. Wiley: Chichester.

Dahl MS, Nielsen J, Mojtabai R. 2010. The effects of becoming an entrepreneur on the use of psychotropics among entrepreneurs and their spouses. *Scandinavian Journal of Public Health* **38**: 857–863.

Dolinsky AL, Caputo RK. 2003. Health and female self-employment. *Journal of Small Business Management* **41**: 233–241.

Ekelund J, Johansson E, Järvelin MR, Lichtermann D. 2005. Self-employment and risk aversion—evidence from psychological test data. *Labour Economics* **12**: 649–659.

European Commission. 2004. Action plan: the European agenda for entrepreneurship at ftp://ftp.cordis.europa.eu/pub/incubators/docs/action_plan_on_entrepreneurship.pdf

Evans D, Leighton L. 1989. Some empirical aspects of entrepreneurship. *American Economic Review* **79**: 519–535.

Finkelstein A, Taubman S, Wright B, Bernstein M, Gruber J, Newhouse JP, Allen H, Baicker K, the Oregon Health Study Group. 2012. The Oregon Health Insurance Experiment: Evidence from the First Year. *Quarterly Journal of Economics* **127**(3): 1057–1106.

Forman-Hoffman VL, Richardson KK, Yankey JW, Hillis SL, Wallace RB, Wolinsky FD. 2008. Retirement and weight changes among men and women in the Health and Retirement Study. *Journal of Gerontology* **638**: S146–S153.

Fuchs VR. 1982. Self-employment and labor force participation of older males. *Journal of Human Resources* **17**: 339–357.

Gielnik MM, Zacher H, Frese M. 2012. Focus on opportunities as a mediator of the relationship between business owners' age and venture growth. *Journal of Business Venturing* **27**: 127–142.

Gilbert BA, Audretsch DB, McDougall PP. 2004. The emergence of entrepreneurship policy. *Small Business Economics* **22**: 313–323.

Granger CWJ. 1969. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* **39**: 424–428.

Hébert RF, Link AN. 1989. In search of the meaning of entrepreneurship. *Small Business Economics* **1**: 39–49.

Jamal M. 1997. Job stress, satisfaction, and mental health: An empirical examination of selfemployed and non-self-employed Canadians. *Journal of Small Business Management* **35**: 48–57.

Juster TF, Suzman R. 1995. An overview of the Health and Retirement Study. *Journal of Human Resources* **20**: 7–56.

Karasek RA. 1979. Job demands, job decision latitude, and mental strain: Implications for job redesign. *Administrative Science Quarterly* **24**: 285–308.

Karasek RA, Theorell T. 1990. Healthy work: Stress, productivity, and the reconstruction of working life. Basic Books: New York.

DOI: 10.1002/hec

Klapper L, Amit R, Guillen MF, Quesada JM. 2007. Entrepreneurship and firm formation across countries. *World Bank Policy Research Working Paper* **4313**.
Koellinger PD, Thurik AR. 2012. Entrepreneurship and the business cycle. *Review of Economics and Statistics* **94**: 1143–1156.
Lewin-Epstein N, Yuchtman-Yaar E. 1991. Health risks of self-employment. *Work and Occupations* **18**: 291–312.
Lleras-Muney A. 2005. The relationship between education and adult mortality in the United States. *Review of Economic Studies* **72**: 189–221.
Nicolaou N, Shane S, Cherkas L, Hunkin J, Spector TD. 2008. Is the tendency to engage in entrepreneurship genetic? *Management Science* **54**: 167–179.
Parker SC. 2009. The economics of entrepreneurship. Cambridge University Press: Cambridge.
Parslow RA, Jorm AF, Christensen H, Rogers B, Strazdins L, D'Souza RM. 2004. The association between work stress and mental health: A comparison of organizationally employed and self-employed individuals. *Work and Stress* **18**: 231–244.
Prottas DJ, Thompson CA. 2006. Stress, satisfaction, and the work–family interface: A comparison of self-employed business owners, independents, and organizational employees. *Journal of Occupational Health Psychology* **11**: 366–378.
Ravesteijn B, van Kippersluis H, van Doorslaer E. 2013. The Contribution of Occupation to Health Inequality. In *Research on Economic Inequality*, Pedro Rosa Dias, Owen O'Donnell (eds.), Health and Inequality, vol. 21. Emerald Group Publishing: Bingley, UK.
Stephan U, Roesler U. 2010. Comparison of entrepreneurs' and employee's health in a national representative sample. *Journal of Occupational and Organizational Psychology* **83**: 717–738.
Stowasser T, Heiss F, McFadden D, Winter J. 2011. Healthy, wealthy and wise? Revisited: An analysis of the causal pathways from socio-economic status to health. *National Bureau of Economic Research (NBER) Working Paper* **17273**.
Tetrick LE, Slack K, Da Silva N, Sinclair RR. 2000. A comparison of the stress-strain process for business owners and nonowners: Differences in job demands, emotional exhaustion, satisfaction, and social support. *Journal of Occupational Health Psychology* **5**: 464–476.
Theorell T, Karasek RA. 1996. Current issues relating to psychosocial job strain and cardiovascular disease research. *Journal of Occupational Health Psychology* **1**: 9–26.
Torres O. 2012. La Santé du Dirigeant: de la Souffrance Patronage à l'Entrepreneuriat Salutaire. De Boeck: Brussels.
Van der Loos MJHM, Rietveld CA, Eklund N, Koellinger PD, Rivadeneira F, Abecasis GR, … Thurik AR. 2013. The molecular genetic architecture of self-employment. *PloS One* **8**: e60542.
Van Praag, CM, van Ophem H. 1995. Determinants of willingness and opportunity to start as an entrepreneur. *Kyklos* **48**: 513–540.
Verheul I, Thurik AR, Hessels J, Van der Zwan PW. 2010. Factors influencing the entrepreneurial engagement of opportunity and necessity entrepreneurs. Research Reports EIM Zoetermeer.
Yoon J, Bernell SL. 2013. The effect of self-employment on health, access to care, and health behavior. *Health* **5**: 2116–2127.
Zissimopoulos J, Karoly LA. 2007. Transitions to self-employment at older ages: The role of wealth, health, health insurance, and other factors. *Labour Economics* **14**: 269–295.

DOI: 10.1002/hec