UNIVERSITÉ
de Cergy-Pontoise

PhD Thesis

Causality in social sciences : some applications in micro-econometrics applied to health and labour economics

Richard Kouamé MOUSSA

ThEMA, Université de Cergy-Pontoise

ENSEA, Côte d'Ivoire.

Committee

Alain CARPENTIER, Referee, Research Director, INRA Rennes

Eric DELATTRE, Thesis supervisor, Associate Professor, Université de Cergy-Pontoise

Jacques Loesse ESSO, Member, Professor, ENSEA Abidjan

Christophe MULLER, President of the committee, Professor, Université de Aix-Marseille

Diane SKATUN, Referee, Professor, HERU University of Aberdeen

December 15, 2016

Remerciements

Mes remerciements vont à l'endroit de :

- Mon directeur de thèse, Eric Delattre pour la confiance et l'excellent encadrement dont il m'a fait profiter.
- Ma co-autrice du second chapitre, Mareva Sabatier pour la bonne collaboration.
- Aux illustres membres du jury pour l'intérêt accordé à ce travail.
- A mes parents, mes amis et collègues de l'ENSEA pour le soutien sans faille dont j'ai bénéficié.
- A ma très chère Anna Amy Soumahoro.

Mentions légales

L'Université de Cergy-Pontoise (UCP) n'entend accorder aucune approbation ni réprobation aux propos contenus dans cette thèse. Ces propos doivent être considérés comme propres à leur auteur.

Introduction

Background

The establishment of causality among social phenomena is one of the most important and controversial topics in social sciences. These social phenomena can be defined in terms of exposure variables and outcomes of interest. Establishing causality requires that researchers observe variation in the exposure variable suspected to induce the change in the outcome of interest, and to measure the change in the outcome variable. It is important that confounding be appropriately controlled for to avoid establishing spurious relationship between exposure and outcome variables. The literature suggests two research approaches to exploring causality: i. experiments and ii. statistical research.

Experiments are conducted in controlled environments such that results are not influenced by confounding factors. In other words, other known and unknown variables different from the causative variable. As an example, these experiments are used in clinical studies where patients are randomly assigned to active and control groups such that the known and unknown variables that might influence the outcome of interest are equally distributed. Under these circumstances, the change observed in the outcome variable is directly attributed to the exposure variable.

For secondary data based research, statistical techniques can used to control for potential confounders.

Problem statement and potential solutions

In social sciences, researchers face several challenges in establishing causality, which require the improvement of existing statistical techniques or developing new ones. The first challenge encountered when establishing causality is the issue of endogeneity, i.e. other variables than the exposure and outcome variables may simultaneously affect the exposure variables and outcomes of interest or some exposure variables may be highly correlated, leading to a misidentification of the real causal links. Endogeneity can be illustrated with the establishment of causal links between health and job. Health status is measured as a self-reported variable. For job, researchers use employment status or job market position. Many individual characteristics such school level, gender, marital status, or age affect simultaneously health and job outcomes. In addition, both health status and job status affect each other.

Second, researchers have to account for heterogeneity while assessing causal links between exposure and outcome variables. Indeed, a healthy individual has not the same probability to enter job market at each age. Since long-term unemployment reduces one chances to enter job market. Job effects on individual health are not instantaneous and depend on the industry sector. Since the effects of job painfulness or others noisy facts can be cumulative. Thus, the hypodissertation of homogeneous causal links over time or among individuals can be challenged.

Third, the initial conditions issue must be addressed since it affects the causality measure. Since data on the analyzed social phenomena from their origins are not available, the first observable state is usually considered as initial condition. In some cases, individual may benefit from inheritance or from different initial endowment . Then, with the same characteristics, a different initial condition can affect an individual's path. Reporting severe illness at the end of schooling can reduce one's likelihood to enter the job market, and then reduce his abilities to cover health care expenses. Thus, researchers have to account for initial conditions while analyzing causality and the dynamic of individuals' paths. This begs the question how should initial conditions be considered in the causality analysis? Researchers show that initial conditions must be treated as endogenous since treating initial conditions as exogenous means that they have different data generation processes from the rest of the phenomenon. Being ill or having an employment at the

end of schooling cannot be considered as an exogenous phenomenon since these jobs and health events can be dependent of individual inheritance. Modeling initial conditions as exogenous leads to inconsistent estimations in this case.

Fourth, the causal links can be time dependent or individual specific. Many social policies may be implemented sequentially in regard to time. Thus, the estimated causal links can be different between two time periods. Furthermore, relationships between outcomes of interest are often time dependent. As a result, researchers must account for the dynamic of impact.

Fifth, many variables of interest are qualitative in social sciences. Since usual causality tests in literature focus on continuous outcomes. Apart from that, in social sciences, researchers have to deal with very short panel data, since the collection of specific data is recent. Due to these specificities, researchers in social sciences have to find different approaches to address the issue of measuring causality.

Given the above mentioned challenges, specific and innovative approaches are needed to help researchers to correctly assess the causal impact of policies in social sciences in both ex ante and ex post frameworks. In ex post framework, for a bivariate case, a bivariate probit model can be used. This vectorial model has the advantage to overcome the endogeneity problem. Specific equations can be included in this model for each variable of interest at the initial period. This helps to solve the problem of initial condition. Individual heterogeneity can be accounted for by including individual effects in the dynamic equation. For the initial condition, the individual heterogeneity is accounted for by including a linear combination of the individual effects in each initial condition equation. However, this framework can only be designed to account for homogeneous causal links or for individual specific causal links. To account for the dynamic of causal links, researchers use non parametric methods. One way to account for the dynamic of causal links is to estimate and test the causal links for each transition between two different waves of the data. In that case, a Chow-type test can be used to compare estimated coefficients. Another way is to use causality measures such as Kullback causality measures to assess the dynamic of the causal links. In that case, a probability density function is used and all the transitions probabilities are computed at each period and/or for each individual. In ex ante framework,

researchers have to specify economic models that account for all transmissions channels between the variables of interest. Hence, the causal links are measured by simulations at the optimum (on the steady state).

Upon reviewing the literature we came to the following conclusions: i. There are few studies that have been published that have addressed simultaneously the issues of endogeneity, initial conditions as endogeneous and individual heterogeneity while estimating the causal links between health and job; ii. There are few studies that account for the issue of the dynamic of the causal links between two or more variables and their determinants in social sciences; iii. There are few studies that have investigated the relationship between early retirement behaviour and inter temporal utility function for workers.

Research aim and objectives

The aim of this research is to investigate the micro-econometric approaches used to test the causality in various fields in social sciences. These investigations will cover both ex ante and ex post impact assessment methods. To achieve this goal, we set out the following objectives. The first objective is to conduct a thorough literature review on causality test methods. For each method, a technical description, their conditions of application as well as their advantages and disadvantages are provided. The second objective is to establish whether or not there exists bidirectional causality between health status and job status. The third objective is to investigate the dynamic of the causal links between health status and job status. The fourth objective is to investigate the effects between individualsï¿½ health status, estate and preferences about early retirement decisions.

The research was conducted in light of the following questions:

1. What are the causality test or measurement methods existing in the literature ?
2. How should endogeneity, individual heterogeneity and initial conditions be simultaneously accounted for when testing for causality between two or more outcomes of interest?
3. How, when relevant, should the hypodissertation of homogeneous causal links and the

determinants of the nature of causal links be challenged?

4. How to predict workers' probability to retire early based on their health status, financial conditions and preferences for future?

This dissertation contributes towards a better understanding of causality in social sciences. Specifically, we point out how to account for endogeneity, individual heterogeneity, initial conditions and persistence when measuring causality. This leads to a better assessment of causality and ensures robust estimations. We also make a contribution to the literature about the non-homogeneity of causal links holds. We test whether the causal links are time-specific. We propose an alternative specification and a better understanding of the dynamic of the causality over time. Another main contribution of this dissertation relates to the human capital theory. We introduce an innovative way to estimate health investment that is the health production function. We also propose an approach to estimate how much job deteriorates health stock. These two functions are very helpful to better understand the determinants of early retirement decision in elderly workers population.

Framework

Causality in social sciences is of significant interest for policy makers as well as other stakeholders . Given that implemented programs need to be assessed, several methods have been developed to measure the causal impact of these implemented programs. This dissertation firstly makes an overview of causality measurement methods by focusing on their technical issues. Then, we apply some of the described methods to specific fields such as health and job relationships, or early retirement behaviour.

To analyze the causal relationship between health status and job status, we used a French longitudinal dataset: the Health and Professional Path Survey (Enquête Santé et Itinéraire Professionel, SIP-2006). The data was collected in 2006 as a retrospective panel data that cover individual health and professional paths. We used the Survey on Health, Ageing and Retirement in Europe (SHARE) for the early retirement decision analysis. These data are collected every two years starting from 2004. Further details about the data are provided in each chapter.

Organization of the dissertation

This research is organized into five overlapping chapters. An overview of these papers is given below.

The first chapter entitled **"Causality measurement and tests in economics and econometrics: a methodological survey "** gives an overview of technical approaches to test for causality reported in the literature. In this chapter, several definitions of causality are given. We also discriminate between causality and some other concepts such as correlation. Before introducing technical methods to test for causality, we present an overview of public policy assessment methods. Then, the technical issues of causality testing methods are presented according to the simple time series framework, the panel data framework, and the non-parametric framework. We end the chapter by reviewing how causality is tested in the health economics framework.

The second chapter is entitled **"Health status and job status interactions : econometric evidence of causality from a French longitudinal survey "**. This chapter aims to test for Granger causality between health status and job status. For that purpose, we use a bivariate probit model. Our model has the advantage to overcome endogeneity problems when analyzing relationship between health status and job status. We also account for individual heterogeneity and initial conditions that have been proven to be determinant for health and job paths. The main results of this chapter are the bidirectional causality between health status and job status. We also highlight the important roles of individual heterogeneity and initial conditions in this relationship.

The third chapter is entitled **"Dynamic interactions between health and employment statuses: a non-parametric analysis "**. The main technical achievement of this chapter is that we test for Granger causality between health status and job status at each period of professional life, as well as global causality. We prove that the causal links are time-dependent and we highlight individual characteristics that affect these causal links. We used the Kullback causality measures developed by Gouriéroux et al (1986) that is based on the transition probabilities between health and job events. The findings are consistent with those of chapter two. However,

we show that at the beginning (before 11 years) and at the end (after 17 years) of professional life, health status causes job status while only during the same period, job status does not cause health status.

The fourth chapter entitled **"What are the main reasons for early retirement decisions? Lessons from a dynamic structural modeling "** uses the ex-ante impact assessment framework to predict individual early retirement behavior based on health, estate and preferences for future. To that end, we specify and estimate an economic model that is based on individual utility function. In this framework, individual decisions are related to how health stock is consumed (health consumption function) and how productive health expenditures are (health production function). We show that our model disentangles between three groups of workers that are i. workers who choose early retirement, ii. workers who do not choose early retirement, and iii. workers who are uncertain about early retirement. In robustness analyses, we show that our predicted early retirement behavior is a good predictor of observed early retirement.

The last chapter concludes and gives the limitations of this dissertation. It also gives suggestions for future researches. The last part of the dissertation is dedicated to appendices. The first section of these appendices is entitled **"On the estimation of causality in a bivariate dynamic probit model on panel data with Stata software: A technical review "**. It is a research manuscript that deals with technical issues related to the model estimated in chapter 2. A description of the Gauss-Hermite quadrature method for numerical approximation of integral functions is made. In the bivariate case, gradient vector and Hessian matrix are calculated and their implementation within STATA software is described. We end this section with robustness check based on simulated data. The second section of these appendices covers all technical detail about space state model estimation. This appendix is related to the estimations made in chapter 4.

Keywords: Causality; lag and instantaneous causality; Markov; Health economics; Labour economics; Early retirement; Human capital theory; Micro-econometrics

Abstract in French

Contexte

L'établissement de la causalité entre différents phénomènes sociaux est d'une importance capitale mais également un sujet de controverse en sciences sociales. Ces phénomènes sont soit considérés comme variables d'intérêt ou variables de contrôle dans l'analyse. Etablir la causalité requière donc la prise en compte de variables de contrôle susceptibles d'influencer les variables d'intérêt et de mesurer la variation des variables d'intérêt due aux variables de contrôle. Ainsi, il convient de contrôler correctement les facteurs pouvant engendrer de l'endogéneité ainsi que tout autre facteur de confusion afin d'éviter d'établir des relations fallacieuses entre les variables d'intérêt. Deux approches sont suggérées dans la littérature pour analyser la causalité : i. l'expérimentation, ii. les techniques statistiques utilisant des données non contrï¿½lï¿½es.

L'expérimentation est généralement conduite dans un environnement dans lequel tous les facteurs pouvant engendrer de l'endogénéité ainsi que les autres facteurs de confusion sont contrôlés afin qu'ils n'influencent pas les causalités estimées. Ainsi, les variations observées sont directement interprétables en termes de causalité.

Pour les travaux de recherches basés sur les autres données non expérimentales, des techniques statistiques sont utilisées pour tenir compte des facteurs pouvant engendrer de l'endogéneité ainsi que tous les autres facteurs de confusion.

Problématique et idées de solutions

En sciences sociales, les chercheurs font face à de nombreux défis dans l'établissement de la causalité. Ce qui requière l'amélioration des techniques économétriques existantes et le développement de nouvelles techniques.

Le premier défi auquel il faut faire face est le problème de l'endogénéité. En effet, plusieurs variables d'intérêt peuvent s'expliquer simultanément. Plusieurs exemples existent dans la

littérature, tels que les liens emploi-santé.

Le second défi est relatif à l'hétérogénéité. En effet, dans l'analyse des relations emploi-santé, on peut constater qu'un individu, même en bon état de santé, n'a pas les chances d'avoir un emploi quelque soit la période dans sa vie professionnelle. Aussi, les effets de la pénibilité du travail et des autres facteurs de nuisance sur la santé apparaissent avec le temps. Ainsi, l'hypothèse de l'homogénété de la mesure de causalité peut être remise en cause.

Le troisième défi est relatif aux conditions initiales. En effet, les phénomènes ne sont généralement pas observés depuis leurs origines. Ainsi, la première observation est considérée comme la condition initiale. Cependant, les individus peuvent bénéficier de différentes dotations ou conditions initiales pouvant influencer la dynamique de leurs trajectoires. Aussi, supposer les conditions initiales exogènes reviendrait à assumer que les phénomènes observés à la date initiale sont générés par un processus différent de celui qui génère les observations aux autres dates. Ce qui peut conduire à des estimations non consistantes.

Le quatrième défi est relatif à la stabilité des relations de causalité. En effet, la mise en oeuvre séquentielle de plusieurs politiques économiques peut influencer la dynamique de la causalité. Ainsi, les chercheurs doivent la mesurer de manière dynamique.

Le cinquième défi est relatif au type de données utilisées en sciences sociales. En effet, en sciences sociales, les phénomènes étudiés sont mesurés par des variables qualitatives tandis que la littérature sur la causalité se focalise sur des données quantitatives. De plus, les bases de données sont généralement d'horizon temporel très court, ce qui constitue une difficulté supplémentaire.

Etant donnés les défis sus-mentionnés, des approches spécifiques et innovantes sont nécéssaires pour mesurer correctement la causalité en sciences sociales, et ce, en évaluation ex ante comme en évaluation ex post. En évaluation ex post, dans le cas bivarié par exemple, un modèle probit bivarié peut être utilisé. Cette spécification vectorielle a l'avantage de prendre en compte le problème de l'endogénété. Des équations spécifiques peuvent être incluses dans ce modèle pour chacune des variables d'intérêt à la date initiale. Ceci permet de régler le problème de

la prise en compte des conditions initiales. L'hétérogéneité individuelle est prise en compte en incluant des effets individuels dans les équations de dynamique. Pour les conditions initiales, une combinaison linéaire des effets individuels est incluse dans chaque équation. Cependant, ce cadre est idéal pour la prise en compte de la causalité lorsque celle-ci est supposé homogène. Pour tenir compte de la dynamique de la causalité, les chercheurs utilisent des méthodes non paramétriques. La première méthode utilisée pour tenir compte de la dynamique de la relation de causalité consiste à estimer et à tester la relation de causalité pour chaque transition entre les différentes vagues de l'enquête. Dans ce cas, un test du type Test de Chow peut être utilisé pour comparer les coefficients. La seconde approche consiste à utiliser les mesures de causalitè basées sur les informations de Kullback. Dans ce cas, la fonction de densité est utilisée et les probabilités des transitions sont calculées à chaque date et/ou pour chaque individu. En évaluation ex ante, les chercheurs doivent spécifier des modèles économiques qui tiennent compte des chaînes de transmissions entre les variables d'intérêt. Ainsi, les liens de causalité sont mesurées à l'état d'équilibre.

En faisant la revue de littérature, on se rend compte que : i. il y a peu d'études publiées qui tiennent compte simultanément de l'endogénéité, des conditions initiales et de l'hétérogénéité individuelle en estimant la causalité; ii. il y a peu d'études qui traitent de la dynamique des rélations de causalité et de leurs déterminants en sciences sociales; iii. il y a peu d'études qui traitent des relations de comportement de retraite anticipée des employés et leurs préférences intertemporelles.

Objectifs et contributions

L'objectif de cette recherche est d'examiner les approches micro économétriques utilisées pour tester la causalité dans divers champs des sciences sociales. Cet examen couvrira aussi bien le cadre de l'évaluation ex ante que celui de l'évaluation ex post. Pour atteindre cet objectif, les objectifs spécifiques suivants sont formulés. Le premier objectif spécifique est de réaliser une revue exhaustive des méthodes de test et de mesure de la causalité. Pour chaque méthode, une description technique est effectuée et les conditions d'application ainsi que les avantages et limites de la méthode sont présentés. Le second objectif spécifique est de verifier l'existence d'une causalité bidirectionnelle entre santé et emploi. Le troisième objectif spécifique est d'étudier

la dynamique des liens de causalité entre santé et emploi. Le quatrième objectif spécifique est d'examiner les liens entre la santé, le patrimoine, les préférences pour le futur et les décisions de retraite anticipée des employés.

Les travaux ont été conduit de sorte à répondre aux questions spécifiques suivantes :

- Quels sont les méthodes de test et de mesure de la causalité dans la littérature ?
- Comment tenir compte simultanément de l'endogénéité, de l'hétérogénéité individuelle et des conditions initiales en testant la causalité entre des phénomènes d'intérêt ?
- Comment analyser, lorsque cela est nécessaire, la dynamique des relations de causalité ainsi que les déterminants de cette dynamique ?
- Comment prédire la probabilité de retraite anticipée en se basant sur la santé, le patrimoine et les préférence pour le future des employés ?

Cette recherche contribue à améliorer la compréhension et l'analyse de la causalité en sciences sociales. De manière spécifique, nous montrons des voies pour tenir compte compte simultanément de l'endogénéité, de l'hétérogénéité individuelle et des conditions initiales en testant la causalité entre des phénomènes d'intérêt. Ce qui conduit à des estimations plus robustes. Nous apportons également une contribution à la littérature sur l'analyse de la dynamique des relations de causalité. Nous proposons une spécification alternative ainsi qu'une meilleure compréhension de la dynamique des relations de causalité. Une autre contribution majeure de cette recherche porte sur la théorie du capital humain. Nous introduisons une approche innovante pour estimer les investissements en santé à travers une fonction de production de santé. Nous proposons également une approche pour estimer les dépréciations de santé dues à l'emploi. Ces deux dernières fonctions sont par la suite utilisées pour améliorer la compréhension de la décision de retraite anticipée au niveau des employés.

Cadre et données

La mesure de la causalité en sciences sociales est d'une importance capitale pour les pouvoirs publics ainsi que les autres décideurs. Etant donné que les projets et programmes mis en oeuvre doivent être évalués, plusieurs méthodes doivent être développées en vu d'estimer l'effet causal

de ces projets ou programmes. Dans cette recherche, nous faisons une revue des méthodes de test et de mesure de la causalité tout en mettant un accent particulier sur les aspects techniques de celles-ci. Ensuite, certaines des méthodes décrites sont appliquées à des domaines particuliers tels que les relations entre santé et emploi ou les comportements de retraite anticipée.

Pour analyser les relations de causalité entre la santé et l'emploi, nous utilisons les données de l'enquête santé et itinéraire professionnel (SIP-2006). Les données sont collectées de manière rétrospective à partir de 2006 en France et couvrent les parcours professionnels et de santé des individus enquêtés.

Nous utilisons les données de l'enquête Santé, vieillissement et retraite en Europe (SHARE) pour l'analyse de la décision de retraite anticipée. Ces données sont collectées dans certains pays européens chaque deux années depuis 2004. Des détails sur les différentes bases de données sont apportés au niveau de chaque chapitre.

Plan et résumé des articles

Cette recherche est organisée en cinq chapitres. Chaque chapitre constituant un article de recherche à part entière mais lié tout de même aux autres. Un résumé de ces articles est donné ci-dessous.

Le premier chapitre intitulé **"Mesures et tests de causalité en économie et en économétrie : une revue méthodologique"** donne une idée générale des approches de tests de causalité dans la littérature. Dans ce chapitre, différentes définitions de la causalité sont données. Nous faisons également la distinction entre la causalité et des concepts proches tels que la corrélation. Avant d'introduire les méthodes de tests de causalité, nous présentons les cadres et méthodes d'évaluation de politiques publiques. Ensuite, les aspects techniques des méthodes de test de causalité sont présentés en les regroupant selon le contexte de séries temporelles simples, de données de panel, et de méthodes non paramétriques.

Le second chapitre est intitulé **"Interaction entre santé et emploi : enseignements économétrique de causalité sur données françaises"**. Dans ce chapitre, nous testons la

causalité au sens de Granger entre la santé et l'emploi. Nous utilisons à cet effet un modèle probit bivarié. Notre spécification a pour avantage de régler le problème d'endogénéité entre santé et emploi. Nous tenons également compte de l'hétérogénéité individuelle et des conditions initiales dont les effets sur les parcours professionnel et de santé ne sont plus à prouver. Le principal résultat de ce chapitre est la preuve de l'existence d'une causalité bidirectionnelle entre santé et emploi. Nous relevons également les rôles importants joués par l'hétérogénéité individuelle et les conditions initiales.

Le troisième chapitre est intitulé **"Dynamique des interactions entre santé et emploi : une analyse non paramétrique"**. La principale contribution de ce chapitre est que nous testons la causalité au sens de Granger entre santé et emploi à chaque période de la vie professionnelle de même que la causalité globale. Nous montrons que les liens de causalité varient dans le temps et nous identifions les caractéristiques individuelles qui l'influencent. Nous utilisons à cet effet la mesure de causalité de Kullback développée par Gouriéroux et al. (1986) qui est basé sur l'analyse des probabilités de transitions entre les évènements de santé et d'emploi. Nos résultats confirment ceux du chapitre 2. De plus, nous montrons qu'au début (avant 11 années) et à la fin (après 17 ans) de la vie professionnelle, la santé cause l'emploi tandis qu'aux mêmes moments, l'emploi ne cause pas la santé.

Le quatrième chapitre intitulé **"Quelles sont les principales raisons de la retraite anticipée ? Enseignements d'un modèle structurel dynamique"** utilise une approche ex ante pour prédire les comportements individuels de retraite anticipée en se basant sur la santé, le patrimoine et les préférences pour le futur. Nous spécifions et estimons à cet effet, un modèle économique basé sur les fonctions d'utilité intertemporelle. Dans ce cadre, les décisions individuelles sont relatives à comment le stock de santé est consommé (fonction de consommation de santé) et à la productivité des dépenses de santé (fonction de production de santé). A partir de ce modèle, nous distinguons selon leurs caractéristiques, les employés en trois groupes que sont i. les employés qui choisissent la retraite anticipée, ii. les employés qui ne choisissent pas la retraite anticipée, et iii. les employés qui sont indécis quant à la retraite anticipée. L'analyse de la robustesse réalisée permet de montrer que notre prédiction du comportement de retraite permet de cerner la décision effective.

Le dernier chapitre conclut et donne les limites de l'étude. Il donne également des suggestions pour des recherches futures. La dernière partie de cette recherche porte sur les annexes. La première section des annexes est intitulée **"Estimation de la causalité via un modèle probit bivarié dynamique sur données de panel avec le logiciel STATA : Une revue technique"**. C'est un article de recherche qui traite tous les aspects techniques liés à l'estimation du modèle utilisé au chapitre 2. Une description de la méthode de quadrature de Gauss-Hermite pour l'approximation numérique des intégrale y est faite. Dans le cas bivarié, le gradient et la matrice hessienne sont calculés et implémentés sur STATA. Cette premiére partie se termine par une analyse de robustesse faite sur données simulées. La seconde partie des annexes couvre tous les aspects techniques liés à l'estimation d'un modèle d'espace état. Cette annexe est lié aux estimations effectuées au chapitre 4.

Chapter 1

Causality Measurement and Tests in Economics and Econometrics: A methodological Survey[1]

[1]Author : **Richard Moussa**, ThEMA-UCP and ENSEA Abidjan

Abstract

This paper makes a survey of causality measurement approaches. We start by presenting the concept and its history and end with methods description. We point out advantages and limits of each approach. We also make a focus on causality measurement approaches in health economics.

Keywords: Causality; lag and instantaneous causality; Markov; Copula

JEL Classification: B23, B41

Introduction

Causality measurement is a large field in science. Since the first definition of the concept, many studies have addressed the issue of the characterization of the causality. Especially in econometrics, many papers deal with causality measurement according to various approaches. However, the first testable definitions were due to Granger' (1969) and Sims' (1972) works. This paper aims to make a survey of definitions and measurements of causality in literature. This chapter is organized as following : Section 1 presents concepts and history, Section 2 focuses on public policy evaluation methods and their related links with causality, Section 3 makes a survey of causality measurement approaches, and Section 4 concludes.

1.1 Definitions and literature

This section aims to present the notion of causality in literature and to highlight some important distinctions with others concepts like correlation.

1.1.1 Concept and history

The concept of causality has widely been discussed in many fields as a specific relationship between two events A and B. The idea was that A causes B if the realization of A is always followed by that of B. However, the formalization took long before generating a testable definition. In

economics and econometrics, the first testable definition is due to the seminal work of Granger (1969). Granger's definition is expressed in terms of predictability. The definition of causality by Granger distinguishes lag causality from instantaneous causality. Let X and Y denote two variables. X causes (Granger lag causality) Y if the whole history of X helps to better predict the current value of Y than if only the history of Y was used. Granger introduced the concept of feedback between two variables. There is a feedback between X and Y if X lag causes Y and Y lag causes X. X instantaneously causes Y if the current value of X helps to better predict the current value of Y than if only the history of Y was used[2]. This definition has then been improved by several works, particularly by that of Sims (1972). Sims's definition of causality seems quite different of that of Granger. Sims considers that X causes Y if the whole history of X helps to better predict the current and future values of Y than if only the history of Y was used. But equivalences have been established between Granger definition of noncausality and that of Sims (See Bouisson et al., 1986 for further details). However, the testable forms of these two definitions are based on linear equations and available for quantitative processes only. Furthermore, these testable forms require the involved variables to be stationary (Granger, 1969). Thus, many other papers have given different specifications to overcome these problems and extend testable definitions of causality.

Causality has then been analyzed as conditional independence (Florens and Mouchart, 1982; Bouisson et al., 1986; Gouriéroux et al., 1987; Adams et al., 2003; Bouezmarni et al., 2012), as orthogonality by Florens and Mouchart (1985), or as correlation between innovations by Pierce and Haugh (1976). The formalization of the definition of non causality in terms of conditional independence is the following one. X does not cause Y conditionally to a variable Z if $f(Y/X, Z) = f(Y/Z)$, where $f(.)$ stands for the density function. This definition is more general as it covers the linear case and can be analyzed using several approaches based on the characterization of the densityfunction f (Kullback information based on density, copula densities, ... See formalization in Section 1.3 below). In terms of orthogonality, X does not cause Y if the projection of the current value of Y in the vectorial space of history of X and Y is includes in the vectorial space of history of Y. For Pierce and Haugh, X does not cause Y if the current innovation of Y is not correlated with any past innovation of X.

[2]Testable forms are presented and discussed in Section 1.3.

1.1.2 Causality vs Correlation

Correlation refers to a dependence link between two quantitative outcomes[3]. This link can be linear or not, and monotonic or not. Researchers conclude to a correlation between two variables X and Y when X and Y can be linked by a function (says $Y = \phi(X)$ or $X = \phi(Y)$). Thus, (i) if ϕ is a linear function with a positive trend, we say that X and Y have a positive linear correlation, (ii) if ϕ is a linear function with a negative trend, we conclude to a negative linear correlation between X and Y, (iii) if ϕ is non linear with a monotonic (positive or negative) trend, we conclude to a non linear monotonic correlation between X and Y, and (iv) if ϕ is non linear with a non monotonic trend, then we conclude to a non linear correlation between X and Y.

To assess correlation between two variables, several approaches are available in both parametric and non parametric frameworks. We will focus on linear correlation measures as it is straightforward to deal with non linear cases via the transformations $\tilde{Y} = \phi(Y)$ or $\tilde{X} = \phi(X)$ and the use of linear correlation measures on these transformed variables. The most common correlation measure in literature for the parametric framework is that of Pearson (1896). The correlation coefficient is given by :

$$\rho_{X,Y} = \frac{\sum_{i=1}^{N}(x_i - \bar{x}) * (y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i - \bar{x})^2 * \sum_{i=1}^{N}(y_i - \bar{y})^2}} \tag{1.1}$$

Where x's and y's are observed values of X and Y, $\bar{x}$ and $\bar{y}$ are averages of x's and y's respectively, and N is the sample size. The correlation measure $\rho_{X,Y}$ ranges in -1 and 1. The closer $\rho_{X,Y}$ is to 1, the higher the positive correlation between X and Y, the closer $\rho_{X,Y}$ is to -1, the higher the negative correlation between X and Y. A value of $\rho_{X,Y}$ closer to 0 denotes a non correlation between the two outcomes[4]. In non parametric framework, the two major measures

[3]Some extensions of the notion of correlation can be found in literature. For example, researchers mean by biserial correlation, the correlation between a dummy variable and a quantitative one. Correlation between two dummy variables can be analysed by a tetrachoric correlation coefficient. In the case where another variable affects the link between the two interest variable, researches use a partial correlation coefficient that controls for the effects of the additional variable. Dependence between quantitative outcomes have also been addressed with copulas in literature (See Embrechts et al., 2001).

[4]Note that if X and Y are normally distributed (a bivariate normal distribution), $\rho_{X,Y} = 0$ denotes indepen-

of correlation[5] are those of Spearman and Kendall given by :

$$\rho_{Spearman} = 1 - \frac{6 * \sum_{i=1}^{N}(r_i - q_i)^2}{N * (N^2 - 1)} \tag{1.2}$$

$$\tau_{Kendall} = \frac{2 * (R - Q)}{N * (N - 1)} \tag{1.3}$$

Where r_i and q_i denote the rank of observation x_i (respectively y_i) in the x's (respectively y's) sample, R denotes the number of couple (i, j) from which $x_i < x_j$ and $y_i < y_j$, and Q denotes the number of couple (i, j) from which the condition $x_i < x_j$ and $y_i < y_j$ is not verified. All these correlation coefficients range in -1 and 1, and have the same interpretation than above.

Researches test for the significance (H0 : $\rho_{X,Y} = 0$) of $\rho_{X,Y}$ by the use of the following Student test with $N - 2$ degree of freedom[6] :

$$t = \frac{\hat{\rho}}{\sqrt{\frac{1 - \hat{\rho}^2}{N - 2}}} \tag{1.4}$$

Researchers often (when the sample size is low or when $\rho_{X,Y}$ is closer to 0) use a Fisher transformation of $\hat{\rho}$ that is $\hat{z} = \frac{1}{2} * ln(\frac{1 + \hat{\rho}}{1 - \hat{\rho}})$ and use $u = \hat{z} * \sqrt{N - 3} \sim N(0, 1)$ as test statistics. Comparison tests can also be completed between two or more correlation coefficients on independent samples[7]. In a case of two independent samples (H0 : $\rho_1 = \rho_2$), researchers use the following z-test and Chi-square test in a case of K independent samples (H0 : $\rho_1 = \rho_2 = ... = \rho_K$)

$$u = \frac{|\hat{z}_1 - \hat{z}_2|}{\sqrt{\frac{1}{N_1 - 3} + \frac{1}{N_2 - 3}}} \sim N(0, 1)$$

$$\chi = \sum_{k=1}^{K}(N_k - 3)\hat{z}_k^2 - \frac{\left(\sum_{k=1}^{K}(N_k - 3)\hat{z}_k\right)^2}{\sum_{k=1}^{K}(N_k - 3)} \sim \chi^2(K - 1)$$

dence between X and Y.

[5]Copulas are also used to assess the dependence structure between variables.

[6]Note that when the sample size tend to infinity, the statistic is asymptotically normal with mean 0 and variance $\frac{1}{N - 1}$.

[7]Note that there exists statistical tests to compare correlation coefficients between variables observed on the same sample, and tests to compare correlation coefficients between two variables observed on the same sample at two different dates.

Where N_k denotes the size of the sample k, $k = 1...K$.

All approaches described above aim to identify or to measure the link between two outcomes. The conclusion of this kind of analysis will be X and Y are linked or associated, or occur simultaneously. It means that an observed value of X is associated with an observed value of Y (or vice versa). Thus, we can not conclude, via this analysis, that there exists a cause and effect relationship between our two outcomes. Contrarily to a causality analysis in which researchers are able to identify whether X causes Y (or vice versa), i.e the available information on X help to better predict Y (Granger, 1969).

1.2 Public policy assessment

Public policy assessment aims to identify whether the implementation of a specific policy has changed the observed value of an outcome of interest. The effect of an implemented policy should be the difference between the observed outcome and what would have been observed if the policy had not been implemented. Thus, researchers aim to assess what would have been observed if the policy had not been implemented. This issue is addressed according to several approaches that can be grouped in two major frameworks : ex ante and ex post. The ex ante evaluation framework covers all approaches that aim to predict the effect of a policy to be implemented. This framework is very helpful when designing the policy. However, it is based on several hypothesis that might weaken it reliability. Contrarily to this framework, the ex post framework covers all approaches that aim is to assess the effective impact of a policy after its implementation. The policy evaluation framework involves causality measurement since it aims to establish that a specific effect is always observed as consequence of the implemented policy.

In the two following subsections, we present various methods within both ex ante and ex post frameworks. However, we will focus on the ex ante framework as ex post assessment methods involved causality measurement. For practical issues and further details on these approaches, one should refer to Khandker et al. (2010).

1.2.1 Ex ante framework

Ex ante evaluation framework aims to predict the possible impact of a policy to be implemented. In this framework, the environment or the economy is modeled with a set of equations using assumption about individuals behaviours and characteristics, links between individuals and the main economic agent in the policy implementation. Researchers have to identify all key variables including both policy and outcome variables, all transmission channels and make assumptions about each of them. On the basis of these assumptions, researchers build economic models that aim to assess the impact of the policy to be implemented. The impact is determined by measuring how a shock on a policy variable deviates the outcomes for the equilibrium (the equilibrium is assumed to be the current situation without the policy to be implemented). Thus, this framework involves simulations on some policy variables.

Econometric structural modelling that allows assessing impacts have been presented by Heckman (2008). Let's recall the notations. Let $y(s)$ denotes an outcome of interest. Let q_s denotes the set of characteristics modified by an implemented policy. Let x denotes relevant economic variables and u_s unobservable. Let's assume that $y(s)$ is mapped with the characteristics as follows :

$$y(s) = g(q_s, x, u_s) \tag{1.5}$$

Heckman (2008) shows that if an implemented policy can be characterized by a known variation of the vector (q_s, x, u_s) that induces a feasible outcome $y(s)$, then the assessment of the ex ante impact of a policy characterized by q_s is possible.

Many ex ante evaluation methods exist in literature. One may cite for example the ex ante Poverty Impact Assessment (ex ante PIA) that aims to assess the expected effect of pro-poor policies, the dynamic stochastic equilibrium approach. Another framework is Euromod (Tax-benefit microsimulation model for the European Union) that is used to assess the impacts of taxes and benefits on household incomes and work incentives. To make these microsimulation models easier, many software such as General Algebraic Modeling System (GAMS), or Model Generator (MODGEN) have been developed.

1.2.2 Ex post framework

Ex post impact assessment major approaches are randomization, propensity score matching, double differences, instrumental variables modeling and regression by discontinuity. Each of these approaches have some underlying assumptions and are useful under some conditions.

The common concerns about ex post evaluation methods are the determination of the counterfactual and the selection bias issues. The counterfactual denotes the outcome that should had been observed if the beneficiaries had not benefit the project or the policy. The interest of the counterfactual is that it is helpful to determine the impact of the policy. The impact of the policy is defined by the difference between the observed outcome and the conterfactual. Thus, many approaches have been developed to estimate the counterfactual.

The first class of methods consists of measuring the outcome before and after the policy implementation. Then, the outcome before the policy implementation is used as counterfactual. The second class of methods consists of constructing a comparison group. The comparison group should have the same characteristics with the beneficiaries and should not get benefit the policy. Thus, the outcome observed on the comparison group is consider as the counterfactual. However, regardless the counterfactual construction method, to conclude for policy impact, researchers use statistical tests procedures. These statistical tests are based on the underlying hypothesis that both beneficiary and non beneficiary (when comparison group is use) are randomly selected. This assumption is usually not fulfilled because beneficiaries are most often a targeted group or volunteers. Thus, ex post impact evaluation methods differ in how the selection bias problem is addressed.

The randomization method consists of a two-step random selection procedure. In the first step, researchers randomly select a representative sample of individuals. At the second step, researchers randomly divide the sample in two sub-samples (beneficiaries and non beneficiaries). Thus, the estimated impact of policy in this framework is not biased by selection procedure. Some refinements to this pure randomization exist in literature (See Khandker et al., 2010). These refinements consist of modifying the selection procedure at the second step by accounting

for some characteristics. However, randomization is most often infeasible in public policy design since it is difficult to justify the selection procedure.

The propensity score matching (PSM) approach consists of evaluation the counterfactual via a comparison group. The comparison group is constructed on the basis of the probability of being a beneficiary. This approach make the assumption (known as conditional independence assumption) that the probability of being selected as beneficiary only depends on the individual observable characteristics. Furthermore, this approach make the assumption (known as common support or overlap assumption) that with the same characteristics, both probabilities of being a beneficiary or not are non null. This assumption is useful in constructing a comparison group with the closest characteristics that the beneficiaries. To construct the comparison group, many approaches are available to match beneficiaries and non beneficiaries. The nearest-neighbour technique consists of finding one or some non beneficiaries with a probability of being beneficiary closer to that of a beneficiary. The radius matching technique consists of setting a threshold the maximum difference tolerable in the nearest-neighbour technique. The local linear matching and Kernel matching techniques are non-parametric approaches that consist of estimating a locally weighted regression of comparison group outcome near each beneficiary, with and without slope term respectively. The stratification technique consists of matching strata of beneficiaries and non beneficiaries and computing the impact within each strata. All these techniques aim to make the overlap assumption less restrictive. However, the conditional independence is often not satisfied since there may exist some unobservable characteristics that determine beneficiaries.

Contrarily to PSM, the double-difference matching technique consists of computing the impact as the difference between the variations in outcomes for both beneficiaries and non beneficiaries before and after the policy implementation. A baseline and a follow-up survey that covers both beneficiaries and non beneficiaries are needed to complete impact estimation. This approach allows accounting for unobserved heterogeneity in selection. However, the approach is implemented under the underlying assumption that unobserved heterogeneity is time-invariant. A discussion of this latest approach is made by Bertrand et al. (2004).

The instrumental variables approach is designed to allow for time-variant unobserved hetero-

geneity in selection. This approach consists of finding an instrument that determines beneficiary status but is correlated with unobservable that affect outcome. Thus, the main difficulty is about finding a relevant instrument since the use of weak instrument leads to inconsistent impact estimation. Many refinements of this approach are available (see Khandker et al., 2010).

The last approach we introduce right here is the regression discontinuity. This approach allows accounting for both observed and unobserved heterogeneity. It consists of estimating the policy impact near the eligibility threshold, assuming that individuals in both side of the threshold have the same characteristics. A non-parametric local regression is used in the neighbourhood of the threshold to estimate the impact of the implemented policy.

1.3 Causality measurement

As mentioned earlier, causality tests and causality measurement have been addressed by several approaches in econometrics literature since the seminal work of Granger (1969). This section aims to make a survey of these approaches. We first introduce in subsection 1.3.1 the general case applied on time series. Then, we present methods used to test for non causality on panel data in subsection 1.3.2. Subsection 1.3.3 presents and discusses non causality tests in non parametric and qualitative processes frameworks, and subsection 1.3.4 deals with the specific field of health economics. In all subsections below, we present tests in bivariate case. The multivariate case is a straightforward generalization of the bivariate case. Let X and Y denote two quantitative processes, except contrary definition in a subsection. Let $Z = (X, Y)'$ a vector. Let m and n denote two positive scalars.

1.3.1 Time series case

Several approaches have been used to test for non causality. For the definitions given in Section 1.1.1 above, we present formalization and tests used to conclude whether or not for non causality. We then specify the following finite horizon model:

$$A_0 Z_t = \sum_{\tau=1}^{m} A_\tau Z_{t-\tau} + \epsilon_t \tag{1.6}$$

Where $A_0 = \begin{pmatrix} 1 & b_0 \\ a_0 & 1 \end{pmatrix}$, $A_\tau = \begin{pmatrix} a_{1,\tau} & b_{1,\tau} \\ a_{2,\tau} & b_{2,\tau} \end{pmatrix}$, and $\epsilon_t = \begin{pmatrix} \epsilon_t^X \\ \epsilon_t^Y \end{pmatrix}$. To test for Granger lag noncausality[8] in a simple causal model (say $a_0 = b_0 = 0$), one might test whether H0 : $b_{1,\tau} = 0 \; \forall \; \tau = 1, ..., m$ for Y does not cause X, nor H0 : $a_{2,\tau} = 0 \; \forall \; \tau = 1, ..., m$ for X does not cause Y. For Granger instantaneous noncausality, in the equation 1.6 above, exclusion restriction is needed to estimate coefficients. Thus, one may use orthogonalization of the matrix A_0 (that is closer to Wold (1954) causal chain between the two variables). To test for Granger instantaneous noncausality from X to Y, the required exclusion restriction is $b_0 = 0$, thus one may test H0 : $a_0 = 0$[9].

In Sims's (1972) framework, researchers specify and estimate the following model:

$$Z_t = B_0 + \sum_{\tau=-n}^{m} B_\tau Z_{t-\tau} + \epsilon_t \tag{1.7}$$

Where $B_0 = \begin{pmatrix} b_0^1 \\ b_0^2 \end{pmatrix}$, $B_\tau = \begin{pmatrix} 0 & b_{1,\tau} \\ b_{2,\tau} & 0 \end{pmatrix}$, and $\epsilon_t = \begin{pmatrix} \epsilon_t^X \\ \epsilon_t^Y \end{pmatrix}$. Sims noncausality from Y to X is tested in the equation 1.7 with the linear constraint test (a F test) $H0 : b_{1,\tau} = 0 \; \forall \; \tau = -n, ..., -1$, and $H0 : b_{2,\tau} = 0 \; \forall \; \tau = -n, ..., -1$ to test for Sims noncausality from X to Y. Due to the serial autocorrelation in residuals ϵ_t that weaken the F test properties, Sims uses autoregressive filters to transform variables before running regressions and tests.

Pierce and Haugh (1976) test of noncausality is based on innovations. To test for noncausality between X and Y, researchers specify and estimate the autoregressive representation of each series X and Y. Let u and v denote the innovations of X and Y respectively. The autoregressive representation of X and Y are:

$$\begin{aligned} \psi_x(L)u_t &= \varphi_x(L)X_t \\ \psi_y(L)v_t &= \varphi_y(L)Y_t \end{aligned} \tag{1.8}$$

Where L is the lag operator (defined as $L^j u_t = u_{t-j}$), ψ_x and ψ_y are assumed to be invertible.

[8]The tests for Granger lag or instantaneous causality assume that only stationary series are involved. The definition of causality by Granger (1969) is based on the variance of involved series. Thus, non-stationary series cannot be used since their variance are not stable. Some authors provide tests for Granger causality that are robust for non-stationary series (see Toda and Yamamoto, 1995 for further details about this issue).

[9]Reciprocally, to test for Granger instantaneous noncausality from Y to X, the required exclusion restriction is $a_0 = 0$, thus one may test H0 : $b_0 = 0$

The predicted residuals are used to compute cross-correlation terms defined as following:

$$\rho_{u,v}(k) = \frac{\sum_{t=1}^{T} \hat{u}_t \hat{v}_{t-k}}{\left(\sum_{t=1}^{T} \hat{u}_t^2 \sum_{t=1}^{T} \hat{v}_t^2 \right)^{(1/2)}} \tag{1.9}$$

From the analysis of the cross-correlation terms, researchers conclude to instantaneous causality if $\rho_{u,v}(0)$ is not null. However, this analysis does not disentangle in which direction the instantaneous causality holds. Researchers conclude to lag causality from Y to X if $\rho_{u,v}(k)$ is not null for some negatives k, and to lag causality from X to Y if $\rho_{u,v}(k)$ is not null for some positives k.

1.3.2 Panel data case

Non causality test on panel data are more often based on the strong underlying assumption that if the causal link exists, it is the same for all individuals and for all time periods the panel. Based on this assumption, the traditional framework for testing causality in panel data is described below. Assume the following bivariate case (the generalization to multivariate case is straightforward):

$$Z_{i,t} = \sum_{\tau=1}^{m} A_\tau Z_{i,t-\tau} + \eta_i + \epsilon_{i,t} \tag{1.10}$$

Where $\eta_i = (\eta_i^1, \eta_i^2)'$ denotes individual effects vector, and others parameters are the same as in subsection 1.3.1. In order to estimate consistent parameters, researchers differentiate the equation 1.10 and use 2SLS procedure with instrumental variables. The differentiated model is:

$$Z_{i,t} - Z_{i,t-1} = \sum_{\tau=1}^{m} A_\tau (Z_{i,t-\tau} - Z_{i,t-\tau-1}) + \epsilon_{i,t} - \epsilon_{i,t-1} \tag{1.11}$$

Thus, to test for lag noncausality, researchers test whether H0 : $b_{1,\tau} = 0 \; \forall \; \tau = 1, ..., m$ for Y does not cause X, nor H0 : $a_{2,\tau} = 0 \; \forall \; \tau = 1, ..., m$ for X does not cause Y.

This approach has the advantage to save degrees of freedom but assumes the causal link to be homogeneous among the panel. This assumption can be challenged (Nair-Reichert and Weinhold, 2000). Thus, Nair-Reichert and Weinhold (2000) specify a model with individual random causal coefficients. The equation is as following:

$$Z_{i,t} = \sum_{\tau=1}^{m} A_{i,\tau} Z_{i,t-\tau} + \epsilon_{i,t} \tag{1.12}$$

Where $A_{i,\tau} = \begin{pmatrix} a_{1,\tau} & b_{1,i,\tau} \\ a_{2,i,\tau} & b_{2,\tau} \end{pmatrix}$, with $b_{1,i,\tau} = b_{1,\tau} + \xi_{1,i,\tau}$ and $a_{2,i,\tau} = a_{2,\tau} + \xi_{2,i,\tau}$. This approach has the advantage to be well suited for heterogeneous panel data and to allow the analysis of causality distribution among panel. Thus, researchers can conclude to causality with a probability rather than concluding to causality on an heterogeneous dataset.

1.3.3 Non-parametric framework

The underlying assumption that variables involved in a causal relationship are stationary have been challenged in literature. Researchers have shown that testing causality only requires one of the involved variables to be homogeneous Markov chain of some order (Sekkat, 1989). Thus, several approaches have been derived based on the assumption of homogeneous Markov chain. Some of these tests deal with qualitative variables case (Bouissou et al., 1986) and can be extend to non homogeneous Markov chains (Gouriéroux et al., 1987).

Bouissou et al. (1986) derive a log-likelihood ratio (LR) test under the assumption that one of the involved variable is a Markov chain. This test is derived for qualitative processes. Let x_t and y_t be two random qualitative variables with I_t and J_t categories respectively at each period $t = 1, ..., T$. Let $i_1^T = (i_1, ..., i_T)$ and $j_1^T = (j_1, ..., j_T)$ denote all the history of X and Y for an individual. Let n denotes the number of individuals. We also assume X to be an homogeneous Markov chain of order m, with $m < T - 3$. The LR statistics to test whether Y does not cause X under the assumption that X is an homogeneous Markov chain of order m is given by:

$$LR = LR_m^m + \sum_{t=m+1}^{T-1} LR_t^m \tag{1.13}$$

Where

$$LR_m^m = 2 \sum_{i_1^T, j_1^m} n(i_1^T, j_1^m) log\left(\frac{n(i_1^T, j_1^m)}{n(i_1^T)} * \frac{n(i_1^m)}{n(i_1^m, j_1^m)} \right)$$

$$LR_t^m = 2 \sum_{i_1^T, j_1^t} n(i_1^T, j_1^{t-1}) log\left(\frac{n(i_1^T, j_1^t)}{i_1^T, j_1^{t-1}} * \frac{n(i_1^t, j_1^t)}{i_1^t, j_1^{t-1}} \right)$$

And $n(i_1^T, j_1^m)$ denotes the number of individuals with the history i_1^T for X and j_1^m for Y. Under the null hypothesis of noncausality, the statistic test LR is asymptotically a Chi-square with

$ddf = ddf_m^m + \sum_{t=m+1}^{T-1} ddf_t^m$ degrees of freedom. The degree of freedom is given by:

$$ddf_m^m = \left[\left(\prod_{k=1}^{m} J_k\right) - 1\right] * \left[\left(\prod_{k=1}^{T} I_k\right) - \left(\prod_{k=1}^{m} I_k\right)\right]$$

$$ddf_t^m = (J_t - 1) * \left[\prod_{k=1}^{T} I_k * \prod_{k=1}^{t-1} J_k - \prod_{k=1}^{t} I_k * \prod_{k=1}^{t-1} J_k\right]$$

Gouriéroux et al. (1987) also derive a test of noncausality based on the Kullback information criterion under the assumption that one of the involved variables is a Markov chain. This test is available for both qualitative and quantitative processes and whether or not the Markov chain is homogeneous. For quantitative processes case, we assume X and Y to be autoregressive of order m with I and J categories respectively. Let assume the following equations for $l = 1, ..., 4$:

$$X_t = \sum_{k=1}^{m} a_{l,k} X_{t-k} + \sum_{k=k_l}^{m} b_{l,k} Y_{t-k} + u_{l,t} \tag{1.14}$$

$$Y_t = \sum_{k=1}^{m} c_{l,k} Y_{t-k} + \sum_{k=k_l}^{m} d_{l,k} X_{t-k} + v_{l,t} \tag{1.15}$$

Where $u_{l,t}$ is distributed with mean zero and covariance matrix Σ_{u_l}, $v_{l,t}$ is distributed with mean zero and covariance matrix Σ_{v_l}, and for $l = 1, ..., 4$, k_l, $a_{l,k}$, $b_{l,k}$, $c_{l,k}$, and $d_{l,k}$ are given in Table 1.3.3 below :

Table 1.1: Different specifications case

Case	Value of l	value of k_l	Constraints
Finite marginal autoregressive	$l = 1$	$k_1 = 1$	$a_{1,k} = c_{1,k} = 0 \,\forall\, k$
Finite joint autoregressive	$l = 2$	$k_2 = 1$	Some $a_{2,k}, b_{2,k}, c_{2,k}$, and $d_{2,k}$ not null
Finite joint autoregressive with current values	$l = 3$	$k_3 = 0$	Some $a_{2,k}, b_{2,k}, c_{2,k}$, and $d_{2,k}$ not null
Sims joint autoregressive	$l = 4$	$k_4 = -\infty$	Some $a_{2,k}, b_{2,k}, c_{2,k}$, and $d_{2,k}$ not null

Then, to test for noncausality, researchers first estimate the empirical covariances matrices $\hat{\Sigma}_{u_l}$ and $\hat{\Sigma}_{v_l}$. Thus, the test statistics are given by :

- For H0 : X does not cause Y, the test statistic $\hat{C}_{X\,to\,Y} = Tlog\left(\frac{det(\hat{\Sigma}_{v_1})}{det(\hat{\Sigma}_{v_2})}\right)$ has a Chi-square distribution with IJm degrees of freedom under the null hypothesis of noncausality.

- For H0 : Y does not cause X, the test statistic $\hat{C}_{Y\,to\,X} = Tlog\left(\frac{det(\hat{\Sigma}_{u_1})}{det(\hat{\Sigma}_{u_2})}\right)$ has a Chi-square distribution with IJm degrees of freedom under the null hypothesis of noncausality.

- For H0 : no instantaneous causality between X and Y, the test statistic $\hat{C}_{X,Y} = Tlog\left(\frac{det(\hat{\Sigma}_{v_2})}{det(\hat{\Sigma}_{v_3})}\right)$ or $\hat{C}_{X,Y} = Tlog\left(\frac{det(\hat{\Sigma}_{u_2})}{det(\hat{\Sigma}_{u_3})}\right)$ has a Chi-square distribution with IJ degrees of freedom under the null hypothesis of instantaneous noncausality.

For qualitative processes case, we assume X and Y to be random qualitative variables with I and J categories respectively. X and Y are also assumed to be Markov processes of order one. We suppose a panel of n realizations of the processes X and Y. Let define the following empirical frequencies (when the Markov chain are assumed to be homogeneous):

$$\hat{p}(i,j/k,l) = \frac{n\Big((X_t = i, Y_t = j)/(X_{t-1} = k, Y_{t-1} = l)\Big)}{n(X_{t-1} = k, Y_{t-1} = l)} \tag{1.16}$$

$$\hat{\pi}(i,j) = \frac{n(X_t = i, Y_t = j)}{nT} \tag{1.17}$$

$$\hat{p}(i,./k,l) = \sum_{j=1}^{J} \hat{p}(i,j/k,l) \tag{1.18}$$

$$\hat{p}(.,j/k,l) = \sum_{i=1}^{I} \hat{p}(i,j/k,l) \tag{1.19}$$

$$\hat{p}_X(i/k) = \frac{\sum_{j=1}^{J}\sum_{l=1}^{J} \hat{p}(i,j/k,l)\hat{\pi}(k,l)}{\sum_{l=1}^{J} \hat{\pi}(k,l)} \tag{1.20}$$

$$\hat{p}_Y(j/l) = \frac{\sum_{i=1}^{I}\sum_{k=1}^{I} \hat{p}(i,j/k,l)\hat{\pi}(k,l)}{\sum_{k=1}^{I} \hat{\pi}(k,l)} \tag{1.21}$$

The first empirical frequency denotes the probability of the transition from the state (k,l) to the state (i,j), the second one denotes the probability of the state (i,j), the third and the fourth denote the probability of transition from the state (k,l) to the event $X_t = i$ and $Y_t = j$ respectively, the fifth and the sixth denote the probability of transition from the event $X_{t-1} = k$ to $X_t = i$ and from the event $Y_{t-1} = l$ to $Y_t = j$ respectively. Thus, to test for noncausality, researchers use the test statistics given by :

- For H0 : X does not cause Y, the statistics

$$\hat{C}_{X\,to\,Y} = 2T\sum_{k=1}^{I}\sum_{l=1}^{J}\hat{\pi}(k,l)\Bigg[\sum_{j=1}^{J}\hat{p}(.,j/k,l)log\bigg(\frac{\hat{p}(.,j/k,l)}{\hat{p}_Y(j/l)}\bigg)\Bigg]$$

  has a Chi-square distribution with $J(I-1)(J-1)$ degrees of freedom under the null hypothesis of noncausality.

- For H0 : Y does not cause X, the statistics

$$\hat{C}_{X\,to\,Y} = 2T\sum_{k=1}^{I}\sum_{l=1}^{J}\hat{\pi}(k,l)\Bigg[\sum_{i=1}^{I}\hat{p}(i,./k,l)log\bigg(\frac{\hat{p}(i,./k,l)}{\hat{p}_X(i/k)}\bigg)\Bigg]$$

  has a Chi-square distribution with $I(I-1)(J-1)$ degrees of freedom under the null hypothesis of noncausality.

- For H0 : no instantaneous causality, the statistics

$$\hat{C}_{X\,to\,Y} = 2T\sum_{k=1}^{I}\sum_{l=1}^{J}\hat{\pi}(k,l)\Bigg[\sum_{i=1}^{I}\sum_{j=1}^{J}\hat{p}(i,j/k,l)log\bigg(\frac{\hat{p}(i,j/k,l)}{\hat{p}(i,./k,l)\hat{p}(.,j/k,l)}\bigg)\Bigg]$$

  has a Chi-square distribution with $IJ(I-1)(J-1)$ degrees of freedom under the null hypothesis of no instantaneous causality.

For non homogeneous Markov chains, researchers can define the time variant counterpart of the empirical frequencies $\hat{p}_t(i,j/k,l)$, $\hat{p}_t(i,./k,l)$, $\hat{p}_t(.,j/k,l)$, $\hat{\pi}_t(k,l)$, $\hat{p}_{t,Y}(j/l)$, and $\hat{p}_{t,X}(i/k)$ in the same manner as in equations 1.16-1.21. Thus, at each time period t, the causality measures are given by :

- $\hat{C}_{X\,to\,Y}(t;k,l) = 2n\hat{\pi}_{t-1}(k,l)\Bigg[\sum_{j=1}^{J}\hat{p}_t(.,j/k,l)log\bigg(\frac{\hat{p}_t(j,./k,l)}{\hat{p}_{t,Y}(j/l)}\bigg)\Bigg]$ to test for noncausality from X to Y,
- $\hat{C}_{Y\,to\,X}(t;k,l) = 2n\hat{\pi}_{t-1}(k,l)\Bigg[\sum_{i=1}^{I}\hat{p}_t(i,./k,l)log\bigg(\frac{\hat{p}_t(i,./k,l)}{\hat{p}_{t,X}(i/k)}\bigg)\Bigg]$ to test for noncausality from Y to X,
- $\hat{C}_{X,Y}(t;k,l) = 2n\hat{\pi}_{t-1}(k,l)\Bigg[\sum_{i=1}^{I}\sum_{i=j}^{J}\hat{p}_t(i,j/k,l)log\bigg(\frac{\hat{p}_t(i,j/k,l)}{\hat{p}_t(i,./k,l)\hat{p}_t(.,j/k,l)}\bigg)\Bigg]$ to test for instantaneous noncausality between X and Y.

These statistics have a Chi-square distribution with $(I-1)(J-1)$ degrees of freedom under null hypothesis of noncausality for large n and are asymptotically independent. To test for

global noncausality, researchers use the test statistic $\hat{C}_{X\,to\,Y} = T\sum_{t=1}^{T}\sum_{k=1}^{I}\sum_{l=1}^{J}\hat{C}_{X\,to\,Y}(t;k,l)$ for noncausality from X to Y, $\hat{C}_{Y\,to\,X} = T\sum_{t=1}^{T}\sum_{k=1}^{I}\sum_{l=1}^{J}\hat{C}_{Y\,to\,X}(t;k,l)$ for noncausality from Y to X, and $\hat{C}_{X,Y} = T\sum_{t=1}^{T}\sum_{k=1}^{I}\sum_{l=1}^{J}\hat{C}_{X,Y}(t;k,l)$ for instantaneous noncausality. These test statistics have a Chi-square distribution with $TI(I-1)(J-1)$, $TJ(I-1)(J-1)$, and $TIJ(I-1)(J-1)$ degree of freedom respectively.

Bouezmarni et al. (2012) approach is based on copula densities. They use the characterization of Granger causality in terms of conditional independence. Let X, Y and Z denote random vectors with d_1, d_2, and d_3 components respectively. Let recall the definition of noncausality : X does not cause Y conditionally to Z if $f(Y/X,Z) = f(Y/Z)$. From this definition, the authors suggest testing the following hypothesis :

$$H0 : f(y,X,Z) * f(Z) = f(y,Z) * f(X,Z)\ \forall\, y \tag{1.22}$$

To test this hypothesis above, the copula characterization of probability density functions (say, $f(x,y,z) = f(x) * f(y) * f(z) * c(F(x),F(y),F(z))$, where c is the copula density and F the probability distribution function, noted pdf below) is used. Thus, the hypothesis 1.22 above can be rewritten in terms of copula density as following :

$$H0 : c(F(x),F(y),F(z)) * c(F(z)) = c(F(y),F(z)) * c(F(x),F(z))\ \forall\, y \tag{1.23}$$

The test statistic $\hat{H}$ is the Hellinger distance between the two functions at each side of equation 1.23 and is given by :

$$\hat{H} = \frac{1}{T}\sum_{t=1}^{T}\left(1 - \sqrt{\frac{\hat{c}(F_T(y_t),F_T(z_t)) * \hat{c}(F_T(x_t),F_T(z_t))}{\hat{c}(F_T(x_t),F_T(y_t),F_T(z_t)) * \hat{c}(F_T(z_t))}}\right)^2 \tag{1.24}$$

Where $\hat{c}$ denotes empirical copula and F_T denotes empirical pdf. If $\hat{H}$ is closer to zero, researchers conclude to noncausality. Bouezmarni et al. (2012) derive a asymptotic distribution of the test statistic BRT that follows a standard normal distribution under null hypothesis. The process (X,Y,Z) is assumed to be strictly stationary and $\beta-$mixing, c is assumed to be twice differentiable, and the bandwidth parameter k is chosen so that $k = O(T^{\xi})$, with $\xi \in [\frac{2}{d+4};\frac{2}{d}]$. This statistic is given by :

$$BRT = \frac{Tk^{-d/2}}{\sqrt{2}(\pi/)^{d/2}}\left(4\hat{H} - 2^{-d}T\left(\frac{\pi}{k}\right)^{d/2} - \hat{B}_1T^{-1}k^{(d_2+d_3)/2} - \hat{B}_2T^{-1}k^{(d_1+d_3)/2} - \hat{B}_3T^{-1}k^{d_3/2}\right) \tag{1.25}$$

Where $d = d_1 + d_2 + d_3$, $\hat{G}_t = (F_T(x_t), F_T(y_t), F_T(z_t))$, and

$$\hat{B}_1 = -2^{-(d_2+d_3-1)}\pi^{(d_2+d_3)/2} + \frac{1}{T}\sum_{t=1}^{T} \frac{\prod_{j=1}^{d_2+d_3}\left(4\pi\hat{G}_{j,t}(1-\hat{G}_{j,t})\right)^{-1/2}}{\hat{c}(F_T(y_t), F_T(z_t))}$$

$$\hat{B}_2 = -2^{-(d_1+d_3-1)}\pi^{(d_1+d_3)/2} + \frac{1}{T}\sum_{t=1}^{T} \frac{\prod_{j=1}^{d_1+d_3}\left(4\pi\hat{G}_{j,t}(1-\hat{G}_{j,t})\right)^{-1/2}}{\hat{c}(F_T(x_t), F_T(z_t))}$$

$$\hat{B}_3 = 2^{-(d_3-2)}\pi^{(d_3)/2} + \frac{1}{T}\sum_{t=1}^{T}\left[\frac{\prod_{j=1}^{d_3}\left(4\pi\hat{G}_{j,t}(1-\hat{G}_{j,t})\right)^{-1/2}}{\hat{c}(F_T(z_t))} - 2\frac{\hat{c}(F_T(y_t))\prod_{j=1}^{d_3}\left(4\pi\hat{G}_{j,t}(1-\hat{G}_{j,t})\right)^{-1/2}}{\hat{c}(F_T(y_t), F_T(z_t))}\right.$$
$$\left. - 2\frac{\hat{c}(F_T(x_t))\prod_{j=1}^{d_3}\left(4\pi\hat{G}_{j,t}(1-\hat{G}_{j,t})\right)^{-1/2}}{\hat{c}(F_T(x_t), F_T(z_t))}\right] \; if \; d_3 > 1$$

$$\hat{B}_3 = 2^{-(d_3-2)}\pi^{(d_3)/2} - 1 \; if \; d_3 = 1$$

1.3.4 Some specific cases for causality in health economics

One of the particularities of health economics analysis is the use of ordinal or qualitative outcomes and short panel data. One of the pioneer paper on this field is that of Adams et al. (2003). Their paper deal with a panel data with 3 waves. Adams et al. use a Chow-type test to conclude to invariance property of causal coefficient for transition between waves 1 and 2, and waves 2 and 3. Let Y denotes a vector of K interest variables that can be binomial, ordered discrete or continue. The authors assume Y to be a Markov chain with order one, and assumed a Wold causal chain among elements of Y. Formally, their model is the following one. The estimated models are the following :

$$Y^*_{k,i,t} = \sum_{j=1}^{k-1}\alpha_{j,k}Y_{j,i,t} + \sum_{j=1}^{K}\beta_{j,k}Y_{j,i,t-1} + \delta_{k,i} + \epsilon_{k,i,t} \tag{1.26}$$

Where $Y_{k,i,t}$, $k = 1, ..., K$, denotes the k^{th} component of Y. $Y^*_{k,i,t}$ is the latent variable (when dealing with binomial or ordered discrete variables) or the observed $Y_{k,i,t}$ in the case of continue component. The authors conclude to causality when the causal coefficients $\beta_{j,k}$ are significant and stable (invariant) within the two transitions (wave 1 to 2 and wave 2 to 3).

Another important paper that deals with causality in health economics is that of Michaud and Van Soest (2008). They use a panel data with 6 waves to investigate causal links between

household' wealth and spouses health. The authors construct health index using principal component analysis on a set of health indicators including self-reported health. Thus, the three variables of interest in their study are continuous variables. Then, the authors estimate a vector autoregressive model that accounts for instantaneous causal links as well as lag causal links. Let Y denotes a vector of three variables (Male and female health indexes, and household wealth). The estimated model is the following one :

$$\Gamma Y_{i,t} = AX_{i,t} + \sum_{k=1}^{p} \Phi_k Y_{i,t-k} + \eta_i + \epsilon_{i,t} \tag{1.27}$$

Where X is a vector characteristics, η_i is the vector of individual effects. To estimate this model, authors used generalized method of moments. To solve identification problems in such a model, authors used as instrument for wealth inheritance, and onsets of critical health condition for instrumenting health changes. The identification of instantaneous causal links between spouses health, authors used onsets of health conditions as instruments. However, the authors estimated separably each equations.

Conclusion

Causality measurement has generated a large literature in many fields. This large interest on the causality issue is due to the important role of causality in economic analysis. Many specific approaches have been developed in both parametric and non-parametric framework. However, as researchers are always interested in impact evaluation, the use of causality measurement approaches and the improvement and development of newer approaches is still a growing field. The major challenge of this dissertation is to propose three innovative contributions to the field.

Chapter 2

Health Condition and Job Status Interactions: Econometric Evidence of Causality from a French Longitudinal Survey[1] [2]

[1]Authors : **Eric Delattre**, ThEMA-UCP; **Richard Moussa**, ThEMA-UCP and ENSEA Abidjan; and **Mareva Sabatier**, IREGE, Université Savoie Mont Blanc, France.

[2]The authors greatfully acknowledge the Risk Foundation (Health, Risk and Insurance Chair, Allianz) for financial support and the Centre Maurice Halbwachs (Réseau Quetelet) for access to the SIP 2007 data set (Santé et itinéraire professionnel - 2007. DARES producteur. Centre Maurice Halbwachs diffuseur). They also thank IRDES for the opportunity to use its mapping code of illnesses, according to ICD codes.

Abstract

This article investigates the causal links between health and employment status. To disentangle correlation from causality effects, the authors leverage a French panel survey to estimate a bivariate dynamic probit model that can account for the persistence effect, initial conditions, and unobserved heterogeneity. The results highlight the crucial role of all three components and reveal strong dual causality between health and employment status. The findings clearly support demands for better coordination between employment and health public policies.

Keywords: health and job causality, bivariate dynamic probit model, Gauss-Hermite method

JEL Classification: I10, J6, C3, C51

Introduction

Health changes and labour market instability both have important impacts on individual well-being, which strongly guide policy makers in defining rules for health insurance, unemployment benefits, and/or retirement. A substantial empirical literature stresses the links between health and labour market risks, yet the precise relationship between the two phenomena remains unclear, leaving the design of appropriate public policies uncertain as well, especially because policies in the labour market can produce health effects (and vice versa).

Early empirical studies focused on one-way causality, such that health conditions explained labour market transitions or *vice versa*. For example, in Berkowitz and Johnson's (1974) pioneering study, people's health determines their labour participation decisions, and Stern (1989) confirms that disabilities strongly affect labour participation. As an endowment of human capital, health determines productivity and preferences for work versus leisure (Grossman, 1972). Moreover, two complementary results emerge from a literature review (Currie and Madiran, 1999). First, poor health affects everyone's labour choices, but the impact is especially powerful among the elderly, such that health problems significantly increase choices to retire (Sickles and Taubman, 1986; Bound, 1991; Cai and Kalb, 2006,2007; Christensen and Kallestrup-Lamp, 2012),

and retirement decisions often represent an attempt to preserve health (Coe and Zamarro, 2008). Second, the impact of a person's health varies with the type of health deterioration. Chronic diseases, such as cancer (Eichenbaum-Voline et al., 2008), diabetes (Bastida and Pagan, 2002; Brown et al., 2005), mental illness (Butterworth et al., 2006), and disabilities (Stern, 1989), seem to have the strongest effect on individual transitions in the labour market.

In addition, employment status has implications for health. For example, unemployment and inactivity slightly increase the risks of cardiovascular diseases (Jin et al., 1995), cancer, or mental illnesses (Brenner, 2002, Llena-Nozal, 2009). Morris et al. (1994) using British data and Mathers and Schofield (1998) using Australian data confirm that a loss of employment increases mortality risk. Mesrine (2000) shows that this impact is even greater following long spells of unemployment. The pecuniary and non-pecuniary effects of inactivity and unemployment on health help explain these empirical findings. Unemployment usually decreases the health care resources available to the person, so it can affect health over the long-term. In addition, unemployment and non-participation in the labour market damage people's self-esteem (Brenner, 2002; Llena-Nozal, 2009) and decrease their sense of well-being (Winkelman and Winkelman, 1998; Clark et al. 2001). Persistent unemployment and inactivity thus create threatening conditions for health. Conversely, being employed can have some deleterious effects on health, such as by increasing the risk of stress, professional illness and work accidents. Thus, Debrand (2011) uses economic data to argue that bad working conditions and work pain cause damage to people's health. Using a matching approach with the French Health Survey 2002, Debrand shows that workers exposed to poor working conditions consult physicians 25% more than those who are not. Hamon-Cholet and Sandret (2007) similarly find, with French data, that noisy jobs increase the professional accident rate to 25%.

However, the links between health and labour status may be more complex than a one-way form of causality. Recently, some authors have emphasized the need to correct for endogeneity between health (Madden, 2004; Brown et al., 2005; Haan and Myck, 2009) and labour market transitions (Rietveld et al., 2015). Neglecting endogeneity can cause strong estimate biases. For example, Caroli and Godard's (2014) analyses of the European Working Conditions data set indicate that the fear of involuntary job loss has health impacts, such as headaches, eye-strain,

and skin problems. Without controlling for the endogeneity of job insecurity, job insecurity degrades all health indicators. This endogeneity of health and job risks likely reflects two main sources. First, unobserved heterogeneity, such as that due to lifestyle or individual preferences, can influence both health and labour market processes (Cai, 2010). Second, measurement errors in self-reported health surveys or using poor health as a reason to justify unemployment, might create substantial endogeneity biases (Zhang et al., 2008).

Another major source of endogeneity is likely to be reciprocity : Labour activities and health affect each other. Few studies take this simultaneity into account, though Haan and Myck (2009) propose a bivariate model with a lagged dependent variable to analyze dynamics in health and labour market risk. This approach offers the advantage of addressing endogeneity problems and allowing for a dynamic analysis. Accordingly, these authors show that recent health conditions affect current labour market risk, and vice versa, and that this dynamic is strongly persistent. Such persistence effects also may be due to favorable or unfavorable initial conditions for health and employment (Heckman, 1981; Arulampalam and Stewart, 1995), and Haan and Myck (2009) do not address these potential contingencies. Neglecting these initial conditions could bias estimates of the simultaneity effect between health and employment status.

Finally, we lack clear definitions of all the links between health and job risks. With this article, we propose an innovative methodology for identifying and assessing all the complex links between health and employment paths. With our modeling approach, we can jointly estimate the two phenomena. We assume sequential causality, as in Alessie et al. (2004), or Lindeboom and Kerkhofs (2009) or Haan and Mynck (2009), such that the most recent health status can influence the current labour market status, and the last event in the labour market affects the current period health status. We also account for unobserved heterogeneity and persistence in the two processes over time (Adams et al., 2003). Finally, following Wooldridge (2005), we control for initial conditions.

Unlike previous empirical work, we aim to establish whether true causality exists between health and employment, as well as to define its meaning and scope , such that we can derive insights and guidance for economic policies. If health and employment are independent, policy makers can use disconnected instruments. If single causation exists instead (e.g., job transitions

explain health paths but health does not affect job risks), it will be necessary to monitor the effects of an employment-centered policy on health. Finally, if dual causality exists, only the joint design of health and employment policies can improve health and employment simultaneously.

The estimates in this study feature a sample of French individuals who completed the Santé et Itinéraire Professionnel (SIP) survey (DARES[3] , DREES[4], 2006). This survey (see Section 1) indicates, for each year since the participant finished school until 2006, all individual events related to health and labour market status. With this long panel data, we can better control for unobserved heterogeneity compared with using cross-sectional data. Moreover, this survey provides empirical evidence of the links between health and labour market paths in France, whereas prior literature has focussed on U.S., British, or Australian data. Significant institutional differences (in terms of legislation regulating the labour market and rules governing health systems) exist across these countries, which limits the generalizability of the results obtained in English-speaking countries to the French case. Focusing on the French case thus might provide new insights and clarify the links between health and labour market transitions, by addressing them in a different kind of health care system.

Section 1 presents the relevant data for this analysis. Section 2 outlines the innovative methodology we have implemented to investigate the complex links between health and labour market transitions. After we present and discuss the results in Section 3, we conclude with some implications and directions for further research.

2.1 French longitudinal survey on health and work: SIP

Conducted in 2006 by DARES and DREES, the Santé et Itinéraire Professionnel (SIP) survey gathered information about 13,991 individuals, aged from 20 to 74 years (Mermilliod, 2012). This survey describes individual paths on the job market and health status. Each respondent provides the information about previous conditions. The survey data also include socioeconomic information, such as gender, age, grades, income, and ethnicity.

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

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

Because we seek to analyze events during people's professional lives, we exclude those who never entered the job market. We also exclude those who entered before 1962, to observe macroeconomic conditions that may affect individual transitions in the labour market. After dropping observations with missing data, we obtained a sample of 10,569 persons who provided detailed information about their participation in the labour market and their health status, spanning the full professional path of each individual, from the end of schooling to retirement. On average, each respondent thus provides information about a period of 26 years[5]. Pooling the data across all years produces a dataset with 255,206 observations.

For each year of professional life, we distinguish four categories for job status :

- Long time period employments, which last at least five years.
- Short time period employments, which last less than five years.
- Unemployment periods, which last more than one year.
- Out of job market time periods, which last more than one year.

With the first two items, we define all respondents who report being employed in a long-term or short-term job as employed for that given year. Our definition of employed people is thus quite expansive, because non-employment status covers both unemployment and non-participation. In addition, the SIP survey does not offer a means to observe short-term (shorter than one year) unemployment or inactivity. Being employed during a particular year in the survey does not imply that individuals were employed for the entire year though, so measurement errors could arise for the labour market status variable. To avoid this bias, and as robustness tests, we also consider long-term inactivity and unemployment status. These two items also are binary variables, equal to 1 if the respondent is inactive or unemployed for the entire given year.

Moreover, participants self-report whether they have encountered illnesses during a given year. With these data, we can construct a health indicator as a binary variable, equal to 1 if the respondent reports any illness. For a better understanding of health status, we also create a more qualitative indicator, similar to Christensen and Kallestrup-Lamp (2012). For each

[5]Excluding the initial lagged period.

illness reported in the survey, we know the corresponding World Health Organization's ICD[6]. That code also reveals an indicator of severity and an indicator of disability according to the mapping created by the Institut de Recherche et de Documentation en Économie de la Santé (IRDES). The severity index indicates if the illness is related to a risk of death; the disability index determines if the illness affects the person's daily life. With this information, we create binary dummy variables to establish whether the risk of death is large (`rdeath=1`) and whether the disability index is large (`disab=1`). In turn, we create a percentage measure to reflect the extent to which each situation occurs over the course of the respondent's full working life.

Because we know the length of each respondent's professional life, we can calculate synthetic indicators of the professional and health paths: the percentage of professional life with at least one illness and the share of employment, unemployment, and out-of-job market periods in professional life (see Table 2.1).

Table 2.1: Descriptive statistics for labour market and health paths

Indicators	Means	Std. Err.
Number of years per individual	26.994	12.070
Share of employment periods in professional life	0.863	0.237
Share of unemployment periods in professional life	0.034	0.093
Share of out-of-job market time periods in professional life	0.103	0.219
Share of years with at least one illness in professional life	0.1795	0.295
Share of years with at least one illness with disability	0.028	0.135
Share of years with at least one illness with risk of death	0.019	0.165

Notes: Number of individuals: 10,569

As this table shows (means in column 2 and standard deviations in column 3), employment periods represent a large fraction of the professional life. Only 3.4% of professional life involved long-term unemployment, and 10.3% occurred out of the job market. Illness periods represented almost 18% of the professional life.

[6]International Statistical Classification of Diseases and Related Health Problems - 10th Revision (ICD-10)

Moreover, exploiting the longitudinal dimension of our data, we examine the conditional outcome in period t, conditional on the respondents' self-assessed statuses in the labour market and health in period $t-1$ (Table 2.2). We find considerable persistence in both the labour market and health paths. For example, conditional on being employed in $t-1$, about 97.8% of respondents report being employed in t (on pooled sample).

Table 2.2: Transitions in labour market and health status

	Status at t					
Status at $t-1$	Employed	Unemployed	Out of labour market	Ill	Ill with disability	Ill with risk of death
Employed	0.978	0.011	0.011	0.213	0.028	0.018
Unemployed	0.331	0.622	0.047	0.324	0.036	0.025
Out of labour market	0.081	0.005	0.914	0.289	0.044	0.030
Ill	0.809	0.043	0.148	0.986	0.125	0.082
Ill with disability	0.782	0.039	0.179	0.982	0.982	0.337
Ill with risk of death	0.770	0.043	0.187	0.970	0.512	0.970
Not Ill	0.879	0.022	0.099	0.017	0.003	0.002

Table 2.1 also presents the labour force status against lagged self-reported health, using the pooled sample. It highlights the negative relationship between poor health and employment. Respondents who declare a disease in $t-1$ are more likely to be unemployed or out of the labour market in t. But these statistics also suggest evidence of a reverse link, as suggested in prior literature. Table 2.1 also contains the health status against the lagged labour market indicators, using the pooled sample. Finally, persistence and simultaneity seem to characterize health and labour market processes.

In addition, some individual attributes can be observed[7]. Table 2.3 provides the information pertaining to these variables for the pooled sample and for sub-samples defined according the labour market and health status.

According to these descriptive statistics, persons who do not participate to the labour market

[7]Among all these variables, only three (age, number of children, and marital status) vary over time.

Table 2.3: Socioeconomic characteristics

	Employed	Unemployed	Out of labour market	Ill	Ill with		Pooled sample
					disability	risk of death	
Men	0.508	0.364	0.095	0.426	0.516	0.481	0.460
Not French*	0.108	0.141	0.195	0.103	0.059	0.08	0.119
Couple	0.705	0.618	0.808	0.734	0.712	0.634	0.713
Number of children	1.257	1.379	2.020	1.613	1.609	1.561	1.350
No grade	0.068	0.134	0.190	0.089	0.101	0.092	0.084
High School grade	0.537	0.543	0.518	0.536	0.555	0.511	0.534
College grade	0.161	0.162	0.141	0.167	0.175	0.181	0.158
Undergraduate studies	0.095	0.068	0.073	0.083	0.083	0.076	0.092
Graduate studies	0.140	0.093	0.077	0.126	0.087	0.14	0.132
Number of obs.	220,812	8,335	31,817	54,989	7,257	4,830	255,206

*: Refers to the individual's nationality.

in a given year are more likely to have certain specific characteristics. As expected, females, less educated people, and those with children are more likely to be out of the labour market. Conversely, among the employed, we count more men and people with academic degrees. Table 2.3 also shows that female, French people and those with academic degrees report more numerous illness periods. These statistics do not necessarily mean that respondents suffer poorer health; they might just be more concerned about their health and thus declare more illnesses.

Finally, these descriptive statistics argue for taking simultaneity and persistence effects into account to obtain a robust analysis of causality links between health and employment status. We present an econometric framework to fulfil that goal.

2.2 Econometric framework

2.2.1 Testing causality: general approach

We first define two dependent variables: health condition ($h = 1$ if an illness is declared, $h = 0$ otherwise) and job status ($w = 1$ if employed, long or short time periods, $w = 0$ otherwise).

From the SIP data set, we can observe h and w for each individual i and each year t. Thus, we model the interactions between h_{it} and w_{it} while accounting for two issues : the path dynamics of each event (and particularly the inertia of each path) and the link between each path. In Figure 2.1, we present all the links that may exist between the two events over time.

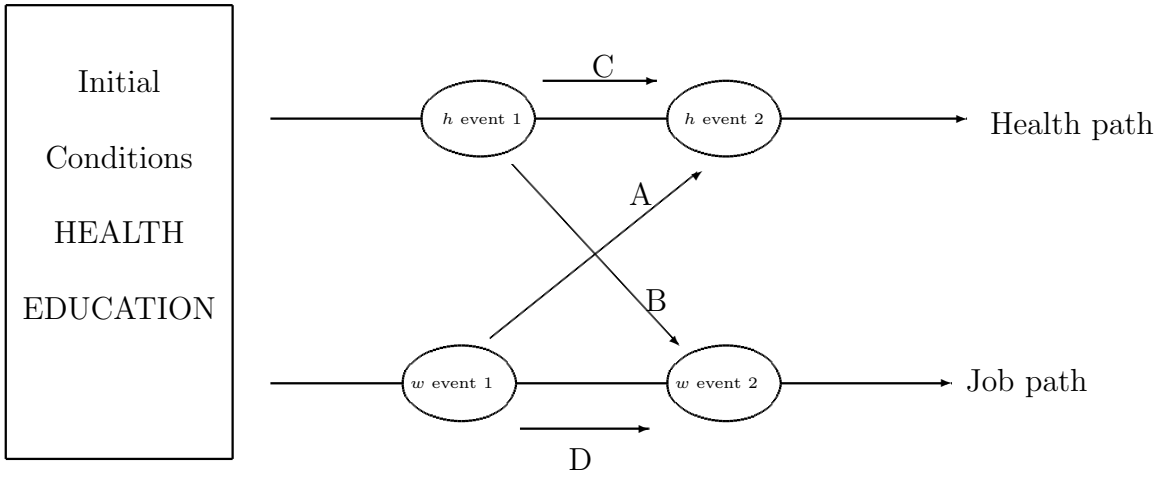


Figure 2.1: Dynamics of health and job status

In the basic example in Figure 2.1, four different interactions appear. Links A and B represent the effect of a health outcome (job status) at time $t-1$ on job status (health outcome) at time t. Inertia can also exist (links C and D, such that the probability of being in a good health condition at time $t-1$ influences the health condition at time t). Finally, various sets of control variables may influence h and w.

To identify all these links clearly, we used the causality concept, introduced by Granger (1969). It defines better predictability for a variable Y according to the use of its lag values, the lag value of another variable Z, and some controls X. Granger (1969) distinguishes instantaneous causality, such that Z_t is causing Y_t (if Z_t is included in the model, it improves the predictability of Y_t) from lag causality , in which case the lag values of Z improve the predictability of Y_t. In this section, we rule out instantaneous causality and deal with lag causality for one period.

The one-period Granger causality also can be regarded as conditional independence. Without loss of generality, we present the univariate case for time series. Let Y_t and Z_t denote some dependent variables and X_t denote a set of controls variables. One-period Granger non-causality from Z to Y is the conditional independence of Y_t from Z_{t-1} conditional on X_t and Y_{t-1}. Therefore, Granger non-causality from Z to Y is:

$$f(Y_t|Y_{t-1}, X_t, Z_{t-1}) = f(Y_t|Y_{t-1}, X_t). \tag{2.1}$$

Note that the same kind of relationship can be written for Granger non-causality from Y to Z. Because Y_t and Z_t are binary outcome variables, we can use latent variables (Y^* and Z^*), with the assumption that Y and Z have a positive outcomes (equal to 1) if their latent variables are positive. The latent variables are defined as follows :
For the left-hand term of Equation 2.1:

$$Y_t^* = X_t\beta_1 + \delta_{11}Y_{t-1} + \delta_{12}Z_{t-1} + \epsilon_t^1 \tag{2.2}$$

$$Z_t^* = X_t\beta_2 + \delta_{21}Y_{t-1} + \delta_{22}Z_{t-1} + \epsilon_t^2 \tag{2.3}$$

For the right-hand term of the Equation 2.1:

$$Y_t^* = X_t\beta_1 + \delta_{11}Y_{t-1} + \epsilon_t^1 \tag{2.4}$$

$$Z_t^* = X_t\beta_2 + \delta_{21}Z_{t-1} + \epsilon_t^2 \tag{2.5}$$

where

$$\begin{pmatrix} \epsilon^1 \\ \epsilon^2 \end{pmatrix} \rightsquigarrow N(0, \Sigma_\epsilon) \; with \; \Sigma_\epsilon = \begin{pmatrix} 1 & \rho_\epsilon \\ \rho_\epsilon & 1 \end{pmatrix}. \tag{2.6}$$

To fit the joint distribution of Y and Z conditional on X (such that we estimate a bivariate model), we need to analyze four available situations: $(Y = Z = 1)$, $(Y = Z = 0)$, $(Y = 1; Z = 0)$, and $(Y = 0; Z = 1)$. For each of these situations, we have:

$$p(Y_t = 1, Z_t = 1|X_t) = p(\epsilon_t^1 > -X_t\beta_1 - \delta_{11}Y_{t-1} - \delta_{12}Z_{t-1}, \epsilon_t^2 > -X_t\beta_2 - \delta_{21}Y_{t-1} - \delta_{22}Z_{t-1}) \tag{2.7}$$

$$p(Y_t = 0, Z_t = 0|X_t) = p(\epsilon_t^1 < -X_t\beta_1 - \delta_{11}Y_{t-1} - \delta_{12}Z_{t-1}, \epsilon_t^2 < -X_t\beta_2 - \delta_{21}Y_{t-1} - \delta_{22}Z_{t-1}) \tag{2.8}$$

$$p(Y_t = 1, Z_t = 0|X_t) = p(\epsilon_t^1 > -X_t\beta_1 - \delta_{11}Y_{t-1} - \delta_{12}Z_{t-1}, \epsilon_t^2 < -X_t\beta_2 - \delta_{21}Y_{t-1} - \delta_{22}Z_{t-1}) \tag{2.9}$$

$$p(Y_t = 0, Z_t = 1|X_t) = p(\epsilon_t^1 < -X_t\beta_1 - \delta_{11}Y_{t-1} - \delta_{12}Z_{t-1}, \epsilon_t^2 > -X_t\beta_2 - \delta_{21}Y_{t-1} - \delta_{22}Z_{t-1}) \tag{2.10}$$

By supposing that $q_t^1 = 2Y_t - 1$ and $q_t^2 = 2Z_t - 1$, we can rewrite these probabilities as:

$$p(Y_t, Z_t|X_t) = \Phi_2\Big(q_t^1(X_t\beta_1 + \delta_{11}Y_{t-1} + \delta_{12}Z_{t-1}), q_t^2(X_t\beta_2 + \delta_{21}Y_{t-1} + \delta_{22}Z_{t-1}), q_t^1 q_t^2 \rho_\epsilon\Big) \quad (2.11)$$

Where Φ_2 denotes the standard normal probability distribution function. Testing for Granger non-causality in this specification involves testing $\delta_{12} = 0$ for the prediction that Z is not causing Y and testing $\delta_{21} = 0$ for the prediction that Y is not causing Z.

2.2.2 Testing causality: panel data case

Two main approaches are available for panel data like the SIP survey. The first assumes that the causal effect is not the same for all individuals in the panel (Nair-Reichert and Weinhold, 2001). The specifications for the latent variables are:

$$Y_{it}^* = X_t\beta_1 + \delta_{11,i}Y_{i,t-1} + \delta_{12,i}Z_{i,t-1} + \eta_i^1 + \zeta_{it}^1 \quad (2.12)$$

$$Z_{it}^* = X_t\beta_2 + \delta_{21,i}Y_{i,t-1} + \delta_{22,i}Z_{i,t-1} + \eta_i^2 + \zeta_{it}^2, \quad (2.13)$$

where $(\eta_i^1, \eta_i^2)'$ denote the individual random effects that are the zero mean and covariance matrix Σ_η, and $(\zeta_{it}^1, \zeta_{it}^2)'$ denote the idiosyncratic shocks that are the zero mean and covariance matrix Σ_ζ, with

$$\Sigma_\eta = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_\eta \\ \sigma_1\sigma_2\rho_\eta & \sigma_2^2 \end{pmatrix} and\ \Sigma_\zeta = \begin{pmatrix} 1 & \rho_\zeta \\ \rho_\zeta & 1 \end{pmatrix}. \quad (2.14)$$

In this approach, testing Granger non-causality is equivalent to testing $\delta_{12,i} = 0, i = 1, ..., N$ for the prediction that Z is not causing Y and to testing $\delta_{21,i} = 0, i = 1, ..., N$ for the prediction that Y is not causing Z.

The second approach, which we use here, assumes that the causal effects, if they exist, are the same for all individuals in the panel. With the same notation, the latent variables are:

$$Y_{it}^* = X_t\beta_1 + \delta_{11}Y_{i,t-1} + \delta_{12}Z_{i,t-1} + \eta_i^1 + \zeta_{it}^1 \quad (2.15)$$

$$Z_{it}^* = X_t\beta_2 + \delta_{21}Y_{i,t-1} + \delta_{22}Z_{i,t-1} + \eta_i^2 + \zeta_{it}^2 \quad (2.16)$$

Testing for Granger noncausality is equivalent to testing $\delta_{12} = 0$ for the prediction that Z is not causing Y and to testing $\delta_{21} = 0$ for the prediction that Y is not causing Z.

2.2.3 Dealing with initial conditions

For the first wave of the panel (initial condition), we lack data for the previous state on Y and Z (we have no information on $Y_{i,0}$ and $Z_{i,0}$), so we cannot evaluate $P(Y_{i1}, Z_{i1}|Y_{i,0}, Z_{i,0}, X_i)$. By ignoring it

in the individual overall likelihood, we also ignore the data generation process for the first wave of the panel. We suppose the data generating process of the first wave of the panel is exogenous or in equilibrium. These assumptions hold only if the individual random effects are degenerated. Otherwise, the initial conditions (first wave of the panel) can be explained by the individual random effects, whereas ignoring them leads to inconsistent parameter estimates (Heckman, 1981).

The solution proposed by Heckman (1981) for the univariate case and generalized by Alessie et al. (2004) involves estimating a static equation for the first wave of the panel (i.e., we do not introduce lagged dependent variables). In this static equation, the random effects are a linear combination of the random effects in the next wave of the panel, and idiosyncratic error terms may have a different structure from the idiosyncratic error terms in the dynamic equation. Formally, the latent variables for the first wave of the panel are:

$$Y_{i1}^{*} = X_i^1\gamma_1 + \lambda_{11}\eta_i^1 + \lambda_{12}\eta_i^2 + \epsilon_i^1 \tag{2.17}$$

$$Z_{i1}^{*} = X_i^2\gamma_2 + \lambda_{21}\eta_i^1 + \lambda_{22}\eta_i^2 + \epsilon_i^2 \tag{2.18}$$

where $(\epsilon_i^1, \epsilon_i^2)'$ denote the idiosyncratic shocks, which include the zero mean and covariance matrix Σ_ϵ with $\Sigma_\epsilon = \begin{pmatrix} 1 & \rho_\epsilon \\ \rho_\epsilon & 1 \end{pmatrix}$.
Because η^1 and η^2 are individual random effects on Y and Z, λ_{12} and λ_{21} can be interpreted as the influence of the Y random individual effects (Z random individual effects) on Z (on Y) for the first wave of the panel.

2.2.4 Estimation methods for health and job paths

Finally, because we want to estimate the dynamics of health (h) and job status (w), we set the following equations for each time period ($t > 1$):

$$h_{it}^{*} = X_t\beta_1 + \delta_{11}h_{i,t-1} + \delta_{12}w_{i,t-1} + \eta_i^1 + \zeta_{it}^1 \tag{2.19}$$

$$w_{it}^{*} = X_t\beta_2 + \delta_{21}h_{i,t-1} + \delta_{22}w_{i,t-1} + \eta_i^2 + \zeta_{it}^2 \tag{2.20}$$

and for the initial conditions:

$$h_{i1}^{*} = X_i^1\gamma_1 + \lambda_{11}\eta_i^1 + \lambda_{12}\eta_i^2 + \epsilon_i^1 \tag{2.21}$$

$$w_{i1}^{*} = X_i^2\gamma_2 + \lambda_{21}\eta_i^1 + \lambda_{22}\eta_i^2 + \epsilon_i^2 \tag{2.22}$$

In Equations 2.19 to 2.22, many characteristics simultaneously affect health and labour market processes. To achieve the estimations, we also need at least two exclusion restrictions. The variable for the labour market status equation is the national unemployment rate (source: INSEE). The exclusion restriction for health status is set according to the physician per population ratio, also known as the medical density (Delattre and Dormont 2003). Equations 2.19 and 2.20 can be consistently estimated under assumption that η_i and ζ_{it} have symmetric distribution (See Heckman, 1981).

The individual level likelihood is given by :

$$L_i = \int_{\mathbb{R}^2} \Phi_2(q_{i0}^1 h_i^0, q_{i0}^2 w_i^0, q_{i0}^1 q_{i0}^2 \rho_\epsilon) \prod_{t=2}^{T_i} \Phi_2(q_{it}^1 \bar{h}_{it}, q_{it}^2 \bar{w}_{it}, q_{it}^1 q_{it}^2 \rho_\zeta) \phi(\eta_i, \Sigma_\eta) d\eta_i^1 d\eta_i^2 \tag{2.23}$$

Where

$$\begin{aligned}
q_{it}^1 &= 2y_{it}^1 - 1 \ \ \forall\, i,t \\
q_{it}^2 &= 2y_{it}^2 - 1 \ \ \forall\, i,t \\
h_i^0 &= Z_i^1 \gamma_1 + \lambda_{11}\eta_i^1 + \lambda_{12}\eta_i^2 \\
w_i^0 &= Z_i^2 \gamma_2 + \lambda_{21}\eta_i^1 + \lambda_{22}\eta_i^2 \\
\bar{h}_{it} &= X_{it}^1 \beta_1 + \delta_{11} h_{i,t-1} + \delta_{12} w_{i,t-1} + \eta_i^1 \\
\bar{w}_{it} &= X_{it}^2 \beta_2 + \delta_{21} h_{i,t-1} + \delta_{22} w_{i,t-1} + \eta_i^2
\end{aligned}$$

Because the likelihood function has an intractable form (integral function), it is impossible to estimate this likelihood with the usual methods. We therefore use numerical integration methods.

There are two main methods to estimate our likelihood function: the Gauss-Hermite quadrature (GHQ) and the maximum simulated likelihood (MSL). To choose a method, we consider the accuracy and the computing time requirement. For our estimations, we chose the adaptative Gauss-Hermite quadrature proposed by Liu and Pierce (1994)[8].

2.3 Results

We present econometric results in Table 2.4. In Table 2.4, columns (1) and (2) contain the results from bivariate probit regressions for Equations 2.19 and 2.20. In columns (1') and (2'), we also provide the

[8]Moussa and Delattre (2015) provide more details about how to make this choice.

univariate probit regressions (with no correlation between the two equations) for these equations. We do the same in Table 2.5 for the initial conditions (Equations 2.21 and 2.22).

The results clearly reveal persistence effects in the health ($\delta_{11} = 3.8243$) and employment ($\delta_{22} = 2.7444$) paths. As Haan and Myck (2009) suggest, we thus confirm the need to study these phenomena dynamically to explain the situation for each individual in terms of her or his health and employment at time t. Evidence for persistence effects also comes from the influence of initial conditions, which depend on various covariates (see Table 2.5).

We also find the expected, well-known effects of socio economic variables on initial health and employment status. Men are less likely to declare an illness and to be employed than women. Elderly people have worse health and job statuses than young people. People without French nationality report less illness and poorer job statuses. Family life also affects health and job conditions: Living as a couple lowers the probability of illness and job stability. The more children in the household, the more illness people experience, and the worse job conditions. Education level creates big differences. More educated people have a lower probability of illness and more likely to be employed. We include some interaction terms between gender and some others socio economic variables to account for gender discrimination in terms of health and job statuses[9]. Tables 2.14 and 2.15 contain estimated coefficients for the models with interacted variables with gender. We find that for men, living as a couple increases the probability of job stability but has no significant effect on the probability to report illness. Men with higher school grade are less likely to report illness than women. For initial condition, we also find that for men, living in couple and having an higher number of children increase the probability to enter job market. However, the higher the school grade for men at the initial state, the lower the probability to enter job market.

The main focus of this paper is on the causality between health and employment status. The bivariate estimates in Table 2.4 offer strong support. The impact of job status on health is reflected by the coefficient $\delta_{12} = 0.2288$, such that people who have a job at time $t-1$ are more likely to report an illness in the next period t (with an increase of 0.0842 in probability to report an illness at t, see Table 2.8 for marginal effects). Two factors could explain these results. First, it could highlight a job quality effect. If being employed involves poor conditions, employment status could readily increase the probability of illness, as argued by Debrand (2001). Unfortunately the SIP survey does not identify

[9]We are grateful to the reviewers of Journal of Applied Econometrics for their helpful comments on these issues.

longitudinal job quality, so we cannot identify the distinct effect of good or poor working conditions. Second, in France, the health care and insurance system is generous for employed people. For example, they may make regular appointments with their physician, which gives them access to more efficient health monitoring. As a result, they may be more likely to detect and report a disease.

Reporting an illness at time $t-1$ lowers the probability of having a job at time t ($\delta_{21} = -0.1927$). The marginal effect of illness on the probability of having a job is -0.0723. This result illustrates that an illness often makes it difficult to stay in a job or to find a new job (Currie and Madrian, 1999). Our main contribution is thus to conclude that health and employment status do not have a one way causality path but instead show a dual causality effect.

This result derives from taking into account three sources of bias, as described in Section 2.2: persistence effects, initial conditions, and unobserved heterogeneity. If all these biases were neglected, as in univariate probit models (columns (1') and (2) of Table 2.4), estimates of the causality effects between health and employment status would be biased. In our case, we would have wrongly concluded that being employed in previous year has no effect on health.

Finally, the existence of the causality between health and employment status also appears evident in Table 2.5. The coefficients λ_{11} and λ_{22} are both significant, confirming the need to integrate unobserved individual effects η in our model. In addition, the coefficient $\lambda_{12} > 0$ shows that the unobserved individual effect explaining job status (η^2) influences the value of health status at time $t = 1$. The method we have developed here is based on the existence of a correlation between unobservable variables in Equations 2.19 and 2.20 and those of Equations 2.21 and 2.22. Table 2.4 gives the values of these correlations. In equations for time $t > 1$ and the initial conditions, correlations between idiosyncratic components are not significant. Therefore, the main unobserved heterogeneity, responsible for the correlation, can be captured with individual-specific effects. In the main equations ($t > 1$), the correlation between individual-specific effects is negative. Therefore, we call for bivariate panel models to avoid any bias in the estimates. We also establish that individual unobserved factors that explain the probability of having a job ($w = 1$) are negatively correlated with individual unobserved factors that explain the probability of declaring an illness ($h = 1$). Among these unobserved factors, individual intrinsic motivation to job and job satisfaction appear to influence individuals' health[10].

Taking advantage of the two other indicators of illness (risk of death and disability, Table 2.2),

[10]Such as mental health (Faragher et al., 2005; Nadinloyi et al., 2013)

in Tables 2.6-2.7, we provide the estimation results with these variables. Using these two additional measures of self-reported illnesses gives support to our main results even if we cannot evaluate the bias in our measures (see Benitez-Silva et al., 2004). Table 2.6 contains the bivariate results for the indicator of disability (columns 1 and 2) and risk of death indicator (columns $1'$ and $2'$). The causality from poor health to job status is confirmed by the coefficients $\delta_{21} = -0.4418$ for the disability index and $\delta_{21} = -0.4981$ for the risk of death. Turning to marginal effects, we find that the same and even a stronger effect of health status on the probability of having a good quality job emerges, compared with the previous health indicator (the marginal effects are -0.1932 for the risk of death indicator, -0.1707 for the indicator of disability). The impact of health on job status is confirmed by the coefficients $\delta_{21} = -0.4418$ for the disability index and $\delta_{21} = -0.4981$ for the risk of death. The same and even a stronger effect of health status on the probability of having a good quality job emerges, compared with the previous health indicator $\delta_{21} = -0.1927$. When looking at the impact of job status on health, we find no significant effect, in contrast with our prior result. We offer two possible interpretations: First, good jobs provide access to better health coverage and increase the probability of reporting an illness (of any kind). Second, having a job is correlated with poor working conditions. When we control for the severity of health conditions, we find additional support for the first interpretation. Even if people appear induced to report an illness when they have a good job and insurance coverage, the illnesses they report are not particularly severe. We find that the marginal effects (see Tables 2.8-2.13) of having job on the probability to report an illness high disability index is close to $1.9e-6$ while this effect is negative $-1.6e-7$ on the probability to report an illness with risk of death, contrarily to illness regardless the degree of severity 0.0842. Tables 2.14 and 2.15 provide estimated coefficients for our two indicators of illness that account for interaction between gender and others socio economic variables. We find that the higher school grade effect for men does not remain significant. We also find a weak evidence that men with high number of children are less likely to report an illness with risk of disability than women. In terms of initial conditions, there is no discrimination between men and women for the probability of reporting any king of illness.

As with the main health indicator (Table 2.4), we find a significant correlation between individual-specific effects of health and the job status equations (Table 2.6). The interpretation of the positive sign of these correlations is rather complex. Some unobservable factors that explain the probability of having a job and severe health conditions simultaneously also correlate positively, such as the existence of specific policies designed to protect the job status of disabled persons.

Finally, and contrary to Haan and Myck's (2009; page 1124) claim that "accounting for unobserved

heterogeneity reduces the magnitude of the estimated coefficients on the lagged endogenous variables and significantly reduces the persistence of both processes", our estimates clearly show that causality links (A and B, Figure 2.1) are rather strong, regardless of the illness severity.

Conclusion

This article has examined the relationship between health and labour market paths. Many previous econometric results fail to account for all the links between health and job market status and thus cannot prove any causality. Instead, we propose a new method based on a bivariate dynamic probit model that acknowledges the simultaneity effects between the two phenomena, persistence effects, the role of the initial conditions, and the influence of unobserved heterogeneity. Using a French longitudinal survey we analyze complex interlinks between past and current levels of health and labour market paths. Our results regarding the causality between our two economic outcomes are innovative, due to the novel econometric methodology and the data set we use.

We demonstrate persistence in both processes. Being ill at $t-1$ is a significant determinant of current health status. Simultaneously, we observe the same persistence in labour market paths. We also confirm the impact of initial conditions, which depend on an individual attributes and macroeconomic conditions.

Taking advantage of this original econometric modelling, which allows us to distinguish between correlation and causality effects, we highlight some significant causal effects between employment and health processes. Being ill at $t-1$ is a significant determinant of current labour market status, and lagged employment has a positive effect on the probability of being ill at time t. In addition, we find an influence of unobserved heterogeneity on the causality effects. These effects are strengthened by the existence of individual-specific effects, which are correlated. When taking these effects into account in our bivariate model, we avoid many biases that univariate modelling cannot avoid.

Finally, our econometric methodology gives us robust estimates of the complex links between health and employment status. Our results therefore argue for a joint design, in France, of health and employment public policies taking interactions between health and employment into account.

Table 2.4: Estimates of health and job status interactions.

Part A: dynamic equations.

	Bivariate estimations		Univariate estimations	
Variables	h : *health*	w : *work*	h : *health*	w : *work*
	(1)	(2)	(1')	(2')
h_{-1}	3.8243^{***} (0.0225)	-0.1927^{***} (0.0138)	4.2513^{***} (0.0154)	-0.2704^{***} (0.0151)
w_{-1}	0.2298^{***} (0.0235)	2.7444^{***} (0.0126)	0.0179 (0.0190)	2.7844^{***} (0.0137)
Gender (male)	-0.1571^{***} (0.0169)	0.8095^{***} (0.0137)	-0.0436^{***} (0.0127)	0.5463^{***} (0.0160)
Age	0.0373^{***} (0.0010)	-0.0227^{***} (0.0007)	0.0114^{***} (0.0008)	-0.0089^{***} (0.0007)
Not French^{+}	-0.0481^{*} (0.0250)	-0.3435^{***} (0.0162)	-0.0217 (0.0191)	-0.2375^{***} (0.0211)
Couple	-0.0205 (0.0181)	-0.1211^{***} (0.0137)	-0.0401^{***} (0.0149)	-0.0991^{***} (0.0142)
Number of children	0.0261^{***} (0.0071)	-0.0758^{***} (0.0052)	0.0168^{***} (0.0057)	-0.0487^{***} (0.0058)
No grade	0.3808^{***} (0.0372)	-0.9158^{***} (0.0269)	0.1246^{***} (0.0279)	-0.5823^{***} (0.0324)
College grade	0.3067^{***} (0.0267)	-0.5544^{***} (0.0217)	0.0913^{***} (0.0194)	-0.2787^{***} (0.0239)
High school grade	0.2427^{***} (0.0317)	-0.3639^{***} (0.0253)	0.0811^{***} (0.0232)	-0.1825^{***} (0.0282)
Undergraduate studies	0.0856^{**} (0.0369)	-0.1470^{***} (0.03)	0.0287 (0.0269)	-0.0925^{***} (0.0326)
Ref : Graduate studies	-	-	-	-
Medical density	-0.0009^{***} (0.0003)	-	0.0018^{***} (0.0003)	-
Unemployment rate	-	0.0714^{***} (0.0022)	-	0.0249^{***} (0.0024)
Intercept	-2.6603^{***} (0.0563)	-1.4684^{***} (0.0319)	-2.8887^{***} (0.0464)	-0.3645^{***} (0.0353)
Covariance matrix	$\sigma_1 = 1.3631^{***}$ (0.0184), $\sigma_2 = 1.7269^{***}$ (0.0161)		-	
	$\rho_\eta = -0.8259^{***}$ (0.0054), $\rho_\zeta = 0.0275$ (0.0174)		-	

The estimated standard errors are within parenthesis.

***: Significant at the 1% level; **: Significant at the 5% level.

*: Significant at the 10% level. +: Refers to the individual's nationality.

Table 2.5: Estimates of health and job status interactions.

Part B: the initial conditions.

	Bivariate estimations		Univariate estimations	
Variables	h : *health*	w : *work*	h : *health*	w : *work*
	Initial conditions			
Gender	-0.2425^{***} (0.0616)	0.1555^{***} (0.0319)	-0.1744^{***} (0.0457)	0.144^{***} (0.0317)
Age	-0.0048 (0.0157)	0.0318^{***} (0.0082)	-0.0127 (0.0120)	0.0347^{***} (0.0081)
$Not\ French^{+}$	-0.227^{**} (0.1009)	-0.4552^{***} (0.0434)	-0.2062^{***} (0.0771)	-0.452^{***} (0.0432)
Couple	0.0347 (0.08)	0.1579^{***} (0.0468)	0.0427 (0.0607)	0.1526^{***} (0.0466)
Number of children	-0.0213 (0.1354)	-0.5407^{***} (0.0619)	0.0143 (0.0986)	-0.5478^{***} (0.0616)
No grade	0.1925 (0.1666)	-0.592^{***} (0.0863)	0.0625 (0.1277)	-0.5782^{***} (0.0858)
College grade	0.0659 (0.1180)	-0.102 (0.0655)	0.0114 (0.0893)	-0.083 (0.0652)
High school grade	-0.0508 (0.1149)	-0.2165^{***} (0.0622)	-0.0772 (0.0868)	-0.2014^{***} (0.0619)
Undergraduate studies	-0.0836 (0.1157)	-0.0045 (0.0661)	-0.0477 (0.0863)	-0.0001 (0.0659)
Ref : Graduate studies	-	-	-	-
Medical density	0.0005 (0.0008)	-	0.0026^{***} (0.0006)	
Unemployment rate	-	-0.0001 (0.0048)	-	-0.0064 (0.0045)
$Ill\ before\ prof.\ life$	0.3626^{***} (0.0122)	-0.0018^{***} (0.0047)	0.3465^{***} (0.0090)	-0.0031 (0.0044)
Intercept	-1.5796^{***} (0.3553)	0.429^{**} (0.1864)	-1.8018^{***} (0.2623)	0.5483^{***} (0.1839)
λ_{11}	1.2085^{***} (0.0639)	-		
λ_{12}	0.3967^{***} (0.0557)	-		
λ_{21}	-	0.0324 (0.0296)		
λ_{22}	-	0.1242^{***} (0.0261)		
Covariance matrix	$\rho_\epsilon = 0.00227$ (0.0460)		-	

The estimated standard errors are within parenthesis.

***: Significant at the 1% level, **: Significant at the 5% level.

*: Significant at the 10% level, +: Refers to the individual's nationality.

Table 2.6: Estimates of health and job status interactions.

Part A: dynamic equations

	Disability index		Risk of death	
Variables	h : *disab*	w : *work*	h : *rdeath*	w : *work*
	(1)	(2)	(1')	(2')
h_{-1}	4.0503^{***} (0.0498)	-0.4418^{***} (0.0328)	3.7859^{***} (0.0502)	-0.4981^{***} (0.0366)
w_{-1}	0.0247 (0.0554)	2.737^{***} (0.0127)	-0.0026 (0.0565)	2.7359^{***} (0.0127)
Gender	0.0432^{***} (0.0376)	0.8349^{***} (0.0142)	-0.0256 (0.0395)	0.8325^{***} (0.0142)
Age	0.0066^{***} (0.0023)	-0.025^{***} (0.0007)	0.0129^{***} (0.0023)	-0.0249^{***} (0.0007)
$Not\ French^{+}$	-0.1478^{**} (0.0595)	-0.3253^{***} (0.0165)	-0.1146^{*} (0.0618)	-0.3219^{***} (0.0165)
Couple	-0.0239 (0.0418)	-0.1207^{***} (0.0138)	-0.1464^{***} (0.0425)	-0.1236^{***} (0.0138)
Number of children	0.008 (0.0161)	-0.072^{***} (0.0053)	0.0093 (0.0167)	-0.0731^{***} (0.0053)
No grade	0.0603 (0.0837)	-0.9518^{***} (0.0276)	-0.0412 (0.0837)	-0.961^{***} (0.0276)
College grade	0.0354 (0.0617)	-0.5797^{***} (0.0224)	-0.0594 (0.0602)	-0.584^{***} (0.0224)
High school grade	0.1002 (0.0722)	-0.3874^{***} (0.026)	-0.0561 (0.0731)	-0.3914^{***} (0.026)
Undergraduate studies	0.0537 (0.0839)	-0.1516^{***} (0.0312)	-0.0304 (0.0846)	-0.1598^{***} (0.0312)
Ref : Graduate studies	-	-	-	-
Medical density	0.0059^{***} (0.0008)	-	0.0057^{***} (0.0009)	-
Unemployment rate	-	0.0727^{***} (0.0023)	-	0.0726^{***} (0.0023)
Intercept	-5.5495^{***} (0.1396)	-1.395^{***} (0.0324)	-5.5168^{***} (0.1539)	-1.3894^{***} (0.0323)
Covariance matrix	$\sigma_1 = 1.0683^{***}$ (0.0123), $\sigma_2 = 1.701^{***}$ (0.0163)		$\sigma_1 = 1.0143^{***}$ (0.0122), $\sigma_2 = 1.701^{***}$ (0.0163)	
	$\rho_\eta = 0.2708^{***}$ (0.032), $\rho_\zeta = 0.0468$ (0.0482)		$\rho_\eta = 0.2284^{***}$ (0.0328), $\rho_\zeta = 0.0175$ (0.05)	

The estimated standard errors are within parenthesis.

***: Significant at the 1% level, **: Significant at the 5% level.

*: Significant at the 10% level, +: Refers to the individual's nationality.

Table 2.7: Estimates of health and job status interactions.

Part B: the initial conditions

	Disability index		Risk of death	
Variables	h : *disab*	w : *work*	h : *rdeath*	w : *work*
	Initial conditions			
Gender	0.099 (0.1597)	0.16*** (0.032)	−0.133 (0.2002)	0.1612*** (0.032)
Age	0.059 (0.0418)	0.0312*** (0.0082)	0.0274 (0.0529)	0.0311*** (0.0082)
$Not\,French^{+}$	−0.6606 (0.4388)	−0.4523*** (0.0435)	−0.9179 (0.5886)	−0.4547*** (0.0436)
Couple	−0.2377 (0.2411)	0.161*** (0.0469)	0.1184 (0.2623)	0.1596*** (0.047)
Number of children	−0.4369 (0.4986)	−0.5376*** (0.0621)	−0.4753 (0.5308)	−0.537*** (0.0622)
No grade	0.8817** (0.4354)	−0.5952*** (0.0865)	0.3794 (0.5088)	−0.5952*** (0.0867)
College grade	0.6666** (0.3344)	−0.1079 (0.0656)	0.1311 (0.3905)	−0.1091* (0.0658)
High school grade	0.2806 (0.3247)	−0.2203*** (0.0623)	−0.1777 (0.3767)	−0.2211*** (0.0624)
Undergraduate studies	0.4078 (0.3213)	−0.0052 (0.0662)	−0.1988 (0.3912)	−0.0061 (0.0664)
Ref : Graduate studies	-	-	-	-
Medical density	0.0085*** (0.0021)	-	0.0111*** (0.0028)	
Unemployment rate	-	0.0028 (0.0048)	-	0.002 (0.0048)
$Ill\,before\,prof.life$	0.1403*** (0.0137)	−0.0012 (0.0045)	0.1381*** (0.0167)	−0.001 (0.0045)
Intercept	−8.3097*** (1.0974)	0.3943** (0.1864)	−8.3287*** (1.3685)	0.4034** (0.1867)
λ_{11}	1.6651*** (0.1381)	-	1.9251*** (0.1822)	
λ_{12}	−0.0464 (0.1057)	-	0.0427 (0.127)	
λ_{21}	-	0.0117 (0.032)		−0.0582* (0.0339)
λ_{22}	-	0.1276*** (0.0212)		0.1377*** (0.021)
Covariance matrix	$\rho_\epsilon = -0.1515$ (0.1161)		$\rho_\epsilon = 0.0175$ (0.1522)	

The estimated standard errors are within parenthesis.

***: Significant at the 1% level, **: Significant at the 5% level.

*: Significant at the 10% level, +: Refers to the individual's nationality.

Table 2.8: Estimates marginal effects on joint, marginal and conditional probabilities.

Part A: dynamic equations.

Variables	$p(h=1,w=1)$	$p(h=0,w=1)$	$p(h=1,w=0)$	$p(h=0,w=0)$	$p(h=1)$	$p(w=1)$	$p(h=1\|w=1)$	$p(w=1\|h=1)$
h_{-1}	0.5094^{***} (0.0011)	0.3601^{***} (0.0011)	-0.5817^{***} (0.0011)	-0.2878^{***} (0.0011)	0.8695^{***} (0.0005)	-0.0723^{***} (0.0008)	-0.0884^{***} (0.0008)	0.8667^{***} (0.0004)
w_{-1}	0.2955^{***} (0.0059)	-0.2113^{***} (0.0059)	0.4618^{***} (0.0059)	-0.546^{***} (0.0059)	0.0842^{***} (0.0014)	0.7573^{***} (0.0128)	0.7635^{***} (0.0125)	0.0654^{***} (0.0012)
Gender (male)	0.0658^{***} (0.0042)	-0.125^{***} (0.0042)	0.2212^{***} (0.0042)	-0.162^{***} (0.0042)	-0.0592^{***} (0.002)	0.287^{***} (0.0079)	0.2846^{***} (0.0077)	-0.064^{***} (0.0019)
Age	0.0062^{***} (0.0006)	0.0079^{***} (0.0006)	-0.0146^{***} (0.0006)	0.0005 (0.0006)	0.0141^{***} (0.0007)	-0.0084^{***} (0.0003)	-0.0085^{***} (0.0003)	0.0143^{***} (0.0007)
$Not\ French^{+}$	-0.0587^{***} (0.0074)	0.0406^{***} (0.0074)	-0.0728^{***} (0.0074)	0.0909^{***} (0.0074)	-0.0181^{***} (0.0023)	-0.1315^{***} (0.0174)	-0.1301^{***} (0.0173)	-0.0162^{***} (0.002)
Couple	-0.0215^{***} (0.0055)	0.0137^{**} (0.0055)	-0.0226^{***} (0.0055)	0.0304^{***} (0.0055)	-0.0078^{***} (0.0019)	-0.0441^{***} (0.0116)	-0.0434^{***} (0.0114)	-0.0071^{***} (0.0017)
Number of children	-0.0038 (0.0024)	0.0136^{***} (0.0024)	-0.0241^{***} (0.0024)	0.0143^{***} (0.0024)	0.0099^{***} (0.0015)	-0.0279^{***} (0.004)	-0.0277^{***} (0.004)	0.0103^{***} (0.0014)
No grade	-0.0776^{***} (0.0127)	0.2265^{***} (0.0127)	-0.2754^{***} (0.0127)	0.1264^{***} (0.0127)	0.1489^{***} (0.0087)	-0.3529^{***} (0.019)	-0.3552^{***} (0.0191)	0.1556^{***} (0.0083)
College grade	0.0034 (0.01)	0.1116^{***} (0.01)	-0.2033^{***} (0.01)	0.0883^{***} (0.01)	0.1151^{***} (0.0075)	-0.1999^{***} (0.0136)	-0.1996^{***} (0.0134)	0.1187^{***} (0.0073)
High school grade	0.0013 (0.0122)	0.0925^{***} (0.0122)	-0.1403^{***} (0.0122)	0.0465^{***} (0.0122)	0.0938^{***} (0.0101)	-0.1391^{***} (0.015)	-0.1398^{***} (0.0148)	0.0964^{***} (0.0098)
Undergraduate studies	-0.0003 (0.0136)	0.033^{**} (0.0136)	-0.055^{***} (0.0136)	0.0223 (0.0136)	0.0327^{***} (0.0106)	-0.0553^{***} (0.0178)	-0.0554^{***} (0.0175)	0.0337^{***} (0.0104)
Medical density	-0.0002 (0.0001)	-0.0001 (0.0001)	0.0002 (0.0001)	0.0001 (0.0001)	-0.0003 (0.0002)	$-$	$6.45e-6$ ($4.51e-6$)	-0.0003 (0.0002)
Unemployment rate	0.0097^{***} (0.0004)	-0.0097^{***} (0.0004)	0.0166^{***} (0.0004)	-0.0166^{***} (0.0004)	$-$	0.0263^{***} (0.0012)	0.026^{***} (0.0012)	-0.0004^{***} (0.00002)

The estimated standard errors are within parenthesis; ***: Significant at the 1% level; **: Significant at the 5% level; *: Significant at the 10% level.

+: Refers to the individual's nationality; h : *health*; w : *work*

Table 2.9: Estimates marginal effects on joint, marginal and conditional probabilities.

Part B: initial conditions.

Variables	$p(h=1,w=1)$	$p(h=0,w=1)$	$p(h=1,w=0)$	$p(h=0,w=0)$	$p(h=1)$	$p(w=1)$	$p(h=1\|w=1)$	$p(w=1\|h=1)$
Gender (male)	-0.0286^{***} (0.007)	-0.0094 (0.007)	0.0678^{***} (0.007)	-0.0297^{***} (0.007)	-0.0381^{***} (0.0075)	0.0392^{***} (0.0086)	0.0387^{***} (0.0079)	-0.0387^{***} (0.0075)
Age	0.00004 (0.0004)	-0.0008^{*} (0.0004)	0.008^{***} (0.0004)	-0.0073^{***} (0.0004)	-0.0008^{***} (0.0003)	0.0081^{***} (0.0022)	0.0078^{***} (0.0019)	-0.0008^{***} (0.0003)
Not French^{+}	-0.0349^{***} (0.006)	0.0029 (0.006)	-0.099^{***} (0.006)	0.131^{***} (0.006)	-0.032^{***} (0.0057)	-0.1339^{***} (0.0331)	-0.1282^{***} (0.0319)	-0.0318^{***} (0.0057)
Couple	0.008 (0.0054)	-0.0024 (0.0054)	0.0301^{***} (0.0054)	-0.0357^{***} (0.0054)	0.0056^{***} (0.0038)	0.0382^{***} (0.0231)	0.0364^{***} (0.0219)	0.0054^{***} (0.0037)
Number of children	-0.0144^{***} (0.0051)	0.011^{**} (0.0051)	-0.1227^{***} (0.0051)	0.1261^{***} (0.0051)	-0.0034^{***} (0.0011)	-0.1371^{***} (0.0493)	-0.1316^{***} (0.0474)	-0.0027^{***} (0.0009)
No grade	0.0075 (0.0165)	0.0266 (0.0165)	-0.1912^{***} (0.0165)	0.1571^{***} (0.0165)	0.0342^{***} (0.0124)	-0.1837^{***} (0.0701)	-0.1798^{***} (0.0685)	0.0357^{***} (0.0122)
College grade	0.0066 (0.0127)	0.0039 (0.0127)	-0.0325^{**} (0.0127)	0.022^{*} (0.0127)	0.0105^{***} (0.0122)	-0.0259^{***} (0.0294)	-0.0252^{***} (0.028)	0.0107^{***} (0.0121)
High school grade	-0.0113 (0.0073)	0.0034 (0.0073)	-0.0473^{***} (0.0073)	0.0552^{***} (0.0073)	-0.0079^{***} (0.005)	-0.0585^{***} (0.0425)	-0.0562^{***} (0.041)	-0.0077^{***} (0.0048)
Undergraduate studies	-0.0107 (0.0211)	-0.002 (0.0211)	0.0096 (0.0211)	0.0031 (0.0211)	-0.0127^{***} (0.025)	-0.0012^{***} (0.0019)	-0.0007^{***} (0.001)	-0.0128^{***} (0.0253)
Medical density	$7.82e-6$ (0.0008)	0.0001 (0.0008)	$-7.82e-6$ (0.0008)	-0.0001 (0.0008)	0.0001^{***} (0.0003)	–	$-1.69e-6^{***}$ (7.52e−6)	0.0001^{***} (0.0008)
Unemployment rate	$3.08e-6$ (0.0048)	$-3.08e-6$ (0.0048)	0.00003 (0.0048)	-0.00003 (0.0048)	–	0.00004^{***} (0.0048)	0.00004^{***} (0.0048)	$-2e-7^{***}$ (1.79e−8)
Illness before prof. life	0.0482^{***} (0.0033)	0.0094^{***} (0.0033)	-0.0487^{***} (0.0033)	-0.009^{***} (0.0033)	0.0577^{***} (0.0085)	-0.0005^{***} (0.0085)	-0.0021^{***} (0.0085)	0.0582^{***} (0.0085)

The estimated standard errors are within parenthesis; ***: Significant at the 1% level; **: Significant at the 5% level; *: Significant at the 10% level.

+: Refers to the individual's nationality; h : *health*; w : *work*

Table 2.10: Estimates marginal effects on joint, marginal and conditional probabilities.

Part A: dynamic equations.

Variables	$p(h=1,w=1)$	$p(h=0,w=1)$	$p(h=1,w=0)$	$p(h=0,w=0)$	$p(h=1)$	$p(w=1)$	$p(h=1\|w=1)$	$p(w=1\|h=1)$
h_{-1}	0.217^{***} (0.016)	0.2089^{***} (0.016)	-0.3877^{***} (0.016)	-0.0382^{**} (0.016)	0.4259^{***} (0.0292)	-0.1707^{***} (0.0032)	-0.2038^{***} (0.0027)	0.4345^{***} (0.0293)
w_{-1}	0.00001^{***} ($3.67e{-7}$)	-0.00001^{***} ($3.67e{-7}$)	0.7629^{***} ($3.67e{-7}$)	-0.7629^{***} ($3.67e{-7}$)	$1.89e-6^{***}$ ($5.09e{-8}$)	0.7629^{***} (0.0195)	0.7875^{***} (0.0175)	$-2.51e-6^{***}$ ($1.10e{-8}$)
Gender (male)	$7.53e-6^{***}$ ($4.31e{-7}$)	$-4.07e-6^{***}$ ($4.31e{-7}$)	0.292^{***} ($4.31e{-7}$)	-0.292^{***} ($4.31e{-7}$)	$3.46e-6^{***}$ ($2.28e{-7}$)	0.292^{***} (0.0133)	0.2749^{***} (0.0118)	$2.68e-6^{***}$ ($1.88e{-7}$)
Age	$2.13e-7^{**}$ ($9.40e{-8}$)	$3.11e-7^{***}$ ($9.40e{-8}$)	-0.0091^{***} ($9.40e{-8}$)	0.0091^{***} ($9.40e{-8}$)	$5.24e-7^{***}$ ($1.08e{-7}$)	-0.0091^{***} (0.0011)	-0.0086^{***} (0.0011)	$5.85e-7^{***}$ ($1.11e{-7}$)
Not French^{+}	$-8.03e-6^{***}$ ($8.67e{-7}$)	$-1.43e-6^{*}$ ($8.67e{-7}$)	-0.1233^{***} ($8.67e{-7}$)	0.1233^{***} ($8.67e{-7}$)	$-9.46e-6^{***}$ ($1.09e{-6}$)	-0.1233^{***} (0.0207)	-0.1167^{***} (0.0199)	$-9.78e-6^{***}$ ($1.14e{-6}$)
Couple	$-2.16e-6^{***}$ ($7.70e{-7}$)	$2.10e-7$ ($7.70e{-7}$)	-0.0434^{***} ($7.70e{-7}$)	0.0434^{***} ($7.70e{-7}$)	$-1.95e-6^{***}$ ($6.91e{-7}$)	-0.0434^{**} (0.0172)	-0.0406^{**} (0.0162)	$-1.91e-6^{***}$ ($6.79e{-7}$)
Number of children	$-3.80e-9$ ($2.69e{-7}$)	$6.38e-7^{**}$ ($2.69e{-7}$)	-0.0262^{***} ($2.69e{-7}$)	0.0262^{***} ($2.69e{-7}$)	$6.34e-7^{***}$ ($2.00e{-7}$)	-0.0262^{***} (0.0072)	-0.0248^{***} (0.0069)	$7.59e-7^{***}$ ($1.92e{-7}$)
No grade	$-4.50e-6^{***}$ ($1.17e{-6}$)	$9.83e-6^{***}$ ($1.17e{-6}$)	-0.3657^{***} ($1.17e{-6}$)	0.3657^{***} ($1.17e{-6}$)	$5.33e-6^{***}$ ($6.72e{-7}$)	-0.3657^{***} (0.0378)	-0.3627^{***} (0.0395)	$7.37e-6^{***}$ ($5.27e{-7}$)
College grade	$-1.56e-6^{*}$ ($8.52e{-7}$)	$4.36e-6^{***}$ ($8.52e{-7}$)	-0.2063^{***} ($8.52e{-7}$)	0.2063^{***} ($8.52e{-7}$)	$2.79e-6^{***}$ ($4.22e{-7}$)	-0.2063^{***} (0.0311)	-0.1947^{***} (0.0303)	$3.65e-6^{***}$ ($3.38e{-7}$)
High school grade	$2.99e-6$ ($2.13e{-6}$)	$6.25e-6^{***}$ ($2.13e{-6}$)	-0.1472^{***} ($2.13e{-6}$)	0.1471^{***} ($2.13e{-6}$)	$9.24e-6^{***}$ ($2.29e{-6}$)	-0.1472^{***} (0.0309)	-0.1422^{***} (0.0301)	0.00001^{***} ($2.32e{-6}$)
Undergraduate studies	$2.12e-6$ ($2.66e{-6}$)	$2.57e-6$ ($2.66e{-6}$)	-0.0565^{***} ($2.66e{-6}$)	0.0565^{***} ($2.66e{-6}$)	$4.69e-6$ ($2.98e{-6}$)	-0.0565^{*} (0.0323)	-0.0542^{*} (0.0307)	$5.19e-6^{*}$ ($3.05e{-6}$)
Medical density	$3.34e-7$ ($4.55e{-7}$)	$1.39e-7$ ($4.55e{-7}$)	$-3.34e-7$ ($4.55e{-7}$)	$-1.39e-7$ ($4.55e{-7}$)	$4.73e-7$ ($6.45e{-7}$)	–	-0.0001^{***} (-0.00002)	$5.01e-7$ ($6.83e{-7}$)
Unemployment rate	$4.56e-7^{***}$ ($1.83e{-8}$)	$-4.56e-7^{***}$ ($1.83e{-8}$)	0.0265^{***} ($1.83e{-8}$)	-0.0265^{***} ($1.83e{-8}$)	–	0.0265^{***} (0.0012)	0.025^{***} (0.001)	$-8.73e-8^{***}$ (0.0023)

The estimated standard errors are within parenthesis; ***: Significant at the 1% level; **: Significant at the 5% level; *: Significant at the 10% level.

+: Refers to the individual's nationality; h : *Ill with disability*; w : *work*

Table 2.11: Estimates marginal effects on joint, marginal and conditional probabilities.

Part B: initial conditions.

Variables	$p(h=1,w=1)$	$p(h=0,w=1)$	$p(h=1,w=0)$	$p(h=0,w=0)$	$p(h=1)$	$p(w=1)$	$p(h=1\|w=1)$	$p(w=1\|h=1)$
Gender (male)	$1.38e-8$ (1.34e−8)	$1.20e-9$ (1.34e−8)	0.0415^{***} (1.34e−8)	-0.0415^{***} (1.34e−8)	$1.50e-8$ (1.44e−8)	0.0415 (0.0284)	0.0361 (0.0244)	$1.53e-8$ (1.47e−8)
Age	$7.55e-9$ (3.16e−8)	$1.09e-9$ (3.16e−8)	0.0081^{***} (3.16e−8)	-0.0081^{***} (3.16e−8)	$8.64e-9$ (3.67e−8)	0.0081^{***} (0.002)	0.0069 (0.4842)	$8.92e-9$ (3.64e−8)
Not French^{+}	$-3.36e-8^{***}$ (1.14e−9)	$-5.18e-9^{***}$ (1.14e−9)	-0.136^{***} (1.14e−9)	0.136^{***} (1.14e−9)	$-3.88e-8^{***}$ (1.46e−9)	-0.136^{**} (0.0683)	-0.1184^{*} (0.0608)	$-4.01e-8^{***}$ (1.54e−9)
Couple	$-2.08e-8^{***}$ (7.80e−9)	$-4.16e-9$ (7.80e−9)	0.0401^{***} (7.80e−9)	-0.0401^{***} (7.80e−9)	$-2.49e-8^{***}$ (8.62e−9)	0.0401 (0.0269)	0.0362 (0.0224)	$-2.60e-8^{***}$ (8.86e−9)
Number of children	$-5.78e-8$ (0.4986)	$-6.83e-9$ (0.4986)	-0.1403 (0.4986)	0.1403 (0.4986)	$-6.46e-8$ (0.4986)	-0.1403^{*} (0.0825)	-0.1024 (0.2549)	$-6.57e-8^{***}$ (2.14e−9)
No grade	$1.26e-6$ (2.06e−6)	$5.90e-7$ (2.06e−6)	-0.1886^{***} (2.06e−6)	0.1886^{***} (2.06e−6)	$1.85e-6$ (2.82e−6)	-0.1886^{**} (0.0783)	-0.179^{**} (0.0723)	$1.97e-6$ (2.96e−6)
College grade	$1.41e-7$ (2.65e−7)	$2.74e-8$ (2.65e−7)	-0.0282^{***} (2.65e−7)	0.0282^{***} (2.65e−7)	$1.68e-7$ (3.14e−7)	-0.0282^{*} (0.0151)	-0.0283^{**} (0.0117)	$1.75e-7$ (3.26e−7)
High school grade	$5.84e-8$ (9.28e−8)	$1.55e-8$ (9.28e−8)	-0.0613^{***} (9.28e−8)	0.0613^{***} (9.28e−8)	$7.39e-8$ (1.10e−7)	-0.0613 (0.0515)	-0.0562 (0.0453)	$7.75e-8$ (1.14e−7)
Undergraduate studies	$1.41e-7$ (3.18e−7)	$2.49e-8$ (3.18e−7)	-0.0014^{***} (3.18e−7)	0.0014^{***} (3.18e−7)	$1.65e-7$ (3.75e−7)	-0.0014 (0.001)	-0.0033^{***} (0.001)	$1.71e-7$ (3.88e−7)
Medical density	$2.e-13$ (3.95e−9)	$1.25e-9$ (3.95e−9)	0 (3.95e−9)	$-1.25e-9$ (3.95e−9)	$1.25e-9$ (2.14e−3)	–	$7.68e-6$ (0.0021)	$7.68e-6$ (0.0021)
Unemployment rate	$8.59e-12$ (0.0048)	$-8.59e-12$ (0.0048)	0.0007 (0.0048)	-0.0007 (0.0048)	–	0.0007 (0.0048)	0.0003 (0.0048)	$-1.33e-11$ (0.0048)
Illness before prof. life	$1.75e-8$ (0.0137)	$3.06e-9$ (0.0137)	-0.0003 (0.0137)	0.0003 (0.0137)	$2.06e-8$ (−1.39e−7)	-0.0003 (0.0091)	-0.0014 (0.0091)	$2.13e-8$ (1.98e−7)

The estimated standard errors are within parenthesis; ***: Significant at the 1% level; **: Significant at the 5% level; *: Significant at the 10% level.

+: Refers to the individual's nationality; h : *Ill with disability*; w : *work*

Table 2.12: Estimates marginal effects on joint, marginal and conditional probabilities.

Part A: dynamic equations.

Variables	$p(h=1,w=1)$	$p(h=0,w=1)$	$p(h=1,w=0)$	$p(h=0,w=0)$	$p(h=1)$	$p(w=1)$	$p(h=1\|w=1)$	$p(w=1\|h=1)$
h_{-1}	0.1551^{***} (0.0144)	0.162^{***} (0.0144)	-0.3483^{***} (0.0144)	0.0312^{**} (0.0144)	0.3171^{***} (0.0273)	-0.1932^{***} (0.004)	-0.2243^{***} (0.0034)	0.3253^{***} (0.0276)
w_{-1}	0.00001^{***} ($2.50e{-7}$)	-0.00001^{***} ($2.50e{-7}$)	0.7633^{***} ($2.50e{-7}$)	-0.7633^{***} ($2.50e{-7}$)	$-1.61e-7^{***}$ ($2.50e{-9}$)	0.7633^{***} (0.0199)	0.7882^{***} (0.0179)	$-3.96e-6^{***}$ ($-4.70e{-8}$)
Gender (male)	$2.69e-6^{***}$ ($2.38e{-7}$)	$-4.25e-6^{***}$ ($2.38e{-7}$)	0.291^{***} ($2.38e{-7}$)	-0.291^{***} ($2.38e{-7}$)	$-1.56e-6^{***}$ ($8.87e{-8}$)	0.291^{***} (0.0141)	0.2741^{***} (0.0125)	$-2.40e-6^{***}$ ($5.87e{-8}$)
Age	$4.39e-7^{***}$ ($1.45e{-7}$)	$3.49e-7^{**}$ ($1.45e{-7}$)	-0.009^{***} ($1.45e{-7}$)	0.009^{***} ($1.45e{-7}$)	$7.87e-7^{***}$ ($1.90e{-7}$)	-0.009^{***} (0.0009)	-0.0086^{***} (0.0009)	$8.59e-7^{***}$ ($1.99e{-7}$)
Not French^{+}	$-5.29e-6^{***}$ ($6.91e{-7}$)	$-6.11e-7$ ($6.91e{-7}$)	-0.1219^{***} ($6.91e{-7}$)	0.1219^{***} ($6.91e{-7}$)	$-5.90e-6^{***}$ ($8.17e{-7}$)	-0.1219^{***} (0.0228)	-0.1155^{***} (0.022)	$-6.04e-6^{***}$ ($8.44e{-7}$)
Couple	$-8.17e-6^{***}$ ($1.20e{-6}$)	$-2.26e-6^{*}$ ($1.20e{-6}$)	-0.0444^{***} ($1.20e{-6}$)	0.0444^{***} ($1.20e{-6}$)	-0.00001^{***} ($1.58e{-6}$)	-0.0444^{***} (0.0095)	-0.0404^{***} (0.0087)	-0.00001^{***} ($1.66e{-6}$)
Number of children	$5.30e-8$ ($2.13e{-7}$)	$5.13e-7^{**}$ ($2.13e{-7}$)	-0.0266^{***} ($2.13e{-7}$)	0.0266^{***} ($2.13e{-7}$)	$5.66e-7^{***}$ ($1.74e{-7}$)	-0.0266^{***} (0.0073)	-0.0251^{***} (0.0071)	$6.68e-7^{***}$ ($1.57e{-7}$)
No grade	$-6.01e-6^{***}$ ($5.60e{-7}$)	$3.66e-6^{***}$ ($5.60e{-7}$)	-0.369^{***} ($5.60e{-7}$)	0.369^{***} ($5.60e{-7}$)	$-2.35e-6^{***}$ ($2.33e{-7}$)	-0.369^{***} (0.0385)	-0.3649^{***} (0.0403)	$-1.55e-6^{***}$ ($1.40e{-7}$)
College grade	$-5.37e-6^{***}$ ($6.04e{-7}$)	$1.69e-6^{***}$ ($6.04e{-7}$)	-0.2076^{***} ($6.04e{-7}$)	0.2076^{***} ($6.04e{-7}$)	$-3.68e-6^{***}$ ($4.10e{-7}$)	-0.2076^{***} (0.0295)	-0.1948^{***} (0.0287)	$-3.37e-6^{***}$ ($3.72e{-7}$)
High school grade	$-3.91e-6^{***}$ ($7.40e{-7}$)	$7.38e-7$ ($7.40e{-7}$)	-0.1486^{***} ($7.40e{-7}$)	0.1486^{***} ($7.40e{-7}$)	$-3.17e-6^{***}$ ($6.14e{-7}$)	-0.1486^{***} (0.0345)	-0.1419^{***} (0.0337)	$-3.04e-6^{***}$ ($5.88e{-7}$)
Undergraduate studies	$-1.95e-6^{*}$ ($1.09e{-6}$)	$1.88e-7$ ($1.09e{-6}$)	-0.0596^{***} ($1.09e{-6}$)	0.0596^{***} ($1.09e{-6}$)	$-1.76e-6^{*}$ ($9.83e{-7}$)	-0.0596 (0.0374)	-0.0563 (0.0357)	$-1.73e-6^{*}$ ($9.66e{-7}$)
Medical density	$2.46e-7$ ($3.83e{-7}$)	$1.02e-7$ ($3.83e{-7}$)	$-2.46e-7$ ($3.83e{-7}$)	$-1.02e-7$ ($3.83e{-7}$)	$3.48e-7$ ($5.42e{-7}$)	–	-0.0001^{***} (-0.00001)	$3.69e-7$ ($5.74e{-7}$)
Unemployment rate	$3.45e-7^{***}$ ($1.38e{-8}$)	$-3.45e-7^{***}$ ($1.38e{-8}$)	0.0264^{***} ($1.38e{-8}$)	-0.0264^{***} ($1.38e{-8}$)	–	0.0264^{***} (0.0012)	0.0249^{***} (0.001)	$-6.71e-8$ (0.0023)

The estimated standard errors are within parenthesis; ***: Significant at the 1% level; **: Significant at the 5% level; *: Significant at the 10% level.

+: Refers to the individual's nationality; h : *Ill with risk of death*; w : *work*

Table 2.13: Estimates marginal effects on joint, marginal and conditional probabilities.

Part B: initial conditions.

Variables	$p(h=1,w=1)$	$p(h=0,w=1)$	$p(h=1,w=0)$	$p(h=0,w=0)$	$p(h=1)$	$p(w=1)$	$p(h=1\|w=1)$	$p(w=1\|h=1)$
Gender (male)	$-1.48e-10^{*}$ (8.57e−11)	$-3.40e-11$ (8.57e−11)	0.0417^{***} (8.57e−11)	-0.0417^{***} (8.57e−11)	$-1.81e-10^{**}$ (9.22e−11)	0.0417 (0.0306)	0.0367 (0.0258)	$-1.91e-10^{**}$ (9.47e−11)
Age	$3.52e-11$ (0.0529)	$5.63e-12$ (0.0529)	0.0081 (0.0529)	-0.0081 (0.0529)	$4.08e-11$ (1.92e−10)	0.0081^{**} (0.0037)	0.0009 (0.0305)	$4.06e-11$ (1.85e−10)
Not French^{+}	$-3.62e-10$ (0.5886)	$-5.47e-11$ (0.5886)	-0.1365 (0.5886)	0.1365 (0.5886)	$-4.17e-10$ (0.5886)	-0.1365^{*} (0.0729)	-0.1152 (0.0766)	$-4.33e-10$ (0.0436)
Couple	$1.91e-10$ (3.05e−10)	$1.80e-11$ (3.05e−10)	0.0396^{***} (3.05e−10)	-0.0396^{***} (3.05e−10)	$2.09e-10$ (3.30e−10)	0.0396 (0.0406)	0.0335 (0.0341)	$2.15e-10$ (3.38e−10)
Number of children	0 (0.5308)	$-1.24e-9$ (0.5308)	-0.1398 (0.5308)	0.1398 (0.5308)	$-1.24e-9$ (−1.09e−6)	-0.1398^{*} (0.0837)	4.9516^{***} (0.2965)	$3.81e-11^{*}$ (2.00e−11)
No grade	$1.18e-9$ (2.13e−9)	$5.65e-10$ (2.13e−9)	-0.1883^{***} (2.13e−9)	0.1883^{***} (2.13e−9)	$1.75e-9$ (2.71e−9)	-0.1883 (0.1352)	-0.1725 (0.1255)	$1.89e-9$ (2.86e−9)
College grade	$1.56e-10$ (4.64e−10)	$3.30e-11$ (4.64e−10)	-0.0284^{***} (4.64e−10)	0.0284^{***} (4.64e−10)	$1.89e-10$ (5.30e−10)	-0.0284 (0.0562)	-0.0253 (0.0473)	$1.98e-10$ (5.50e−10)
High school grade	$-1.59e-10$ (9.67e−11)	$-1.96e-11$ (9.67e−11)	-0.0614^{***} (9.67e−11)	0.0614^{***} (9.67e−11)	$-1.78e-10$ (1.12e−10)	-0.0614 (0.0738)	-0.0527 (0.0644)	$-1.84e-10$ (1.16e−10)
Undergraduate studies	$-1.55e-10$ (1.84e−10)	$-2.57e-11$ (1.84e−10)	-0.0016^{***} (1.84e−10)	0.0016^{***} (1.84e−10)	$-1.80e-10$ (2.15e−10)	-0.0016 (0.0030)	-0.0004 (−0.0018)	$-1.88e-10$ (2.24e−10)
Medical density	0 (0.0028)	$1.52e-11$ (0.0028)	$-4.e-13$ (0.0028)	$-1.52e-11$ (0.0028)	$1.52e-11$ (0.0028)	$-$	0 (0.0028)	0 (0.0028)
Unemployment rate	$-4.30e-12$ (0.0048)	$4.30e-12$ (0.0048)	0.0005 (0.0048)	-0.0005 (0.0048)	$-$	0.0005 (0.0048)	-0.0200^{**} (0.0048)	$-5.36e-12$ (0.0048)
Illness before prof. life	$1.44e-10^{*}$ (7.55e−11)	$2.89e-11$ (7.55e−11)	-0.0003^{***} (7.55e−11)	0.0003^{***} (7.55e−11)	$1.73e-10$ (1.67e−2)	-0.0003 (0.0106)	-0.0195^{*} (0.0106)	$1.76e-10^{***}$ (3.65e−14)

The estimated standard errors are within parenthesis; ***: Significant at the 1% level; **: Significant at the 5% level; *: Significant at the 10% level.

+: Refers to the individual's nationality; h : *Ill with risk of death*; w : *work*

Table 2.14: Estimates of health and job status interactions with cross terms.

Part A: dynamic equations

	Illness		Disability index		Risk of death	
Variables	h : *health*	w : *work*	h : *disab*	w : *work*	h : *rdeath*	w : *work*
h_{-1}	3.8216^{***} (0.0226)	-0.2085^{***} (0.0139)	4.0523^{***} (0.0499)	-0.4731^{***} (0.0328)	3.7859^{***} (0.0502)	-0.5299^{***} (0.0365)
w_{-1}	0.2225^{***} (0.0238)	2.7262^{***} (0.0127)	0.0275 (0.0562)	2.723^{***} (0.0128)	0.0019 (0.0577)	2.7215^{***} (0.0128)
Age	0.0374^{***} (0.001)	-0.0247^{***} (0.0007)	0.0072^{***} (0.0023)	-0.0269^{***} (0.0007)	0.013^{***} (0.0023)	-0.0269^{***} (0.0007)
Not French^{+}	-0.0493^{**} (0.025)	-0.3426^{***} (0.0163)	-0.1607^{***} (0.0597)	-0.3394^{***} (0.0166)	-0.1259^{**} (0.0622)	-0.3375^{***} (0.0166)
Gender (male)	-0.0951^{**} (0.0395)	0.2837^{***} (0.0289)	0.0665 (0.0909)	0.3039^{***} (0.0297)	0.0311 (0.0933)	0.2957^{***} (0.0297)
Couple	-0.0365 (0.0235)	-0.3873^{***} (0.0168)	-0.0692 (0.0573)	-0.3857^{***} (0.0169)	-0.1329^{**} (0.0555)	-0.3917^{***} (0.017)
Male $*$ *Couple*	0.0358 (0.0371)	0.7932^{***} (0.0302)	0.1042 (0.0843)	0.7945^{***} (0.0305)	-0.0238 (0.0869)	0.7998^{***} (0.0305)
Number of children	0.0278^{***} (0.0088)	-0.0795^{***} (0.006)	0.0318 (0.0209)	-0.0779^{***} (0.006)	0.0186 (0.0209)	-0.0776^{***} (0.006)
Male $*$ *Number of child.*	-0.0031 (0.0129)	-0.0028 (0.0105)	-0.0511^{*} (0.0294)	-0.012 (0.0107)	-0.0242 (0.0315)	-0.013 (0.0107)
No grade	0.2986^{***} (0.0446)	-0.9276^{***} (0.0303)	0.0413 (0.1069)	-0.9484^{***} (0.0308)	-0.0599 (0.1038)	-0.9503^{***} (0.0308)
College grade	0.2464^{***} (0.0321)	-0.5452^{***} (0.0237)	0.0104 (0.0773)	-0.5744^{***} (0.0243)	-0.0591 (0.0737)	-0.5743^{***} (0.0243)
High school grade	0.2058^{***} (0.0337)	-0.3707^{***} (0.0261)	0.0729 (0.0781)	-0.3875^{***} (0.0268)	-0.0569 (0.0779)	-0.3857^{***} (0.0268)
Undergraduate studies	0.0653^{*} (0.0376)	-0.1301^{***} (0.0305)	0.0255 (0.0863)	-0.1401^{***} (0.0317)	-0.0286 (0.0863)	-0.1414^{***} (0.0317)
Ref : Graduate studies	-	-	-	-	-	-
Male $*$ *School grade*	-0.051^{***} (0.015)	-0.0063 (0.013)	-0.0104 (0.0332)	0.0132 (0.0141)	0.0001 (0.0337)	0.0151 (0.0142)
Medical density	-0.0009^{***} (0.0003)	–	0.0057^{***} (0.0008)	–	0.0057^{***} (0.0009)	–
Unemployment rate	–	0.0732^{***} (0.0023)	–	0.0751^{***} (0.0023)	–	0.0752^{***} (0.0023)
Intercept	-2.5989^{***} (0.0604)	-1.1899^{***} (0.0348)	-5.5109^{***} (0.1512)	-1.1263^{***} (0.0353)	-5.558^{***} (0.1635)	-1.122^{***} (0.0353)
Covariance matrix	$\sigma_1 = 1.3679^{***}$ (0.0184), $\sigma_2 = 1.7238^{***}$ (0.0162)		$\sigma_1 = 1.0634^{***}$ (0.0122), $\sigma_2 = 1.6888^{***}$ (0.0164)		$\sigma_1 = 1.0132^{***}$ (0.012), $\sigma_2 = 1.6884^{***}$ (0.0164)	
	$\rho_\eta = -0.8239^{***}$ (0.0054), $\rho_\zeta = 0.0174$ (0.0249)		$\rho_\eta = 0.2839^{***}$ (0.032), $\rho_\zeta = 0.0468$ (0.0483)		$\rho_\eta = 0.2451^{***}$ (0.0325), $\rho_\zeta = 0.025$ (0.0502)	

The estimated standard errors are within parenthesis; ***: Significant at the 1% level

**: Significant at the 5% level; *: Significant at the 10% level; +: Refers to the individual's nationality.

Table 2.15: Estimates of health and job status interactions with cross terms.

Part B: the initial conditions

	Illness		Disability index		Risk of death	
Variables	h : *health*	w : *work*	h : *disab*	w : *work*	h : *rdeath*	w : *work*
Age	-0.0055 (0.0158)	0.031^{***} (0.0082)	0.0571 (0.0418)	0.0304^{***} (0.0082)	0.024 (0.0531)	0.0303^{***} (0.0083)
Not French^{+}	-0.2273^{**} (0.101)	-0.4549^{***} (0.0437)	-0.669 (0.4378)	-0.4539^{***} (0.0438)	-1.0945^{*} (0.6058)	-0.4563^{***} (0.0438)
Gender (male)	-0.2567^{**} (0.1049)	0.5179^{***} (0.0542)	0.0758 (0.273)	0.5157^{***} (0.0543)	-0.3258 (0.3297)	0.5184^{***} (0.0544)
Couple	-0.0254 (0.0958)	-0.0192 (0.0559)	-0.238 (0.3047)	-0.0164 (0.056)	0.0507 (0.3176)	-0.018 (0.0561)
Male $*$ *Couple*	0.1859 (0.1694)	0.538^{***} (0.1054)	-0.0689 (0.492)	0.5372^{***} (0.1054)	0.3158 (0.5275)	0.5383^{***} (0.1056)
Number of children	-0.037 (0.1642)	-0.6761^{***} (0.0748)	-0.2841 (0.5259)	-0.6713^{***} (0.0749)	-0.297 (0.535)	-0.6699^{***} (0.075)
Male $*$ *Number of children*	0.0411 (0.2891)	0.4718^{***} (0.1452)	-5.6504 (2888.779)	0.4661^{***} (0.1453)	-6.3479 (4291.568)	0.4633^{***} (0.1455)
No grade	0.1805 (0.1831)	-1.0426^{***} (0.0994)	0.9295^{*} (0.5166)	-1.0407^{***} (0.0995)	0.5275 (0.5598)	-1.0413^{***} (0.0997)
College grade	0.0629 (0.1306)	-0.4353^{***} (0.0758)	0.7047^{*} (0.388)	-0.4371^{***} (0.0759)	0.2636 (0.4272)	-0.4393^{***} (0.076)
High school grade	-0.0538 (0.1206)	-0.4353^{***} (0.0677)	0.2907 (0.3458)	-0.436^{***} (0.0678)	-0.0964 (0.3931)	-0.4371^{***} (0.0679)
Undergraduate studies	-0.081 (0.1176)	-0.0996 (0.0681)	0.3926 (0.3267)	-0.0998 (0.0682)	-0.1939 (0.3975)	-0.1004 (0.0683)
Ref : Graduate studies	-	-	-	-	-	-
Male $*$ *Schoolgrade*	-0.0108 (0.0514)	-0.2778^{***} (0.0273)	0.0293 (0.1394)	-0.2733^{***} (0.0273)	0.0982 (0.1617)	-0.2741^{***} (0.0274)
Medical density	0.0007 (0.0008)	–	0.0084^{***} (0.0021)	–	0.0119^{***} (0.0029)	–
Unemployment rate	–	0.0017 (0.0048)	–	0.0039 (0.0048)	–	0.0033 (0.0049)
Illness before prof. life	0.3626^{***} (0.0122)	-0.0025 (0.0048)	0.1412^{***} (0.0137)	-0.0017 (0.0045)	0.1376^{***} (0.0168)	-0.0015 (0.0045)
Intercept	-1.582^{***} (0.3594)	0.7729^{***} (0.1906)	-8.3301^{***} (1.1098)	0.741^{***} (0.1904)	-8.5935^{***} (1.4329)	0.7481^{***} (0.1908)
λ_{11}	1.2064^{***} (0.0634)	-	1.0634^{***} (0.1365)	-	1.9091^{***} (0.185)	-
λ_{12}	0.4028^{***} (0.0555)	-	0.0021 (0.1089)	-	0.1543 (0.1285)	-
λ_{21}	-	0.0307 (0.0296)	-	0.0159 (0.0324)	-	-0.0558 (0.0339)
λ_{22}	-	0.10^{***} (0.026)	-	0.1001^{***} (0.0214)	-	0.1129^{***} (0.0213)
Covariance matrix	$\rho_\epsilon = 0.0192$ (0.0461)		$\rho_\epsilon = -0.1614$ (0.1148)		$\rho_\epsilon = 0.0049$ (0.1496)	

The estimated standard errors are within parenthesis; ***: Significant at the 1% level

**: Significant at the 5% level; *: Significant at the 10% level; +: Refers to the individual's nationality.

Table 2.16: Estimates of health and job status interactions with cross terms.

Part A: dynamic equations

	Bivariate estimations		Univariate estimations	
Variables	h : *health*	w : *work*	h : *health*	w : *work*
	(1)	(2)	(1')	(2')
h_{-1}	3.8216^{***} (0.0226)	-0.2085^{***} (0.0139)	4.2512^{***} (0,0154)	$-0,2793^{***}$ (0,0153)
w_{-1}	0.2225^{***} (0.0238)	2.7262^{***} (0.0127)	$0,0179$ (0,0194)	$2,7522^{***}$ (0,0137)
Age	0.0374^{***} (0.001)	-0.0247^{***} (0.0007)	$0,0114^{***}$ (0,0008)	$-0,0104^{***}$ (0,0007)
$Not\ French^{+}$	-0.0493^{**} (0.025)	-0.3426^{***} (0.0163)	$-0,0216$ (0,0191)	$-0,2477^{***}$ (0,0214)
Gender(*Male*)	-0.0951^{**} (0.0395)	0.2837^{***} (0.0289)	$-0,0251$ (0,0306)	$0,2039^{***}$ (0,0327)
Couple	-0.0365 (0.0235)	-0.3873^{***} (0.0168)	$-0,0488^{**}$ (0,0195)	$-0,3197^{***}$ (0,0177)
Male $*$ *Couple*	0.0358 (0.0371)	0.7932^{***} (0.0302)	$0,0257$ (0,0305)	$0,664^{***}$ (0,0314)
Number of children	0.0278^{***} (0.0088)	-0.0795^{***} (0.006)	$0,0212^{***}$ (0,0071)	$-0,0502^{***}$ (0,0067)
Male $*$ *Number of children*	-0.0031 (0.0129)	-0.0028 (0.0105)	$-0,0107$ (0,0104)	$-0,0139$ (0,0113)
No grade	0.2986^{***} (0.0446)	-0.9276^{***} (0.0303)	$0,1014^{***}$ (0,0335)	$-0,6564^{***}$ (0,0379)
College grade	0.2464^{***} (0.0321)	-0.5452^{***} (0.0237)	$0,0742^{***}$ (0,0237)	$-0,3277^{***}$ (0,0277)
High school grade	0.2058^{***} (0.0337)	-0.3707^{***} (0.0261)	$0,0702^{***}$ (0,0248)	$-0,2148^{***}$ (0,0298)
Undergraduate studies	0.0653^{*} (0.0376)	-0.1301^{***} (0.0305)	$0,0227$ (0,0273)	$-0,1045^{***}$ (0,0333)
Ref : *Graduate studies*	-	-	-	-
Male $*$ *Schoolgrade*	-0.051^{***} (0.015)	-0.0063 (0.013)	$-0,0134$ (0,0109)	$-0,0543^{***}$ (0,0139)
Medical density	-0.0009^{***} (0.0003)	–	$0,0018^{***}$ (0,0003)	–
Unemployment rate	–	0.0732^{***} (0.0023)	–	$0,0264^{***}$ (0,0024)
Intercept	-2.5989^{***} (0.0604)	-1.1899^{***} (0.0348)	$-2,8737^{***}$ (0,0496)	$-0,0917^{***}$ (0,0395)
Covariance matrix	$\sigma_1 = 1.3678^{***}$ (0.0184), $\sigma_2 = 1.7238^{***}$ (0.0162)		-	
	$\rho_\eta = -0.8239^{***}$ (0.0054), $\rho_\zeta = 0.0249$ (0.0174)		-	

The estimated standard errors are within parenthesis.

***: Significant at the 1% level; **: Significant at the 5% level.

*: Significant at the 10% level. +: Refers to the individual's nationality.

Table 2.17: Estimates of health and job status interactions.

Part B: the initial conditions.

	Bivariate estimations		Univariate estimations	
Variables	h : *health*	w : *work*	h : *health*	w : *work*
	Initial conditions			
Age	-0.0055 (0.0158)	0.031^{***} (0.0082)	$-0,0134$ (0,012)	$0,0332^{***}$ (0,0082)
$Not\,French^{+}$	-0.2273^{**} (0.101)	-0.4549^{***} (0.0437)	$-0,2068^{***}$ (0,077)	$-0,4536^{***}$ (0,0435)
$Gender(Male)$	-0.2567^{**} (0.1049)	0.5179^{***} (0.0542)	$-0,248^{***}$ (0,0796)	$0,5188^{***}$ (0,054)
Couple	-0.0254 (0.0958)	-0.0192 (0.0559)	$-0,0088$ (0,0733)	$-0,0221$ (0,0557)
$Male * Couple$	0.1859 (0.1694)	0.538^{***} (0.1054)	$0,1636$ (0,1264)	$0,5365^{***}$ (0,1053)
$Number\,of\,children$	0.0411 (0.2891)	-0.6761^{***} (0.0748)	$0,0235$ (0,1218)	$-0,6883^{***}$ (0,0745)
$Male * Number\,of\,children$	-0.037 (0.1642)	0.4718^{***} (0.1452)	$-0,0514$ (0,2064)	$0,4892^{***}$ (0,1448)
$No\,grade$	0.1805 (0.1831)	-1.0426^{***} (0.0994)	$0,1031$ (0,1403)	$-1,0393^{***}$ (0,099)
$College\,grade$	0.0629 (0.1306)	-0.4353^{***} (0.0758)	$0,0443$ (0,0992)	$-0,4263^{***}$ (0,0756)
$High\,school\,grade$	-0.0538 (0.1206)	-0.4353^{***} (0.0677)	$-0,057$ (0,0912)	$-0,4265^{***}$ (0,0675)
$Undergraduate\,studies$	-0.081 (0.1176)	-0.0996 (0.0681)	$-0,0357$ (0,0876)	$-0,0975$ (0,068)
Ref : $Graduate\,studies$	-	-	-	-
$Male * Schoolgrade$	-0.0108 (0.0514)	-0.2778^{***} (0.0273)	$0,0254$ (0,0384)	$-0,2832^{***}$ (0,0272)
$Medical\,density$	0.0007 (0.0008)	–	$0,0027^{***}$ (0,0006)	–
$Unemployment\,rate$	–	0.0017 (0.0048)	–	$-0,0033$ (0,0046)
$Illness\,before\,prof.\,life$	0.3626^{***} (0.0122)	-0.0025 (0.0048)	$0,3465^{***}$ (0,009)	$-0,0032$ (0,0045)
Intercept	-1.582^{***} (0.3594)	0.7729^{***} (0.1906)	$-1,809^{***}$ (0,2644)	$0,8691^{***}$ (0,1881)
λ_{11}	1.2064^{***} (0.0634)	-		
λ_{12}	0.4028^{***} (0.0555)	-		
λ_{21}	-	0.0307 (0.0296)		
λ_{22}	-	0.1^{***} (0.026)		
Covariance matrix	$\rho_\epsilon = 0.0192$ (0.0461)		-	

The estimated standard errors are within parenthesis.

***: Significant at the 1% level, **: Significant at the 5% level.

*: Significant at the 10% level, +: Refers to the individual's nationality.

Table 2.18: Estimates of health and job status interactions with cross terms.

Part A: dynamic equations

	Disability index		Risk of death	
Variables	h : *disab*	w : *work*	h : *rdeath*	w : *work*
h_{-1}	4.0523^{***} (0.0499)	-0.4731^{***} (0.0328)	3.7859^{***} (0.0502)	-0.5299^{***} (0.0365)
w_{-1}	0.0275 (0.0562)	2.723^{***} (0.0128)	0.0019 (0.0577)	2.7215^{***} (0.0128)
Age	0.0072^{***} (0.0023)	-0.0269^{***} (0.0007)	0.013^{***} (0.0023)	-0.0269^{***} (0.0007)
Not French^{+}	-0.1607^{***} (0.0597)	-0.3394^{***} (0.0166)	-0.1259^{**} (0.0622)	-0.3375^{***} (0.0166)
Gender (*male*)	0.0665 (0.0909)	0.3039^{***} (0.0297)	0.0311 (0.0933)	0.2957^{***} (0.0297)
Couple	-0.0692 (0.0573)	-0.3857^{***} (0.0169)	-0.1329^{**} (0.0555)	-0.3917^{***} (0.017)
Male $*$ *Couple*	0.1042 (0.0843)	0.7945^{***} (0.0305)	-0.0238 (0.0869)	0.7998^{***} (0.0305)
Number of children	0.0318 (0.0209)	-0.0779^{***} (0.006)	0.0186 (0.0209)	-0.0776^{***} (0.006)
Male $*$ *Number of child.*	-0.0511^{*} (0.0294)	-0.012 (0.0107)	-0.0242 (0.0315)	-0.013 (0.0107)
No grade	0.0413 (0.1069)	-0.9484^{***} (0.0308)	-0.0599 (0.1038)	-0.9503^{***} (0.0308)
College grade	0.0104 (0.0773)	-0.5744^{***} (0.0243)	-0.0591 (0.0737)	-0.5743^{***} (0.0243)
High school grade	0.0729 (0.0781)	-0.3875^{***} (0.0268)	-0.0569 (0.0779)	-0.3857^{***} (0.0268)
Undergraduate studies	0.0255 (0.0863)	-0.1401^{***} (0.0317)	-0.0286 (0.0863)	-0.1414^{***} (0.0317)
Ref : *Graduate studies*	-	-	-	-
Male $*$ *School grade*	-0.0104 (0.0332)	0.0132 (0.0141)	0.0001 (0.0337)	0.0151 (0.0142)
Medical density	0.0057^{***} (0.0008)	–	0.0057^{***} (0.0009)	–
Unemployment rate	–	0.0751^{***} (0.0023)	–	0.0752^{***} (0.0023)
Intercept	-5.5109^{***} (0.1512)	-1.1263^{***} (0.0353)	-5.558^{***} (0.1635)	-1.122^{***} (0.0353)
Covariance matrix	$\sigma_1 = 1.0634^{***}$ (0.0122), $\sigma_2 = 1.6888^{***}$ (0.0164)		$\sigma_1 = 1.0132^{***}$ (0.012), $\sigma_2 = 1.6884^{***}$ (0.0164)	
	$\rho_\eta = 0.2839^{***}$ (0.032), $\rho_\zeta = 0.0468$ (0.0483)		$\rho_\eta = 0.2451^{***}$ (0.0325), $\rho_\zeta = 0.025$ (0.0502)	

The estimated standard errors are within parenthesis; ***: Significant at the 1% level

**: Significant at the 5% level; *: Significant at the 10% level.

+: Refers to the individual's nationality.

Table 2.19: Estimates of health and job status interactions with cross terms.

Part B: the initial conditions

	Disability index		Risk of death	
Variables	$h : disab$	$w : work$	$h : rdeath$	$w : work$
Age	0.0571 (0.0418)	0.0304*** (0.0082)	0.024 (0.0531)	0.0303*** (0.0083)
$Not\ French^{+}$	−0.669 (0.4378)	−0.4539*** (0.0438)	−1.0945* (0.6058)	−0.4563*** (0.0438)
Gender (male)	0.0758 (0.273)	0.5157*** (0.0543)	−0.3258 (0.3297)	0.5184*** (0.0544)
Couple	−0.238 (0.3047)	−0.0164 (0.056)	0.0507 (0.3176)	−0.018 (0.0561)
$Male * Couple$	−0.0689 (0.492)	0.5372*** (0.1054)	0.3158 (0.5275)	0.5383*** (0.1056)
Number of children	−0.2841 (0.5259)	−0.6713*** (0.0749)	−0.297 (0.535)	−0.6699*** (0.075)
$Male * Number\ of\ children$	−5.6504 (2888.779)	0.4661*** (0.1453)	−6.3479 (4291.568)	0.4633*** (0.1455)
No grade	0.9295* (0.5166)	−1.0407*** (0.0995)	0.5275 (0.5598)	−1.0413*** (0.0997)
College grade	0.7047* (0.388)	−0.4371*** (0.0759)	0.2636 (0.4272)	−0.4393*** (0.076)
High school grade	0.2907 (0.3458)	−0.436*** (0.0678)	−0.0964 (0.3931)	−0.4371*** (0.0679)
Undergraduate studies	0.3926 (0.3267)	−0.0998 (0.0682)	−0.1939 (0.3975)	−0.1004 (0.0683)
Ref : Graduate studies	-	-	-	-
$Male * Schoolgrade$	0.0293 (0.1394)	−0.2733*** (0.0273)	0.0982 (0.1617)	−0.2741*** (0.0274)
Medical density	0.0084*** (0.0021)	–	0.0119*** (0.0029)	–
Unemployment rate	–	0.0039 (0.0048)	–	0.0033 (0.0049)
$Illness\ before\ prof.\ life$	0.1412*** (0.0137)	−0.0017 (0.0045)	0.1376*** (0.0168)	−0.0015 (0.0045)
Intercept	−8.3301*** (1.1098)	0.741*** (0.1904)	−8.5935*** (1.4329)	0.7481*** (0.1908)
λ_{11}	1.0634*** (0.1365)	-	1.9091*** (0.185)	-
λ_{12}	0.0021 (0.1089)	-	0.1543 (0.1285)	-
λ_{21}	-	0.0159 (0.0324)	-	−0.0558 (0.0339)
λ_{22}	-	0.1001*** (0.0214)	-	0.1129*** (0.0213)
Covariance matrix	$\rho_\epsilon = -0.1614$ (0.1148)		$\rho_\epsilon = 0.0049$ (0.1496)	

The estimated standard errors are within parenthesis; ***: Significant at the 1% level

**: Significant at the 5% level; *: Significant at the 10% level.

+: Refers to the individual's nationality.

Table 2.20: Marginal effects on probabilities of positive outcomes.

Part A: dynamic equations

	Illness		Disability index		Risk of death	
Variables	$h : health$	$w : work$	$h : disab$	$w : work$	$h : rdeath$	$w : work$
h_{-1}	0.8697^{***} (0.0004)	-0.0775^{***} (0.0007)	0.4279^{***} (0.0291)	-0.1818^{***} (0.0034)	0.3148^{***} (0.027)	-0.2045^{***} (0.0042)
w_{-1}	0.0815^{***} (0.0011)	0.7625^{***} (0.0097)	$2.12e-6^{***}$ ($5.84e{-8}$)	0.7694^{***} (0.0192)	$1.15e-7^{***}$ ($4.52e{-9}$)	0.7697^{***} (0.0197)
Age	0.0141^{***} (0.0006)	-0.009^{***} (0.0003)	$5.78e-7^{***}$ ($1.14e{-7}$)	-0.0096^{***} (0.0011)	$7.69e-7^{***}$ ($1.77e{-7}$)	-0.0096^{***} (0.001)
$NotFrench^{+}$	-0.0185^{***} (0.0019)	-0.13^{***} (0.0143)	-0.00001^{***} ($1.09e{-6}$)	-0.1274^{***} (0.0204)	$-6.19e-6^{***}$ ($7.92e{-7}$)	-0.1266^{***} (0.0226)
$Gender(male)$	-0.0591^{***} (0.0062)	0.2951^{***} (0.0138)	$4.33e-6^{**}$ ($1.98e{-6}$)	0.3035^{***} (0.0251)	$-1.19e-6^{*}$ ($6.81e{-7}$)	0.3022^{***} (0.0283)
Couple	-0.0077^{***} (0.0013)	-0.0099 (0.0135)	$-1.8e-6^{**}$ ($8.78e{-7}$)	-0.0089 (0.0226)	$-9.92e-6^{***}$ ($9.21e{-7}$)	-0.0102 (0.0195)
Number of children	0.01^{***} (0.0015)	-0.0294^{***} (0.0041)	$6.90e-7$ ($6.79e{-7}$)	-0.0299^{***} (0.0071)	$4.53e-7$ ($3.53e{-7}$)	-0.0299^{***} (0.0081)
No grade	0.1163^{***} (0.0073)	-0.3569^{***} (0.0207)	$3.58e-6^{***}$ ($5.54e{-7}$)	-0.3638^{***} (0.0484)	$-3.21e-6^{***}$ ($3.71e{-7}$)	-0.3645^{***} (0.0466)
College grade	0.0926^{***} (0.0068)	-0.1942^{***} (0.0151)	$8.36e-7^{***}$ ($1.63e{-7}$)	-0.2012^{***} (0.0389)	$-3.56e-6^{***}$ ($4.78e{-7}$)	-0.2009^{***} (0.0348)
High school grade	0.0793^{***} (0.0085)	-0.1405^{***} (0.0152)	$6.53e-6^{***}$ ($1.75e{-6}$)	-0.1456^{***} (0.0348)	$-3.12e-6^{***}$ ($6.44e{-7}$)	-0.1448^{***} (0.0361)
Undergraduate studies	0.0249^{***} (0.0089)	-0.0483^{***} (0.0171)	$2.15e-6$ ($1.67e{-6}$)	-0.0514 (0.0374)	$-1.62e-6$ ($1.03e{-6}$)	-0.0519 (0.0372)
Medical density	-0.0003 (0.0002)	–	$4.6e-7$ ($1.25e{-6}$)	–	$3.4e-7$ ($9.63e{-7}$)	–
Unemployment rate	–	0.0266^{***} (0.0012)	–	0.0269^{***} (0.0023)	–	0.0269 (0.0353)

The estimated standard errors are within parenthesis; ***: Significant at the 1% level

**: Significant at the 5% level; *: Significant at the 10% level; +: Refers to the individual's nationality.

Table 2.21: Marginal effects on probabilities of positive outcomes

Part B: the initial conditions

	Illness		Disability index		Risk of death	
Variables	h : *health*	w : *work*	h : *disab*	w : *work*	h : *rdeath*	w : *work*
Age	-0.0009^{**} (0.0004)	0.0075^{***} (0.0007)	$3.42e-9$ (0.0291)	0.0075^{**} (0.0034)	$8.45e-12$ (0.027)	0.0075^{*} (0.0042)
$NotFrench^{+}$	-0.0317^{***} (0.0011)	-0.1288^{***} (0.0097)	$-1.58e-8$ ($5.84e{-8}$)	-0.1312^{***} (0.0192)	$-8.4e-11$ ($4.52e{-9}$)	-0.132^{***} (0.0197)
$Gender(male)$	-0.0372^{***} (0.0006)	0.0328^{***} (0.0003)	$-8.11e-9$ ($1.14e{-7}$)	0.035^{***} (0.0011)	$-1.03e-10$ ($1.77e{-7}$)	0.0353^{***} (0.001)
Couple	0.0096^{***} (0.0019)	0.0509^{***} (0.0143)	$-1.11e-8$ ($1.09e{-6}$)	0.0529^{***} (0.0204)	$7.53e-11$ ($7.92e{-7}$)	0.0526^{**} (0.0226)
Number of children	-0.0029 (0.0062)	-0.1114^{***} (0.0138)	$-1.73e-7$ ($1.98e{-6}$)	-0.1139^{***} (0.0251)	$-6.21e-10$ ($6.81e{-7}$)	-0.1138^{***} (0.0283)
No grade	0.0315^{***} (0.0013)	-0.3459^{***} (0.0135)	$1.05e-6$ ($8.78e{-7}$)	-0.3498^{***} (0.0226)	$8.29e-10$ ($9.21e{-7}$)	-0.3499^{***} (0.0195)
College grade	0.0099^{***} (0.0015)	-0.1062^{***} (0.0041)	$7.95e-8$ ($6.79e{-7}$)	-0.1094^{***} (0.0071)	$7.55e-11$ ($3.53e{-7}$)	-0.1099^{***} (0.0081)
High school grade	-0.0083 (0.0073)	-0.12^{***} (0.0207)	$3.28e-8$ ($5.54e{-7}$)	-0.1229^{**} (0.0484)	$-1.96e-11$ ($3.71e{-7}$)	-0.1232^{***} (0.0466)
Undergraduate studies	-0.0122^{*} (0.0068)	-0.025^{*} (0.0151)	$6.47e-8$ ($1.63e{-7}$)	-0.0257 (0.0389)	$-3.13e-11$ ($4.78e{-7}$)	-0.0258 (0.0348)
Medical density	0.0001 (0.0085)	–	$5.01e-10$ ($1.75e{-6}$)	–	$3.6e-12$ ($6.44e{-7}$)	–
Unemployment rate	–	0.0004 (0.0171)	–	0.001 (0.0374)	–	0.0008 (0.0372)
Illness before prof. life	0.0571^{***} (0.0002)	-0.0006^{***} (0.0002)	$8.49e-9$ ($1.25e{-6}$)	-0.0004 (0.0045)	$1.01e-7$ ($9.63e{-7}$)	-0.0004 (0.0045)

The estimated standard errors are within parenthesis; ***: Significant at the 1% level

**: Significant at the 5% level; *: Significant at the 10% level; +: Refers to the individual's nationality.

Chapter 3

Dynamic Interactions between Health and Employment Statuses : A Non-parametric Analysis[1]

[1]Authors : **Richard Moussa**, ThEMA-UCP and ENSEA Abidjan; **Eric Delattre**, ThEMA-UCP

Abstract

Despite numerous sociological results, there are 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. This 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

Introduction

Relationship between health condition and job status has been analyzed according to several approaches in both the economic and sociology 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 (see 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 is to estimate a bivariate model. Barnay (2015) reviews key papers on bidirectional causal links between health condition and job status on European data. The key methodological approaches are the following one. 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 sexes. 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 condition 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

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

[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.

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 fully been 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 a grouping 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 grouping 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 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.

3.1 Econometrics Model

3.1.1 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 then if not. 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 works. 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 variable. 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 :

$$Y_t = \alpha_1 Z_t + \delta_{11} Y_{t-1} + \delta_{12} Z_{t-1} + \beta_1 X_t + \epsilon_t^1 \tag{3.1}$$

$$Z_t = \alpha_2 Y_t + \delta_{21} Y_{t-1} + \delta_{22} Z_{t-1} + \beta_2 X_t + \epsilon_t^2 \tag{3.2}$$

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

$$Y_{it} = \alpha_1 Z_{i,t} + \delta_{11} Y_{i,t-1} + \delta_{12} Z_{i,t-1} + \beta_1 X_{i,t} + \eta_i^1 + \zeta_{it}^1 \tag{3.3}$$

$$Z_{it} = \alpha_2 Y_{i,t} + \delta_{21} Y_{i,t-1} + \delta_{22} Z_{i,t-1} + \beta_2 X_{i,t} + \eta_i^2 + \zeta_{it}^2 \tag{3.4}$$

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 δ_{12} and δ_{21} if we wish to test for lagged non-causality, and on α_1 and α_2 if we wish to test for instantaneous non-causality. Note that when one does not account for instantaneous non-causality, α_1 and α_2 are null in the two models above (Equations 1 to 4). Generalization of this framework to more than two variables and up to one lag order can be found in Michaud and Van Soest (2008). In such a case, authors use generalized moment method to estimate the parameters.

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 δ_{12} and δ_{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). This specification has the advantage to give a better assessment of the distribution of the causal effect among individuals but as we can see, the number of causal coefficients to be estimated grows up from 4 to 4N, it induces a lost of degrees of freedom; that can be worse for short panels.

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) approach is 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.

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

In all specifications above, when dependent variables are qualitative outcomes (inducing that error term ζ can not be treated as normal), it is common to use latent variables and probit probabilities[5]. Another way to deal with qualitative dependent variables is to construct index based on multiple correspondence analysis (Michaud and Van Soest, 2008). 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 suppose 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 Granger causality using the Hellinger distance under the assumption that the interest process is β-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 suppose that causal links between our two binary dependent variables change over time and we need to assess the causal links between states. Full description of the test procedure is provided in the following section.

[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

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

3.1.2 Model specification

Let us $W_{i,t}$ denote job status and $H_{i,t}$ denote 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 :

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

Testing for Granger non-causality on this qualitative process can be done by the use of a non-parametric test : the Kullback causality measure by Gouriéroux, Monfort and Renault (1987). We assume our process to be an homogeneous (among individuals) Markov chain of order one. Formally, we first assume $p_{i,t}(s_{i,t}|s_{i,t-1}) = p_t(s_t|s_{t-1})$ for all individual. Thus, the test statistics used for this purpose is (for the non-causality from W to H):

$$\hat{C}_{W\,to\,H} = \frac{1}{T}\sum_{t=1}^{T}\hat{C}_{W\,to\,H}(t) = \frac{1}{T}\sum_{t=1}^{T}\sum_{w=0}^{1}\sum_{h=0}^{1} I\hat{\pi}_{t-1}(w,h)\hat{C}_{W\,to\,H}(t,w,h) \tag{3.5}$$

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 :

$$\hat{C}_{W\,to\,H}(t,w,h) = \sum_{s^2_t=0}^{1}\hat{p}_t((.,s^2_t)|(w,h))log\frac{\hat{p}_t((.,s^2_t)|(w,h))}{\hat{p}_{H,t}(s^2_t|h)} \tag{3.6}$$

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

$$\hat{C}_{W,H}(t,w,h) = \sum_{s^1_t=0}^{1}\sum_{s^2_t=0}^{1}\hat{p}_t((s^1_t,s^2_t)|(w,h))log\frac{\hat{p}_t((s^1_t,s^2_t|w,h))}{\hat{p}_t((s^1_t,.)|(w,h))\hat{p}_t((.,s^2_t)|(w,h))} \tag{3.7}$$

With

$$\hat{p}_t((.,s^2_t)|(w,h)) = \sum_{s^1_t=0}^{1}\hat{p}_t((s^1_t,s^2_t)|(w,h))$$

$$\hat{p}_t((s^1_t,.)|(w,h)) = \sum_{s^2_t=0}^{1}\hat{p}_t((s^1_t,s^2_t)|(w,h))$$

$$\hat{p}_{H,t}(s^2_t|h) = \frac{\sum^1_{s^1_{t-1}=0}\sum^1_{s^1_t=0}\hat{p}_t((s^1_t,s^2_t)|(s^1_{t-1},h))\hat{\pi}_{t-1}(s^1_{t-1},h)}{\sum^1_{s^1_{t-1}=0}\hat{\pi}_{t-1}(s^1_{t-1},h)}$$

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 :

$$\hat{C}_{W\,to\,H}(t) = \sum_{h=0}^{1}\sum_{w=0}^{1} 2I\hat{\pi}_{t-1}(w,h)\hat{C}_{W\,to\,H}(t,w,h) \tag{3.8}$$

$\hat{C}_{W\,to\,H}(t)$ 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

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

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

and

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

H is given by :

$$Ctr_{W\,to\,H}^{(w,h)} = \frac{2I\hat{\pi}_{t-1}(w,h)\hat{C}_{W\,to\,H}(t,w,h)}{\hat{C}_{W\,to\,H}(t)} \tag{3.9}$$

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 lag causality measure from H to W.

3.1.3 Transition probabilities estimation

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

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

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 the fact that they was in state (w,h) at $t-1$, and $N_{t-1}(w,h)$ denotes the number of individual in 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 :

$$P\Big(s_{it} = s_k|s_{i,t-1}, X_{it}\Big) = \frac{exp((s_{i,t-1}, X_{it})'\beta_k)}{1 + \sum_{j=1}^{3} exp((s_{i,t-1}, X_{it})'\beta_j)} \; with \; k = 1, ..., 4 \tag{3.10}$$

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

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 last 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 role in professional path (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.

3.2 Data and related statistics

3.2.1 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

[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 invited 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

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.

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). The use of qualitative variable for measuring job (extensive margin) instead of quantitative ones such as number of hours worked (intensive margin) is not very restrictive. In fact the job impact on health condition is mainly due to whether individuals have choice or not on their amount of work (Caroli and Bassanini, 2015). In 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). These desegregations have been done in order to account for the severity of illness in the causal links analysis. The disability index and the risk of death variable are constructed 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.

3.2.2 Descriptive statistics on states and transitions

On the overall observation period, reported illness represents around 22% of observations when employment is approximately 86% of observations. 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 3.1(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 report 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 3.1(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 of nature, 16 transitions are available. But, we will 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 :

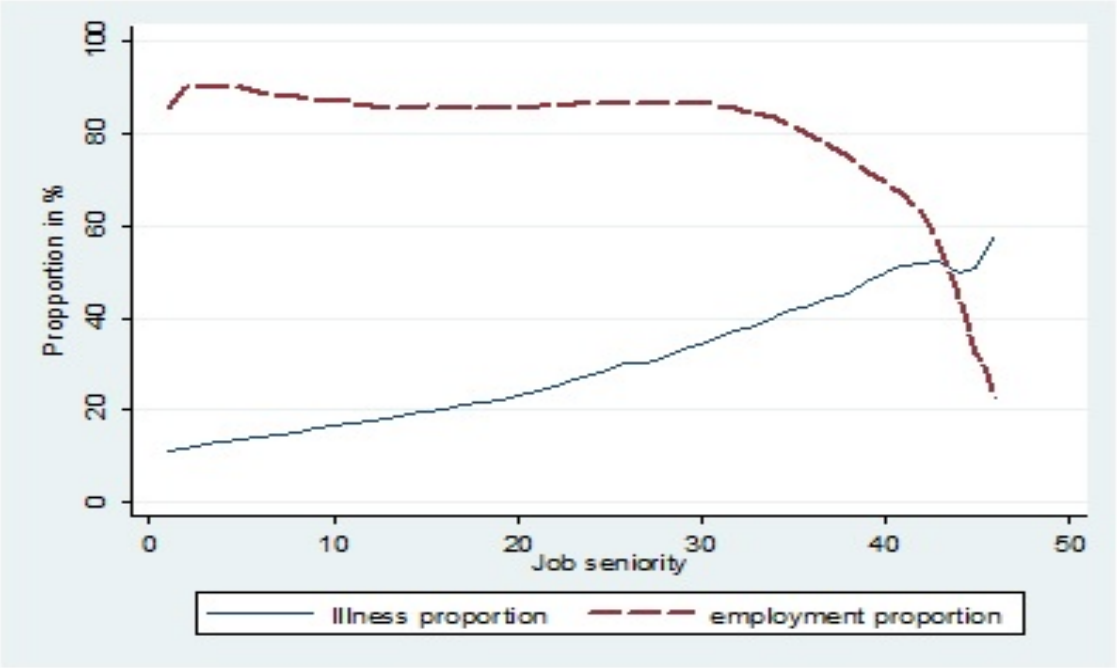
- From **healthy and unemployed** to **healthy and employed** : that can be analyzed as being healthy promotes entering work (Cai and Kalb 2006 for econometric evidences or Benjamin and Wilson 2005 for sociological evidences).

- 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).

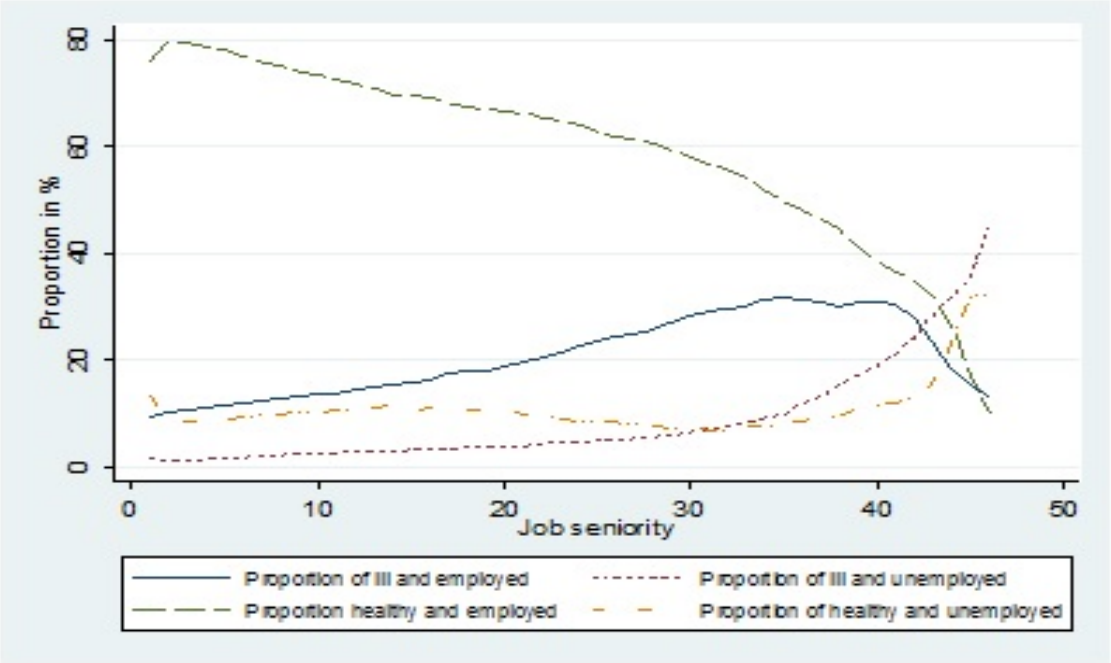
- From **healthy and employed** to **ill and employed** : it denotes, *ceteris paribus*, that working

[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 3.1: Evolution of proportion of individuals in each state of nature

conditions and working painfulness degrade health condition (Debrand 2011 for some evidence). 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 3.2.

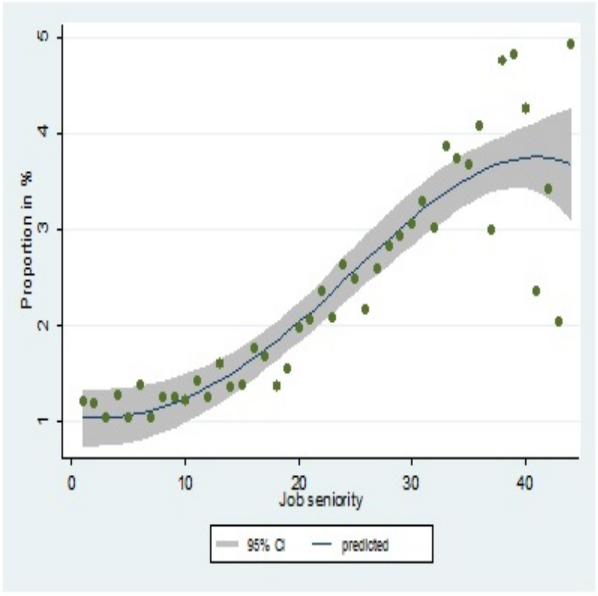
We compute in Figure 3.2, the transition probabilities between states and fit a non-linear curve with confident interval. It clearly appears that transition probabilities between states of nature are not homogeneous during the professional life. They have different evolution directions during professional life.

- From **healthy and unemployed** to **healthy and employed** (Figure 3.2(d)). 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.

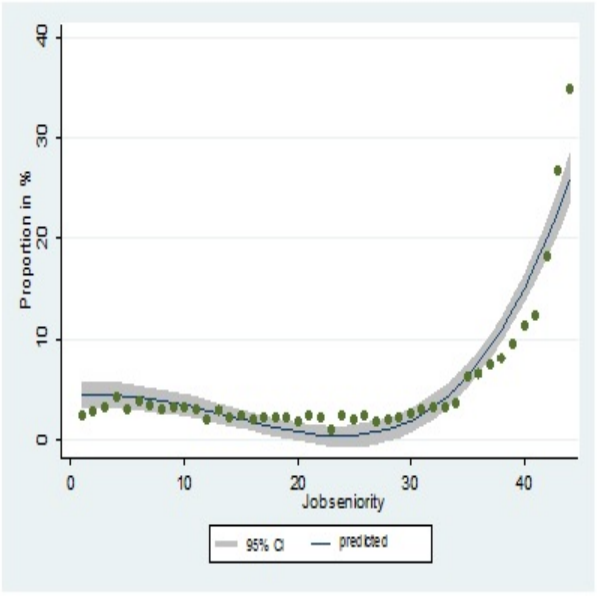
- From **ill and employed** to **ill and unemployed** (Figure 3.2(b)). At the beginning of the professional life, there are no evidence that being ill for an individual induces lost of his job (around 2.5%). This transition remains lower than 4% until 30 years of professional life. But, after 30 years of professional life, this transition rises exponentially.

- From **healthy and employed** to **ill and employed** (Figure 3.2(a)). 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.

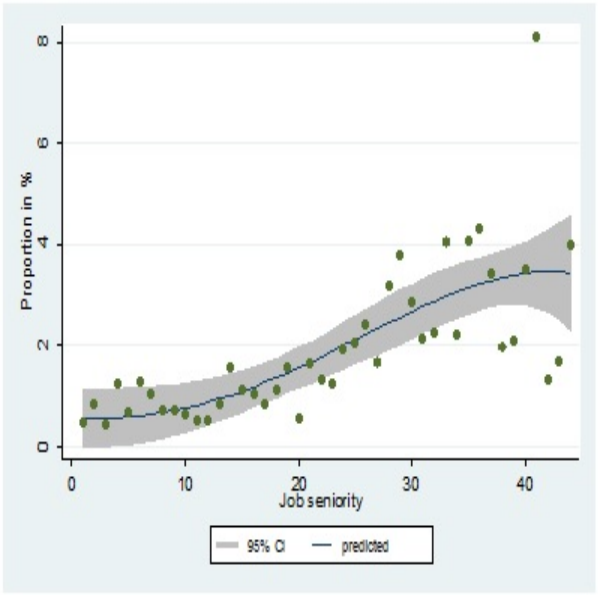
- From **healthy and unemployed** to **ill and unemployed** (Figure 3.2(c)). This transition is less than 1% at the beginning of professional life. As an individual progresses in his professional life, the likelihood of transition from "healthy and unemployed" to "ill and unemployed" increases and reaches 4% at the end of their professional life.



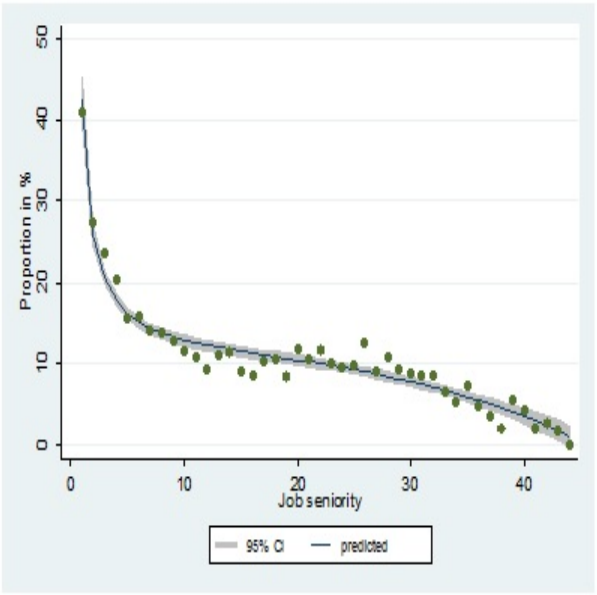
(a) Transition from healthy and employed to ill and employed



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



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



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

Figure 3.2: Dynamic of the transitions between some states of nature.

This analysis remains the same regardless of the professional life beginning period. When we consider the cohorts of individuals with professional life beginning 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 3.1 for the first period of professional life and for the 10th year of professional life.

Table 3.1: 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										

As we can see in Table 3.1, at the beginning of professional 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.

3.3 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].

3.3.1 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 3.2[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 3.2[18]. As we can see on Table 3.2, we can conclude at 5% significance level, the rejection 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

[16]The 45 estimated multinomial logistic models for the case 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 alternative approach to estimate probabilities instead of using empirical probabilities. However we present results of only one of them in appendix (table 3.4.

[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 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.

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 3.3[19], non-causality from health condition to job 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

[19]Figure 3.3(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 3.3(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 line 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 3.2: 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

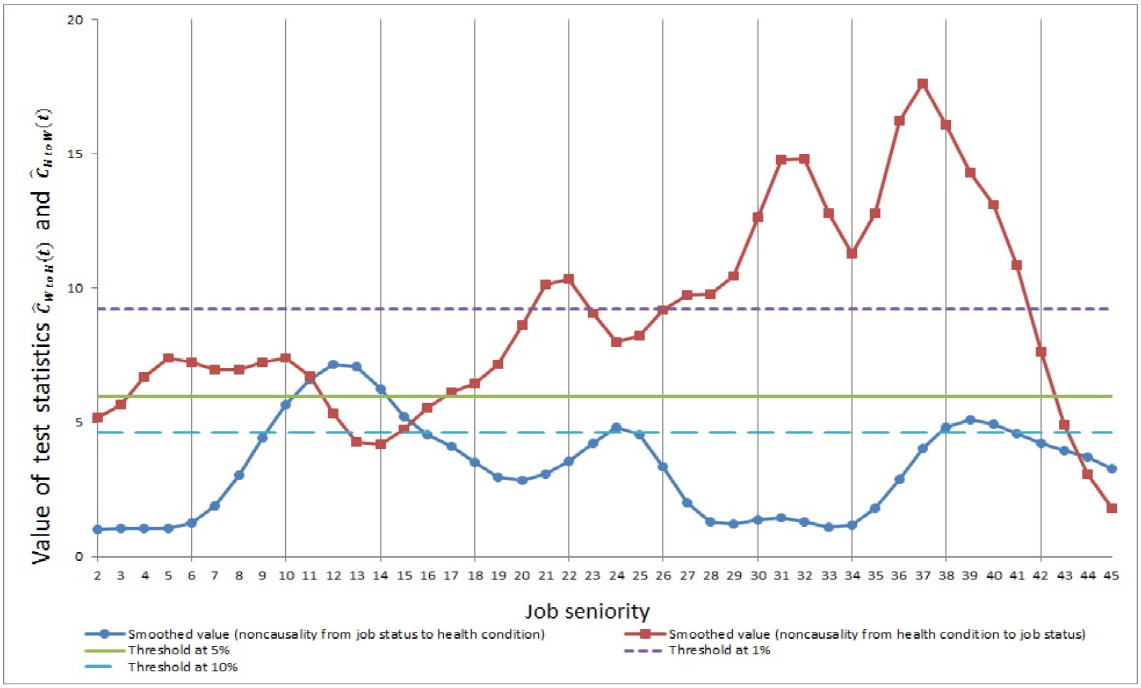
the smoothed curve that at 5% significance level, job status causes health condition during the 11^{th} and the 15^{th} years of professional life. But at 10% significance level, we can conclude that during the period between the 9^{th} to 17^{th} year of professional life, job status causes health condition. This finding illustrates the cumulative effects of job status on health condition (Barnay, 2015). When we consider illness with large disability index and illness with large risk of death, as we can see from Figure 3.3(b) and Figure 3.3(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 3.3(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}, 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

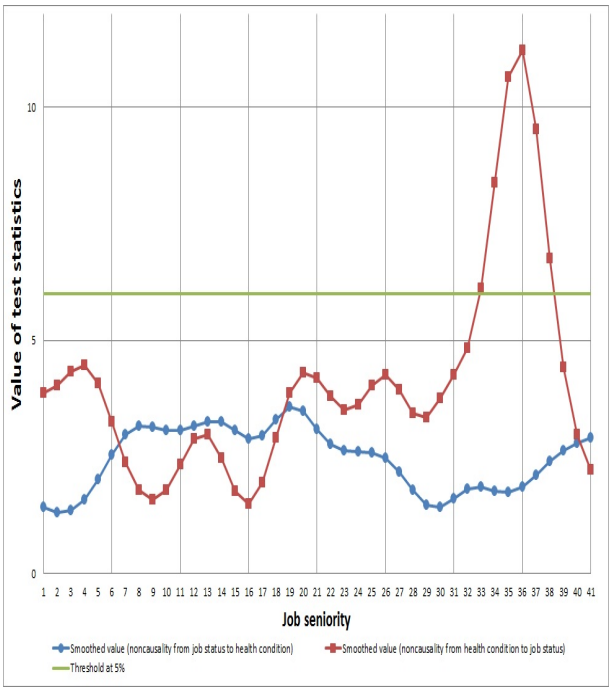
life. During this period, health significantly causes job status but the contrary is not significant. At each time period, we have only one unidirectional causal link that is significant. However, this finding does not remain true when we consider illness with large disability index or illness with large risk of death. Only health condition causes job status after 33 years and after 27 years of professional life respectively in these cases.

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.

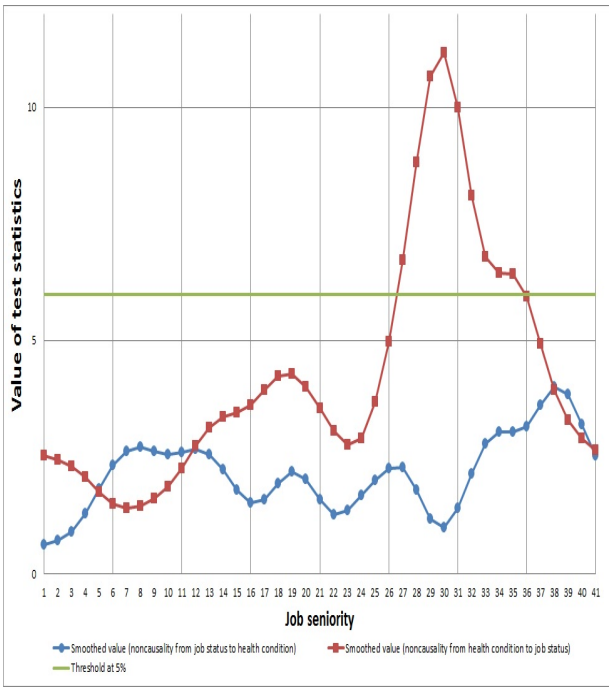
For a better analysis of those causal links, we will compute in the next section, the contribution of each states to the causal links at all periods. We also compute causality measures at individual level and estimate a model that aims to assess which of individual characteristics affect the causal link.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



(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

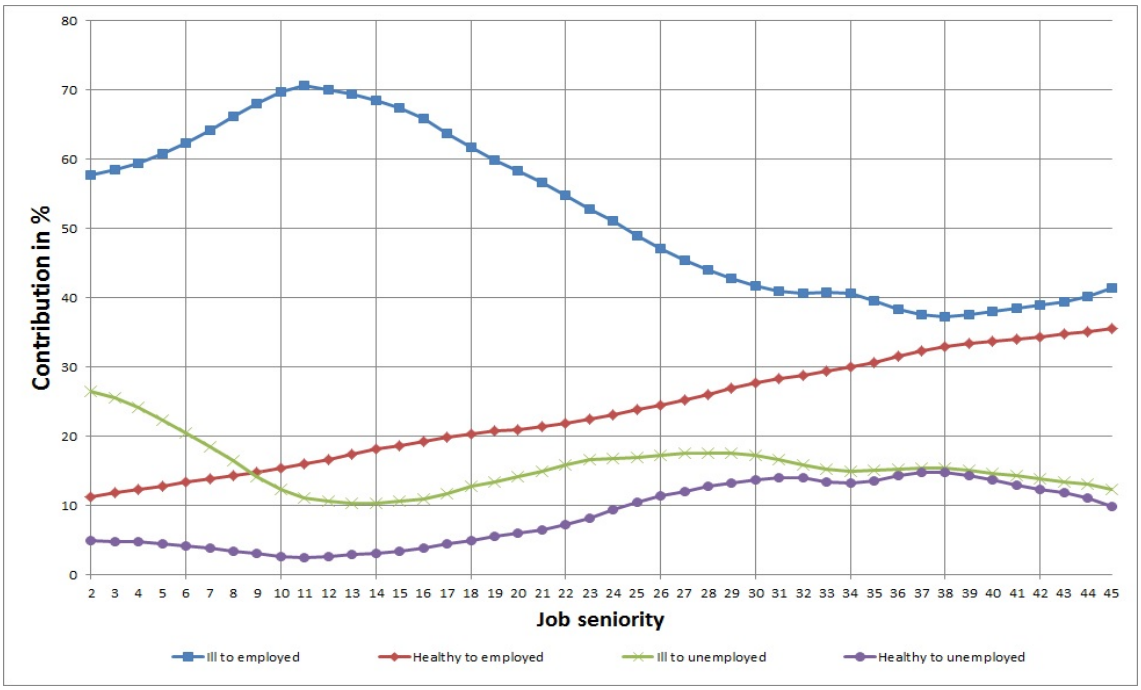
Figure 3.3: Dynamic of causality links between health and job

3.3.2 Contributions of states to causal links

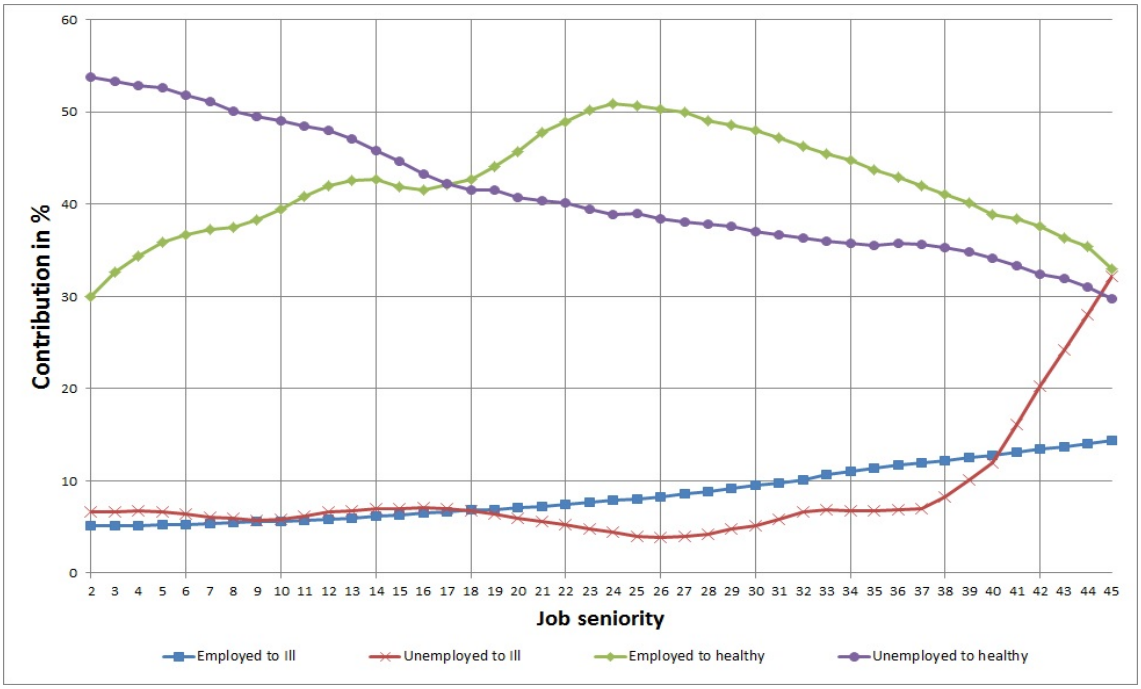
In order to give a better analysis of the dynamic of causal links over professional life (see Figure 3.2), we compute the contributions ($Ctr^{(w,h)}_{W\,to\,H}$ see Equation 3.9) of states of nature to the causal link $\hat{C}_{W\,to\,H}(t)$ at each time period. Dynamic of contributions are shown in Figure 3.4.

For the causal link from health condition to job status (see Figure 3.4(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 year of professional life.

For the causal link from job status to health condition (see Figure 3.4(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.



(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 3.4: Dynamic of contribution of states to causality links between health and job

3.3.3 Contributions of individual characteristics to causal links

This section aims to highlight which individual characteristics affect significantly the causal links. For that 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 3.8. $\hat{C}_{W\,to\,H}(i,t)$ is given by :

$$\hat{C}_{W\,to\,H}(i,t) = \sum_{h=0}^{1}\sum_{w=0}^{1} 2I\hat{\pi}_{i,t-1}(w,h)\hat{C}_{W\,to\,H}(i,t,w,h) \tag{3.11}$$

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 3.3. 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 3.2(a), 3.2(b) and 3.2(d) 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 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).

[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 $Threshold_{\hat{C}_{W\,to\,H}} = -\frac{-0.01933}{-2*0.00073}$

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. Researches show 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, the higher the school grade is, the lower the causal link from job to health is. 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, foreign, 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 (2010)'s 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.

Table 3.3: Effects of individual characteristics on causal links

Variable	$\hat{C}_{H\,to\,W}$	$\hat{C}_{W\,to\,H}$
	Coef.	Coef.
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)
ρ_u (variance due to u_i)	0.41508^{***}	0.0283^{***}

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

$***$: significant at 1% ; $**$: significant at 5%

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.

Appendices

Table 3.4: 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				

Table 3.5: Hausman test for IIA assumption

Omitted	Chi-square	Decision
IU^{*}	1.556	IIA met
NIE^{*}	4.484	IIA met
NIU^{*}	0.234	IIA met

Chapter 4

Early Retirement Decision : What are the Main Reasons for Early Retirement Decisions? Lessons from a Dynamic Structural Modelling[1]

[1]Authors : **Eric Delattre**, ThEMA-UCP; **Richard Moussa**, ThEMA-UCP and ENSEA Abidjan

Abstract

Early retirement has many causes according to economic and sociological literature. These causes may be the preference for leisure, financial and health conditions, and social environment. In our paper, we aim to specify and estimate an econometric model to assess the early retirement decision-making process for aged workers. We specify a worker's utility function from which we derive worker's probability to retire earlier that depends on his or her health stock, estate value and preference for future. We also estimate two functions : an health production by investment and an health consumption by working that are key factors in the individual's decision to retire earlier. Thus, we show that our model disentangles between three groups of workers : (i) those who choose early retirement (20.07%), (ii) those who will never choose early retirement (44.97%) and (iii) those who are uncertain about early retirement (34.96%). We also show that our computed early retirement probability is a good predictor of early retirement as it is causal for observed early retirement.

Keywords: Early retirement, Health, Estate value, Working condition, QALY, Space-state model, Utility function, Causality

JEL Classification: C32, C38, C51, D81, D91, I10, J24, J26

Introduction

There is a large literature in sociology and economics about the early retirement decision. These studies highlight the preference for leisure, the good financial conditions, the individual health conditions, the social environment, and the working environment as main factors of early retirement decision.

The individual preferences for leisure is related to the financial condition. For an individual, if early retirement does not deteriorate her financial condition then she is more likely to retire earlier; and this likelihood is greater when she prefers leisure (Brothers, 2000). In individual social environment, the retirement status of spouse specially and that of family members and neighbours in general can increase the likelihood of an individual to retire earlier (Brothers, 2000). Individuals with many post-retirement opportunities are more likely to retire earlier. These post-retirement opportunities can be related to the education level, the unemployment rate in the region, or the industry sector of worker. Individual health condition is one of the major factor that determines the labour force participation. The likelihood to continue working cannot be satisfy for an individual in very bad health condition. Thus, the perceived ability to remain in job market and the good working condition reduce the probability that individual early retires.

However, the use of relevant micro economics datasets to analyze the theoretical findings is recent. With the collection of recent specific datasets like the survey of health, ageing and retirement in Europe (SHARE), some applied micro economics papers have addressed the early retirement issue. These papers commonly use as depend variable the binary variable that captures if or not the individual looks for early retirement. The dependent variable can also be constructed as a binary variable that is one if retirement age is under 65. One of the determinant that are commonly underlined in the literature is related to working conditions. Even if Quinn (1977) finds that there is no evidence of the influence of job characteristics and financial variables on early retirement of white married men in the US, many recent studies challenge this finding. Bazzoli (1985) finds that economic variables play more important role than health in retirement decision-making process. With the first wave of the SHARE dataset, Debrand and Blanchet (2008) show that being satisfy with job reduces the probability to look for early retirement. Mein et al (2000) also show on a British dataset that less satisfied workers are more likely to retire earlier. Early retirement is also specific to activity sector (Dorn and Sousa-Poza 2004). Alhawarin (2014) shows that workers in army and security forces sector in Jordan are more likely to retire earlier. Pollak (2012), by the use of a panel dataset from SHARE, shows that health status, job satisfaction and working condition are the major factors that explain the fact that individual looks for early retirement or not. She also highlights the important role of rewards in keeping in labour force older workers even with disabilities. Siegrist et al (2006) also show that effort reward imbalance and poor quality of work are main factors that explain that workers look for early retirement. The workload is also an important determinant of early retirement (Boumans et al, 2008). There are also empirical evidences that early retirement is related to earnings. Workers with higher-paid employment are more likely to retire earlier (Mein et al, 2000). Dorn and Sousa-Poza (2004), on the Swiss Labour force survey dataset, show that wage rate has a non linear effect on early retirement. Both workers with high and low wage are more likely to keep working. Dorn and Sousa-Poza also highlight the important role played by the coverage in the social security system. Quinn (1977) finds that eligibility to social security lower the probability to participate to labour market. Another main factor of early retirement to be highlighted is the post retirement opportunities for early retired. These opportunities are related with unemployment rate, school grade and activity sector (Brothers 2000) or to demographics characteristics such as living in couple and spouse employment status (Jiménez-Martin et al, 2015). Workers that retire earlier continue working after retirement (30% of them, see Dorn and Sousa-Poza 2004), even in jobs with a degree of informality (Alhawarin, 2014).

Health is also an important determinant for early retirement decision. Both current health condition and perceived future health condition play an important role in the decision of early retirement. Workers retire if they have poor health (Galama et al, 2013) or if they think that their future health condition will not allow them to continue working. Bazzoli (1985) suggests the use of current health status in addition of perceived future health limitations to better assess effect of health on early retirement. By analyzing a set of married white men aged between 58 and 63 extracted from the US social security administration's retirement history study, Quinn (1977) finds that health limitations lower the probability to participate to labour market. Coe and Zamarro (2008) use the SHARE dataset to show that retirement has a health-preserving effect on general health. They find that

being retired reduces the probability to report bad health condition. Health problems increase the probability of early retirement (Albuquerque, 2009). Disability, severity of health shock, increased rate of sickness absence and alcohol abuse are strong determinants for early retirement (Szubert and Sobala 2005, Jiménez-Martin et al, 2006). Health shocks can also induce working hours reduction (Cai et al, 2006). The fear that health condition limits working abilities increases the probability to look for early retirement (Debrand and Blanchet, 2008). Boumans et al (2008) focus on Belgian older nurses and show that perceived health condition is a major determinant of early retirement.

Social environment of workers plays an important role in the decision of early retirement. The household wealth has a negative impact on early retirement (Alhawarin 2014). The family size (Albuquerque 2009 and Alhawarin 2014) is also a determinant of early retirement. However, the effect of the family size can be different among countries. Alhawarin (2014) finds that the family size increases the probability of early retirement in Jordan while Albuquerque (2009) shows that small family size increases the probability of early retirement. Another social environment variable that is determinant for early retirement is the partner employment status. This is important because couples coordinate their retirement decision (Albuquerque 2009). Many studies underlined the important role of partner employment status in early retirement decision (Dorn and Sousa-Poza 2004, Szubert and Sobala 2005, Boumans et al 2008, Albuquerque 2009, Jiménez-Martin et al, 2015). Workers with retired partner are more likely to retire earlier.

These findings confirm the economic and sociological theory. It appears clearly that individual will retire earlier if (i) his health condition limits his capability to continue working, (ii) his perceived future health condition does not allow him to continue working, or (iii) he has a job with low quality. In this paper, we aim to assess the early retirement decision process for aged workers. For this purpose, we specify and estimate a micro economics model that accounts for workers financial, health, and working conditions and some socio economics variables. Our model provides us an estimation of the individual retirement probability at each time period. By the use of a European panel dataset, our estimates provide a tool to assess the early retirement decision among a large set of countries with different health and retirement systems. In Section 1, we describe our methodology. We present the dataset and some related descriptive statistics in Section 2. Section 3 presents our results and discussion, then we conclude by giving some implications of our findings.

4.1 The economic model

In this section, we present the theoretical framework of the current paper. We first describe the dynamic of health stock equation and the dynamic of estate equation and we end by specifying the model and its constraints.

4.1.1 Health equation

Based on Grossman's (1972) theory on health capital, we propose a health stock dynamic equation for workers. The original model proposed by Grossman (1972) is the following:

$$H_t - H_{t-1} = I_{t-1} - \delta_{t-1} * H_{t-1} \tag{4.1}$$

where δ_{t-1} denotes the health depreciation rate at time $t-1$, H_t is the health stock at t and I_{t-1}, the investment in health. In our context, as we focus on workers, it is possible to decompose health depreciation rate into three sources that are:

- Health depreciation rate due to working condition: this depreciation rate is time-variant as well as working condition can be improved or can be deteriorated among time. Many changes have been done for facilitating job condition at governments level (working legislation) and at firms level. The depreciation rate due to working condition is also individual specific and specific to economic sector.
- Health depreciation rate due to ageing: it is well known that as individual age increases, he becomes more vulnerable to illnesses. Health depreciation rate due to ageing is not invariant during individual life. Many demographic and medical studies underlined that at the beginning and at the end of life, individuals are more vulnerable. To model this fact, investigators generally include in their model a quadratic specification for age effects.
- Natural health depreciation rate: individual health naturally declines as soon as he reports an illness once. Individual health might deteriorate due to the long run impact of illnesses on individual health. This depreciation can be growing up or slowing down among time period.

Thus, from the original demand health stock dynamic equation, δ for an individual i at date t can be desegregated as follows:

$$\delta_{i,t} = \alpha_1 C_{i,t} + \alpha_2 Agesq_{i,t} + \alpha_3 Age_{i,t} + \alpha_4 H_{i,t} \tag{4.2}$$

Where $C_{i,t}$, $Age_{i,t}$, $Agesq_{i,t}$ and $H_{i,t}$ are respectively the working condition, age, the square of age and the health stock of individual i at the date t. The interest of equation 4.2 is that it allows to model the health depreciation rate due to working condition as an input consumed in a wealth production function. Thus, the reciprocal of this wealth production function (says g) will denote the earnings from the job : $\alpha_1 C_{i,t} H_{i,t} = g(W_{i,t})$, where $W_{i,t}$ denotes the job revenue for individual i at date t. The health depreciation function is a marginally increasing function in health.

To adjust his health stock to a desired level, worker can also invest in his health stock. This investment can be separate into two components according to his interest to current health stock level or his current health stock depreciation rate. We distinguish : (i) Investment in health care that is made when individual report an illness and (ii) Investment in health prevention that is made to prevent health to depreciate.

Notice that the decision to invest in health does not only depend on the need to adjust the health stock. It also depends on individual financial situation. An invest in health $I_{i,t}$ is produced by a function of the amount spent in health (Hall and Jones, 2007). Let f denotes the health production function by investment in health care or prevention, and $Exph_{i,t}$ the amount of estate that the individual i spends for his health at t. We can then rewrite I_t as follows:

$$I_{i,t} = f(Exph_{i,t}) \tag{4.3}$$

Let $p_{i,t}$ denotes the probability that individual i decides to retire at t.The expected health stock at $t+1$ is given by:

$$H_{i,t+1} - H_{i,t} = I_{i,t} - \Big((1-p_{i,t})\alpha_1 C_{i,t} + \alpha_2 Agesq_{i,t} + \alpha_3 Age_{i,t} + \alpha_4 H_{i,t}\Big) H_{i,t} \tag{4.4}$$

From Equation 4.4, the retirement probability (regardless the earning from current work) is given by:

$$p_{i,t} = \frac{1}{\alpha_1 C_{i,t}}\left(\gamma_{i,t+1} - \frac{I_{i,t}}{H_{i,t}} + \delta_{i,t}\right) \tag{4.5}$$

Where $\gamma_{i,t} = \dfrac{H_{i,t} - H_{i,t-1}}{H_{i,t-1}}$ denotes the gross health stock growth rate. This probability $p_{i,t}$ is a decreasing function of the anticipated gross health stock growth rate and the depreciation rate. The retirement probability is also a decreasing function of working condition and investment share.

4.1.2 Estate equation

Individual estate/worth includes financial assets (saves, amount on bound, stocks and mutual fund, value of whole life policies, amount on retirement account) and real assets (amount if selling cars, houses and owned firms) and debts. Estate accumulation process includes earnings from current job or revenue from pension if individual is retired. It also includes at each period expenditures in goods and others services. At date t, individual receives an interest π from his total estate at $t-1$. Thus, individual estate accumulation dynamic equation is the following:

$$E_{i,t} = (1+\pi)E_{i,t-1} + W_{i,t} - Exp_{i,t} \tag{4.6}$$

where $E_{i,t}$, $W_{i,t}$ and $Exp_{i,t}$ are respectively estate, revenue from current job status and total expenditures including the amount spent in health investment $Exph_{i,t}$ for individual i at date t.

When worker retires, he perceives a pension (says $P_{i,t}$) that is an amount lower than his earnings from his last job. In many European retirement systems, this pension is a share of the last job income. Let ω denotes the share of last job income that individual i receives at the date t as his retirement pension if he retires, then $P_{i,t} = \omega W_{i,T_R}$ (where T_R denotes the individual retirement date if he is retired or the current date if not). The ratio ω can be larger than one as its numerator includes all sources of retirement income (Brothers, 2000). Thus, $W_{i,t}$ can be re-expressed as $W_{i,t} = \omega W_{i,t} + (1-\omega)W_{i,t} * \mathbb{I}_{job}$.

Let $p_{i,t-1}$ denotes the probability that individual i decides at $t-1$ to retire at t. Thus the expected estate accumulation dynamic equation can be rewritten as follows:

$$E_{i,t} = (1+\pi)E_{i,t-1} + \omega W_{i,t} + (1-p_{i,t-1})(1-\omega)W_{i,t} - Exp_{i,t} \tag{4.7}$$

For simplicity, we can also suppose the reimbursement to be a part of expenditures and then $E_{i,t} = (1 + \pi)E_{i,t-1}$ can be interpreted as estate accumulation between $t-1$ and t.
From the equation 4.7, the probability of retirement is given by (regardless the health condition):

$$p_{i,t-1} = \frac{W_{i,t} - Exp_{i,t} - E_{i,t} + (1+\pi)E_{i,t-1}}{(1-\omega)W_{i,t}} \tag{4.8}$$

This probability is a decreasing function of estate growth rate and expenditures. The retirement probability is a decreasing function of retirement pension share (ω) only if earnings from current job and the estate accumulated between $t-1$ and t can not cover expenditures and reimbursement at t.

4.1.3 Model and constraints

There are two main factors concerning health stock and estate value at the period t on which individual can make a decision to maximize his or her utility. These controls are:

- **Decision to retire**: if individual decides to work, his or her health stock decreases by $\alpha_1 C_{i,t} H_{i,t}$ due to health condition and his or her estate value increase of $(1-\omega)W_{i,t}$. If individual decides to retire, he or she preserves his or her health stock from a decrease of $\alpha_1 C_{i,t} H_{i,t}$ for a lost of $(1-\omega)W_{i,t}$ in estate value. Thus, for a unit decrease in health due to work, individual estate increases by $\frac{(1-\omega)W_{i,t}}{\alpha_1 C_{i,t-1} H_{i,t-1}}$. As g denotes the health depreciation function due to job, we can express the health depreciation saved by a retirement as $\alpha_1 C_{i,t} H_{i,t} = g\big((1-\omega)W_{i,t}\big)$. The second approach to model this situation is to suppose that workers can decide to reduce their per week working time from the full time work to a partial time work with a certain rate in order to preserve health decrease. In both cases, the workers are in an adjudication situation between health preservation and estate accumulation. However, as we are interested on retirement (individual decides either to continue working or not), this second case will be treated idly.
- **Amount invested in health**: investment in health stock that aims to slow down health depreciation rate by a payment for care or for health preservation involves reducing estate level to earn a compensation of health stock depreciation. If worker decides not to invest in his health stock ($I_{i,t-1} = 0$), his health stock decreases by $\delta_{i,t-1} * H_{i,t-1}$. A unit increase in health investment expenditures $Exph_{i,t-1}$ increases the health stock by $f^{'}(Exp_{i,t-1})$.

It denotes that there is a substitution rate between estate and health stock as worker can substitute a part of his estate (by payment for care or prevention) in health stock and he might also grant an health decrease due to working condition in order to maintain unchanged his estate level.

We suppose that worker has a utility function that only depends on his estate and his health stock $u(H_{i,t}, E_{i,t})$. We will analyze both separable and non-separable in health stock and estate utility function case. The separable

and non-separable utility function have respectively the following forms:

$$u(H_{i,t}, E_{i,t}) = \frac{E_{i,t}^{1-\lambda}}{1-\lambda} + a * \frac{H_{i,t}^{1-\gamma}}{1-\gamma} \tag{4.9}$$

$$u(H_{i,t}, E_{i,t}) = \left[(1-\nu) * H_{i,t}^{1-\eta} + \nu * E_{i,t}^{1-\eta}\right]^{\frac{1}{1-\eta}} \tag{4.10}$$

Where a, γ, λ, ν and η are positive. Let T_i^R denotes the difference between the individual i's age and the legal age of retirement in his country, and T_i^c the difference between individual i's age and the life expectancy at the legal retirement age in his country. We assume that worker i's future flow of utility discount factor at period t is $r_{i,t}$. The total lifetime utility can be disaggregated into two time periods (the period until the legal retirement age, and the period between the legal retirement age and the end of life). Thus, at each date t, the total lifetime utility is given by:

$$u_t(H_i, E_i) = \sum_{\tau=0}^{T_i^R} r_{i,t}^{\tau} u_{i,t+\tau}(H_{i,t+\tau}, E_{i,t+\tau}) + \sum_{\tau=T_i^R}^{T_i^c} r_{i,t}^{\tau} u_{i,t+\tau}(H_{i,t+\tau}, E_{i,t+\tau}) \tag{4.11}$$

At each period t, individual's controls that described his state are the amount invested in health stock ($Exph_{i,t}$) and the decision to retire that determines his earnings level. An individual optimal state at t is a value of $Exph_{i,t}$ and $(1-\omega)W_{i,t}$ that maximizes his utility. Thus, the worker's program is the following :

$$u = \underset{((1-\omega)W_{i,t}, Exph_{i,t})}{Max} \left\{ \sum_{\tau=0}^{T_i^c} r_{i,t}^{\tau} u_{i,t+\tau}(H_{i,t+\tau}, E_{i,t+\tau}) \right\} \tag{4.12}$$

Subject to

$$\begin{aligned} H_{i,t+\tau+1} = I_{i,t+\tau} + (1-\delta_{i,t+\tau}) * H_{t+\tau} \geqslant H_{min} \quad \forall \tau = 1, ..., T_c \\ E_{i,t+\tau} + W_{i,t+\tau} \geqslant Exp_{i,t+\tau} \quad \forall \tau = 1, ..., T_c \\ I_{i,t+\tau} = f(Exph_{i,t+\tau}) \ \forall \tau = 1, ..., T_c \end{aligned}$$

Where H_{min} denotes the health stock level under which retirement is imposed either by worker or by firm even if worker want to continue working. The first constraint means that individual health stock must always be greater than a vital minimum. The second constraint means that individual estate and earnings for the current job must cover his expenditures. The last constraint is a health production function (investment in health). The set of constraints can be rewritten as follows :

$$\begin{aligned} H_{min} \leqslant f(Exph_{i,t+\tau}) + \left(1 - g\big((1-\omega)W_{i,t+\tau}\big) - \alpha_2 Agesq_{i,t+\tau} - \alpha_3 Age_{i,t+\tau} - \alpha_4 H_{i,t+\tau}\right) * H_{t+\tau} \\ \forall \tau = 1, ..., T_c \\ Exp_{i,t+\tau} \leqslant E_{i,t+\tau} + W_{i,t+\tau} \ \ \forall \tau = 1, ..., T_c \end{aligned}$$

4.1.4 Model solving

Let θ_1 and θ_2 denote the Lagrange multiplier respectively on estate stock and on health. The Lagrangian of the problem described in section 2.3 is given by[2]:

$$L = L\Big(Exph_{i,t}, (1-\omega)W_{i,t}\Big) = \sum_{\tau=0}^{T_i^c} r_{i,t}^{\tau} u_{i,t+\tau}(H_{i,t+\tau}, E_{i,t+\tau}) + \theta_1 \sum_{\tau=0}^{T_i^c} \Big(E_{i,t+\tau} + W_{i,t+\tau} - Exp_{i,t+\tau}\Big) + \theta_2 \sum_{\tau=0}^{T_i^c} \Big(f(Exph_{i,t+\tau}) + \big(1 - g\big((1-\omega)W_{i,t+\tau}\big) - \alpha_2 Agesq_{i,t+\tau} - \alpha_3 Age_{i,t+\tau} - \alpha_4 H_{i,t+\tau}\big) * H_{i,t+\tau}\Big)$$

Due to the fact that the health stock at $t+1$ is a function of the current depreciation rate and the current investment, the retirement decision does not affect also health stock at $t+1$. But when a worker decides at t to retire at $t+1$, his decision affects his health stock at $t+2$. In terms of estate accumulation, a retirement decision taken at t for $t+1$ directly affects the estate level from date $t+1$ until the legal retirement age. The dynamic in individual state (health stock and estate) if he retires at $t+1$ is the following one:

$$H_{i,t+\tau} = f(Exph_{i,t+\tau-1}) + \Big(1 - \alpha_2 Agesq_{i,t+\tau-1} - \alpha_3 Age_{i,t+\tau-1} - \alpha_4 H_{i,t+\tau-1}\Big) * H_{t+\tau-1} \; \forall \tau = 2, ..., T_i^c$$
$$E_{i,t+\tau} = (1+\pi)E_{i,t+\tau-1} + P_{i,t+r} - Exp_{i,t+\tau} \; \forall \tau = 1, ..., T_i^c$$

Let u_t^R and u_t^J denote respectively the utility at t if the individual retires and if he continues working. Thus, the program can be rewritten as follows:

$$L = \Big(u_{i,t}(H_{i,t}, E_{i,t}) + p_{i,t} \sum_{\tau=1}^{T_i^c} r_{i,t}^{\tau} u_{i,t+\tau}^R(H_{i,t+\tau}, E_{i,t+\tau}) + (1-p_{i,t}) \sum_{\tau=1}^{T_i^c} r_{i,t}^{\tau} u_{i,t+\tau}^J(H_{i,t+\tau}, E_{i,t+\tau})\Big) + \theta_1 \sum_{\tau=0}^{T_i^c} \Big(E_{i,t+\tau} + W_{i,t+\tau} - Exp_{i,t+\tau}\Big) + \theta_2 \sum_{\tau=0}^{T_i^c} \Big(f(Exph_{i,t+\tau}) + \big(1 - g\big((1-\omega)W_{i,t+\tau}\big) - \alpha_2 Agesq_{i,t+\tau} - \alpha_3 Age_{i,t+\tau} - \alpha_4 H_{i,t+\tau}\big) * H_{i,t+\tau}\Big)$$

The first order conditions with respect to $Exph_{i,t}$ and $(1-\omega)W_{i,t}$ allow to derive the following expressions (we assume that $P_{i,t+\tau} = \omega W_{i,t}, \; \forall \tau \geqslant 1$ when individual i decides at t to retire):

- For a separable utility function (in this case $u'_{t+1,H} = u'^R_{t+1,H} = u'^J_{t+1,H}$):

$$r_{i,t} = \frac{E_{i,t}^{-\lambda}}{a * f'(Exph_{i,t}) * H_{i,t+1}^{-\gamma}} \tag{4.13}$$
$$p_{i,t} = \frac{g'((1-\omega)W_{i,t}) * H_{i,t} - f'(Exph_{i,t})}{g'((1-\omega)W_{i,t}) * H_{i,t} - \sum_{\tau=1}^{T_i^c} r_{i,t}^{\tau-1} \frac{E_{i,t+\tau}^{-\lambda}}{a * H_{i,t+1}^{-\gamma}}}$$

[2]We can make the assumption that H_{min} is zero

- For a non-separable utility function:

$$r_{i,t} = \frac{\nu * E_{i,t}^{-\eta} * u_{i,t}^{\eta}}{(1-\nu) * f^{'}(Exph_{i,t}) * H_{i,t+1}^{-\eta} * \left(u_{i,t+1,J}^{\eta} + p_{i,t} * (u_{i,t+1,R}^{\eta} - u_{i,t+1,J}^{\eta}) \right)} \tag{4.14}$$

$$p_{i,t} = \frac{g^{'}((1-\omega)W_{i,t}) * H_{i,t} * u_{i,t+1,J}^{\eta} - f^{'}(Exph_{i,t}) * \left(u_{i,t+1,J}^{\eta} + p_{i,t} * (u_{i,t+1,R}^{\eta} - u_{i,t+1,J}^{\eta}) \right)}{g^{'}((1-\omega)W_{i,t}) * H_{i,t} * u_{i,t+1,J}^{\eta} - \frac{\nu}{1-\nu} * \sum_{\tau=1}^{T_i^c} r_{i,t}^{\tau-1} * \frac{E_{i,t+\tau}^{-\eta}}{H_{i,t+1}^{-\eta}} * u_{i,t+\tau,R}^{\eta}}$$

Notice that if we relax the assumption that $P_{i,t+\tau} = \omega W_{i,t} \, \forall j \geqslant 1$, the factor $\sum_{\tau=1}^{T_i^c} r_{i,t}^{\tau-1} u_{t+\tau,E}^{'R}$ of the retirement probability denominator is reduced to $u_{t+1,E}^{'R}$. It denotes that when worker's retirement pension is a fixed share of his last revenue from job, he accounts for the current discounted value of marginal utilities of this revenue until the end of life. However, he only accounts for the current value of marginal utility of his current revenue when his pension is null or is an unknown share of his last revenue.

At the optimum, the individual discount factor is the ratio between the marginal utility of job revenue at t and marginal utility of health stock at $t+1$. It means that individual will prefer future when his marginal utility of his job revenue is closer or larger than his marginal utility of health stock. Individual's probability to retire depends on the difference between the marginal depreciation of health due to work and the marginal health investment productivity and the difference between the marginal depreciation of health due to work and the discounted future flows of marginal utility of job revenue divided by the marginal utility of health stock.

For the separable utility function case, we use the probability in Equation 4.13 to derive conditions under which workers will not choose early retirement (see areas 1 and 2 in Figure 4.1), conditions under which workers will choose early retirement (see areas 3 and 4 in Figure 4.1), and conditions under which worker's decision is uncertain (see areas 5 and 6 in Figure 4.1).

Proposition 1: Worker will never choose early retirement (Areas 1 and 2 in Figure 4.1) if:

$$\sum_{\tau=1}^{T_i^c} r_{i,t}^{\tau-1} \frac{E_{i,t+\tau}^{-\lambda}}{a * H_{i,t+1}^{-\gamma}} < g^{'}((1-\omega)W_{i,t}) * H_{i,t} < f^{'}(Exph_{i,t})$$

$$or \sum_{\tau=1}^{T_i^c} r_{i,t}^{\tau-1} \frac{E_{i,t+\tau}^{-\lambda}}{a * H_{i,t+1}^{-\gamma}} > g^{'}((1-\omega)W_{i,t}) * H_{i,t} > f^{'}(Exph_{i,t})$$

If the marginal productivity of health expenditures is greater than the marginal health depreciation due to job and if the discounted future flow of marginal utility of estate is lower than the current marginal utility of health stock, worker will prefer continue working. However, even if marginal productivity of health expenditures is lower than the marginal health depreciation due to job, worker will continue working if the discounted future flow of marginal utility of estate is greater than the current marginal utility of health stock. We distinguish two groups of workers among those who do not choose early retirement. The former group is that of workers who enjoy continue working as their marginal health depreciation due to working condition and their marginal utility of

estate are very low, and their marginal productivity of health expenditures is very high. The later group is that of workers who are constraint to continue working in order to increase their future flow of utility. Their marginal health depreciation due to job is very large but their current estate is too low to allow them to retire without affecting their financial condition.

Proposition 2: Worker will choose early retirement (Areas 3 and 4 in Figure 4.1) if:

$$g^{'}((1-\omega)W_{i,t}) * H_{i,t} > \sum_{\tau=1}^{T_i^c} r_{i,t}^{\tau-1} \frac{E_{i,t+\tau}^{-\lambda}}{a * H_{i,t+1}^{-\gamma}} > f^{'}(Exph_{i,t})$$

$$or\ g^{'}((1-\omega)W_{i,t}) * H_{i,t} < \sum_{\tau=1}^{T_i^c} r_{i,t}^{\tau-1} \frac{E_{i,t+\tau}^{-\lambda}}{a * H_{i,t+1}^{-\gamma}} < f^{'}(Exph_{i,t})$$

If the discounted future flow of marginal utility of estate and the current marginal utility of health are fairly identical and the marginal productivity of health expenditures is lower than the marginal health depreciation due to job, worker will prefer an early retirement. However, even if the marginal productivity of health expenditures is higher than the marginal health depreciation due to job, worker will choose early retirement if the discounted future flow of marginal utility of estate is fairly identical with the current marginal utility of health. Two groups of workers can be distinguished among this category of workers. The former is that of workers who choose early retirement because the marginal health depreciation granted to earn an additional estate can not be covered by the health stock produced by the same amount invested in health. Furthermore, their marginal utility of estate is too large. The later group is that of workers who choose early retirement because they enjoy leisure. They are in good health condition, their marginal productivity of an additional health expenditures covers the marginal health depreciation due to working condition as they are in good working condition and their job does not highly depreciate their health. Furthermore, their financial condition can not be considerably deteriorated if they retire.

Proposition 3: Worker is uncertain about early retirement (Areas 5 and 6 in Figure 4.1) if:

$$\sum_{\tau=1}^{T_i^c} r_{i,t}^{\tau-1} \frac{E_{i,t+\tau}^{-\lambda}}{a * H_{i,t+1}^{-\gamma}} > f^{'}(Exph_{i,t}) > g^{'}((1-\omega)W_{i,t}) * H_{i,t}$$

$$or\ \sum_{\tau=1}^{T_i^c} r_{i,t}^{\tau-1} \frac{E_{i,t+\tau}^{-\lambda}}{a * H_{i,t+1}^{-\gamma}} < f^{'}(Exph_{i,t}) < g^{'}((1-\omega)W_{i,t}) * H_{i,t}$$

If the marginal productivity of health expenditures is greater than the marginal health depreciation due to job and the discounted future flow of marginal utility of estate is greater than the current marginal utility of health, worker will have uncertainty about continuing working. Likewise, if the marginal health depreciation due to job is greater than the marginal productivity of health expenditures, worker will have uncertainty about continuing working if the discounted future flow of marginal utility of estate is lower than the current marginal utility of health. We also distinguish two groups of workers among this category. The former is that of workers who are uncertain because they have a fairly good health condition but a worse financial condition. Thus, even if their marginal productivity of an additional health expenditures covers the marginal health depreciation granted to earn an additional revenue, they still uncertain as retirement deteriorates their financial condition. The later

group is that of workers who are uncertain because they have a fairly good financial and health conditions but their marginal health depreciation due to job is higher than their marginal productivity of health expenditures.

Notice that individuals in areas 1, 3 and 5 (respectively in areas 2, 4 and 6) have the same conditions in terms of health production and consumption. Workers in areas 1, 3 and 5 have a marginal health depreciation due to job larger than their marginal productivity of health expenditures. Thus, they should normally retire earlier. But those in area 1 will not retire earlier due to their high preference for future. While workers in area 5 are uncertain as they have a low preference for future. The same analysis can be made for workers in areas 2, 4 and 6 who should normally continue working as their marginal productivity of health expenditures is greater than the marginal health depreciation due to job. But, workers in area 4 will choose to retire earlier as they have a relative higher preference for future while those in area 6 are uncertain as they have a very low estate value.

Figure 4.1: Conditions for early retirement

4.2 Data and descriptive statistics

In this section, we firstly give a description of the dataset we use for the analysis herein : the Survey on Health, Ageing and Retirement in Europe (SHARE). We then briefly present health and working conditions indexes construction methods before characterizing workers with respect to their health condition, their job characteristics and their social and financial situation.

4.2.1 Data set

The data set we use in this paper is an appended dataset of waves of the SHARE data set. SHARE[3] is a longitudinal survey conducted each two years in European countries. It provides information on aged health condition, economic and social situation. For description, our interest variable is related to early retirement. It measures individual hopes[4] in terms of retirement. Some other important variables on health conditions, health care consumption, job characteristics, social variables, and financial condition (real and financial assets, and debts) are available.

Individual health condition is described by self assessed health in general, both observed physical and mental health illnesses, and health care consumption. Working environment is described by some variables related to working condition, revenue and pension.

As we are only interested on workers, we exclude from the dataset, all individuals that are non-workers as they first appear in the panel. We also exclude all individuals with only one observation period, as we are interested on the dynamic. After these cleaning up and the rest of data processing, we extract a subset of dataset that contains 17,568 individuals who are observed from 2 to 4 times (2.75 periods on average). Thus, the pooled dataset contains 44,331 observations.

[3]The SHARE data collection has been primarily funded by the European Commission through the 5th Framework Programme (project QLK6-CT-2001-00360 in the thematic programme Quality of Life), through the 6th Framework Programme (projects SHARE-I3, RII-CT-2006-062193, COMPARE, CIT5- CT-2005-028857, and SHARELIFE,CIT4-CT-2006-028812) and through the 7th Framework Programme (SHARE-PREP, N° 211909, SHARE-LEAP, N° 227822 and SHARE M4, N° 261982). Additional funding from the U.S. National Institute on Aging (U01 AG09740-13S2, P01 AG005842, P01 AG08291, P30 AG12815, R21 AG025169, Y1-AG-4553-01, IAG BSR06-11 and OGHA 04-064) and the German Ministry of Education and Research as well as from various national sources is gratefully acknowledged (see www.share-project.org for a full list of funding institutions)

[4]the question asked is : *... look for early retirement in main job ?*

4.2.2 Health stock and working condition indexes estimation

Both health and working condition are described by a set of categorical variables in SHARE dataset. Each of these categorical variable describes a specific dimension of the aggregate. For an individual characterization we need to aggregate all dimensions of the concept to get a continue variable. This aggregation will lead to a composite index that describes the situation of each individual related to the considered concept. The main problem that has to be challenged is that of the weighting set we use to aggregate the dimensions.

The index construction based on the categorical variables that describe each of the concepts above can be down by the use of the multiple correspondence analysis (MCA) method[5]. But in our case, we use the ordinary probit model. We choose this approach because we have a self-reported global health condition and a self-reported global work satisfaction that are respectively variables of 5 levels scale (ordered from 5 (excellent) to 1 (poor)) and 4 levels scale (ordered from 4 (strongly agree) to 1 (strongly disagree)). These variables have been used to create a continue health and working condition variables. The approach we use consist in estimating an ordered probit model on the overall dataset (see Cutler and Richardson (1997) for further details on this approach). Let h^* and y^* denote respectively a latent variable that measures health and working condition, X_1 denotes a set of demographic variables such as age, gender and country dummies, X_2 denotes a set of demographic variables such as age, gender, school grade, and health condition, D denotes diseases that have been observed by a doctor[6], M denotes mental health condition variables[7], and C denotes the working condition variables[8]. The estimated models are :

$$(1.2.1)\begin{cases} h^* = \beta_d D + \beta_m M + \beta_1 X_1 + \epsilon^1 \\ h = j \;\; if \;\; c^1_{j-1} \leq h^* < c^1_j \;\; for \;\; j = 1, ..., 5 \;\; with \; c^1_0 = -\infty \, and \, c^1_5 = +\infty \end{cases}$$

$$(1.2.2)\begin{cases} y^* = \beta_c C + \beta_2 X_2 + \epsilon^2 \\ y = j \;\; if \;\; c^2_{j-1} \leq y^* < c^2_j \;\; for \;\; j = 1, ..., 4 \;\; with \; c^2_0 = -\infty \, and \, c^2_4 = +\infty \end{cases}$$

The results of the estimated ordered probit models are in appendix 3 in Table 4.11 for health condition and Table 4.12 for working condition. For each of these two continue variables that values range from -9.5 to -0.28

[5]see Volle (1997) and Bry (1999) for further details on factorial analysis framework.

[6]For physical health, the available variables are answers to the question *Doctor told you had :*, and the items are : heart attack, hypertension, cholesterol, stroke, diabetes, lung disease, asthma, arthritis, osteoporosis, cancer, ulcer, Parkinson disease, cataract, hip or femoral fracture.

[7]For mental health, we use some self-reported variables that are : being sad or depressed, hopes for future, felt would rather be dead, trouble sleeping, less interest in things, irritability, lost of appetite, fatigue, concentration on reading and entertainment, enjoyment, tearfulness.

[8]Working condition is described by 9 variables related to : job physically demanding, time pressure due to heavy workload, little freedom to decide how work is down, opportunity to develop skills, receiving support in difficult situation, receiving recognition for work, adequate earnings or salary, poor security, and poor prospects for job advancement.

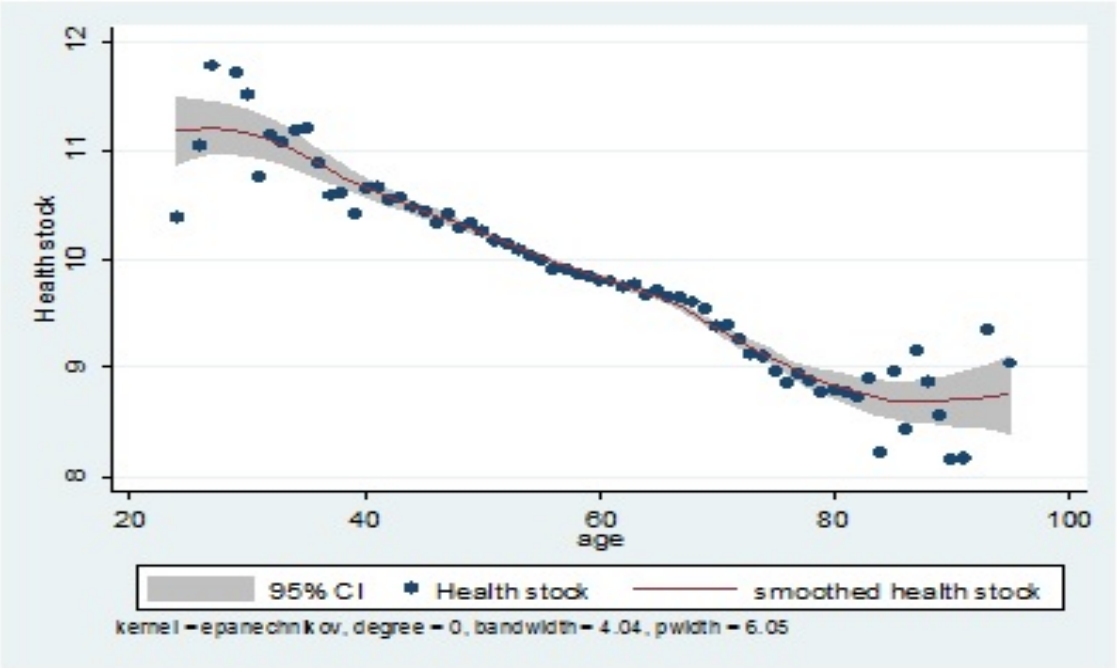


Figure 4.2: Evolution of health stock among age

for health stock and from -2.2 to 2.33 for working condition, we add a scalar to allow them to be positive. For health condition, we estimate the score for an individual that have all of physical and mental diseases we used as control in our ordered probit. The scalar we get is -12.2 and we add it to the health stock to get it positive. We did the same thing for working condition and the scalar we got is -2.84. The health stock ranges from 2.7 to 11.92 while working condition index ranges from 0.64 to 5.17. Note that higher values of health stock denote healthier individual and higher values of working condition index denote that worker is in better working conditions.

4.2.3 Analysis of health stock and job satisfaction

This subsection aims to give a description of the health stock. We highlight the differences between workers and retired people in health stock by testing for equality between these two groups. We also analyze the health stock differences between workers looking for early retirement and those who do not. Health stock is 9.3 at average in the pooled dataset[9]. However, it declines slowly from 9.6 at the first wave to 9.16 at the last wave. Health stock is also significantly lower for retired (9.02 vs 9.95) than workers. We also find a significant difference in health stock between workers looking for early retirement (9.86 vs 10.01 in average on pooled dataset) than those who do not. However, as we can see from Figure 4.2 the estimated health stock fulfils the common finding that is health stock declines with age.

Figure 4.4 in appendix 5 shows the dynamic of job satisfaction among age. It appears that elders are most

[9]see Table 4.13 in appendix 4 for full statistics on health stock among waves and on overall dataset

satisfied than younger workers. Job satisfaction index registers a 33% growth from 50 to 75 years old. It denotes that workers are more likely to remain on job market if they are satisfied from their main job.

4.2.4 Some determinants of early retirement in Europe

This subsection aims to compare workers that look for early retirement with those who do not. The comparison is made on some key socio economical variables that might have influence on the early retirement according to the literature. We perform a mean t-test (or proportion z-test when we analyze proportion) to confirm or not the equality between groups of workers looking for early retirement and those who do not.

As we can see from the Table 4.1, workers that are looking for early retirement are more likely to retire at the following period. At each wave of the dataset, we find a significant difference between the group of workers that look for early retirement and those who do not. Globally, this difference is around 5%.

Table 4.1: Retirement among European aged workers

Retired	Early^{+}	Wave 2	Wave 4	Wave 5	Overall
Proportion of	Yes	0.1722 (0.0074)	0.348 (0.0098)	0.1959 (0.0059)	0.2268 (0.0043)
workers that	No	0.1342 (0.0058)	0.2676 (0.0078)	0.1579 (0.0043)	0.1774 (0.0033)
retire	Difference	0.038*** (0.0092)	0.0804*** (0.0125)	0.038*** (0.0072)	0.0494*** (0.0053)

$^{+}$: worker looked for early retirement at last wave.

*** : significant at 1% level. Standard errors are in parenthesis.

The proportion of workers looking for early retirement slightly declines among time. From 43.3% in the first wave (2004), this proportion reaches 39.9% in the last wave of the survey (2013). As we can see from Table 4.14 in appendix 4, workers looking for early retirement are significantly younger (1.2 year lower) than those who do not. Another relevant item is the fact who workers that are afraid that their health limits their ability to work are also significantly more likely to look for early retirement (38% on pooled dataset) that those who are not (18% on pooled dataset). It also clearly appears that workers looking for early retirement are less satisfied of their job (around 6% lower in pooled dataset) than those who do not and are in worst working condition than those who do not (working condition index is 0.38 point lower in pooled dataset). However, contrarily to the common understanding, workers with high school level (undergraduate and graduate studies) are less likely to look for early retirement than not (around 10% lower).

From the financial condition, contrarily to the common understanding, workers looking for early retirement are not in better situation that those who do not (see Table 4.15 in appendix 4 for further details). Their annual earnings from job is in average significantly lower than that of workers who do not look for early retirement

(difference of 2,714.66€). The average amount on their bank account is also significantly lower than that of workers who do not look for early retirement (difference of 7,207.03€). These evidences mean that workers with low earnings and savings prefer preserving their health instead of continuing working. But the more consistent variable that workers account for is the percentage of salary that worker will receive as pension if he retires. We find that workers looking for early retirement are those with higher (39.25% vs 32.83%) percentage of salary to be received as pension. When we turn to out-of-pocket health expenditures, we highlight from the pooled dataset that there is a significant difference (26.45€)between workers who do not look for early retirement and those who do.

Post retirement opportunities are determinant for retirement (Brothers, 2000; Dorn and Sousa-Poza, 2004; Alhawarin, 2014). Proportion of retired who continue working among elders workers still growing (from 3.2% in the first wave of SHARE to 10.1% in the last wave). In the last wave of the survey[10], countries that are most concerned are Estonia (28.4%), Israel (21.3%), France (12.8%), and Switzerland (8.8%). Female workers are significantly less concerned than male in countries such as Germany (2.9% vs 5.4%), Sweden (1.1% vs 4.7%), Italy (1.7% vs 4.7%), Denmark (2.6% vs 5.6%), Greece (1.4% vs 4.7%), Switzerland (5.9% vs 8.8%), Belgium (1% vs 1.7%), and Israel (9.4% vs 12.7%). While in countries such as Estonia (29.3% vs 24.8%), Poland (3.1% vs 0.8%), and Austria (8.8% vs 2.1%), female workers are significantly most concerned than male. Retired workers proportion in older workers population is significantly higher among workers with undergraduate level (8.6%) or graduate level (7.3%) than workers with college degree (6.0%) or workers with no grade (7.1%). Female workers are significantly less concerned among workers with graduate level (6.4% vs 8.3%) and most concerned among workers with undergraduate level (9.6% vs 7.3%). For workers with college degree or no grade, there are no significant differences between male and female.

4.3 Empirical models and results

The proposed model has numerous parameters that have to be estimated. The main parameters we will discuss are the share of health depreciation due to working condition, the health consumption function for a worker and the health production function. Before estimating these two functions, we must firstly perform an estimation of $I_{i,t}$ and $\delta_{i,t}$ from the Grossman model.

4.3.1 Empirical estimation of Grossman's model

Based on the Grossman original model given by equation 4.1, many empirical works have proposed reformulations for empirical estimation of reduced forms of demand for health and demand for health care equations. Wagstaff (2002) underlines consistency problems with these empirical works. These empirical estimations lead to wrong signs of estimated coefficients (that are not consistent with the predicted signs in the Grossman's theory) that

[10]See Tables 4.16 and 4.17 in appendix 5 for further statistics on post retirement employment.

are due to inappropriate assumptions when moving from theoretical to empirical model[11]. To overcome this problem, Wagstaff (2002) assumes the desired health stock to be $H_t^* = \beta X_t + u_t$ where X_t is a set of exogenous variables. Wagstaff also assumes that individuals are not able to adjust instantaneously the health stock. Then he includes a fraction μ (between 0 and 1) that denotes the instantaneous adjustment rate of the desired health stock.

In our case, we construct an individual health stock index at each time period. Thus, only health investment I_t is unobserved. As δ is assumed to be individual and time variant, the equation to estimate is the following one:

$$H_{i,t} = I_{i,t-1} + (1 - \delta_{i,t-1})H_{i,t-1} + \xi_{i,t} \tag{4.15}$$

Where and $\xi_{i,t}$ denotes error terms, that are assumed to have a random effects model structure $\left(\xi_{i,t} = \xi_i^1 + \xi_{i,t}^2\right.$ with the individual effects $\xi_i^1 \sim \mathrm{N}(0,\sigma_1^2)$, the idiosyncratic error $\xi_{i,t}^2 \sim \mathrm{N}(0,\sigma_2^2)$ and ξ_i^1 supposed to be independent of $\left.\xi_{i,t}^2\right)$. This is a dynamic model with hidden factors $I_{i,t-1}$ and $\delta_{i,t-1}$. As coefficients $I_{i,t-1}$ and $\delta_{i,t-1}$ to be estimated are time variant and individual specifics, the model can not be estimated by a least square regression.

Thus, we use the space-state models framework that is helpful for that purpose (Peyrache and Rambaldi, 2012). From the specification in equation 4.15 above, two hidden states equations have to be defined: for the health investment $I_{i,t-1}$ and the health depreciation rate $\delta_{i,t-1}$. To achieve this goal, we make two assumptions. The first one is related to health investment. Individual health investment is assumed to have the following specification:

$$I_{i,t} = a_{1,0} + a_{1,1}I_{i,t-1} + a_{1,2}\delta_{i,t-1} + \xi_{i,t}^I \tag{4.16}$$

Where the error terms $\xi_{i,t}^I$ are assumed to be randomly distributed $\xi_{i,t}^I \sim \mathrm{N}(0,\sigma_I^2)$. It means that individual health investment accounts for the last period health depreciation rate, the last period health investment and the last period health adjustment coefficient. The second assumption is related to health depreciation rate that is assumed to have the following specification:

$$\delta_{i,t} = a_{2,0} + a_{2,1}I_{i,t-1} + a_{2,2}\delta_{i,t-1} + \xi_{i,t}^\delta \tag{4.17}$$

Where the error terms $\xi_{i,t}^\delta$ are assumed to be randomly distributed $\xi_i^{\delta_1} \sim \mathrm{N}(0,\sigma_\delta^2)$. By putting together the measurement equation in 4.15 and the two states equations 4.16 and 4.17, the overall state-space model to be estimated has the following form:

$$(3.1)\begin{cases} H_{i,t} &= I_{i,t-1} + (1 - \delta_{i,t-1})H_{i,t-1} + \xi_{i,t}\,, \forall t \geqslant 1 \\ I_{i,t-1} &= a_{1,0} + a_{1,1}I_{i,t-2} + a_{1,2}\delta_{i,t-2} + \xi_{i,t-1}^I\,, \forall t \geqslant 2 \\ \delta_{i,t-1} &= a_{2,0} + a_{2,1}I_{i,t-2} + a_{2,2}\delta_{i,t-2} + \xi_{i,t-1}^\delta\,, \forall t \geqslant 2 \end{cases}$$

The matrix state-space representation for the system 4.3.1 above is the following one :

$$\begin{aligned} H_{i,t} &= H_{i,t-1} + B_{i,t-1}\Gamma_{i,t-1} + \xi_{i,t}\,, \forall t \geqslant 1 \\ \Gamma_{i,t-1} &= A_0 + A_1\Gamma_{i,t-2} + \Xi_{i,t-1}\,, \forall t \geqslant 2 \end{aligned} \tag{4.18}$$

[11] See Wagstaff (2002) for further discussions on consistency problems with empirical reformulation of Grossman's model

With $A_0 = \begin{pmatrix} a_{1,0} \\ a_{2,0} \end{pmatrix}$, $A_1 = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}$, $\Gamma_{i,t-1} = \begin{pmatrix} I_{i,t-1} \\ \delta_{i,t-1} \end{pmatrix}$, $\Xi_{i,t-1} = \begin{pmatrix} \xi^I_{i,t-1} \\ \xi^\delta_{i,t-1} \end{pmatrix}$, and the transpose of B, $B'_{i,t-1} = \begin{pmatrix} 1 \\ -H_{i,t-1} \end{pmatrix}$.

Ξ and ξ are supposed to be uncorrelated (i.e the model is causal and invertible), and the covariance matrix structure for the errors vector Ξ in state equation is defined by :

$$\Sigma_\Xi = \begin{pmatrix} \sigma_I^2 & \rho_{I,\delta}\sigma_I\sigma_\delta \\ \rho_{I,\delta}\sigma_I\sigma_\delta & \sigma_\delta^2 \end{pmatrix}$$

To estimate the state-space model in Equation 4.18, we use a Kalman Filter algorithm[12] to provide value of state variables (I and δ). For initialization of the Kalman filter, we use :

$$\hat{\Gamma}_{i,1/1} = E(\Gamma_{i,1}/H_{i,1}) = m_\Gamma$$
$$\Sigma_{i,1/1} = V(\Gamma_{i,1}/H_{i,1}) = P_\Gamma$$

With $m_\Gamma = E(\Gamma_{i,1})$ and $P_\Gamma = V(\Gamma_{i,1})$, that are parameters for initial states. Droesbeke et al (2013) argue that m_Γ can be any real value vector and $P_\Gamma = \lambda I$ with the scalar λ very large and I the identity matrix. This approach that consists to set a large λ can be inappropriate (De Jong 1988, 1991a, 1991b). Thus, De Jong (1991a) proposes a diffuse Kalman filter or to model the state space model as diffuse (De Jong, 1991b) and some algorithms to solve the model. These specifications allow to estimate the model without setting P_Γ. Even if we suppose the model not to be diffuse, the approach by De Jong improves the Kalman filter by including a recursion.

The individual level log-likelihood function can be rewritten as follows (further details in appendix 1) :

$$LL_i = \frac{1}{2}\Bigg(- log(1 + \sigma_1^2 \sum_{t=1}^{T} M_{i,t/t-1}^{-1}) - \sum_{t=1}^{T}\Big[M_{i,t/t-1}^{-1} h_{i,t}^2 + log(2\,\pi M_{i,t/t-1})\Big] + \frac{\sigma_1^2\Big(\sum\limits_{t=1}^{T} M_{i,t/t-1}^{-1}\, h_{i,t}\Big)^2}{(1 + \sigma_1^2 \sum\limits_{t=1}^{T} M_{i,t/t-1}^{-1})}\Bigg) \quad (4.19)$$

Where $h_{i,t} = H_{i,t} - \hat{H}_{i,t/t-1} + \xi_i^1 = H_{i,t} - B_{i,t-1}\hat{\Gamma}_{i,t-1/t-1}$, , and $det(M_{t/t-1})$ denotes the determinant of matrix $M_{t/t-1}$ that is a scalar ($det(M_{i,t/t-1}) = M_{i,t/t-1}$) as we deal with one measurement equation. For likelihood calculation, we use $\hat{H}_{i,t/t-1}$ and $M_{i,t/t-1}$ provided by the Kalman filter. The parameters of the model described in equation 4.18 that have to be estimated are scalars σ_1 and σ_2, and matrices A_0, A_1, and Σ_Ξ. The maximization algorithm has two major steps that are iterated until convergence :

- for a fixed value of model's parameters, use the Kalman filter to estimate $\hat{H}_{i,t/t-1}$ and $M_{i,t/t-1}$, then compute the log-likelihood LL_i
- improve the model parameters to maximize the log-likelihood LL_i

Estimates results are in Table 4.2. We find a significant bidirectional causal link between investment and depreciation. Health stock depreciation causes negatively health investment and health investment causes positively health stock depreciation. The higher health stock depreciation is, the lower the health investment will be. This finding denotes that the older are less likely to demand for health care when they health stock depreciation rate

[12]Further details on the Kalman filter derivation are given in appendix 2

is high. Conversely, the higher the health investment is, the higher the health stock depreciation rate will be. An increase in demand for health care for older augurs an increase in health depreciation rate.

Table 4.2: Estimated coefficients of the state space model

Variable	Measurement equation $H_{i,t}$	Investment equation	Depreciation equation
	$I_{i,t-1} + (1 - \delta_{i,t-1})H_{i,t-1} + \xi_{i,t}$	$I_{i,t}$	$\delta_{i,t}$
$I_{i,t-1}$	–	-0.8293^{***} (0.0352)	1.2696^{***} (0.1412)
$\delta_{i,t-1}$	–	-0.0448^{**} (0.0212)	0.0688 (0.0730)
Intercept	–	-0.00004 (0.0579)	-0.00004 (0.0567)
Variance covariance structure			
σ_1	0	–	–
σ_2	0.00096	–	–
σ_I	–	5.4931^{***} (0.1303)	–
σ_δ	–	–	5.1494^{***} (0.1281)
$\rho_{I,\delta}$	–	0.9999^{***} (0.1082)	

*** : significant at 1% level, ** : significant at 5% level, Standard errors are in parenthesis.

4.3.2 Health production and health consumption function

In this section, we estimate an health consumption and health production functions. These functions are key functions in the retirement process as they determine the individual's discount factor and it's probability to retire earlier.

As we described in section 2, individual can invest in health and this investment can be interpreted as input for an health production function. Let assume the health production function to be $\hat{I}_{i,t} = f(Exph_{i,t}) = A_{0,i}Exph_{i,t}^{\varphi}$ where $\hat{I}_{i,t}$ denotes the produced health by an invested health expenditures $Exph_{i,t}$, and $A_{0,i}$ denotes individual and country specifics variables that are determinant of health investment. These controls are individual characteristics (age, gender, school grade, marital status), individual behaviour (smoking, drinking daily more than 2 glasses of alcohol, visiting doctors), individual health shocks that are captured by the fact of being patient, the frequency of being patient and the nights stayed in hospital. We also include country dummies that are supposed to capture technological differences between countries in terms of medical improvements.In our data set, health expenditures account for expenditures for inpatient care, outpatient care, prescribed drugs, nursing

home, day-care and home-care. Thus the log-linear form we estimate is :

$$\hat{I}_{i,t} = \varphi log(Exph_{i,t}) + c_0 + c_1 Age_{i,t} + c_2 Male_i + c_3 Grade_i + c_4 Couple_{i,t} + c_5 Smoke_{i,t} \quad (4.20)$$
$$+ c_6 Drink_{i,t} + c_7 doctor_{i,t} + c_8 Patient_{i,t} + c_9 Tpatient_{i,t} + c_{10} Hnights + c_{11} Country_i + \epsilon_{i,t}$$

Estimates results are in Table 4.3. Most of our estimated coefficients have the expected sign. The significant coefficients for country dummies involve medical technological differences across European countries. We find a positive, less than one and significant elasticity of health expenditures. Individual characteristics also affect the health production function. Being in couple reduces significantly health investment. We also find a weak evidence (significant at 10% level) that male invests more than female. But, contrarily to Wagstaff (2002), we find that ageing has a positive and significant effect on health investment. This denotes that older invest more than younger in health care. A part from individual with no school grade, the higher the school grade is, the lower the demand for health care is. However, for individual with no school grade, we find a weak evidence that the demand for health care is higher than that of individual with graduate studies level. Contrarily to Grossman (1999) and Wagstaff (2002), our specification gives an effect of education that is consistent with the original health investment model (Grossman, 1972). Turning to behavioural variables, we can see that drinking alcohol has no significant effect on health investment. However, smoking increases the demand for health care. Seeing doctor or being patient increase the health investment but the frequency of being patient and the total nights in hospital decrease the health investment.

For the health consumption function, we assume that each worker uses a share of is health stock as input to earn a wage at the end of an production process. As earnings are subject to the number of hours worked and in our dataset, more than 91% of worker have a permanent job, it is unnecessary to account for the number of hours worked as input of the wage equation. Thus we assume wage to be function with health as input $\bar{H}_{i,t} = g\Big((1-\omega)W_{i,t}\Big) = A_{1,i}\Big((1-\omega)W_{i,t}\Big)^{\theta}$ where $\bar{H}_{i,t}$ denotes the health depreciation due to working condition (the health stock used to earn a wage that is the difference between the wage earned and the retirement pension), and $A_{1,i}$ includes controls such as industry sector, country dummy, age, gender and school grade that are discriminating factor of wage. The health depreciation due to working condition $\bar{H}_{i,t}$ is estimated from the Equation 4.2 which econometric form is :

$$\hat{\delta}_{i,t} = \alpha_w Work_{i,t} + \alpha_c C_{i,t} * Work_{i,t} + \alpha_3 Age_{i,t} + \alpha_4 H_{i,t} + \alpha_5 Country + u^{\delta}_{i,t} \quad (4.21)$$

where $u^{\delta}_{i,t}$ are error terms, $Work_{i,t}$ is a dummy that is one if individual i works at t, and $\hat{\delta}_{i,t}$ is provided by the estimated state-space model. Thus, the effect of job condition on health depreciation is given by $\alpha_{1,i,t} = \alpha_w + \alpha_c C_{i,t}$. We include country dummies to account for country heterogeneity in terms of working condition. Due to the fact that the dataset covers only aged worker, the age square effect is not significant. Thus, we exclude the square of age in the estimated model.

Estimates results are in Table 4.4. Estimated coefficients have the expected signs. We find a strong evidence that health depreciation is higher for workers and that the better the working condition is, the lower the health depreciation is. That is a strong result and it is consistent with previous literature (Debrand and Blanchet, 2008).

We also find a positive and significant effect of ageing on health depreciation. Included country dummies have the expected signs and significant coefficients. The higher the health stock is, the higher the health depreciation is. This finding denotes that a health shock has an higher effect on healthier.

Thus the log-linear form that will be estimated the following one :

$$\bar{H}_{i,t} = \theta log\big[(1-\omega)W_{i,t}\big] + b_0 + b_1 Age_{i,t} + b_2 Male_i + b_3 Grade_i + b_4 Country_i + \xi^w_{i,t} \tag{4.22}$$

The proportion of missing data for the percentage of salary to receive as pension is large (around 39%). Neglecting workers for who this variable is missing will considerably drop our estimate sample as this variable is important for the early retirement probabilities computation. Thus, we complete a multiple imputation technique to predict these missing value. For the sample on which the percentage of salary to receive as pension is observed, we estimate a model that explain the later variable with observed individual characteristics such as school grade, salary and country dummies that account for pension regulation across countries. This approach is relevant because the observed variability in the percentage of salary received as pension is due to individual heterogeneity (variance between = 91.2% of the overall variance). The results of the estimated model for imputation are in Table 4.18 in appendix 6. After imputation, the individual heterogeneity is 90.5% of the overall variance and the average percentage of salary as pension is 56.06% versus 55.26% before imputation.

The estimation results of health depreciation function are in Table 4.5. The estimated coefficients have the expected signs. The health depreciation due to working condition is higher for male and the lower the pension share and the school level are, the higher the depreciation due to working condition is. This denotes that as workers with lower school level have job with low security, high physical pressure and low working condition, then the effect of job on their health is higher.

4.3.3 Estimation of utility functions parameters

The utility functions parameters estimation is based on the approach used by Hall and Jones (2007). We specify two utility functions[13] : a separable utility function (with three parameters γ, λ, and a), and a non separable utility function (with two parameters ν and η). The estimation of these parameters is done by following three steps :

- Estimate QALY for all available diseases in each country for cohorts of individual. In our paper, we use four cohorts of individuals in each country : individuals aged 50 to 59 years old, 60 to 64 years old, 65 to 74 years old, and over 75 years old. This grouping is used because SHARE data collection focuses on aged people and it allows to have enough individual by group and country to estimate the model. In each ordered probit regression, we include reported diseases and some demographic characteristics such as age, sex, the square of age and cross terms between age and sex, square of age and sex.

[13]See equations 4.9 for further details

- For each estimation done in the previous step, by country, we keep only QALY that are significant in each cohort and are decreasing by age. We also estimate with the SHARE dataset, the average of estate by country and age group. Then we use data from mortality table (from EuroStat database and Israel national statistics bureau) to compute the age-specific state of health by age group.
- The last step consists in solving the following equations for the separable and non separable utility function cases.

$$\frac{u(H_{50-59}, E_{50-59})}{Q_{50-59}} = \frac{u(H_{60-64}, E_{60-64})}{Q_{60-64}} = \frac{u(H_{65-74}, E_{65-74})}{Q_{65-74}} = \frac{u(H_{75+}, E_{75+})}{Q_{75+}}$$

The estimated parameters are provided by country in Table 4.6.

4.3.4 Estimation of retirement probabilities

In this section, we present our calculated discount factor and early retirement probabilities computed using a separable utility function. The computation of $r_{i,t}$ and $p_{i,t}$ at the date t involves the use of the health stock at $t+1$ and the dynamic of estate from $t+1$ until worker dies[14]. Health stock at $t+1$ is known with the individual characteristics at t. For estate equation, earnings after retirement are given by a share of annual salary received as pension. The remaining estate value after expenditures at t is supposed to be appreciated at the interest rate in the country. For the expenditures level after retirement, many papers address these issues. We use the results of Fisher et al (2005) who find in their research that consumption expenditures decline by 2.5% at the retirement and by 1% per year after retirement.

Statistics on preferences for future are provided in Table 4.7. The preference for future across elder workers in Europe is highly volatile. On average, 66.56% of aged workers in Europe have a low preference for future (on average, $r_{i,t} = 0.16755$ with a standard deviation of 0.23963). However it exists high volatility across countries. In countries such as Germany, France, Netherlands, Spain, Greece, Israel and Poland, at least 70% of aged workers have a low preference for future while in countries such Estonia, Switzerland and Slovenia, less than 50% of aged workers have a low preference for future. Individual preference for future can also be higher than one (for 33.44% of the estimation sample). These individuals are characterized by a relatively high health stock but a very low estate value. Thus, for them, as they expect a long and healthier time to live, their hope on future is high because, in addition to the pension they will have at retire, they can continue working to earn additional income that will increase their estate and then, their utility.

Early retirement probabilities are analyzed according to the different areas in Figure 4.1. From descriptive statistics presented in Table 4.8, we can see that worker who will not choose early retirement (areas 1 and 2 described in Section 2.4) are 44.97% of workers. Workers in area 1 (17.67% of workers) have higher preference for future (on average $r_{i,t} = 266.35$), but a lower estate value (on average 196,003 €) and a lower health stock (on average, their health stock is 9.7919) than worker in area 2 (27.3% of workers). Workers who will choose early retirement (areas 3 and 4 described in Section 2.4) represent 20.07% of workers. Those of the latter in

[14]The individual survival is assumed to be the life expectancy at his age in his country.

area 3 (2.42% of workers) are characterized by a higher preference for future (on average,$r_{i,t} = 11.27$), a higher estate value (on average 319,991.6 €) but a lower health stock level (on average, their health stock is 9.938) than workers in area 4. The last group of workers is that of workers with uncertainty on early retirement (areas 5 and 6 described in Section 2.4). This group represent 34.96% of workers. Those of this group that are in area 5 are characterized by a higher early retirement probability (on average 0.7912 with a standard error of 0.0022) than those in area 6 (on average 0.0893 with a standard error of 0.0015). Workers in area 5 have lower preference for future (on average $r_{i,t} = 0.1296$) and lower health stock (on average their health stock is 10.0466), but higher estate value (on average 536,234.9 €) than those in area 6. Across groups, we can see that workers who will choose early retirement are significantly in better health condition but in worse financial condition than those who are uncertain about their early retirement decision. The latter are also significantly in better health condition but in worse financial condition than those who will not choose early retirement. The same analysis can be made about the preference for future. Workers who will choose early retirement have significantly less preference for future than those who are uncertain and the latter have significantly less preference for future than those who will not choose early retirement.

4.3.5 Robustness check and causality analysis

The robustness check for our model will consist in showing the ability of our model to distinguish between two groups of individual across workers with uncertainty about their early retirement decision. For that purpose, we plot the density of the estimated early retirement probability for workers with uncertainty about early retirement. The plotted density is shown in Figure 4.3. As we can see, the density has two peaks, the first one around the early retirement probability of 0.07 that is the higher and second one around the early retirement probability of 0.95 that is the lower peak. In addition, early retirement probability values between 0.1 and 0.85 have lower and closer to zero densities. This plot gives a first level validation of the accuracy of our model.

The second robustness check test we do is to test whether the calculated early retirement probability causes the transition from work to early retirement. For that purpose, we estimate a dynamic probit model with the transition from work to early retirement as dependant variable. We estimate 3 models : the first one with only the last period early retirement probability as explanatory variable, the second one by including school grade that account for the post retirement opportunities, and the third one by including country dummies and school grade to account for country heterogeneities in terms of retirement policy and in terms of job opportunities after retirement. The estimated model is given by :

$$\begin{cases} ER^*_{i,t} = d_0 + d_1 p_{i,t-1} + d_2 grade_i + d_3 country_i + \epsilon_{i,t} \\ ER_{i,t} = 1 \;\; if \;\; ER^*_{i,t} > 0 \end{cases}$$

Where $ER_{i,t}$ is 1 if individual i retire earlier at t. Thus, the calculated early retirement probability is causal for early retirement if d_1 is significantly different from zero. Estimate results are in Table 4.9 for the model with the lag of the calculated early retirement status and in Table 4.10 for the model with the lag calculated early retirement probability. As we can see, in all estimated models, the calculated early retirement probability and

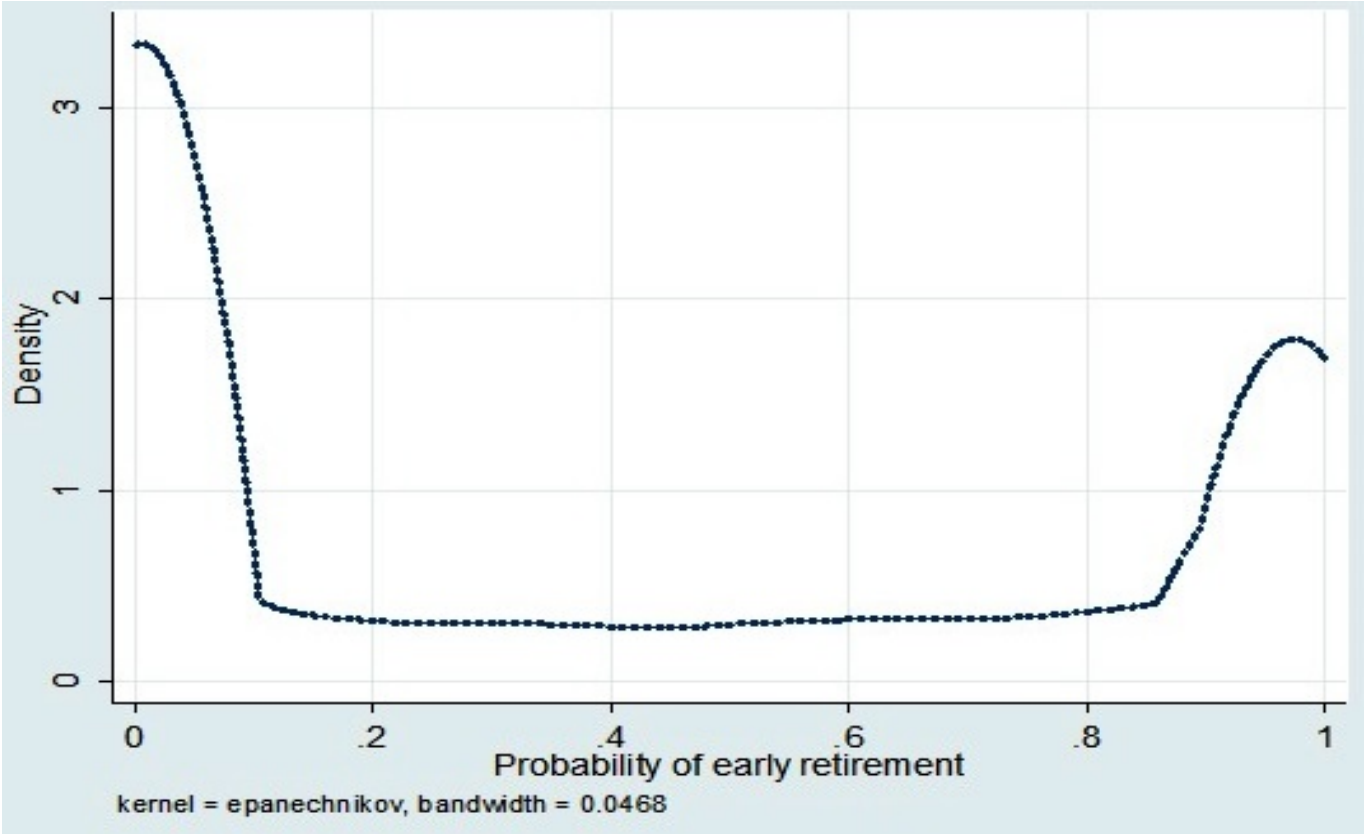


Figure 4.3: Density of early retirement probability

the calculated early retirement status causes the transition from work to early retirement. These causal links remain significant when we control for school level and country heterogeneity. These findings strengthen the reliability of our approach and our computed early retirement probabilities.

Conclusion

This paper analyzes the early retirement decision-making process among older workers in Europe. Several previous papers focus on this issue by analyzing the health effects, financial effects, or both. Some papers highlight the important roles of worker's environment and institutional regulations. However, these previous papers use a binary outcome model to assess the effects of those key variables on the probability to retire earlier. In our approach, we specify a worker's utility function depending on his or her health, estate, institutional framework by the use of the share of salary as pension, and preference for future. This specification allows us to assess the early retirement decision-making process by accounting for, not only the current health condition and estate value, but for the whole discounted lifetime utility. Estimations are done with four waves of the SHARE dataset.

Our framework is innovative. We estimate health investment and health depreciation from the Grossman's model using a space-state approach and we use these estimations to estimate a health production and health consumption function that are key in early retirement decision-making process. Contrarily to previous literature on demand for health care equation, this approach lead to expected signs for all determinants.

From our model, we predict for each individual and at each period, the probability that workers retire early with regards to their financial, health and socioeconomics conditions. These early retirement probabilities are function of (i) the marginal productivity of health expenditures, (ii) the marginal health depreciation due to working condition, and (iii) the discounted future marginal utility of estate divided by the current marginal utility of health. We show that our approach is robust as it disentangles between three categories of workers : those who will not choose early retirement, those who will choose early retirement, and those who are uncertain about early retirement. We also show that our calculated early retirement probabilities are good predictor of observed individual early retirement. Finally, this framework allows us to investigate on the effects of public policies such as (i) predicting, by simulations, the probability of early retirement with respect to health, estate value, and pension share, and (ii) predicting how public health policy or retirement policies may affect retirement behaviour.

Appendices

Appendix 3 : Estimation of health stock and working condition index

Table 4.3: Health production function estimation

Variable	Coefficients	Variable	Coefficients
Log of health expenditures	0.000057^{***} (0.00002)	Country	ref : Estonia
Age	0.00011^{***} ($6.46e{-6}$)	Austria	-0.00008 (0.00018)
Male	0.00013^{*} (0.00007)	Germany	0.0024^{***} (0.0002)
Couple	-0.0003^{***} (0.000089)	Sweden	0.00135^{***} (0.00017)
Ever smoke	0.00016^{**} (0.00007)	Netherlands	0.00136^{***} (0.00018)
Drink alcohol	0.00012 (0.00009)	Spain	0.00168^{***} (0.0002)
Doctor	0.00006^{***} ($5.05e{-6}$)	Italy	0.00063^{***} (0.00019)
Be patient	0.00043^{**} (0.00018)	France	0.00033^{**} (0.00017)
Times being patient	-0.00042^{***} (0.00009)	Denmark	0.00088^{***} (0.00017)
Nights in hospital	-0.00002^{**} ($9.42e{-6}$)	Greece	-0.00001 (0.00023)
Grade	ref : graduate studies	Switzerland	0.00068^{***} (0.00017)
no grade	-0.00025^{*} (0.00013)	Belgium	0.00011 (0.00016)
college degree	0.00028^{***} (0.00008)	Israel	-0.00237^{***} (0.00018)
undergraduate studies	-0.00014 (0.00017)	Czech Republic	0.00135^{***} (0.00019)
Intercept	-0.00667^{***} (0.00042)	Slovenia	0.00029 (0.00029)
$\sigma_\mu = 0$, $\sigma_e = 0.0124$			

*** : significant at 1% level, ** : significant at 5% level, * : significant at 10% level, Standard errors are in parenthesis.

Table 4.4: Health depreciation explanatory factors

Variable	Coefficient	Variable	Coefficient
Work	0.0037^{***} (0.00064)	Country	Ref = Estonia
Work*Condition	-0.00097^{***} (0.00017)	Austria	-0.0076^{***} (0.00057)
Age	0.00009^{***} (0.00002)	Germany	-0.00847^{***} (0.00055)
Health stock	0.00557^{***} (0.00016)	Sweden	-0.01133^{***} (0.00053)
		Netherlands	-0.00951^{***} (0.00052)
		Spain	-0.00672^{***} (0.00055)
		Italy	-0.00662^{***} (0.00054)
		France	-0.00588^{***} (0.00048)
		Denmark	-0.01137^{***} (0.00054)
		Greece	-0.00957^{***} (0.00074)
		Switzerland	-0.01049^{***} (0.00054)
		Belgium	-0.0079^{***} (0.0005)
		Israel	-0.0033^{***} (0.00057)
		Czech Republic	-0.0052^{***} (0.00053)
		Poland	-0.00051^{***} (0.00094)
Intercept	-0.05446^{***} (0.00215)	Slovenia	-0.00455^{***} (0.00087)
$\sigma_\mu = 0$, $\sigma_e = 0.0138$			

*** : significant at 1% level, ** : significant at 5% level

Standard errors are in parenthesis.

Table 4.5: Health consumption function

Variable	Coefficient	Variable	Coefficient
Wage : $log\big[(1-\omega)W_{i,t}\big]$	0.00015^{***} $(3.89e{-}6)$	country	ref : Estonia
age	-0.00052^{***} $(2.6e{-}6)$	Austria	-0.00225^{***} (0.00008)
male	0.00091^{***} (0.00003)	Germany	0.00019^{**} (0.00008)
Grade	ref : graduate studies	Sweden	-0.00094^{***} (0.00007)
no grade	0.00288^{***} (0.00005)	Netherlands	-0.00168^{***} (0.00007)
college degree	0.00165^{***} (0.00003)	Spain	-0.00062^{***} (0.00008)
undergraduate studies	0.00274^{***} (0.00005)	Italy	0.00045^{***} (0.00008)
		France	-0.00045^{***} (0.00007)
		Denmark	-0.00256^{***} (0.00007)
		Greece	0.00345^{***} (0.0001)
		Switzerland	-0.00274^{***} (0.00007)
		Belgium	-0.00135^{***} (0.00007)
		Israel	0.0004^{***} (0.00009)
		Czech Republic	-0.00032^{***} (0.00008)
		Slovenia	-0.00193^{***} (0.00012)
Intercept	0.0326^{***} (0.00017)	Poland	0.00261^{***} (0.00013)
$\sigma_\mu = 0.0229$, $\sigma_e = 0.0054$, $\rho = 0.1524$			

*** : significant at 1% level, ** : significant at 5% level, Standard errors are in parenthesis.

Table 4.6: Utility functions parameters (with share)

Country	Separable function			Non separable function	
	γ	λ	a	ν	η
Austria	1.105350	2.120083	0.000162	0.036219	1.066892
Germany	1.282327	2.167405	0.000266	0.196694	1.067830
Sweden	1.235924	2.165392	0.0002	0.221287	1.067209
Netherlands	1.341981	2.167014	0.000344	0.339188	1.069074
Spain	1.241212	2.314658	0.000179	0.252536	1.065737
Italy	1.240786	2.134855	0.000204	0.187802	1.058653
France	1.484212	2.114771	0.000327	0.402915	1.062368
Denmark	1.169949	2.082954	0.000190	0.252544	1.065339
Greece	1.254943	2.535495	0.000136	0.252195	1.070837
Switzerland	1.051630	1.965825	0.000201	0.249546	1.057752
Belgium	1.210802	2.182727	0.000069	0.261289	1.065980
Israel	1.201668	2.245111	0.000459	0.115228	1.056579
Czech Republic	1.151662	2.228266	0.000160	0.036270	1.072141
Poland	1.308289	2.449786	0.000462	0.076810	1.074665
Slovenia	1.029674	1.888861	0.000464	0.081811	1.070760
Estonia	1.023071	2.058305	0.000152	0.074303	1.071437

Table 4.7: Preference for future across European countries

Country	Population	Low ($r_{i,t} < 1$)		High ($r_{i,t} > 1$)	
	Proportion(in%)	%	*Average*	%	*Average*
Austria	4.61	2.76	0.21301 (0.25814)	1.85	464.62 (3273.347)
Germany	5.3	3.74	0.17314 (0.22755)	1.55	459.2 (3404.512)
Sweden	10.08	5.86	0.21836 (0.24735)	4.22	290.5 (2330.292)
Netherlands	8.18	7.07	0.09752 (0.19306)	1.11	699.77 (5019.927)
Spain	4.94	4.3	0.0996 (0.19075)	0.64	548.74 (3734.307)
Italy	5.08	3.02	0.21695 (0.27218)	2.06	478.17 (4380.884)
France	9.6	7.52	0.12266 (0.21442)	2.08	494.25 (3997.869)
Denmark	10.08	5.91	0.20999 (0.25577)	4.17	141.21 (1661.173)
Greece	5.54	5.13	0.06239 (0.14002)	0.41	344.96 (2921.652)
Switzerland	7.54	3.23	0.25173 (0.28794)	4.31	777.1 (5364.31)
Belgium	10.97	6.92	0.25032 (0.26481)	4.05	578.78 (4729.252)
Israel	6.37	5.63	0.0999 (0.18414)	0.74	233.66 (2226.619)
Czech Republic	4.66	2.87	0.23729 (0.26131)	1.79	301.56 (2377.2)
Poland	1.23	0.89	0.1345 (0.2069)	0.34	717.75 (4591.578)
Slovenia	1.22	0.61	0.20732 (0.2578)	0.61	132.84 (909.072)
Estonia	4.61	1.08	0.31709 (0.28705)	3.52	697.69 (4723.899)
Overall (Obs.: 20,782)	100	66.56	0.16755 (0.23963)	33.44	474.09 (3871.625)

Standard deviations are in parenthesis

Table 4.8: Characterization of early retirement across European countries

Groups	Areas	Prop.$^{+}$ (in %)	Mean $p_{i,t}$	Mean $r_{i,t}$	Mean Estate	Mean health
Not choose	Area 1	17.67	0	266.35 (17.6386)	196,003 (1,568.45)	9.7919 (0.0062)
early	Area 2	27.30	0	0.0489 (0.0005)	649,296.6 (5,639.91)	10.1936 (0.0035)
retirement	Difference	44.97	–	266.3*** (14.1842)	−453,293.7*** (7,124.15)	−0.4017*** (0.0066)
Choose	Area 3	2.42	1	11.2668 (1.9061)	319,991.6 (5,979.89)	9.938 (0.0149)
early	Area 4	17.65	1	0.1544 (0.0013)	265,844.9 (1,700.57)	10.2554 (0.0046)
retirement	Difference	20.07	–	11.1124*** (0.7053)	54,146.72*** (5,100.38)	−0.3174*** (0.0136)
Uncertain	Area 5	15.08	0.7912 (0.0022)	0.1296 (0.0012)	536,234.9 (6,011.08)	10.0466 (0.0053)
early	Area 6	19.88	0.0893 (0.0015)	116.35 (9.2775)	164,875.2 (1,284.52)	10.122 (0.005)
retirement	Difference	34.96	0.7019*** (0.0026)	−116.22*** (10.644)	371,362.8*** (5,439.89)	−0.0755*** (0.0074)
Overall	Obs.: 20,782	100	–	–	380,209.9 (1,959.87)	9.2994 (1.0717)
Comparison across groups						
Choose early retire		20.07	1	1.4919 (0.2309)	272,366.9 (1,664.55)	10.2172 (0.0045)
Not choose early retirement		44.97	0	104.63 (6.9529)	471,223.2 (3,626.88)	10.0358 (0.0033)
Difference		–	–	−103.14*** (10.4074)	−198,856.3*** (5,541.59)	0.1814*** (0.0058)
Choose early retire		20.07	1	1.4929 (0.2309)	272,366.9 (1,664.55)	10.2172 (0.0045)
Uncertain early retirement		34.96	0.3922 (0.0022)	66.1737 (5.2808)	325,090.7 (2,861.81)	10.0895 (0.0037)
Difference		–	–	−64.6808*** (6.9698)	−52,723.85*** (3,982.18)	0.1277*** (0.0059)
Uncertain early retirement		34.96	0.3922 (0.0022)	66.1737 (5.2808)	325,090.7 (2,861.81)	10.0895 (0.0037)
Not choose early retirement		44.97	0	104.63 (6.9529)	471,223.2 (3,626.88)	10.0358 (0.0033)
Difference		–	–	−38.4548*** (9.1584)	−146,132.5*** (4,825.58)	0.0537*** (0.005)

*** : significant at 1% level. Standard errors are in parenthesis. $^{+}$: For proportions, the values in the difference cells denote the overall proportion for the 2 compared areas.

Table 4.9: Causal link between calculated early retirement status and retirement

Transition from work to early retirement	Model 1	Model 2	Model 3
Calculated early retirement status	Ref = Not choose early retirement at $t-1$		
Choose early retirement at $t-1$	0.1437^{***} (0.0158)	0.1412^{***} (0.0158)	0.1262^{***} (0.0165)
Uncertain early retirement at $t-1$	0.2154^{***} (0.0128)	0.212^{***} (0.0128)	0.1502^{***} (0.0136)
Grade	Ref = Graduate studies		
No grade	–	0.0915^{***} (0.019)	0.1622^{***} (0.02)
College degree	–	0.0689^{***} (0.0119)	0.0711^{***} (0.0123)
Undergraduate studies	–	0.0704^{***} (0.0178)	0.1255^{***} (0.0187)
Intercept	-1.3775^{***} (0.0062)	-1.427^{***} (0.01)	-2.0295^{***} (0.0291)
Country fix effects	NO	NO	YES

*** : significant at 1% level. Standard errors are in parenthesis.

Table 4.10: Causal link between calculated early retirement probability and early retirement

Transition from work to early retirement	Model 1	Model 2	Model 3
$p_{i,t-1}$	0.0899^{***} (0.0264)	0.089^{***} (0.0265)	0.0857^{***} (0.0308)
Grade	Ref = Graduate studies		
No grade	–	-0.0203 (0.0399)	0.0769^{*} (0.0418)
College degree	–	0.154^{***} (0.0277)	0.1673^{***} (0.0287)
Undergraduate studies	–	0.0088 (0.0385)	-0.0029 (0.0402)
Intercept	-1.1968^{***} (0.0172)	-1.2666^{***} (0.0255)	-1.0694^{***} (0.077)
Country fix effects	NO	NO	YES

*** : significant at 1% level; * : significant at 10% level. Standard errors are in parenthesis.

Table 4.11: Ordered probit estimates of health

Variables^{+}	Coefficients	Variables	Coefficients	Variables	Coefficients
Heart attack	-0.5448^{***} (0.0125)	Sad or depressed	-0.1651^{***} (0.0088)	Germany	-0.3083^{***} (0.0277)
Hypertension	-0.2924^{***} (0.0088)	No hopes for future	-0.2195^{***} (0.0106)	Sweden	0.4359^{***} (0.0254)
Cholesterol	-0.0884^{***} (0.0096)	Rather be dead	-0.2663^{***} (0.0157)	Netherlands	0.0566^{**} (0.0247)
Stroke	-0.6507^{***} (0.0209)	Trouble sleeping	-0.233^{***} (0.0086)	Spain	-0.3816^{***} (0.0244)
Diabetes	-0.5315^{***} (0.0134)	Less interest in things	-0.1589^{***} (0.0142)	Italy	-0.2341^{***} (0.0242)
Lung disease	-0.559^{***} (0.017)	Irritability	-0.0698^{***} (0.0089)	France	-0.0942^{***} (0.0221)
Arthritis	-0.435^{***} (0.0101)	Lost of appetite	-0.3035^{***} (0.0143)	Denmark	0.5837^{***} (0.026)
Osteoporosis	-0.3641^{***} (0.0126)	Fatigue	-0.4051^{***} (0.0085)	Greece	0.2173^{***} (0.0288)
Cancer	-0.5786^{***} (0.0183)	No conc. in entertainment	-0.1418^{***} (0.0131)	Switzerland	0.3581^{***} (0.0251)
Ulcer	-0.1946^{***} (0.018)	No conc. in reading	-0.1537^{***} (0.0126)	Belgium	0.1542^{***} (0.022)
Parkinson disease	-1.1373^{***} (0.0496)	No enjoyment	-0.1636^{***} (0.0114)	Israel	-0.0481^{*} (0.0278)
Cataract	-0.0354^{**} (0.0141)	Tearfulness	-0.0274^{***} (0.0096)	Czech Republic	-0.4814^{***} (0.0235)
Fracture	-0.3312^{***} (0.0279)	Cut 1	-5.1334^{***} (0.0463)	Poland	-0.9733^{***} (0.0344)
Age	-0.0275^{***} (0.0006)	Cut 2	-3.4868^{***} (0.0444)	Slovenia	-0.467^{***} (0.0306)
Male	-0.1688^{***} (0.0606)	Cut 3	-1.8771^{***} (0.0431)	Estonia	-1.0942^{***} (0.0231)
Age*Male	0.0019^{**} (0.0009)	Cut 4	-0.6280^{***} (0.0426)	Austria	Reference

$^{+}$: dependent variable is self-reported health evaluated on a 5-level scale : Excellent, Very good, . Good, Fair, and Poor. *** : significant at 1% level, ** : significant at 5% level, * : significant at 10% level. Standard errors are in parenthesis.

Table 4.12: Ordered probit estimates of Working condition

Variable[+]	Coefficients
Job physically demanding	−0.0227 (0.0205)
Time pressure/heavy workload	−0.1184*** (0.0196)
Little freedom to decide how to do the work	−0.2826*** (0.0222)
No opportunity to develop new skills	−0.4611*** (0.0223)
No receive support in difficult situation	−0.4159*** (0.0232)
No receive recognition for the work	−0.5804*** (0.0237)
Salary or earnings are not adequate	−0.2938*** (0.0207)
Poor job security	−0.2654*** (0.0228)
Poor prospects for job advancement	−0.2608*** (0.0207)
Age	0.0238*** (0.002)
Male	−0.0607*** (0.0209)
Undergraduate or graduated studies	0.0545** (0.023)
Very good health[++]	−0.2219*** (0.0293)
Good health[++]	−0.4179*** (0.0291)
Fair health[++]	−0.5359*** (0.0355)
Poor health[++]	−0.6466*** (0.0673)
Cut 1	−3.2466*** (0.1215)
Cut 2	−2.0563*** (0.1164)
Cut 3	0.3171*** (0.1142)

[+] Dependent variable is Job satisfaction evaluated on a 4-level scale Strongly agree, agree, disagree, strongly disagree. *** : significant at 1% level, ** : significant at 5% level, [++] reference in Excellent health. Standard errors are in parenthesis.

Appendix 4 : Descriptive statistics on health, financial situation and early retirement in Europe

Table 4.13: Health stock level

Variable	Modalities	Wave 1	Wave 2	Wave 4	Wave 5	Overall
Mean of health stock		9.599 (0.9484)	9.4627 (1.0292)	9.1703 (1.098)	9.1598 (1.0868)	9.2994 (1.0717)
Job	Retired	9.3158 (0.0089)	9.1792 (0.0082)	8.9038 (0.0068)	8.9254 (0.0066)	9.0207 (0.0038)
Status	Worker	10.1707 (0.0077)	10.1225 (0.0076)	9.7998 (0.0081)	9.8232 (0.0083)	9.9481 (0.0042)
	Difference	-0.8549^{***} (0.0118)	-0.9433^{***} (0.0112)	-0.896^{***} (0.0105)	-0.8977^{***} (0.0106)	-0.9274^{***} (0.0056)
Look for	Yes	10.090 (0.0123)	9.9852 (0.0127)	9.7269 (0.0122)	9.7329 (0.0129)	9.8606 (0.0065)
early	No	10.2361 (0.0098)	10.2231 (0.0092)	9.8512 (0.0108)	9.8837 (0.0108)	10.0113 (0.0055)
retirement	Difference	-0.1461^{***} (0.0157)	-0.2378^{***} (0.0157)	-0.1243^{***} (0.0163)	-0.1508^{***} (0.0169)	-0.1507^{***} (0.0085)
Afraid health	Yes	9.9168 (0.0169)	9.8367 (0.0172)	9.5209 (0.0168)	9.5241 (0.0183)	9.6698 (0.009)
limits ability	No	10.273 (0.0082)	10.2315 (0.0079)	9.9169 (0.009)	9.9276 (0.0091)	10.0564 (0.0045)
to work	Difference	-0.3563^{***} (0.0187)	-0.3948^{***} (0.0189)	-0.396^{***} (0.0081)	-0.4035^{***} (0.0205)	-0.3866^{***} (0.0101)

*** significant at 1% level. Standard errors are in parenthesis.

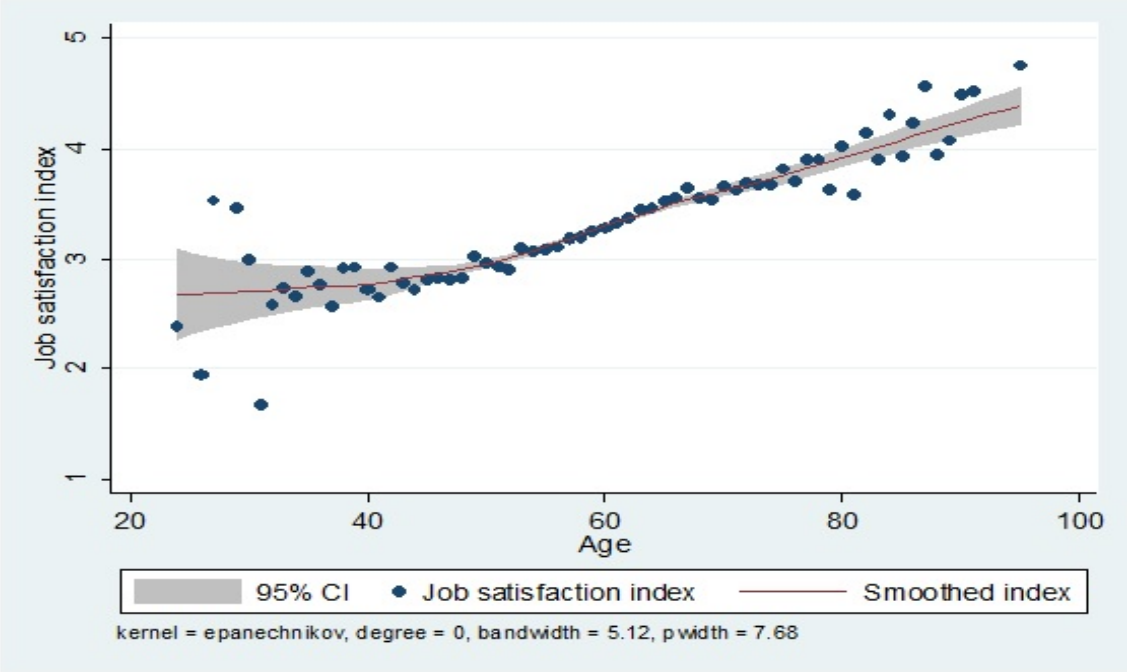


Figure 4.4: Evolution of job satisfaction index among age

Table 4.14: Characterization of early retirement in Europe : Part 1

Variable	Early^{+}	Wave 1	Wave 2	Wave 4	Wave 5	Overall
Proportion of worker	in %	43.26	41.85	39, 94	38.88	40.7
looking for early retirement	Obs.	6, 840	8, 105	11, 428	9, 886	36, 259
Proportion of smoker	Yes	0.2815 (0.0083)	0.2621 (0.0076)	0.2397 (0.0063)	0.1964 (0.0064)	0.242 (0.0035)
	No	0.2363 (0.0068)	0.2166 (0.006)	0.2073 (0.0049)	0.1802 (0.0049)	0.207 (0.0028)
	Difference	0.0452*** (0.0106)	0.0454*** (0.0096)	0.0324*** (0.0079)	0.0162** (0.008)	0.035*** (0.0044)
Proportion of male	Yes	0.5465 (0.0092)	0.5139 (0.0086)	0.4748 (0.0074)	0.4737 (0.008)	0.4979 (0.0041)
	No	0.4945 (0.008)	0.4963 (0.0073)	0.4599 (0.006)	0.4547 (0.0064)	0.4727 (0.0034)
	Difference	0.052*** (0.0122)	0.0176 (0.0113)	0.0149 (0.0095)	0.0191* (0.0103)	0.0252*** (0.0053)
Proportion of worker that	Yes	0.3721 (0.0089)	0.3721 (0.0083)	0.3935 (0.0072)	0.373 (0.0078)	0.379 (0.004)
are afraid that health limits	No	0.1927 (0.0063)	0.1842 (0.0056)	0.1916 (0.0047)	0.1592 (0.0047)	0.1811 (0.0026)
ability to work	Difference	0.1794*** (0.0108)	0.1879*** (0.0099)	0.2019*** (0.0085)	0.2138*** (0.0088)	0.1979*** (0.0048)
Proportion of worker	Yes	0.8364 (0.0068)	0.8275 (0.0065)	0.7627 (0.0063)	0.7575 (0.0069)	0.791 (0.0033)
in couple	No	0.8093 (0.0063)	0.8088 (0.0057)	0.748 (0.0052)	0.7458 (0.0056)	0.7718 (0.0029)
	Difference	0.0271*** (0.0094)	0.0187** (0.0087)	0.0147* (0.0082)	0.0118 (0.0089)	0.0193*** (0.0044)
Mean of age in year	Yes	55.61 (0.0814)	55.84 (0.0766)	55.96 (0.0664)	57.3 (0.0711)	56.21 (0.037)
	No	56.18 (0.0861)	56.74 (0.0768)	57.37 (0.0699)	58.78 (0.0718)	57.42 (0.0383)
	Difference	−0.57*** (0.1185)	−0.91*** (0.1084)	−1.42*** (0.0964)	−1.48*** (0.101)	−1.21*** (0.0533)
Proportion of workers that	Yes	0.8628 (0.0063)	0.8567 (0.006)	0.8718 (0.0049)	0.8574 (0.0155)	0.6692 (0.0039)
are satisfy of their job	No	0.9531 (0.0034)	0.9635 (0.0027)	0.9672 (0.0021)	0.9626 (0.0066)	0.7292 (0.003)
	Difference	−0.0903*** (0.0068)	−0.1068*** (0.0061)	−0.0954*** (0.0049)	−0.1052 (0.015)	−0.0599*** (0.0049)

$^{+}$: look for early retirement. Standard errors are in parenthesis. *** : significant at 1% level.

** : significant at 5% level, * : significant at 10% level

Table 4.15: Characterization of early retirement in Europe : Part 2

Variable	Early^{+}	Wave 1	Wave 2	Wave 4	Wave 5	Overall
Mean of annual	Yes	25,565.87 (180.83)	17,885.61 (138.73)	19,502.44 (130.43)	21,209.61 (144.76)	20,913.96 (74.83)
earnings from	No	28,305.94 (187.18)	20,288.7 (136.21)	22,479.46 (128.97)	24,145.13 (139.52)	23,628.62 (73.79)
employment in €	Diff.	$-2,740.09^{***}$ (265.63)	$-2,403.09^{***}$ (199.28)	$-2,977.02^{***}$ (190.29)	$-2,935.52^{***}$ (208.60)	$-2,714.66^{***}$ (108.33)
Amount in bank	Yes	12,336.73 (198.78)	15,486.33 (252.69)	17,082.53 (303.44)	15,111.29 (243.92)	15,238.72 (133.67)
account in €	No	15253.45 (208.16)	23,741.11 (313.56)	24,741.46 (311.09)	23,644.07 (283.19)	22,445.74 (149.55)
	Diff.	$-2,916.72^{***}$ (294.52)	$-8,254.77^{***}$ (425.84)	$-7,658.93^{***}$ (453.93)	$-8,532.79^{***}$ (402.79)	$-7,207.03^{***}$ (211.12)
Out-of-pocket	Yes	293.96 (4.90)	269.03 (5.14)	-	435.79 (6.71)	344.46 (3.48)
health expenditures	No	312.22 (5.00)	279.44 (3.59)	-	463.96 (5.53)	370.91 (3.02)
in €	Diff.	-18.26^{**} (7.13)	-10.41^{*} (6.07)	-	-28.17^{***} (8.76)	-26.45^{***} (4.64)
Proportion of	Yes	0.2406 (0.0079)	0.2565 (0.0075)	0.1933 (0.0058)	0.2734 (0.0072)	0.2382 (0.0035)
undergraduate	No	0.3187 (0.0075)	0.366 (0.007)	0.2880 (0.0055)	0.3805 (0.0062)	0.3367 (0.0032)
studies at least	Diff.	-0.0781^{***} (0.011)	-0.1095^{***} (0.0105)	-0.0948^{***} (0.0083)	-0.1071^{***} (0.0098)	-0.0985^{***} (0.0049)
Percentage of	Yes	33.3586 (0.7757)	36.3031 (0.672)	41.5433 (0.5732)	43.7965 (0.6365)	39.251 (0.329)
salary to be	No	28.5844 (0.634)	29.7371 (0.5162)	33.6839 (0.4709)	37.5918 (0.4963)	32.8254 (0.2615)
received as pension	Diff.	4.7742^{***} (1.0018)	6.566^{***} (0.8474)	7.8594^{***} (0.7418)	6.2047^{***} (0.8071)	6.4256^{***} (0.4203)
Mean of job	Yes	2.7745 (0.0133)	2.6746 (0.0126)	2.6664 (0.0113)	3.6753 (0.0079)	2.9527 (0.0067)
satisfaction and	No	3.1979 (0.0105)	3.1496 (0.0092)	3.1138 (0.0082)	3.7612 (0.0051)	3.3341 (0.0046)
condition index	Diff.	-0.4234^{***} (0.017)	-0.475^{***} (0.0156)	-0.4474^{***} (0.0139)	-0.1406^{***} (0.0094)	-0.3814^{***} (0.0081)

$^{+}$: look for early retirement. Standard errors are in parenthesis. *** : significant at 1% level.

** : significant at 5% level, * : significant at 10% level

Appendix 5 : Descriptive statistics on post retirement employment

Table 4.16: Proportion of retired workers among aged worker per country, school grade and gender

	Wave 1			Wave 2			Wave 4			Wave 5			Overall		
Country	*Male*	*Female*	*Total*	*Male*	*Female*	*Total*	*Male*	*Female*	*Total*	*Male*	*Female*	*Total*	*Male*	*Female*	*Total*
Per country															
Austria	0,8	6,7	3,4	0	7,3	3,4	3,3	10,3	6,9	1,6	7,8	4,8	2,1	8,8	5,5
Germany	5,3	2,6	4	3,3	1,7	2,5	8	4,4	5,9	7,4	3,6	5,2	5,4	2,9	4,1
Sweden	1,5	0,4	0,9	4	1	2,3	5,9	2,5	4,1	12,1	1,8	6,7	4,7	1,1	2,8
Netherlands	1,7	1	1,3	1,2	1,7	1,4	1,5	1	1,2	4,4	1,9	3,1	2,1	1,4	1,7
Spain	2,6	4,4	3,3	1,3	1,9	1,5	2,2	2,2	2,2	3,2	2,5	2,8	2,3	2,6	2,4
Italy	3,8	2,6	3,3	4,6	1	3	6,4	1,3	4,1	3,3	2,3	2,8	4,7	1,7	3,3
France	4,5	6,2	5,5	6,8	7,7	7,3	10,2	8,9	9,5	15,5	10,7	12,8	9,7	8,6	9,1
Denmark	3,3	2,8	3	4	1,8	2,9	6,3	2,3	4,3	8,1	3,8	5,9	5,6	2,6	4,1
Greece	3,4	0,8	2,4	6	2	4,5	–	–	–	–	–	–	4,7	1,4	3,4
Switzerland	9,1	5,6	7,4	6,9	3,7	5,3	7,7	6	6,8	10,8	7	8,8	8,8	5,9	7,3
Belgium	0,9	1,1	1	1,8	0,3	1,1	1,9	1,1	1,5	2	0,9	1,4	1,7	0,9	1,3
Israel	8,3	5,1	6,6	11,3	7	9,1	0	0	0	21,8	20,8	21,3	12,7	9,4	11,1
Czech Republic	–	–	–	4,5	3,4	4	3,4	4,9	4,2	5	4,7	4,8	4,2	4,5	4,4
Poland	–	–	–	0	2,2	1	2	4,7	3,3	–	–	–	0,8	3,1	1,9
Slovenia	–	–	–	–	–	–	0,5	0	0,3	2,9	0	1,4	1,6	0	0,8
Estonia	–	–	–	–	–	–	24,6	28,2	26,7	25,1	30,6	28,4	24,8	29,3	27,5
Total	3,6	2,9	3,2	4,4	2,9	3,7	7,8	8,6	8,2	10,2	10	10,1	6,8	6,8	6,8

Table 4.17: Proportion of retired workers among aged worker per country and school grade

	Wave 1			Wave 2			Wave 4			Wave 5			Overall		
Country	*Male*	*Female*	*Total*	*Male*	*Female*	*Total*	*Male*	*Female*	*Total*	*Male*	*Female*	*Total*	*Male*	*Female*	*Total*
Per school grade															
no grade	7,1	5,8	6,3	12,5	3,3	6,5	–	–	–	3,8	12,8	9,2	8	6,5	7,1
college degree	2,6	2,9	2,7	4	3,4	3,7	8	8,2	8,1	7,9	8,9	8,4	5,8	6,3	6
undergraduate studies	6,5	4	5,2	3,8	2,4	3,1	6,4	9,1	7,9	14,2	18,1	16,4	7,3	9,6	8,6
graduate studies	3,5	1,7	2,6	5,1	2,2	3,6	9	8,7	8,8	13,6	10,2	11,7	8,3	6,4	7,3
Total	3,1	2,6	2,8	4,4	2,9	3,7	7,8	8,6	8,2	10,2	10	10,1	6,7	6,8	6,8

Appendix 6 : Multiple imputation model for pension rate

Table 4.18: Outputs of multiple imputation model

Variable	Coefficient
Log of wage	-0.00056^{***} (0.00009)
Grade	Ref = Graduate studies
No grade	0.02295^{***} (0.00151)
College degree	0.022^{***} (0.00113)
Undergraduate studies	0.01702^{***} (0.00125)
Intercept	0.55217^{***} (0.00422)
Country	ref : Estonia
Austria	0.13666^{***} (0.00502)
Germany	0.03179^{***} (0.00542)
Sweden	-0.00509 (0.00498)
Netherlands	-0.02272^{***} (0.00534)
Spain	0.26443^{***} (0.00631)
Italy	0.15523^{***} (0.00557)
France	0.0625^{***} (0.00476)
Denmark	-0.25459^{***} (0.00478)
Greece	0.11541^{***} (0.0072)
Switzerland	-0.15168^{***} (0.00482)
Belgium	0.06925^{***} (0.00482)
Israel	-0.26053^{***} (0.00625)
Czech Republic	-0.02318^{***} (0.00507)
Slovenia	0.08582^{***} (0.00667)
Poland	0.13705^{***} (0.00732)
$\sigma_\mu = 0.17921$, $\sigma_e = 0.08981$, $\rho = 0.79929$	

*** : significant at 1% level, ** : significant at 5% level, Standard errors are in parenthesis.

Conclusion

Results review and contributions

This dissertation consists of five research manuscripts that contribute towards the same goal. The aim of the dissertation was to investigate the micro-econometric approaches for testing and measuring causal links in various fields in social sciences, using both ex-ante and ex-post frameworks. This aim has been reached through four objectives covered by the five research manuscripts herein. The first manuscript reviewed technical issues while testing for causality, with a focus on causality measurement in health economics. The second manuscript investigated causal links between health status and job status by accounting simultaneously for endogeneity, individual heterogeneity and initial conditions. The third manuscript dealt with the dynamic of causal links and its determinants. The fourth manuscript investigated an ex-ante structural model to predict early retirement behaviour among aged workers. The fifth manuscript presented technical issues related to the estimation of causal links in the second manuscript.

The main results of the dissertation are discussed. We show that health and job are mutually causal and the dynamic is persistent. We also highlight the important role of initial conditions in the dynamic of the causal links between health and job. We show evidence that health status causes job status at the beginning and at the end of professional life and during the same period, job status does not cause health status. In terms of early retirement behaviour, we show that the marginal health depreciation due to job, the marginal productivity of health expenditures and the preference for future determine individual probability to retire earlier.

The major contributions of this dissertation are methodological. We first make a systematic technical review of causality tests and measurement methods in literature. We propose and test innovative approaches of causality measurement. We introduce an approach that has the advantage to account for endogeneity, individual heterogeneity and initial conditions while testing for causality. We also propose a non-parametric framework that helps to estimate dynamic causal links and to investigate their determinants. A part from that, we propose an innovative approach to estimate a health depreciation and health investment functions. Then, we show how these two functions and the preferences for future affect the early retirement behaviour.

Limitations

While analyzing causal links through several approaches described above, some limitations were encountered. First, the dataset used for the second and third manuscripts is a retrospective panel data. In this type of collection procedures, individuals are asked to provide information from the date they left school till the date of survey. This involves many collection biases in data. Furthermore, there may exist a selection bias due to the selection of individuals alive. Since we are analyzing relationships between health and job statuses, individuals who died before the data collection are of interest[15] because they may be dead due to a worst health status. The dataset does not allow to observe and analyze infra-annual transition. Second, in all implemented approaches above, we assumed the processes to be of lag one. This assumption can be challenged, since the optimal lag length that determines a phenomenon can be more than one time period.

[15]We are grateful to the reviewers of the Journal of Applied Econometrics for this comment on our research manuscript

Bibliography

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.

Albuquerque P., Arcanjo M. and Escaria V. (2009). Early retirement in Portugal. *Working Paper WP 39/2009/DE/SOCIUS*.

Alessie R., Hochguertel S. and van Soest A. (2004). Ownership of stocks and mutual funds: a panel data analysis. *Review of Economics and Statistics*, 86:783–796.

Alhawarin, I. (2014). Patterns and determinants of early retirement: The case of Jordanian men. *Jordan Journal of Economic Sciences*, 1(1):49–65.

Arulampalam, W. and Stewart, M. (1995). The determinants of individual unemployment durations in a era of high unemployment. *Economic Journal*, 105(429):321–332.

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. (2015). Health, work and working conditions: a review of the European economic literature. *European Journal of Health Economics*.

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*.

Bastida E. and Pagan J.A. (2002). The impact of diabetes on adult employment and earnings of mexican americans: Findings from a community based study. *Health Economics*, 11(5):403–413.

Bazzoli, G. J. (1985). The early retirement decision: New empirical evidence on the influence of health. *The Journal of Human Resources*, 20(2):214–234.

Benitez-Silva H., Buchinsky M., Chan H.M., Cheidvasser S. and Rust J. (2004). How large is the bias in self-reported disability ? *Journal of Applied Econometrics*, 19:649–670. DOI: 10.1002/jae.671.

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

Berkowitz M. and Johnson W.G. (1974). Health and labor force participation. *Journal of Human Resources*, 9(1):117–128.

Bertrand M., Duflo E., and Mullainathan S. (2004). How much should we trust differences-in-differences estimates ? *The Quaterly Journal of Economics*, pages 249–275.

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*, 30(2):275–287.

Bouissou M. B. Laffont J. J. and Vuong Q. H. (1986). Tests of noncausality under Markov assumptions for qualitative panel data. *Econometrica*, 54(2):395–414.

Boumans N. P.G., De Jong Ad H.J., and Vanderlinden L. (2008). Determinants of early retirement intentions among Belgian nurses. *Journal of Advanced Nursing*, 63(1):64–74.

Bound, J. (1991). Self-reported versus objective measures of health in retirement models. *Journal of Human Resources*, 26(1):106–138.

Brenner, M.H. (2002). Employment and public health. Final report to the European Commission directorate general employment, industrial relations and social affairs. *European Commission. Brussels*.

Brothers, L. S. (2000). How do you choose a retirement age? *Retirement Needs Framework: SOA Monograph M-RS00*, 1:11–24.

Brown, H.S., Pagan, J.A. and Bastida, E. (2005). The impact of diabetes on employment: Genetic IVs in a bivariate probit. *Health Economics*, 14:537–544.

Bry X. (1999). Analyses factorielles simples. *Economica*.

Butterworth, P., Gill, S., Rodgers, B., Anstey, K., Villamil, E., and Melzer, D. (2006). Retirement and mental health: analysis of the Australian national survey of mental health and well-being. *Social Science and Medicine*, 62:1179–1191.

Cai, L. (2010). The relationship between health and labour force participation: Evidence from a panel data simultaneous equation model. *Labour Economics*, 17:17–90.

Cai, L. and Kalb, G. (2006). Health status and labour force participation: evidence from Australia. *Health Economics*, 15:241–261.

Cai, L. and Kalb, G. (2007). Health status and labour force status of older working-age Australian men. *Australian Journal of Labour Economics*, 10:227–252.

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. (2015). Does job insecurity deteriorate health ? *Cahiers de la Chaire Santé*, (21).

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

Christensen, B. J. and Kallestrup-Lamb, M. (2012). The impact of health changes on labor supply: evidence from merged data on individual objective medical diagnosis codes and early retirement behavior. *Health Economics*, 21:56–100.

Clark, A., Gerogellis, Y. and Sanfey, P. (2001). Scarring: the psychological impact of past unemployment. *Economica*, 68:221–241.

Coe N.B. and Zamarro G. (2008). Retirement effects on health in Europe. *RAND working paper*, (588).

Currie, J. and Madrian, B. (1999). Health, health insurance and the labor market. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, pages 3310–3415. Elsevier.

Cutler D. and Richardson E. (1997). Measuring the Health of the U.S. population. *Brookings Papers : Microeconomics*.

De Jong, P. (1988). The likelihood for a state space model. *Biometrika*, 75(1):165–169.

De Jong, P. (1991a). The diffuse Kalman filter. *The Annals of Statistics*, 19(2):1073–1083.

De Jong, P. (1991b). Stable algorithms for the state space model. *Journal of Time Series Analysis*, 12(2):143–157.

Debrand, T. (2011). L'influence des conditions de travail sur les dépenses de santé. *IRDES working paper*, 41.

Debrand, T. and Blanchet, D. (2008). Aspiration à la retraite, santé et satisfaction au travail : une comparaison européenne. *IRDES Working Paper*.

Delattre, E. and Dormont, B. (2003). Fixed fees and physician-induced demand: A panel data study on French physicians. *Health Economics*, 12:741–754.

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*.

Dorn, D. and Sousa-Poza, A. (2004). The determinants of early retirement in Switzerland. *Forschungsinstitut für Arbeit und Arbeitsrecht*.

Droesbeke, J.J. and Saporta, G. and Thomas-Agnan, C. (2013). *Modèles à variables latentes et modèles de mélange*. Editions Technip.

Eichenbaum-Voline, S., Malavolti, L., Paraponaris, A. and Ventelou, B. (2008). Cancer et activité professionnelle. *Revue de l'OFCE*, 104:105–134.

Faragher E.B. , Cass M. and Cooper C.L. (2005). The relationship between job satisfaction and health: a meta-analysis. *Occupational and Environmental Medicine*, 62:105–12.

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.

Florens, J.-P. and Mouchart, M. (1982). A note on noncausality. *Econometrica*, 50(3):583–591.

Florens, J.-P. and Mouchart, M. (1985). A linear theory for noncausality. *Econometrica*, 53(1):157–176.

Galama T., Kapteyn A., Fonseca R., and Michaud P. C. (2013). A health production model with endogenous retirement. *Health Economics*, 22:883–902.

Gould W. Pitblado J. and Poi B. (2010). *Maximum Likelihood Estimation With Stata*. STATA PRESS, Fourth edition.

Gouriéroux, C. and Monfort, A. (1997). Simulated-based econometric methods. *Oxford University Press*.

Gouriéroux C., Monfort A., and Renault E. (1987). Kullback causality measures. *Annales d'Economie et de Statistique*, 6(7):369–410.

Granger C.W.J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424–438.

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy*, 80:223–255.

Grossman, M. (1999). The human capital model of the demand for health. *NBER Working paper series*, Working paper 7078:1–98.

Haan, P. and Myck, M. Dynamics of health and labor market risks.

Hall R. and Jones C. (2007). The Value of Life and the Rise in Health Spending. *Quarterly Journal of Economics*.

Hamon-Cholet, S. and Sandret N. (2007). Accidents et conditions de travail. *Premières Synthèses*, 31.

Heckman, J. J. (1981). Statistical models for discrete panel data. In Manski, C. and McFadden, D., editors, *Structural Analysis of Discrete Data with Econometric Applications*, pages 114–178. MIT Press.

Heckman, J. J. (2008). Econometric causality. *IZA Discussion Paper*, (3525):1–53.

Jackel, P. (2005). A note on multivariate gauss-hermite quadrature. *www. pjaeckel. webspace. virginmedia.*

Jiménez-Martin S., Labeaga J. M., and Prieto C. V. (2006). A sequential model of older workers' labor force transitions after a health shock. *Health Economics*, 15:1033–1054.

Jin, R.L., Shah, C.P. and Svoboda, T.J. (1995). The impact of unemployment on health: a review of the evidence. *Canadian Medical Association Journal*, 153:529–540.

Khandker R. Shahidur, Gayatri B. Koolwal, and Hussain A. Samad. Handbook on impact evaluation : Quantitative methods and practices. DOI: 10.1596/978-0-8213-8028-4.

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

Lindeboom, M. and Kerkhofs, M. Health and work of the elderly: subjective health measures, reporting errors and endogeneity in the relationship between health and work. *Journal of Applied Econometrics*, 24:1024–1046. DOI: 10.1002/jae.1077.

Liu Q. and Pierce D.A. (1994). A note on Gauss-Hermite quadrature. *Biometrika*, (3):624–629.

Llena-Nozal, A. (2009). The effect of work status and working conditions on mental health in four OECD countries. *National Institute Economic Review*, 9:72–87.

Madden, D. (2004). Labour market discrimination on the basis of health: an application to UK data. *Applied Economics*, 36:421–442.

Mathers, C.D. and Schofield D.J. (1998). The health consequences of unemployment: the evidence. *Medical Journal of Australia*, 168:178–182.

Mein G., Martikainen P., Stansfeld S. A., Brunner E. J., Fuhrer R., and Marmot M. G. (2000). Predictors of early retirement in British civil servants. *Oxford Journals - Medicine and Health*, 29:529–536.

Mermilliot C. (2012). L'impact de la mise en place d'un suivi d'adresse entre les deux vagues de l'enquète santé et itinéraire professionnel (SIP). *Document de travail DRESS*.

Mesrine, A. (2000). La surmortalité des chômeurs : un effet catalyseur du chômage ? *Economie et Statistique*, 334.

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. DOI : 10.1016/j.jhealeco.2008.04.002.

Michel Volle (1997). Analyses des données. *Economica*.

Miranda, A. (2011). Migrant networks, migrant selection and high school graduation in Mexico. *Research in Labor Economics*.

Morris, J., Cook, D. and Shaper, G. (1994). Loss of employment and mortality. *British Medical Journal*, 308:1134–1139.

Moussa, R. and Delattre, E. (2015). On the estimation of causality in a bivariate dynamic probit model on panel data with Stata software. A technical review. *ThEMA working paper*.

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

Nadinloyi K. , Sadeghi H. and Hajloo N. (2013). Relationship between job satisfaction and employees' mental health. *Procedia - Social and Behavioral Sciences*, 84(0):293 – 297.

Nair-Richert, U. and Weinhold D. (2001). Model evaluation and causality testing in short panels: The case of infrastructure provision and population growth in the Brazilian Amazon. *Journal of Regional Science*, 41(4):639–657.

Naylor, J. C. and Smith, A. F. M. (1982). Applications of a method for the efficient computation of posterior distributions. *Royal Statistical Society*.

Peyrache A. and Rambaldi A. N. (2012). A state-space stochastic frontier panel data model. *Centre for Efficiency and Productivity Analysis, Working Paper No. WP01/2012*.

Pierce, D. A. and Haugh, L. D. (1976). Causality in temporal systems: Characterizations and a survey. *Journal of Econometrics*, 5:265–293.

Pollak, C. (2012). Employed and Happy despite Weak Health? Labour Market Participation and Job Quality of Older Workers with Disabilities. *IRDES Working Paper*.

Quinn, J. F. (1977). Microeconomic determinants of early retirement: A cross-sectional view of white married men. *The Journal of Human Resources*, 12(3):329–346.

Rietveld, C. A., van Kippersluis, H., and Thurik, A. R. (2015). Self-employment and health: Barriers or benefits? *Health Economics*, 24(10):1302–1313.

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

Sekkat, K. (1989). L'analyse de la causalité comme méthode de détermination des filières industrielles. *Annales d'Economie et de Statistique*, (14):191–223.

Sickles, R. and Taubman, P. (1986). An analysis of the health and retirement status of the elderly. *Econometrica*, 54:1339–1356.

Siegrist J., Morten Wahrendorf, Olaf von dem Knesebeck, Hendrik Jurges and Axel Borsch-Supan (2006). Quality of work, well-being, and intended early retirement of older employees. baseline results from the SHARE study. *European Journal of Public Health*, 17(1):62–68.

Sims, C. A. (1972). Money, income, and causality. *The American Economic Review*, 64(4):540–552.

Stern, S. (1989). Measuring the effect of disability on labour force participation. *Journal of Human Resources*, 24:361–395.

Szubert, Z. and Sobala, W. (2005). Current determinants of early retirement among blue collar workers in Poland. *International Journal of Occupational Medicine and Environmental Health*, 18(2):177–184.

Toda, H. and Yamamoto (1995). Statistical inference in vector autoregressions with possibly integrated processes. *Journal of Econometrics*, (66):225–250.

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

Wagstaff A. (2002). The Demand for Health: an Empirical Reformulation of the Grossman Model. *Econometric Analysis of Health Data*, pages 15–23.

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.

Wooldridge, J. (2005). Simple solution to the initial conditions problem in dynamic, nonlinear panel-data models with unobserved heterogeneity. *Journal of Applied Econometrics*, 20:739–54.

Zhang X., Zhao X. and Harris A. (2008). Chronic diseases and labour force participation in Australia. *Journal of Health Economics*, 28:91–108.

Appendix A

On the Estimation of Causality in a Bivariate Dynamic Probit Model on Panel Data with Stata Software. A Technical Review[1] [2]

[1]Authors : **Richard Moussa**, ThEMA-UCP and ENSEA Abidjan; **Eric Delattre**, ThEMA-UCP

[2]We greatfully acknowledge the Centre Maurice Halbwachs (Réseau Quetelet) for access to the SIP 2007 data set (Santé et itinéraire professionnel - 2007. DARES producteur. Centre Maurice Halbwachs diffuseur).

Abstract

In order to assess causality between binary economic outcomes, we consider the estimation of a bivariate dynamic probit model on panel data that has the particulary to account the initial conditions of the dynamic process. Due to the untractable form of the likelihood function that is a two dimensions integral, we use an approximation method : the adaptative Gauss-Hermite quadrature method as proposed by Liu and Pierce (1994). For the accuracy of the method and to reduce computing time, we derive the gradient of the log-likelihood and the hessian of the integrand. The estimation method has been implemented using the d1 method of Stata software. We made an empirical validation of our estimation method by applying on simulated data set. We also analyze the impact of the number of quadrature points on the estimations and on the estimation process duration. We then conclude that when exceeding 16 quadrature points on our simulated data set, the relative differences in the estimated coefficients are around 0.01% but the computing time grows up exponentially.

Keywords: Causality; Bivariate Dynamic Probit;
Gauss-Hermite Quadrature; Simulated Likelihood; Gradient; Hessian

JEL Classification: C5; C6

Introduction

Testing Granger causality has generated a large set of paper in the literature. The larger part of this literature concerns the case where we have continuous dependent variables. For binary outcomes, there is also a way to consider the causality problem. As described by Adams, McFadden and alii (2003) for a vector of dependant variables, the one order Granger causality can be analyse as a probability conditional independence given a set of exogenous variables and the first order lagged dependent variables. And for a binary outcome in the dependent vector, one can use a probit probability that implies the use of latent variable.

For panel data case, as the one way fix effects model estimated on a finite sample has necessarily inconsistent estimators (Heckman, 1981), the random effect model is used. Due to the fact that we aim to test for one order Granger causality, lagged dependent variables are included as explanatory variables. For the first wave of the panel, we do not have previous values for the dependent variables, and treating them casually or as exogenous leads to inconsistent estimators (Heckman 1981). So we specify an other equation for initial conditions as described by Alessie (2004). The equation is allowed to have different explanatory variables and different idiosyncratic error terms from the dynamic equation.

This specification leads to a likelihood function with an untractable form that is a two dimensions integral with a large set of parameters to be estimated. The estimation of this likelihood function requires the use of numerical

approximation of integral function such as maximum simulated likelihood (see Gouriéroux and Monfort 1993 for more details) or Gauss-Hermite quadrature (for more details see Naylor and Smith 1982, Liu and Pierce 1994, Jackel 2005).

In this paper, we discuss on the problem of testing Granger causality with a bivariate dynamic probit model taking into account the initial condition. The organization of this paper is the following one. In section 1 we explain the causality test method for bivariate probit model in panel data. In section 2, we describe the estimation method available when the likelihood function has an untractable form (two dimensions integral in our case). Section 3 presents the calculation of the gradient with respect to the model parameters and the calculation of the hessian matrix with respect to the random effects vector. In section 4, we present a robustness analysis of our selected estimation method by doing some simulations[3].

A.1 Testing causality with a bivariate dynamic probit model

This section aims to describe causality test method in the case of binary variables. We start by presenting the general approach in time series before introducing panel data case. We end this section by a discussion on the initial condition problem.

A.1.1 Testing causality : general approach

Causality concept was introduced by Granger (1969) as a better predictability of a variable Y by the use of it lag values, the lag value of an other variable Z and some controls X. In his paper, Granger (1969) distinguishes instantaneous causality that means Z_t is causing Y_t (if Z_t include in the model it improves the predictability of Y_t than if not) from lag causality that means lag values of Z improve the predictability of Y_t. In this section, we rule out the instantaneous causality and deal with lag causality of one period.

The one period Granger causality can be rephrase in terms of conditional independence. Without lost of generality, we present the univariate case for time series. Let's Y_t and Z_t denote some dependent variables and X_t denote a set of controls variables. One period Granger non-causality from Z to Y is the conditional independence of Y_t from Z_{t-1} conditionally to X_t and Y_{t-1}. More clearly, Granger non-causality from Z to Y is :

$$f(Y_t|Y_{t-1}, X_t, Z_{t-1}) = f(Y_t|Y_{t-1}, X_t) \tag{A.1}$$

Note that the same kind of relationship can be written for Granger non-causality from Y to Z. As Y_t and Z_t are binary outcome variables, we can use latent variables (Y^* and Z^* respectively) and make the assumption that

[3]For each section, specifics notations are down at the beginning of the section. Otherwise, in general $f(x)|_{x=a}$ denote the value of the function or the matrix f at the point a. When not specify, $|a|$ denote the integer part of the scalar a.

Y and Z have positive outcomes (equals to 1) if their latent variable is positive. The latent variables are defined as follows :

For the left term of the equation A.1 ($f(Y_t|Y_{t-1}, X_t, Z_{t-1})$) :

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

For the right term of the equation A.1 ($f(Y_t|Y_{t-1}, X_t)$) :

$$Y_t^* = X_t\beta_1 + \delta_{11}Y_{t-1} + \epsilon_t^1$$
$$Z_t^* = X_t\beta_2 + \delta_{21}Z_{t-1} + \epsilon_t^2$$

where

$$\begin{pmatrix}\epsilon^1\\ \epsilon^2\end{pmatrix} \rightsquigarrow N(0, \Sigma_\epsilon) \; with \; \Sigma_\epsilon = \begin{pmatrix}1 & \rho_\epsilon\\ \rho_\epsilon & 1\end{pmatrix}$$

To fit the joint distribution of Y and Z conditionally to X (meaning that we estimate a bivariate model), we need to analyze four available situations that are ($Y = Z = 1$), ($Y = Z = 0$), ($Y = 1; Z = 0$) and ($Y = 0; Z = 1$). For each of these situations, we have :

$$P\Big(Y_t = 1, Z_t = 1|X_t\Big) = P\Big(\epsilon_t^1 > -X_t\beta_1 - \delta_{11}Y_{t-1} - \delta_{12}Z_{t-1}, \epsilon_t^2 > -X_t\beta_2 - \delta_{21}Y_{t-1} - \delta_{22}Z_{t-1}\Big)$$
$$P\Big(Y_t = 0, Z_t = 0|X_t\Big) = P\Big(\epsilon_t^1 < -X_t\beta_1 - \delta_{11}Y_{t-1} - \delta_{12}Z_{t-1}, \epsilon_t^2 < -X_t\beta_2 - \delta_{21}Y_{t-1} - \delta_{22}Z_{t-1}\Big)$$
$$P\Big(Y_t = 1, Z_t = 0|X_t\Big) = P\Big(\epsilon_t^1 > -X_t\beta_1 - \delta_{11}Y_{t-1} - \delta_{12}Z_{t-1}, \epsilon_t^2 < -X_t\beta_2 - \delta_{21}Y_{t-1} - \delta_{22}Z_{t-1}\Big)$$
$$P\Big(Y_t = 0, Z_t = 1|X_t\Big) = P\Big(\epsilon_t^1 < -X_t\beta_1 - \delta_{11}Y_{t-1} - \delta_{12}Z_{t-1}, \epsilon_t^2 > -X_t\beta_2 - \delta_{21}Y_{t-1} - \delta_{22}Z_{t-1}\Big)$$

As we can see, by supposing $q_t^1 = 2Y_t - 1$ and $q_t^2 = 2Z_t - 1$, we can rewrite the probabilities above as :

$$P\Big(Y_t, Z_t|X_t\Big) = \Phi_2\Big(q_t^1(X_t\beta_1 + \delta_{11}Y_{t-1} + \delta_{12}Z_{t-1}), q_t^2(X_t\beta_2 + \delta_{21}Y_{t-1} + \delta_{22}Z_{t-1}), q_t^1q_t^2\rho_\epsilon\Big)$$

where $\Phi_2()$ stands for the bivariate normal c.d.f.

Then testing Granger non-causality in this specification is testing $\delta_{12} = 0$ for Z is not causing Y and testing $\delta_{21} = 0$ for Y is not causing Z.

A.1.2 Testing causality : Panel data case

For panel data case, two major approaches can be used. The first one is to consider that causal effect is not the same for all individuals in the panel (Weinhold, 2000). This approach is useful when individuals are heterogeneous or when the causal effect is not homogenous. The specification for latent variables are :

$$Y_{it}^* = X_t\beta_1 + \delta_{11,i}Y_{i,t-1} + \delta_{12,i}Z_{i,t-1} + \eta_i^1 + \zeta_{it}^1$$
$$Z_{it}^* = X_t\beta_2 + \delta_{21,i}Y_{i,t-1} + \delta_{22,i}Z_{i,t-1} + \eta_i^2 + \zeta_{it}^2$$

Where $(\eta_i^1, \eta_i^2)'$ denote the individual random effects which are zero mean covariance matrix Σ_η and $(\zeta_{it}^1, \zeta_{it}^2)'$ denote the idiosyncratic shocks which are zero mean and covariance matrix Σ_ζ with

$$\Sigma_\eta = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_\eta \\ \sigma_1\sigma_2\rho_\eta & \sigma_2^2 \end{pmatrix} and \ \Sigma_\zeta = \begin{pmatrix} 1 & \rho_\zeta \\ \rho_\zeta & 1 \end{pmatrix}$$

In this approach, testing Granger non-causality is equivalent to test $\delta_{12,i} = 0, i = 1, ..., N$ for Z is not causing Y and to test $\delta_{21,i} = 0, i = 1, ..., N$ for Y is not causing Z.
The second approach (that is on use in this paper) is to suppose the causal effects, if it exists, is the same for all individuals in the panel. With the same notation that the previous case, the latent variables are :

$$Y_{it}^* = X_t\beta_1 + \delta_{11}Y_{i,t-1} + \delta_{12}Z_{i,t-1} + \eta_i^1 + \zeta_{it}^1$$
$$Z_{it}^* = X_t\beta_2 + \delta_{21}Y_{i,t-1} + \delta_{22}Z_{i,t-1} + \eta_i^2 + \zeta_{it}^2$$

Then testing Granger non-causality is equivalent to test $\delta_{12} = 0$ for Z is not causing Y and to test $\delta_{21} = 0$ for Y is not causing Z.

A.1.3 Dealing with initial conditions

For the first wave of the panel (initial condition), due to the fact that we do not have data for the previous state on Y and Z (no values for $Y_{i,0}$ and $Z_{i,0}$) we are not able to evaluate $P(Y_{i1}, Z_{i1}|Y_{i,0}, Z_{i,0}, X_i)$. By ignoring it in the individual overall likelihood, we ignore the data generation process for the first wage of the panel. This means that we suppose the data generating process of the first wave of the panel to be exogenous or to be in equilibrium. These assumptions hold only if the individual random effects are degenerated. If not, the initial condition (the first wave of the panel) are explained by the individual random effects and ignoring it leads to inconsistent parameter estimates (Heckman, 1981).

The solution proposed by Heckman (1981) for the univariate case and generalized by Alessie (2004) is to estimate a static equation for the first wave of the panel (meaning that we do not introduce lagged dependent variables). In this static equation, the random effects are a linear combination of the random effects in the next wave of the panel and idiosyncratic error terms may have different structure from the idiosyncratic error terms in the dynamic equation. Formally, the latent variables for the first wave of the panel are defined as follows :

$$Y_{i,1}^* = X_i^1\gamma_1 + \lambda_{11}\eta_i^1 + \lambda_{12}\eta_i^2 + \epsilon_i^1$$
$$Z_{i,1}^* = X_i^2\gamma_2 + \lambda_{21}\eta_i^1 + \lambda_{22}\eta_i^2 + \epsilon_i^2$$

Where $(\epsilon_i^1, \epsilon_i^2)'$ denote the idiosyncratic shocks which are zero mean and covariance matrix Σ_ϵ with $\Sigma_\epsilon = \begin{pmatrix} 1 & \rho_\epsilon \\ \rho_\epsilon & 1 \end{pmatrix}$.
As η^1 and η^2 are individual random effects respectively on Y and Z, λ_{12} and λ_{21} can be interpreted as the influence of the Y random individual effects (respectively Z random individual effects) on Z (respectively on Y) at the first wave of the panel.

A.2 Estimation methods

Due to the fact that the likelihood function has an untractable form (an integral function), it is impossible to estimate this likelihood by usual methods. We then deal with numerical integration methods that are numerical approximation method for an integral. In this section we describe two major methods and argue for one of them to estimate our likelihood function.

A.2.1 Gauss-Hermite quadrature method

Gauss-Hermite quadrature is a numerical approximation method use to close the value of an integral function. The default approach is relative to an univariate integral of the form :

$$\int_{\mathbb{R}} f(x)exp(-x^2)dx \tag{A.2}$$

With the Gaussian factor $exp(-x^2)$. But without this factor, one can use the Gauss-Hermite quadrature by using a straightforward transformation that is to multiply and divide the integrand $f(x)$ by a Gaussian factor $exp(-x^2)$. Then the integral above can be approximated using :

$$\int_{\mathbb{R}} f(x)exp(-x^2)dx = \sum_{q=1}^{Q} w_q * f(x_q) \tag{A.3}$$

Where x_q, $q = 1, ..., Q$ are nodes from the Hermite polynomial and w_q, $q = 1, ..., Q$ are corresponding weights.

This approximation supposes that the integrand can be well approximated by an $2Q + 1$ order polynomial and that the integrand is sampled on a symmetric range centered in zero. So, for suitable results, these two assumptions may be taken into account.

For the first one, finding best number of quadrature point can be achieve numerically. For the accuracy of the approximation, it is required to choose the best number of quadrature points. To do this, one can start with a number $\bar{q}$ of quadrature points and increase it and see if it significantly changes the result, and repeat this process until convergence in terms of overall likelihood value variation and estimated coefficients variation. But it is also important to take into account the fact that increasing number of quadrature point also increase computing time. An example of the impact of number of quadrature points on estimated results is given in section 5.

For the problem of suitable sampling range, the solution of using the adaptative Gauss-Hermite quadrature was proposed by Naylor and Smith (1982) and by Liu and Pierce (1994). In this approach, in fact of using $exp(-x^2)$ as a gaussian factor to multiply and divide the integrand, they use a gaussian density $\phi(\mu, \sigma)$ of mean μ and variance σ^2. That means (see Naylor and Smith, 1982) :

$$\int_{\mathbb{R}} \frac{f(x)\phi(x, \mu, \sigma)}{\phi(x, \mu, \sigma)}dx = \sum_{q=1}^{Q} w_q^* g(x_q^*) \tag{A.4}$$

Where $\frac{f(x)}{\phi(x,\mu,\sigma)}$.

Then the sampling range is transformed and the new nodes are $x_q^* = \mu + \sqrt{2}\sigma x_q$ and weights are $w_q^* = \sqrt{2}\sigma w_q exp(x_q^2)$. For Naylor and Smith (1982), one can choose the normal density with posterior mean and variance equal respectively to μ and σ. For the implementation, we can start with $\mu = 0$ and $\sigma = 1$ and at each iteration of the likelihood maximization process, calculate the posterior weighted mean and variance of the quadrature points and use them to calculate the nodes and weights for the next iteration. For Liu and Pierce (1994), one can choose μ to be the mode of the integrand and σ to be the square of the hessian of the log of integrand taken in the mode.

$$\sigma = \left(- \frac{\delta^2}{\delta x^2} log(f(x))|_{x=\hat{x}} \right)^{-1/2} \tag{A.5}$$

For the multivariate integral case, the same approach is used. Without lost of generality, we discuss the bivariate case that can be apply to others multivariate cases. The function to approximate is written as follows :

$$\int_{\mathbb{R}^2} f(x,y)dxdy \tag{A.6}$$

With the assumption of independence between x and y (that can be overcome by using a Cholesky decomposition $x' = x$ and $y' = \rho x' + y$, see Naylor and Smith (1982) or Jackel (2005) for more precision on these Cholesky transformation or other transformations that can lead to similar results) the integral above can be approximated by :

$$\int_{\mathbb{R}^2} \frac{f(x,y)\phi(x,\mu,\sigma)\phi(y,\mu,\sigma)}{\phi(x,\mu,\sigma)\phi(y,\mu,\sigma)} dxdy = \sum_{q_1=1,q_2=1}^{Q} w_{q_1}^* w_{q_2}^* g(x_{q_1}^*, y_{q_1}^*) \tag{A.7}$$

Where $\frac{f(x,y)}{\phi(x,\mu,\sigma)\phi(y,\mu,\sigma)}$.

And in this case, the nodes and weights are derived as follows :

$$\begin{pmatrix} x_{q_1}^* \\ y_{q_1}^* \end{pmatrix} = \hat{x} + \sqrt{2} * \left(- \frac{\delta^2}{\delta x^2} log(f(x,y))|_{x,y=\hat{x}} \right)^{-1/2} * \begin{pmatrix} x_{q_1} \\ y_{q_1} \end{pmatrix} \tag{A.8}$$

and

$$\begin{pmatrix} w_{q_1}^* \\ w_{q_2}^* \end{pmatrix} = 2 * \left| - \frac{\delta^2}{\delta x^2} log(f(x,y))|_{x,y=\hat{x}} \right|^{-1/2} * \begin{pmatrix} w_{q_1} exp(x_{q_1}^2) \\ w_{q_2} exp(x_{q_2}^2) \end{pmatrix} \tag{A.9}$$

Where $|A|$ denote the determinant of the matrix A.

Jackel (2005) also suggests that for the nodes with low weights (when contributions to the integral value are not significative) we can prune the range from those nodes in order to save calculation time. That means to set a scalar $\tau = \frac{w_1 w_{|(Q+1)/2|}}{Q}$ and drop all nodes with weights lower than this scalar.

A.2.2 Maximum simulated likelihood method

Maximum Simulated Likelihood method was introduced by Gouriéroux and Monfort (1993) as a solution to maximization problems that have an integral as objective function. In this approach, the likelihood function is

supposed to be defined as :

$$f(x,y) = \int_{\mathbb{R}^2} f^*(x,y,u_1,u_2)g(u_1,u_2)du_1du_2 \tag{A.10}$$

where $g(u_1,u_2)$ is a probability distribution function, $f^*(x,y,u_1,u_2)$ is called simulator and denotes the function from which the mean value at some draws u_1 and u_2 gives an approximation of the overall likelihood. Without lost of generality, we only define the two dimensions case that can be generalized to fewer or larger dimensions integral. For this kind of likelihood function, Gouriéroux and Monfort (1993) proposed as simulator the function $f^*(x,y,u_1,u_2)$ with u_1 and u_2 drawn from the same probability distribution function g (the probability distribution function of the individual random effects). Then the overall likelihood function can be approximated by :

$$f(x,y) = \frac{1}{D}\sum_{d=1}^{D} f^*(x,y,u_{1d},u_{2d}) \tag{A.11}$$

Where D denotes the number of draws.

To implement this method, we start by simulating a bivariate normal draw $N(0,I_2)$ and we give them the (u_1,u_2) covariance matrix structure. Then we calculate the value of the simulator at these transformed draws and we repeat D times. The overall likelihood is the mean of the simulator value at each transformed draw. At each iteration, once the random effects covariance matrix is calculated, we apply it to the simulated first normal draws to transform them in draws of the random effects and use them to calculate the likelihood. This process is repeated until convergence.

The simulated likelihood estimator is consistent and asymptotically equivalent to the likelihood estimator (Gouriéroux and Monfort, 1993) if the number of draws tend to infinity faster than $\sqrt{N}$.

A.2.3 GHQ or MSL : what method to choose ?

As described above, they are two major methods to estimate our likelihood function. To choose which method to implement, we deal with the accuracy and the computing time requirement.
For our estimations, we choose the adaptative Gauss-Hermite quadrature proposed by Liu and Pierce (1994) for three main reasons.

- Our dataset is an unbalanced panel data with 10,311 individuals observed in mean over 26 years, that leads 272,465 observations. Due to the fact that the simulated likelihood method requires that the number of draw D be larger than the square of the number of observations, we do not use it to avoid waste of time in computing process.
- The Gauss-Hermite quadrature requires that we find the best number of quadrature Q that is the one for whom the integrand can be well approximated by an $2Q+1$ order polynomial. If Q is small, that reduces computing time. For our estimations, that are achieved in general for Q between 8 and 14. It means

that at each iteration, for the likelihood value calculation, we do a weighted sum of between $8^2 = 64$ and $14^2 = 196$ terms.

- Using the Gauss-Hermite quadrature method reduces computing time but this computing time remains very long if the integrand is not sampled at the suitable range (meaning that the adaptative method has not been used). And in this case, the maximization process spends between two and three weeks before achieving convergence on an Intel Core i7 computer at 3.4 GHz with 8 GB of RAM memory. By applying the adaptative Gauss-Hermite quadrature, the computing time is significatively reduced and then, we spend between two and three days for achieving convergence on the same computer.

Note that the reduced convergence time mentioned above is in part due to the implementation of the first order derivatives of the likelihood function. Using the overall log-likelihood approximated by the Liu and Pierce adaptative Gauss-Hermite quadrature method, we can get derivatives with respect of all model parameters. The implementation of these derivative in the maximization process allows us to used the Stata's d1 method. The convergence time saved by this method is clearly enormous. On our overall data set, with 8 quadratures points, when we use a non adaptative quadrature method, the convergence is not achieved : after 3 weeks of computation, the model underflows. When we use the Liu and Pierce adaptative Gauss-Hermite quadrature, but without implementing the first order derivatives, the estimation process takes 11 days and 10 hours to achieve convergence. When we use the Liu and Pierce adaptative Gauss-Hermite quadrature with implemented the first order derivatives, the estimation process achieve convergence only after 1 day and 17 hours, clearly faster ...

A.3 Chosen method requirements

In this section we describe some requirements of the selected method that is the adaptative Gauss-Hermite Quadrature. The first one is the fact that the adaptative Gauss-Hermite quadrature requires to derive the hessian of the log of the integrand (Liu and Pierce, 1994). The second one is that we derive the gradient of the overall likelihood function in order to use Stata's d1 method (see Gould et alii, 2010) for more accuracy and more speed in the calculations.

A.3.1 Gradient vector calculation

The gradient of the overall log-likelihood function has been calculated to speed up the maximization process. This will allow us to use the Stata's d1 method that requires the implementation of the gradient vector in addition to the overall log-likelihood. The overall likelihood function for an individual i is :

$$L_i = \int_{\mathbb{R}^2} \Phi_2(q_{i0}^1 h_i^0, q_{i0}^2 w_i^0, q_{i0}^1 q_{i0}^2 \rho_\epsilon) \prod_{t=2}^{T_i} \Phi_2(q_{it}^1 \bar{h}_{it}, q_{it}^2 \bar{w}_{it}, q_{it}^1 q_{it}^2 \rho_\zeta) \phi(\eta_i, \Sigma_\eta) d\eta_i^1 d\eta_i^2 \tag{A.12}$$

Where

$$
\begin{aligned}
q_{it}^1 &= 2y_{it}^1 - 1 \ \ \forall\, i,t \\
q_{it}^2 &= 2y_{it}^2 - 1 \ \ \forall\, i,t \\
h_i^0 &= Z_i^1\gamma_1 + \lambda_{11}\eta_i^1 + \lambda_{12}\eta_i^2 \\
w_i^0 &= Z_i^2\gamma_2 + \lambda_{21}\eta_i^1 + \lambda_{22}\eta_i^2 \\
\bar{h}_{it} &= X_{it}^1\beta_1 + \delta_{11}h_{i,t-1} + \delta_{12}w_{i,t-1} + \eta_i^1 \\
\bar{w}_{it} &= X_{it}^2\beta_2 + \delta_{21}h_{i,t-1} + \delta_{22}w_{i,t-1} + \eta_i^2
\end{aligned}
$$

Using the Liu and Pierce adaptative Gauss-Hermite quadrature method, the overall likelihood function is given by (we use the same notation that those used in section A.2) :

$$
L_i = \sum_{k=1,j=1}^{Q} w_k^* w_j^* \Phi_2(q_{i0}^1 h_i^0, q_{i0}^2 w_i^0, q_{i0}^1 q_{i0}^2 \rho_\epsilon) \prod_{t=2}^{T_i} \Phi_2(q_{it}^1 \bar{h}_{it}, q_{it}^2 \bar{w}_{it}, q_{it}^1 q_{it}^2 \rho_\zeta) \phi(\eta_i, \Sigma_\eta) \Big|_{\eta_i^1 = x_k^*, \eta_i^2 = x_j^*} \tag{A.13}
$$

To get the gradient vector, the log-likelihood above must be derive with respect to 13 parameters that are : $\bar{\beta}_1 = (\beta_1, \delta_{11}, \delta_{12})'$, $\bar{\beta}_2 = (\beta_2, \delta_{21}, \delta_{22})'$, γ_1, γ_2, λ_{11}, λ_{12}, λ_{21}, λ_{22}, σ_1, σ_2, ρ_η, ρ_ζ, and ρ_ϵ.
Let's l_{kj} denote :

$$
l_{kj} = \Phi_2(q_{i0}^1 h_i^0, q_{i0}^2 w_i^0, q_{i0}^1 q_{i0}^2 \rho_\epsilon) \prod_{t=2}^{T_i} \Phi_2(q_{it}^1 \bar{h}_{it}, q_{it}^2 \bar{w}_{it}, q_{it}^1 q_{it}^2 \rho_\zeta) \phi(\eta_i, \Sigma_\eta) \Big|_{\eta_i^1 = x_k^*, \eta_i^2 = x_j^*}
$$

Then the first order derivatives with respect to each α of the 13 parameters is given by :

$$
\frac{\partial log(L_i)}{\partial \alpha} = \sum_{k=1,j=1}^{Q} \frac{\partial l_{kj}/\partial \alpha}{L_i}
$$

With respect to $\bar{\beta}_1$ the first order derivative is :

$$
\frac{\partial l_{kj}}{\partial \bar{\beta}_1} = l_{kj} \sum_{t=2}^{T_i} \frac{q_{it}^1 \phi(q_{it}^1 \bar{h}_{it}) \Phi_1(\frac{q_{it}^2 \bar{w}_{it} - q_{it}^2 \rho_\zeta \bar{h}_{it}}{\sqrt{1-\rho_\zeta^2}})}{\Phi_2(q_{it}^1 \bar{h}_{it}, q_{it}^2 \bar{w}_{it}, q_{it}^1 q_{it}^2 \rho_\zeta)}
$$

With respect to $\bar{\beta}_2$ the first order derivative is :

$$
\frac{\partial l_{kj}}{\partial \bar{\beta}_2} = l_{kj} \sum_{t=2}^{T_i} \frac{q_{it}^2 \phi(q_{it}^2 \bar{w}_{it}) \Phi_1(\frac{q_{it}^1 \bar{h}_{it} - q_{it}^1 \rho_\zeta \bar{w}_{it}}{\sqrt{1-\rho_\zeta^2}})}{\Phi_2(q_{it}^1 \bar{h}_{it}, q_{it}^2 \bar{w}_{it}, q_{it}^1 q_{it}^2 \rho_\zeta)}
$$

With respect to γ_1 the first order derivative is :

$$
\frac{\partial l_{kj}}{\partial \gamma_1} = l_{kj} \frac{q_{i0}^1 \phi(q_{i0}^1 h_i^0) \Phi_1(\frac{q_{i0}^2 w_i^0 - q_{i0}^2 \rho_\epsilon h_i^0}{\sqrt{1-\rho_\epsilon^2}})}{\Phi_2(q_{i0}^1 h_i^0, q_{i0}^2 w_i^0, q_{i0}^1 q_{i0}^2 \rho_\epsilon)}
$$

With respect to γ_2 the first order derivative is :

$$
\frac{\partial l_{kj}}{\partial \gamma_2} = l_{kj} \frac{q_{i0}^2 \phi(q_{i0}^2 w_i^0) \Phi_1(\frac{q_{i0}^1 h_i^0 - q_{i0}^1 \rho_\epsilon w_i^0}{\sqrt{1-\rho_\epsilon^2}})}{\Phi_2(q_{i0}^1 h_i^0, q_{i0}^2 w_i^0, q_{i0}^1 q_{i0}^2 \rho_\epsilon)}
$$

With respect to λ_{11} the first order derivative is :

$$\frac{\partial l_{kj}}{\partial \lambda_{11}} = l_{kj} \frac{q^1_{i0} x^*_k \phi(q^1_{i0} h^0_i) \Phi_1(\frac{q^2_{i0} w^0_i - q^2_{i0} \rho_\epsilon h^0_i}{\sqrt{1-\rho^2_\epsilon}})}{\Phi_2(q^1_{i0} h^0_i, q^2_{i0} w^0_i, q^1_{i0} q^2_{i0} \rho_\epsilon)}$$

With respect to λ_{12} the first order derivative is :

$$\frac{\partial l_{kj}}{\partial \lambda_{12}} = l_{kj} \frac{q^1_{i0} x^*_j \phi(q^1_{i0} h^0_i) \Phi_1(\frac{q^2_{i0} w^0_i - q^2_{i0} \rho_\epsilon h^0_i}{\sqrt{1-\rho^2_\epsilon}})}{\Phi_2(q^1_{i0} h^0_i, q^2_{i0} w^0_i, q^1_{i0} q^2_{i0} \rho_\epsilon)}$$

With respect to λ_{21} the first order derivative is :

$$\frac{\partial l_{kj}}{\partial \lambda_{21}} = l_{kj} \frac{q^2_{i0} x^*_k \phi(q^2_{i0} w^0_i) \Phi_1(\frac{q^1_{i0} h^0_i - q^1_{i0} \rho_\epsilon w^0_i}{\sqrt{1-\rho^2_\epsilon}})}{\Phi_2(q^1_{i0} h^0_i, q^2_{i0} w^0_i, q^1_{i0} q^2_{i0} \rho_\epsilon)}$$

With respect to λ_{22} the first order derivative is :

$$\frac{\partial l_{kj}}{\partial \lambda_{22}} = l_{kj} \frac{q^2_{i0} x^*_j \phi(q^2_{i0} w^0_i) \Phi_1(\frac{q^1_{i0} h^0_i - q^1_{i0} \rho_\epsilon w^0_i}{\sqrt{1-\rho^2_\epsilon}})}{\Phi_2(q^1_{i0} h^0_i, q^2_{i0} w^0_i, q^1_{i0} q^2_{i0} \rho_\epsilon)}$$

With respect to σ_1 the first order derivative is :

$$\frac{\partial l_{kj}}{\partial log(\sigma_1)} = l_{kj} * \left(-1 + \frac{(x^*_k/\sigma_1)^2 - \rho_\eta x^*_k x^*_j/(\sigma_1 \sigma_2)}{1-\rho^2_\eta} \right)$$

With respect to σ_2 the first order derivative is :

$$\frac{\partial l_{kj}}{\partial log(\sigma_2)} = l_{kj} * \left(-1 + \frac{(x^*_j/\sigma_2)^2 - \rho_\eta x^*_k x^*_j/(\sigma_1 \sigma_2)}{1-\rho^2_\eta} \right)$$

With respect to ρ_η the first order derivative is :

$$\frac{\partial l_{kj}}{\partial log(\frac{1+\rho_\eta}{1-\rho_\eta})^{1/2}} = l_{kj} * \left(\rho_\eta - \frac{\rho_\eta((x^*_k/\sigma_1)^2 + (x^*_j/\sigma_2)^2) - (1+\rho^2_\eta) x^*_k x^*_j/(\sigma_1 \sigma_2)}{1-\rho^2_\eta} \right)$$

With respect to ρ_ζ the first order derivative is :

$$\frac{\partial l_{kj}}{\partial log(\frac{1+\rho_\zeta}{1-\rho_\zeta})^{1/2}} = l_{kj} \sum_{t=2}^{T_i} \frac{q^1_{it} q^2_{it} \phi(q^1_{it} \bar{h}_{it}, q^2_{it} \bar{w}_{it}, q^1_{it} q^2_{it} \rho_\zeta)}{\Phi_2(q^1_{it} \bar{h}_{it}, q^2_{it} \bar{w}_{it}, q^1_{it} q^2_{it} \rho_\zeta)}$$

With respect to ρ_ϵ the first order derivative is :

$$\frac{\partial l_{kj}}{\partial log(\frac{1+\rho_\epsilon}{1-\rho_\epsilon})^{1/2}} = l_{kj} \frac{q^1_{i0} q^2_{i0} \phi(q^1_{i0} h^0_i, q^2_{i0} w^0_i, q^1_{i0} q^2_{i0} \rho_\epsilon)}{\Phi_2(q^1_{i0} h^0_i, q^2_{i0} w^0_i, q^1_{i0} q^2_{i0} \rho_\epsilon)}$$

Remarks :

- For σ_1, σ_2, ρ_η, ρ_ζ, and ρ_ϵ, we used some transformations on parameters to insure that in the maximization process, all σ remain positive and all ρ between -1 and 1 at all iteration. For σ we use exponential transformation then in the derivation, we derive with respect to $log(\sigma)$. For ρ we use arctangency transformation (i.e. $\frac{exp(2\rho)-1}{exp(2\rho)+1}$) then in the derivation, we derive with respect to $log\left(\frac{1+\rho}{1-\rho}\right)^{1/2}$.

- To easily derive a bivariate normal probability with zero mean, variance one and correlation ρ, we can transform it into an integral that integrand is a product of an univariate normal density and an univariate normal probability as follows :

$$\Phi_2(x,y,\rho) = \int_{-\infty}^{y} \phi(v)\Phi\left(\frac{x-\rho v}{\sqrt{1-\rho^2}}\right)dv = \int_{-\infty}^{x} \phi(u)\Phi\left(\frac{y-\rho u}{\sqrt{1-\rho^2}}\right)du.$$

- Given the transformation above, the first order derivatives of $\Phi_2(x,y,\rho)$ with respect to x and y are respectively given by :

$$\frac{\partial\Phi_2(x,y,\rho)}{\partial x} = \phi(x)\Phi\left(\frac{y-\rho x}{\sqrt{1-\rho^2}}\right)$$
$$\frac{\partial\Phi_2(x,y,\rho)}{\partial y} = \phi(y)\Phi\left(\frac{x-\rho y}{\sqrt{1-\rho^2}}\right)$$

A.3.2 Hessian matrix calculation

For the requirement of the adaptative Gauss-Hermite quadrature method, we need to derive the Hessian matrix of the log of the integrand function with respect to the random effects vector. In this section, $\phi(x)$ denotes the univariate normal density function, $\phi(x,y,\rho)$ denote the bivariate normal density with correlation ρ, $\Phi_1(x)$ denote the univariate normal probability function, and $\Phi_2(x,y,\rho)$ denote the bivariate normal probability function with correlation ρ.

The individual likelihood function is defined as follows :

$$L_i = \int_{\mathbb{R}^2} \Phi_2(q_{i0}^1 h_i^0, q_{i0}^2 w_i^0, q_{i0}^1 q_{i0}^2 \rho_\epsilon) \prod_{t=2}^{T_i} \Phi_2(q_{it}^1 \bar{h}_{it}, q_{it}^2 \bar{w}_{it}, q_{it}^1 q_{it}^2 \rho_\zeta)\phi(\eta_i, \Sigma_\eta) d\eta_i^1 d\eta_i^2 \tag{A.14}$$

Where

$$\begin{aligned}
q_{it}^1 &= 2y_{it}^1 - 1 \ \ \forall\, i,t \\
q_{it}^2 &= 2y_{it}^2 - 1 \ \ \forall\, i,t \\
h_i^0 &= Z_i^1\gamma_1 + \lambda_{11}\eta_i^1 + \lambda_{12}\eta_i^2 \\
w_i^0 &= Z_i^2\gamma_2 + \lambda_{21}\eta_i^1 + \lambda_{22}\eta_i^2 \\
\bar{h}_{it} &= X_{it}^1\beta_1 + \delta_{11}h_{i,t-1} + \delta_{12}w_{i,t-1} + \eta_i^1 \\
\bar{w}_{it} &= X_{it}^2\beta_2 + \delta_{21}h_{i,t-1} + \delta_{22}w_{i,t-1} + \eta_i^2
\end{aligned}$$

where the log of the integrand is $log(g(\eta_i^1,\eta_i^2)) = log\left(\Phi_2(q_{i0}^1 h_i^0, q_{i0}^2 w_i^0, q_{i0}^1 q_{i0}^2 \rho_\epsilon)\prod_{t=2}^{T_i} \Phi_2(q_{it}^1 \bar{h}_{it}, q_{it}^2 \bar{w}_{it}, q_{it}^1 q_{it}^2 \rho_\zeta) g(\eta_i, \Sigma_\eta)\right)$. We derive from this function the Hessian matrix by calculating $-\frac{\delta^2}{\delta(\eta_i^1)^2}log(g(\eta_i^1,\eta_i^2))$, $-\frac{\delta^2}{\delta(\eta_i^2)^2}log(g(\eta_i^1,\eta_i^2))$ and $-\frac{\delta^2}{\delta\eta_i^1\delta\eta_i^1}log(g(\eta_i^1,\eta_i^2))$.

The first order derivatives are given by :

$$-\frac{\partial}{\partial \eta_i} log(g) = -\frac{\Phi_2^{'}(q_{i0}^1 h_i^0, q_{i0}^2 w_i^0, q_{i0}^1 q_{i0}^2 \rho_\epsilon)}{\Phi_2(q_{i0}^1 h_i^0, q_{i0}^2 w_i^0, q_{i0}^1 q_{i0}^2 \rho_\epsilon)} - \sum_{t=2}^{T_i} \frac{\Phi_2^{'}(q_{it}^1 \bar{h}_{it}, q_{it}^2 \bar{w}_{it}, q_{it}^1 q_{it}^2 \rho_\zeta)}{\Phi_2(q_{it}^1 \bar{h}_{it}, q_{it}^2 \bar{w}_{it}, q_{it}^1 q_{it}^2 \rho_\zeta)} + \frac{\eta_i^1/\sigma_1^2 - \rho \eta_i^2/(\sigma_1 \sigma_2)}{1-\rho_\eta^2} \tag{A.15}$$

With respect to η_i^1 we have :

$$\begin{aligned}
\Phi^{'}_{2\eta_i^1}(q_{i0}^1 h_i^0, q_{i0}^2 w_i^0, q_{i0}^1 q_{i0}^2 \rho_\epsilon) &= q_{i0}^1 \lambda_{11} \phi(q_{i0}^1 h_i^0) \Phi_1\left(\frac{q_{i0}^2 w_i^0 - q_{i0}^2 \rho_\epsilon h_i^0}{\sqrt{1-\rho_\epsilon^2}}\right) \\
&\quad + q_{i0}^2 \lambda_{21} \phi(q_{i0}^2 w_i^0) \Phi_1\left(\frac{q_{i0}^1 h_i^0 - q_{i0}^1 \rho_\epsilon w_i^0}{\sqrt{1-\rho_\epsilon^2}}\right) \\
\Phi^{'}_{2\eta_i^1}(q_{it}^1 \bar{h}_{it}, q_{it}^2 \bar{w}_{it}, q_{it}^1 q_{it}^2 \rho_\zeta) &= q_{it}^1 \phi(q_{it}^1 \bar{h}_{it}) \Phi_1\left(\frac{q_{it}^2 \bar{w}_{it} - q_{it}^2 \rho_\zeta \bar{h}_{it}}{\sqrt{1-\rho_\zeta^2}}\right)
\end{aligned}$$

And with respect to η_i^2 we have :

$$\begin{aligned}
\Phi^{'}_{2\eta_i^2}(q_{i0}^1 h_i^0, q_{i0}^2 w_i^0, q_{i0}^1 q_{i0}^2 \rho_\epsilon) &= q_{i0}^1 \lambda_{12} \phi(q_{i0}^1 h_i^0) \Phi_1\left(\frac{q_{i0}^2 w_i^0 - q_{i0}^2 \rho_\epsilon h_i^0}{\sqrt{1-\rho_\epsilon^2}}\right) \\
&\quad + q_{i0}^2 \lambda_{22} \phi(q_{i0}^2 w_i^0) \Phi_1\left(\frac{q_{i0}^1 h_i^0 - q_{i0}^1 \rho_\epsilon w_i^0}{\sqrt{1-\rho_\epsilon^2}}\right) \\
\Phi^{'}_{2\eta_i^2}(q_{it}^1 \bar{h}_{it}, q_{it}^2 \bar{w}_{it}, q_{it}^1 q_{it}^2 \rho_\zeta) &= \phi(q_{it}^2 \bar{w}_{it}) \Phi_1\left(\frac{q_{it}^1 \bar{h}_{it} - q_{it}^1 \rho_\zeta \bar{w}_{it}}{\sqrt{1-\rho_\zeta^2}}\right)
\end{aligned}$$

The second order derivatives are given by :

$$\begin{aligned}
-\frac{\partial^2}{\partial (\eta_i^1)^2} log(g) = &- \frac{\Phi^{''}_{2\eta_i^1}(q_{i0}^1 h_i^0, q_{i0}^2 w_i^0, q_{i0}^1 q_{i0}^2 \rho_\epsilon) \Phi_2(q_{i0}^1 h_i^0, q_{i0}^2 w_i^0, q_{i0}^1 q_{i0}^2 \rho_\epsilon)}{\Phi_2^2(q_{i0}^1 h_i^0, q_{i0}^2 w_i^0, q_{i0}^1 q_{i0}^2 \rho_\epsilon)} \\
&+ \frac{\Phi^{'2}_{2\eta_i^1}(q_{i0}^1 h_i^0, q_{i0}^2 w_i^0, q_{i0}^1 q_{i0}^2 \rho_\epsilon)}{\Phi_2^2(q_{i0}^1 h_i^0, q_{i0}^2 w_i^0, q_{i0}^1 q_{i0}^2 \rho_\epsilon)} \\
&- \sum_{t=2}^{T_i} \left(\frac{\Phi^{''}_{2\eta_i^1}(q_{it}^1 \bar{h}_{it}, q_{it}^2 \bar{w}_{it}, q_{it}^1 q_{it}^2 \rho_\zeta) \Phi_2(q_{it}^1 \bar{h}_{it}, q_{it}^2 \bar{w}_{it}, q_{it}^1 q_{it}^2 \rho_\zeta)}{\Phi_2^2(q_{it}^1 \bar{h}_{it}, q_{it}^2 \bar{w}_{it}, q_{it}^1 q_{it}^2 \rho_\zeta)} \right. \\
&\left. - \frac{\Phi^{'2}_{2\eta_i^1}(q_{it}^1 \bar{h}_{it}, q_{it}^2 \bar{w}_{it}, q_{it}^1 q_{it}^2 \rho_\zeta)}{\Phi_2^2(q_{it}^1 \bar{h}_{it}, q_{it}^2 \bar{w}_{it}, q_{it}^1 q_{it}^2 \rho_\zeta)} \right) + \frac{1}{\sigma_1^2(1-\rho_\eta^2)}
\end{aligned} \tag{A.16}$$

$$\begin{aligned}
-\frac{\partial^2}{\partial (\eta_i^2)^2} log(g) = &- \frac{\Phi^{''}_{2\eta_i^2}(q_{i0}^1 h_i^0, q_{i0}^2 w_i^0, q_{i0}^1 q_{i0}^2 \rho_\epsilon) \Phi_2(q_{i0}^1 h_i^0, q_{i0}^2 w_i^0, q_{i0}^1 q_{i0}^2 \rho_\epsilon)}{\Phi_2^2(q_{i0}^1 h_i^0, q_{i0}^2 w_i^0, q_{i0}^1 q_{i0}^2 \rho_\epsilon)} \\
&+ \frac{\Phi^{'2}_{2\eta_i^2}(q_{i0}^1 h_i^0, q_{i0}^2 w_i^0, q_{i0}^1 q_{i0}^2 \rho_\epsilon)}{\Phi_2^2(q_{i0}^1 h_i^0, q_{i0}^2 w_i^0, q_{i0}^1 q_{i0}^2 \rho_\epsilon)} \\
&- \sum_{t=2}^{T_i} \left(\frac{\Phi^{''}_{2\eta_i^2}(q_{it}^1 \bar{h}_{it}, q_{it}^2 \bar{w}_{it}, q_{it}^1 q_{it}^2 \rho_\zeta) \Phi_2(q_{it}^1 \bar{h}_{it}, q_{it}^2 \bar{w}_{it}, q_{it}^1 q_{it}^2 \rho_\zeta)}{\Phi_2^2(q_{it}^1 \bar{h}_{it}, q_{it}^2 \bar{w}_{it}, q_{it}^1 q_{it}^2 \rho_\zeta)} \right. \\
&\left. - \frac{\Phi^{'2}_{2\eta_i^2}(q_{it}^1 \bar{h}_{it}, q_{it}^2 \bar{w}_{it}, q_{it}^1 q_{it}^2 \rho_\zeta)}{\Phi_2^2(q_{it}^1 \bar{h}_{it}, q_{it}^2 \bar{w}_{it}, q_{it}^1 q_{it}^2 \rho_\zeta)} \right) + \frac{1}{\sigma_2^2(1-\rho_\eta^2)}
\end{aligned} \tag{A.17}$$

$$
\begin{aligned}
-\frac{\partial^2}{\partial\eta_i^1\delta\eta_i^2}log(g) = & -\frac{\Phi^{''}_{2\eta_i^1\eta_i^2}(q_{i0}^1h_i^0,q_{i0}^2w_i^0,q_{i0}^1q_{i0}^2\rho_\epsilon)\Phi_2(q_{i0}^1h_i^0,q_{i0}^2w_i^0,q_{i0}^1q_{i0}^2\rho_\epsilon)}{\Phi_2^2(q_{i0}^1h_i^0,q_{i0}^2w_i^0,q_{i0}^1q_{i0}^2\rho_\epsilon)} \\
& +\frac{\Phi^{'}_{2\eta^1}(q_{i0}^1h_i^0,q_{i0}^2w_i^0,q_{i0}^1q_{i0}^2\rho_\epsilon)\Phi^{'}_{2\eta^2}(q_{i0}^1h_i^0,q_{i0}^2w_i^0,q_{i0}^1q_{i0}^2\rho_\epsilon)}{\Phi_2^2(q_{i0}^1h_i^0,q_{i0}^2w_i^0,q_{i0}^1q_{i0}^2\rho_\epsilon)} \\
& -\sum_{t=2}^{T_i}\Bigg(\frac{\Phi^{''}_{2\eta_i^1\eta_i^2}(q_{it}^1\bar{h}_{it},q_{it}^2\bar{w}_{it},q_{it}^1q_{it}^2\rho_\zeta)\Phi_2(q_{it}^1\bar{h}_{it},q_{it}^2\bar{w}_{it},q_{it}^1q_{it}^2\rho_\zeta)}{\Phi_2^2(q_{it}^1\bar{h}_{it},q_{it}^2\bar{w}_{it},q_{it}^1q_{it}^2\rho_\zeta)} \\
& -\frac{\Phi^{'}_{2\eta^1}(q_{it}^1\bar{h}_{it},q_{it}^2\bar{w}_{it},q_{it}^1q_{it}^2\rho_\zeta)\Phi^{'}_{2\eta^2}(q_{it}^1\bar{h}_{it},q_{it}^2\bar{w}_{it},q_{it}^1q_{it}^2\rho_\zeta)}{\Phi_2^2(q_{it}^1\bar{h}_{it},q_{it}^2\bar{w}_{it},q_{it}^1q_{it}^2\rho_\zeta)}\Bigg)-\frac{\rho_\eta}{\sigma_1\sigma_2(1-\rho_\eta^2)}
\end{aligned}
\tag{A.18}
$$

Where

$$
\Phi^{''}_{2\eta_i^1}(q_{it}^1\bar{h}_{it},q_{it}^2\bar{w}_{it},q_{it}^1q_{it}^2\rho_\zeta)=-\bar{h}_{it}\Phi^{'}_{2\eta_i^1}(q_{it}^1\bar{h}_{it},q_{it}^2\bar{w}_{it},q_{it}^1q_{it}^2\rho_\zeta)-\rho_\zeta\phi_{\eta_i^1}(q_{it}^1\bar{h}_{it},q_{it}^2\bar{w}_{it},q_{it}^1q_{it}^2\rho_\zeta) \tag{A.19}
$$

$$
\Phi^{''}_{2\eta_i^2}(q_{it}^1\bar{h}_{it},q_{it}^2\bar{w}_{it},q_{it}^1q_{it}^2\rho_\zeta)=-\bar{w}_{it}\Phi^{'}_{2\eta_i^2}(q_{it}^1\bar{h}_{it},q_{it}^2\bar{w}_{it},q_{it}^1q_{it}^2\rho_\zeta)-\rho_\zeta\phi_{\eta_i^1}(q_{it}^1\bar{h}_{it},q_{it}^2\bar{w}_{it},q_{it}^1q_{it}^2\rho_\zeta) \tag{A.20}
$$

$$
\Phi^{''}_{2\eta_i^1\eta_i^2}(q_{it}^1\bar{h}_{it},q_{it}^2\bar{w}_{it},q_{it}^1q_{it}^2\rho_\zeta)=q_{it}^1q_{it}^2\rho\zeta\phi_{\eta_i^1}(q_{it}^1\bar{h}_{it},q_{it}^2\bar{w}_{it},q_{it}^1q_{it}^2\rho_\zeta) \tag{A.21}
$$

$$
\begin{aligned}
\Phi^{''}_{2\eta_i^1}(q_{i0}^1h_i^0,q_{i0}^2w_i^0,q_{i0}^1q_{i0}^2\rho_\epsilon) = & (2\lambda_{11}\lambda_{21}-\rho_\epsilon(\lambda_{11}^2+\lambda_{21}^2))\phi(q_{i0}^1h_i^0,q_{i0}^2w_i^0,q_{i0}^1q_{i0}^2\rho_\epsilon) \\
& -\lambda_{11}^2h_i^0\phi(q_{i0}^1h_i^0)\Phi_1\Bigg(\frac{q_{i0}^2w_i^0-\rho_\epsilon q_{i0}^2h_i^0}{\sqrt{1-\rho_\epsilon^2}}\Bigg) \\
& -\lambda_{21}^2w_i^0\phi(q_{i0}^2w_i^0)\Phi_1\Bigg(\frac{q_{i0}^1h_i^0-\rho_\epsilon q_{i0}^1w_i^0}{\sqrt{1-\rho_\epsilon^2}}\Bigg)
\end{aligned}
$$

$$
\begin{aligned}
\Phi^{''}_{2\eta_i^2}(q_{i0}^1h_i^0,q_{i0}^2w_i^0,q_{i0}^1q_{i0}^2\rho_\epsilon) = & (2\lambda_{12}\lambda_{22}-\rho_\epsilon(\lambda_{12}^2+\lambda_{22}^2))\phi(q_{i0}^1h_i^0,q_{i0}^2w_i^0,q_{i0}^1q_{i0}^2\rho_\epsilon) \\
& -\lambda_{12}^2h_i^0\phi(q_{i0}^1h_i^0)\Phi_1(\frac{q_{i0}^2w_i^0-\rho_\epsilon q_{i0}^2h_i^0}{\sqrt{1-\rho_\epsilon^2}}) \\
& -\lambda_{22}^2w_i^0\phi(q_{i0}^2w_i^0)\Phi_1(\frac{q_{i0}^1h_i^0-\rho_\epsilon q_{i0}^1w_i^0}{\sqrt{1-\rho_\epsilon^2}})
\end{aligned}
$$

$$
\begin{aligned}
\Phi^{''}_{2\eta_i^1\eta_i^1}(q_{i0}^1h_i^0,q_{i0}^2w_i^0,q_{i0}^1q_{i0}^2\rho_\epsilon) = & q_{i0}^1q_{i0}^2(\lambda_{11}\lambda_{22}+\lambda_{12}\lambda_{21}-\rho_\epsilon(\lambda_{11}\lambda_{12}+\lambda_{21}\lambda_{22}))* \\
& \phi_2(q_{i0}^1h_i^0,q_{i0}^2w_i^0,q_{i0}^1q_{i0}^2\rho_\epsilon)-\lambda_{11}\lambda_{12}h_i^0\phi(q_{i0}^1h_i^0)\Phi_1\Bigg(\frac{q_{i0}^2w_i^0-\rho_\epsilon q_{i0}^2h_i^0}{\sqrt{1-\rho_\epsilon^2}}\Bigg) \\
& -\lambda_{21}\lambda_{22}w_i^0\phi(q_{i0}^2w_i^0)\Phi_1\Bigg(\frac{q_{i0}^1h_i^0-\rho_\epsilon q_{i0}^1w_i^0}{\sqrt{1-\rho_\epsilon^2}}\Bigg)
\end{aligned}
$$

Then, the Hessian matrix is given by :

$$
H=\begin{pmatrix} -\frac{\delta^2}{\delta(\eta_i^1)^2}log(g) & -\frac{\delta^2}{\delta\eta_i^1\delta\eta_i^2}log(g) \\ -\frac{\delta^2}{\delta\eta_i^1\delta\eta_i^2}log(g) & -\frac{\delta^2}{\delta(\eta_i^2)^2}log(g)\end{pmatrix} \tag{A.22}
$$

As described in section A.2.1, after having derived this Hessian matrix, we calculate its value at the mode of the integrand and use it to resample the integrand.

A.4 Robustness analysis based on simulations

This section aims to insure that the implemented method gives suitable results. We consider that the implemented method give us suitable results if for a given relationship between variables, by applying the estimation method on these variables we find approximatively the same coefficients. To reach this goal, we perform a robustness analysis on the estimation method. This robustness analysis is an empirical one based on simulations. We use two different approaches for that.

The first approach is to simulate bivariate binary variables by specifying a relationship between some explanatory variables (it means that we fix coefficients of explanatory variables) and estimate this relationship with the implemented method in order to compare the results with the relationship specified before. In the second approach, we introduce new variables (that were not used in the data generating process) when estimating the relationship with the implemented method and compare the new results with the first ones. The implemented method is robust when it is able to correctly estimate the relationship specified even if we introduce other variables and also to estimate non significant coefficients to those other variables. Finally, the method we make use of to check for the robustness is the same that in Miranda (2011).

As the estimation method implemented is a numerical approximation method, the results will depend on the selected number of quadrature points. We deal with the incidence of number of quadrature points on results in the last part of this section. For a better analysis of the results we also add the standard errors of each estimated coefficients.

A.4.1 Simulated relationship between real variables

In this section, we use variables from the French SIP (Santé et Itinéraire Professionnel) survey data set and we simulate error terms and a relationship between some selected variables. The subset of the database use for this section is an unbalanced panel of 1202 individuals with total waves per individual between 5 and 10 waves.

We fix the error terms parameters as $\sigma_1 = 2.1$, $\sigma_2 = 3.1$, $\rho_\eta = 0.7$, $\rho_\zeta = 0.5$ and $\rho_\epsilon = 0.4$.

We simulate idiosyncratic errors vectors $\zeta = (\zeta_1, \zeta_2)'$ and $\epsilon = (\epsilon_1, \epsilon_2)'$ as bivariate normal variables with zero mean, variance equal to 1 and covariances respectively equal to ρ_ζ and ρ_ϵ. We also simulate individual random effects as bivariate normal variables with zero mean, covariance equals to ρ_η and variance equals to σ_1^2 for the first component of the random effects vector and equals to σ_2^2 for the second component of the random effects vector. It has been done as follows :

$$\epsilon_1 = rnormal()$$
$$\epsilon_2 = rnormal() * \sqrt{1-\rho_\epsilon^2} + \rho_\epsilon \epsilon_1$$
$$\zeta_1 = rnormal()$$
$$\zeta_2 = rnormal() * \sqrt{1-\rho_\zeta^2} + \rho_\zeta \zeta_1$$

As individuals effects are time invariant, we simulate η as follows :

$$\eta_1 = rnormal(0, \sigma_1)\, if\ date = 1$$
$$\eta_2 = rnormal(0, \sigma_2) * \sqrt{1-\rho_\eta^2} + \rho_\eta \frac{\sigma_2}{\sigma_1}\eta_1\, if\ date = 1$$
$$\eta_1 = \eta_1[1]\, if\ date \neq 1$$
$$\eta_2 = \eta_2[1]\, if\ date \neq 1$$

Where $rnormal(\mu, \sigma)$ denote the random normal density with mean μ and standard deviation σ and $rnormal()$ denote the random normal density with mean zero and standard deviation 1.

For the initial condition, the simulated relationship is :

$$y_1^* = -0.2 + 0.3illness - 0.2unemployment + 0.4\eta_1 - 0.5\eta_2 + \epsilon_1$$
$$y_2^* = 2 - 0.2illness - 0.08age + 0.3\eta_1 + 0.5\eta_2 + \epsilon_2$$
$$y_1 = \mathbb{I}(y_1^* > 0)$$
$$y_2 = \mathbb{I}(y_2^* > 0)$$

For t>1, we specify the following relationship :

$$y_{1t}^* = 1.9 + 0.3y_{1,t-1} + 0.1y_{2,t-1} - 0.05Male_t - 0.2unemployment_t + \eta_1 + \zeta_{1t}$$
$$y_{2t}^* = -0.4 - 0.1y_{1,t-1} + 0.4y_{2,t-1} + 0.05Male_t - 0.5density_t + \eta_2 + \zeta_{2t}$$
$$y_{1t} = \mathbb{I}(y_{1t}^* > 0)$$
$$y_{2t} = \mathbb{I}(y_{2t}^* > 0)$$

Estimation results for 16 quadrature points are displayed in table A.1. For all equations, we give the coefficients that are used in the DGP and those that are estimaed by our program. As we can see, all the coefficients from the DGP are very closed from the estimates ones.

Table A.1: Simulated data set estimation's results

	Equation 1		*Equation* 2	
	DGP	$Estimated\,coef.$	DGP	$Estimated\,coef.$
	(1)	(2)	(1')	(2')
	Dynamic Equation			
$y1_{-1}$	0.3	0.2195^{***} (0.05)	-0.1	-0.0051 (0.0567)
$y2_{-1}$	0.1	0.1267^{**} (0.0513)	0.4	0.4926^{***} (0.061)
$Gender = Male$	-0.05	-0.0554 (0.0521)	0.05	0.073 (0.0594)
$Medical\,density$	–	–	0.5	0.5687 (1.1111)
$Unemployment\,rate$	-0.2	-0.1682^{***} (0.0269)	–	–
$Intercept$	1.9	2.3113^{***} (0.2667)	-0.4	-0.4677 (2.122)
	Initial Conditions			
$Illness\,before\,prof.\,life$	0.3	0.3032^{***} (0.0283)	-0.2	-0.1624^{***} (0.0221)
Age	–	–	-0.08	-0.093^{***} (0.0202)
$Unemployment\,rate$	-0.2	-0.144^{**} (0.057)	–	–
$Intercept$	-0.2	-0.7331 (0.6194)	2	2.6757^{***} (0.4591)
λ_1	0.4	0.2581^{***} (0.0651)	0.3	0.2660^{***} (0.0463)
λ_2	-0.5	-0.5168^{***} (0.0753)	0.5	0.7022^{***} (0.0598)
	Covariance matrix structure			
	DGP		$Estimated\,coef.$	
	(4)		(5)	
σ_1	2.1		2.4399^{***} (0.1034)	
σ_2	3.1		2.7649^{***} (0.1365)	
ρ_η	0.7		0.7188^{***} (0.0212)	
ρ_ζ	0.5		0.5290^{***} (0.0419)	
ρ_ϵ	0.4		0.6972^{***} (0.1378)	

Estimated standard deviations for estimated coefficients are given within parenthesis.

$***$: significant at the 1% level.

$**$: significant at the 5% level.

A.4.2 Simulated relationship with additional variables

In this section, we keep the same DGP than in section A.4.1 and we add other variables in the model that we estimate in order to evaluate the robustness of the estimation method by the fact that all estimated coefficients for variables in the DGP should remain the same and the added variables coefficients should not significative. We introduce two variables *rural* and *nationality* (*not French*) in the dynamic equations of the regression.

Results are in table A.2. Columns 1 and 2 in table A.2 are the same than corresponding columns in table A.1. We provide in table A.2, column 3, the new results with the additional variables in order to compare with previous estimates[4]. As we can see in the table A.2, the coefficients estimated (using again 16 quadrature points) for those variables are not significant and all initial coefficients in the model remain sensibly the same.

A.4.3 Impact of number of quadrature points on estimated results

As the accuracy of the method depends on the number of quadrature points used for the likelihood calculation, we can try to see how it affects the results when this number increases. For doing so, we fit the same model with different numbers of quadrature points and we calculate the relative difference in log-likelihood and in estimated parameters.
We fit some models by using the same simulated relationship between variables as in section A.4.1.

The results are displayed in the table A.3 for dynamic equations and in the table A.4 for initial conditions equations and errors terms covariance matrix structure.

As we can see from tables A.3 and A.4, by increasing the number of quadrature points the changes in results decline and the relative differences are around 0.01% for significant coefficients and 0.1% or at most 1% for non significant coefficients. After 16 quadrature points, the relative differences in log-likelihood and in estimated coefficients become fewer as we increase the number of quadrature points. The estimations with 22 quadrature points are closer to those with 24 quadrature points than the others. So when we increase the number of quadrature points the changes in estimated coefficients are not significant but the computing time grows up exponentially. For these models, estimation time on an i5 core computer at 2.5 GHz with 6 GB of RAM memory for the different number of quadrature points are given in table A.5.

[4]We do the same with columns 1', 2' of tables A.1 and A.2 (new results are in column 3') and with columns 4 and 5 of both tables (new results in column 6).

Table A.2: Simulated data set with added variables estimation's results

	Equation 1			*Equation* 2		
	DGP	*coef.*	*coef.*	*DGP*	*coef.*	*coef.*
	(1)	(2)	(3)	(1')	(2')	(3')
	Dynamic Equation					
$y1_{-1}$	0.3	0.2195^{***} (0.05)	0.2184^{***} (0.05)	-0.1	-0.0051 (0.0567)	-0.0052 (0.0568)
$y2_{-1}$	0.1	0.1267^{**} (0.0513)	0.1283^{**} (0.0513)	0.4	0.4926^{***} (0.061)	0.4944^{***} (0.0612)
$Gender = Male$	-0.05	-0.0554 (0.0521)	-0.0571 (0.0521)	0.05	0.073 (0.0594)	0.0751 (0.0596)
$Medical\,density$	–	–	–	0.5	0.5687 (1.1111)	0.5567 (1.1112)
$Unemployment\,rate$	-0.2	-0.1682^{***} (0.0269)	-0.1698^{***} (0.0269)	–	–	–
$Not\,French$	–	–	0.1246 (0.0956)	–	–	0.0015 (0.1076)
$rural$	–	–	0.0743 (0.0628)	–	–	0.0283 (0.0719)
$Intercept$	1.9	2.3113^{***} (0.2667)	2.2994^{***} (0.2667)	-0.4	-0.4677 (2.122)	-0.4527 (2.1215)
	Initial Conditions					
$Illness\,before\,prof.life$	0.3	0.3032^{***} (0.0283)	0.3032^{***} (0.0283)	-0.2	-0.1624^{***} (0.0221)	-0.1627^{***} (0.0221)
Age	–	–	–	-0.08	-0.093^{***} (0.0202)	-0.0932^{***} (0.0202)
$Unemployment\,rate$	-0.2	-0.144^{**} (0.057)	-0.144^{**} (0.057)	–	–	–
$Intercept$	-0.2	-0.7331 (0.6194)	-0.7335 (0.6195)	2	2.6757^{***} (0.4591)	2.6803^{***} (0.4595)
λ_1	0.4	0.2581^{***} (0.0651)	0.2582^{***} (0.0653)	0.3	0.266^{***} (0.0463)	0.267^{***} (0.0464)
λ_2	-0.5	-0.5168^{***} (0.0753)	-0.5171^{***} (0.0754)	0.5	0.7022^{***} (0.0598)	0.703^{***} (0.0599)
	Covariance matrix structure					
	DGP	*Estimated coef.*	*Estimated coef.*			
	(4)	(5)	(6)			
σ_1	2.1	2.4399^{***} (0.1034)	2.4353^{***} (0.1032)			
σ_2	3.1	2.7649^{***} (0.1365)	2.763^{***} (0.1366)			
ρ_η	0.7	0.7188^{***} (0.0212)	0.7187^{***} (0.0212)			
ρ_ζ	0.5	0.529^{***} (0.0419)	0.5301^{***} (0.0419)			
ρ_ϵ	0.4	0.6972^{***} (0.1379)	0.697^{***} (0.1378)			

Estimated standard deviations for estimated coefficients are given within parenthesis.

$***$: significant at the 1% level.

$**$: significant at the 5% level.

Table A.3: Impact of the number of quadrature points on estimation results. Part A

	DGP	$Q = 10$	$Q = 16$	$Q = 22$	$Q = 24$
$Log - likelihood$		-8212.05	-8211.26	-8301.71	-8301.24
$y1$	Dynamic equation				
$y1_{-1}$	0.3	0.2754^{***} (0.0489)	0.2195^{***} (0.05)	0.2206^{***} (0.052)	0.2131^{***} (0.0527)
$y2_{-1}$	0.1	0.1376^{***} (0.0483)	0.1267^{**} (0.0513)	0.1196^{**} (0.0554)	0.1010^{*} (0.0568)
$Gender = Male$	-0.05	-0.0580 (0.0479)	-0.0554 (0.0521)	-0.0732 (0.058)	-0.0599 (0.0604)
$Unemployment\,rate$	-0.2	-0.1509^{***} (0.0262)	-0.1682^{***} (0.0269)	-0.1792^{***} (0.0273)	-0.1810^{***} (0.0275)
$Intercept$	1.9	2.3270^{***} (0.2598)	2.3113^{***} (0.2667)	2.3089^{***} (0.2726)	2.30^{***} (0.2753)
$y2$	Dynamic equation				
$y1_{-1}$	-0.1	0.0224 (0.0541)	-0.0051 (0.0567)	-0.0136 (0.0594)	-0.0191 (0.0605)
$y2_{-1}$	0.4	0.5851^{***} (0.0596)	0.4926^{***} (0.0610)	0.4846^{***} (0.0642)	0.4752^{***} (0.0650)
$Gender = Male$	0.05	0.0570 (0.0542)	0.0730 (0.0594)	0.0817 (0.0650)	0.0725 (0.0673)
$Medical\,density$	0.5	1.3305 (1.0685)	0.5687 (1.1111)	0.4874 (1.1357)	0.3549 (1.1473)
$Intercept$	-0.4	-1.7595 (2.040)	-0.4677 (2.1220)	-0.4064 (2.1704)	-0.1492 (2.1936)

Estimated standard deviations for estimated coefficients are given within parenthesis.

$***$: significant at the 1% level.

$**$: significant at the 5% level.

Table A.4: Impact of the number of quadrature points on estimation results. Part B

	DGP	$Q = 10$	$Q = 16$	$Q = 22$	$Q = 24$
$y1$	Initial conditions				
$Illness\,before\,prof.\,life$	0.3	0.3005^{***} (0.0278)	0.3032^{***} (0.0283)	0.3022^{***} (0.0282)	0.3026^{***} (0.0284)
$Unemployment\,rate$	−0.2	-0.1592^{***} (0.0573)	-0.1440^{**} (0.0570)	-0.1437^{**} (0.0571)	-0.1431^{**} (0.0572)
$Intercept$	−0.2	−0.6120 (0.6197)	−0.7331 (0.6194)	−0.7065 (0.6187)	−0.7153 (0.6188)
λ_{11}	0.4	0.2608^{***} (0.0644)	0.2581^{***} (0.0651)	0.2584^{***} (0.0658)	0.2628^{***} (0.0664)
λ_{12}	−0.5	-0.5076^{***} (0.0723)	-0.5168^{***} (0.0753)	-0.5051^{***} (0.0744)	-0.5019^{***} (0.0741)
$y2$	Initial conditions				
Age	−0.08	-0.0859^{***} (0.0196)	-0.0930^{***} (0.0202)	-0.0929^{***} (0.0205)	-0.0943^{***} (0.0207)
$Illness\,before\,prof.\,life$	−0.2	-0.1593^{***} (0.0221)	-0.1624^{***} (0.0221)	-0.1648^{***} (0.0225)	-0.1650^{***} (0.0226)
$Intercept$	2	2.7329^{***} (0.4483)	2.6757^{***} (0.4591)	2.5788^{***} (0.4644)	2.5904^{***} (0.4676)
λ_{21}	0.3	0.2689^{***} (0.0467)	0.2660^{***} (0.0463)	0.2691^{***} (0.0474)	0.2679^{***} (0.0475)
λ_{22}	0.5	0.7136^{***} (0.0607)	0.7022^{***} (0.0598)	0.7008^{***} (0.0625)	0.6932^{***} (0.0626)
	Covariance matrix structure				
σ_1	2.1	2.5202^{***} (0.1053)	2.4399^{***} (0.1034)	2.3920^{***} (0.1047)	2.3898^{***} (0.1051)
σ_2	3.1	2.7012^{***} (0.1307)	2.7649^{***} (0.1365)	2.7928^{***} (0.1444)	2.8281^{***} (0.1468)
ρ_η	0.7	0.7380^{***} (0.0206)	0.7188^{***} (0.0212)	0.7143^{***} (0.0219)	0.7162^{***} (0.0219)
ρ_ζ	0.5	0.5451^{***} (0.0411)	0.5290^{***} (0.0419)	0.5225^{***} (0.0423)	0.5145^{***} (0.0424)
ρ_ϵ	0.4	0.6550^{***} (0.1394)	0.6972^{***} (0.1378)	0.6996^{***} (0.1381)	0.6944^{***} (0.1371)

Estimated standard deviations for estimated coefficients are given within parenthesis.

$***$: significant at the 1% level.

$**$: significant at the 5% level.

$*$: significant at the 10% level.

Table A.5: Computing time for different number of quadrature points

$Quad.\,points$	10	16	22	24
$Comp.\,time\,(in\,min.)$	83	190	450	480

Conclusion

This paper describes the bivariate dynamic probit model with endogenous initial condition starting by justifying the econometric specification of the model, giving the estimation method and its requirements and ending by presenting a robustness analysis. We calculate derivatives of the log-likelihood function with respect to the 13 parameters in the model. For the use of the adaptative Gauss-Hermite quadrature, we also calculate the hessian matrix with respect to individual random effects vector.

The implementation has been done using Stata software. We wrote 2 ado-files for this purpose. We use Stata's d1 method for the maximization process. For the use of this method, we implement the gradient vector for the 13 parameters and we also implement the hessian matrix with respect the random effects vector in order to use the adaptative Gauss-Hermite quadrature. We also wrote two others ado-files for the estimation of the bivariate probit for panel data and the bivariate dynamic probit without initial condition for panel data. These ado-files are written using the same method (Stata's d1 method) with the adaptative Gauss-Hermite quadrature.

Due to the fact that the integration is bi-dimensional, estimation time is very long and still increasing when the quadrature point or the number of observation or the number of variable increase. For an estimated model, one should insure that when increasing the number of quadrature point, the computed results don't change significantly before using them. It means that the relative difference in the results must be around 0.1% or fewer, and if so, we can conclude that the results remain stable when increasing the number of quadrature points. And it means that there is no need to increase the number of quadrature points that will increase computing time but will not improve significantly the results.

Appendix B

Appendix for Chapter 4

B.1 Log-likelihood derivation

This section aims to show details of the calculation of the straightforward form of the likelihood function used for the maximization algorithm. Note that as the measurement variable $h_{i,t}$ is a unidimensional vector, the matrix $M_{i,t/t-1}$ is a scalar.

$$
\begin{aligned}
L_i &= \int_{\mathbb{R}} \phi_{\sigma_1}(\xi_i^1) \prod_{t=1}^{T} \frac{1}{\sqrt{2\,\pi det(M_{i,t/t-1})}} exp\Big(-\frac{1}{2}\,(h_{i,t}-\xi_i^1)\,M_{i,t/t-1}^{-1}\,(h_{i,t}-\xi_i^1)'\Big)\,d\xi_i^1 \\
&= \frac{1}{\sigma_1\sqrt{2\pi}} \prod_{t=1}^{T} \frac{1}{\sqrt{2\,\pi det(M_{i,t/t-1})}} \int_{\mathbb{R}} exp\bigg(-\frac{1}{2}\Big[\sum_{t=1}^{T} M_{i,t/t-1}^{-1}\,(h_{i,t}-\xi_i^1)^2 + (\frac{\xi_i^1}{\sigma_1})^2\Big]\bigg)\,d\xi_i^1 \\
&= \frac{1}{\sigma_1\sqrt{2\pi}} \prod_{t=1}^{T} \frac{1}{\sqrt{2\,\pi det(M_{i,t/t-1})}} exp\bigg(-\frac{1}{2}\sum_{t=1}^{T} M_{i,t/t-1}^{-1} h_{i,t}^2\bigg) * \\
&\quad \int_{\mathbb{R}} exp\bigg(-\frac{1}{2}\Big[\Big(\sum_{t=1}^{T} M_{i,t/t-1}^{-1} + \frac{1}{\sigma_1^2}\Big)(\xi_i^1)^2 - 2\xi_i^1 \sum_{t=1}^{T} M_{i,t/t-1}^{-1} h_{i,t}\Big]\bigg)\,d\xi_i^1 \\
&= \frac{1}{\sigma_1\sqrt{2\pi}} \prod_{t=1}^{T} \frac{1}{\sqrt{2\,\pi det(M_{i,t/t-1})}} exp\bigg(-\frac{1}{2}\sum_{t=1}^{T} M_{i,t/t-1}^{-1} h_{i,t}^2\bigg) exp\bigg(\frac{\big(\sum\limits_{t=1}^{T} M_{i,t/t-1}^{-1}\,h_{i,t}\big)^2}{2(\frac{1}{\sigma_1^2} + \sum\limits_{t=1}^{T} M_{i,t/t-1}^{-1})}\bigg) * \\
&\quad \int_{\mathbb{R}} exp\bigg(-\frac{1}{2}\Big[\Big(\sum_{t=1}^{T} M_{i,t/t-1}^{-1} + \frac{1}{\sigma_1^2}\Big)\Big(\xi_i^1 - \frac{\sum\limits_{t=1}^{T} M_{i,t/t-1}^{-1} h_{i,t}}{\sum\limits_{t=1}^{T} M_{i,t/t-1}^{-1} + \frac{1}{\sigma_1^2}}\Big)^2\Big]\bigg)\,d\xi_i^1 \\
&= \frac{1}{\sqrt{1+\sigma_1^2 \sum\limits_{t=1}^{T} M_{i,t/t-1}^{-1}}} \prod_{t=1}^{T} \Big(\frac{1}{\sqrt{2\,\pi det(M_{i,t/t-1})}}\Big) exp\bigg(-\frac{1}{2}\sum_{t=1}^{T} M_{i,t/t-1}^{-1} h_{i,t}^2\bigg) exp\bigg(\frac{\big(\sum\limits_{t=1}^{T} M_{i,t/t-1}^{-1}\,h_{i,t}\big)^2}{2(\frac{1}{\sigma_1^2} + \sum\limits_{t=1}^{T} M_{i,t/t-1}^{-1})}\bigg)
\end{aligned}
$$

The link between the last two equalities is established by the use of the following relationship (Gauss integral) :

$$\int_{\mathbb{R}} exp\bigg(-a(x-b)^2\bigg)dx = \int_{\mathbb{R}} exp\bigg(-ay^2\bigg)dy = \sqrt{\frac{\pi}{a}}$$

B.2 Kalman filter derivation

In this section, we present details of calculation of the Kalman filter's tools applied on the following state-space model matrix representation :

$$H_{i,t} = B_{i,t-1}\Gamma_{i,t-1} + H_{i,t-1} + \xi_{i,t}, \forall t \geqslant 1$$
$$\Gamma_{i,t-1} = A_1\Gamma_{i,t-2} + A_0 + \Xi_{i,t-1}, \forall t \geqslant 2$$

The calculations are inspired by the chapter 11 of the book of Droesbeke et al (2013). We start with the notations below :

$$\begin{cases} (a_1)\ \hat{\Gamma}_{i,t-1/t} &= E(\Gamma_{i,t-1}/H_{i,1},...,H_{i,t}) \\ (a_2)\ \Sigma_{i,t-1/t} &= V(\Gamma_{i,t-1}/H_{i,1},...,H_{i,t}) \\ (b_1)\ \hat{\Gamma}_{i,t/t} &= E(\Gamma_{i,t}/H_{i,1},...,H_{i,t}) \\ (b_2)\ \Sigma_{i,t/t} &= V(\Gamma_{i,t}/H_{i,1},...,H_{i,t}) \\ (c_1)\ \hat{H}_{i,t/t-1} &= E(H_{i,t}/H_{i,1},...,H_{i,t-1}) \\ (c_1)\ M_{i,t/t-1} &= V(H_{i,t}/H_{i,1},...,H_{i,t-1}) \\ Q_t &= V(\Xi_{i,t}) = \Sigma_\Xi \\ R_t &= V(\xi_{i,t}) = \sigma_1^2 + \sigma_2^2 \end{cases}$$

The first step consists of the calculation of (a_1) and (a_2). The probability distribution function ℓ of $\left(\Gamma_{i,t-1}/H_{i,1},...,H_{i,t-1}\right)$ is (recurrence hypothesis) :

$$\ell(\Gamma_{i,t-1}/H_{i,1},...,H_{i,t-1}) = N(\hat{\Gamma}_{i,t-1/t-1}, \Sigma_{i,t-1/t-1})$$

And the probability distribution function of $H_{i,t}/\Gamma_{i,t-1}, H_{i,1},...,H_{i,t-1}$ is :

$$\ell(H_{i,t}/\Gamma_{i,t-1}, H_{i,1},...,H_{i,t-1}) = N(B_{i,t-1}\Gamma_{i,t-1} + H_{i,t-1}, R_t)$$

As $B_{i,t-1}\Gamma_{i,t-1} + H_{i,t-1} = B_{i,t-1}\hat{\Gamma}_{i,t-1/t-1} + H_{i,t-1} + B_{i,t-1}(\Gamma_{i,t-1} - \hat{\Gamma}_{i,t-1/t-1})$, where $B\hat{\Gamma}_{i,t-1/t-1} + H_{i,t-1}$ denotes the mean of $\left(H_{i,t}/H_{i,1},...,H_{i,t-1}\right)$, and by using the theorem 1 in chapter 11 of the book of Droesbeke et al (2013), we can deduce the probability distribution function of $\left(H_{i,t}, \Gamma_{i,t-1}/H_{i,1},...,H_{i,t-1}\right)$ as :

$$\ell(H_{i,t}, \Gamma_{i,t-1}/H_{i,1},...,H_{i,t-1}) = N(m_{H,\Gamma}, V_{H,\Gamma})$$

Where $m_{H,\Gamma} = \begin{pmatrix} B_{i,t-1}\Gamma_{i,t-\hat{1}/t-1} + H_{i,t-1} \\ \hat{\Gamma}_{i,t-1/t-1} \end{pmatrix}$, and $V_{H,\Gamma} = \begin{pmatrix} R_t + B_{i,t-1}\Sigma_{i,t-1/t-1}B' & B_{i,t-1}\Sigma_{i,t-1/t-1} \\ \Sigma_{i,t-1/t-1}B'_{i,t-1} & \Sigma_{i,t-1/t-1} \end{pmatrix}$. Thus, by using the theorem 2 in chapter 11 of the book of Droesbeke et al (2013), we can deduce explicit forms of (a_1) and (a_2) as :

$$\hat{\Gamma}_{i,t-1/t} = \hat{\Gamma}_{i,t-1/t-1} + \Sigma_{i,t-1/t-1}B'_{i,t-1}(R_t + B_{i,t-1}\Sigma_{i,t-1/t-1}B'_{i,t-1})^{-1}(H_{i,t} - B_{i,t-1}\hat{\Gamma}_{i,t-1/t-1} - H_{i,t-1})$$
$$\Sigma_{i,t-1/t} = \Sigma_{i,t-1/t-1} - \Sigma_{i,t-1/t-1}B'_{i,t-1}(R_t + B_{i,t-1}\Sigma_{i,t-1/t-1}B'_{i,t-1})^{-1}B_{i,t-1}\Sigma_{i,t-1/t-1}$$

For terms in (b_1) and (b_2), we start by calculating the probability distribution function of $\left(\Gamma_{i,t-1}/H_{i,1},...,H_{i,t}\right)$ and $\left(\Gamma_{i,t-1}/H_{i,1},...,H_{i,t}\right)$. They are :

$$\ell(\Gamma_{i,t-1}/H_{i,1},...,H_{i,t}) = N(\hat{\Gamma}_{i,t-1/t},\Sigma_{i,t-1/t})$$
$$\ell(\Gamma_{i,t}/\Gamma_{i,t-1},H_{i,1},...,H_{i,t}) = N(A_0+A_1\Gamma_{i,t-1},Q_t)$$

Here too, we use the theorem 1 and the fact that $A_0+A_1\Gamma_{i,t-1} = A_0+A_1\hat{\Gamma}_{i,t-1/t}+A_1(\Gamma_{i,t-1}-\hat{\Gamma}_{i,t-1/t})$ to deduce the probability distribution function of $\left(\Gamma_{i,t-1},\Gamma_{i,t}/H_{i,1},...,H_{i,t}\right)$ as :

$$\ell(\Gamma_{i,t-1},\Gamma_{i,t}/H_{i,1},...,H_{i,t}) = N(m_{\Gamma_t,\Gamma_{t-1}},V_{\Gamma_t,\Gamma_{t-1}})$$

Where $m_{\Gamma_t,\Gamma_{t-1}} = \begin{pmatrix} \hat{\Gamma}_{i,t-1/t} \\ A_0+A_1\hat{\Gamma}_{i,t-1/t} \end{pmatrix}$, and $V_{\Gamma_t,\Gamma_{t-1}} = \begin{pmatrix} \Sigma_{i,t-1/t} & \Sigma_{i,t-1/t}A_1^{'} \\ A_1\Sigma_{i,t-1/t} & Q_t+A_1\Sigma_{i,t-1/t}A_1^{'} \end{pmatrix}$. Then we can deduce explicit forms of (b_1) and (b_2) as (and this relation proves the recurrence hypothesis) :

$$\hat{\Gamma}_{i,t/t} = A_0+A_1\hat{\Gamma}_{i,t-1/t}$$
$$\Sigma_{i,t/t} = Q_t+A_1\Sigma_{i,t-1/t}A_1^{'}$$

Then we can deduce explicit forms of (c_1) and (c_2) by using the derived probability distribution function of $\left(H_{i,t},\Gamma_{i,t-1}/H_{i,1},...,H_{i,t-1}\right)$ as :

$$\hat{H}_{i,t/t-1} = B_{i,t-1}\hat{\Gamma}_{i,t-1/t-1}+H_{i,t-1}$$
$$M_{i,t/t-1} = R_t+B_{i,t-1}\Sigma_{i,t-1/t-1}B_{i,t-1}^{'}$$

When we apply the Kalman filter to our model, we obtain the following estimation for parameters :

$$\begin{cases} \hat{\Gamma}_{i,t-1/t} &= \hat{\Gamma}_{i,t-1/t-1}+\Sigma_{i,t-1/t-1}B_{i,t-1}^{'}(R_t+B_{i,t-1}\Sigma_{i,t-1/t-1}B_{i,t-1}^{'})^{-1}(H_{i,t}-B_{i,t-1}\hat{\Gamma}_{i,t-1/t-1}-H_{i,t-1}) \\ \Sigma_{i,t-1/t} &= \Sigma_{i,t-1/t-1}-\Sigma_{i,t-1/t-1}B_{i,t-1}^{'}(R_t+B_{i,t-1}\Sigma_{i,t-1/t-1}B_{i,t-1}^{'})^{-1}B_{i,t-1}\Sigma_{i,t-1/t-1} \\ \hat{\Gamma}_{i,t/t} &= A_0+A_1\hat{\Gamma}_{i,t-1/t} \\ \Sigma_{i,t/t} &= Q_t+A_1\Sigma_{i,t-1/t}A_1^{'} \\ \hat{H}_{i,t/t-1} &= B_{i,t-1}\hat{\Gamma}_{i,t-1/t-1}+H_{i,t-1} \\ M_{i,t/t-1} &= R_t+B_{i,t-1}\Sigma_{i,t-1/t-1}B_{i,t-1}^{'} \end{cases}$$

Contents

List of Tables

List of Figures