Thema

UMR 8184

THéorie Économique, Modélisation et Applications

THEMA Working Paper n°2017-07
Université de Cergy-Pontoise, France

# Dynamic interactions between health and employment statuses : a non-parametric analysis

Richard Moussa, Eric Delattre

UNIVERSITÉ de Cergy-Pontoise

CNRS CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Mars 2017

# Dynamic interactions between health and employment statuses : a non-parametric analysis

Richard Moussa [*]

Eric Delattre[†]

March 14, 2017

**Abstract**

Despite numerous sociological results, there is few econometric evidence on the causal links between health condition and job status. It is important to investigate the stability of these causal links during one's professional life. Papers that treat causal links between health and job statuses, make the assumption that causal links are identical over time. This could lead to a weak assessment of the causal effects. In this paper, we use a non-parametric approach, the Kullback causality measure, to test for causal links among time periods as well as global causal links. Our approach is more robust than the ones available and allows the determination of the effects of individual characteristics on causal links. We find significant reciprocal causal links between health condition (regardless of disease severity) and job status. However, job status does not cause both illness with large disability index and illness with large risk of death. These findings confirm evidence from the literature. However, analyzing the dynamic of the evolution of causal links between job status health condition regardless of severity allows us to conclude that job status only causes health between the $11^{th}$ and the $17^{th}$ year of professional life while only at the same period, health condition does not cause job status.

**Keywords: Causality; Markov chain; Kullback Information; Health; Employment**

**JEL Classification: C14, C25, D31, I10, J20**

[*]ThEMA, Université de Cergy-Pontoise, France and ENSEA, Côte d'Ivoire.

[†]ThEMA, Université de Cergy-Pontoise.

# Introduction

Relationship between health condition and job status has been analyzed according to several approaches in both the economic[1] and sociological literature. As it is well known that health is a key factor of job status and it's transitions among professional life (Grossman, 1972), the link between health and job status has been firstly analyzed as a one-way causal link with health explaining job status. But many studies (Stern, 1989; Haan and Myck, 2009; Caroli and Godard, 2014; Delattre et al. 2015) show that, when analyzing job status, health may not be treated as exogenous. This may lead to biased estimations of health impact on job status.

To overcome the problem of endogeneity in the relationship between health and job status and to allow a causal analysis, two approaches have generally been used. The first approach makes use of instrumental variables methods. Caroli and Godard (2014) show that without instrumenting job insecurity, job seems to deteriorate all health indicators but the instrumenting approach shows that only few health indicators are deteriorated by job[2]. This approach helps to solve endogeneity problem but allows analyzing only a one-directional causality.

The second way to deal with endogeneity problems and two-directional causality is to estimate a bivariate model. Cai (2010) uses a simultaneous equations model approach on Australian panel data and shows that health status affects positively job status and employment affects positively women's health but negatively men's health. Barnay and Legendre (2012) use a bivariate ordered probit model to show that there is bidirectional instantaneous causality between health status and employment status, and these results are true for both genders. Haan and Myck (2009) used a bivariate dynamic logistic model on German socioeconomic panel data. They show that both last health condition and last labour market risk affect the current labour market risk and health condition, and that the dynamic is persistent. Delattre et al. (2015) use a bivariate dynamic probit model on the French longitudinal data on health and professional path to show that health causes job status and vice versa. Besides the specified model (probit vs. logit), their approach differs from that of Haan and Myck (2009) by the treatment of the initial conditions

[1]Barnay (2016) surveys European literature on health and job conditions' relationships.

[2]Stern (1989) also uses instrumental variable approach to reach the same goal.

in the estimated model[3]. Haan and Myck (2009) treat the initial conditions as exogenous when Delattre et al. (2015) treat the initial conditions as endogenous. However, both approaches allow analyzing a bi-directional causality.

The model specification made by Delattre et al. (2015) aims to overcome specification problems that may often lead to misjudgement of causal links. However, the approaches above have an underlying hypothesis in which the causal links between health and job status is homogeneous among individual's professional life. But, there are some sociological and very few econometric evidence that question this assumption.

As mentioned by Waddell and Burton (2006), health selection for entering work is less important since younger people are assumed healthier. Unemployment effects on younger people well-being are different from those that older because younger people often receive parental support and are assumed to have less financial and social commitments than their counterpart. In the same way, Lakey et al. (2001) argue that health effects of unemployment are more severe on older workers than younger. Haan and Myck (2009) show that health condition is particularly important for employment after age 50. As a result, for a full assessment of the causal links between health and job statuses, one may account for time and individual heterogeneity of causal links by analyzing the evolution of the causal links among individuals' professional life. This suggests that there may exist some individual characteristics that affect the causal links. Thus, one may also account for these factors.

The latest point has not been fully discussed in the literature about causality. Researchers mainly focus on determining whether there exists a causal link between a set of variables as the causal link is supposed to be homogeneous among time and individual. But, there are few papers that have addressed the issue of characteristics that affect the causal link. In macroeconomics, with time series, researchers often use the regime shift framework to show that policies implementation affects the causal link between a set of variables (Firouz, 2011; Balcilar et al., 2015). On cross-sectional data, researchers estimate the causal links among some clusters created by a

[3] Initial condition refers to an individual health and job statuses the first time he/she has been observed in the dataset. In our dataset, this date is the date when an individual leaves school.

categorical variable. Cai (2010) uses gender as grouping variable and shows that employment is positively instantaneously causal for female's health and negatively causal for male's health. Salm (2009) finds that for near elderly employees, job loss does not cause any physical or mental illness. These two approaches are almost identical since they divide the sample in sub-samples according to some categorical variables before estimating causal links on each sub-samples. The lack of attempts to identify which individual characteristics affect the causal links is due to the assumption that causal links are homogeneous on the overall sample or by sub-samples. Thus, the estimation of a time-varying or individual-specific or both causal links allows addressing the issue of which characteristics have significant impact on the causal links.

In this paper, the hypothesis we aim to test is that the causal links between health condition and job status, if they can be proved in general on the overall observation period, are not homogeneous during professional life. We propose a non-parametric analysis of the causal link based on a Kullback causality measure developed by Gouriéroux et al. (1987). This approach is applicable to qualitative outcomes and allows the assessment of the causal links evolution among time periods as well as global causal links. This approach has three major advantages : (i) it is not biased by a misspecification problem and remains robust even if the causal links are nonlinear, (ii) it allows testing for causality at each time period as well as the global causality on overall time period, and (iii) it allows analyzing the contribution of each state to the causal links at each period and the effects of individual characteristics on the causal links.

This paper begins by giving an overview of the literature on causality test methods and the description of our methodology in Section 1. In Section 2, we present the dataset and some related descriptive statistics. Section 3 presents the results and we conclude in Section 4.

# 1 Econometrics Model

## 1.1 General framework of causality test methods

The original conception of Granger non-causality is the better predictability of a variable $Y$ by the use of lagged values of $Z$. Granger (1969) distinguishes lag causality from instantaneous causality. Instantaneous causality from $Z_t$ to $Y_t$ denotes that the knowledge of $Z_t$ improves the predictability of $Y_t$. This definition is not often used in applied studies. Then, the most common definition in literature is the lag causality that denotes that the use of lagged values of $Z_t$ improves the predictability of $Y_t$.

There are various approaches in the literature to test for Granger non-causality. It can be achieved by specifying a dynamic relationship model between variables or in terms of probability as conditional independence between variables. For quantitative time series or quantitative panel data the common approach is to consider that causality between variables, when it exists, is the same for all periods or all individuals. This assumption is abridged by Weinhold and Nair-Reichert (2000) in the following terms : *"either causality occurs everywhere or it occurs nowhere"*.

Without lost of generality, we present a bivariate case that can be easily generalized to multivariate case. The specification that allows testing for Granger lag and instantaneous causality for time series case is :

$$Y_t = \alpha_1 Z_t + \delta_{11} Y_{t-1} + \delta_{12} Z_{t-1} + \beta_1 X_t + \epsilon_t^1 \quad (1)$$
$$Z_t = \alpha_2 Y_t + \delta_{21} Y_{t-1} + \delta_{22} Z_{t-1} + \beta_2 X_t + \epsilon_t^2 \quad (2)$$

With traditional assumption of normality on $\epsilon = (\epsilon^1, \epsilon^2)$. For panel data with a one way error component model :

$$Y_{it} = \alpha_1 Z_{i,t} + \delta_{11} Y_{i,t-1} + \delta_{12} Z_{i,t-1} + \beta_1 X_{i,t} + \eta_i^1 + \zeta_{it}^1 \quad (3)$$
$$Z_{it} = \alpha_2 Y_{i,t} + \delta_{21} Y_{i,t-1} + \delta_{22} Z_{i,t-1} + \beta_2 X_{i,t} + \eta_i^2 + \zeta_{it}^2 \quad (4)$$

Also with standards assumptions of normality on $\eta = (\eta^1, \eta^2)$ and $\zeta = (\zeta^1, \zeta^2)$. The non-causality test in these models consists in a linear constraints test on $\delta_{12}$ and $\delta_{21}$ if we wish to test for lag non-causality, and on $\alpha_1$ and $\alpha_2$ if we wish to test for instantaneous non-causality. Note that when one does not account for instantaneous non-causality, $\alpha_1$ and $\alpha_2$ are null in the two models above (Equations 1 to 4).

The causal effect can be different from an individual to another in a panel or from a time period to an other. This can be true in heterogeneous datasets (see Weinhold and Nair-Reichert, 2000) or when the causal effect is not homogeneous. In the case of individual specific causal links, coefficients $\delta_{12}$ and $\delta_{21}$ are different for each individuals[4] or more generally, one can assume a distribution on these coefficients. Then, researchers use the Mixed Fixed and Random model framework to estimate coefficients (Hsiao et al., 1989; Weinhold and Nair-Reichert, 2000). This specification has the advantage to give a better assessment of the distribution of the causal effect among individuals.

Causal link can also be time dependent (Adams et al., 2003; Balcilar et al., 2015), it denotes that $Z$ can be causal for $Y$ at certain time periods but not at all, specially for lag causality case. This may happen when there are some policy interventions that alter $Z$'s distribution or when the process meets an equilibrium after a while (it denotes that $Z$ becomes not causal for $Y$). In this case, one may account for these time-specific causal effects when testing for non-causality. Adams et al. (2003) and Balcilar et al. (2015) approaches are an application of the *regime shift model* to causality test. For panel data case, it consists to consider that there is a causal link between variables when there is a conditional independence between these variables and when the invariance property is reached. It means that the causal link is assumed to be true when it remains stable from a panel wave to an other. Thus a Chow type-test is run to address this issue.

In all specifications above, when dependent variables are qualitative outcomes (inducing that error term $\zeta$ can not be treated as normal), it is common to use latent variables and probit probabilities[5]. As these approaches deal with a parametric framework and specified models have linear forms, it is well known that any misspecification or nonlinear causal links may lead to wrong conclusion on Granger non-causality. To overcome the problem of misspecification and nonlinear case and the problem of degree of freedom reduction by the time or individual specific causal effect when testing for Granger non-causality, some non-parametric approaches have been developed. All those tests assume the processes to be Markov of a fixed order p and are more robust.

For time series[6], Bouezmarni et al. (2012) propose a non-parametric copula-based test for

[4]See Delattre et al. (2015) for further details.

[5]Delattre et al. (2015) give further details on theses specifications.

[6]See Bouezmarni et al., (2012) for literature on others non-parametric approaches for testing conditional independence

Granger causality using the Hellinger distance under the assumption that the interest process is $\beta$-mixing. They derive a test statistic that follows a standard normal distribution under the null hypothesis of conditional independence.

Bouissou et al. (1986) derive a non-causality test for qualitative processes on panel data. To test for the Granger non-causality of $Y$ on $Z$, they derive a log-likelihood ratio test (LR Test) based on the assumption that $Z$ is a Markov chain of order one. Another approach is to test Granger non-causality by using the Kullback causality measure[7]. This approach can also be applied for qualitative interest variables on panel data. The main advantages of both Gouriéroux et al.'s (1987) and Bouissou et al.'s (1986) approaches are that (i) they allow examining how causal links vary through time periods as well as the global causal links on the overall observation period, and (ii) they are non-parametric approaches and only based on the assumption that interest variables are a Markov chain of order one.

It is the latest approach that is used in this paper. We use this approach because we assume that the causal links between our two binary dependent variables may change over time and we need to assess the causal links between health or job statuses. In the following section, a full description of the test process is provided.

## 1.2 Model specification

Let $W_{i,t}$ denotes job status and $H_{i,t}$ denotes the health status of individual $i$ at the period $t$ of his professional life. A state of nature is given by a realisation of $W_{i,t}$ and $H_{i,t}$, denotes $s_{i,t} = (s^1_{i,t}, s^2_{i,t}) = (w, h) \in \{(1,1); (1,0); (0,1); (0,0)\}$. Transition probability between a state of nature at the period $t-1$ and the new state of nature at the period $t$ of individual professional life is given by:

$$p_{i,t}(s_{i,t}|s_{i,t-1}) \quad = \quad P\Big((W_{i,t}, H_{i,t}) = s_{i,t}|(W_{i,t-1}, H_{i,t-1}) = s_{i,t-1}\Big) \tag{5}$$

Testing for Granger non-causality on this qualitative process can be done by the use of a non-parametric test: the Kullback causality measure developed by Gouriéroux et al. (1987). The rationale is that the independence between two variables $X_t$ and $Y_t$ conditionally to $X_{t-1}$ and

[7]See Gouriéroux et al. (1987) for more details

$Y_{t-1}$ is equivalent to non-causality from $X_t$ to $Y_t$, non-causality from $Y_t$ to $X_t$, and instantaneous non-causality between $X_t$ and $Y_t$. We first assume our process to be an homogeneous (among individuals) Markov chain of order one. Formally, we assume $p_{i,t}(s_{i,t}|s_{i,t-1}) = p_t(s_t|s_{t-1})$ for all individuals. Thus, the test statistics used for this purpose is (for the non-causality from $W$ to $H$):

$$\hat{C}_{W\,to\,H} = \frac{1}{T}\sum_{t=1}^{T}\hat{C}_{W\,to\,H}(t) = \frac{1}{T}\sum_{t=1}^{T}\sum_{w=0}^{1}\sum_{h=0}^{1} I\hat{\pi}_{t-1}(w,h)\hat{C}_{W\,to\,H}(t,w,h) \tag{6}$$

Where $\hat{\pi}_{t-1}(w,h) = \hat{P}\Big((W_{t-1},H_{t-1}) = (w,h)\Big)$, $I$ denotes the number of individuals. To test for lag non-causality, we use:

$$\hat{C}_{W\,to\,H}(t,w,h) = \sum_{s_t^2=0}^{1}\hat{p}_t((.,s_t^2)|(w,h))log\frac{\hat{p}_t((.,s_t^2)|(w,h))}{\hat{p}_{H,t}(s_t^2|h)} \tag{7}$$

To test for instantaneous non-causality, instead of $\hat{C}_{W\,to\,H}(t,w,h)$ we use:

$$\hat{C}_{W,H}(t,w,h) = \sum_{s_t^1=0}^{1}\sum_{s_t^2=0}^{1}\hat{p}_t((s_t^1,s_t^2)|(w,h))log\frac{\hat{p}_t((s_t^1,s_t^2|w,h))}{\hat{p}_t((s_t^1,.)|(w,h))\hat{p}_t((.,s_t^2)|(w,h))} \tag{8}$$

Where

$$\begin{aligned}
\hat{p}_t((.,s_t^2)|(w,h)) &= \sum_{s_t^1=0}^{1}\hat{p}_t((s_t^1,s_t^2)|(w,h)) \\
\hat{p}_t((s_t^1,.)|(w,h)) &= \sum_{s_t^2=0}^{1}\hat{p}_t((s_t^1,s_t^2)|(w,h)) \\
\hat{p}_{H,t}(s_t^2|h) &= \frac{\sum_{s_{t-1}^1=0}^{1}\sum_{s_t^1=0}^{1}\hat{p}_t((s_t^1,s_t^2)|(s_{t-1}^1,h))\hat{\pi}_{t-1}(s_{t-1}^1,h)}{\sum_{s_{t-1}^1=0}^{1}\hat{\pi}_{t-1}(s_{t-1}^1,h)}
\end{aligned}$$

Asymptotically, $2T\hat{C}_{W\,to\,H}$ has a chi-square distribution with $2T$ degrees of freedom under null hypothesis for testing non-causality from $W$ to $H$, $2T\hat{C}_{W,H}$ has a chi-square distribution with $4T$ degrees of freedom under null hypothesis for testing instantaneous non-causality between $W$ and $H$.

As described by Gouriéroux et al. (1987), $\hat{C}_{W\,to\,H}(t,w,h)$ is a causality measure for the state ($w$ , $h$) for the transition between periods $t-1$ and $t$. When this measure is near zero, it denotes a non-causality from $W$ to $H$ for the state ($w$ , $h$). The test statistics $2I\hat{\pi}_{t-1}(w,h)\hat{C}_{W\,to\,H}(t,w,h)$

has asymptotically a chi-square distribution with 1 degree of freedom under null hypothesis of non-causality, and for each $w$ and $h$ the statistics $2I\hat{\pi}_{t-1}(w,h)\hat{C}_{W\,to\,H}(t,w,h)$ are asymptotically independent for each time period. It means that as we can test Granger non-causality for the overall observation period, we can also test for Granger non-causality at each observation period between specific states of nature.

The global causality measure at the period $t$ from job status (W) to health condition (H) is given by:

$$\hat{C}_{W\,to\,H}(t) = \sum_{h=0}^{1}\sum_{w=0}^{1} 2I\hat{\pi}_{t-1}(w,h)\hat{C}_{W\,to\,H}(t,w,h) \tag{9}$$

$\hat{C}_{W\,to\,H}(t)$ has asymptotically a chi-square distribution with 2 degrees of freedom under null hypothesis of non-causality from $W$ to $H$. For the global instantaneous causality measure at period $t$ between job status (W) and health condition (H), we use the statistic $\hat{C}_{W,H}(t) = \sum_{h=0}^{1}\sum_{w=0}^{1} 2I\hat{\pi}_{t-1}(w,h)\hat{C}_{W,H}(t,w,h)$ that has asymptotically a chi-square distribution with 4 degrees of freedom under null hypothesis of instantaneous non-causality between health condition and job status. Note that a similar statistics can be derived for testing Granger non-causality from $H$ to $W$[8]. The contributions of each state of nature to the causal links can be derived from the global causality measure at each time period. For a state $(w,h)$, the contribution to the causality measure from $W$ to $H$ is given by:

$$Ctr^{(w,h)}_{W\,to\,H} = \frac{2I\hat{\pi}_{t-1}(w,h)\hat{C}_{W\,to\,H}(t,w,h)}{\hat{C}_{W\,to\,H}(t)} \tag{10}$$

This statistic allows determining at each period, states from which causal links mainly depend. It allows us to give an analysis of the causal links structure among individuals professional life. Note that the same statistic can be written for the instantaneous causality measure and for the

[8]In this case, the test statistic is $2I\hat{\pi}_{t-1}(w,h)\hat{C}_{H\,to\,W}(t,w,h)$ where

$$\hat{C}_{H\,to\,W}(t,w,h) = \sum_{s_t^1=0}^{1} \hat{p}_t((s_t^1,.)|(w,h))log\frac{\hat{p}_t((s_t^1,.)|(w,h))}{\hat{p}_{W,t}(s_t^1|w)}$$

and

$$\hat{p}_{W,t}(s_t^1|w) = \frac{\sum_{s_{t-1}^2=0}^{1}\sum_{s_t^2=0}^{1}\hat{p}_t((s_t^1,s_t^2)|(w,s_{t-1}^2))\hat{\pi}_{t-1}(w,s_{t-1}^2)}{\sum_{s_{t-1}^2=0}^{1}\hat{\pi}_{t-1}(w,s_{t-1}^2)}$$

lag causality measure from $H$ to $W$.

### 1.3 Transition probabilities estimation

As we explained above, the predicted probabilities used by Gouriéroux et al. (1986) are computed as empirical frequencies :

$$\hat{p}_t((s_t^1, s_t^2)|(w,h)) \quad = \quad \frac{N((s_t^1, s_t^2)|(w,h))}{N_{t-1}(w,h)} \tag{11}$$

Where $N((s_t^1, s_t^2)|(w,h))$ denotes the number of individual in state $(s_t^1, s_t^2)$ at $t$ conditionally to their last period's state $(w,h)$ at $t-1$, and $N_{t-1}(w,h)$ denotes the number of individuals with the state $(w,h)$ at $t-1$. As we have a panel dataset, the estimation of the transition matrix at each period (which components are the probabilities $\hat{p}_t(s_t|s_{t-1})$ with $s_t = (s_t^1, s_t^2) = (w,h) \in \{(1,1);(1,0);(0,1);(0,0)\}$.) and of the marginal probabilities $\hat{\pi}_t(w,h)$ can be achieved by using a multinomial logistic model. We use the multinomial logistic model because this specification allows us to control for individual characteristics that can affect the estimated probabilities. At each time period of the professional life, we specify the following model :

$$P\Bigg(s_{it} = s_k|s_{i,t-1}, X_{it}\Bigg) \quad = \quad \frac{exp((s_{i,t-1}, X_{it})'\beta_k)}{1+\sum_{j=1}^{3} exp((s_{i,t-1}, X_{it})'\beta_j)} \; with \; k = 1, ..., 4 \tag{12}$$

$$\sum_{k=1}^{4} P\Bigg(s_{it} \quad = \quad s_k|s_{i,t-1}, X_{it}\Bigg) = 1, with\, s_k \in \{(1,1);(1,0);(0,1);(0,0)\}$$

The individual characteristics used in this specification are socioeconomic individual characteristics, illness type and job characteristics. With this specification, not only we are able to compute predicted probabilities following Gouriéroux et al. (1986), but we can also point out characteristics that affect these probabilities. This approach has three major advantages. Firstly, as we suppose the process to be a Markov chain of order one, we control for initial conditions by taking them into account for the first transition. Furthermore, at each period, initial conditions are supposed to be the previous period conditions. This dynamic of initial conditions allows accounting for changes in individual specific conditions that affect the professional path. Notice that initial conditions play important an role in professional paths (Delattre et al., 2015). Secondly, because we control for individual characteristics, transition probabilities are not the same for each individual as when we use the empirical frequencies. Thirdly, our approach avoid the cases of 0/0 probabilities that may occur with empirical frequencies[9].

As we specify a multinomial logistic model, we have to test for the underlying hypothesis of independence of irrelevance alternatives (IIA)[10]. We achieve that goal by using the Hausman IIA test statistic that has a chi-square distribution. At each period, the predicted probabilities are computed for all transitions between states of nature and all individuals. The transition matrix components are then given by the mean of it values for all individual at the considered period.

# 2 Data and related statistics

## 2.1 Dataset

The dataset that has been used for this paper is from the French survey on health and work (Enquête Santé et Itinéraire Professionnel (SIP 2006)). It is a retrospective[11] survey achieved by DARES[12] and DREES[13] in 2006 that provide information on the health condition and the job status for individuals aged between 20 and 74 years old in 2006. It also provides information on individual socio-economic statuses. All this information is gathered from starting work the first time to 2006. After data processing, which consisted on treating missing data and dropping individual with professional life starting before 1962, the subset that has been used is a panel dataset on 10,942 individuals for an observation time varying between 2 and 45 years.

[9]Note that the use of empirical frequencies approach may often lead to a 0/0 probability. To overcome this problem, Bouissou et al. (1986) use the convention that $0/0 = 0$ and argue for that. With our approach, this case can not appear as at each time period we estimate, with respect to individuals characteristics, the probabilities of different states. These estimated probabilities are strictly positive and different from 1.

[10]which means that adding or removing a state of nature or changing it characteristics in the specified model does not change probabilities ratios between states.

[11]Individuals in 2006 are asked to provide information on each year of their professional and social life since the beginning of their professional life. Since respondents may have problems to remember events of their life, this approach induces collection bias that can affect estimation results, particularly in a parametric framework.

[12]Direction de l'Animation de la Recherche, des Etudes et des Statistiques, the statistical bureau of the French administration for Labour Affairs

[13]Direction de la Recherche, des Etudes, de l'Evaluation et des Statistiques, the statistical bureau of the French administration for Health Affairs

The variables of interest are health condition (a binary variable that takes the value 1 if individual reported an illness in the year and 0 otherwise, regardless to the illness type) and job status (also a binary variable that takes the value 1 if individual is employed and 0 otherwise, employment includes both short and long term employment). For the econometric analysis, we estimate the causal links for the health condition variable, for the variable of health with large disability index (a binary variable that takes the value 1 if individual reports illness with large disability index) and for the variable of health with large risk of death (a binary variable that takes the value 1 if individual reports illness with large risk of death). The disability index and the risk of death variable are designed by the use of the International Statistical Classification of Diseases and Related Health Problems $10^{th}$ Revision (ICD-10). Note that as health variable is self-reported, it may induce some endogeneity bias (Bound, 1991).

The other socio-economic variables used as controls are gender, school grades, age, living with a partner, number of children, national unemployment rate, number of illness period before entering the job market, illness type and medical density in individual's area.

## 2.2 Descriptive statistics on states and transitions

For all individuals on the overall observation period, reported illness represents 22% and reported employment 86%. Since our two dependent variables are binary, we have four possible states of nature during individual professional life. Those states are:

- being healthy and employed: 68.3% on the overall observation period;
- being healthy and unemployed: 9.8% on the overall observation period;
- being ill and employed: 17.7% on the overall observation period;
- being ill and unemployed: 4.2% on the overall observation period.

Reported illness or unemployment rates are time-dependent. As we can see from Figure 1(a), at the entrance in professional life, individuals are most likely to be healthy (only 11% report an illness) and employed (85.3% are employed). But, during professional life, reporting an illness becomes more frequent and one half of individuals reports an illness at the end of their professional life. Contrariwise, the employment rate grows up during the 5 first years of professional life to reach 90% before it decreases slightly and remains stable around 86% from the $10^{th}$ to the $30^{th}$ year of professional life[14]. After that, it gradually declines to reach 22.5% at earlier years. The same analysis can be done for the different states (evolutions are given in Figure 1(b)). We can notice that reporting healthy and employed rate's decline from 80% at the beginning of professional life to around 20% at the earlier years[15], while reporting ill and employed rate's grows up from 10% at the beginning of professional life to 31% after 35 years before declining until the end of professional life. Transitions between these states are dynamic during individuals professional life. From 4 states, 16 transitions are available. But, we focus on 4 transitions that are most common in the literature, not because they occur most often, but because of their economic relevance. These transitions raise economic questions such as (i) maintaining ill workers on the job market, (ii) protecting workers from illnesses due to work accidents or other sources, (iii) unemployment effects on health, and (iv) health effects on the likelihood of getting a job. There are :

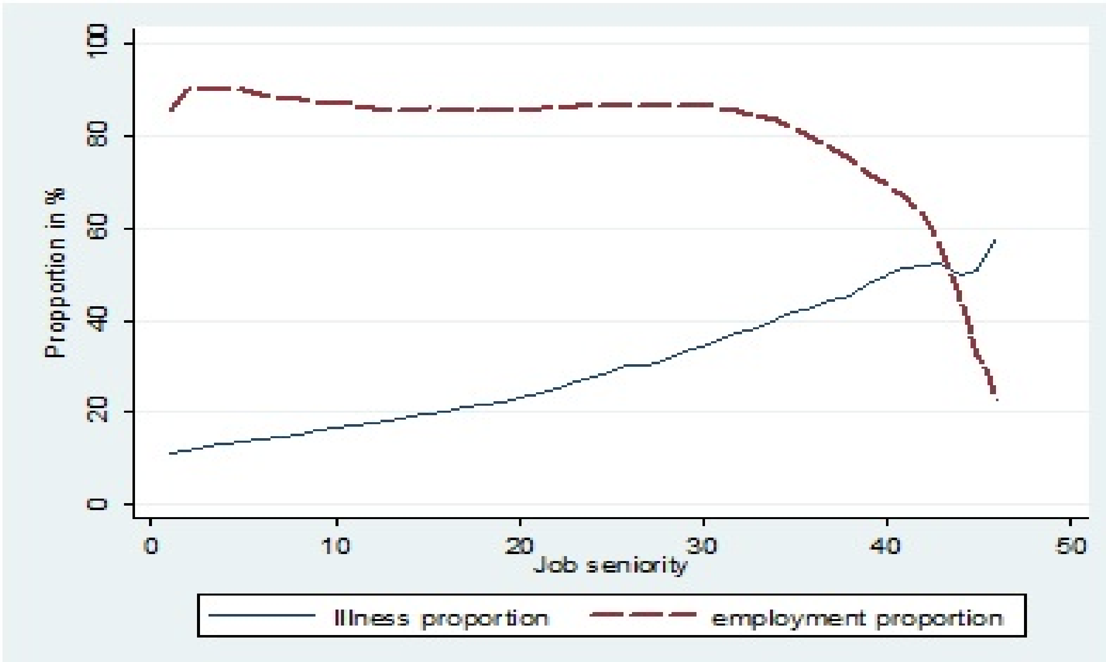
- From healthy and unemployed to healthy and employed: that can be analyzed as being healthy promotes entering work (Cai and Kalb, 2006 for econometric evidences or Benjamin and Wilson, 2005 for sociological evidences).

- From ill and employed to ill and unemployed: that can be interpreted as illness induces lost of job or illness can reduce work abilities (Stern, 1989; Waddell and Burton, 2006 for more details).

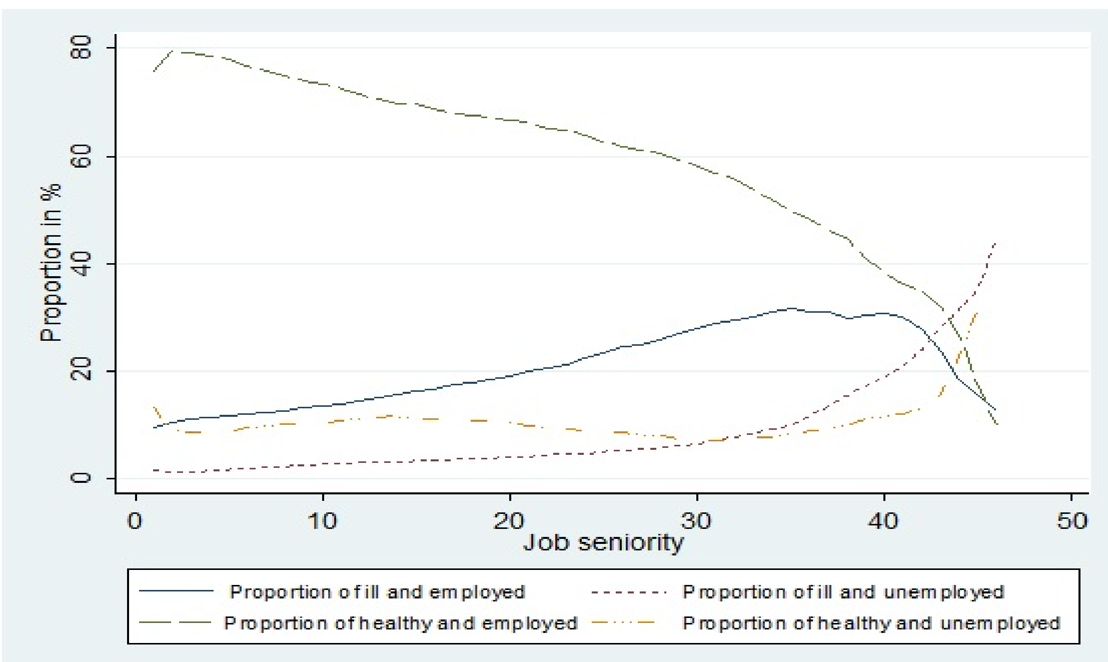
- From healthy and employed to ill and employed: it denotes, *ceteris paribus*, that working conditions and working painfulness degrade health condition (Debrand 2011 for some ev-

[14]In our data, while employment rate remains stable across time periods, one can notice that there are high changes in employment types. At the beginning of professional life, short-term employments are most common, among 52% while long-term employments are less common 35%. But, only after 3 year, the trend changes. Then, short-term employments rates decline and reach 15% after 10 years and remain stable (while declining very slowly) till the end of professional life. At the same time, long-term employments rates grow 75% after 15 years and remain stable until the end of professional life. It means that after 10 years, individuals in short-term employments have considerably reduced chances to move to long-term employments.

[15]These statistics illustrate the findings commonly underlined in literature : health and age are negatively correlated. Employment has often worse consequences on health (Debrand, 2011; Caroli and Godard, 2014).




(a) Probabilities of illness and employment




(b) Probabilities of each states

Figure 1: Evolution of proportion of individuals in each state of nature

idence). Caroli and Godard (2014) also show that the fear of involuntary job loss affects health.

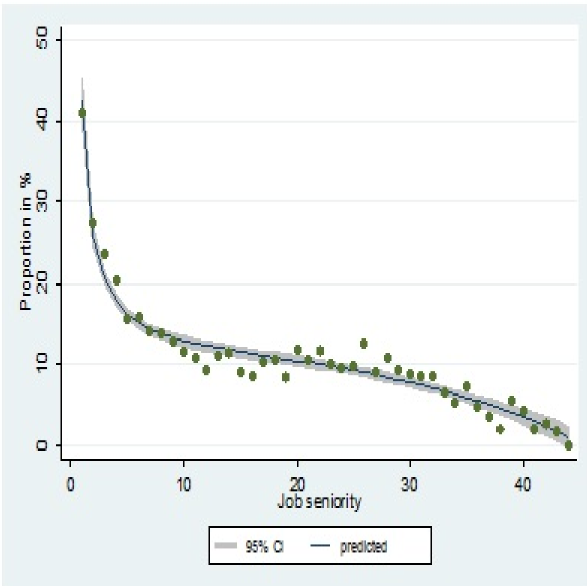
- From healthy and unemployed to ill and unemployed: it means that unemployment can affect mental health and also physical health for some pecuniary reasons (See Murphy and Athanasou, 1999 for sociological evidence; Case and Deaton, 2005 for econometric evidence).

The dynamic of these four specific transitions in states of nature are described in the Figure 2 below. We compute in Figure 2, the transition probabilities between states and fit a non-linear curve with a 5 % level confident interval. It clearly appears that transition probabilities between states of nature are not homogeneous during the professional life. They have different paths during professional life.

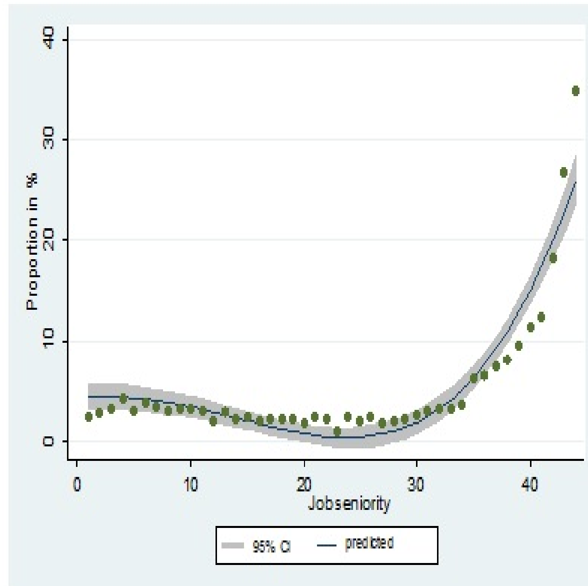
- From healthy and unemployed to healthy and employed (Figure 2(a)). At the beginning of the professional life, this transition is very easy for an individual (with a probability around 40%). The transition probability decreases quickly at 15% only after 5 years of professional life. During individual professional life, this transition gradually becomes more and more difficult (the probability is approximatively 10% between the $10^{th}$ and $20^{th}$ years of professional life). After 30 years of professional life, the probability becomes less than 5%. It suggests that the effect of health on individual chances to access the job market declines gradually during professional life, and after 30 years of professional life, this effect is quite null.

- From ill and employed to ill and unemployed (Figure 2(b)). At the beginning of the professional life, there are no evidence that being ill for an individual induces lost of his job (around 2.5%). This transition remains lower than 4% until 30 years of professional life. But, after 30 years of professional life, this transition rises exponentially.

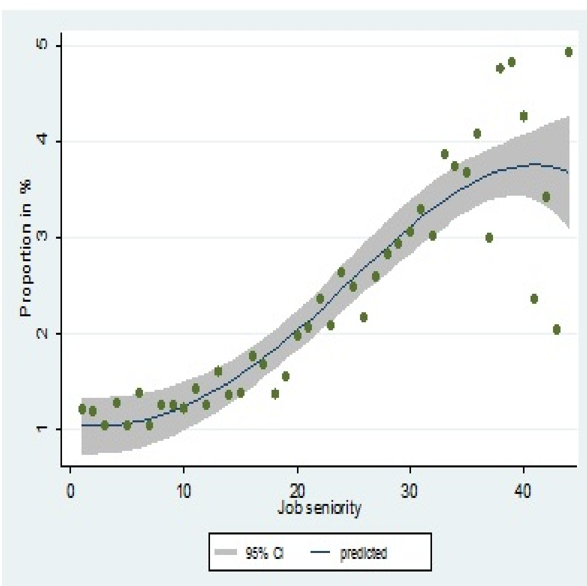
- From healthy and employed to ill and employed (Figure 2(c)). Around 1% at the beginning of individual's professional life, this transition rises gradually during professional life. It remains less than 2% until 20 years of professional life, and after the rise becomes more important and reach 4% at 30 years of professional life.



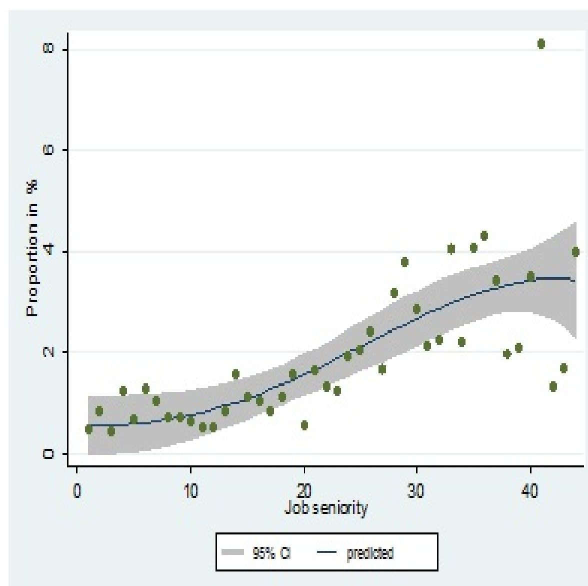

(a) Transition from healthy and unemployed to healthy and employed




(b) Transition from ill and employed to ill and unemployed




(c) Transition from healthy and employed to ill and employed




(d) Transition from healthy and unemployed to ill and unemployed

Figure 2: Dynamic of the transitions between some states of nature.

- From healthy and unemployed to ill and unemployed (Figure 2(d)). This transition is less than 1% at the beginning of professional life. As an individual progresses in his professional life, the likelihood of transition from "healthy and unemployed" to "ill and unemployed" increases and reaches 4% at the end of their professional life.

This analysis remains the same regardless of the professional life beginning period. When we consider the cohorts of individuals whose professional life begins between 1960 and 1969, or 1970 and 1979, or 1980 and 1989, or 1990 and 1999 the trend seems to have the same structure during professional life.

Individual socio economic characteristics that may impact health condition and job status are described for individuals of the data set in the Table 1 for the first period of professional life and for the 10th year of professional life. As we can see in Table 1, at the beginning of professional

Table 1: Socio-economic characteristics at the $1^{st}$ and $10^{th}$ periods of professional life

| Variables | $IE^*$ | | $IU^*$ | | $NIE^*$ | | $NIU^*$ | | All | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $1^{st}$ | $10^{th}$ | $1^{st}$ | $10^{th}$ | $1^{st}$ | $10^{th}$ | $1^{st}$ | $10^{th}$ | $1^{st}$ | $10^{th}$ |
| Gender = men (%) | 41.1 | 46.9 | 32.0 | 13.0 | 47.3 | 52.0 | 40.1 | 6.5 | 45.6 | 45.6 |
| $Not\,French^+$ (%) | 7.3 | 8.8 | 6.4 | 10.9 | 9.8 | 11.2 | 23.2 | 23.1 | 11.3 | 12.1 |
| Couple (%) | 26.0 | 76.3 | 16.9 | 76.5 | 19.1 | 76.5 | 16.0 | 87.6 | 19.3 | 77.6 |
| Number of child | 0.03 | 1.02 | 0.10 | 1.60 | 0.04 | 1.0 | 0.09 | 1.8 | 0.05 | 1.10 |
| No grade (%) | 5.2 | 5.7 | 16.9 | 16.6 | 5.5 | 6.4 | 16.2 | 17.5 | 7.0 | 7.8 |
| High school grade (%) | 39.7 | 46.1 | 48.3 | 55.5 | 49.0 | 51.0 | 46.7 | 55.3 | 47.8 | 50.9 |
| College grade (%) | 18.6 | 19.0 | 21.5 | 13.8 | 16.7 | 16.8 | 17.8 | 12.6 | 17.1 | 16.5 |
| Undergraduate studies (%) | 13.7 | 11.7 | 6.4 | 6.9 | 12.0 | 10.7 | 8.1 | 7.4 | 11.6 | 10.4 |
| Graduate studies (%) | 22.8 | 17.6 | 7.0 | 7.3 | 16.8 | 15.1 | 11.3 | 7.3 | 16.5 | 14.4 |
| Number of observations | 1,032 | 1,269 | 172 | 247 | 8,296 | 6,843 | 1,442 | 972 | 10,942 | 9,331 |

$*$ IE: Ill and Employed, IU : Ill and Unemployed, NIE : healthy and Employed, NIU : healthy and Unemployed, + : Refers to individual's nationality

life, individuals who are not French are often healthy but unemployed, and those with graduate studies levels or in couple are most commonly ill and employed. But 10 years after, we can see that female are most likely unemployed, even ill (around 87%) or not (around 95%) and people with no grade are most often unemployed while those with graduate studies level are most commonly employed but they are no evidence for those with high school or college grades. The proportion of individuals that are not French remains nearly the same over time, for the different states of nature.

# 3 Results

In this section we present results for global non-causality test from health condition to job status, from job status to health condition, and for instantaneous non-causality between health condition and job status. These global non-causality tests are done with each of the three variables of health (health regardless severity of illness, illness with large disability index and illness with large risk of death). We also analyze the dynamic of the causal links over time and the contribution of different states of nature to the causal link through time. We end this section by analyzing the effect of individual characteristics on the causal links[16].

## 3.1 Dynamic of causal links between health condition and job status

Results for global non-causality tests between health condition in general and job status are displayed in the first part of Table 2[17]. Results for global non-causality tests between health condition (illness with large disability index) and job status, and between health condition (illness with large risk of death) are displayed respectively in the second part and the third part of Table 2[18]. As we can see on Table 2, we can conclude at 5% significance level, the rejection

[16]The 45 estimated multinomial logistic models for the cases with health regardless severity of illness and the 84 estimated multinomial logistic models for the cases with illness with large disability index and illness with large risk of death that we use to compute probabilities of states and transition probabilities and their related IIA assumption tests are not discussed herein as we do not focus on. They are used as an alternative approach to estimate probabilities instead of using empirical probabilities. However we present results of only one of them in appendix (table 4).

[17]Note that for non-causality, as we have at most 45 observation periods (from 1961 to 2006), the global causality statistics are computed over 44 time periods. Then, under null hypothesis, the statistic follows a chi-square distribution with 88 degrees of freedom for non-causality from $H$ to $W$ or vice versa, and a chi-square with 176 degrees of freedom for instantaneous non-causality.

[18]Note that for non-causality, as we have at most 41 observation periods due to lack of reported illness with large disability index or large risk of death between 42 and 45 years of professional life, the global causality statistics are computed over 41 time periods. Then under null hypothesis, the statistic follows a chi-square

of null hypothesis of non-causality from health condition to employment and vice versa. So, in general health condition causes job status and vice versa. When we consider illness with large disability index, we find that health causes job status but job status does not cause health at 5% level. However, there is a weak evidence of a causal link from job status to health (at 10% level). For illness with large risk of death, health condition causes job status but there is no evidence that job status causes health condition, even at 10% level. These findings are consistent with the previous literature (see Delattre et al., 2015 for further details).

The null hypothesis of instantaneous non-causality can not be rejected at 5% significance level but is rejected at 10% significance level. This finding is consistent with previous literature (Cai, 2010). It means that there are weak evidence that job status events cause instantaneously health status and vice versa. This can be analysed as health condition does not strongly matter for the current job status as it does for the next job status. The same analysis can be made for the effects of job status on health condition. Thus, health effects on job status and job status effects on health condition are not strongly instantaneous. The weak evidence of instantaneous causality between health and job can be the fact of job protection and adjustment issues. Firstly, employees with long term employment contract are less vulnerable to health events than those with short term employment. Secondly, the effects of unemployment on health is weakly instantaneous as a short term unemployment can not strongly affect the financial condition and the ability to cover health expenditures. If we consider illness with large disability index or illness with large risk of death, we find that there is no instantaneous causal link between health condition and job status. As we can see on Figure 3[19], non-causality from health condition to job

distribution with 82 degrees of freedom for non-causality from $H$ to $W$ or vice versa, and a chi-square with 164 degrees of freedom for instantaneous non-causality.

[19] Figure 3(a) shows the dynamic of causal links from health condition (illness regardless severity) to job status and from job status to health condition (illness regardless severity). Figure 3(b) shows the dynamic of causal links from health condition (illness with large disability index) to job status and from job status to health condition (illness with large disability index). At each time period in individual professional life, we compute and represent the values of the Kullback causality measure $\sum_{h=0}^{1}\sum_{w=0}^{1} 2I\hat{\pi}_{t-1}(w,h)\hat{C}_{W\,to\,H}(t,w,h)$ and $\sum_{h=0}^{1}\sum_{w=0}^{1} 2I\hat{\pi}_{t-1}(w,h)\hat{C}_{H\,to\,W}(t,w,h)$ that have asymptotically a chi-square distribution with 2 degrees of freedom. Threshold lines for 1%, 5% and 10% are also drawn to allow easy comparison of causality measure to these thresholds at each time of professional life. Areas above threshold lines denote non-causality rejection areas.

Table 2: Global causality tests between health and job

| non-causality | Test Statistic | Threshold at 5% | Threshold at 10% |
|---|---|---|---|
| Part 1 : Illness in general (regardless severity) | | | |
| From $H$ to $W$ | 388.4921 | 110.898 | 105.3723 |
| From $W$ to $H$ | 147.6943 | 110.898 | 105.3723 |
| Instantaneous | 206.0259 | 207.9547 | 200.4315 |
| Part 2 : Illness with large disability index $Hdisab$ | | | |
| From $Hdisab$ to $W$ | 167.139 | 104.1387 | 98.7803 |
| From $W$ to $Hdisab$ | 101.5401 | 104.1387 | 98.7803 |
| Instantaneous | 82.7351 | 194.8825 | 187.5959 |
| Part 3 : Illness with large risk of death $Hrisk$ | | | |
| From $Hrisk$ to $W$ | 173.4521 | 104.1387 | 98.7803 |
| From $W$ to $Hrisk$ | 87.9921 | 104.1387 | 98.7803 |
| Instantaneous | 53.7852 | 194.8825 | 187.5959 |

status and vice versa change over individual professional life. At 5% significance level, when we analyze the smoothed curve, we can conclude that even if health condition generally causes job status on the overall professional life, during the first two years of professional life, the period between the $11^{th}$ and the $17^{th}$ year of professional life and after 42 years of professional life, health condition does not cause job status. At 1% significance level, health condition causes job status only between the $20^{th}$ to $22^{th}$ and the $26^{th}$ to $41^{th}$ year of the professional life. For causal link from job status to health condition, we can see from the smoothed curve that at 5% significance level, job status causes health condition during the $11^{th}$ and the $15^{th}$ years of professional life. But at 10% significance level, we can conclude that during the period between the $9^{th}$ to $17^{th}$ year of professional life, job status causes health condition. When we consider illness with large disability index and illness with large risk of death, as we can see from Figure 3(b) and Figure 3(c) respectively, we find that health condition causes job status only from the $33^{th}$ to $38^{th}$, and from $27^{th}$ to $36^{th}$ year of professional life respectively. However, job status does not cause health condition at any period of professional life when we consider illness with large disability index or illness with large risk of death.

Causal links from health condition to job status and from job status to health condition have inverse trends during professional life. As we can see in Figure 3(a), from the $11^{th}$ to the $17^{th}$ year of professional life, when causal link from health condition to job status tends to be non significant, the causal link from job status to health condition becomes significant. From the $18^{th}$ to the $40^{th}$ year of professional life, we can observe the opposite situation. Causal link from job status to health condition remains not significant when causal link from health condition to job status remains significant with greater significance level. The same conclusion can be observed at the beginning of professional life till 10 years of professional life. During this period, health significantly causes job status but the contrary is not significant. At each time period, we have only one unidirectional causal link that is significant. However, this finding does not remain true when we consider illness with large disability index or illness with large risk of death. Only health condition causes job status after 33 years and after 27 years of professional life respectively in these cases. The results in Figure 3(b) and Figure 3(c) point out the shortness of the periods that are driving the overall causal links (from the $33^{th}$ to the $38^{th}$ year of professional life for illnesses with large disability index, and from the $27^{th}$ to the $36^{th}$ year of professional life for illnesses with large risk of death).

Thus, we can deduce that being healthy matters for job status during the 10 first years (for entering the labour market, so entering work for unhealthy is more difficult than for healthy) and after 17 years of professional life (to stay in the labour market, it means that after 17 years, the job market tends to eject unhealthy workers). However at the middle of professional life (i.e during the $11^{th}$ and $17^{th}$ year), job status is not caused by health condition. Then, we can deduce that job status causes health condition during this period. It means that job status effects on health condition are not immediate and seem to be a delayed phenomenon. It may exist an accumulation process of job status effects on health that becomes significant after 10 years in professional life.

For a better analysis of those causal links, we will compute in the next section, the contribution of each states to the causal links at all periods. We also compute causality measures at individual level and estimate a model that aims to assess which of individual characteristics affect the causal link. These two analysis are done only for the causal links between health condition regardless severity and job status.

## 3.2 Contributions of states to causal links

The aim of this section is to highlight which states are driving the overall causal links at each time period. For this purpose, we compute the contribution of each state to the causal links. For example, for the causality from $H$ to $W$, the contribution of the state $(w, h)$ is $Ctr^{(w,h)}_{H\,to\,W}$ (see Equation 10). Then, 4 contributions are available at each period which sum is equal to one. The dynamic of these contributions are presented in Figure 4 below.

For the causal link from health condition to job status (see Figure 4(a)), the largest contribution to the causal link is the causality measure for the state `"ill and employed"` for each transition during observation period. This contribution is in average 52.3% of the causal link. At the beginning of professional life, this contribution is around 60% and continues growing until the $10^{th}$ year of professional life where it reach a climax of 70%. After 10 year of professional life, this contribution starts declining until the $30^{th}$ year of professional life and remains stable till individual leaves job market. The second most important contribution is the causality measure for the state `" healthy and employed"`, that contributes in average for 23.6% of the causal link. This contribution is the only one that increases during professional life, from 10% at the beginning of professional life to around 35%. The contribution of the causality measure for the state `"ill and unemployed"` is larger at the beginning of professional (around 25%) but it declines very quickly and remains stable around 12% just after 10 years of professional life.

For the causal link from job status to health condition (see Figure 4(b)), the two largest contributions to the causal link are the causality measure for the states `"healthy and employed"` and `"healthy and unemployed"`, both with approximatively 41% of causal link. But the contribution to causality measure for the state `"healthy and unemployed"` remains descending during all professional life while the contribution of causality measure of `"employed and healthy"` grows during the 25 first years of professional life before declining till the exit from the job market. At the beginning of professional life, the contribution of the causality measure of the state `"healthy and unemployed"` is around 55% and it declines progressively and reaches 30%. The contribution of the state `"healthy and employed"` is around 30% at the beginning of professional life, and grows progressively to reach 51% after 25 years of professional life before starting a decline phase to reach 32% at the exit of job market. In relatively low proportion, the contribution of the state `"ill and employed"` remains growing from around 5% at the beginning of professional life to approximatively 15% at the exit of job market.

### 3.3 Contributions of individual characteristics to causal links

This section aims to highlight which individual characteristics affect significantly the causal links. The rationale is to disentangle, through a decomposition of the causality measures, individual characteristics that are driving the causal links. This approach is similar to that of Doorslaer and Koolman (2004) on income-related health inequalities indexes. For this purpose, we compute the causality measures at individual level $\hat{C}_{W\,to\,H}(i,t)$ and $\hat{C}_{H\,to\,W}(i,t)$. These causality measures are the generalized forms of the causality measures in Equation 9. $\hat{C}_{W\,to\,H}(i,t)$ is given by :

$$\hat{C}_{W\,to\,H}(i,t) = \sum_{h=0}^{1}\sum_{w=0}^{1} 2I\hat{\pi}_{i,t-1}(w,h)\hat{C}_{W\,to\,H}(i,t,w,h) \tag{13}$$

Where $\hat{C}_{W\,to\,H}(i,t,w,h)$ is computed by replacing the probabilities $\hat{\pi}_{i,t-1}(w,h)$, $\hat{p}_t((.,s_t^2)|(w,h))$, $\hat{p}_t((s_t^1,.)|(w,h))$ and $\hat{p}_{H,t}(s_t^2|h)$ by their corresponding individual level values. Then we regress these individual level causality measures on the individual characteristics. Estimation results are presented in Table 3. We include the square of age in the regressions in order to account for the fact that it might exist a nonlinear effect of age on causal links.

Our results suggest a significant nonlinear effect of age on both causal links from health to job status and vice versa. Causal link from health to job status is decreasing until 39[20] years old and increasing after that. The causal link from job status to health stills decreasing among professional life. These results are consistent with the trend observed in Figures 2(c), 2(b) and 2(a) about the dynamics of the probabilities of transition. Being ill significantly decreases the causality from health to job status and increases the causal link from job to health. These findings reflect two facts : (i) for younger, being healthy promotes ceteris paribus entering job market (Cai and Kalb, 2006; Benjamin and Wilson, 2005) and (ii) for elders, illness can lead to

[20] This value is calculated by dividing minus the coefficient of age by twice the coefficient of the square of age. $Threshold_{\hat{C}_{H\,to\,W}} = -\frac{-0.70269}{2*0.00891}$ and $Threshold_{\hat{C}_{W\,to\,H}} = -\frac{-0.01933}{-2*0.00073}$

loss of employment (Stern, 1989; Waddell and Burton, 2006), specially for illnesses with high degree of severity or for less protected jobs (as short term job).

Long term employment decreases significantly the causality from health to job while short term employment increases the causality from health to job. However both long and short term employment decrease significantly the causality from job to health. Unemployment decreases the causality from health to job and increases the causality from job to health. These findings are consistent with the previous literature. Researchers highlight that for pecuniary reasons, unemployment reduces individual ability to face health shocks, thus individual health condition (Winkelman and Winkelman, 1998). This result also involves that the negative effect of job status on health condition (Debrand, 2011) is inhibited by the positive one (the pecuniary effect of labour market participation on individual wealth and health).

The lower the school grade is, the higher the causality from health to job is. The causality from job to health is positive for individuals with no grade and negative for individuals with college degree comparatively to individual with graduates studies. However, there is no significant difference between individuals with high school degree, undergraduate studies and those with graduate studies. These findings are consistent with previous literature on the links between health, employment and school grade. It is shown in the literature that individuals with higher school level have good job status and have more possibilities of employment when they are unemployed. Thus, job status is not causal for health in that case. For individuals with lower school level, the underlying intuition is that in the case of unemployment, even if they are healthy, they have less opportunities of employment. Thus, health condition is less causal for job status in that case. By including interaction terms between gender and education level, we find that contrarily to females, for males, having a school grade reduces the causal link from job to health. However, for the causal link from health to job, we do not find significant discriminant effects between males and females in terms of school grade.

The causality measure from health to job is lower for individuals in couple and for males, and for foreigners. But we find a significant discriminant effect between males and females in couple. The causal link from health to job is higher for males in couple and for males with higher number of children than for females with the same characteristics. Turning to causal link from job to health, we find that this causal link is higher for males, foreigners, individuals in couples and those with higher number of children. However, contrarily to females with the same characteristics, males with higher number of children, or males in couple have a significantly lower causal measure from job to health. These results generalize Cai's (2010) findings. In addition to the fact that with the same characteristics, job is negatively causal for male's health than female's one, we also show evidence that the reciprocal is true : health condition is negatively causal for male's job status.

## Conclusion

The literature on health and job status highlights that there are reasons to suppose that the causal links between health condition and job status do not remain stable among professional life. Thus, besides of the global causal links on the overall professional life, we should test for causal links at each period of professional life. This paper explores a non-parametric approach based on the Kullback causality measures (by Gouriéroux et al., 1987) to test for both Granger instantaneous and lag non-causality between health condition and job status. This approach has the advantages to be more robust than the traditional parametric framework, to give an assessment of the dynamic of causal links between the two outcomes as well as the overall causal links, and to estimate the effect of individual characteristics on causal links. Thus, we complete an innovative causality analysis that can not be done by the usual parametric framework.

Our results confirm the findings in literature that both health condition and job status are causal for each others with a relative high significant level. But if we focus on illness with large disability index or illness with large risk of death, we only conclude to a significant unidimensional causal link from health condition to job status. We also find a weak evidence of instantaneous causal link between health condition and job status. We highlight that the causal link from job status to health condition is significant only between the $11^{th}$ and the $17^{th}$ year of professional life, while only at the same period causal link from health to job status becomes insignificant. These results are consistent with our methodological approach in which we assumed that causal links are not homogeneous among professional life. We also highlight individual characteristics effects on the causal links, that is an original and innovative approach in causal links analysis. The results of this analysis are consistent with previous literature. We find a negative effect of unemployment on the causality from job to health and a positive effect of unemployment, school grade and gender on the causality from health to job. We also find that both causal links from job to health and from health to job are higher at the beginning and at the end of professional life.

This paper enhances the common understanding of the causal links between health condition and job status. Our paper, by the use of a robust approach, clearly gives periods of professional life from which health events cause job events and vice versa in France. It also highlights which individual characteristics rise both causal links. Policy makers should account for these periods and characteristics for public policies in health and employment.

# References

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

Balcilar M., Gungor H., and Hammoudeh S. (2015). The time-varying causality between spot and futures crude oil prices: A regime switching approach. *International Review of Economics and Finance*, 40:51–71. DOI : 10.1016/j.iref.2015.02.008.

Barnay, T. (2016). Health, work and working conditions: a review of the European economic literature. *European Journal of Health Economics*, 17:693–709. DOI: 10.1007/s10198-015-0715-8.

Barnay T., and Legendre F. (2012). Simultaneous causality between health status and employment status within the population aged 30-59 in France. *Travail, Emploi et Politiques Publiques, Working Paper number 2012-13.*

Benjamin, K. and Wilson, S. (2005). Facts and misconceptions about age, health status and employability. *Health and Safety Laboratory.*

Bouezmarni T., Rombouts J. V.K., and Taamouti A. (2012). Nonparametric copula-based test for conditional independance with applications to Granger causality. *Journal of Business and Economic Statistics.*

Bouissou M. B. Laffont J. J. and Vuong Q. H. (1986). Tests of noncausality under Markov assumptions for qualitative panel data. *Econometrica.*

Bound, J. (1991). Self-reported vs. objective measures of health in retirement models. *National Bureau of Economic Research.*

Cai, L. (2010). The relationship between health and labor force participation : Evidence from a panel data simultaneous equation model. *Labour Economics.*

Cai, L. and Guyonne, K. (2006). Health status and labour force participation : evidence from Australia. *Journal of Health Economics.*

Caroli E. and Bassanini A. (2015). Is work bad for health? the role of constraint vs choice. *Annals of Economics and Statistics*, (119-120):13–37.

Caroli E. and Godard M. (2014). Does job insecurity deteriorate health ? *Health Economics*, 25:131–147. DOI: 10.1002/hec.3122.

Case, A. and Deaton, A. (2005). Broken down by work and sex : How our health declines. *National Bureau of Economic Research.*

Coe, N. B. and Zamarro, G. (2008). Retirement effect on health in Europe. *RAND Labor and Population working paper series.*

Debrand, T. (2011). L'influence des conditions de travail sur les dépenses de santé. *IRDES Working paper.*

Delattre E., Moussa R. and Sabatier M. (2015). Health condition and job status interactions : some econometrical evidences of causality on a French longitudinal survey. *THEMA Working Paper.*

Doorslaer, E. V. and Koolman, X. (2004). Explaining the differences in income-related health inequalities across european countries. *Health Economics*, 13:609–628. DOI:10.1002/hec.918.

Firouz, F. (2011). Causal relationship between energy consumption (ec) and gdp: A markov-switching (ms) causality. *Energy*, 36:4165–4170. DOI : 10.1016/j.energy.2011.04.027.

Gouriéroux C., Monfort A., and Renault E. (1987). Kullback causality measures. *Annales d'Economie et de Statistique.*

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica.*

Hann, P. and Myck, M. (2009). Dynamics of health and labor market risks. *Journal of Health Economics.*

Hsiao, C., Mountain, D., Chan, M. W. L., and Tsui, K. Y. (1989). Modeling ontario regional electricity system demand using a mixed fixed and random coefficients approach ? *Regional Science and Urban Economics*, 19:565–587.

Lakey J., Mukherjee A., and White M. (2001). Youth unemployment, labour market programmes and health. *Policy Studies Institute.*

Michaud P.C., and Van Soest A. (2008). Health and wealth of elderly couples: causality tests using dynamic panel data models. *Journal of Health Economics*, 27(5):1312–1325.

Murphy, G. C. and Athanasou, J. A. (1999). The effect of unemployment on mental health. *Journal of Occupational and Organizational Psychology.*

Salm, M. (2009). Does job loss cause ill health ? *Journal of Health Economics*, 18(9):1075–1089. DOI: 10.1002/hec.1537.

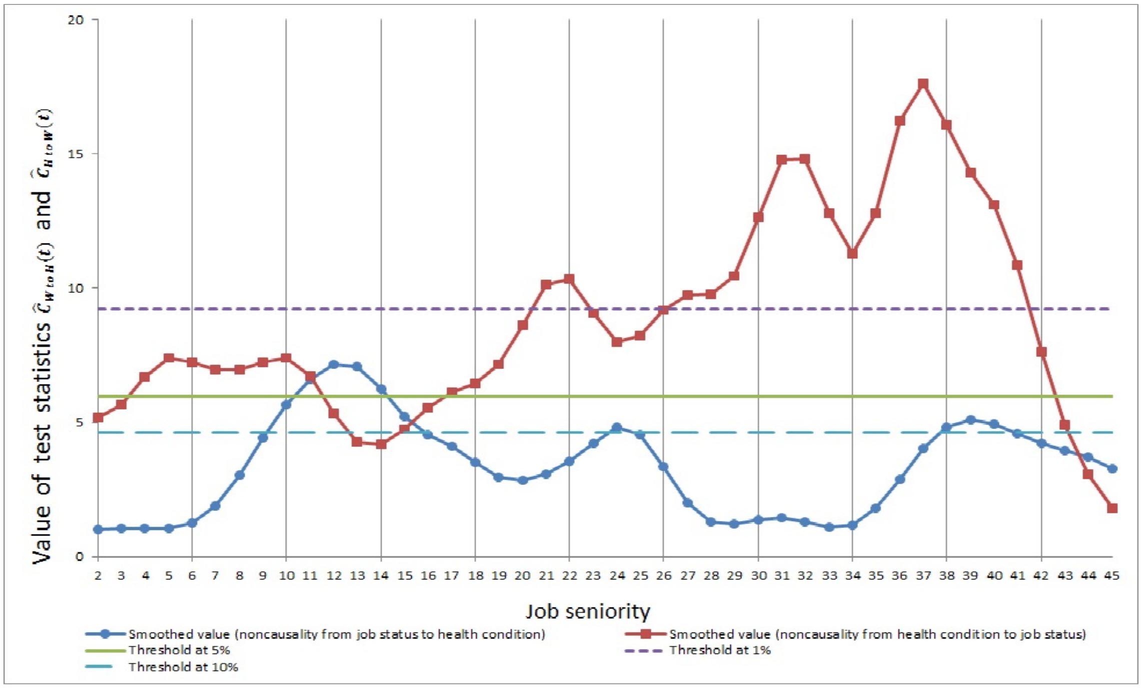
Stern, S. (1989). Measuring the effect of disability on labor force participation. *The Journal of Human Resources.*

Waddell, G. and Burton, K. (2006). Is work good for your health and well-being? *The Stationery Office.*

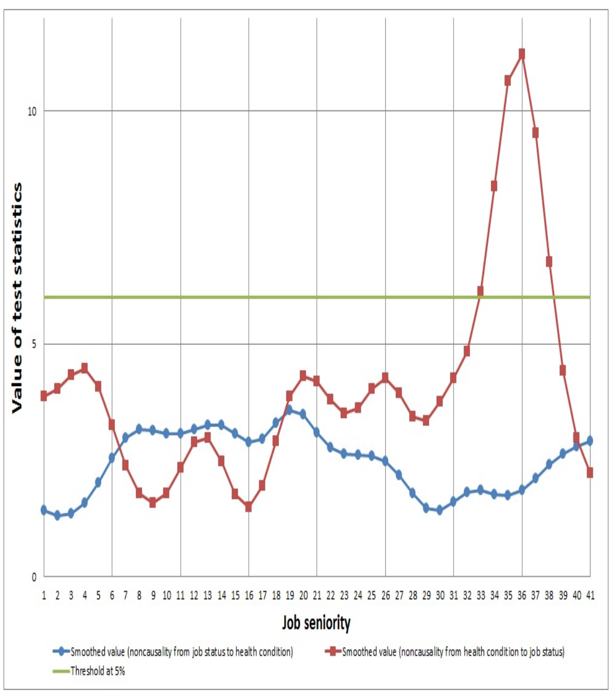
Weinhold, D. and Nair-Reichert, A. (2000). Causality tests for cross-country panels : new look at fdi and economic growth in developing countries. *Oxford Bulletin of Economics and Statistics, 63(2), 153-171.*

Winkelmann, L. and Winkelmann, R. (1998). Why are the unemployed so unhappy? Evidence from panel data. *Economica*, 65:1–15.

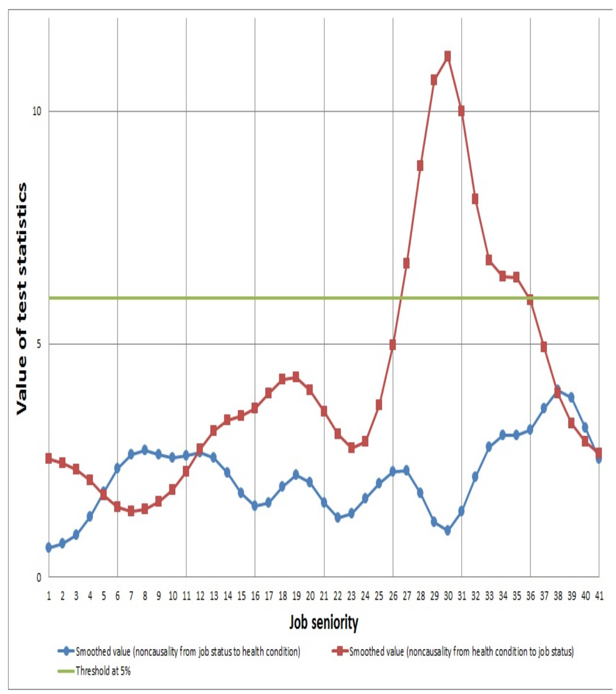
# Appendices




(a) Dynamic of causality links between health (regardless severity) and job

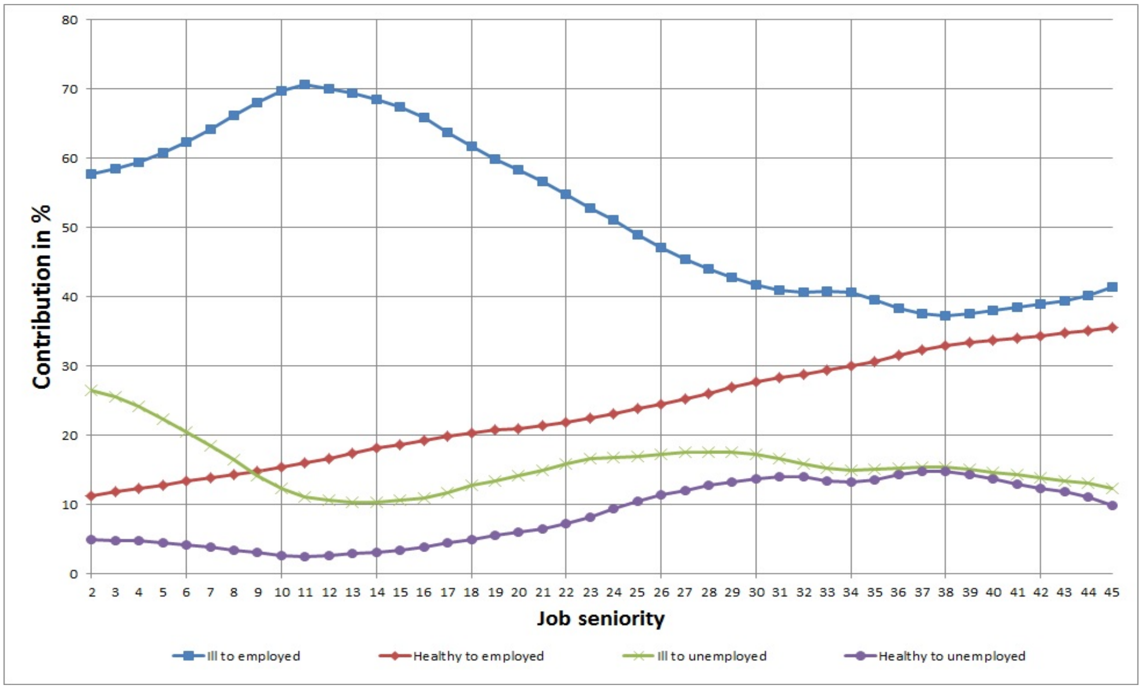



(b) Dynamic of causality links between health (illness with large disability index) and job

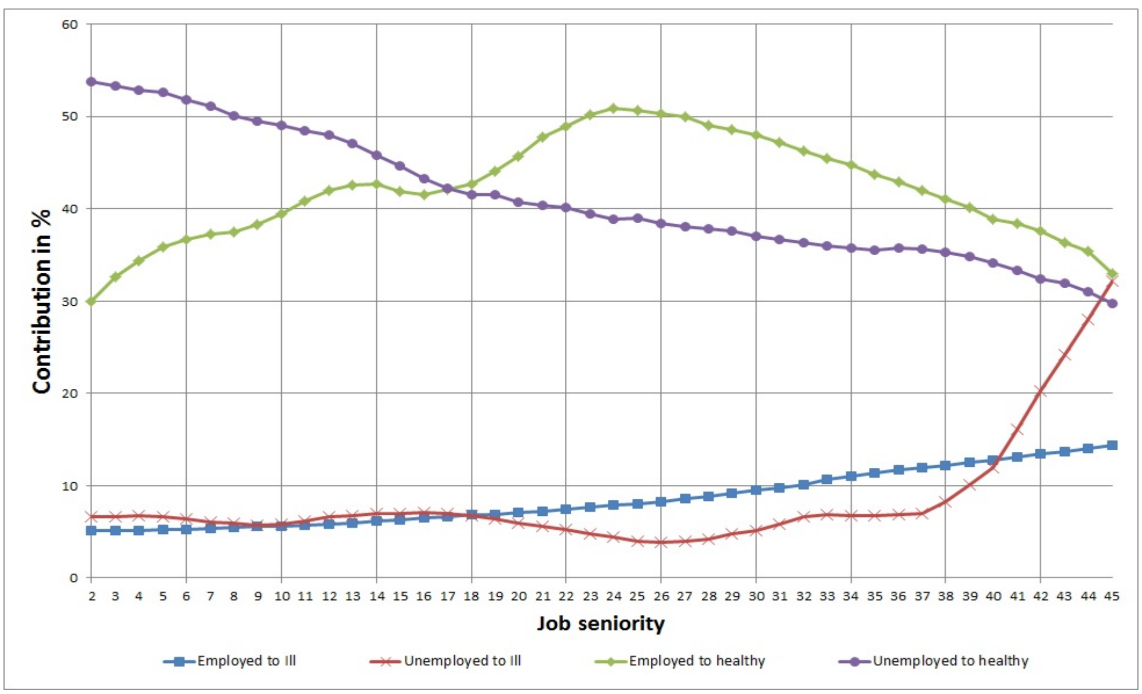



(c) Dynamic of causality links between health (illness with large risk of death) and job

Figure 3: Dynamic of causality links between health and job




(a) Dynamic of contribution of states to causality from Health condition to Job status




(b) Dynamic of contribution of states to causality from Job status to Health condition

Figure 4: Dynamic of contribution of states to causality links between health and job

Table 3: Effects of individual characteristics on causal links

| Variable | $\hat{C}_{H\,to\,W}$ | $\hat{C}_{W\,to\,H}$ |
|---|---|---|
| Age | $-0.70169^{***}$ (0.01184) | $-0.01933^{**}$ (0.00954) |
| Square of age | $0.00891^{***}$ (0.00015) | $-0.00073^{***}$ (0.00012) |
| Ill (ref = healthy) | $-0.54105^{***}$ (0.05798) | $2.66832^{***}$ (0.03716) |
| Job status | ref = inactive (out of labour market) | |
| Long term employee | $-0.40362^{***}$ (0.05590) | $-2.88559^{***}$ (0.04348) |
| Short term employee | $0.96165^{***}$ (0.05810) | $-2.24175^{***}$ (0.04769) |
| Unemployed | $0.35830^{***}$ (0.09095) | $-0.17121^{**}$ (0.07656) |
| School grade | ref = graduate studies | |
| No grade | $2.89954^{***}$ (0.31731) | $0.78338^{***}$ (0.09236) |
| College degree | $1.59345^{***}$ (0.21477) | $-0.45652^{***}$ (0.06450) |
| High school degree | $4.42229^{***}$ (0.22256) | $-0.05982$ (0.06684) |
| Undergraduate degree | $0.58470^{***}$ (0.23517) | $0.0426$ (0.07202) |
| Male | $-0.97484^{**}$ (0.22044) | $0.39750^{***}$ (0.07544) |
| Male*School grade | $-0.13192$ (0.09885) | $-0.14146^{***}$ (0.02917) |
| Not French | $-1.30767^{***}$ (0.18787) | $0.55205^{***}$ (0.05268) |
| Number of children | $0.03517$ (0.02821) | $0.42344^{***}$ (0.01849) |
| Male*Number of children | $0.25469^{***}$ (0.03401) | $-0.23654^{***}$ (0.02509) |
| Couple | $-0.76347^{***}$ (0.05744) | $0.26812^{***}$ (0.04507) |
| Male*Couple | $0.55564^{***}$ (0.08616) | $-0.45119^{***}$ (0.06814) |
| Intercept | $15.60896^{***}$ (0.27536) | $4.76891^{***}$ (0.17739) |
| $\rho_u$ (variance due to $u_i$) | $0.41508^{***}$ | $0.0283^{***}$ |

Obs. = 261,654 ; Number of individual = 10,811

Standard errors are in parenthesis; $***$ : significant at 1%

$**$ : significant at 5%; $*$ : significant at 10%

Table 4: Multinomial logistic model at the $6^{th}$ year of professional life

| $Variables^{+++}$ | $IE^{++}$ | $IU$ | $NIE$ | $NIU$ |
|---|---|---|---|---|
| $IE_{-1}$ | | $-0.9598^{**}$ (0.4399) | $-8.6521^{***}$ (0.2619) | $-7.2475^{***}$ (0.7247) |
| $IU_{-1}$ | | $3.9876^{***}$ (0.4705) | $-7.9575^{***}$ (1.0272) | $-4.4715^{***}$ (1.0384) |
| $NIU_{-1}$ | | $3.868^{***}$ (1.1546) | 0.4021 (1.0116) | $5.1677^{***}$ (1.0117) |
| $NIE_{-1}$ | reference | | | |
| $Age$ | | $-0.0944$ (0.07) | $-0.025$ (0.05) | $-0.1525^{***}$ (0.0567) |
| $Male$ | | $-0.9894^{***}$ (0.2907) | $-0.0027$ (0.1969) | $-1.1521^{***}$ (0.2348) |
| $NotFrench^{+}$ | | $-0.1359$ (0.4487) | $-0.1632$ (0.3028) | 0.5391 (0.3294) |
| $Child$ | | $0.4673^{***}$ (0.16) | $-0.2692^{**}$ (0.1265) | $0.4874^{***}$ (0.1397) |
| $Rural$ | | 0.1124 (0.2766) | 0.057 (0.2149) | $-0.336$ (0.249) |
| $Medicaldensity$ | | 0.0152 (0.0109) | $-0.0033$ (0.0082) | $-0.005$ (0.0096) |
| $Unemploymentrate$ | | $-0.2251^{*}$ (0.1254) | 0.0214 (0.0943) | 0.0241 (0.1097) |
| $Nograde$ | | 0.8162 (0.713) | 0.2558 (0.5787) | 0.2323 (0.6376) |
| $Collegegrade$ | | 0.5592 (0.5618) | 0.184 (0.3884) | $-0.0416$ (0.4435) |
| $Highschoolgrade$ | | 0.0054 (0.5391) | $-0.3169$ (0.3545) | $-0.5023$ (0.414) |
| $Undergraduatestudies$ | | $-0.5732$ (0.6303) | $-0.0263$ (0.363) | $-0.1303$ (0.4312) |
| $Graduatestudies$ | reference | | | |
| $intercept$ | | $-1.0454$ (1.9303) | $5.6083^{***}$ (1.4051) | $5.4912^{***}$ (1.5863) |
| Number of observations = 10,130 ; Pseudo R2 = 0.724 ; Log-likelihood = -2118.46 | | | | |

$+++$ IE: Ill and Employed, IU : Ill and Unemployed, NIE : healthy and Employed, NIU : healthy and Unemployed ; $*$ 10% significance level, $*5$ 5% significance level, $***$ 1% significance level; $+$ : Refers to individual's nationality; $++$ Base outcome