I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11305

The Effect of Self-Employment on Health: Evidence from Longitudinal Social Security Data

Judite Gonçalves
Pedro S. Martins

JANUARY 2018

DISCUSSION PAPER SERIES

IZA DP No. 11305

The Effect of Self-Employment on Health: Evidence from Longitudinal Social Security Data

Judite Gonçalves
Queen Mary University of London and Nova SBE

Pedro S. Martins
Queen Mary University of London, Nova SBE and IZA

JANUARY 2018

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Effect of Self-Employment on Health: Evidence from Longitudinal Social Security Data

The growth of novel flexible work formats raises a number of questions about their effects upon health and the potential required changes in public policy. However, answering these questions is hampered by lack of suitable data. This is the first paper that draws on comprehensive longitudinal administrative data to examine the impact of self-employment in terms of health. It also considers an objective measure of health – hospital admissions – that is not subject to recall or other biases that may affect previous studies. Our findings, based on a representative sample of over 100,000 individuals followed monthly from 2005 to 2011 in Portugal, indicate that the likelihood of hospital admission of self-employed individuals is about half that of wage workers. This finding holds even when accounting for a potential self-selection of the healthy into self-employment. Similar results are found for mortality rates.

JEL Classification: I18, J24

Keywords: self-employment, hospitalization, sick leave, mortality

Corresponding author:
Pedro S. Martins
School of Business and Management
Queen Mary University of London
Mile End Road
London E1 4NS
United Kingdom

E-mail: p.martins@qmul.ac.uk

1. Introduction

Many governments pay close attention to the self-employed because of the positive association between self-employment and economic growth (e.g. Koellinger and Thurik, 2012). Currently, the self-employed already represent a non-negligible 14% of the labor force in European countries (Eurostat, 2017). In addition, the ongoing growth of the 'gig' (or 'platform') economy, caracterised by more flexible work formats, distinct from formal employer-employee relationships framed by employment law, raises a number of questions about public policy (e.g. Garben 2017, European Commission 2017). One important issue concerns the impact of the novel jobs created by this economy —as well as all other, more established forms of self-employment— upon workers' health. Indeed, key occupational characteristics, such as job control and job demand, vary significantly between self-employment and wage work. At the same time, self-employment jobs are characterized in most countries by little or no protection from public policies or regulations, in terms of social security, employment law or collective bargaining. This highly flexible context, where workers benefit from little or no insurance and are potentially subject to considerable volatility in work levels, can have additional implications for health.

Job control stands for decision authority, e.g. the freedom to decide what work to do, when and at what pace, which will reduce work-related stress. Job demand, on the other hand, represents sources of stress at work, such as receiving a lot of work and or being subject to little time to carry out specific tasks. This Job Demand-Job Control theory, originally proposed by Karasek (1979) —see also Karasek and Theorell (1990) and Theorell and Karasek (1996) — suggests that, compared to wage work, self-employment is associated with both higher job control and higher job demand (e.g. Prottas and Thompson, 2006; Stephan and Roesler, 2010). In fact, self-employed individuals are not subject to orders from others higher up the organizational hierarchy, so they have more decision authority.

However, their income and assets directly hinge on their ability to work and work effort, with greater exposure to unantecipated demand shocks, leaving them subject to more volatile workload and income flows, which represents a source of stress. Work-related stress, in turn, impacts negatively on health and well-being in general and may increase incidence of disease (e.g. depression, cardiovascular problems), absence from work due to sickness, use of health care services, and mortality. Stress is also associated with unhealthy behavior such as smoking and drinking, which may also be detrimental for health. Given these two opposite mechanisms —higher job demand and higher job control—, whether self-employment has a positive or detrimental effect on health is a public policy question that can only be fully addressed through empirical evidence of a causal nature.

There are two main empirical challenges to the identification of the effect of self-employment on health: reverse causality and individual unobserved heterogeneity. Reverse causality has to do with the possibility that individuals become self-employed or wage workers at least partly for health-related reasons. On the one hand, self-employment may attract individuals that are healthier on average because healthier individuals tend to be more able to focus on business opportunities or may have easier access to financing (e.g. Gielnik et al., 2012). Additional reasons are that income when self-employed tends to be more closely linked to one's ability to work than when a wage worker, and that access to sickness benefits is harder for the self-employed. All these factors suggest a positive (self-)selection of the healthy into self-employment.

On the other hand, individuals with health problems may have greater difficulties in finding a wage job, particularly if those health issues are visible to the employer, which may then push such individuals into self-employment (e.g. Zissimopoulos and Karoly, 2007). Several individual traits that are difficult to measure may be related to both health and self-employment decisions. Examples include perseverance, risk aversion, and genetics. Earlier life circumstances such as childhood health also influence adult health, employment, and socioeconomic status (Case et al., 2005; Case and Paxson, 2010). Taken together, these traits and earlier circumstances mean that self-employed individuals and wage workers may have different health profiles —the so-called individual unobserved heterogeneity.

The empirical literature on self-employment and health is growing but still scarce. Most of it is plagued by the endogeneity issues mentioned above, which are difficult to tackle without longitudinal data. A recent study finds significantly lower work-related stress among self-employed individuals without employees compared to wage workers, using longitudinal data from Australia and controlling for individual fixed-effects (Hessels et al., 2017). Previous studies on self-employment and stress provide contradictory findings, but most of them are based on cross-sectional data and use descriptive methods (see Hessels et al., 2017, Table 1, for a review). In a recent study, self-employed individuals appear healthier than wage workers (Rietveld et al., 2015). While the positive association between self-employment and health holds when the authors control for reverse causality, it vanishes when they control for individual unobserved heterogeneity. This finding suggests a positive selection of the healthy into self-employment. That study considers subjective health measures, including self-reported number of conditions, overall health, and mental health. It uses longitudinal survey data representative of the population 50+ in the US. The results may therefore not be generalizable for a broader working-age population, in particular for younger workers that may be overrepresented amongst 'gig economy' jobs.

Another study by Yoon and Bernell (2013) relies on cross-sectional survey data representative of the adult population in the US and adopts an instrumental variable approach. The authors find that self-employment has a positive impact on several health indicators, namely the absence of chronic conditions such as hypertension and diabetes. They find no effects on other health outcomes, including perceived physical health and mental health. Regarding more objective indicators, a five-year follow-up study of the total working population in Sweden finds that self-employed individuals have lower average risk of mortality compared to wage workers, even when controlling for several potential confounders (Toivanen et al., 2016). Overall, there is little robust evidence on the effect of self-employment on health. Most of the literature does not take endogeneity into account, as longitudinal data or instrumental variables are seldom available.

In this paper we address some of the limitations of the existing research. We assess the impact of self-employment on the likelihood of hospitalization, based on a large sample of administrative social security records representative of the active population in Portugal. We observe more than 132,000 self-employed individuals and wage workers over a period of up to 84 months, between January 2005 and December 2011. Our contribution to the literature on the relationship between self-employment and health is three-fold. First, we exploit the longitudinal nature of our data to tackle specifically the endogeneity of the decision to become self-employed. Using individual fixed-effects models, we compare different health outcomes of the same person over time, in particular when the individual switches between self-employment and wage work.

Second, we look at an objective health measure —hospitalizations—, based on administrative records (see Dobkin et al. (Forthcoming) for a very recent study that also considers hospitalizations, even from a different perspective —that of its economic impacts). This outcome is therefore strictly comparable across individuals and time periods and not subject to recall bias, as may happen with self-reported indicators in survey data. We also investigate the effect of self-employment on mortality, another objective measure of health. Third, we consider the whole working population regardless of age, i.e. a sample not limited to older workers where self-employment itself may be less widespread and representative. We also explore if there are heterogeneous effects by gender, age, and nationality.

Also related to our study is research closer to Labour Economics that assesses the relationship between employment protection and absenteeism. The literature tends to find that stronger employment protection (against dismissal) leads to higher absenteeism. For example, Riphahn (2004) finds that when public sector workers in Germany reach tenure and become virtually 'un-dismissable', the annual number of days of (self-reported) sickness absence increases significantly. The author identifies the effect based on survey data and a difference-in-differences strategy, using private sector employees as the control group. Similarly, the average number of days of sick leave per week significantly increases immediately after the probation period ends for white-collar workers of a large Italian bank (Ichino and Riphahn, 2005). The authors interpret these findings as a form of moral hazard, whereby workers reduce effort when it becomes harder for their employer to fire them.

In contrast, using longitudinal survey data and an instrumental variable strategy, a recent study finds that perceived employment insecurity has negative consequences on mental health and general well-being, which in turn may contribute to more sickness absence episodes (Cottini and Ghinetti, 2017). Self-employment may be seen as lower protection employment, as it will typically not be subject to any form of employment law. We contribute to the literature on employment protection and absenteeism by comparing the effect of self-employment on the likelihood of standard sick leave to its effect on the likelihood of hospitalization. In contrast to standard sick leave, hospitalizations are not subject to moral hazard given their acute nature. Thus, the difference in the effects of self-employment on the two outcomes —standard sick leave and hospitalizations— can be regarded as evidence of the extent of moral hazard created by employment protection.

Our study is based on data from Portugal, where statutory sick leave covers both wage workers and self-employed individuals. As in many European countries, to deter moral hazard, wage workers face a three-day gap (i.e. waiting period) from the onset of the sickness episode until the social security benefit can be paid. For the self-employed, this waiting period is much longer, at thirty days. However, in cases where the sickness episode corresponds to a hospitalization, there is no waiting period for either wage workers or the self-employed. For nearly the entire period under analysis here (Sep 2005-Dec 2011), the replacement rate of the Portuguese sickness benefit was equal to 65% of forgone wages for the first 90 days of sick leave, 70% from the 91st to the 365th day, and 75% from the 365th day onwards.[1]

The remaining of this paper is as follows: in Section 2 we present the dataset used as well as our empirical strategy. In Section 3 we present our results, considering different outcomes and specifications as well as a number of robustness checks. Finally, in Section 4 we discuss our findings.

2. Methods

2.1. Data

We use data made available by the Social Security public agencies of Portugal (ISS and IISS). The dataset is a 1% random sample of the entire population which makes income-related payments to or receives benefits from the Social Security over the period 2005-2011. The Social Security records include salaries from (self-)employment, as well as sickness, unemployment, maternity, and other Social Security benefits (see Martins (2016) for more details about this dataset).

For the purposes of this study, we consider individuals on a monthly basis —i.e. for up to 84 periods, from January 2005 to December 2011— and use information on whether they are wage workers or self-employed, as well as whether they have a sickness episode that leads to a sickness benefit being paid in a specific month (both types of workers pay Social Security contributions and receive sickness and other benefits from Social Security). The data distinguish between standard sickness spells (implying absence from work) and more serious cases that correspond to hospitalizations, as the applicable regulation is different, as discussed before. The dataset includes 132,141 different individuals, for a total of 7,018,732 individual-month observations. The average

[1] For the first eight months of 2005, the replacement rate was 55% of forgone wages for the first 30 days of sick leave, and 60% from the 31st to the 90th day. Sickness benefits are granted for a maximum of 1,095 days for wage workers and 365 days for self-employed individuals (Decree-law 28/2004 of Feb 4, Decree-law 133/2012 of Jun 27, and Decree-law 146/2005 of Aug 26).

number of monthly observations of each individual is 67. Moreover, in each month, we observe on average 83,628 different individuals. Additional individual information available in the dataset includes gender, age, and nationality (Portuguese or foreign).

Descriptive statistics by type of employment are shown in Table 1. The self-employed account for 4.21% of the person-month observations, or almost 300,000 observations. Based on our representative data, we find that the average monthly incidence rate of hospitalization is 0.09% among the self-employed, compared to 0.17% —nearly twice as much— among wage workers. The self-employed also have lower monthly incidence of standard sick leave than wage workers. Here, the gap is even larger —0.32% compared to 2.13% (i.e. standard sick leave is nearly seven times more likely amongst wage workers). These differences are also clear from Figure 1, which presents the monthly incidence rates of hospitalization and standard sick leave over the entire period, for both the self-employed and wage workers. Table 1 also shows the average number of days of hospitalization and standard sick leave, the incidence rates of hospitalization and standard sick leave over the previous 12 months, as well as the average number of months over the previous 12 months with a hospitalization or standard sickness episode. We also find that the proportion of women is slightly lower among the self-employed than among wage workers (49% versus 53%), that the self-employed are on average older (43 versus 37 years old), and that the proportion of foreigners is also sightly lower among the self-employed (10% versus 12%).

2.2. Identification and empirical strategy

This study focuses on the effect of self-employment on the likelihood of hospitalization. Our main results relate to hospitalizations rather than standard —i.e. non-hospitalization— sickness episodes for two main reasons. First, in Portugal as in other countries, self-employed individuals face a much longer (thirty-day) waiting period before they can receive (non-hospitalization) sickness benefits, whereas wage workers wait only three days. This means that in the data, we observe sickness episodes that last at least four days for wage workers, but only sickness episodes that last at least thirty-one days for the self-employed (of which the first 30 days are not eligible for sickness benefits). Thus, all things equal, on average the sickness events of the self-employed that are registered in the dataset are much more selected —and severe. Furthermore, the different waiting periods may entail different incentives for wage workers and self-employed individuals. That is, wage workers may engage more often in moral hazard (i.e. 'cheat' by going on sick leave when they are not really sick), as they face much lower opportunity costs —fewer days without income.[2] In striking contrast, there is no waiting period for either wage workers or self-employed individuals in the case of hospitalization. The second reason for our focus on hospitalizations is that due to their specific, extreme nature, they are less likely to be timed deliberately by individuals and therefore less likely to be artificial episodes of sickness. Thus, hospitalization is a significantly more objective

[2]In some cases, collective bargaining provisions may lead to the payment of the first three days of absence too.

measure of health, and hospitalization events should be strictly comparable between wage workers and self-employed individuals.

To determine the effect of self-employment on the likelihood of hospitalization, we estimate several specifications of a linear probability model, as in the equation below:

$$Pr[\text{hosp}_{i,t} = 1] = \alpha + \delta_0 \text{self-employed}_{i,t} + \delta_1 \text{self-employed}_{i,t-1} + \gamma \text{hosp12}_{i,t} + \beta X_{i,t} + \tau_t + \mu_i + \varepsilon_{i,t} \quad (1)$$

The dependent variable, $\text{hosp}_{i,t}$, is binary and indicates whether individual i is hospitalized in month t or not. The first, most parsimonious specification includes only an intercept, α, and the self-employment indicator, $\text{self-employed}_{i,t}$, which takes value one if individual i is self-employed in month t and zero if he or she is a wage worker (Model 1). $\varepsilon_{i,t}$ denotes the usual error term. In this simple specification, the coefficient δ_0 gives the unadjusted average difference in the incidence rate of hospitalization between wage workers and the self-employed.

Next, we add a lag of self-employment to the model, $\text{self-employed}_{i,t-1}$ (Model 2). In this specification, δ_1 indicates the association between self-employment and the likelihood of hospitalization in the following period, taking into account a potentially delayed effect.[3] The third specification controls for observed individual characteristics —gender, age, and nationality—, included in the vector $X_{i,t}$, as well as time (month and year) fixed-effects, denoted by τ_t (Model 3).

As mentioned in the Introduction, to identify the causal effect of self-employment on the likelihood of hospitalization, we must rule out reverse causality and unobserved individual heterogeneity. Thus, in the fourth specification, we control also for individual time-invariant heterogeneity through individual fixed effects, denoted μ_i (Model 4). This is an important specification, as it allows us to compare the likelihood of hospitalization for the same individual over the up to 84 month-long period covered by our data. Identification in this specification comes from individuals that are observed under the two types of employment over that period, i.e. switchers between self-employment and wage work. Lastly, to rule out reverse causality, we include in the fifth and last specification a binary indicator, $\text{hosp12}_{i,t}$, which takes value one if individual i was hospitalized at least once in the previous 12 months. In sum, Models 1-3 are pooled OLS models, and Models 4 and 5 are fixed-effects panel data models.

We also estimate the same model specifications to assess the effects of self-employment on the duration of hospitalizations and likelihood of standard sick leave. In the case of duration of hospitalizations, the analysis is based on individual-month observations with a hospitalization, a much smaller sample. The natural logarithm of the number of hospitalization days is used as the dependent variable, to account for non-linearities and limit the influence of outliers.

To assess whether the effects of self-employment on the likelihood of hospitalization differ

[3] We also estimated finite distributed lag models of higher orders but found that additional lags of the self-employment indicator were not statistically significant.

according to gender, age, and nationality, we estimate separate regressions for men and women, under and over 45 year olds, natives and foreigns. Furthermore, we investigate whether the effects are symmetric according to whether individuals move to or from self-employment by interacting contemporaneous and lagged self-employment with indicators of type of transition (wage worker to self-employed or vice-versa).

Finally, two methodological notes are in order. First, the linear probability model was preferred to the logit model due to computational efficiency, statistical power, and straightforward interpretation of the coefficients.[4] Both models produce identical estimated effects of self-employment on the likelihood of hospitalization. Moreover, the linear probability model does not produce out-of-interval predictions. Second, we considered dynamic panel data models as a way to deal with reverse causality. However, such models proved unfeasible as hospitalization is a relatively rare event and, consequently, the coefficients associated with lags of the dependent variable were not statistically significant in any specification attempted.

2.3. Analysis of mortality rates

We also investigate the effect of self-employment on mortality. To do this, we aggregated our data to construct a new dataset of person-year observations, as we only know the year in which the individual passes away (and not the year and the month, as in the case of hospitalizations). We create a binary dependent variable that takes value one if individual i passes in year $t+1$. The self-employment indicator takes value one if the individual is self-employed during the whole year t. The model specifications are the same ones as described above, adjusted for the annual frequency considered here. In the most complete specification, we include the self-employment dummy, age, time fixed effects, individual fixed effects to control for time-invariant individual characteristics, and the indicator of any hospitalization to (partly) control for time-variant individual health. All explanatory variables are measured in year t. We exclude observations for the year in which the person passes, as we don't know when in the year that event takes place. Moreover, we consider the sensitivity of our results to a definition of self-employed individuals in year t as those who were self-employed during at least 7 months of the year.

3. Results

3.1. Effects on hospitalization

Table 2 shows the effect of self-employment on the likelihood of hospitalization. The unadjusted average monthly difference in the likelihood of hospitalization between wage workers and self-employed individuals is -0.00076 ($p < 0.01$; Model 1). In relative terms, this difference is

[4]The fixed-effects logit model has lower statistical power because it excludes individuals that do not change employment status.

considerable: the self-employed are 46% less likely than wage workers to have a hospitalization. When we include lagged self-employment in the model, we see that the association is actually lagged one period: being self-employed is associated with a 0.086 percentage point (51%) lower likelihood of hospitalization in the following month ($p < 0.05$; Model 2). Controlling for gender, age, nationality, and time fixed effects actually makes the association between self-employment and likelihood of hospitalization even more negative (Model 3). These results also indicate that women, older workers, and natives exhibit higher rates of hospitalization. Moreover, the estimated coefficient on lagged self-employment in the model with individual fixed effects is -0.00079 ($p < 0.05$; Model 4). This suggests that individual time-invariant unobserved characteristics associated with both self-employment and likelihood of hospitalization are not driving our results. Lastly, controlling for any hospitalization in the previous 12 months has little impact on the magnitude of the coefficient, although it becomes less significant ($p < 0.1$; Model 5). This is possibly a result of the smaller sample size, which decreases by more than 2 million observations, as the first 12 months of data for each individual cannot be used. In the end, controlling for observable and unobservable time-invariant individual characteristics, time fixed effects, and reverse causality leaves the estimated coefficient on lagged self-employment almost unchanged: -0.00092 compared to -0.00086 (Model 5 versus Model 2). This indicates that self-employment has a positive impact on health: the positive association between the two variables is not (only) a matter of the healthy self-selecting into self-employment.

Regarding heterogeneous effects, we find that the negative effect of self-employment on the likelihood of hospitalization is larger among women and foreigners, while it does not differ much between individuals under and above the age of 45 (Table 3). We also find heterogeneous effects depending on whether individuals switch over time from wage work to self-employment or from self-employment to wage work. Table 4 presents the results. The difference between Specifications 1 and 2 is that Specification 1 includes triple interactions between (lagged) self-employment, a dummy that identifies individuals who switch to self-employment, and a dummy that identifies individuals who switch from self-employment. Thus, Specification 1 allows us to distinguish between the effects of self-employment among those who only switch once during the whole period from wage work to self-employment, those who switch once from self-employment to wage work, and those who transition to-and-fro. Table 5 presents the number of observations in each of these groups. Specification 2 only distinguishes between the effect of switching to/from self-employment —whether it happens once or more than once. The negative effect of self-employment on the likelihood of hospitalization found above (Table 2) comes essentially from individuals who switch from self-employment to wage work. There is not strong evidence of a significant effect among individuals who switch from wage work to self-employment, or among individuals who switch to-and-fro, although the sample sizes for these two groups are smaller.

We also find some evidence of a negative effect of self-employment on the duration of hospitalization when controlling for individual fixed-effects (Table 6, Models 4-5). We refrain from making definitive conclusions based on this result because it is identified based on few individuals who switch between self-employment and wage work and have at least one hospitalization under each employment situation.

3.2. Effects on standard sick leave

Table 7 presents the effect of self-employment on the likelihood of standard sick leave. On average, the self-employed have a 1.813 percentage point (85%) lower likelihood of sick leave than wage workers ($p < 0.01$; Model 1). Adding the estimated coefficients on self-employment and lagged self-employment in Model 2 gives the same average difference. The coefficients remain practically the same when controlling for gender, age, nationality, and time fixed effects (Model 3). When we control for individual time-invariant characteristics and any sick leave in the previous 12 months, the coefficients become smaller in absolute terms, but remain significantly negative ($p < 0.01$; Model 5). Adding the two coefficients on self-employment and its lag gives a difference in the likelihood of sick leave between the self-employed and wage workers of 1.359 percentage points —i.e. the self-employed are 66% less likely to have a sick leave episode, a sizable difference. Thus, as happens for hospitalizations, selection of the healthy into self-employment doesn't (fully) explain the negative association between self-employment and likelihood of sick leave.

However, as explained in Section 2.2, we cannot say that self-employment is good for health based on this result. The association between self-employment and likelihood of sick leave estimated in Model 5 may be partly due to the different waiting periods that self-employed individuals and wage workers face until they can claim sickness benefits. Indeed, the different waiting periods imply that one cannot observe sick leave spells of self-employed individuals that last for less than 31 days. Moreover, it is likely that there is a higher proportion of fraudulent sickness spells among wage workers. These facts are consistent with the negative association found. In sum, given the institutional context and data available, one cannot fully disentangle the effect of self-employment on the likelihood of sick leave that is due to health differences between the self-employed and wage workers from the confounding effect of the heterogeneous waiting periods.

3.3. Effects on mortality

The analysis of mortality rates is based on a necessarily smaller sample than was available for the previous analyses that were based on person-month observations. Still, the sample has nearly 700,000 person-year observations. Table 8 presents the effect of self-employment on the likelihood of mortality in the following year. We find that the unadjusted average difference in the likelihood of mortality between wage workers and self-employed individuals is statistically zero ($p > 0.1$; Model 1). However, when adjusted for age, gender, nationality, and time fixed effects, that difference becomes statistically significant and equal to -0.00077 —i.e. self-employed individuals are 0.077 percentage points (73%) less likely to die in the following year than wage workers ($p < 0.01$; Model

2). We also find evidence that mortality is lower for women, not statistically different between native and foreign workers, and (unsurprisingly) increases with age. When controlling only for individual fixed effects, the association between self-employment and likelihood of mortality is even more negative and still statistically significant ($p < 0.1$; Model 3). When more controls are included, the estimated coefficient on self-employment remains about the same magnitude, but loses its statistical significance. Although we do not find conclusive evidence of a negative effect of self-employment on mortality, we regard these results as consistent with our main findings for hospitalizations. They are not as statistically precise, most likely because of the lower relative frequency of deaths and the consequently smaller sample used.

3.4. Robustness checks

As a robustness check, we repeat our main analysis of the effect of self-employment on the likelihood of hospitalization on data aggregated from the person-month to the person-quarter level. The results of this analysis are reported in Table 9. We find quantitatively the same effects as reported in Table 2 —i.e. coefficients are roughly multiplied by three, as expected. For example, the likelihood of hospitalization is 0.214 percentage points (46%, as with monthly observations) lower among self-employed individuals than among wage workers ($p < 0.01$; Model 1).

We also find that the effect of self-employment on mortality is robust to an alternative definition of self-employment: being self-employed during at least seven months of the year instead of during the entire year. The estimated coefficients are slightly smaller in absolute terms, as may be expected, but remain negative and equally significant (results available upon request).

4. Discussion and conclusion

It is probably as challenging as it is important to determine whether self-employment is good or detrimental for health. Indeed, the potential self-selection of the healthy into or out of self-employment is difficult to rule out empirically. However, separating the effect of self-employment on health from that selection effect is crucial to inform public policy decisions. Moreover, as the 'gig economy' grows around the world, causal evidence about its health implications become more pressing. In this study, we provide causal estimates of the impact of self-employment on health by taking advantage of a large longitudinal sample of Social Security records representative of the working population in Portugal. Moreover, we focus on the specific and original dimension of hospitalizations, which can be particularly insightful in this debate.

We find that self-employed individuals have about 50% lower likelihood of hospitalization than wage workers, an effect that remains when controlling for the endogeneity of self-employment decisions. This is in contrast with the results of Rietveld et al. (2015), who find a negative association between self-employment and health that is fully explained by a selection effect. The different results between the two studies may be due to the type of health measures and samples used. While we focus on administrative records of hospitalizations and consider the whole working population, Rietveld et al. (2015) draw on survey-based subjective health measures and focus on the 50+ population. On the other hand, our results are consistent with those of Yoon and Bernell (2013), who find a positive impact of self-employment on more objective health measures, namely absence of specific health conditions.

We also find that the likelihood of standard sick leave is lower among self-employed individuals than among wage workers: 85% lower when unadjusted; 66% lower when controlling for reverse causality and individual heterogeneity. That is, we find some evidence of a selection effect in this case, but not enough to explain entirely the negative association between self-employment and likelihood of sick leave. However, we cannot attribute the remaining association (entirely) to health benefits of self-employment, because we have a confounding factor —different waiting periods for self-employed individuals and wage workers. Lastly, we also find evidence of a negative effect of self-employment on mortality, consistently with our remaining results and with the limited previous literature on this topic (Toivanen et al., 2016).

Our results also contribute to an additional literature that finds that employment protection legislation can lead to absenteeism. This follows from the fact that self-employment involves little or no legal protection against dismissals or any form of risk sharing with employees (or clients). For instance, we find a negative relationship between self-employment and likelihood of sick leave, which is consistent with a moral hazard interpretation, at least if disregarding for a moment the legal differences in waiting times before Social Security protection is made available. However, such negative relationship still exists when looking only at sick leave episodes that correspond to hospitalizations, which should not be subject to moral hazard and where the legal differences above do not apply. This suggests that the relationship between employment protection and absenteeism may not be entirely explained by moral hazard as implied before (e.g. Riphahn, 2004). On the other hand, our results are in contrast to the study by Cottini and Ghinetti (2017), who use an instrumental variable approach and find that perceived employment insecurity has negative consequences upon mental health and general well-being.

The results of this study indicate that the beneficial effects of higher job control when self-employed exceed the detrimental effects of higher job demand. Our results may also reflect changes in the type of work *per se* when individuals switch to/from self-employment. Unfortunately, we cannot explore this issue with the available data, as we do not know the type of job/industry of self-employed individuals. Nevertheless, it seems plausible that in most cases individuals do not switch to drastically different jobs when switching between self-employment and wage work —e.g. from blue-collar to white-collar jobs. Even if individuals change industry when switching to/from self-employment, most often they are likely to remain in relatively similar occupations. Still, one's inability to investigate the potentially different effects of self-employment by industry remains a limitation of this study, driven by the lack of data on the occupations of the self-employed. For instance, manufacturing workers —typically wage workers— may be more prone to injuries at work, which may partly drive our results. Regarding differences by gender, it is important to note that hospitalizations related with risky pregnancies and childbirth are not included in the dataset. Therefore, potential differences in fertility rates between women who are self-employed or wage workers do not drive our results.

In conclusion, this study provides causal evidence of a positive effect of self-employment on health. Such impact may be at least partly explained by greater control by the individual over different aspects of the working life associated with this type of employment. One important dimension of the ongoing debate about the 'future of work' is precisely how to increase protection for workers under flexible contracts, such as those that increasingly emerge under the 'gig economy' (e.g. Garben 2017, European Commission 2017). This may involve multiple policy dimensions such as Social Security, employment law and collective bargaining. Our results indicate that the current concerns about the lack of protection of the self-employed may be somewhat exaggerated, at least as far as health is concerned.

Acknowledgements

We thank financial support from the European Union (grant VS/2016/0340), data access by Social Security (Portugal), and comments from Elena Cottini, Kevin Denny, Rudolf Winter-Ebmer, and participants at seminar at Nova School of Business and Economics (Nova Healthcare Initiative).

References

Case, A., Fertig, A., Paxson, C. 2005. The lasting impact of childhood health and circumstance. Journal of Health Economics, 24, 365–389.

Case, A., Paxson, C. 2010. Causes and consequences of early-life health. Demography, 47, S65–S85, URL: http://link.springer.com/10.1353/dem.2010.0007.

Cottini, E., Ghinetti, P. 2017. Employment insecurity and employees' health in Denmark. Health Economics, 1–14.

Dobkin, C., Finkelstein, A., Kluender, R., Notowidigdo, M. The economic consequences of hospital admissions. American Economic Review (Forthcoming), URL: http://www.nber.org/papers/w22288.pdf.

European Commission. 2017. Proposal for a directive of the European Parliament and of the Council on transparent and predictable working conditions in the European Union. URL: http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52017PC0797.

Eurostat. 2017. Employment by sex, age, and professional status. URL: http://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do.

Garben, S. 2017. Protecting workers in the online platform economy: an overview of regulatory and policy developments in the EU.Technical report, European Agency for Safety and Health at Work, Luxemburg.

Gielnik, M. M., 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, URL: `http://dx.doi.org/10.1016/j.jbusvent.2010.05.002`.

Hessels, J., Rietveld, C. A., Zwan, P. V. D. 2017. Self-employment and work-related stress: the mediating role of job control and job demand. Journal of Business Venturing, 32, 178–196, URL: `http://dx.doi.org/10.1016/j.jbusvent.2016.10.007`.

Ichino, A., Riphahn, R. T. 2005. The effect of employment protection on worker effort: absenteeism during and after probation. Journal of the European Economic Association, 3, 120–143.

Karasek, R. A. 1979. Job demands, job decision latitude, and mental strain: implications for job redesign. Administrative Science Quarterly, 24, 285–308.

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

Koellinger, P. D., Thurik, A. R. 2012. Entrepreneurship and the business cycle. The Review of Economics and Statistics, 94, 1143–1156.

Martins, P. S. 2016. Working to get fired? Regression discontinuity effects of unemployment benefit eligibility on prior employment duration. IZA Discussion Paper Series, 10262.

Prottas, D. J., Thompson, C. A. 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.

Rietveld, C. A., Kippersluis, H., Thurik, A. R. 2015. Self-employment and health: barriers or benefits?. Health Economics, 24, 1302–1313.

Riphahn, R. T. 2004. Employment protection and effort among German employees. Economic Letters, 85, 353–357.

Stephan, U., Roesler, U. 2010. Health of entrepreneurs versus employees in a national representative sample. Journal of Occupational and Organizational Psychology, 83, 717–738.

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

Toivanen, S., Griep, R. H., Mellner, C., Vinberg, S., Eloranta, S. 2016. Mortality differences between self-employed and paid employees: a 5-year follow-up study of the working population in Sweden. Occupational and Environmental Medicine, 73, 627–636.

Yoon, J., Bernell, S. L. 2013. The effect of self-employment on health, access to care, and health behavior. Health, 05, 2116–2127, URL: http://www.scirp.org/journal/PaperInformation.aspx?PaperID=41092{&}{#}abstract.

Zissimopoulos, J. M., Karoly, L. A. 2007. Transitions to self-employment at older ages: the role of wealth, health, health insurance and other factors. Labour Economics, 14, 269–295.

Figures and Tables

Figure 1: Probability of hospitalization or standard sickness episode by type of employment, Jan 2005-Dec 2011

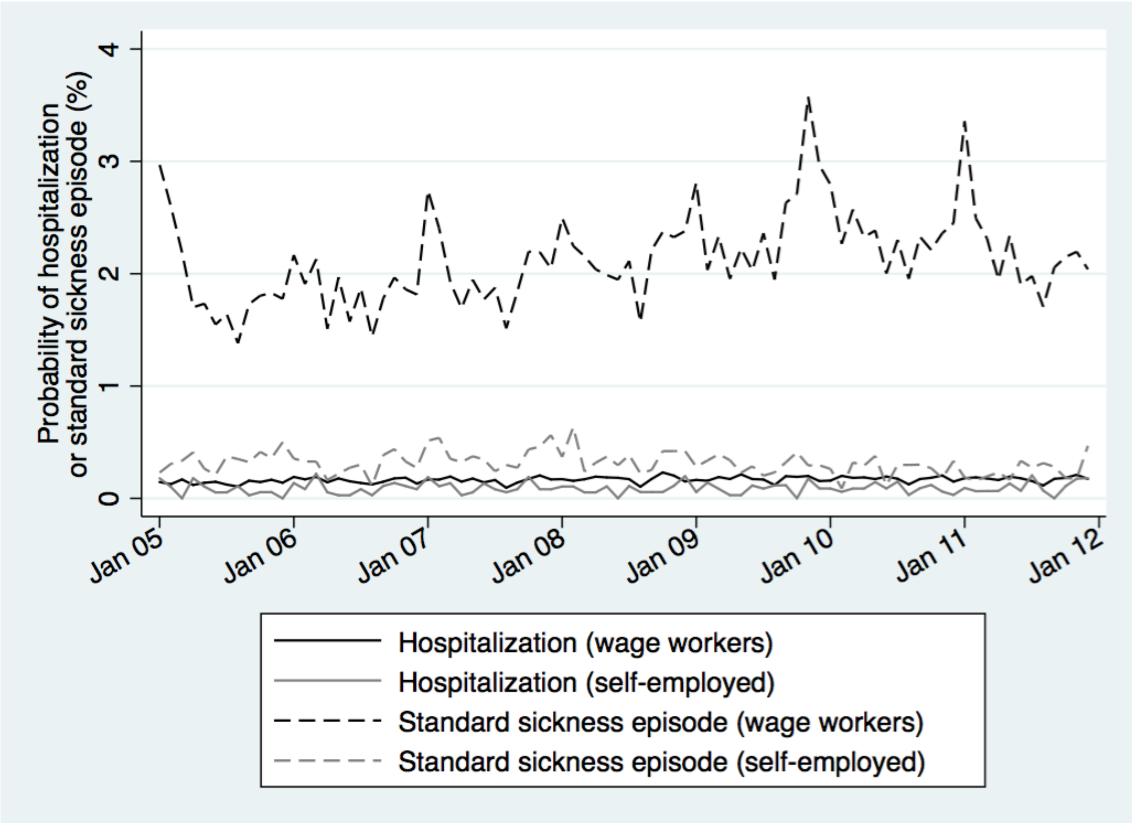

Table 1: Descriptive statistics by type of employment

	Self-employed	Wage workers	Total
Any hospitalization in the following month	0.09%	0.17%	0.16%
Average number of days of hospitalization[1]	9.30	2.27	2.32
	(28.57)	(14.85)	(15.01)
	[0; 358]	[0; 1,095]	[0; 1,095]
Any hospitalization in the previous 12 months	0.69%	1.56%	1.52%
Months with any hospitalization in the previous 12[1]	0.01	0.02	0.02
	(0.08)	(0.13)	(0.13)
	[0; 2]	[0; 10]	[0; 10]
Any standard sickness episode in the following month	0.32%	2.13%	2.05%
Average number of days of sick leave[1]	51.28	20.73	20.97
	(69.73)	(46.58)	(46.88)
	[0; 564]	[0; 1,259]	[0; 1,259]
Any standard sick leave in the previous 12 months	2.19%	16.40%	15.85%
Months with any standard sick leave in the previous 12[1]	0.02	0.21	0.20
	(0.17)	(0.55)	(0.54)
	[0; 5]	[0; 12]	[0; 12]
Female	48.83%	53.15%	52.97%
Age[1]	42.80	37.09	37.33
	(11.24)	(10.56)	(10.65)
	[18; 70]	[18; 70]	[18; 70]
Foreign	10.03%	12.06%	11.97%
Observations	295,819	6,722,913	7,018,732
	4.21%	95.79%	100.00%

[1]Continuous variable with standard deviation in parentheses and minimum and maximum values in brackets.

Table 2: Effect of self-employment on the likelihood of hospitalization

	Model 1	Model 2	Model 3	Model 4	Model 5
Self-employed (t)	-0.00076***	0.00007	-0.00006	0.00001	0.00017
	(0.00006)	(0.00039)	(0.00039)	(0.00039)	(0.00052)
Self-employed (t-1)	–	-0.00086**	-0.00093**	-0.00079**	-0.00092*
		(0.00038)	(0.00038)	(0.00040)	(0.00052)
Any hospitalization in last 12 months	–	–	–	–	-0.01894***
					(0.00046)
Female	–	–	0.00053***	–	–
			(0.00004)		
25-34 years old	–	–	0.00018***	-0.00020**	-0.00022
			(0.00005)	(0.00009)	(0.00015)
35-44 years old	–	–	0.00081***	0.00002	0.00029
			(0.00006)	(0.00014)	(0.00021)
45-54 years old	–	–	0.00119***	0.00009	0.00027
			(0.00007)	(0.00020)	(0.00030)
55-70 years old	–	–	0.00140***	0.00055*	0.00081*
			(0.00009)	(0.00029)	(0.00042)
Foreign	–	–	-0.00011**	–	–
			(0.00006)		
Time FE	–	–	Yes	Yes	Yes
Individual FE	–	–	–	Yes	Yes
Constant	0.00166***	0.00169***	0.00049***	0.00114***	0.00190***
	(0.00002)	(0.00002)	(0.00013)	(0.00016)	(0.00024)
Observations	7,018,732	6,693,712	6,693,712	6,693,712	4,386,427
R^2	0.000	0.000	0.000	0.000	0.002

Robust standard errors in parentheses. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$

Table 3: Heterogeneous effects of self-employment on the likelihood of hospitalization

	Men	Women	Under 45 years old	45+ years old	Native	Foreign
Model 2						
Self-employed (t)	-0.00005	0.00019	0.00001	-0.00014	0.00014	-0.00045***
	(0.00074)	(0.00032)	(0.00038)	(0.00112)	(0.00044)	(0.00008)
Self-employed (t-1)	-0.00054	-0.00114***	-0.00091**	-0.00088	-0.00088**	-0.00081***
	(0.00074)	(0.00032)	(0.00038)	(0.00112)	(0.00044)	(0.00006)
Observations	3,160,919	3,532,793	4,972,053	1,721,659	5,900,613	793,099
R^2	0.000	0.000	0.000	0.000	0.000	0.000
Model 4						
Self-employed (t)	-0.00001	0.00001	0.00003	0.00012	0.00006	-0.00032
	(0.00075)	(0.00035)	(0.00040)	(0.00111)	(0.00045)	(0.00020)
Self-employed (t-1)	-0.00048	-0.00107***	-0.00083**	-0.00078	-0.00078*	-0.00081***
	(0.00076)	(0.00035)	(0.00040)	(0.00116)	(0.00045)	(0.00022)
Observations	3,160,919	3,532,793	4,972,053	1,721,659	5,900,613	793,099
R^2	0.000	0.000	0.000	0.000	0.000	0.000
Model 5						
Self-employed (t)	-0.00010	0.00044	0.00017	0.00008	0.00027	-0.00062*
	(0.00083)	(0.00061)	(0.00068)	(0.00031)	(0.00058)	(0.00037)
Self-employed (t-1)	-0.00067	-0.00118*	-0.00106	-0.00050	-0.00093	-0.00087**
	(0.00084)	(0.00063)	(0.00069)	(0.00038)	(0.00059)	(0.00042)
Observations	2,144,077	2,242,350	3,179,571	1,206,856	3,909,682	476,745
R^2	0.002	0.002	0.003	0.003	0.002	0.003
Model 5 without self-employment indicator in t						
Self-employed (t-1)	-0.00075***	-0.00078**	-0.00090***	-0.00043	-0.00068***	-0.00142**
	(0.00029)	(0.00037)	(0.00026)	(0.00062)	(0.00025)	(0.00072)
Observations	2,144,077	2,242,350	3,179,571	1,206,856	3,909,682	476,745
R^2	0.002	0.002	0.003	0.003	0.002	0.003
Simple panel regression (FE)						
Self-employed (t-1)	-0.00050***	-0.00109***	-0.00080***	-0.00080**	-0.00076***	-0.00102***
	(0.00017)	(0.00021)	(0.00014)	(0.00039)	(0.00015)	(0.00032)
Observations	3,160,919	3,532,793	4,972,053	1,721,659	5,900,613	793,099
R^2	0.000	0.000	0.000	0.000	0.000	0.000
Simple linear regression (OLS)						
Self-employed (t-1)	-0.00059***	-0.00095***	-0.00090***	-0.00101***	-0.00074***	-0.00125***
	(0.00008)	(0.00009)	(0.00006)	(0.00011)	(0.00006)	(0.00012)
Observations	3,160,919	3,532,793	4,972,053	1,721,659	5,900,613	793,099
R^2	0.000	0.000	0.000	0.000	0.000	0.000

Robust standard errors in parentheses. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$

Table 4: Effects of switching to vs. from self-employment on the likelihood of hospitalization

	Specification 1	Specification 2
Individual switches to self-employment and		
is self-employed in t	-0.00052*	-0.00067
	(0.00027)	(0.00052)
is self-employed in t-1	-0.00020	-0.00007
	(0.00029)	(0.00053)
Individual switches from self-employment and		
is self-employed in t	0.00066***	0.00061**
	(0.00023)	(0.00024)
is self-employed in t-1	-0.00109***	-0.00106***
	(0.00028)	(0.00027)
Individual switches to-and-fro and		
is self-employed in t	-0.00022	–
	(0.00075)	
is self-employed in t-1	0.00017	–
	(0.00079)	

Similarly to Model 5, Specifications 1 and 2 control for age, time and individual fixed effects, and any hospitalization in last 12 months. Robust standard errors in parentheses. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$

Table 5: Number of observations according to switches to vs. from self-employment

	Doesn't switch to self-employment	Switches to self-employment	Total
Doesn't switch to wage work	6,683,528	97,978	6,781,506
Switches to wage work	125,044	112,182	237,226
Total	6,808,572	210,160	7,018,732

Table 6: Effect of self-employment on the duration of hospitalization

	Model 1	Model 2	Model 3	Model 4	Model 5
Self-employed (t)	0.34810***	0.58524***	0.53278**	0.16755	Dropped[1]
	(0.06236)	(0.21656)	(0.22924)	(0.24201)	
Self-employed (t-1)	–	-0.23677	-0.28452	-1.56319***	-1.93747***
		(0.22124)	(0.23302)	(0.28062)	(0.39535)
Any hospitalization in last 12 months	–	–	–	–	0.08756
					(0.05533)
Female	–	–	-0.03933*	–	–
			(0.02130)		
25-34 years old	–	–	-0.03529	0.00661	-0.39761
			(0.04172)	(0.23787)	(0.32966)
35-44 years old	–	–	0.11712***	0.01833	-0.38482
			(0.04114)	(0.28883)	(0.39662)
45-54 years old	–	–	0.26590***	0.35021	-0.00423
			(0.04316)	(0.35178)	(0.47743)
55-70 years old	–	–	0.27521***	0.46500	0.29440
			(0.05443)	(0.40414)	(0.53082)
Foreign	–	–	-0.11886***	–	–
			(0.03885)		
Time FE	–	–	Yes	Yes	Yes
Individual FE	–	–	–	Yes	Yes
Constant	2.94695***	2.93973***	2.83582***	3.18876***	2.61267***
	(0.01096)	(0.01099)	(0.10512)	(0.34607)	(0.43560)
Observations	11,399	11,061	11,061	11,061	7,802
R^2	0.003	0.003	0.035	0.068	0.093

[1]Dropped due to collinearity. Robust standard errors in parentheses. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$

Table 7: Effect of self-employment on the likelihood of sick leave

	Model 1	Model 2	Model 3	Model 4	Model 5
Self-employed (t)	-0.01813***	-0.00750***	-0.00749***	-0.00560***	-0.00592***
	(0.00017)	(0.00107)	(0.00107)	(0.00111)	(0.00164)
Self-employed (t-1)	–	-0.01072***	-0.01100***	-0.00605***	-0.00767***
		(0.00107)	(0.00107)	(0.00109)	(0.00162)
Any sick leave in last 12 months	–	–	–	–	-0.01731***
					(0.00035)
Female	–	–	0.01038***	–	–
			(0.00019)		
25-34 years old	–	–	-0.00002	0.00131***	0.00116*
			(0.00026)	(0.00041)	(0.00064)
35-44 years old	–	–	0.00183***	-0.00035	-0.00015
			(0.00029)	(0.00055)	(0.00082)
45-54 years old	–	–	0.00402***	-0.00015	-0.00010
			(0.00033)	(0.00076)	(0.00107)
55-70 years old	–	–	0.00363***	0.00117	0.00182
			(0.00039)	(0.00099)	(0.00138)
Foreign	–	–	-0.00278***	–	–
			(0.00027)		
Time FE	–	–	Yes	Yes	Yes
Individual FE	–	–	–	Yes	Yes
Constant	0.02129***	0.02129***	0.01982***	0.02325***	0.02047***
	(0.00010)	(0.00010)	(0.00060)	(0.00068)	(0.00088)
Observations	7,018,732	6,693,712	6,693,712	6,693,712	4,386,427
R^2	0.001	0.001	0.003	0.001	0.002

Robust standard errors in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 8: Effect of self-employment on the likelihood of mortality in year t+1

	Model 1	Model 2	Model 3	Model 4	Model 5
Self-employed in year t	-0.00004	-0.00077***	-0.00094*	-0.00083	-0.00081
	(0.00025)	(0.00025)	(0.00052)	(0.00052)	(0.00052)
Any hospitalization in year t	–	–	–	–	0.00227***
					(0.00061)
Female	–	-0.00105***	–	–	–
		(0.00008)			
25-34 years old	–	0.00010	–	-0.00095***	-0.00095***
		(0.00007)		(0.00014)	(0.00014)
35-44 years old	–	0.00057***	–	-0.00166***	-0.00166***
		(0.00009)		(0.00024)	(0.00024)
45-54 years old	–	0.00175***	–	-0.00083**	-0.00083**
		(0.00015)		(0.00041)	(0.00041)
55-70 years old	–	0.00370***	–	0.00188***	0.00188***
		(0.00028)		(0.00070)	(0.00070)
Foreign	–	-0.00009	–	–	–
		(0.00011)			
Year FE	–	Yes	–	Yes	Yes
Individual FE	–	–	Yes	Yes	Yes
Constant	0.00106***	0.00059***	0.00108***	-0.00002	-0.00004
	(0.00004)	(0.00011)	(0.00001)	(0.00017)	(0.00017)
Observations	696,676	696,676	696,676	696,676	696,676
R^2	0.000	0.001	0.000	0.002	0.002

Robust standard errors in parentheses. $*p < 0.1$, $**p < 0.05$, $***p < 0.01$

Table 9: Effect of self-employment on the likelihood of hospitalization (quarterly observations)

	Model 1	Model 2	Model 3	Model 4	Model 5
Self-employed (t)	-0.00214***	-0.00060	-0.00104*	-0.00097	-0.00096
	(0.00017)	(0.00063)	(0.00063)	(0.00067)	(0.00094)
Self-employed (t-1)	–	-0.00163***	-0.00177***	-0.00125*	-0.00142
		(0.00061)	(0.00061)	(0.00066)	(0.00091)
Any hospitalization in last 12 months	–	–	–	–	-0.05394***
					(0.00094)
Female	–	–	0.00148***	–	–
			(0.00010)		
25-34 years old	–	–	0.00049***	-0.00060**	-0.00080**
			(0.00014)	(0.00026)	(0.00039)
35-44 years old	–	–	0.00235***	0.00004	0.00039
			(0.00016)	(0.00039)	(0.00055)
45-54 years old	–	–	0.00339***	0.00031	0.00083
			(0.00019)	(0.00057)	(0.00078)
55-70 years old	–	–	0.00393***	0.00148*	0.00207*
			(0.00025)	(0.00081)	(0.00109)
Foreign	–	–	-0.00036**	–	–
			(0.00015)		
Time FE	–	–	Yes	Yes	Yes
Individual FE	–	–	–	Yes	Yes
Constant	0.00462***	0.00486***	0.00176***	0.00380***	0.00593***
	(0.00005)	(0.00005)	(0.00025)	(0.00035)	(0.00050)
Observations	2,481,892	2,264,626	2,264,626	2,264,626	1,751,168
R^2	0.000	0.000	0.001	0.000	0.009

Robust standard errors in parentheses. $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$