TI 2013-129/V
Tinbergen Institute Discussion Paper

Self-employment and Health: Barriers or Benefits?

Cornelius A. Rietveld[1]
Hans van Kippersluis[1,2]
A. Roy Thurik[1,2,3,4]

[1] *Erasmus School of Economics, Erasmus University Rotterdam, The Netherlands;*
[2] *Tinbergen Institute;*
[3] *Panteia, Zoetermeer, The Netherlands;*
[4] *GSCM-Montpellier Business School, France.*

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at http://www.tinbergen.nl

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Duisenberg school of finance is a collaboration of the Dutch financial sector and universities, with the ambition to support innovative research and offer top quality academic education in core areas of finance.

DSF research papers can be downloaded at: http://www.dsf.nl/

Duisenberg school of finance
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 8579

Self-employment and Health: Barriers or Benefits?

Cornelius A. Rietveld[1], Hans van Kippersluis[1,2], and A. Roy Thurik[1,2,3,4]

[1]Department of Applied Economics, Erasmus School of Economics, Erasmus University Rotterdam, 3000 DR, Rotterdam, The Netherlands

[2]Tinbergen Institute Rotterdam, 3062 PA, Rotterdam, The Netherlands

[3]Panteia, Zoetermeer 2701 AA, The Netherlands

[4]GSCM-Montpellier Business School, Montpellier 34185, France

Abstract

The self-employed are often reported to be healthier than wage workers; 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, causal 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.

JEL codes: I19, L26

Keywords: Health, HRS, selection, self-employment

Version: September 2013

Acknowledgements: The Health and Retirement Study is sponsored by the National Institute on Aging (grant number NIA U01AG009740) and is conducted by the University of Michigan. Van Kippersluis gratefully acknowledges funding from the National Institute on Aging (NIA) under grant R01AG037398. The authors are grateful for the valuable comments of Martin Carree and André van Stel on earlier versions of this paper. All errors remain with the authors.

Corresponding author: Cornelius A. Rietveld, Department of Applied Economics, Erasmus School of Economics, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR, Rotterdam, The Netherlands, nrietveld@ese.eur.nl, tel.: +31104088935, fax: +31104089141

1. Introduction

Many governments stimulate self-employment (Gilbert et al., 2004; European Commission, 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 ageing 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, (i) scarce, (ii) conflicting, and, partly due to the cross-sectional nature of existing analyses, (iii) poorly understood (Torres, 2012). This is surprising given the quantitative and qualitative 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 wage workers (Buttner, 1992; Jamal, 1997; Lewin-Epstein and Yuchtman-Yaar, 1991; Parslow et al., 2004; Dahl et al., 2010). All cited studies emphasize structural differences between self-employment and wage work to explain the difference in health between the self-employed and wage workers. The self-employed operate their business independently, without the control of a supervisor, while wage workers 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).

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. Such a selection mechanism is more difficult to reveal because longitudinal data are required.

In this paper, we use the rich dataset of 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 wage workers for all three available measures of health: 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-effect 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 (Altonji et al., 2005).

This paper is organized as follows. Section two discusses related literature, after which section three describes the data. In section four, the empirical methods are explained, after which section five presents the results. Section six 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 wage workers. Contemporaneous "reverse causality" could occur where individuals decide to quit or enter self-employment for health reasons, and pre-determined individual cognitive and non-cognitive skills could

simultaneously affect health and self-employment decisions, resulting in a spurious association between self-employment and health. We will denote the latter as the "selection effect". We discuss each of these correlations below.

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 to wage workers, 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 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 wage workers (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).

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

The reverse causality effect

Self-employment can attract individuals with a specific health profile, different from that of wage workers. First, ill health decreases the ability to focus on business opportunities (Gielnik et al., 2012): seeing opportunities and also pursuing them are only possible for healthy individuals. Wage work may therefore be a more natural

choice for less healthy individuals. Second, compared to wage work, income in self-employment hinges much more on the individual ability to work. For example, in most Western countries, social security benefits related to illness are much lower for the self-employed than for wage workers (European Commission, 2004). This makes choosing self-employment a less attractive option for less healthy individuals. Third, access to start-up financing is crucial for business start-up (Klapper et al., 2007; Beck and Demirguc-Kunt, 2006). Health will influence the investment risk, which will be considered by investors, and imposes as such an access burden for those who want to become self-employed but do not have perfect health.

The aforementioned arguments all would suggest a positive selection of healthier individuals into self-employment; however, particularly those with health problems may have strong difficulties finding suitable wage work. Employers may discriminate against them in the job-selection procedure, which could push them into so-called necessity self-employment (Verheul et al., 2010). Indeed, using the HRS, Zissimopoulos and Karoly (2007) show that having a health limitation is a pull factor into self-employment. However, although Fuchs (1982) shows that good health is a predictor for continuing work at older ages among males, his results also show that having a health limitation is not associated with the transition from wage work to self-employment. Evans and Leighton (1989) and Van Praag and Van Ophem (1995) show that having a health limitation is not associated with the choice for self-employment. Thus, the empirical evidence about health as an explanatory variable for the choice to become self-employed is mixed and is unfortunately only based on a single binary variable that indicates whether an individual has health limitations or not.

The selection effect

Self-employment is also associated with certain sociodemographic characteristics that independently affect health and health behavior (Lewin-Epstein and Yuchtman-Yaar, 1991). One obvious characteristic is age. In most empirical studies, the relationship between age and the decision to start a business follows an inverse U-shaped pattern (Parker, 2009). The increasing rate of self-employment over age is well documented (Fuchs, 1982; Zissimopoulos and Karoly, 2007). Two health-related reasons can be given for these observations. First, because ill-health is known to be an important cause for exiting the workforce (Bound et al, 1999), if the self-

employed are healthier than wage workers, then the positive correlation between age and the self-employment rate can be explained by wage workers leaving the labor market sooner than the self-employed. Second, a switch from wage work to self-employment can be seen as a type of partial retirement and an alternative to exiting the workforce entirely because self-employment involves much more flexibility regarding workload (Quinn, 1980). However, the role of health in this decision is ambiguous. Those individuals in good health can decide to work somewhat less instead of retiring completely; however, those individuals in ill health can also decide to work somewhat less because they are no longer able to perform fulltime wage work.

Further-known characteristics affecting both self-employment and health are, for example, education (Blanchflower, 2000; Lleras-Muney, 2005), perseverance (Beugelsdijk and Noordhaven, 2005), and risk aversion (Ekelund et al., 2005), which implies that there are several reasons to expect that individuals entering self-employment have a different health profile than wage workers. In what 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.[1] 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, such that we are able to answer the question whether self-employment makes the self-employed healthier than wage workers.

3. Data

Our study uses the Health and Retirement Study, a longitudinal panel study that surveys a representative sample of Americans over the age of 50 every two years. The dataset has three advantages: First, the HRS is a population-wide study and thus includes both the self-employed and wage workers. 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

[1] An exception is Dolinksy 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.

participation rates. Third, the dataset includes information on several health measures. We use the HRS RAND v.L dataset, which consists of ten biennial waves of data collection (1992 – 2010).

We study three health indicators as dependent variables: number of health conditions, self-reported health and mental health. We dichotomize these measures to ensure compatibility for all of our empirical methods. We note that the results are not qualitatively different if we keep the original categorical measures and are thus not driven by this dichotomization.

The number of health conditions is measured using a 9-point scale, indicating for a set of 8 common chronic diseases (arthritis; cancer; diabetes; heart problems; high blood pressure; lung disease; psychiatric problems; stroke) how many of these a doctor has *ever* told the respondent that he or she has. Our binary variable *No Health Conditions* takes the value 1 if a person has none of the mentioned diseases and 0 otherwise. Self-reported health is measured on a 5-point Likert scale, with 1: Excellent, 2: Very good, 3: Good, 4: Fair, and 5: Poor. Our binary variable *Self-reported Health* takes the value 1 if self-reported health is Excellent or Very good and 0 otherwise. Mental health is measured on a 9-point CESD (Center for Epidemiological Studies Depression Scale) scale, ranging from 0 (absence of depression symptoms) to 8 (presence of all measured depression symptoms). CESD is consistently measured in wave 2-10 (wave 1 uses a different scale); therefore, we only use the variables of wave 2-10. Our variable *Mental Health* takes 1 if CESD equals 0 and 0 otherwise.

Our main independent variable is the binary variable *Self-employment*. In each wave, those who identified as self-employed or to be running their own business are coded as 1, and those who identified as working for someone else are coded as 0.[2] In addition, we have the following demographic control variables: *Gender* (0: female, 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 father* (0 – 17+ years), and *Years of education 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, 3: Tertiary

[2] 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.

sector)[3], *Job type* (1: White collar, 2: Blue collar, 3: Other)[4], and *Working hours* (1: 0-10, 2: 11-30, 3: 31-50, 4: 51+) are constructed to control for heterogeneity within *Self-employment*. We refer to these three variables as the employment controls.

The HRS RAND v.L dataset includes 30,671 individuals, with potentially 306,710 person-year observations. However, not all individuals participated in all 10 waves of data collection, and often only a subset of the full questionnaire is answered by a participant. The restriction to person-year observations with complete information on health, demographics and employment results in a sample size of 55,689 for *No Health Conditions* and *Self-reported Health*. A subsample of 45,622 observations is available for *Mental Health* because this variable has no observations in wave 1. Descriptive statistics of the sample are presented in Table 1. In total, there are 43,438 person-year observations for wage workers and 12,251 for the self-employed. Differences in health between the self-employed and wage workers are small but apparent. Differences in the mean values of the control variables indicate the necessity to control for these observables.

[3] The primary sector includes agriculture, forestry, fishing, and mining (& construction). The secondary sector includes manufacturing, utilities, and construction. The tertiary sector includes all other job industries.
[4] We followed Forman-Hoffman et al. (2008) in the construction of this categorical variable.

Table 1. Descriptive statistics of the analysis sample. Mean values are reported, and standard deviations are given in parentheses. For the categorical employment controls, percentages are given per category.

	Wage workers	Self-employed
Health measures		
No Health Conditions (0/1)	0.33 (0.47)	0.33 (0.47)
Self-reported Health (0/1)	0.56 (0.50)	0.59 (0.49)
Mental Health (0/1)	0.54 (0.50)	0.57 (0.49)
Demographic controls		
Gender (0: female, 1: male)	0.44 (0.50)	0.62 (0.49)
Age (years)	58.04 (7.04)	61.20 (8.45)
Race (0: white, 1: non-white)	0.18 (0.38)	0.11 (0.31)
Years of education (0-17+ years)	13.17 (2.83)	13.47 (2.85)
Years of education father (0-17+ years)	9.70 (3.93)	10.02 (3.85)
Years of education mother (0-17+ years)	10.05 (3.53)	10.41 (3.46)
Employment controls		
Industry		
Primary sector	4.80%	16.83%
Secondary sector	16.58%	9.33%
Tertiary sector	78.63%	73.84%
Job type		
White collar	65.20%	63.59%
Blue collar	33.39%	29.72%
Other	1.40%	6.69%
Working hours		
0-10	4.55%	13.65%
11-30	18.43%	30.13%
31-50	68.45%	36.82%
51+	8.57%	19.40%
Sample size	43,438	12,251

4. Method

Pooled regressions controlling for observables

First, we compare the average health status of the self-employed with that of wage workers. Using pooled logit regression, we explain *No Health Conditions* (0 / 1), *Self-reported Health* (0 / 1) and *Mental Health* (0 / 1). In these models, a significant positive coefficient for *Self-employment* means that the self-employed are healthier than wage workers. 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 wage workers. We keep the regression sample constant across the three model specifications by including only the set of person-year observations eligible for inclusion in model specification three.

Regressions controlling for lagged health and time-invariant unobservables

Next, we perform longitudinal analyses to further investigate the association between self-employment and health. The cross-sectional results could be due to 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 unsuitable for inclusion in longitudinal analyses.[5] 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*.

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. Therefore, a further challenge is to establish whether the association between self-employment and health

[5] *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 completely recovered from a heart attack will always be seen in the data as having at least one health condition.

is the result of unobserved "third" factors that influence both. 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 even 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 in line 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.

Pooled regressions controlling for unobservables

The fixed-effect 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 five percent in every wave), resulting in a small and possibly non-random 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 using 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.[6] 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 (see 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.

5. Results

In Table 2, we present the regression results of the pooled logit models explaining *No Health Conditions* (*N* = 55,696), *Self-reported Health* (*N* = 55,691) and *Mental Health* (*N* = 45,627). The coefficients for *Self-employment* are all positive in the univariate regressions and significant for *Self-reported Health* and *Mental Health*. Odds ratios are respectively 1.06 for *No Health Conditions,* 1.12 for *Self-reported Health,* and 1.13 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, however, significant. For *Mental Health,* the coefficient also becomes smaller. In contrast to the result for *Self-reported Health*, the significance of the *Self-employment* coefficient disappears. Interestingly, the adjustment for demographics and employment characteristics renders the *Self-employment* coefficient significant in the *No Health Conditions* regression.[7] Altogether, we find that the self-employed are in better health than wage workers, although the difference in mental health is not statistically significant once demographic variables are controlled for.

We repeated the estimation of the model for *No Health Conditions* with demographic and employment controls to investigate which of the diseases underlying this variable are associated with self-employment. We find that from the 8 underlying diseases, arthritis (p = 0.003), and high blood pressure (p < 0.001) are significantly

[6] 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.

[7] This outcome is primarily the result of the inclusion of *Age* and *Age-squared* in the regression (coefficient for *Self-employment* is 0.23 with *p*-value < 0.001 if we only control for these two variables and lowers when more controls are included). The large effect of *Age* on *No Health Conditions* is explained by the fact that *No Health Conditions* is strictly increasing from 0 to 1 with age.

negatively associated; the self-employed have these conditions less often than wage workers. The other health conditions, cancer (p = 0.90), diabetes (p = 0.07), heart problems (p = 0.07), lunge disease (p = 0.08), psychiatric problems (p = 0.87), and stroke (p = 0.37) are not significantly associated at the five-percent level.

Table 2. Regression coefficients for *Self-employment* in the logit models explaining *No Health Conditions, Self-reported Health* and *Mental Health*.

	No Health Conditions	Self-reported Health	Mental Health
Pooled Logit (N)	*55,689*	*55,689*	*45,622*
Univariate regression	0.06 (0.04)	0.11*** (0.03)	0.12*** (0.03)
+ Demographic controls	0.19*** (0.04)	0.08* (0.04)	0.00 (0.03)
+ Employment controls	0.19*** (0.04)	0.07* (0.04)	0.01 (0.04)

Note. Standard errors in parentheses and clustered per individual; $^{*}p < 0.05$. $^{**}p < 0.01$. $^{***}p < 0.001$.

Panel 1 of Table 3 reports results of a logit regression with a lagged dependent variable to establish whether the cross-sectional results are the result of reverse causality between self-employment and health. We find qualitatively the same results as those presented in Table 2.[8] The coefficient for *Self-employment* is significant in the regressions for *Self-reported Health* (N = 39,599). For *Mental Health* (N = 31,485), only the univariate model shows this association. We conclude that for these two health measures, the association from Table 2 cannot be entirely due to reverse causality.

[8] We repeated the 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 2.

Table 3. 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, Lag Health) (N)	*39,599*	*31,485*
Univariate regression	0.08** (0.03)	0.08** (0.03)
+ Demographic controls	0.08* (0.04)	0.00 (0.03)
+ Employment controls	0.08* (0.03)	0.01 (0.03)
Fixed Effect Logit (N)	*28,732*	*27,975*
Univariate regression	0.03 (0.07)	-0.11 (0.07)
+ Demographic controls	0.03 (0.07)	-0.11 (0.07)
+ Employment controls	0.03 (0.07)	-0.12 (0.07)
Bivariate Probit (N)	*55,689*	*45,622*
$\rho = 0.00$	0.05*** (0.01)	0.01 (0.02)
$\rho = 0.10$	-0.11*** (0.01)	-0.15*** (0.01)
$\rho = 0.20$	-0.27*** (0.01)	-0.31*** (0.01)
Equal selection	-0.07*** (0.01)	-0.24*** (0.01)

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

The results of the fixed-effects panel regressions are in panel 2 of Table 3. Note that the sample size is substantially smaller here because in the fixed-effects logit regression, the individuals without a change in the dependent variable are dropped.[9] The associations for *Self-reported Health* (*N* = 28,732) and *Mental Health* (*N* = 27,975) 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.[10] 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 2 and statistically significant at the five-percent level up to

[9] A fixed-effect OLS regression that includes all person-year observations gives qualitatively the same results. Moreover, the regression results from Table 2 remain qualitatively the same if the regression analyses are restricted to the person-year observations that are included in the fixed-effect logit regressions.

[10] We investigated whether health-related attrition out of work was different between the self-employed and wage workers. 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 five-percent 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 wage workers.

the second (*Self-reported Health*) and fifth 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.

Using the methodology of Altonji et al. (2005), we investigate to what extent selection of unobserved variables is responsible for the cross-sectional association presented in Table 2. The bivariate probit results are given in the bottom panel of Table 3. 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 2 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 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 a "worst case" scenario, where the selection on observable variables is assumed to be similar to the selection on unobservable variables. Under this scenario of equal selection, the parameter ρ is estimated to be 0.07 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 'types' 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 3, 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.

Moreover, when the assumption is imposed that the selection on unobserved factors is as large as the selection on observed factors, the association even becomes negative, which suggests 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-effect 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, causal effect of self-employment on health is non-positive, possibly zero, and, if anything, negative.

6. Conclusion

It is notoriously difficult to discriminate between a contextual (benefit) effect of self-employment on health and health-related selection (barrier) 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 wage workers, 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 entirely 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 periods, females only, study on the relation between self-employment and health by Dolinksy 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. It has to be kept in mind that this conclusion is tentative and based upon relatively strong assumptions on the amount of selection on the basis of unobserved individual characteristics. Nonetheless, the results do raise some puzzling evidence: monetary

returns to self-employment have been shown to be lower than for wage work (Hamilton, 2000), while this paper suggests that presumed non-monetary benefits of self-employment in terms of health are in fact non-existent and potentially even negative. Further research on this emerging "self-employment puzzle" is clearly warranted.

Further research is also needed to identify the factors influencing both self-employment and health. Apart from traditional and more obvious correlations of 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.

For now, 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 due to 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 relatively healthy individuals self-select into self-employment. 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).

7. References

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

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

Audretsch DB, Thurik AR. What is new about the new economy: sources of growth in the managed and entrepreneurial economies, Industrial and Corporate Change 2001;10(1), 267-315.

Audretsch DB, Thurik AR. Capitalism and democracy in the 21st century: From the managed to the entrepreneurial economy. Journal of Evolutionary Economics 2000;10(1-2); 17-34.

Audretsch DB, Keilbach M. Entrepreneurship capital and economic performance. Regional Studies 2004;38(8); 949-959.

Beck T, Demirguc-Kunt A. Small and medium-size enterprises: Access to finance as a growth constraint. Journal of Banking & Finance 2006;30(11), 2931–2943.

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

Blanchflower, D.G. Self-employment in OECD countries. Labour Economics 2000;7(5); 471-505.

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

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

Bjuggren CM, Johansson D, Stenkula M. Using self-employment as proxy for entrepreneurship: some empirical caveats. International Journal of Entrepreneurship and Small Business 2012;17(3); 290-303.

Buttner EH. Entrepreneurial stress: is it hazardous to your health? Journal of Managerial Issues 1992;4(2); 223-240.

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

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

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

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

Dolinsky AL, Caputo RK. Health and female self-employment. Journal of Small Business Management 2003;41(3); 233-241.

Ekelund J, Johansson E, Järvelin MR, Lichtermann D. Self-employment and risk aversion—evidence from psychological test data. Labour Economics 2005;12(5); 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. Some empirical aspects of entrepreneurship. American Economic Review 1989;79; 519-535.

Forman-Hoffman VL, Richardson KK, Yankey JW, Hillis SL, Wallace RB, Wolinsky FD. Retirement and weight changes among men and women in the Health and Retirement Study. Journal of Gerontology 2008;638(3); S146-S153.

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

Gielnik MM, Zacher H, Frese M. Focus on opportunities as a mediator of the relationship between business owners' age and venture growth. Journal of Business Venturing 2012;27; 127–142.
Gilbert BA, Audretsch DB, McDougall PP. The emergence of entrepreneurship policy. Small Business Economics 2004;22; 313-323.
Gwatkin DR. Health inequalities and the health of the poor: What do we know? What can we do? Bulletin of the World Health Organization 2000;78(1).
Granger CWJ. Investigating causal relations by econometric models and cross-spectral methods. Econometrica 1969;39; 424-428.
Hamilton, BH. Does entrepreneurship pay? An empirical analysis of the returns to self-employment, Journal of Political Economy 2000;108; 604-631.
Hébert RF, Link AN. In search of the meaning of entrepreneurship. Small Business Economics 1989;1, 39–49.
Jamal M. Job stress, satisfaction, and mental health: An empirical examination of selfemployed and non-self-employed Canadians. Journal of Small Business Management 1997;35; 48–57.
Juster TF, Suzman R. An overview of the Health and Retirement Study. The Journal of Human Resources 1995;20; 7-56.
Karasek RA, Theorell T. Healthy work: Stress, productivity, and the reconstruction of working life. Basic Books: New York; 1990.
Karasek RA. Job demands, job decision latitude, and mental strain: Implications for job redesign. Administrative Science Quarterly 1979;24; 285–308.
Kirzner IM. Entrepreneurial discovery and the competitive market process: An Austrian approach, Journal of Economic Literature 1997;35; 60-85.
Klapper L, Amit R, Guillen MF, Quesada JM. Entrepreneurship and firm formation across countries. World Bank Policy Research Working Paper 4313; 2007.
Koellinger PD, Thurik AR. Entrepreneurship and the business cycle. Review of Economics and Statistics 2012;94(4); 1143-1156.
Lewin-Epstein N, Yuchtman-Yaar E. Health risks of self-employment. Work and Occupations 1991; 18; 291-312.
Lleras-Muney, A. The relationship between education and adult mortality in the United States. Review of Economic Studies 2005; 72(1); 189-221.
Nicolaou N, Shane S, Cherkas L, Hunkin J, Spector TD. Is the tendency to engage in entrepreneurship genetic? Management Science 2008;54; 167–179.
Parker SC. The economics of entrepreneurship, Cambridge University Press: Cambridge; 2009.
Parslow RA, Jorm AF, Christensen H, Rogers B, Strazdins L, D'Souza RM. The association between work stress and mental health: A comparison of organizationally employed and self-employed individuals. Work and Stress 2004;18(3); 231-244.
Prottas DJ, Thompson CA. Stress, satisfaction, and the work–family interface: A comparison of self-employed business owners, independents, and organizational employees. Journal of Occupational Health Psychology 2006;11(4); 366–378.
Quinn JF. Labor-force participation patterns of older self-employed workers. Social Security Bulletin 1980;43; 17-28.
Ravesteijn, B., Van Kippersluis, H. Van Doorslaer, E. The Contribution of Occupation to Health Inequality. In: Research on Economic Inequality, Volume 21: Health and Inequality, Pedro Rosa Dias and Owen O'Donnell (volume eds.), Emerald Group Publishing.
Schumpeter JA. The Theory of Economic Development, Harvard University Press: Cambridge; 1934.
Shane S, Venkataraman S. 2000. The promise of entrepreneurship as a field of research. Academy of Management Review 2000;26(1); 217–226.
Stephan U, Roesler U. Comparison of entrepreneurs' and employee's health in a national representative sample. Journal of Occupational and Organizational Psychology 2010;83(3), 717-738.

Stowasser T, Heiss F, McFadden D, Winter J. Healthy, wealthy and wise? Revisited: An analysis of the causal pathways from socio-economic status to health. Working Paper 17273, National Bureau of Economic Research (NBER) 2011.

Tetrick LE, Slack K, Da Silva N, Sinclair RR. 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 2000;5; 464–476.

Theorell T, Karasek RA. Current issues relating to psychosocial job strain and cardiovascular disease research. Journal of Occupational Health Psychology 1996;1; 9–26.

Torres O. La Santé du Dirigeant: de la Souffrance Patronage à l'Entrepreneuriat Salutaire, De Boeck: Brussels; 2012.

Van der Loos MJHM, Rietveld CA, Eklund N, Koellinger PD, Rivadeneira F, Abecasis GR, … Thurik AR. The molecular genetic architecture of self-employment. PLOS ONE 2013;8(4), e60542.

Van Praag, CM, van Ophem H. Determinants of willingness and opportunity to start as an entrepreneur. Kyklos 1995;48; 513-540.

Verheul I, Thurik AR, Hessels J, Van der Zwan PW. Factors influencing the entrepreneurial engagement of opportunity and necessity entrepreneurs. Research Reports EIM Zoetermeer. 2010.

Zissimopoulos J, Karoly LA. Transitions to self-employment at older ages: The role of wealth, health, health insurance, and other factors. Labour Economics 2007;14; 269-295.