Thèse de doctorat Novembre/ 2025

Université Paris - Panthéon - Assas

École doctorale d'Économie, Gestion, sciences de l'Information et de la Communication - ED 455

Thèse de doctorat en Sciences Économiques

soutenue le 25 Novembre 2025

Three essays on the determinants and consequences of employee presenteeism in France



Théo DENIS

Sous la direction de Mme. Sandra CAVACO, Maître de conférences - HDR, et M. Joseph LANFRANCHI, Maître de conférences, à l'Université Paris - Panthéon - Assas

Membres du jury :
Thomas Barnay, Professeur à l'Université Paris-Est Créteil, Président
Nathalie Havet, Professeure à l'Université de Lyon, rapporteure
Yannick L'Horty, Professeur à l'Université Gustave Eiffel, rapporteur
Liliane Bonnal, Professeure à l'Université de Poitiers, examinatrice

Avertissement

L'université n'entend donner aucune approbation ni improbation aux opinions émises dans cette thèse ; ces opinions doivent être considérées comme propres à leur auteur.

Remerciements

En tout premier lieu, je tiens à exprimer ma profonde gratitude envers Sandra Cavaco et Joseph Lanfranchi pour leur encadrement, leurs conseils, leurs relectures attentives ainsi que pour tous nos échanges, qui n'ont fait qu'enrichir et améliorer cette thèse. Plus encore, s'agissant d'un travail consacré à la santé et la qualité de vie au travail, je leur suis infiniment reconnaissant pour leur bienveillance et leur confiance tout au long de ces années. L'autonomie dont j'ai bénéficié m'a largement permis de trouver un équilibre entre mon travail de thèse et ma vie personnelle, me permettant de concrétiser un certain nombre de projets.

J'adresse également toute ma reconnaissance aux chercheurs du LEMMA qui ont pris le temps de lire et de discuter ce travail de recherche. Je pense à Anne-Laure Samson, José De Sousa, Nicolas Frémeaux, Paul Maarek, et notamment à Ali Skalli, dont les échanges ont été particulièrement riches et instructifs. Je remercie aussi Christina Katt-Pawlowitsch pour nos discussions à propos de la théorie du signal, concernant un chapitre qui n'aura malheureusement pas vu le jour.

Je ne remercierai jamais assez Chloé, ma compagne dans la vie, qui a partagé et supporté les contraintes qu'impose un travail de thèse pendant six longues années. Elle aura cependant toujours été à mes côtés, su écouter mes doutes et m'accompagner dans mes choix.

Je veux également adresser une pensée toute particulière à mes parents, qui ont toujours cru en moi, même lorsque beaucoup auraient perdu patience bien plus tôt. Je leur serai toujours reconnaissant de ne jamais m'avoir imposé des choix qui n'auraient pas été les miens, et, au contraire, d'avoir toujours soutenu mes décisions de manière inconditionnelle. Je les remercie pour l'éducation qu'ils m'ont transmise.

Enfin, mes remerciements vont à mes amis, pour leur soutien et leur intérêt constant à l'égard de l'avancée de ma thèse, pour m'avoir si souvent demandé "alors, tu la finis quand cette thèse ?" La voici enfin "cette thèse" tant attendue de tous ! Je souhaite également remercier, avec toute mon amitié, mes chers collègues et amis doctorants du LEMMA, avec qui j'ai eu le plaisir de partager le bureau doctorant, lieu de passionnantes discussions. Je remercie enfin le LEMMA, dirigé par Antoine Billot, de m'avoir accueilli et donné l'opportunité de réaliser cette thèse.

Je souhaite dédier cette thèse à tous ceux qui pensent ne pas avoir le choix. Jean-Paul Sartre disait si simplement : “La liberté est choix”. Il suffit simplement de sauter le pas !

Titre et résumé :

Trois essais sur les déterminants et les conséquences du présentéisme des salariés en France

Dans un contexte de hausse des charges de sécurité sociale en France, comprendre les mécanismes expliquant l'augmentation de l'absence est essentiel. Au-delà du coût important de l'absence, il convient de considérer celui du présentéisme, c'est-à-dire le fait de fournir un effort alors qu'on ne devrait pas. Cette thèse étudie deux formes de présentéisme : le présentéisme maladie (travailler en étant malade) et le présentéisme compétitif (travailler plus que nécessaire). En outre, elle propose une analyse empirique conjointe des comportements d'absence et de présentéisme. Ainsi, le premier chapitre explore les effets des caractéristiques de l'emploi sur la survenue de la maladie, ainsi que sur les durées cumulées d'absence et de présence en cas de maladie. Il montre que certains facteurs préalablement identifiés comme affectant le présentéisme et l'absence concernent en réalité la maladie elle-même. Par ailleurs, la durée d'absence dépend surtout des caractéristiques du contrat de travail, tandis que le présentéisme maladie est davantage lié aux conditions de travail. Le deuxième chapitre étudie les effets à long terme des choix entre absence et présence en cas de maladie sur la santé future. Celui-ci confirme que le présentéisme maladie nuit à la santé future, contrairement à l'absence, qui, bien que coûteuse à court terme, est bénéfique pour le rétablissement de la santé à long terme. Enfin, le troisième chapitre examine quant à lui les effets du présentéisme compétitif sur la carrière et la santé. Si ce comportement est effectivement associé à une progression de carrière en termes de salaire et de niveau professionnel, celui-ci peut néanmoins dégrader la santé et conduire à une augmentation de la durée d'absence future.

Descripteurs :

Présentéisme, Absentéisme, Santé, Bien-être, Motivation, Carrière, Signal, Endogénéité, Copules.

Abstract:

In a context of rising social security costs in France, it is essential to understand the mechanisms behind the increase in absence. Beyond the significant cost of absence, it is important to consider the cost of presenteeism, i.e. the act of working when one should not. This thesis examines two forms of presenteeism: sickness presenteeism (working while sick) and competitive presenteeism (working more than necessary). It also offers a joint empirical analysis of absenteeism and presenteeism behaviours. The first chapter explores the effects of employment characteristics on illness, as well as on the cumulative duration of absenteeism and presenteeism in the event of illness. It shows that some factors previously identified as affecting presenteeism and absenteeism are in fact related to the illness itself. Furthermore, the duration of absence depends mainly on the characteristics of the employment contract, while sickness presenteeism is more closely related to working conditions. The second chapter examines the long-term effects of choices between absence and attendance in the event of illness on future health. It confirms that sickness presenteeism is detrimental to future health, unlike absence, which, although costly in the short term, is beneficial to recovery in the long run. Finally, the third chapter examines the effects of competitive presenteeism on career and health. While this behaviour is indeed associated with career advancement in terms of wages and occupational status, it can nevertheless damage health and lead to an increase in the future duration of sickness absence.

Keywords:

Presenteeism, Absenteeism, Health, Well-being, Motivation, Career, Signal, Endogeneity, Copula.

Contents

Contents 9

General introduction 13

1 Presenteeism: definitions and challenges . . . 16

1.1 The first approaches to presenteeism . . . 17

1.2 Emerging issues following the seminal articles . . . 20

2 Definitions and terminology . . . 23

2.1 Presenteeism . . . 23

2.2 Sickness presenteeism . . . 24

2.3 Competitive presenteeism . . . 26

2.4 Presenteeism and moral hazard . . . 28

3 Research framework . . . 30

Chapter 1 - A new empirical model of the determinants of sickness and the choice between presenteeism and absence . . . 32

Chapter 2 - All work and no rest makes Jack a sick boy: A causal analysis of sickness absence and presenteeism on health . . . 34

Chapter 3 - Is presenteeism such a valuable signal for employees? Analysis on wages, career, health and well-being . . . 37

Bibliography 41

Chapter 1 - A new empirical model of the determinants of sickness and the choice between presenteeism and absence 47

1 Introduction . . . 48

2 Data and measurement . . . 53

2.1 Data and sample of study . . . 53

2.2 Measure of sickness absence and presenteeism . . . 53

2.3 Explanatory variables . . . 56
3 An empirical model of attendance decisions and econometric methodology . . . 57
3.1 Count-data model with endogenous participation . . . 59
3.2 A copula approach for SA and SP days with endogenous participation . . . 61
4 Results . . . 64
4.1 Model choice . . . 64
4.2 Participation bias . . . 68
4.3 Determinants of the feeling of illness . . . 72
4.4 Determinants of the duration of sickness presenteeism and absence. . . . 75
4.5 Additional results . . . 79
5 Concluding remarks and discussion . . . 84

Bibliography **89**

Appendices **95**
1 Descriptive statistics . . . 97
2 AIC and BIC value tables for every SA and SP duration models . . . 99
3 Estimates of the models with different sets of variables . . . 102
4 Estimates of models with various restrictions of the sample of study 108
5 Models with different choice of copula . . . 115
6 Models with alternative exclusion restrictions . . . 120
7 Additional controls for past health and employment history . . . 126

Chapter 2 - All work and no rest makes Jack a sick boy: A causal analysis of sickness absence and presenteeism on health **129**
1 Introduction . . . 130
2 Review of the literature . . . 135
2.1 The costs of sickness presenteeism and absenteeism . . . 135
2.2 The effects of sickness presenteeism and absenteeism over time 137
3 Empirical intuition of the effects of SA and SP over time . . . 140
4 Data and methodology . . . 145

4.1 Data and sample of study . . . 145
4.2 Measures . . . 146
4.3 Empirical methodology . . . 150
5 Results . . . 157
5.1 Descriptive analysis . . . 158
5.2 Inferential analysis . . . 161
6 Discussion and conclusion . . . 171

Bibliography **179**

Appendices **185**
1 Descriptive statistics . . . 187
2 Interaction results between self-rated health and both SP and SA . . . 191

Chapter 3 - Is presenteeism such a valuable signal for employees? Analysis on wages, career, health and well-being **193**
1 Introduction . . . 194
2 Review of the literature . . . 197
2.1 The career concern . . . 197
2.2 The health consequences of over-work . . . 201
3 Data and method . . . 203
3.1 Sample restriction . . . 204
3.2 The outcome variables . . . 205
3.3 The identification of competitive presenteeism . . . 207
3.4 Additional covariates . . . 214
4 Method and econometric models . . . 215
4.1 The lagged models . . . 216
4.2 Exploring causality . . . 218
5 Results . . . 226
5.1 Preliminary analysis . . . 226
5.2 Inferential analysis . . . 229
5.3 The moderated effects of prosocial and work pressure characteristics . . . 238
5.4 Additional results . . . 241

6 Conclusion and discussion 243

Bibliography **249**

Appendices **255**

1 Descriptive statistics 257

1.1 Tables of descriptive statistics of explanatory variables 257

1.2 Graphics of the MCA to create the competitive presenteeism index 259

2 Complete tables of the main results 261

2.1 Results of the pooled models 261

2.2 Results of the two stages Copula Endogeneity Correction models 262

2.3 Instrumental variable models 264

3 Mediated effects of work pressure and intrinsic motivations 267

3.1 The mediated effects of work pressure 267

3.2 The mediated effects of intrinsic motivations 271

4 Heterogenous effects 275

4.1 Gender 275

4.2 Public sector 277

4.3 Occupation (blue or white collars) 279

4.4 Healthcare sector subset 281

4.5 Public healthcare sector 282

5 Direct effects of the variables included in the CP indicator 284

General conclusion **289**

1 Main contributions 290

2 Limitations and future research 294

Bibliography **299**

List of Tables **311**

List of Figures **316**

General introduction

In the context of a sharp increase in social security costs, sick leave and workplace accidents are widely blamed. A report by the French Cour des Comptes (2024) highlights, for example, a considerable increase in the cost of the French social security system, representing €12 billion in 2022, an increase of 56% since 2017. In addition, a report by the Inspection générale des finances and l'Inspection générale des affaires sociales (Triolle et al., 2024) indicates that sickness absence increased by 63% between 2019 and 2022 in the French public sector, a much larger increase than in the private sector, which has not been spared either, with an increase of 39%. Since 2019, the rate of absenteeism due to long-term illness (90 days or more) in the private sector has been rising steadily, with an increase of more than 30% between 2019 and 2023[1]. Reducing the duration and frequency of absences and promoting worker health are therefore key concerns in the context of preserving the social security system.

However, should these costs be reduced through policies that discourage absenteeism, as recommended by the French Cour des Comptes, in favour of greater attendance? If so, what levers could be used? First of all, to answer this question, it is important to identify the factors that lead to an increase in the length of absences. Furthermore, from a broader perspective, it is worth considering the consequences that a strong incentive for attendance might have.

It is indeed important to consider the antagonistic counterpart of absenteeism: presenteeism at work. Implementing incentives for attendance could lead workers to over-respond by showing up to work even when it would have been justified for them to stay home and rest. In such a case, is it appropriate to encourage such behaviour? What would be the results, the long-term consequences and the costs? At first glance, ever-increasing effort or unfailing presence might be associated with higher production. However, if repeated too frequently, such intense effort could ultimately be detrimental to health. A constant and sustained investment in one's work could heighten stress and fatigue, leading to an overall deterioration of workers' health. Moreover, these incentives to exert effort could drive employees to attend work even when they are ill.

1. Baromètre Ayming/AG2R20 du 15 avril 2024 : "Ayming et AG2R LA MONDIALE dévoilent les premiers chiffres de leur 16e Baromètre de l'Absentéisme® et de l'Engagement 2024"

Such behaviours – reinforced by technological and organizational changes and by a competitive environment that already fosters work intensification (Gallie & Zhou, 2013; Green et al., 2021) – could push employees into a vicious circle in which their health would continuously deteriorate. In the end, this would lead to a higher probability of absence, longer spells of absence (Bergström et al., 2009b), and potentially more burnout (Demerouti et al., 2009) and workplace accidents. Ultimately, all of this would raise costs both for firms and for the social security system, and thus would not resolve the issue raised by the French Cour des comptes. For instance, according to Evans-Lacko & Knapp (2016), the costs associated with sickness presenteeism are estimated to be five to ten times higher than those associated with sickness absence. In this context, this thesis seeks to better understand the associations between absenteeism and over-investment behaviours. What are the determinants that foster one or the other, and what are the relationships between them?

These attendance behaviours, which may be regarded as inappropriate, characterise what is referred to as presenteeism. For several years now, research on presenteeism has been attracting growing interest across several disciplines, including economics and human resource management, but also sociology, psychology, and organizational medicine. Because of this multidisciplinary nature, the literature addressing this phenomenon is relatively extensive (Karanika-Murray & Cooper, 2018). Even today, there is no clear consensus within the academic community regarding the definition to be adopted (Lohaus & Habermann, 2019). For instance, two terms – *"presenteeism"* and *"sickness presenteeism"* – are commonly used to describe what appears to be the same behaviour, namely attending work while sick. However, the original meaning of the term presenteeism is broader, thereby opening different avenues for research on worker behaviour.

According to the original definition by Cooper (1996), presenteeism would correspond to any situation in which a worker devotes time and effort to their job even though it would have been entirely legitimate to spend that time on leisure (meaning anything not related to employment). Thus, presenteeism is not solely a behaviour associated with illness (Ashby & Mahdon, 2010). Rather, the term can also apply to a pattern of over-investment in work – whether in terms of time or

effort – motivated by performance within a competitive career framework (Simpson, 1998). Such behaviour is referred to as *"competitive presenteeism"*, an aspect of presenteeism that remains under-explored in the academic literature, and which will be the focus of a dedicated chapter in this thesis.

Literature reviews and meta-analyses devoted to presenteeism (Patel et al., 2023; Ruhle et al., 2019; Lohaus & Habermann, 2019; Navarro et al., 2019; Kinman, 2019; McGregor et al., 2018; Karanika-Murray & Cooper, 2018; Miraglia & Johns, 2016; Skagen & Collins, 2016; Garrow, 2016; Johns, 2010; Schultz & Edington, 2007) highlight the fact that there is a wide diversity of interpretations of what exactly this term covers. This has led to the development of a sparse literature and a profusion of definitions that can refer to behaviours that are sometimes very different, even contradictory. This polysemy affecting presenteeism in the academic field may stem from cultural, societal or economic differences around the world (Karanika-Murray & Cooper, 2018), making comparisons between countries difficult.

Given these particularities, how can these difficulties be addressed and a research framework in economics be proposed that is compatible with the various forms of presenteeism? What behaviour are we talking about? Is it a single behaviour, or are there in fact several types of presenteeism? In the latter case, what are the fundamental characteristics that allow each type of presenteeism to be identified? The following section presents previous research and organises the different concepts and their definitions, in order to propose a clear research framework and open avenues for addressing some of the questions raised.

1 Presenteeism: definitions and challenges

Research on presenteeism shows a certain convergence towards the analysis of presenteeism associated with illness. However, as we shall see below, this convergence is only partial and masks other possible meanings of the term presenteeism. Recent literature actually mentions two common definitions of presenteeism, one linked to illness (Aronsson et al., 2000; Johns, 2010), the other to productivity loss (A. M. Collins et al., 2017; Ishimaru et al., 2020). The second approach defines presenteeism as any situation in which an individual is unproductive at work,

with illness being one of the reasons for the decline in productivity. Furthermore, while presenteeism has regularly been described as a behaviour with harmful consequences for both the health and productivity of workers, recent literature has begun to study the beneficial effects of presenteeism (Patel et al., 2023). This shows a certain tendency to explore mechanisms outside the initial field of study relating to the definitions established in previous literature. Thus, with the aim of developing a structured and clear analytical framework, a core behaviour associated with the term *presenteeism* is defined below. Then, taking into account previous literature, two specific forms of presenteeism are also characterised and analysed: *sick presenteeism* and *competitive presenteeism*. Finally, a cheating behaviour associated with the issue of moral hazard in the context of presenteeism will also be presented.

1.1 The first approaches to presenteeism

The term "presentee" first appeared in 1892, in Mark Twain's novel *The American Claimant*, presumably in opposition to "absentee" [2]. Several decades passed before the concept and term presenteeism appeared in the business press in the 1930s, where it referred to the simple fact of being fully present at work. This interpretation of presenteeism was put forward by Canfield & Soash (1955), in strict opposition to absenteeism, as the "demonstration of impeccable attendance". In this view, the absenteeism rate was contrasted with the presenteeism rate, the latter corresponding to the ratio of time spent at work to total working time. This initial interpretation of presenteeism prevailed until the early 1990s (Smith, 1970; Stolz, 1993). From that point onwards, researchers began to conceptualise and study presenteeism as a behaviour in its own right. In other words, whereas presenteeism was primarily defined as a measure, in opposition to absenteeism, it is now defined as the decision to attend work under certain conditions.

Thus, the first comprehensive definition of presenteeism was provided by Cooper (1996). He defined presenteeism as "being at work when you should be at home either because you are ill or because you are working such long hours that you are no

2. "At last Gwendolen got some peace of mind by sending word to the Thompsons, in the neighborhood, that she was coming there to dinner. She wouldn't be reminded, at that table, that there was an absentee who ought to be a presentee — a word which she meant to look out in the dictionary at a calmer time."

longer effective". This definition highlights three specific points: a behaviour (going to work when one should not), two causes (illness or excessive working hours), and a consequence (loss of efficiency). As we shall see later, it is likely from these three elements raised in this definition that the multitude of interpretations surrounding presenteeism today has emerged. The counterpart to this was the development of multiple definitions associated with the term presenteeism, sometimes linked to different underlying concepts.

Two years after Cooper, Simpson (1998), who examined the effect of organisational changes on the workload and the behaviour of managers, found that restructuring was associated with the need to demonstrate commitment to one's job by working visibly longer hours. The fear of dismissal and uncertainty about promotion opportunities would, in turn, lead to a form of presenteeism defined as "the tendency to stay at work beyond the time needed for effective performance of the job".

This definition focuses on the worker's marginal efficiency, that is, on the worker's actual performance associated with an additional hour of work, given the amount of working time already provided. The author suggests that this inefficient behaviour is linked to the competitive environment in which the worker operates, compelling them to demonstrate visible and sustained attendance. Workers thus compete with one another to remain at work for as long as possible, despite their lack of effectiveness. A behaviour she called *"competitive presenteeism"*. This form is not related to the individual's state of health. However, similar to sickness presenteeism, which is regarded as detrimental to productivity, the behaviour she outlines may also have counterproductive effects for the organisation as well as for the worker. We shall return to this point in greater detail later.

In line with Simpson, by examining managers' behaviour in response to organisational changes when firms make use of redundancies, Worrall et al. (2000) highlight that "managers attend work to demonstrate their indispensability by working visibly long hours". An interesting detail emphasised by these authors is the visibility of the behaviour. The idea that presenteeism functions primarily as a signal to the employer emerged relatively early in the literature, notably with the paper by Hemp (2004).

For their part, Aronsson et al. (2000) focused the definition of presenteeism on the health status. They define, through the new term *sickness presenteeism*, the behaviour of an individual who, "despite complaints and ill health that should prompt rest and absence from work, still turns up at their job". Unlike previous definitions, this latest definition is based on a central behaviour without attaching a particular objective to it or presupposing any consequences of this behaviour. In addition, the use of a new term specifically related to illness makes it possible to identify and frame the issues of interest associated with this concept. Nevertheless, this definition only addresses part of the field of study concerning presenteeism related to illness. Later, Chapman (2005) defined presenteeism as "the extent to which symptoms, ailments and illnesses affect the productivity of individuals who choose to remain at work". Thus, the author reintroduces into the definition of sickness presenteeism given by Aronsson et al. (2000) the fact that coming to work in poor health leads to a decline in productivity.

Through an extensive review of the literature, Johns (2010) finally summarises the definition of presenteeism as "attending work while ill". However, his literature review focuses mainly on presenteeism related to illness or the decline in productivity associated with a behaviour defined as such. It thus omits a field of literature related to workload, as initially mentioned in the definition given by Cooper (1996). Furthermore, while Aronsson et al. (2000) proposed the use of the term "sickness presenteeism", the definition given by Johns (2010) uses only the term "presenteeism". However, this choice confines this term to a behaviour associated solely with illness. Yet, the same year, Ashby & Mahdon (2010) emphasised the fact that presenteeism is not only related to illness, opening the way to two fields of research based on the same term but underpinned by two different concepts.

Finally, presenteeism can also be interpreted in a very different way from the previous forms. In the introduction of his paper, Hemp (2004) quotes Woody Allen, who once said that "80% of success in life can be attributed simply to showing up". While the article by Hemp (2004) focuses mainly on sickness presenteeism, Woody Allen's quote describes a relatively different behaviour. It emphasises more the fact of showing up at work without intentionally making the effort necessary to perform one's tasks. Finally, relatively close to this latter conception of presenteeism,

D'Abate & Eddy (2007) choose to define what they call *"non-work presenteeism"* as spending time on personal tasks during working hours, ultimately leading to a decline in productivity. However, this behaviour has not received much attention from the academic community. Nevertheless, the idea that presenteeism serves solely as a visible signal appears implicitly in many research papers. This is the case in the articles by Simpson (1998) and Worrall et al. (2000) on competitive presenteeism, but also in the papers by Aronsson et al. (2000) and Johns (2010) on sickness presenteeism. This shows the importance of clearly distinguishing and defining the underlying motivations of individuals to engage in presenteeism behaviour. Is it a behaviour that describes the pursuit of effort in the event of illness, of over-effort through hard work, or simply the act of showing up without providing the expected effort?

These different visions associated with presenteeism reveal the need to establish a clear research framework between motivations, behavioural characteristics and possible consequences. Furthermore, are there any links between these different behaviours? The literature relating to these three dimensions of presenteeism is presented below.

1.2 Emerging issues following the seminal articles

As mentioned above, three types of presenteeism behaviour have been studied in the literature. These definitions are based on determinants, objectives, concepts or consequences in order to define a specific behaviour. The definitions used in the studies appear to be directly linked to the objectives pursued by the researchers (Karanika-Murray & Cooper, 2018). The first definition developed by Cooper (1996) thus opened up a wide range of possible interpretations of presenteeism. As a result, some of the research focused on sickness presenteeism (Aronsson & Gustafsson, 2005; Hansen & Andersen, 2008; Bergström et al., 2009b; Johns, 2010; Arnold, 2016; Gustafsson Sendén et al., 2016; Kinman & Wray, 2018; Helgesson et al., 2021; Biron et al., 2022; Marie & Val Castello, 2023), and on the other hand, on the productivity loss associated with both "illness" and "competitive" presenteeism (Simpson, 1998; Burton et al., 2004; Hemp, 2004; J. Collins et al., 2005; Chapman, 2005; Goetzel et al., 2009; Zhang et al., 2015; Allen et al., 2018; Malmberg-Ceder et al., 2020).

It should be noted that recent literature still mainly distinguishes between two characteristics of the definition given by Cooper (1996): working while ill (Johns, 2010) and working while experiencing a decline in productivity (A. M. Collins et al., 2017; Ishimaru et al., 2020). However, various literature reviews and meta-analyses on the subject (Miraglia & Johns, 2016; Skagen & Collins, 2016; Karanika-Murray & Cooper, 2018; Lohaus & Habermann, 2019; Navarro et al., 2019; Patel et al., 2023) highlight the fact that sickness presenteeism has received much more attention than competitive presenteeism or non-work presenteeism. Indeed, as the first type of presenteeism is conceptually opposed to sickness absenteeism, researchers have taken a particular interest in assessing the determinants and costs of sickness presenteeism in order to compare them with those of absenteeism. The definition given by Johns (2010) is currently the most widely used in the literature, regardless of the objectives pursued (Zhou et al., 2016; A. M. Collins et al., 2017; Nordenmark et al., 2019; Kinman, 2019; Caers et al., 2019), thus neglecting the interest that may be taken in over-work, over-investment or over-effort linked to a competitive presenteeism, or even the fact of showing up at work solely for the sake of being seen.

However, this definition, which is based solely on illness, is much narrower than the one originally developed by Cooper (1996). Furthermore, it leads to the term "sickness presenteeism", previously developed by Aronsson et al. (2000), not being used, even though it would be the most appropriate for the concept in question. Unfortunately, recent research has not been open to analysing a plurality of types of presenteeism as described by Cooper (1996). Even recently, the special issue published in the Journal of Organisational Behaviour [3], although entitled "Presenteeism at Work", devoted its entire issue to sickness presenteeism.

On the other hand, the two definitions of presenteeism in the Cambridge (2020) [4] and Oxford (2020) [5] dictionaries clearly emphasise that it is linked to showing up at work primarily to signal one's presence and perceived importance to the company, or to demonstrate intense effort. However, these do not provide information on the nature of the effort behind this behaviour. Is it a productive behaviour or, as implied

3. Presenteeism at Work: Developments and Avenues for Future Research, Volume 44, Issue 6 in the Journal of Organisational Behaviour (2023).

4. "the act of staying at work longer than usual to show that one is working hard and is important to one's employer".

5. "the practice of spending more time at work than necessary, as specified in one's contract, in order to show that one is working hard".

by Hemp (2004), a passive and unproductive behaviour associated with simply being seen?

It is preferable not to attach any motivations or consequences to the definition of presenteeism, an approach also adopted by Johns (2010) in relation to sickness presenteeism. As pointed out by Kinman (2019), it does not seem relevant to focus on the worker's loss of productivity. Indeed, presenteeism – whether driven by illness or competitiveness – can be a behaviour involving a productive effort. Therefore, if reduced productivity is included in the definition, an individual who attends work while sick but remains fully productive would not be considered a presenteeist.

If this is the case with regard to the consequences of presenteeism, the same applies to its objectives. Indeed, the motivations of an individual to engage in presenteeism due to illness (Aronsson et al., 2000) seem to be of a different nature from those leading to "non-work" presenteeism (D'Abate & Eddy, 2007). Conversely, the motivations associated with competitive presenteeism (Simpson, 1998) may be closely related to those that lead to sickness presenteeism.

We can therefore see from all these definitions that the inconsistency between the various studies conducted on presenteeism behaviour stems from different interpretations of the behaviour itself. The choice to include a consequence or motivation in the definition leads to a restriction of the scope of analysis. Intrinsically linking a characteristic to the definition makes it impossible, by design, to study its invalidity. The terms and characteristics used to define a behaviour are therefore of crucial importance in determining the direction that research on the subject will take. For instance, recent research is increasingly focusing on the potential benefits of presenteeism (Wang et al., 2022; Patel et al., 2023): maintaining social ties with colleagues, interest in participating in a productive and/or social objective, but also the fact that, from the employer's point of view, presenteeism may simply be a desirable behaviour and preferable to absenteeism.

While it can be considered that a consensus has emerged since Johns (2010) on the interest of studying presenteeism linked to illness, this should not exclude the study of other forms. This emerging consensus around illness-related presenteeism, which nevertheless uses different terminology to express the same behaviour, isolates a field of research that appears to be broader. This review of the definitions of

presenteeism highlights the importance and necessity, before any further analysis, of developing a clear definition for each form of presenteeism, but also of reorganising the overall empirical research framework in which these behaviours are classified.

2 Definitions and terminology

Following the observations made in the literature review above, this section first proposes to define a generic behaviour associated solely with the term *presenteeism*, then defines two specific behaviours, *sickness presenteeism* and *competitive presenteeism*, as well as a cheating behaviour on presenteeism. The question of the effort underlying the behaviour will also be addressed. This will lead us to consider the issue of moral hazard in the context of presenteeism. In order to avoid restricting the definitions, no motivation or consequence for the behaviours is presupposed, a methodology also supported by Johns (2010), McGregor et al. (2018) and Caers et al. (2019). It is therefore up to each researcher to study all the possible motivations and consequences associated with the different forms of presenteeism. Finally, unlike previous research that defines presenteeism to reflect the observations, we have chosen to define the behaviours using a single concept that allows them to be distinguished from one another. This will enable us to study, within a clearly defined research framework, the different issues specific to each behaviour and those that may be common to all of them.

2.1 Presenteeism

We have found that using the term presenteeism alone has led to various interpretations of the behaviour without reaching a consensus on a single definition. This does not allow us to account for all the concepts associated with this behaviour. Consequently, we have chosen to consider the term presenteeism as a root, defining it in the broadest possible way. To do this, we drew on the first definition provided by Cooper (1996) to establish our initial proposal:

***Proposition* 1.** *Presenteeism* refers to the behaviour of individuals who, in the context of their work, make an effort when they should not.

Thus, in its broadest sense, presenteeism corresponds to making an effort at work when one should not. The empirical or theoretical study of such a definition is not really feasible because its scope of interpretation is too broad. The challenge of analysing such behaviour lies in defining what the phrase 'working when one should not' means. Furthermore, it is important to note that we have chosen not to link the act of presenteeism to workplace. Doing so would have effectively excluded a number of professions as well as teleworking. Until the Covid-19 crisis, very few studies had focused on presenteeism behaviour outside the workplace, particularly in the context of working from home[6]. However, working from home does require a redefinition of absenteeism and presenteeism centred around the effort. A new aspect that is incorporated into the definition of this behaviour. Presenteeism defined in this way implies that workers who engage in this behaviour do so in a committed manner in order to perform the tasks for which they are employed. Once presenteeism has been defined in this way, we can focus on more specific presenteeism behaviours. Below, we will define only the two forms of presenteeism most commonly discussed in the literature: sickness presenteeism and competitive presenteeism. Finally, in order to better understand the different orientations of the literature, we will look at the issue of moral hazard associated with a cheating behaviour associated with presenteeism.

2.2 Sickness presenteeism

In light of the existing literature, there appears to be general agreement on the definition of presenteeism when it relates to health status. As firstly described by Aronsson et al. (2000) under the term "sickness presenteeism", and later summarised by Johns (2010) simply as "presenteeism", this behaviour – defined as attending work despite illness – remains subject to debate, particularly regarding the terminology to be used and the theoretical framework it belongs to. By adopting the term presenteeism to refer to illness-related behaviour, Johns (2010) effectively rules out its use to describe more general forms of the behaviour.

6. For some examples of recent research on this topic, see the papers by Brosi & Gerpott (2022), Schmitz et al. (2023), Fiorini (2024) and Nowrouzi-Kia et al. (2024).

Since the term presenteeism has been defined in proposal 1 as the root concept under study, we prefer to associate the term sickness presenteeism to refer specifically to behaviour linked to the individual's state of health. Thus, in proposal 1, the verbal phrase "when they should not" is materialised here by the situation in which an individual is ill and should be resting rather than working. Furthermore, in order to restrict the scope of study as little as possible, the infinitive "to work" is preferred to "to attend". This means that we focus on sickness presenteeism as the act of making an effort and not simply the act of attending the workplace. Focusing on the effort made by the individual allows for a more precise understanding of the behaviour. Indeed, simply showing up does not provide information about the behaviour adopted by the individuals when they are at their workplace. This lack of precision leaves ambiguity as to whether sickness presenteeism could be the simple act of showing up while sick without voluntarily making the necessary effort.

With regard to the behaviour we wish to define here, the aim is to understand the extent to which individuals make an effort or invests themselves in their job even though they should not given their state of health, which would justify their absence. The following definition of sickness presenteeism is therefore proposed:

Proposition 2. *Sickness presenteeism* refers to the behaviour of individuals who decide to work even though they do not feel sufficiently well to do so.

An important aspect in defining this behaviour is to understand what being ill means. Karanika-Murray & Cooper (2018) identify four dimensions of illness: occurrence, periodicity, and the two main forms of illness, mental and physical. Garrow (2016) further highlights severity and duration as important factors influencing the decision to attend work. However, both studies overlook the importance of one's perception of his or her state of health in the decision to work or not, which we want to emphasise here. As Lohaus & Habermann (2019) explain, the subjective perception of illness is a crucial determinant in the decision to engage in specific behaviours. Capturing the individual's state of mind at the time of the decision-making is therefore essential. For example, an individual may not be ill in the medical sense but may nevertheless feel ill on a regular basis. Conversely, an individual may be ill but not feel that the illness is disabling in terms of effort.

Such examples underline the importance of subjective health perception in shaping behaviour. This is why we propose to define sickness presenteeism through the lens of the individual's perception of illness.

Furthermore, as previously indicated, no assumptions are made regarding the impact of sickness presenteeism on productivity. Indeed, an important factor remains the worker's intention. Sickness presenteeism, as suggested above with regard to presenteeism in general, is a behaviour in which individuals invest themselves in the aim of making an effort. If the productivity associated with this behaviour falls below a "normal" level, it is assumed to result from illness rather than from a deliberate intention not to work. Furthermore, it is possible that a sick worker may be able to perform his tasks as efficiently as usual, at least in the short term.

2.3 Competitive presenteeism

As previously highlighted by Ashby & Mahdon (2010), presenteeism is not only a behaviour associated with illness. Indeed, it should be noted that the first definition proposed by Cooper (1996) refers to the fact that this behaviour may be linked to excessive working hours. This second interpretation of presenteeism was addressed by Simpson (1998) and corresponds to staying at work longer than necessary. The author associates this concept with the term *competitive presenteeism*, referring to the fact that workers are in constant competition with each other.

Thus, the verbal phrase: "when they should not" in proposition 1 is materialised in this case by the fact that the individual is making a greater effort than they should. Simpson (1998), in line with the view of Cooper (1996), who argues that this behaviour leads to a loss of productivity, also associates the fact that these hours worked are not effective. However, staying longer at work does not provide any information about the behaviour that the individual will adopt during these extra hours. In other words, it does not indicate whether this is a behaviour of effort or a cheap signal.

Therefore, following our reasoning, we again prefer to use the infinitive "to work" instead of "to attend" in order to emphasise that competitive presenteeism is not simply a cheap signal but rather a behaviour that is costly in effort. We consider that individuals, in competition with their colleagues, remain at their workplace in

order to work more than they should, with the aim of signalling their importance to the company, probably with a view to obtaining a promotion. The objective would therefore to make an extra effort to increase productivity, rather than simply showing up for the purpose of being seen. This choice allows us to broaden the significance of such behaviour, which may include excessive work behaviours such as taking work home in the evenings or at weekends, constantly taking work calls, working unpaid overtime, etc.

Furthermore, it is important to clearly define what "working more than necessary" means. Indeed, this refers to the fact that there is a work situation or intensity of effort required to perform a task that is considered as "normal" or as a "reference". In the first case, it is possible to empirically assess a so-called normal situation. This would correspond to the average or median level of effort or number of hours worked by an employee or a group of employees with similar employment characteristics. In this way, the behaviour of working more than necessary would be linked to a statistical measure of the deviation from the average for the profession, the sector or the individual himself. It would therefore be possible to measure the intensity of engagement in competitive presenteeism behaviour. In the second case, the measurement would be linked to a reference defined by the employment contract, either explicitly or implicitly. It would therefore be the deviation between the measurement and the statutory level specified in the contract that would establish an individual's level of investment in competitive presenteeism behaviour.

The choice between these two measures may depend on the type of contract to which the worker is subject. In France, for example, executives do not always have a specified number of contractual working hours. As a result, it is not always possible to rely on a contractual reference to assess the involvement in competitive presenteeism. To avoid restricting our definition, we therefore retain the idea that individuals work more than they should, without linking this to a specific measure. Given the data available to the authors, the method of measurement is deliberately left open. On this basis, the following definition of competitive presenteeism is proposed:

***Proposition* 3.** *Competitive presenteeism* refers to the behaviour of individuals who work more than they should.

Finally, in each of the above definitions, we orient the intention of the worker who engages in presenteeism towards a productivist vision. However, the various measures of the phenomenon already existing in the literature do not allow us to identify the worker's intention. In other words, when we say that the individuals come to work sick for instance, we do not know whether they are doing so in order to make an effort or simply to demonstrate their presence. The same is true for competitive presenteeism: we do not know whether the individuals are working longer hours during which they are making an effort (which may be less effective due to fatigue), or whether they are intentionally working fewer hours (the inefficiency then stemming from the worker's decision not to make the necessary effort). For both sickness presenteeism and competitive presenteeism, we would then have a productive behaviour (production of a voluntary effort) and an unproductive behaviour (non-production of a voluntary effort). The following section thus addresses the issue of the observability of the effort behind presenteeism.

2.4 Presenteeism and moral hazard

It is important to highlight the existence of a behaviour that is ultimately quite different from the idea of presenteeism defined above, but which is nevertheless often mentioned in the literature. Similar to the shirking behaviour associated with absenteeism (Lazear, 1979), it is also possible to cheat on presenteeism. Indeed, a behaviour often associated with presenteeism, introduced by Hemp (2004) through the quote from Woody Allen, consists of going to work for the sole purpose of being present. This behaviour is frequently implied in definitions of sickness presenteeism or competitive presenteeism. For example, in the case of sickness presenteeism, this behaviour would allow the employee to alert their employer that they are coming to work even though they are ill. However, the interpretation here would be that this employee, who comes to work ill, is not actually working, and is doing so intentionally. Furthermore, in the case of competitive presenteeism, this would correspond to the employee working overtime solely for the purpose of being seen at work without actually putting in any effort during these extra working hours.

These interpretations of presenteeism behaviour do not correspond, in this case, to the proposals 2 and 3, insofar as sickness presenteeism and competitive presen-

teeism are associated with the act of making a productive effort. In other words, being present at work is assumed to be active rather than passive. On the contrary, in the definition given by D'Abate & Eddy (2007), which reorients the interpretation of presenteeism as simply being at work without devoting time and energy to it, the attendance is passive. The behaviour described by these authors ultimately corresponds to a cheating behaviour related to presenteeism, which refers to shirking as described in the literature about reduced effort (Shapiro & Stiglitz, 1984) or absenteeism (Lazear, 1979).

The issue of cheating on effort is not new. It essentially reflects a problem of moral hazard linked to the unobservability of individual effort at work. When effort cannot be directly observed, workers may remain longer in the workplace while engaging in personal activities. In a competitive environment, it is easy to imagine that a worker might adopt presenteeism behaviours – whether sickness-related or competitive – without actually providing the corresponding effort, simply to be seen and avoid being perceived as uncommitted. In this case, the only cost of being present would be compensated by the signalling benefits it provides. Moreover, as suggested by D'Abate & Eddy (2007), this cost of presence may be further reduced when working time is reallocated to personal tasks. Ultimately, such behaviour is only possible in contexts where neither performance, which could serve as a basis for piece-rate compensation, nor individual effort is directly observable, unlike physical presence at work.

The analysis of this behaviour is constrained by the identification of the worker's intention and their actual productive performance in the specific context of presenteeism, and will therefore not be studied in this thesis. Indeed, we do not have data containing any indication of the worker's individual productivity. Nevertheless, it is worth considering the possibility of cheating on presenteeism in order to gain a comprehensive understanding of the mechanisms associated with the behaviour and to propose a comprehensive analytical framework for future research.

3 Research framework

This thesis is exclusively based on an empirical approach to studying the determinants and consequences of presenteeism and absenteeism. The development of research over the last two decades on presenteeism, which is often defined in various ways, has not facilitated the understanding of the phenomenon. However, a consensus on its definition is essential, as this is the basis for the development of empirical measures and tools for its assessment (Karanika-Murray & Cooper, 2018). A precise definition of each form of presenteeism, established in relation to a reference behaviour, facilitates the organisation of the research areas associated with them.

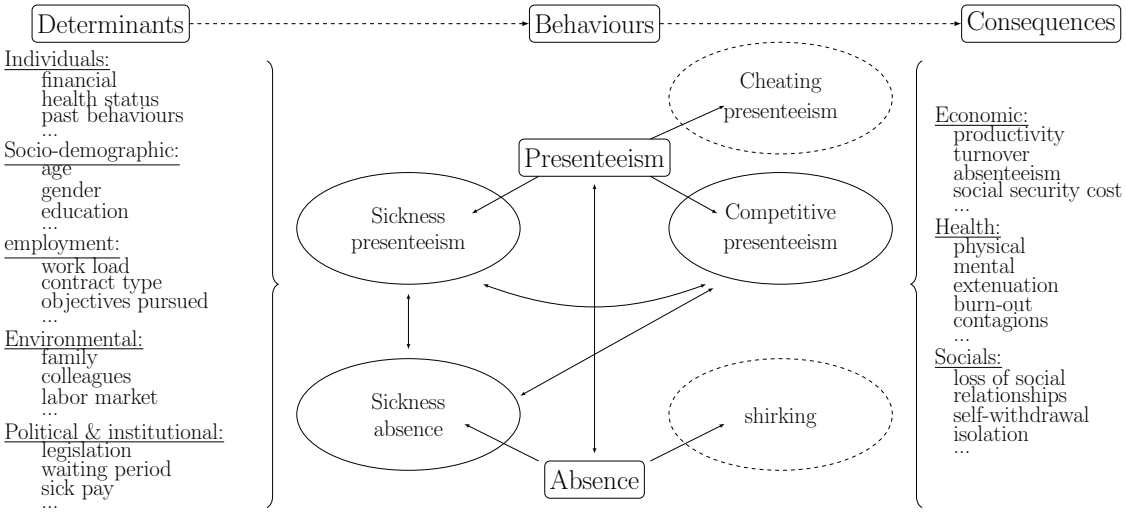

Figure 1 – Empirical research framework

Figure 1 above presents an analytical framework identifying the interactions that may exist between behaviours, their determinants and their consequences (in a non-exhaustive way). This provides an overview and helps to organise empirical research. For instance, research can focus on the relationships between the different behaviours studied. These include the relationship between presenteeism and absence in general, between sickness presenteeism and sickness absence more specifically, or between competitive presenteeism and sickness presenteeism. It would also be interesting to study the link between these behaviours and the cheating behaviour. As the latter is more difficult to identify in the data, this implies a more theoretical approach to the study of this link. Research can also focus on the determinants and consequences of these behaviours, given the links they may have.

Some determinants may be common to several forms of presenteeism as well as to absenteeism. For example, sickness presenteeism may be closely linked to competitive presenteeism, insofar as similar motivations may lead to both of them. Believing that their presence demonstrates their commitment to their employer, workers may adopt competitive presenteeism which, when they fall ill, leads them to engage in sickness presenteeism rather than taking sick leave. On the other hand, mechanisms that encourage attendance, such as waiting days or the level of sick pay, are supposed to have an effect on sickness absenteeism and, consequently, on sickness presenteeism, but should not be associated with competitive presenteeism.

Nevertheless, it is also possible that the motivations are very different. Sickness presenteeism could be explained more by characteristics of pressure at work, leading to involuntary behaviour driven by the occurrence of illness, while competitive presenteeism is supposed to be behaviour motivated by the pursuit of rewards in terms of career and wages, reflecting voluntary behaviour.

In fact, the same factor may be important in explaining two behaviours that are similar but nevertheless different in nature. If work pressure is indeed correlated with a competitive environment, individuals may invest more in their job in order to obtain a promotion, but, in the event of illness, suffer this pressure and prevent themselves from taking time off. On the other hand, the consequences of each type of presenteeism behaviour can be radically different for the employer. While there are regular references to a decline in productivity among employees who engage in sickness presenteeism, employees who engage in competitive presenteeism may, on the contrary, produce more than what is stipulated in their contract. However, the consequences of these two behaviours on the individual's health can be very similar. Sickness presenteeism can worsen the individual's health in the long term, or even lengthen their recovery period. Similarly, over-commitment to work through competitive presenteeism can also lead to health problems in the future (e.g. exhaustion, stress and burnout).

The aim of this thesis is therefore to explore various aspects relating to the research framework outlined above. It comprises three chapters, which are briefly presented below. The first chapter examines the determinants of both sickness absenteeism and sickness presenteeism. It seeks to determine whether the same

determinants can explain both behaviours or whether different factors are involved. In a similar vein, the second chapter examines the link between sickness presenteeism and absence. The question raised in this chapter is that of the relative consequences of these two behaviours on health in the longer term. Finally, the last chapter examines the economic and health consequences of competitive presenteeism. This behaviour has received very little attention in the literature, compared to sickness presenteeism. This final chapter aims to assess the extent to which over-investment in work can be perceived as a valuable signal in terms of career advancement, while analysing its potentially harmful effects on health.

In addition, this thesis uses recent statistical methods based on copulas to address some specific issues. Copulas are used in the first chapter to analyse endogenous participation associated with illness. It is assumed that characteristics related to sickness absence and sickness presenteeism primarily explain the individual's state of health. Not taking these effects into account would lead to a bias in estimating the impact of these characteristics on the behaviours studied. In the second and third chapters, copulas are used to address the endogeneity of the explanatory variables. Given the lack of information preventing the use of instrumental variables, as well as recent debates on the validity of the instruments, the use of copulas seems particularly appropriate.

Chapter 1 - A new empirical model of the determinants of sickness and the choice between presenteeism and absence

This first chapter, co-authored with Joseph Lanfranchi and published in Labour [7], looks at the factors that influence both sickness presenteeism and absenteeism. Despite the extensive research devoted to both sickness absence and presenteeism, the two bodies of literature have largely evolved separately (Halbesleben et al., 2014). Only a limited number of studies have examined how worker characteristics and employment conditions jointly influence both behaviours (Hansen & Andersen, 2008; Leineweber et al., 2012; Arnold & de Pinto, 2015; Gerich, 2016). The objective here is to identify, using the same sample of observations, which employment charac-

7. Denis, T. and Lanfranchi, J. (2025), A New Empirical Model of the Determinants of Sickness and the Choice Between Presenteeism and Absence. Labour, 39: 61-87. https://doi.org/10.1111/labr.12286

teristics explain both sickness presenteeism and absenteeism, and conversely, which factors are associated with only one of these behaviours. This distinction is essential when seeking to design an effective human resources policy, whether the goal is to encourage absence in the event of illness or, on the contrary, to promote attendance, thereby avoiding an excessive loss of productivity for instance. Failing to take into account the causes of sickness presenteeism when developing a policy aimed at reducing absence could render such a policy counterproductive if it simultaneously increases the costs associated with sickness presenteeism. Indeed, given the costs linked to both behaviours, firms must carefully target the factors that encourage employees to adopt the desired one.

Furthermore, this chapter proposes to correct the estimation of the effects of employment characteristics on both behaviours in terms of their influence on health. Indeed, the study of sickness absence and presenteeism requires taking into account the endogenous selection bias related to individuals' health. The probability of falling ill is certainly not exogenous to employment characteristics and working conditions (Green et al., 2021; Havet & Plantier, 2023). In this way, the two-stage model proposed in this chapter makes it possible to identify, on the one hand, the determinants of illness and, on the other hand, the determinants that more regularly lead to the choice of absence or presence in case of illness. To do this, data from the 2016 'working conditions and psychosocial risks' survey, based on a representative sample of the French population, are used. Like in the paper by Arnold & de Pinto (2015), this analysis focuses on the annual duration of sickness absence and presenteeism. From an economic point of view, studying duration appears to be more relevant than studying the frequency of behaviours. Finally, the use of a copula-based approach (Marra & Wyszynski, 2016) allows for great flexibility in estimating the distributions of sickness absence and sickness presenteeism, in addition to considering endogenous participation bias.

First of all, the results indicate that without the correction associated with endogenous participation, the estimates would have been biased and the conclusions regarding the determinants involved in the durations of sickness absence and presenteeism would have been different. Among other findings, it is important to emphasise the importance of implementing occupational health protection policies, given the

large number of factors related to employment and working conditions that are primarily associated with illness. By reducing the likelihood of illness, the company prevents any risk of its employees engaging in sickness presenteeism or taking sick leave. Our results thus reveal that certain determinants previously identified in the literature as affecting sickness absenteeism and presenteeism are ultimately only associated with illness. Among these, constraints on the pace of work (due to a machine, external demand, standards, etc.) are only associated with a higher probability of illness, without impacting either the duration of absence or the duration of sickness presenteeism.

Finally, the results indicate that the determinants associated with the duration of sickness absence or sickness presenteeism are not of the same nature. The duration of sickness absence is more closely associated with the characteristics of the employment contract, while the determinants associated with the duration of sickness presenteeism relate to the employees' working conditions. This reveals that it is more complex for firms to take effective action on presenteeism than on absenteeism. Indeed, reducing the duration of sickness presenteeism must be achieved through prevention policies and effective training of managers to identify risk situations, rather than through mechanisms associated with the employment contract, as is the case for absence. Reducing the duration of presenteeism must therefore be encouraged by a work environment free of pressure, with a good division of workload and mutual support between employees as well as between supervisors and their subordinates. These results provide a better understanding of how to address sickness presenteeism and absenteeism by clearly identifying the drivers that can influence one or both. In this respect, this study sheds new light on the characteristics that companies wishing to establish effective human resources policies should target.

Chapter 2 - All work and no rest makes Jack a sick boy: A causal analysis of sickness absence and presenteeism on health

In the second chapter, the focus is on the consequences of sickness absenteeism and presenteeism in terms of future health. In particular, sickness presenteeism is

considered to be a risk factor for future health (Bergström et al., 2009a; Gustafsson & Marklund, 2011; Dellve et al., 2011) and increases the risk of burnout (Demerouti et al., 2009) as well as the likelihood of taking sick leave (Bergström et al., 2009b; Janssens et al., 2013). On the contrary, the intuition put forward is that absenteeism is beneficial to health by allowing time for physical and mental rest (Aronsson et al., 2000). For example, Aboagye et al. (2019) showed that exhaustion was associated with past sickness presenteeism but not with absenteeism. Thus, a part of future absenteeism could be avoided by good management of presence and absence in the event of illness.

However, in the literature cited above, when the estimate includes past absenteeism as an explanatory variable in the models, a negative relationship with future health emerges. This result would indicate that workers with the highest absenteeism rates would see their health deteriorate further, which appears in contradiction with the proposed intuition. It is therefore assumed that sickness absence and presenteeism are two indicators of health. As such, since these behaviours are fundamentally linked to illness, the more frequently and for longer periods individuals exhibit these behaviours, the more this would indicate poor health. Conversely, individuals in good health would not often report being absent or present when ill, given their robustness to illness. This principle is similar to the 'healthy worker effect' in organisational medicine (Haan & Myck, 2009) or the 'healthy user effect' in pharmacoepidemiology (Shrank et al., 2011). The robustness or fragility of individuals, which would explain their level of sickness absence, sickness presenteeism and health, is not directly observable in the data. This therefore creates an endogeneity problem. In this chapter, it is therefore assumed that this negative relationship between past absenteeism and future health is the result of the endogeneity associated with this behaviour. Similarly, we may ask to what extent the endogeneity associated with sickness presenteeism affects the estimates.

This chapter therefore explores the causal relationship between longer durations of sickness absence and presenteeism and future health indicators such as duration of absence, total duration of illness (calculated as the sum of the duration of absence and presenteeism), self-reported health status and mental health. To do this, the two waves of the 2013 and 2016 'working conditions and psychosocial risks' survey

are used. These two waves constitute a panel covering a sample of the French working population. In order to explore and control for the effect of the endogeneity of behaviours on future health, we propose to use a copula-based approach, which does not require any instruments (Yang et al., 2024).

Without correcting for endogeneity, the results, which appear to be consistent with previous studies, show positive correlations between both past sickness absence and future absence and total sickness duration, and between sickness presenteeism and these same variables. When endogeneity is taken into account, the results confirm the intuition that absence is a protective behaviour for health. An additional day of absence in the initial period would lead to a 1.7% decrease in the future duration of sickness absence. This effect also appears to be decreasing and monotonic. On the other hand, the duration of sickness presenteeism does not appear to lead to an increase in the duration of future absenteeism, but has a U-shaped effect on the total duration of future illness, mainly driven by future sickness presenteeism. Thus, unlike absence, a long period of presenteeism would lead to a deterioration in individual health and an increase in the total duration of illness. The decreasing effect observed for shorter periods of sickness presenteeism could be associated with the fact that this behaviour served as a signal in the context of competition for promotion or to avoid dismissal. In this case, once the promotion has been obtained or the period of redundancies has ended, the worker may reduce his duration of sickness presenteeism. Furthermore, this U-shaped effect can be explained by the high heterogeneity within the population of individuals who report a positive duration of presenteeism. Taking into account the heterogeneity associated with self-reported health status, in addition to endogeneity, the results confirm an increase in the duration of future absences and the total duration of future illness associated with a longer duration of sickness presenteeism at baseline. This detrimental effect is all the more significant when the baseline self-reported health status is good. Workers who initially report very good health would see their future total illness duration increase by nearly 20% for each additional day of sickness presenteeism.

Furthermore, the increase in the total duration of future illness, caused by a longer duration of past sickness presenteeism, would be driven by the increase in the future duration of sickness presenteeism rather than by the increase in the future

duration of absence, indicating a long-term preference for presenteeism. In addition to these results, longer duration of presenteeism at baseline is also associated with a significant decline in psychological well-being, unlike absence. Finally, all the parameters capturing endogeneity show the expected signs, probably indicating the persistent effects of poor health on future health.

These results once again offer interesting insights for guiding corporate human resources policies. A cost-benefit analysis would be necessary to arbitrate between absence and presence in case of illness. However, it seems relevant to assess the long-term effects of each behaviour. This cost-benefit arbitration must be carried out in an intertemporal manner. Sickness absence appears to be a good way to maintain a healthier workforce in the long term, unlike sickness presenteeism. Thus, a policy that encourages attendance must take into account the fact that, in the event of illness, it is important to encourage absence, otherwise there is a risk of increasing the long-term cost of this policy. This chapter contributes to the literature by exploring the endogeneity associated with sickness absence and presenteeism, revealing the beneficial effect of absence in the event of illness and confirming the harmful effect of sickness presenteeism.

Chapter 3 - Is presenteeism such a valuable signal for employees? Analysis on wages, career, health and well-being

Finally, unlike the previous chapters, the third chapter does not focus on sickness presenteeism but on competitive presenteeism. Indeed, as indicated by Ashby & Mahdon (2010), presenteeism is not a behaviour solely associated with illness. However, its other forms have not received much attention in previous literature. Competitive presenteeism can be interpreted, for example, as taking work home in the evenings and weekends or staying at work to do overtimes that are not compensated, either in wages or in kind. If employees engage in over-investment behaviours when it is not directly required by the firm, it is probably because they nevertheless perceive it as useful. The theoretical framework proposed for analysing this behaviour is thus based on the signalling theory (Spence, 1973).

The intuition which is tested in this chapter is the signalling value of competitive presenteeism. In an environment where employees are constantly competing

with each other to advance their careers within their company (Holmstrom, 1982), it would be in their interest to demonstrate their willingness to work hard, their commitment and devotion to their work. The problem a firm faces when promoting employees is selecting the right candidate. Workers who wish to demonstrate their abilities by working harder than necessary send a signal that they are more willing to make an effort than their colleagues.

This behaviour may, however, affect the individual's social relationships outside the workplace and their health if they are too dedicated to their work. Therefore, what are the consequences of such behaviour on health and well-being? For example, Avgoustaki & Frankort (2019) and Avgoustaki & Cañibano (2020) highlight the negative effects of overtime and work intensity on health and stress. In addition, according to the French Caisse Nationale d'Assurance Maladie (National Health Insurance Fund), consumption of antidepressants increased by 9.7% between 2013 and 2019, representing a cost of €1,214 million in the latter year. No less than 16% of workers in the sample studied in this thesis report often or always working unpaid overtime. 14% also report often or always taking work home in the evening. Finally, 40% reported thinking about their work regularly or always outside of working hours. Thus, the prevalence of competitive presenteeism behaviour, combined with its potential consequences on health, particularly mental health, makes it a public health issue, just like sickness presenteeism.

This chapter therefore proposes to measure the intensity of competitive presenteeism behaviour using a continuous synthetic indicator. This indicator is constructed on the basis of the first factorial axis of a multiple correspondence analysis that incorporates a set of variables that best reflect the behaviour. This strategy is also assumed to be more closely related to the multidimensional aspect of presenteeism. Next, the effect of this indicator on the individual's career and health is explored. With regard to career, the effect of this indicator on hourly wage and occupational hierarchy is studied. The effect of the indicator is also tested on self-reported health, mental well-being and the duration of sick leave. The use of the last three waves of the "working conditions and psychosocial risks" survey is particularly useful for controlling for unobserved heterogeneity. In addition, attention will be paid to the causal relationship between competitive presenteeism and the ex-

plained variables using two different approaches: one using copulas, the other using instrumental variables.

The results show that past competitive presenteeism can increase future hourly wages by up to 3.4% and the probability of hierarchical advancement by 10%. These results take individual heterogeneity into account and are robust to a wide range of control variables. This leads to the conclusion that this behaviour is indeed an effective signal for a better career. However, the results also confirm that over-investment in the job is detrimental to health, with a 25% decrease in the probability of reporting better health in the future, as well as an increase in the duration of sick leave of nearly 19% on average. Furthermore, the results concerning self-reported health tend to show a U-shaped effect, particularly after controlling for endogeneity. This positive effect observed among the most invested workers could be attributed to their state of mind and psychology: knowing that they are rewarded by their employer, they no longer feel entitled to report poor health. However, these effects on self-reported health deserve to be explored in more detail in future research. Finally, no significant association appears to exist between competitive presenteeism and mental well-being.

These results therefore indicate that while over-investment in work is well rewarded in terms of wages and career advancement, the importance of long-term health effects should not be overlooked. From the employer's point of view, this behaviour is associated with a sharp increase in future sickness absence.

Bibliography

Aboagye, E., Björklund, C., Gustafsson, K., Hagberg, J., Aronsson, G., Marklund, S., ... Bergström, G. (2019). Exhaustion and impaired work performance in the workplace: Associations with presenteeism and absenteeism. *Journal of Occupational and Environmental Medicine*, *61*, 1.

Allen, D., Hines, E. W., Pazdernik, V., Konecny, L. T., & Breitenbach, E. (2018). Four-year review of presenteeism data among employees of a large united states health care system: a retrospective prevalence study. *Human Ressources for Health*, *16*(1), 59.

Arnold, D. (2016). Determinants of the annual duration of sickness presenteeism: Empirical evidence from european data. *Labour*, *30*(2), 198–212.

Arnold, D., & de Pinto, M. (2015). How are work-related characteristics linked to sickness absence and presenteeism? theory and data. *Schmollers Jahrbuch*, *135*, 465-498.

Aronsson, G., & Gustafsson, K. (2005). Sickness presenteeism: Prevalence, attendance-pressure factors, and an outline of a model for research. *Journal of Occupational and Environmental Medicine*, *47*, 958–966.

Aronsson, G., Gustafsson, K., & Dallner, M. (2000). Sick but yet at work. an empirical study of sickness presenteeism. *J Epidemiol Community Health*, *54*, 502–509.

Ashby, K., & Mahdon, M. (2010). Why do employees come to work when ill: An investigation into sickness presence in the workplace. *London: The Work Foundation*.

Avgoustaki, A., & Cañibano, A. (2020). Motivational drivers of extensive work effort: Are long hours always detrimental to well-being? *Industrial Relations: A Journal of Economy and Society*, *59*(3), 355-398.

Avgoustaki, A., & Frankort, H. T. W. (2019). Implications of work effort and discretion for employee well-being and career-related outcomes: An integrative assessment. *ILR Review*, *72*(3), 636-661.

Bergström, G., Bodin, L., Hagberg, J., Aronsson, G., & Josephson, M. (2009b). Sickness presenteeism today, sickness absenteeism tomorrow? a prospective study on sickness presenteeism and future sickness absenteeism. *Journal of Occupational and Environmental Medicine*, *51*(6), 629–638.

Bergström, G., Bodin, L., Hagberg, J., Lindh, T., Aronsson, G., & Josephson, M. (2009a). Does sickness presenteeism have an impact on future general health? *International archives of occupational and environmental health*, *82*(10), 1179–1190.

Biron, C., Karanika-Murray, M., & Ivers, H. (2022). The health-performance framework of presenteeism: A proof-of-concept study. *Frontiers in Psychology*, *Volume 13 - 2022*.

Brosi, P., & Gerpott, F. H. (2022). Stayed at home—but can't stop working despite being ill?! guilt as a driver of presenteeism at work and home. *Journal of Organizational Behavior*.

Burton, W. N., Pransky, G., Conti, D. J., Chen, C.-Y., & Edington, D. W. (2004). The association of medical conditions and presenteeism. *Journal of Occupational and Environmental Medicine*, *46*, S38–S45.

Caers, R., Akgul, K. L., Baert, S., Feyter, T. D., & Couck, M. D. (2019). Too sick or not too sick?: The importance of stress and satisfaction with supervisor support on the prevalence of sickness presenteeism. *International Journal of Occupational Safety and Ergonomics*.

Cambridge. (2020). Presenteeism definition.

Canfield, G., & Soash, D. (1955). Presenteeism: A constructive view. *Personal Journal*, *34*.

Chapman, L. S. (2005). Presenteeism and its role in worksite health promotion. *American Journal of Health Promotion*, *19*(4), 1–8.

Collins, A. M., Cartwright, S., & Cowlishaw, S. (2017). Sickness presenteeism and sickness absence over time: A uk employee perspective. *Work & Stress*, 1-16.

Collins, J., Baase, C., Sharda, C., Ozminkowski, R., Nicholson, S., Billotti, G., ... Berger, M. (2005). The assessment of chronic health conditions on work performance, absence, and total economic impact for employers. *Journal of occupational and environmental medicine / American College of Occupational and Environmental Medicine*, *47*, 547-57.

Cooper, C. L. (1996). Hot under the collar. *Times Higher Education Supplement*, 15.

Cour des Comptes. (2024). *Communiqué de presse. arrêt de travail pour maladie : Mise au point du premier président de la cour des comptes* (Rapport technique). Consulté sur https://www.ccomptes.fr/fr/documents/70036

D'Abate, C. P., & Eddy, E. R. (2007). Engaging in personal business on the job: Extending the presenteeism construct. *Human Resource Development Quarterly*, *18*(3), 361–383.

Dellve, L., Hadzibajramovic, E., & Ahlborg, G. (2011). Work attendance among healthcare workers: Prevalence, incentives, and long-term consequences for health and performance. *Journal of advanced nursing*, *67*, 1918-29.

Demerouti, E., Le Blanc, P. M., Bakker, A. B., Schaufeli, W. B., & Hox, J. (2009). Present but sick: a three-wave study on job demands, presenteeism and burnout. *Career Development International*, *14*(1), 50–68.

Evans-Lacko, S., & Knapp, M. (2016). Global patterns of workplace productivity for people with depression: absenteeism and presenteeism costs across eight diverse countries. *Social Psychiatry and Psychiatric Epidemiology Journal*, *51*, 1525–1537.

Fiorini, L. A. (2024). Remote workers' reasons for changed levels of absenteeism, presenteeism and working outside agreed hours during the covid-19 pandemic. *SAGE Open*, *14*(1), 21582440241240636.

Gallie, D., & Zhou, Y. (2013). Job control, work intensity, and work stress. In *Economic crisis, quality of work, and social integration: The european experience.* Oxford University Press.

Garrow, V. (2016). Presenteeism: A review of current thinking. *Institute for Employment Studies, Report 507*.

Gerich, J. (2016). Determinants of presenteeism prevalence and propensity: Two sides of the same coin? *Archives of Environmental & Occupational Health*, *71*(4), 189–198.

Goetzel, R. Z., Carls, G. S., Wang, S., Kelly, E., Mauceri, E., Columbus, D., & Cavuoti, A. (2009). The relationship between modifiable health risk factors and medical expenditures, absenteeism, short-term disability, and presenteeism among employees at novartis. *Journal of Occupational and Environmental Medicine*, *51*(4), 487–499.

Green, F., Felstead, A., Gallie, D., & Henseke, G. (2021). Working still harder. *ILR Review*, *75*(2), 458–487.

Gustafsson, K., & Marklund, S. (2011). Consequences of sickness presence and sickness absence on health and work ability: A swedish prospective cohort study. *International Journal of Occupational Medicine and Environmental Health*, *24*, 153–165.

Gustafsson Sendén, M., Schenck-Gustafsson, K., & Fridner, A. (2016). Gender differences in reasons for sickness presenteeism - a study among gps in a swedish health care organization. *Annals of Occupational and Environmental Medicine*, *28*, 50.

Haan, P., & Myck, M. (2009). Dynamics of health and labor market risks. *Journal of Health Economics*, *28*(6), 1116-1125.

Halbesleben, J. R., Whitman, M. V., & Crawford, W. S. (2014). A dialectical theory of the decision to go to work: Bringing together absenteeism and presenteeism. *Human Resource Management Review*, *24*, 177–192.

Hansen, C. D., & Andersen, J. H. (2008). Going ill to work – what personal circumstances, attitudes and work-related factors are associated with sickness presenteeism? *Social Science & Medicine*, *67*, 956–964.

Havet, N., & Plantier, M. (2023). The links between difficult working conditions and sickness absences in the case of french workers. *Labour*, *37*, 160–195.

Helgesson, M., Gustafsson, K., Marklund, S., Aronsson, G., & Leineweber, C. (2021). Sickness absence and sickness presence among health and care employees in sweden - health complaints, health behavior, and future long-term sickness absence. *Journal of occupational and environmental medicine*, *Publish Ahead of Print*.

Hemp, P. (2004). Presenteeism: At work – but out of it. *Harvard Business Review*, *82*(10), 49–58.

Holmstrom, B. (1982). Moral hazard in teams. *The Bell Journal of Economics*, *13*(2), 324–340.

Ishimaru, T., Mine, Y., & Fujino, Y. (2020). Two definitions of presenteeism: sickness presenteeism and impaired work function. *Occupational Medicine*.

Janssens, H., Clays, E., De Clercq, B., De Backer, G., & Braeckman, L. (2013). The relation between presenteeism and different types of future sickness absence. *Journal of occupational health*, *55*.

Johns, G. (2010). Presenteeism in the workplace: A review and research agenda. *Journal of Organizational Behavior*, *31*, 519–542.

Karanika-Murray, M., & Cooper, C. L. (2018). Presenteeism: an introduction to prevailing global phenomenon. *Cambridge Core*, 1–34.

Kinman, G. (2019). Sickness presenteeism at work: prevalence, costs and management. *British Medical Bulletin*, 1–10.

Kinman, G., & Wray, S. (2018). Presenteeism in academic employees - occupational and individual factors. *Occupational Medicine*, *68*, 46–50.

Lazear, E. P. (1979). Why is there mandatory retirement? *Journal of Political Economy*, *87*(6), 1261–1284.

Leineweber, C., Westerlund, H., Hagberg, J., Svedberg, P., & Alexanderson, K. (2012). Sickness presenteeism is more than an alternative to sickness absence: results from the population-based slosh study. *International Archives of Occupational and Environmental Health*.

Lohaus, D., & Habermann, W. (2019). Presenteeism: A review and research directions. *Human Ressource Management Review*, *29*, 43–58.

Malmberg-Ceder, K., Vuorio, T., Korhonen, P., Kautiainen, H., Soinila, S., & Haanpää, M. (2020). The impact of self-reported recurrent headache on absenteeism and presenteeism at work among finnish municipal female employees. *Journal of Pain Research*, *13*, 2135-2142.

Marie, O., & Val Castello, J. (2023). Sick leave cuts and (unhealthy) returns to work. *Journal of Labor Economics*, *41*, 923–956.

Marra, G., & Wyszynski, K. (2016). Semi-parametric copula sample selection models for count responses. *Computational Statistics and Data Analysis*, *104*, 110-129.

McGregor, A., Sharma, R., Magee, C., Caputi, P., & Iverson, D. (2018). Explaining variations in the findings of presenteeism research: A meta-analytic investigation into the moderating effects of construct operationalizations and chronic health. *Journal of Occupational Health Psychology*, *23*(4), 584–601.

Miraglia, M., & Johns, G. (2016). Going to work ill: A meta-analysis of the correlates of presenteeism and a dual-path model. *Journal of Occupational Health Psychology*, *21*(3), 261–283.

Navarro, A., Salas-Nicás, S., Llorens, C., Moncada, S., Molinero-Ruíz, E., & na, D. M. (2019). Sickness presenteeism: Are we sure about what we are studying? a research based on a literature review and an empirical illustration. *American Journal of Industrial Medicine*, 1–10.

Nordenmark, M., Hagqvist, E., & Vinberg, S. (2019). Sickness presenteeism among self-employed and employed in northwestern europe – the importance of time demands. *Safety and Health at Work*.

Nowrouzi-Kia, B., Haritos, A. M., Long, B.-Z. S., Atikian, C., Fiorini, L. A., Gohar, B., ... Bani-Fatemi, A. (2024). Remote work transition amidst covid-19: Impacts on presenteeism, absenteeism, and worker well-being—a scoping review. *PLOS ONE*, *19*(7), 1-20.

Oxford. (2020). Presenteeism definition. *Oxford Learners Dictionaries*.

Patel, C., Biron, M., Cooper, S. C., & Budhwar, P. S. (2023). Sick and working: Current challenges and emerging directions for future presenteeism research. *Journal of Organizational Behavior*, *44*(6), 839-852.

Ruhle, S., Breitsohl, H., Aboagye, E., Baba, V., Biron, C., Correia Leal, C., ... Muschalla, B. (2019). "to work, or not to work, that is the question" - recent trends and avenues for research on presenteeism. *European Journal of Work and Organizational Psychology*, *29*, 344-363.

Schmitz, H., Bauer, J. F., & Niehaus, M. (2023). Working anytime and anywhere -even when feeling ill? a cross-sectional study on presenteeism in remote work. *Safety and Health at Work*, *14*(4), 375-383.

Schultz, A. B., & Edington, D. W. (2007). Employee health and presenteeism: A systematic review. *Journal of Occupational Rehabilitation*, *17*, 547–579.

Shapiro, C., & Stiglitz, J. E. (1984). Equilibrium unemployment as a worker descipline device. *The American Economic Review*, *74*(3), 433–444.

Shrank, W., Patrick, A., & Brookhart, M. (2011). Healthy user and related biases in observational studies of preventive interventions: A primer for physicians. *Journal of general internal medicine*, *26*, 546-50.

Simpson, R. (1998). Presenteeism, power and organizational change: Long hours as a career barrier and the impact on the working lives of women managers. *British Journal of Management*, *9*(s1), 37–50.

Skagen, K., & Collins, A. M. (2016). The consequences of sickness presenteeism on health and wellbeing over time: A systematic review. *Social Sciences & Medicine*, *161*, 169–177.

Smith, D. J. (1970). Absenteeism and presenteeism in industry. *Archives of Environmental Health*, *21*, 670–677.

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, *87*(3), 355–374.

Stolz, R. (1993). Reducing turnover through incentive programs. *Cornell Hotel and Restaurant Administration Quarterly*, *34*, 79–80.

Triolle, A., Lefevre, N., Leclercq, M., Hairault, A., Lianova, R., & Pavis, I. (2024). *Revue de dépenses relative à la réduction des absences dans la fonction publique* (Rapport technique). Inspection générale des finances et Inspection générale des affaires sociales.

Wang, M., Lu, C.-q., & Lu, L. (2022). The positive potential of presenteeism: An exploration of how presenteeism leads to good performance evaluation. *Journal of Organizational Behavior*, *44*.

Worrall, L., Cooper, C., & Campbell, F. (2000). The new reality for uk managers: Perpetual change and employment instability. *Work, Employment & Society*, *14*, 647–668.

Yang, F., Qian, Y., & Xie, H. (2024). Addressing endogeneity using a two-stage copula generated regressor approach. *Journal of Marketing Research*, *0*(0), 00222437241296453.

Zhang, W., Sun, H., Woodcock, S., & Anis, A. (2015). Illness related wage and productivity losses: Valuing 'presenteeism'. *Social Science & Medicine*, *147*, 62–71.

Zhou, Q., Martinez, L. F., Ferreira, A. I., & Rodrigues, P. (2016). Supervisor support, role ambiguity and productivity associated with presenteeism: A longitudinal study. *Journal of Business Research*, *69*, 3380–3387.

Chapter 1

A new empirical model of the determinants of sickness and the choice between presenteeism and absence

Denis, T. and Lanfranchi, J. (2025), A New Empirical Model of the Determinants of Sickness and the Choice Between Presenteeism and Absence. Labour, 39: 61-87. https://doi.org/10.1111/labr.12286

1 Introduction

Several recent analyses have highlighted the detrimental influence that employment increasingly exerts on employees' health, emphasizing the need to strengthen its rewarding and protective aspects (Hudson et al., 2019; Sorensen et al., 2021). The combination of persistent physical hardship and growing time constraints imposed on employees contributes to a deterioration in occupational health (Green et al., 2021; Kelly & Moen, 2020). This decline is further exacerbated by the harmful influence of psychosocial work factors (Niedhammer et al., 2013; Havet & Plantier, 2023) as well as disruptive organizational changes (Ben Halima et al., 2023).

This accumulation of hardships therefore leads to illness, leaving the employee with two main alternatives: either sickness absence (SA), staying home to take care of himself (Treble & Barmby, 2011), or sickness presenteeism (SP), going to work despite feeling ill (Johns, 2010; Miraglia & Johns, 2016; Lohaus & Habermann, 2019). Both behaviors share a common feature: they generate costs not only for the sick employee but also for employers and society at large.

Beyond the pain caused by illness, the main cost of SA for the employee is the potential loss of a portion of her or his salary if sick leave pay is less than their regular earnings. At the company level, the cost of absence is evaluated by losses in terms of productivity, knowing that these are all the more important when the production technology of the company exhibits strong complementarity between labour and capital or between employees (Coles et al., 2007; Heywood et al., 2008; Grinza & Rycx, 2020). Finally, the significant increase in the cost of compensation for sick leave poses a challenge to the sustainability of health insurance systems; for example, daily sickness benefit payments in France increased by more than 16% between 2010 and 2017.

Although it is relatively straightforward to identify the factors that can impact a firm when an employee is absent, it is much more challenging to do the same for the costs that can be generated by SP. However, several studies have shown that coming to work when feeling ill causes a greater drop in productivity than that caused by SA (Schultz & Edington, 2007). Therefore, it would be appropriate for sick employees to take time off from work, since when they come to work they are at least 30% less productive, considering the work done under normal conditions (Hemp, 2004).

Evans-Lacko & Knapp (2016) note that in eight countries, the costs associated with SP were five to ten times higher than those associated with SA. In addition, loss of productivity may extend to the collective level given the risks of contamination in the workplace when employees choose to attend with a transmissible disease like a flu (Barmby & Larguem, 2009). Further, unlike SA, SP might have broader hidden costs for individuals and firms, such as long-term effects on physical or mental health leading to future SA (Skagen & Collins, 2016). Taken together, SA and SP may have high economic and social costs (Pauly et al., 2008).

Knowing the job determinants of SA and SP would make it possible to fine-tune firms' practices and policies that promote a reduction in absences and those that protect the organization from excessive attendance of unhealthy employees. Marie & Val Castello (2023) show that the reduction of paid sick leave lowered absences but, at the same time, seems to have extended relapses and accidents at work in the Spanish public sector. In contrast, as shown by Pichler & Ziebarth (2015) and Pichler et al. (2021), the introduction of paid sick leave in some American cities and states resulted in a significant diminution in influenza-like illnesses in the whole population probably caused by the decrease in sickness presenteeism.

The main objective of this chapter is therefore to empirically investigate which employment characteristics and working conditions are likely to worsen employees' state of health on the one hand, and to encourage employees to favour SP or SA on the other. From a personnel manager's perspective, it would thus be possible to address the former by means of preventive policies to protect employees' health and, on the other hand, to utilize the latter to design mechanisms that incentivize absence or presence based on their respective costs.

This chapter offers two original features. First, using the 2016 survey "Working conditions and psychosocial risks" , the determinants of both SA and SP annual duration are identified using two count data models correcting for the endogenous participation related to illness. Employment characteristics may influence the likelihood of ever being sick separately from the intensive margin decision to stay at home for treatment or work sick for any number of days. Second, to do that, we chose to follow a modern methodology proposed by Marra & Wyszynski (2016) based on a

"Conditions de travail et risques psychosociaux" conducted by the French Ministry of Labour

flexible copula approach. This enables to experiment various possible copulae representing the joint distribution of the probability of feeling ill and the number of days of SA or SP, and estimate the dependence parameter between these outcomes. This approach is also especially flexible, as at least four different marginal distributions of the count outcomes can be used to model employees' responses.

Despite the extensive research conducted on SA and SP , the two bodies of literature rarely overlap. This chapter therefore belongs to this limited literature that attempt to explain and estimate the duality of the employee's decision regarding whether to work when feeling ill.

Two main reasons explain the previous attempts to consider SA and SP simultaneously. Given the high costs of both phenomena for the employee, employer, and society, it seemed relevant to study the determinants of their prevalence. Moreover, the authors have questioned the validity of the substitutability hypothesis between these two behaviours. Strict substitution would occur when a change in a work-related variable implies a variation of opposite sign between SA and SP. If this relationship is imperfect, it suggests the existence of opportunities for firm policies to influence one or another of the two behaviours.

In fact, previous econometric studies seem to invalidate the strong substitution hypothesis. For example, in the "illness flexibility" model proposed by Johansson & Lundberg (2004), the choice made by employees who feel ill between SA and SP would be influenced by the latitude granted in their work pace, as well as by all factors that make absence costly in a broad sense. However, empirically, the ability to adjust one's effort positively affects the frequency of absence for female employees but has no effect on attendance frequency regardless of gender. In recent studies, Arnold & de Pinto (2015) found that a majority of the work-related determinants is either associated to the duration of SA or SP without affecting the other, while Gerich (2016) has shown among the determinants that significantly influence both the number of sickness absence and presenteeism days most of them display correlations of the same sign with these two phenomena. These results are in line with the bulk of studies conducted in Northern European countries, which have found a

For general overviews of the literature results regarding SA see Beemsterboer, Stewart, Groothoff, & Nijhuis (2009) and for those regarding SP see Miraglia & Johns (2016); Karanika-Murray & Cooper (2018); Lohaus & Habermann (2019).

positive correlation between the frequencies of SA and SP (see, for example, Hansen & Andersen (2008) and Leineweber et al. (2012)).

One plausible explanation for these results lies in the need to examine the factors driving the feeling of illness to explain the duration of SA and SP. First of all, individual and employment characteristics may affect the frequency of SA or SP, both directly, by encouraging employees to rest or come to work sick, and indirectly, by influencing the probability of illness. For instance, job insecurity is a stress-inducing factor that can lead to illness and prompt sick employees to show their commitment to their job by attending work. Hansen & Andersen (2008) attempt to determine which of these characteristics act as 'double risk factors' for presenteeism while controlling for employees' self-reported health levels. However, this method does not consider the fact that health levels are not exogenous and that self-reported health is not a measure of the day-to-day feeling of illness.

Furthermore, the perception of being sick is inherently subjective and likely to be influenced by the resources employees have in their work. Thus, Johansson & Lundberg (2004) consider the possibility that employees who have a high level of autonomy in determining their work pace or the intensity of their effort would be less inclined to self-diagnose as sick and therefore less likely to declare themselves as engaged in sickness presenteeism. Once again, in order to address this issue, it would be necessary to estimate an econometric model that explains the perception of being sick as a prerequisite to estimating the determinants of SA and SP.

This idea was developed theoretically by Arnold & de Pinto (2015), who proposed a model in which the effect of employment characteristics on the prevalence of SA and SP can follow two paths. Firstly, they influence the threshold health level that makes the employee indifferent between absence and attendance, and secondly, they also have an impact on the subjective feeling of being ill. In the latter case, this feeling is influenced by the sources of hardship at work worsened by the illness, but also by the incentives to take time off or to work while sick, instituted by the employer according to the anticipated productivity of a sick worker. For example, the company could implement incentives encouraging employees to take time off, even in case of a minor illness, if, considering the organization of work, their presence would be detrimental to their productivity or that of their colleagues. Consequently, the

model shows that it is then possible for SA and SP to be correlated with employment characteristics in the same direction. These results are confirmed by their empirical analysis of the determinants of SA and SP using seemingly unrelated regression models. However, like in most empirical models to our knowledge, the authors neglect to estimate beforehand whether the employee has experienced illness during the observed period leading to likely biases in the final estimation.

In this research, unlike previous studies, we explicitly assume that employment characteristics and working conditions may affect differently the extensive margin, i.e. the probability of illness, and the intensive margin of SA and SP, i.e. their duration. Our results underline the fact that without taking into account the endogeneity of illness, the magnitude of the effects of the studied characteristics on both SA and SP would be greatly biased. In addition, our results highlight several points to be taken into account when designing the firm's personnel policy. Firstly, the importance of implementing preventive occupational health policies because of the large number of factors that deteriorate health like the causes of intense work pace. Secondly, many of the determinants identified as affecting SA or SP in the literature only predict illness in our models. Finally, our results suggest that influencing workers' decision to engage in SA requires policies based on the employment contract, whereas influencing the decision to engage in SP requires management and prevention policies based primarily on working conditions and employee support. So, depending on the objectives pursued by a firm, the control of sickness presenteeism and absence would not be achieved using the same human resources management and work organization practices.

The next section presents our data and questions related to sickness, absenteeism and presenteeism, as well as the explanatory variables employed. Section 3 presents our economic intuition and the methodology of our econometric specification. Then, section 4 presents the results of the study, and finally, we discuss our findings and conclude in section 5.

2 Data and measurement

2.1 Data and sample of study

To analyse SP and SA together, we use the 2016 "Working conditions and psychosocial risks" survey, a national survey carried out by the French Ministry of Labour every three years. This is the eighth wave of the working conditions survey but the second in which the phenomenon of SP was measured. The survey initially covers a sample of 24,640 respondents representative of the French working population aged 15 years and above, regardless of their occupation, employment forms, business sectors, etc. These data contain information related to employees' sociodemographic traits and employment characteristics. Among the latter, working conditions at large are precisely described.

We consider employees aged from 18 to 65 years old who have either a temporary or permanent contract, working at least 10 hours per week but no more than 70 hours. We exclude self-employed individuals, students and apprentices, individuals with government-subsidized jobs and employees without working contracts. Moreover, to avoid measuring long sickness durations due to long-term illness or chronic diseases, we disregard observations with more than a total of 60 days of sickness a year. Finally, we have a dataset composed of 19,331 observations according to the applied restrictions.

2.2 Measure of sickness absence and presenteeism

Researchers traditionally study SP through self-reported assessments. Our SP measure is the annual duration of the phenomenon, similar to that used by Arnold & de Pinto (2015) and Arnold (2016), which was constructed from the European Working Conditions Survey. Respondents are asked to identify if they have ex-

The non-response to the Working conditions and psychosocial risks 2016 survey occurs at two levels: non-response from the household, which was the unit selected for the survey interview (19.5%), and non-response within the household, where among the selected individuals, some did not respond to the questionnaire (5.7%). This non-response was addressed in two steps, at the household level and then at the individual level, both times using the Chi-square Automatic Interaction Detector (CHAID) algorithm, a decision tree technique that classifies households and individuals into homogeneous response groups, dividing the population into homogeneous subpopulations in terms of non-response probability. This method generates weights to correct for total non-response and circumvent representativity bias.

perienced at least one episode of SP over the past twelve months: *Over the past twelve months, have you ever worked (gone to work) when you thought you shouldn't (should have stayed home) because you were sick?*. Then, if individuals had experienced at least one episode of SP, they were asked to evaluate the cumulative number of days over all episodes: *How many times has this happened in the past twelve months?*, with each episode counted in days. This measure has already been used in the literature (Gerich, 2016; Hirsch et al., 2017) in Austria and Germany respectively, allowing comparison of our findings with the previous results found for the European area.

The question related to SA asks respondents to report how many times they had taken sick leave over the past twelve months. Then, if they had taken at least one period of sick leave, the respondents were also asked to evaluate their cumulative number of SA days in the past year: *How many days of absence correspond to these sick leaves?* One limitation of these measures is the total lack of information on the type of illness, both in terms of the feeling of being ill and the official cause of sick leave.

While the number of SA days may be easier to remember than the number of SP days, as sick leaves are often justified by a physician prescription, both of these measures may be subject to memory bias. Furthermore, our measure of SP is founded on a retrospective assessment of a subjective feeling of illness. It is indeed possible that such memory biases are present since, as shown in Figure 1.1 below, peaks in each distribution appear approximately every five to ten days beginning around the fifth day. However, Hillion et al. (2021), who matched the database we use with administrative records , does not found significant difference for several characteristics such as SA days between the self-reported information reported in our data and the administrative ones.

In the overall population, as shown in table 1.1 below, 31.1% of our sample declared to experience at least one day of SA while 44.6% experienced SP at least once. On average, this represents 3.63 days of SA and 2.35 days of SP. For ill individuals, these figures amount to 6.33 days of SA and 4.09 days of SP, meaning that while feeling ill, employees nevertheless spend around 40% of their illness duration

From the Caisse Nationale d'Assurance Maladie, the French national fund for health insurance.

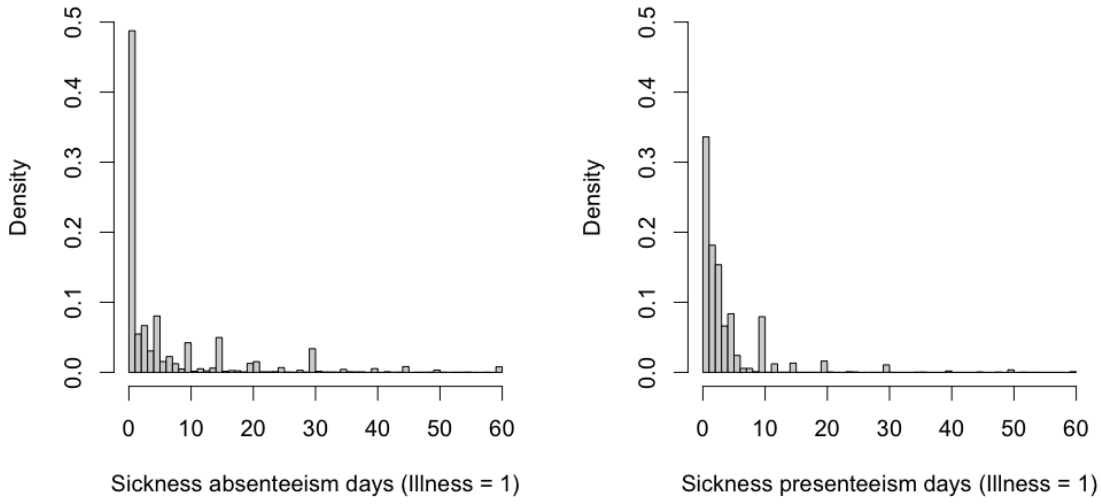

Figure 1.1 – Distribution of sickness absence and presenteeism days of ill individuals

working anyway. These observations are consistent with previous research (Lohaus & Habermann, 2019) and confirm the high occurrence of SP relative to SA in terms of both the amount of time employees spend at work while ill and the number of individuals involved in presenteeism.

Hence, regarding illness, more than half of the employees (57.4%) were ill at least one day over the past year. Among them, see table 1.2, less than a quarter (22.3%) only took sick leave, a large portion (45.7%) only engaged in SP, while 32% engaged in both behaviours. Therefore, when feeling ill, more than three-fourths (77.7%) of the employees go to work on at least one day.

Table 1.1 – Descriptive statistics of dependent variables

Determinants	Mean	St. Dev.	Min	Max
Sickness	0.574	0.495	0	1
Sickness presenteeism	0.446	0.497	0	1
Sickness presenteeism days	2.348	5.164	0	60
Sickness presenteeism days (conditional on illness)	4.093	6.272	0	60
Sickness absenteeism	0.311	0.463	0	1
Sickness absenteeism days	3.634	8.760	0	60
Sickness absenteeism days (conditional on illness)	6.333	10.801	0	60

Table 1.2 – Contingency tables comparing sickness presenteeism and sickness absence

% of total population n = 19,331		Sickness presenteeism 0	Sickness presenteeism 1
Sickness absence	0	42.62%	26.24%
	1	12.80%	18.34%

% of sick population n = 11,092		Sickness presenteeism 0	Sickness presenteeism 1
Sickness absence	0	0.00%	45.73%
	1	22.30%	31.97%

2.3 Explanatory variables

In accordance with our econometric strategy, we select individual and job characteristics that could explain SA and SP duration but that could also be deleterious for illness. We divide these characteristics into three groups of variables: i) individuals' sociodemographic characteristics and health states, ii) employment characteristics and iii) working conditions and individuals' feelings about their jobs. The description of all explanatory variables together with their descriptive statistics can be found in Tables A1 and A2 in the appendix section.

Concerning the individual's sociodemographic traits, we include age and dummies for gender, marital status, the presence of at least one dependent child and education level. In addition to these individual characteristics, we use the reported health status of the employee ranked on a five-point scale from very poor to very good. Furthermore, we control for individuals who declare they suffer from chronic disease, as they represent approximately 28% of our selected sample. We do not include SA days as an explanatory variable in the model explaining SP days, and vice versa, as this would result in a serious endogeneity bias and lead to the estimation of a simultaneous equations model, which is not our objective here.

Regarding the employment characteristics, we include information about the nature of the employment contract and job content: the wage level, working hours, years of seniority, occupational status, work in the public sector, having atypical work schedules and temporary or fixed-term contract, being a supervisor, working alone, and having a quantified performance target. Then, supplementary indicators are used to describe working conditions. Measures of work intensity include a constrained working pace due to the rhythm of a machine, colleagues' work, the extent of demand, norms or short delays of production, and computer or hierarchical control. We introduce further information about workload, such as whether an employee declares a lack of time to be effective and work under pressure. Furthermore, we

consider both employee's autonomy and the extent of the helping effort he receives. Hence, two dummies measure first if the employee can choose how to achieve the stated work objective and second if he must take initiative at work. Additionally, two binary variables evaluate whether he receives help from his supervisor or colleagues. Finally, a dummy variable indicated whether an employee has experienced a significant modification in his work environment, such as a change of function, work organization, redundancy plan, a merger or acquisition of his company, an outsourcing episode, etc.

Lastly, we include variables related to how individuals feel about their job. We control for any tension with the public, the firm's hierarchy or with subordinates or colleagues, the degree of conciliation between private and professional life as reported by the employee and perceived job security based on concern individuals may have regarding their job future. Industry levels and firm size are included to hopefully control for the differences among sectors and companies in terms of specific collective agreements related to sick pay provisions negotiated between unions and companies.

3 An empirical model of attendance decisions and econometric methodology

Our empirical model rests upon the assumption that an employee i decides to attend work to maximize his utility subject to standard budget constraint and the realization of daily random shock about his feeling of illness. Every day, he experiences a new subjective and specific feeling of illness, δ_i, which modifies the difficulty of his work effort and his work utility. Hence, the outcome of the balance between the benefits of working and the costs of effort daily varies in ways that are unpredictable for employers.

Following previous theoretical models by Brown & Sessions (2004) and latter by Arnold & de Pinto (2015), and assuming that the worse the employee's health, the larger his costs of effort, it is possible to demonstrate the existence of two thresholds for illness status in order to explain the choice between SA, SP and healthy working. So, the higher of these thresholds, δ_i^{sa}, makes the employee indifferent between working and being absent, and therefore depends on the costs associated with productive

effort when ill. Hence, the employee will rest at home (SA) if $\delta_i^{sa} \leqslant \delta_i$. Then a second threshold δ_i^{sp} determines whether the employee feels ill or healthy whenever he turns up to work. He will chose to experience sickness presenteeism (working while feeling ill) if $\delta_i^{sp} \leqslant \delta_i < \delta_i^{sa}$ and finally attend work while feeling healthy if $\delta_i < \delta_i^{sp}$. This last self-diagnosed reservation health threshold is certainly impacted by work organisation (decision latitude left to the employee in his or her work, the dependence of the company's productive performance or that of his colleagues on his presence) and the incentives for attendance, whether financial or linked to the threat of job loss. The central question, of course, is how these thresholds vary with wage earnings and the overall non-monetary benefits and costs of work. Figure 1.2 below summarises the behaviour of workers according to their illness thresholds.

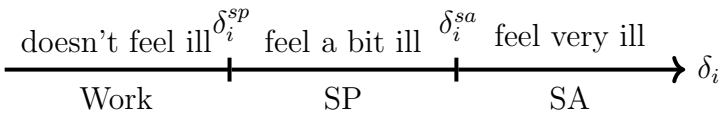

Figure 1.2 – Illness thresholds for which workers decide to be present sick (SP) or absent (SA)

The content of the job in the broad sense will therefore influence the reservation states of health (δ^{sp} and δ^{sa}) by varying the perceived disutility of effort. According to the job demand-resource (JD-R) model (Bakker & Demerouti, 2007), all the resources that enable workers to cope better with their job demands have to be considered. Recently, Böckerman et al. (2020) show the relevance of job design, defined as the combination of job demands and control, for workers' stress and job satisfaction. Hence, a worker's cumulative number of SP days is likely to be higher in a job with extended job demands and high workload. On the other hand, being able to receive help from one's work group (colleagues or supervisors) would reduce the difficulty of managing this workload after a sickness absence. Additionally, if production is organized just in time, employers may deter sickness absenteeism, as the work rhythm of employees would be affected by their colleagues' absence (Coles et al., 2007). Finally, certain features of a job make it more satisfying and motivating, such as interest, autonomy and mission, and can reduce the cost of effort and increase the worker's threshold of illness leading to SA (Jensen et al., 2019; Nikolova & Cnossen, 2020).

It should be noted that, among other things, our empirical model has two notable limitations. The first is the assumption that daily illness shocks are independent of each other. However, several studies have shown that the severity of a current illness can be influenced by the medical management of previous ones. For example, Bergström et al. (2009a) and Janssens et al. (2013) demonstrated that presenteeism at baseline can lead to an increase in absenteeism in the more or less near future. The second limitation concerns the dynamics of employees' decisions between absence and presenteeism in the event of illness, for example when the employer uses an attendance incentive scheme based on past episodes of absence. Barmby et al. (1991) examine the behavior of employees facing an experience-based sick pay scheme, which adjusts sick pay according to previous absence records: the higher an employee's past absences, the lower their current sick pay. Hence, the decision to be absent not only entails an immediate loss of earnings but also reduces the likelihood of receiving illness compensation in the future. In such cases, the decision whether to work or stay home while feeling ill is significantly more complex than depicted in our simplified model. This model aligns with the lack of information about the dynamics of absence and presenteeism in our data.

3.1 Count-data model with endogenous participation

An empirical model explaining the annual duration of SP and SA must recognise that these phenomena are by nature measured by count data, and be therefore estimated using count data models. A first issue concerns the endogeneity of variables explaining the phenomena. We will assume that the characteristics of the job are determined by the employer and therefore are exogeneously imposed to the employee. Of course, this is a strong assumption, but we miss appropriate instruments for the explanatory characteristics of the job while the choice of inappropriate instrument may severely bias the estimates (Kiviet, 2020). In addition, the question arises of the self-selection of employees in certain forms of jobs. Some of their unobservable characteristics, such as health or propensity for effort, may lead them to accept more or less demanding jobs. In order to mitigate estimation bias, Böckerman et al. (2012) proposed to capture these unobservable individual characteristics through employee's health history and labour market experience. We will show in our robust-

ness analysis (section 4.5 below) that the addition of such variables does not alter the association between either SA or SP duration and employment characteristics and working conditions.

Initially, we propose a model in which endogeneity results from endogenous participation in terms of illness. Our theoretical framework implies that the employee's feeling of illness is endogenously related to observable employment characteristics, working conditions and unobservables (such as managerial attendance incentive methods, particular characteristics of the employee such as the severity of his illness episode, his addictive behaviour and lifestyle habits) that could also affect the choice of SA or SP.

First and foremost, we must recognize that although the empirical problem of endogenous participation is inherently different from that of endogenous sample selection, both of these problems involve maximizing the same likelihood function when, as in the framework considered here, it is possible to observe non-negative values for SA or SP days when the individual feels ill (for details, see Bratti & Miranda (2011)). Empirical models have been proposed to correct for endogenous selection in count data models (see Greene (1997); Terza (1998); Bratti & Miranda (2011) for example). These methods share one major drawback: the use of fully parametric models based on the assumption of a Poisson distribution for the count outcome and the assumption that the link between the selection and outcome equations follows a bivariate normal distribution. Of course, if one of these assumptions is incorrect, the obtained estimates will not be consistent.

Consequently, this leads us to follow the econometric approach detailed in Marra & Wyszynski (2016), using an endogenous participation model for count data based on the choice of a copula to represent the association between the participation and outcome equations, these equations explaining the feeling of illness and SA or SP durations. This method appears to be the most appropriate for our observational data. First, it is flexible in that it allows a choice between several margins for the participation and outcome equations, and also between various copula functions to model the dependence between participation and outcome variables. Second, it allows to estimate the parameter of dependence between these two variables and so testing for endogenous participation.

3.2 A copula approach for SA and SP days with endogenous participation

Any non negative number of SP or SA days is observed if an employee i feels sick at least once a year. While we cannot observe his latent daily level of illness δ_i, we observe whether this employee has ever reported being ill. This variable, S_i, is the binary outcome of the latent daily level of illness such as:

$$S_i = \begin{cases} 1 & \text{if } \delta_i = \mathbf{X}_i'\gamma + \mathbf{Z}_i'\phi + \mathbf{R}_i'\upsilon + \varepsilon_i \geqslant \delta_i^{sp} \ , \quad i = 1, ..., N \\ 0 & \text{otherwise} \end{cases} \tag{1.1}$$

where the latent daily level of illness is explained by $\mathbf{X}$, the vector of sociodemographic characteristics of employees, $\mathbf{Z}$, the vector of employment characteristics and $\mathbf{R}$ the exclusion restriction used to secure identification of the model. γ, ϕ and υ are the coefficient vectors of all parametric components, while ε_i is an error term containing every individual specific and unobservable characteristics.

Marra & Wyszynski (2016) conducted simulation studies to evaluate how their copula-based endogenous participation model for count data performed in various scenarios. One set of these simulations demonstrated that even if the simulated data were generated with an exclusion restriction in the participation equation, the estimates of the outcome equation remained similar regardless of the presence or absence of this exclusion restriction. In such cases, the identification of the model likely relies on adequate assumptions about the marginal distributions of the participation and outcome variables. However, in three health economics-related endogenous switching models employing copulae to represent the association between a selection and two outcome equations, Marra et al. (2024) found that, in absence of exclusion restrictions, the estimates of the outcome equations were less precise. Therefore, in our analysis, we introduce episodes from the employee's history before the age of 18 as potential exclusion restrictions. Specifically, three indicators were alternatively used as exclusion restrictions: whether the employee suffered from family conflicts, violence at school or in the neighborhood, or physical and psychological mistreatment. Indeed, numerous studies have consistently identified long-term

correlations between the quality of family relationships during childhood and adolescence and poor emotional and physical health in adulthood (see, e.g., Clark et al. (2010), Landstedt et al. (2015) and Berg et al. (2017)). Similarly, literature reviews on the link between negative life outcomes in adulthood and exposure to mistreatment during youth or violence at school confirm the deleterious long-term impacts of such episodes on health (Norman et al. (2012) and Turanovic (2022)). Based on this evidence, we assume that these variables act as determinants of present health and are linked to current SP and SA annual duration solely through their effect on health. To the best of our knowledge, we have not found any study showing a direct link between work engagement and motivation that are other determinants of SA and SP and past traumatic episodes during youth. All of the exclusion restrictions were found to be significant in the participation equation. However, we present the main results using only family conflicts before 18 as the exclusion restriction since the Wald test returns the higher Chi-square statistics ($\chi^2 = 121.57$) compared to the others . This choice was also motivated by the absence of tests to assess the weakness of the instruments in this type of model (Marra et al., 2024).

Regarding the count outcomes, we assume that $Y_i \sim \mathcal{D}(\mu_i, \sigma, \nu)$, where Y_i is either SA or SP days and follows a discrete distribution $\mathcal{D}$, with:

$$\mu_i = \mathbb{E}(Y_i) = exp\left(\mathbf{X}_i'\alpha + \mathbf{Z}_i'\beta + \mathbf{H}_i'\psi\right) \tag{1.2}$$

including, in addition to the individual's and employment characteristics, $\mathbf{H}$ the employee's subjective assessment of health. Here, $\mathcal{D}$ can be a Poisson, a negative binomial type I, a negative binomial type II or a Poisson inverse Gaussian distribution. σ and ν are scale and shape parameters according to the chosen distribution, while α, β and ψ are the vectors of parameters associated with each vector of characteristics .

Our approach assumes that the numbers of SA or SP days are correlated with the process that generates the feeling of illness in the participation equation. This

Wald tests for violence at school or in the neighborhood and physical or psychological mistreatment as exclusion restrictions provide a χ^2 equal to 24.04, and 41.92 respectively.

For details about the probability mass function and the parameters of each distribution, see Table 2 of Marra and Wyszynski's paper for example.

association between participation and outcome equations can be described using a copula function. Given that $F_s(s_i)$ and $F_y(y_i)$ denote the cumulated distribution functions (cdf) of the feeling of illness S_i and count outcome Y_i , we can find a two-place copula function C such that their joint cdf can be represented by $F(s_i, y_i) = C\left(F_s(s_i), F_y(y_i); \theta\right)$ where θ corresponds to the dependence parameter between the equations. For an easier interpretation, this parameter can be converted into Kendall's τ measure of association that ranges into $[-1, 1]$. One advantage of using copulae is that they allow for modeling the dependence structure between the observed values of the participation variable S_i and count outcome Y_i independently of their marginal distributions. Consequently, the selection of the copula is performed while accounting for potential measurement errors in these variables. In our case, any potential memory bias regarding past episodes of SA and SP has therefore no impact on the effectiveness of the correction for endogenous participation.

The likelihood of the endogenous participation model can be formulated as follows:

$$\begin{aligned} L &= \prod Pr(S_i = 0) \prod Pr(Y_i = y_i, S_i = 1) \\ &= \prod Pr(\delta_i < \delta_i^{sp}) \prod f_{y|s}(y_i|\delta_i \geqslant \delta_i^{sp}) Pr(\delta_i \geqslant \delta_i^{sp}) \end{aligned} \tag{1.3}$$

Our FIML estimation method fully utilizes all available information to obtain correct estimates of coefficients and standard errors of the explanatory variables. The objective is now to find the right model for both SA and SP behaviours. This means that we must find the appropriate margins for the participation and count equations as well as the most appropriate copula linking the distributions. For the participation equation, we consider three possible distributions: logit, probit and cloglog. We have four margins to try for the count distribution: Poisson, negative binomial types I and II and Poisson inverse Gaussian. Furthermore, there are eight copulae to test: Normal, Frank, Ali-Mikhail-Haq (AMH), Farlie-Gumbel-Morgenstern (FGM), Student-t, Clayton, Joe and Gumbel. Additionally, rotated versions of Clayton, Joe and Gumbel exist. These three copulae have a strictly positive dependence parameter θ. Thus, either a 90° or a 270° rotation allows the use of these copulae when the dependence parameter is assumed to be negative. To determine which rotation of the copula we must use, we begin to apply Gaussian,

FGM, AMH, Student-t and Frank copulae to determine if there is a positive or negative association between selection and outcome equations. Therefore, we have 11 copulae to test, giving a total of 132 models to explore.

4 Results

4.1 Model choice

Following the methodology of Marra & Wyszynski (2016), we use the Akaike and Bayesian Information Criteria (AIC and BIC) to preserve the quantity of information when selecting the distributions of the duration of SP, SA and that of the feeling of illness as well as the copula linking these distributions. With a fixed margin (resp. copula) and the number of parameters to be estimated remaining unchanged, the AIC and BIC are valid selection criteria since they lead to the choice of the model with the highest log-likelihood. Both criteria always conduct to the same model choice. Also, we performed the Vuong test to corroborate if our chosen model is closer to the true data generating process (Hasebe & Vijverberg, 2012). In addition, we propose a graphical exploration of the choice of distributions for the number of days of SP and SA using the Q-Q plots. Although the goodness-of-fit of empirical continuous distributions can be assessed using statistical tests, they cannot be used to check the fit of empirical discrete distributions.

Our results indicate that the logit margin for the participation equation always provides the lowest AIC and BIC values for the whole model compared to the probit and cloglog margins . Hence, table 1.3 below presents the AIC for each discrete margin and copula for the SA and SP models. For both models, there is a single distribution associated with the lowest AIC values, regardless of the copula selected. This leads to the selection of a negative binomial type II margin for the SA duration and a Poisson inverse Gaussian margin for the SP duration. The Vuong test also leads to this choice comparing the different margins for the same copula .

In the figure 1.3 below, the post-estimation graphical analysis of the randomized and normalized residuals Q-Q plots of the chosen models highlights a very good fit

AIC and BIC value tables are available in section A of the appendices.

The results of Vuong test are not shown as the R command directly provides the preferred model between two models.

Table 1.3 – AIC values for the discrete marginal distributions and copulae

	Sickness absence (SA)					Sickness presenteeism (SP)			
Copula	Poisson	PIG	NB I	NB II	Copula	Poisson	PIG	NB I	NB II
Normal	172,142	83,413	80,876	80,440	Normal	103,116	78,028	78,674	78,753
Frank	168,526	83,402	80,865	80,424	Frank	102,417	77,996	78,209	78,255
AMH	168,014	83,414	80,872	80,437	AMH	102,817	78,035	78,531	78,572
FGM	167,340	83,398	80,864	80,418	FGM	102,346	78,041	78,523	78,571
Student-t	167,450	83,503	80,878	80,461	Student-t	103,094	**77,958**	78,361	78,458
Clayton	168,810	83,429	80,884	80,453	Clayton 90	101,176	78,013	78,174	78,234
Clayton 180	168,201	83,149	80,825	80,359	Clayton 270	103,319	78,085	78,774	78,808
Joe	168,066	83,095	80,816	**80,348**	Joe 90	103,320	78,067	78,775	78,810
Joe 180	168,681	83,429	80,884	80,453	Joe 270	101,078	78,026	78,156	78,216
Gumbel	168,562	83,297	80,850	80,406	Gumbel 90	103,320	78,019	78,739	78,808
Gumbel 180	168,674	83,429	80,884	80,453	Gumbel 270	102,166	77,991	78,372	78,457

Note: logit in first step

of the empirical distributions of SA and SP with the theoretical quantiles. In other words, these distributions describe the most accurately the generating process of observed values of SA and SP days. For comparison, the figure A1 in the appendix section exhibits how different are the quantiles of the observed and theoretical distribution of the residuals when we assumed the latter to be a Poisson distribution with a Gaussian copula, which clearly does not fit our data .

The choice of these two marginal distributions is consistent with the shape of the cumulative days of SA and SP shown in Figure 1.1. The NBII and PIG distributions have in common that they allow adjustment in the case of strong overdispersion in the SA and SP distributions, but in addition, the PIG model can particularly handle count data that present sharp peaks at 0 and small values such as 1 or 2 days, which is the case here of SP (Hilbe, 2014).

Regarding the copula that link the distribution of illness and the SA or the SP distribution, we also based our choice on the AIC values in table 1.3. Firstly, after applying Gaussian, FGM, AMH, Student-t and Frank copula to our models, we learned that unobservable characteristics in the SA equation are negatively correlated with those of the illness equation, while by contrast, the unobservables of the SP equation are positively related to those that explain the illness feeling. This led us to add the unrotated and 180° rotated Clayton, Joe and Gumbel copula for the SA equation and the 90° and 270° rotated Clayton, Joe and Gumbel copula for the SP equation.

Before modelling the two durations of SA and SP, the graphical analysis of the Q-Q plot of the unconditional distributions has already provided some insight into the inadequacy of the Poisson distribution.

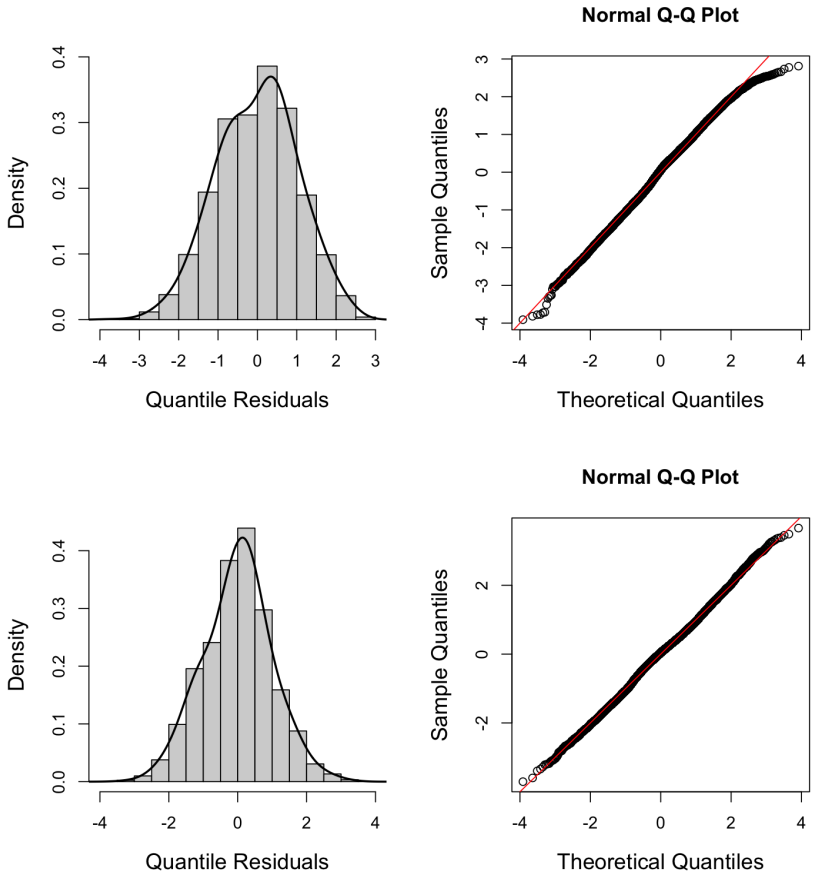

Figure 1.3 – Q-Q plots of the randomized and normalized residuals of the count regressions with endogenous participation. The first row corresponds to the regression of sickness absenteeism days with a negative binomial type II margin, and the second row corresponds to the regression of sickness presenteeism days with a Poisson inverse Gaussian margin.

As in the case of the choice of margins, there exist several goodness-of-fit tests for choosing the right copula when the margins are continuous since the copula is unique. However, this is not true when the two margins are discrete (Cameron et al., 2004) and these tests may not lead to reliable conclusions. To address this issue, we therefore present in section 4.5 some robustness checks about the choice of the copula. However, Marra et al. (2024) find that the choice of other copula does not impact the estimated effects but rather the efficiency by increasing the confidence intervals. In addition, we calculate the Vuong test by comparing each model with different copulae for a given margin. All of these tests lead to the choice of the same copula as the AIC. As a result, the best models according to the AIC are the negative binomial of type II with a Joe copula for the SA model and the Poisson inverse Gaussian with a Student-t copula for the SP model.

The estimates of the two models are reported in Table 1.4 below. Columns (1) and (4) report the estimated coefficients of the count, and columns (2) and (5) report the average marginal effects with their standard deviation calculated by bootstrap on 600 replications. Finally, columns (3) and (6) report the average marginal effects of the participation equation corresponding to the probability of illness.

Table 1.4 – Determinants of sickness presenteeism and absence days with endogenous participation

	Sickness absence			Sickness presenteeism		
Equation	SA days		Illness	SP days		Illness
Distribution	(NB II)		(logit)	(PIG)		(logit)
	Coef.	AME[a]	AME[a]	Coef.	AME[a]	AME[a]
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	**1.407*****	−−	−−	**1.167*****	−−	−−
	(0.367)			(0.282)		
Health (ref: good)						
bad or very bad	**0.587*****	**3.347*****	−−	**0.445*****	**3.300*****	−−
	(0.057)	(0.464)		(0.049)	(0.497)	
rather good	**0.179*****	**0.814*****	−−	**0.286*****	**1.838*****	−−
	(0.034)	(0.175)		(0.027)	(0.201)	
very good	**−0.083****	**−0.353****	−−	**−0.150*****	**−0.862*****	−−
	(0.042)	(0.169)		(0.033)	(0.181)	
Chronic disease	**0.199*****	**0.895*****	−−	**0.063****	**0.385****	−−
	(0.032)	(0.158)		(0.025)	(0.167)	
Occupation (ref: employee)						
labourer	0.055	0.245	0.013	−0.025	−0.153	0.016
	(0.053)	(0.256)	(0.012)	(0.043)	(0.270)	(0.012)
intermediate	0.069	0.309	−0.001	−0.005	−0.032	−0.001
	(0.044)	(0.208)	(0.010)	(0.034)	(0.221)	(0.010)
executive	−0.029	−0.128	−0.018	0.002	0.012	−0.014
	(0.064)	(0.287)	(0.014)	(0.048)	(0.310)	(0.014)
Seniority	0.004	−0.003	**0.001*****	−0.006	−0.008	**0.002*****
	(0.006)	(0.011)	(0.0005)	(0.005)	(0.011)	(0.0005)
Seniority^2	−0.000			0.000		
	(0.000)			(0.000)		
Work alone	**−0.067***	**−0.291***	**0.020****	0.020	0.120	**0.022*****
	(0.035)	(0.170)	(0.008)	(0.027)	(0.178)	(0.008)
Supervisor	**−0.194*****	**−0.789*****	**−0.050*****	−0.016	−0.094	**−0.045*****
	(0.051)	(0.203)	(0.011)	(0.037)	(0.222)	(0.011)
Public sector	**0.156*****	**0.687*****	**0.029*****	−0.002	−0.012	**0.024****
	(0.043)	(0.203)	(0.010)	(0.034)	(0.203)	(0.011)
Temporary contract	**−0.466*****	**−1.671*****	**−0.114*****	0.036	0.220	**−0.112*****
	(0.079)	(0.222)	(0.015)	(0.058)	(0.351)	(0.015)
Working time	**0.067*****	**−0.036*****	0.001	**−0.014***	0.019	0.001
	(0.011)	(0.011)	(0.001)	(0.007)	(0.011)	(0.001)
Working time^2	**−0.001*****			**0.000****		
	(0.000)			(0.000)		
Atypical schedules	**−0.101*****	**−0.447*****	**−0.016****	**0.056****	**0.341****	**−0.014****
	(0.031)	(0.140)	(0.007)	(0.025)	(0.147)	(0.007)
Wage (ref: 1,800-2,499)						
100-1,199	0.085	0.386	−0.005	−0.002	−0.011	−0.005
	(0.067)	(0.321)	(0.015)	(0.052)	(0.326)	(0.015)
1,200-1,499	0.067	0.299	−0.017	−0.012	−0.070	−0.011
	(0.047)	(0.227)	(0.011)	(0.038)	(0.224)	(0.011)
1,500-1,799	0.066	0.294	−0.016	0.051	0.312	−0.014
	(0.043)	(0.215)	(0.010)	(0.034)	(0.218)	(0.010)
2,500-2,999	0.028	0.124	−0.019	0.076	0.478	−0.019
	(0.062)	(0.264)	(0.013)	(0.047)	(0.316)	(0.013)
3,000-20,000	0.011	0.050	**−0.034****	−0.001	−0.006	**−0.034****
	(0.068)	(0.310)	(0.014)	(0.050)	(0.308)	(0.014)
Quantified target	**0.063***	0.279	**0.026*****	**0.086*****	**0.530*****	**0.023*****

	SA days		Illness	SP days		Illness
	Coef.	AME	AME	Coef.	AME	AME
	(0.035)	(0.165)	(0.009)	(0.028)	(0.180)	(0.008)
Work rhythm paced by:						
machine	0.067	0.302	**0.042*****	0.014	0.083	**0.039*****
	(0.047)	(0.238)	(0.012)	(0.038)	(0.235)	(0.012)
external demand	**−0.136*****	**−0.616*****	−0.004	**0.056****	**0.337****	−0.002
	(0.035)	(0.163)	(0.008)	(0.028)	(0.178)	(0.008)
norms	0.016	0.070	**0.014***	0.007	0.045	**0.017****
	(0.033)	(0.146)	(0.008)	(0.026)	(0.159)	(0.008)
computing control	0.010	0.043	0.007	0.003	0.019	0.008
	(0.033)	(0.153)	(0.008)	(0.026)	(0.151)	(0.008)
monitoring	0.047	0.206	**0.033*****	0.041	0.250	**0.031*****
	(0.034)	(0.158)	(0.009)	(0.027)	(0.166)	(0.009)
colleagues relation	−0.009	−0.042	**0.023*****	0.015	0.091	**0.022*****
	(0.034)	(0.165)	(0.008)	(0.026)	(0.161)	(0.008)
Help from supervisors	−0.040	−0.175	**−0.059*****	**−0.075*****	**−0.459*****	**−0.057*****
	(0.031)	(0.147)	(0.007)	(0.025)	(0.149)	(0.007)
Help from colleagues	**0.089***	**0.380***	**0.020***	**−0.126*****	**−0.800*****	**0.021****
	(0.046)	(0.201)	(0.010)	(0.034)	(0.222)	(0.010)
Tensions at work	**0.132*****	**0.567*****	**0.106*****	**0.111*****	**0.660*****	**0.106*****
	(0.034)	(0.149)	(0.007)	(0.028)	(0.156)	(0.007)
Lack of time	0.006	0.027	**0.059*****	**0.099*****	**0.611*****	**0.059*****
	(0.034)	(0.155)	(0.008)	(0.026)	(0.161)	(0.008)
Under pressure	0.013	0.059	**0.071*****	**0.144*****	**0.885*****	**0.068*****
	(0.034)	(0.161)	(0.009)	(0.027)	(0.179)	(0.009)
Choice of methods	**−0.095*****	**−0.430****	−0.013	0.023	0.137	−0.010
	(0.037)	(0.175)	(0.009)	(0.030)	(0.186)	(0.009)
Initiative	−0.034	−0.149	0.010	**0.094*****	**0.573*****	0.009
	(0.031)	(0.144)	(0.008)	(0.024)	(0.152)	(0.008)
Work-to-life conflict	−0.031	−0.135	**0.069*****	0.036	0.223	**0.069*****
	(0.039)	(0.189)	(0.009)	(0.030)	(0.187)	(0.009)
Fear of job future	**−0.062***	−0.269	**0.048*****	**0.092*****	**0.576*****	**0.047*****
	(0.038)	(0.174)	(0.010)	(0.029)	(0.176)	(0.010)
Organizational changes	**0.131*****	**0.571*****	**0.060*****	0.022	0.132	**0.058*****
	(0.031)	(0.137)	(0.007)	(0.024)	(0.151)	(0.007)
Family conflicts			**0.079*****			**0.088*****
			(0.007)			(0.007)
Controls[b]	Yes		Yes	Yes		Yes
Num. obs.	11,073		19,285	11,073		19,285
τ	0.312			−0.341		
	(0.271, 0.343)			(−0.382, −0.289)		
Log Likelihood	−40032.815			−38838.129		

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

[a] Average marginal effects (AME) are calculated. AME for seniority (resp. working time) is calculated at the mean: 15 years (resp. 36.8 hours). Standard deviations are reported in parentheses. The standard deviations of the AME are calculated on 600 bootstrap replications.

[b] We control for individual's characteristics (age, gender, marital status, having children and education level), firm sectors and firm size.

4.2 Participation bias

The first important result is related to the impact of considering the participation bias associated to illness on the estimation of the effects of job features on the duration of SA and SP. The estimates reported in Table 1.4 reveal a significant Kendall

τ, but of opposite sign, measuring the association between unobserved determinants of illness and unobserved determinants of both the duration of SP and SA. Thus, unobservable variables that increase the probability of feeling ill are negatively correlated ($\tau = -0.34$) with unobservable variables that increase the number of days of SP but are positively correlated ($\tau = 0.31$) with unobserved determinants of days of SA. Hence, one possible interpretation is that these unobservables are indicators of the seriousness of illness, information that is missing in our database. Therefore, these results would be compatible with our attendance model, as the more severe the illness experienced, the longer the period of SA and the shorter the period of SP. These significant correlations between the participation and count equations underscore the importance of addressing endogenous participation bias in our econometric model. Given the substantial costs of SA and SP for social insurance systems and firms, It is therefore essential to accurately identify which job characteristics and working conditions affect the duration of SA and SP, as well as to assess the magnitude of their respective impacts.

The Table 1.5 below reports the average marginal effects of the univariate count models (which do not account for endogenous participation) and the bivariate count models (which do account for endogenous participation), along with their differences in absolute value and percentage. A comparison of the estimates from the univariate and bivariate models, for SA and SP, confirms that failing to include the participation equation leads to significantly different conclusions regarding the determinants of SA and SP durations. These results show that the discrepancies between the univariate and bivariate models are greater for SA than for SP.

Table 1.5 – Average marginal effects comparison between univariate and bivariate models

	AME SA days				AME SP days			
	Univariate	Bivariate	Diff.[a]	Diff.[a] (%)	Univariate	Bivariate	Diff.[a]	Diff.[a] (%)
				Health characteristics				
Health (ref: good)								
bad or very bad	**4.254*****	**3.347*****	0.907	27.10	**2.461*****	**3.300*****	−0.839	−25.42
	(0.588)	(0.464)			(0.368)	(0.497)		
rather good	**1.002*****	**0.814*****	0.188	23.10	**1.354*****	**1.838*****	−0.484	−26.33
	(0.226)	(0.175)			(0.137)	(0.201)		
very good	**−0.408***	**−0.353****	−0.055	15.58	**−0.630*****	**−0.862*****	0.232	−26.91
	(0.223)	(0.169)			(0.131)	(0.181)		
Chronic disease	**1.125*****	**0.895*****	0.23	25.70	**0.280*****	**0.385****	−0.105	−27.27
	(0.206)	(0.158)			(0.120)	(0.167)		

	AME SA days				AME SP days			
	Univariate	Bivariate	Diff.[a]	Diff.[a] (%)	Univariate	Bivariate	Diff.[a]	Diff.[a] (%)
Employment characteristics								
Occupation (ref: employee)								
labourer	0.179	0.245	−0.066	−26.94	−0.039	−0.153	0.114	−74.51
	(0.315)	(0.256)			(0.199)	(0.270)		
intermediate	0.439	0.309	0.13	42.07	−0.005	−0.032	0.027	−84.37
	(0.255)	(0.208)			(0.156)	(0.221)		
executive	0.018	−0.128	0.146	−114.06	−0.048	0.012	−0.06	−500.00
	(0.355)	(0.287)			(0.221)	(0.310)		
Seniority	−0.014	−0.003	−0.011	366.67	−0.002	−0.008	0.006	−75.00
	(0.014)	(0.011)			(0.008)	(0.011)		
Work alone	**−0.587*****	**−0.291***	−0.296	101.72	0.138	0.120	0.018	15.00
	(0.213)	(0.170)			(0.125)	(0.178)		
Supervisor	**−0.737*****	**−0.789*****	0.052	−6.59	−0.242	−0.094	−0.148	157.45
	(0.274)	(0.203)			(0.150)	(0.222)		
Public sector	**0.692*****	**0.687*****	0.005	0.73	0.086	−0.012	0.098	−816.67
	(0.256)	(0.203)			(0.143)	(0.203)		
Temporary contract	**−1.541*****	**−1.671*****	0.13	−7.78	−0.196	0.220	−0.416	−189.09
	(0.325)	(0.222)			(0.229)	(0.351)		
Working time	**−0.054*****	**−0.036*****	−0.018	50.00	0.013	0.019	−0.006	−31.58
	(0.014)	(0.011)			(0.008)	(0.011)		
Atypical schedules	**−0.490*****	**−0.447*****	−0.043	9.62	0.144	**0.341****	−0.197	−57.77
	(0.176)	(0.140)			(0.101)	(0.147)		
Wage (ref: 1.800−2.499)								
100−1.199	0.509	0.386	0.123	31.87	−0.095	−0.011	−0.084	763.64
	(0.402)	(0.321)			(0.226)	(0.326)		
1.200−1.499	**0.533***	0.299	0.234	78.26	−0.130	−0.070	−0.06	85.71
	(0.288)	(0.227)			(0.152)	(0.224)		
1.500−1.799	**0.526***	0.294	0.232	78.91	0.159	0.312	−0.153	49.04
	(0.272)	(0.215)			(0.151)	(0.218)		
2.500−2.999	0.315	0.124	0.191	154.03	0.249	0.478	−0.229	47.91
	(0.337)	(0.264)			(0.219)	(0.316)		
3.000−20.000	0.454	0.050	0.404	808.00	−0.183	−0.006	−0.177	2,950.00
	(0.397)	(0.310)			(0.211)	(0.308)		
Quantified target	0.165	0.279	−0.114	−40.86	**0.504*****	**0.530*****	−0.026	−4.91
	(0.205)	(0.165)			(0.128)	(0.180)		
Working conditions								
Work rhythm paced by:								
machine	0.119	0.301	−0.182	−60.46	0.198	0.083	0.115	138.55
	(0.289)	(0.238)			(0.167)	(0.235)		
external demand	**−0.783*****	**−0.616*****	−0.167	27.11	0.174	**0.337****	−0.163	48.37
	(0.201)	(0.163)			(0.124)	(0.178)		
norms	−0.025	0.070	−0.095	−135.71	0.096	0.045	0.051	113.34
	(0.183)	(0.146)			(0.114)	(0.159)		
computing control	0.011	0.043	−0.032	−74.42	0.053	0.019	0.034	178.95
	(0.193)	(0.153)			(0.110)	(0.151)		
monitoring	0.028	0.206	−0.178	−86.40	**0.304*****	0.250	0.054	21.6
	(0.199)	(0.158)			(0.119)	(0.166)		
colleagues relation	−0.227	−0.042	−0.185	440.48	0.186	0.091	0.095	104.40
	(0.204)	(0.165)			(0.112)	(0.161)		
Help from supervisors	0.153	−0.175	0.328	−187.43	**−0.566*****	**−0.459*****	−0.107	23.31
	(0.181)	(0.147)			(0.107)	(0.149)		
Help from colleagues	0.399	**0.380***	0.019	5.00	**−0.533*****	**−0.800*****	0.267	−33.37
	(0.264)	(0.201)			(0.156)	(0.222)		
Tensions at work	−0.154	**0.567*****	−0.721	−127.16	**0.840*****	**0.660*****	0.18	27.27
	(0.189)	(0.149)			(0.104)	(0.156)		
Lack of time	**−0.352***	0.027	−0.379	−1,203.70	**0.624*****	**0.611*****	0.013	2.13
	(0.195)	(0.155)			(0.115)	(0.161)		
Under pressure	**−0.420****	0.059	−0.479	−611.86	**0.931*****	**0.885*****	0.046	5.20
	(0.194)	(0.161)			(0.122)	(0.179)		
Choice of methods	**−0.484****	**−0.430*****	−0.054	12.56	0.020	0.137	−0.117	−85.40

	AME SA days				AME SP days			
	Univariate	Bivariate	Diff.[a]	Diff.[a] (%)	Univariate	Bivariate	Diff.[a]	Diff.[a] (%)
	(0.222)	(0.175)			(0.131)	(0.186)		
Initiative	−0.253	−0.149	−0.104	69.80	**0.477*****	**0.573*****	−0.096	−16.75
	(0.182)	(0.144)			(0.108)	(0.152)		
Work−to−life conflict	**−0.559****	−0.135	−0.424	314.07	**0.412*****	0.223	0.189	84.75
	(0.230)	(0.189)			(0.133)	(0.187)		
Fear of job future	**−0.697*****	−0.269	−0.428	159.11	**0.553*****	**0.576*****	−0.0230	−3.99
	(0.218)	(0.174)			(0.125)	(0.176)		
Organizational changes	0.269	**0.571*****	−0.302	−52.89	**0.309*****	0.132	0.177	134.09
	(0.176)	(0.137)			(0.102)	(0.151)		

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$, standard deviation in parentheses calculated on 600 bootstrap replications

[a] difference is calculated as $\text{AME}_{uni} - \text{AME}_{biv}$ and the percentage difference is calculated as $\text{Diff.}/\text{AME}_{biv}$

Compared to the bivariate models that account for endogenous participation, the univariate models exhibit an average error rate - or average difference in estimates - of 140% for SA and 33.5% for SP. Among the 21 (respectively 17) variables identified as having a significant effect in at least one model for SA (resp. SP), 15 (resp. 8) coefficients were overestimated with an average difference of 62% (resp. 43%) relative to the bivariate model. Conversely, 6 (resp. 9) of the coefficients were underestimated, representing an average difference of 335% (resp. 24.8%) relative to the SA (resp. SP) bivariate model. Even more problematically, many determinants change in significance depending on the model. This is the case for 9 variables (resp. 5) for SA (resp. SP). For this subset of explanatory variables, the average difference between the estimates from the univariate and bivariate models is significantly greater than for the explanatory variables that remain significant in both models. Specifically, this average difference is equal to 292% compared to 25.6% for SA modeling and 69.3% compared to 18.6% for SP modeling.

In more detail, all the effects associated with health variables appear to be overestimated in the univariate model for the duration of absence (23%), while they are underestimated for the duration of SP (−26.5%). In other words, this misinterpretation of the impacts of these variables would lead to an exaggeration of the anticipated benefits of occupational health prevention policies aimed at reducing the duration of absence.

Regarding employment characteristics, the estimation errors between univariate and bivariate models result in an overvaluation of SA duration by 0.3 days, primarily explained by the overestimation of the negative impact of working in isolation.

Furthermore, failing to correct for endogenous illness would result in misidentifying the significant increase in SP duration associated with working atypical hours.

In addition, the most significant differences concern the employee's working conditions and work environment. For instance, the pressure exerted by external demand on the pace of work would have been associated with a reduction in the duration of SA of almost 0.8 days in the univariate model, without affecting the duration of SP. However, the bivariate models indicate a smaller reduction, of 0.6 days, in the duration of SA, but above all accompanied by a significant increase in the duration of SP of 0.34 days. A similar pattern emerges for characteristics such as working under pressure and a lack of time to complete tasks. In the univariate models, these factors significantly reduced the duration of SP (by 0.6 and 0.9 days) while increasing SA duration (by 0.35 and 0.4 days). In contrast, the bivariate models did not identify any significant effects of these factors on the duration of absence. Another example concerns employees receiving support from colleagues. According to the univariate models, this would lead to a reduction in SP of 0.53 days with no impact on absence. However, the bivariate models reveal an increase in absence (significant at 10% level) of 0.38 days and a greater reduction in SP of 0.8 days. While an imbalance between time devoted to work and personal life appears to significantly affect the duration of SA and SP in the univariate models, the bivariate models suggest otherwise. Finally, contrary to the findings of the univariate models, organizational changes seem to increase SA rather than SP, and fear for the future of one's job is not associated with reduced absenteeism in the bivariate model.

Hopefully, these findings may guide companies to assess the effectiveness of various personnel policies. Indeed, depending on the relative costs of a day of sickness absence or presenteeism, firms need to identify the job features that can serve as effective instruments to induce the workforce to adopt appropriate behavior. Misidentifying the key characteristics influencing the decision to be absent or present when sick could lead to inefficient incentive schemes.

4.3 Determinants of the feeling of illness

The estimates of the participation equations related to illness, reported in Columns 3 and 6 in table 1.4, illustrate the extent to which employment characteristics and

working conditions contribute to the feeling of illness. Our results reveal that some variables affect only the likelihood of feeling ill, without affecting the choice between SA and SP, while others act as dual risk factors impacting both the likelihood of illness and the decision between absence and attendance in the event of illness. The former variables highlight opportunities for firms to mitigate SA and SP by directly protecting the workforce against the risk of illness.

Four sets of variables are found to be only associated with the likelihood to feeling ill. Firstly, the more senior the employees, the more likely they are to report illness (+0.2 percentage points, or pp). This relationship can be explained by the rules of the social insurance system in France, which stipulate that the level of sick pay, as well as the duration of the right to this coverage, increases with workers' seniority. The generosity of the sick pay mechanism can also be enhanced by collective agreements at the firm or industry level (see Lanfranchi & Treble (2010) for a presentation). Secondly, employees earning more than 3,000 euros per month are found to be less likely to consider themselves ill (−3.4 pp). Reasons for such a result may be found in the better opportunities that high wage earners possess to access quality healthcare and improve their material living conditions (Barnay, 2016). Thirdly, when the pace of work is constrained by machines, norms, hierarchical control, and relationships with colleagues, the likelihood of feeling ill is found to increase (respectively by +3.9, 1.7, 3.1, and 2.2 pp). This evidence is consistent with the hypothesis that work intensity is detrimental to health and that firms could control SA and SP through the development of preventive policies. For example, when the rhythm imposed by machines is significant, prevention plans detailing how to adopt good postures and avoid repetitive movements, as well as the acquisition of ergonomic equipment that reduces physical stress, are measures that can reduce the risk of illness. Finally, the likelihood of feeling ill is also increased when employees' working hours do not match their family and social commitment (+6.9 pp), a finding that supports the use of individualized flexible work arrangements. This is consistent with the meta analysis provided by Shifrin & Michel (2022) showing that these arrangements are associated with better health.

The other significant variables influence both the feeling of sickness and the duration of SA and/or SP. Five of these relate to specific employment characteristics.

Thus, employees on fixed-term contracts and those working atypical work schedules are less likely to declare themselves ill (respectively by −11.2 and −1.4 pp). This subjective assessment can be explained the desirability to signal their commitment to their job hoping to be offered a permanent contract (Harbaugh & van Den Nouweland, 2005; Amilon & Wallette, 2009). On the contrary, public sector employees and those who work alone are found to be significantly more inclined to self-diagnose as ill (respectively by +2.4 and 2.2 pp). This finding is consistent with the fact that employees who self-select into the public sector have a higher preference for low work effort (Ehlert & García-Morán, 2022) and/or low attendance motivation (Mastekaasa, 2020). On their part, employees who work alone are less likely to receive help if they encounter work-related problems, which can create stress, or if they are victims of an accidental risk at work. Finally, when the firm implement specific quantified targets, employees are more likely to report at least one episode of illness (+2.3 pp) confirming that incentives for performance may have a downside effect on mental and physiological health (Andelic et al., 2024).

Regarding working conditions, our results show that excessive workload and high job demands are related to employees reporting illness more often. Specifically, this is the case for employees who work under pressure (+6.8 pp), suffer from a lack of time to perform tasks efficiently (+5.9 pp) and tensions at work (+10.6 pp). However, these deleterious effects of physical and psychological hardship on employees' perceived health could be mitigated when they receive help from their hierarchy (−5.7 pp), a result which supports the previous conclusions of Lu et al. (2013).

Uncertainty in the working environment also plays an important role in the feeling of illness confirming previous empirical evidence about job insecurity (Lepinteur, 2021). Significant past organizational changes are associated with an increased likelihood of feeling ill (+5.8 pp), possibly due to changes in previous productive routines, physical hazards and/or psychosocial risks (Ben Halima et al., 2023). Once again, this finding illustrates that effective occupational risk prevention is essential to moderate the health impact of future innovations. Furthermore, worrying about the future of one's job appear to increase employees' risk of feeling unwell (+4.7 pp).

4.4 Determinants of the duration of sickness presenteeism and absence.

Estimates of the associations between employment characteristics, working conditions and the duration of SA or SP are given in columns 1 and 4. The corresponding average marginal effects are reported in columns 2 and 5. While we have previously identified the drivers that can be used for health prevention policies, we now identify the characteristics that can be used by firms to design their policies to encourage or discourage absence and presence in case of illness. Given the daily cost of SA and SP, firms can determine the most effective policy instruments for regulating their duration and minimising the associated costs. In this sense, our results make it possible to use each tool appropriately by knowing its effects on the duration of SA and SP.

Health

Once we have controlled for the effects of the variables on the probability of feeling ill, our results confirm a well-known fact in the literature, namely that the level of self-reported health is by far the determinant with the strongest effect on both the duration of SA and SP (Miraglia & Johns, 2016). Our results indicate that when the declared level of health deteriorates from good to rather good, individuals tend to extend their duration of SP by just over 1.8 days, compared with an increase of around 0.8 days of SA. However, for a major decrease in declared health from good to bad or very bad, the lengthening in the duration in days of SA and SP tends to equalise, at around 3.3 days for each.

In contrast, improving health to a very good status decreases the duration of SP more than that of SA (−0.86- and −0.35-day, respectively). This indicates that healthier employees more often choose SP over SA when they feel ill. Finally, employees suffering from a chronic disease increase their duration of SA more than their duration of SP (by 0.90- and 0.39-day, respectively) probably indicative of repeated daily episodes of serious illness. Of course, while it is difficult to infer a direction of causality between these variables, these results reinforce our previous recommendations about the importance of preventive health policies.

Employment characteristics

Among the characteristics of employment, it is noteworthy that neither the salary level nor the tasks related to the employees' occupation influence the average duration of sick leave and presenteeism.

However, four of these employment characteristics are dual risk factors for sick leave duration only. First, performing supervisory tasks reduces the duration of SA by 0.79 days per year without significantly affecting that of SP. Being a supervisor likely entails responsibilities in overseeing the production process and assisting subordinates, which motivates supervisors not to take leave. Second, employees who work alone not only have a higher probability of feeling ill during the year but are also significantly less absent on average, by 0.29 days per year. In the absence of a work group that provides help and support, these employees find it more difficult to take time off to recover. Finally, the association of the following two characteristics with the duration of absence seems to confirm the influence of incentive schemes. Thus, public sector employees, who benefit from greater generosity in terms of payment in case of absence - no waiting days in the French public service at the time of the survey -, are on average 0.69 days more absent than their private sector counterparts. What's more, there is no evidence of a public sector motivation effect inducing workplace attendance during illness. Similarly, employees on temporary contracts, who benefit from less job security but may have the prospect of a permanent contract, have a shorter duration of sick leave, by 1.67 days per year. Our estimates are fairly close to those obtained by Ichino & Riphahn (2005) who, in the context of the probationary period, show that when employees are granted tenure and therefore protected, their weekly absence time decreases by 0.04 days. These results illustrate that attendance incentive systems are potentially effective instruments for employers.

The organization of working time on a daily or weekly basis is a factor that significantly affects the duration of SA and slightly that of SP. Long workweeks and atypical hours tend to decrease the duration of the former and increase that of the latter. These two findings suggest that when the organization of working time imposes long or irregular hours, employees cannot take time off without disrupting the production schedule or transferring their workload to their colleagues.

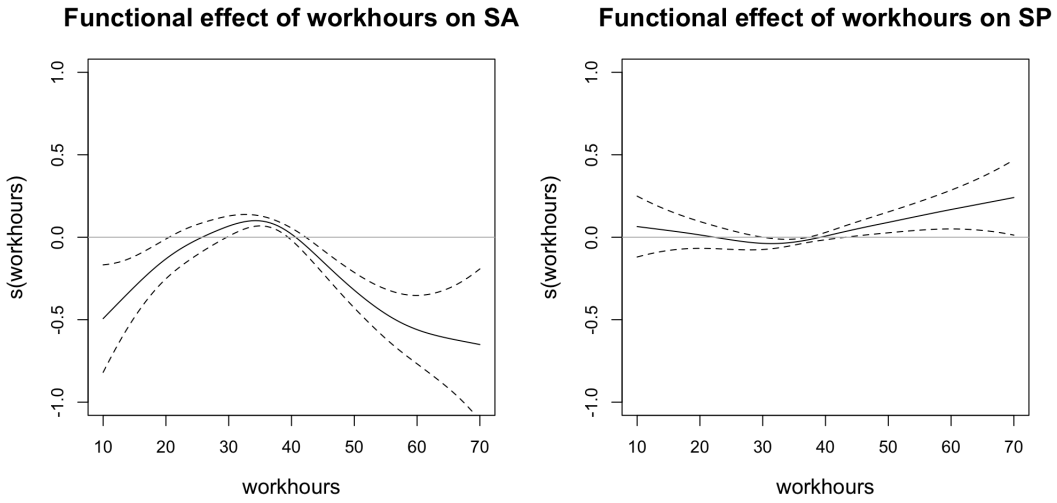

Figure 1.4 – Smooth functional forms of work hours on sickness absence and presenteeism

The relation between working time and duration of SA is concave, while this relation seems convex for SP (only significant at 10%). To better observe the relationship between working time and SA or SP, we model it as a smooth polytomic function to obtain the most accurate functional form. The smooth functions for both outcomes are depicted in figure 1.4 above, in which we can note that the working time beyond which ill employees will either decrease their SA duration or increase their SP duration roughly corresponds to the legal weekly working hours in France, i.e. 35 hours.

Finally, the existence of quantifiable targets at work seems to have the expected effect. In addition to increasing the probability of feeling sick, this incentive system increases the duration of SP by 0.53 days. However, the positive effect of the coefficient on the duration of SA, which is only significant at the 10% level, may imply that the relative costs of attendance and absence vary over the year. It is thus possible that, as the assessment period of performance draws to a close, employees seeking to achieve the required target may substitute SP for SA.

Working conditions

Our estimates show that working conditions are more strongly related to the duration of SP than to that of SA. Specifically, a high workload leads employees to forgo rest, thereby increasing the duration of SP days. This occurs when employees

must work under pressure to complete their tasks or when they lack time to meet set objectives, which increase the duration of SP by 0.89 and 0.61 days respectively. Also, when employees are unable to work at their own pace due to external demand, SP duration is lengthened by 0.34 days while SA duration decreases by 0.62 days. Notably, this constraint on pace does not appear to significantly influence feeling of illness. Once again, the need to meet customer or supplier demands makes it difficult to take time off, even in case of illness. This effect can be exacerbated when employees cannot be replaced, or when their specific skills require her presence. Nonetheless, when the workload can be shared with the help of coworkers or when employees receive support from their hierarchy, annual SP duration is shortened by 0.80 and 0.46 days respectively. Furthermore, help from colleagues is associated with a 0.38-day longer duration of SA (significant at only 10%). These findings are qualitatively in line with those of Arnold & de Pinto (2015). They suggest that, given the costs associated with SA and SP, firms aiming to reduce the SP duration should focus on better organizing employees' workload, encouraging mutual support, and also investing in manager training.

Furthermore, when employees are allowed to choose their working methods, they experience a 0.43-day decrease in SA duration, but there is no significant association with SP. Conversely, when employees are allowed to take initiative in their jobs, their duration of SP increases by 0.57 days without affecting SA duration. Nevertheless, these two job latitude characteristics are not shown to be protective to health in our estimates. This conclusion appears to be consistent with the results by Arnold (2016) regarding SP duration in Europe.

Taken together, these results suggest that when employees are given the opportunity to organize their work, they may substitute SA to SP when feeling ill. This might underline the fact that job latitude, through different channels, functions more as a feature of job demand –highlighting the pressure associated with workers' empowerment and responsibility– rather than as a job resource. Another possible explanation for the increase in SP lies in the positive relationship between intrinsic motivation at work and autonomy.

Employee's environment

Episodes of major change in the work environment are associated with an increase of 0.57 days in SA duration, confirming recent findings showing that extensive organisational changes negatively impact employees' health, leading to long-term sick leaves (Ben Halima et al., 2023). Additionally, as previously shown in the literature (Miraglia & Johns, 2016), our results indicate that uncertainty about the future of one's job would prompt workers to extend their attendance duration by 0.58 days when ill, while also reducing their SA duration by 0.27 days (significant at the 10% level). Finally, it is worth noting that the existence of tensions at work, whether with colleagues, supervisors, clients or subordinates, is associated with increases in SA and SP durations, by 0.57 and 0.66 days respectively. Over the course of the year, employees may feel compelled to come to work sick so as not to exacerbate tensions at work, while in due course the accumulation of such tensions may worsen their health further, leading to longer sick leaves.

4.5 Additional results

In this section, we present some additional results to check the robustness of our statements. All tables reporting the new evidence discussed in this section are provided in the appendices. To evaluate the results reported in these supplementary tables, we have checked whether the new estimated coefficients belong to the confidence interval of the initial estimates . We also investigate for changes in significance in the estimated coefficients. In most cases except those explicitly discussed below, the estimated coefficients associated with the significant variables in each of the following robustness tables do not significantly differ from those in Table 1.4.

Firstly, we examine whether our main results are impaired when sets of likely endogenous regressors are excluded from the models. Among the strong suspects of endogeneity are health characteristics, such as self-reported health and chronic disease. These are particularly likely to be endogenous to SA and SP durations, but also to be correlated with working conditions, as physically and mentally demanding jobs

The confidence intervals of the main results displayed in Table 1.4 are reported in Table A9 in the section B of the appendices.

can influence health. So, our first robustness check involved removing these health characteristics from the count part of the models. The results, presented in table A10 of the supplementary appendices document, show that while the coefficients tend to increase or decrease slightly in absolute value, the significant explanatory variables largely remain significant. One exception is help from colleagues in the model for SA duration, which decreases in significance. Conversely, some explanatory variables gain in significance such as having to achieve quantified target and the existence of work to life conflict in the models of SA and SP, respectively. However, even in these cases, the values remain within the confidence intervals of the initial estimates. Our second robustness check involved separately removing employment characteristics and working conditions from each model (Table A11 and A12). Similar to the results obtained when removing health characteristics, the exclusion of either set of variables generally leads to small changes in the magnitude of the coefficients. Once again, all the estimated coefficients fall within the confidence intervals of the initial results. The source of these minor changes lies in the introduction of omitted variable bias when an explanatory variable of interest is excluded. If a relevant variable is removed from the equation - even though it explains the duration of absence or presenteeism and is also correlated with the other explanatory variables in the model - it increases the endogeneity bias of the remaining explanatory variables. This occurs because its omission causes this variable to be absorbed into the error term of the equation. For instance, failing to control for declared health would amplify (resp. reduce) the effects of variables harmful to (resp. protective of) health, as the impact of poor health would be attributed to these variables. This is likely the case for experiencing work-to-life conflict, which is probably detrimental to mental health, or receiving help from colleagues, which serves as a protective factor against excessive effort. Thus, since fully addressing the endogeneity of all explanatory variables is unfeasible, these supplementary estimations underscore both the robustness of our initial findings and the importance of controlling for all relevant characteristics to minimize potential biases from omitted variables.

Secondly, we check the robustness of the results associated with the sample restrictions imposed on age, working time and the cumulative days of illness. The results are largely robust to inclusion of all employees whatever their age and working hours

(Table A13 and A14). The only noticeable change concerns the influence of working time: while the association between working time and SP duration remains insignificant, that with SA duration is significantly negative. In terms of average marginal effect, an increase of working time leads to a shorter decrease of -0.027 days of SA, instead of -0.034.

When the study samples are modified to include, respectively, employees reporting fewer than 30, more than 120 or up to 365 sick days over the year, we note that the results of the model explaining the duration of SP (table A15) are largely robust. All estimated coefficients belong to the confidence interval of those in Table 1.4. However, such variations in the sample selection rules according to the total number of sick days have a greater impact on the results relating to the duration of SA (Table A16). These robustness results are consistent with our expectations. Indeed, if we include in the sample employees with a higher number of sick days per year, this means that the severity of their illness is on average higher. Hence, according to our theoretical intuition, the more severe the perceived illness (i.e. the higher the illness shock δ_i), the more likely the worker is to choose absence rather than presence . Two factors, very poor declared health and working under pressure, saw their coefficients increase so much that they fell outside the confidence interval of the coefficients in Table 1.4. The average marginal effect of being in poor or very poor health, compared with being in good health, on the duration of SA increases from 3.35 days to 12.2 days when the sample restriction in terms of total duration of illness is extended to 365 days. Further, the average marginal effect of working under pressure increases from 0.06 supplementary days of SA to 1 supplementary day. Accordingly, this empirical evidence seems to show that employees who frequently work under pressure are more likely to suffer from a more serious illness. Finally, a number of determinants become also significant. This is the case for wage levels below 1,500 euros, when the pace of work is constrained by the rhythm of a machine or monitoring, or when workers receive assistance from their superiors or colleagues. While there are indeed some minor changes in the results, the overall logic clearly prevails.

This is confirmed by the sharper increase in the average duration of SA (from 6.33 to 15.80 days) compared to the average duration of SP (from 4.09 to 5.55 days) when employees reporting up to 365 days of sickness are included in the sample.

Thirdly, another methodological issue worth highlighting concerns our choice of copula. Indeed, as summarized in Cameron et al. (2004), when the multivariate cumulated distribution function is continuous, the copula is unique. However, this does not hold for the joint distribution of discrete variables. This led us to question the reliability of our results and compare them to those obtained when choosing the same margins but different copulae incorporating a certain degree of error according to the information criteria. We found that selecting the copula with the second smallest AIC does not change the results at all (Table A17). Each coefficient belongs to the confidence interval of the one presented in the main results. Moreover, deviating significantly from our selection criterion, i.e. choosing the copula associated with the highest AIC, does not radically modify the logic of the results either (Table A18). Some variables change in significance – 4 previously significant determinants (resp. 2) lose significance for SA duration (resp. SP duration) and for those which remain significant the value of the coefficients are systematically modified but always stand within the confidence interval of those presented in Table A9. These results corroborate those previously identified in the case studies presented in the paper of Marra et al. (2024). The authors actually found that the choice of the copula, on the contrary to the importance of choosing the right margins, did not seriously impact the estimates as long as the link between the margins are sufficiently well captured. Somehow, our evidence is generally consistent with the effective use of information criteria to select copula advocated by simulation and empirical case studies (Marra & Wyszynski, 2016; Marra et al., 2024).

Fourthly, as discussed above, we have outlined the various arguments supporting the inclusion of exclusion restrictions. To assess their relevance in our case study, we thus estimate our models without exclusion restrictions (results in Table A19 in the appendices).Furthermore, we also test the robustness of our results to the choice of the exclusion restriction by reporting estimates from models using the two other restrictions available in the data: experience of mistreatment or violence (Tables A20 and A21 in the appendices). We find that the estimated coefficients across these three alternative specifications are very similar to those from the models using the chosen restriction (family conflicts), thereby supporting the consistency of our previous findings. Furthermore, consistent with Marra & Wyszynski (2016), we find

no real change in efficiency with the inclusion or exclusion of our restriction variable when comparing the results in Table 1.4 and Table A19. Thus, whether or not we include an exclusion restriction does not alter our conclusions in any way .

Finally, while the exclusion restrictions used in our econometric model address the endogenous participation problem, they do not insure causality between explanatory variables and the choice of SA and SP. Therefore, it is necessary to explicitly account for the problem of self-selection of employees into jobs where they are better able to cope with the job characteristics and working conditions. Healthier employees are likely to accept more demanding working conditions, whether mentally or physically challenging. As a result, the observed association between working conditions and SA and SP durations could be biased. To address this issue, we follow the method proposed by Llena-Nozal (2009) and Böckerman et al. (2012) controlling for serious health episodes experienced in the last three years. Additionally, the exclusion variables used in the participation equation also allow the estimates to be conditioned on health predictors. However, it is also plausible that other unobservable employee traits influence both their choice of job and their propensity for absence or presenteeism. For example, individuals who have experienced recent episodes of unemployment may be less demanding in terms of employment conditions while simultaneously being more reluctant to take time off, fearing the risk of loosing their job. to account for this, we also condition on employees' short and long episodes of unemployment within the last three years. According to the results reported in Table A22, our estimates of the association between employment characteristics and working conditions with SA and SP duration remain consistent in terms of both significance and the values of the estimated coefficients. This suggests that the inclusion of these variables measuring past health and unemployment does not provide additional control for the self-selection of employees into certain types of employment. It is possible that the healthy worker effect has already been accounted for, first through the exclusion restrictions in the participation equation, and second through the inclusion of self-reported health status in the outcome equations.

The models presenting the main results in Table 1.4 show the best AIC compared to the models including the two other restrictions and that without any restriction.

5 Concluding remarks and discussion

Our goal was to propose an innovative contribution to the understanding of the causes of sickness absence and presenteeism by analysing which employment characteristics and working conditions are primarily associated with employees' feeling of illness and secondary with the choice between work and take time off to rest in case of illness. Previous literature has reviewed the extended costs for employees, firms and social security systems of sickness absence and presenteeism and studied their determinants but has largely neglected the need to empirically model both phenomena simultaneously and conditionally to the realization of the illness shock.

Our theoretical framework assumes that each employee experiences a random daily health shock which, beyond a certain threshold, forces them to decide whether to take sick leave or come to work while being ill. Over the course of a year, these repeated choices result in a certain duration of SP and SA. The empirical equivalent of this decision involves estimating a model that explains the cumulative durations of SA and SP, conditional on a reservation level of illness. Accounting for this condition requires estimating a first equation for endogenous participation in the choice between absence or presence, i.e., modeling the probability of exceeding a severity threshold that leads the employee to reporting ill, and a second equation explaining the duration of sickness absence (SA) or presenteeism (SP). Indeed, the problem of endogeneity, which has been neglected in the literature so far, arises from the fact that working conditions can impact the likelihood of becoming ill separately from the intensive margin decision of staying home or working while sick for any particular number of days.

Therefore, using a copula-based econometric method, we investigated the determinants involved in the total duration of SA and SP while accounting for the endogenous participation bias. To this end, we analysed the French Ministry of Labour's 2016 "Working Conditions and Psychosocial Risks" survey representative of the working population. It extends the working conditions surveys carried out since 1978, which provided the framework for the European Working Conditions Survey. This survey further includes precise information about both sickness absence and presenteeism allowing to identify which employees have ever been ill during the studied year. In addition, our econometric method offers a great flexibility in the

choice of distributions modelling the count outcomes and their link with the endogenous participation process. Unlike traditional models of endogenous participation whose assumptions were not very flexible, it makes it possible to obtain very stable results under different specifications. We have identified potential explanations that can help improve our understanding of sickness absence and attendance thereby guiding personnel policies.

This new method enabled us to pinpoint several areas of improvement in this regard. Firstly, we showed that failing to account for the endogeneity caused by illness would have significantly biased the estimates of the association between employment characteristics, working conditions and SA or SP. Indeed, this would have led us to misidentify the true determinants associated to the duration of SA and SP and thus compromise our conclusions and recommendations regarding corporate policies. Secondly, we need to stress the importance of implementing company health prevention policies. A significant number of employment characteristics and working conditions involved in the duration of SA or SP are primarily associated to illness. More specifically, we were able to identify which of the determinants are single risk factor, only associated with the feeling of illness, but not with the choice between SA and SP. For example, it is the case of constraints on the work pace created by machines, norms, hierarchical control, and relationships with colleagues and conflicts between working hours and family and social time. Using the former to design prevention policies helping employees to deal with these difficult rhythms will ultimately reduce both SA and SP. Conversely, although it is possible to focus on double risk factors to prevent illness, firm managers must be aware of their additional effects on both SA and SP behaviours once the worker is ill.

Finally, depending on the relative costs of SA and SP according to the objectives pursued by the firm, we found that SA and SP cannot be controlled via the same channels. While characteristics of the employment contract and job content can be used to manage absence behaviour, they are often useless to reduce SP. In order to manage SP, firms decision makers need to focus more on job demands that put pressure on work intensity, but also on employee autonomy and support from supervisors and colleagues, that can only be achieved by modifying the management methods of work organization.

A major limitation of our study is the impossibility of establishing causality between our explanatory variables and the annual duration of SA and SP. Unfortunately, given the information available in the data, it is impossible to deal with the endogeneity of all employment characteristics and working conditions because of the lack of appropriate instruments. We have tried to cope with the problem of self-selection of employees in their jobs according to their ability to perform in high demanding situations or their risk propensity, taking into account their past episodes of health problems and labour market mobility. Our results do not appear to be affected by these additions. Moreover, in cases of measurement error due to memory bias in reporting of the duration of SA and SP, the employees' psychological state (whether pessimistic or optimistic) may introduce individual-specific errors and classic problems of heteroscedasticity. However, we have shown that our econometric approach - combining the use of copula to model the association between the distributions of observed illness and durations of SA and SP along with the selection of Negative binomial and Poisson Inverse Gaussian distributions for these durations - is likely the most appropriate for addressing such issues.

However, in order to efficiently treat endogeneity, future studies should use longitudinal data to take into account the individuals' unobserved traits, such as their illness propensity, state of mind (optimism or pessimism) and their risk preferences.

If our detailed data has given us the opportunity to control for the endogeneity of illness and to model its link with SA and SP, some shortcomings of our study open several avenues for future research. To explicitly account for the relationships between unobservable variables that may influence the phenomena of sickness absence and presenteeism, it would be necessary to develop a trivariate model in which illness, durations of SA and SP are jointly modeled. Such a development would allow for a more accurate estimation of their determinants affecting these phenomena. Another potential improvement could stem from the availability of data enabling the modeling of the dynamics of absence and presenteeism choices. Instead of analyzing the duration of these phenomena, a field experiment could record episodes of absence and presenteeism on a daily basis over a given period, allowing the identification of the drivers of attendance or rest decisions conditional on previous episodes.

Finally, future research would have to explore firms' production technologies in more detail. As explained by Coles et al. (2007), when technology requires capital/labour or worker complementarity, absenteeism is more costly for a firm, causing employer to implement incentive mechanisms to reduce absence. Therefore, an employee's choice to increase the duration of SP or decrease the duration of SA, for example, will ultimately depend on the firm's technology and related incentive mechanisms. Hence, to better understand choices between SA and SP, as well as to control for possible self-selection of employees among firms with different technologies, future research should also focus on the organizational characteristics of the firms. The use of employer-employee matched data would therefore be perfectly suited to explain SP and SA according to different production methods.

Bibliography

Amilon, A., & Wallette, M. (2009). Work absence—a signalling factor for temporary workers? *Labour*, *23*, 171-194.

Andelic, N., Allan, J., Bender, K., Powell, D., & I.Theodossiou. (2024). Performance-related pay, mental and physiological health. *Industrial Relations: A Journal of Economy and Society*, *63*(1), 3–25.

Arnold, D. (2016). Determinants of the annual duration of sickness presenteeism: Empirical evidence from european data. *Labour*, *30*(2), 198–212.

Arnold, D., & de Pinto, M. (2015). How are work-related characteristics linked to sickness absence and presenteeism? theory and data. *Schmollers Jahrbuch*, *135*, 465-498.

Bakker, A. B., & Demerouti, E. (2007). The job demands-resources model: state of the art. *Journal of Managerial Psychology*, *22*(3), 309–328.

Barmby, T., & Larguem, M. (2009). Coughs and sneezes spread diseases: An empirical study of absenteeism and infectious illness. *Journal of Health Economics*, *28*(5), 1012-1017.

Barmby, T., Orme, C., & Treble, J. (1991). Worker absenteeism: An analysis using microdata. *The Economic Journal*, *101*, 214-229.

Barnay, T. (2016). Health, work and working conditions: a review of the european economic literature. *The European Journal of Health Economics*, *17*, 693–709.

Beemsterboer, W., Stewart, R., Groothoff, J., & Nijhuis, F. (2009). A literature review on sick leave determinants (1984–2004). *International Journal of Occupational Medicine and Environmental Health*, *22*(2), 169–179.

Ben Halima, M. A., Greenan, N., & Lanfranchi, J. (2023). Getting sick for profit? the impact of cumulative ict and management changes on long term sickness absence. *Journal of Economic Behavior and Organization*, *212*, 659-688.

Berg, N., Kiviruusu, O., Karvonen, S., Rahkonen, O., & Huurre, T. (2017). Pathways from problems in adolescent family relationships to midlife mental health via early adulthood disadvantages–a 26-year longitudinal study. *PLos One*, *12*(5), 460-468.

Bergström, G., Bodin, L., Hagberg, J., Lindh, T., Aronsson, G., & Josephson, M. (2009a). Does sickness presenteeism have an impact on future general health? *International archives of occupational and environmental health*, *82*(10), 1179–1190.

Bratti, M., & Miranda, A. (2011). Endogenous treatment effects for count data models with endogenous participation or sample selection. *Health Economics*, *20*, 1090-1109.

Brown, S., & Sessions, J. G. (2004). Absenteeism, 'presenteeism', and shirking. *Economic Issues*, *9*(1), 15–23.

Böckerman, P., Bryson, A., & Ilmakunnas, P. (2012). Does high involvement management improve worker wellbeing? *Journal of Economic Behavior & Organization*, *84*(2), 660–680.

Böckerman, P., Bryson, A., Kauhanen, A., & Kangasniemi, M. (2020). Does job design make workers happy? *Scottish Journal of Political Economy*, *67*(1), 31-52.

Cameron, A. C., Li, T., Trivedi, P. K., & Zimmer, D. M. (2004). Modelling the differences in counted outcomes using bivariate copula models with application to mismeasured counts. *The Econometrics Journal*, *7*(2), 566–584.

Clark, C., Caldwell, T., Power, C., & Stansfeld, S. A. (2010). Does the influence of childhood adversity on psychopathology persist across the lifecourse? a 45-year prospective epidemiologic study. *Annals of epidemiology*, *20*(5), 385-394.

Coles, M., Lanfranchi, J., Skalli, A., & Treble, J. (2007). Pay, technology, and the cost of worker absence. *Economic Inquiry*, *45*, 268–285.

Ehlert, A., & García-Morán, E. (2022). Workers' self-selection into public sector employment: A tale of absenteeism. *Kyklos*, *75*, 394-409.

Evans-Lacko, S., & Knapp, M. (2016). Global patterns of workplace productivity for people with depression: absenteeism and presenteeism costs across eight diverse countries. *Social Psychiatry and Psychiatric Epidemiology Journal*, *51*, 1525–1537.

Gerich, J. (2016). Determinants of presenteeism prevalence and propensity: Two sides of the same coin? *Archives of Environmental & Occupational Health*, *71*(4), 189–198.

Green, F., Felstead, A., Gallie, D., & Henseke, G. (2021). Working still harder. *ILR Review*, *75*(2), 458–487.

Greene, W. H. (1997). *FIML estimation of sample selection models for count data.* (Discussion Paper EC-97-02, Department of Economics, Stern School of Business, New York University)

Grinza, E., & Rycx, F. (2020). The impact of sickness absenteeism on firm productivity: New evidence from belgian matched employer-employee panel data. *Industrial Relations*, *59*(1), 150–194.

Hansen, C. D., & Andersen, J. H. (2008). Going ill to work – what personal circumstances, attitudes and work-related factors are associated with sickness presenteeism? *Social Science & Medicine*, *67*, 956–964.

Harbaugh, W. T., & van Den Nouweland, A. (2005). Demonstrating worker quality through strategic absenteeism. *International Game Theory Review*, *23*, 189-209.

Hasebe, T., & Vijverberg, W. P. (2012). *A Flexible Sample Selection Model: A GTL-Copula Approach* (IZA Discussion Papers N° 7003). Institute of Labor Economics (IZA).

Havet, N., & Plantier, M. (2023). The links between difficult working conditions and sickness absences in the case of french workers. *Labour*, *37*, 160–195.

Hemp, P. (2004). Presenteeism: At work – but out of it. *Harvard Business Review*, *82*(10), 49–58.

Heywood, J. S., Jirjahn, U., & Wei, X. (2008). Teamwork, monitoring and absence. *Journal of Economic Behavior & Organization*, *68*, 676–690.

Hilbe, J. (2014). *Modelling count data* (1st éd.). New York: Cambridge University Press.

Hillion, M., Mette, C., & Pisarik, J. (2021). L'effet des conditions de travail sur la santé et le recours aux soins. *Preliminary version*, 1–98.

Hirsch, B., Lechmann, D. S. J., & Schnabel, C. (2017). Coming to work while sick: An economic theory of presenteeism with an application to german data. *Oxford Economic Papers*, *69*(4), 1010–1031.

Hudson, H. L., Nigam, J. A., Sauter, S. L., Chosewood, L., Schill, A. L., & Howard, J. E. (2019). *Total worker health: Integrative approaches to safety, health, and well-being* (Rapport technique). American Psychological Association.

Ichino, A., & Riphahn, R. T. (2005). The effect of employment protection on worker effort: Absenteeism during and after probation. *Journal of the European Economic Association*, *3*(1), 120–143.

Janssens, H., Clays, E., De Clercq, B., De Backer, G., & Braeckman, L. (2013). The relation between presenteeism and different types of future sickness absence. *Journal of occupational health*, *55*.

Jensen, U. T., Andersen, L. B., & Holten, A.-L. (2019). Explaining a dark side: Public service motivation, presenteeism, and absenteeism. *Review of Public Personnel Administration*, *39*(4), 487-510.

Johansson, G., & Lundberg, I. (2004). Adjustment latitude and attendance requirements as determinants of sickness absence or attendance. empirical tests of the illness flexibility model. *Social Science & Medicine*, *58*, 1857–1868.

Johns, G. (2010). Presenteeism in the workplace: A review and research agenda. *Journal of Organizational Behavior*, *31*, 519–542.

Karanika-Murray, M., & Cooper, C. L. (2018). Presenteeism: an introduction to prevailing global phenomenon. *Cambridge Core*, 1–34.

Kelly, E. L., & Moen, P. (2020). *Review of overload: How good jobs went bad and what we can do about it* (Rapport technique). Princeton University Press.

Kiviet, J. F. (2020). Testing the impossible: Identifying exclusion restrictions. *Journal of Econometrics*, *218*(2), 294-316.

Landstedt, E., Hammarström, A., & Winefield, H. (2015). How well do parental and peer relationships in adolescence predict health in adulthood? *Scandinavian journal of public health*, *43*(5), e0178136.

Lanfranchi, J., & Treble, J. (2010). Just-in-time production, work organization and absence control. *The Manchester School*, *78*(5), 460–483.

Leineweber, C., Westerlund, H., Hagberg, J., Svedberg, P., & Alexanderson, K. (2012). Sickness presenteeism is more than an alternative to sickness absence: results from the population-based slosh study. *International Archives of Occupational and Environmental Health*.

Lepinteur, A. (2021). The asymmetric experience of gains and losses in job security on health. *Health Economics*, *30*, 2217–2229.

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

Lohaus, D., & Habermann, W. (2019). Presenteeism: A review and research directions. *Human Ressource Management Review*, *29*, 43–58.

Lu, L., Lin, H. Y., & Cooper, C. L. (2013). Unhealthy and present: Motives and consequences of the act of presenteeism among taiwanese employees. *Journal of Occupational Health Psychology*, *18*(4), 406–416.

Marie, O., & Val Castello, J. (2023). Sick leave cuts and (unhealthy) returns to work. *Journal of Labor Economics*, *41*, 923–956.

Marra, G., Radice, R., & Zimmer, D. (2024). A unifying switching regime regression framework with applications in health economics. *Econometric Reviews*, *43*(1), 52-70.

Marra, G., & Wyszynski, K. (2016). Semi-parametric copula sample selection models for count responses. *Computational Statistics and Data Analysis*, *104*, 110-129.

Mastekaasa, A. (2020). Absenteeism in the public and the private sector: Does the public sector attract high absence employees? *Journal of Public Administration Research and Theory*, *30*(1), 60–76.

Miraglia, M., & Johns, G. (2016). Going to work ill: A meta-analysis of the correlates of presenteeism and a dual-path model. *Journal of Occupational Health Psychology*, *21*(3), 261–283.

Niedhammer, I., Chastang, J. F., Sultan-Taïeb, H., Vermeylen, G., & Parent-Thirion, A. (2013). Psychosocial work factors and sickness absence in 31 countries in europe. *The European Journal of Public Health*, *23*, 622–629.

Nikolova, M., & Cnossen, F. (2020). What makes work meaningful and why economists should care about it. *Labour Economics*, *65*, Aricle 101847.

Norman, R. E., Byambaa, M., De, R., Butchart, A., Scott, J., & Vos, T. (2012). The long-term health consequences of child physical abuse, emotional abuse, and neglect: a systematic review and meta-analysis. *PLoS medicine*, *9*(11), e1001349.

Pauly, M. V., Nicholson, S., Polsky, D., Berger, M. L., & Sharda, C. (2008). Valuating reductions in on-the-job illness: 'presenteeism' from managerial and economic perspectives. *Health Economics*, *17*, 469–485.

Pichler, S., Wen, K., & Ziebarth, N. (2021). Positive health externalities of mandating paid sick leave. *Journal of Policy Analysis and Management*, *40*, 715–743.

Pichler, S., & Ziebarth, N. (2015). The pros and cons of sick pay schemes: Testing for contagious presenteeism and noncontagious absenteeism behavior. *Journal of Public Economics*, *156*, 14–33.

Schultz, A. B., & Edington, D. W. (2007). Employee health and presenteeism: A systematic review. *Journal of Occupational Rehabilitation*, *17*, 547–579.

Shifrin, N. V., & Michel, J. (2022). Flexible work arrangements and employee health: A meta-analytic review. *Work & Stress*, *36*, 60–85.

Skagen, K., & Collins, A. M. (2016). The consequences of sickness presenteeism on health and wellbeing over time: A systematic review. *Social Sciences & Medicine*, *161*, 169–177.

Sorensen, G., Dennerlein, T., Peters, S. E., Sabbath, E. L., Kelly, E. L., & Wagner, G. R. (2021). The future of research on work, safety, health and wellbeing: A guiding conceptual framework. *Social Science and Medicine*, *269*, 113593.

Terza, J. V. (1998). Estimating count data models with endogenous switching: Sample selection and endogenous treatment effects. *Journal of Econometrics*, *84*, 129–154.

Treble, J., & Barmby, T. (2011). *Worker absenteeism and sick pay*. Cambridge University Press.

Turanovic, J. J. (2022). Exposure to violence and victimization: Reflections on 25 years of research from the national longitudinal study of adolescent to adult health. *Journal of Adolescent Health*, *71*(6), s14-s23.

Appendices

1 Descriptive statistics

Table A1 – Variables description

Socio-demographics:	
Age	Continuous (from 18 to 65 years old)
Gender	Being a female = 1
Marital status	Being single = 1
Children	Having at least one dependent child = 1
Education	Categories of degree levels: No degree/less than High school diploma/High school diploma/Higher education degree (Ref: High school diploma)
Health state	Health levels: Bad or Very Bad/Rather Good/Good/Very Good (Ref: Good)
Chronic disease	Having a chronic disease = 1
Job characteristics:	
Occupation	Labourer/Employee/ intermediate Profession/ executive (Ref: Employee)
Supervisor	Q: In your main job, do you sometimes supervise the work of other employees? A: Yes, this is my main task/Yes, but it is not my main task/No (Dichotomous: Yes is my main task = 1)
Public sector	Working in the public sector = 1 (State, Territorial Collectivity, etc.)
Seniority	Continuous (from 0 to 49 years)
Temporary contract contract	Dichotomized with fixed-term contracts (CDD and interim) = 1, opposed to permanent contract (CDI)
Atypical hours	Working nights, evenings, Saturdays or Sundays = 1
Work alone	Q: Do you work alone? A: always/often/sometimes/never, dichotomized with always = 1
Working time	Continuous (from 8 to 70 hours/week)
Wage	Categories of wage: 100–1,199/1,200–1,499/1,500–1,799/1,800–2,499/2,500–2,999/3,000–20,000
Quantified targets	Q: Do you have to achieve specific quantified objectives? A: Yes/No (Dichotomous for yes = 1)
Company size	Number of employees From 1 to 9/10-49/50-199/200-499/500+/ NoInfo (Ref = 10-49)
Activity sector	17 groups of sectors + No info (Ref = Public Administration, Education, Health and social activities)
Working conditions	
Rhythm paced by	a machine / external demand / production norms / computerized control / hierarchical monitoring / dependence with colleagues (dichotomous variables)
Lack of time	Q: Do you lack time to do your job properly A: Yes/No
Under pressure	Always or often work under pressure (sometimes not included), dichotomized
Supervisor's help	Q: If you have trouble doing delicate, complicated work, are you helped by your superiors? A: Yes/No
Colleagues' help	Q: If you have trouble doing delicate, complicated work, are you helped by your colleagues? A: Yes/No
Choice of methods	Q: You receive orders, instructions, instructions for use. To do your job correctly, do... A: ...you strictly apply the instructions/...you do otherwise/not applicable (no orders , instructions or instructions), dichotomized for doing otherwise
Initiative	Q: Does your work require you to take Initiatives? A: Always / Often / Sometimes / Never, dichotomized for always and often
Work-to-life conflicts	Q: In general, do your working hours match your social and family commitments outside of your work? A: Very Good/Good/Bad/Very Bad, dichotomized with bad or very bad = 1
Tensions at work	Dichotomous: tensions with the public, the hierarchy, the colleagues or the subordinates = 1 if at least one is true
Fear of job future	Q: For the coming year, do you fear for your job? A: Yes/No
Organizational changes	Having their work environment strongly modified by a change of position or function, technique, restructuring or relocation of the firm, work organization, a layoff plan, a buyout or change of management team or for another reason (= 1 if at least one of the reasons is true)
Exclusion restrictions	
Family conflicts	Q: Have you been strongly affected by family conflicts concerning you or your family? A: Yes/No
Mistreatment	Q: Have you been subjected to physical or psychological mistreatments? A: Yes/No
School violence	Q: Have you experienced conflict or repeated violence at school or in your neighbourhood (racketeering, etc.)? A: Yes/No

Note: Q for question and A for answer

Table A2 – Descriptive statistics of covariables

	Study sample $N = 19,331$				Overall sample $N = 24,640$			
Determinants	Mean	St. Dev.	Min	Max	Mean	St. Dev.	Min	Max
Age	44.216	10.305	18	65	44.744	10.627	15	129
Female	0.564	0.496	0	1	0.553	0.497	0	1
Single	0.229	0.420	0	1	0.234	0.424	0	1
Child	0.608	0.488	0	1	0.594	0.491	0	1
Education:								
no diploma	0.063	0.243	0	1	0.068	0.252	0	1
< high school degree	0.306	0.461	0	1	0.317	0.465	0	1
high school degree	0.183	0.386	0	1	0.183	0.386	0	1
> high school degree	0.449	0.497	0	1	0.431	0.495	0	1
Health:								
bad or very bad	0.035	0.184	0	1	0.052	0.222	0	1
rather good	0.221	0.415	0	1	0.231	0.421	0	1
good	0.518	0.500	0	1	0.503	0.500	0	1
very good	0.226	0.419	0	1	0.215	0.411	0	1
Chronic disease	0.277	0.447	0	1	0.297	0.457	0	1
Occupation:								
labourer	0.173	0.378	0	1	0.176	0.381	0	1
employee	0.317	0.465	0	1	0.332	0.471	0	1
intermediate	0.314	0.464	0	1	0.306	0.461	0	1
executive	0.188	0.391	0	1	0.179	0.383	0	1
Supervisor	0.135	0.341	0	1	0.131	0.338	0	1
Public sector	0.421	0.494	0	1	0.378	0.485	0	1
Seniority	15.179	10.953	0	49	14.827	10.997	0	54
Temporary contract	0.064	0.245	0	1	0.059	0.236	0	1
Atypical schedules	0.561	0.496	0	1	0.595	0.491	0	1
Work alone	0.282	0.450	0	1	0.320	0.466	0	1
Working time	36.819	8.098	10	70	37.472	10.973	1	168
Wage:								
100–1,199	0.121	0.327	0	1	0.148	0.356	0	1
1,200–1,499	0.190	0.392	0	1	0.190	0.392	0	1
1,500–1,799	0.195	0.397	0	1	0.191	0.393	0	1
1,800–2,499	0.281	0.450	0	1	0.268	0.443	0	1
2,500–2,999	0.082	0.274	0	1	0.077	0.267	0	1
3,000–20,000	0.112	0.316	0	1	0.105	0.307	0	1
Quantified targets	0.272	0.445	0	1	0.273	0.446	0	1
Firm size:								
<10	0.141	0.348	0	1	0.141	0.348	0	1
10-50	0.243	0.429	0	1	0.243	0.429	0	1
50-200	0.226	0.418	0	1	0.221	0.415	0	1
200-500	0.130	0.337	0	1	0.127	0.333	0	1
>500	0.211	0.408	0	1	0.207	0.405	0	1
Work paced by:								
machines	0.106	0.307	0	1	0.108	0.310	0	1
external demand	0.699	0.459	0	1	0.702	0.457	0	1
norms	0.448	0.497	0	1	0.444	0.497	0	1
computing control	0.382	0.486	0	1	0.355	0.478	0	1
monitoring	0.281	0.449	0	1	0.257	0.437	0	1
colleagues relation	0.292	0.455	0	1	0.273	0.445	0	1
Tensions at work	0.568	0.495	0	1	0.550	0.497	0	1
Under pressure	0.301	0.459	0	1	0.302	0.459	0	1
Lack of time	0.275	0.446	0	1	0.268	0.443	0	1
Supervisors' help	0.651	0.477	0	1	0.573	0.495	0	1
Colleagues' help	0.835	0.371	0	1	0.739	0.439	0	1
Choice of methods	0.819	0.385	0	1	0.729	0.445	0	1
Initiative	0.456	0.498	0	1	0.494	0.500	0	1
Work-to-family conflicts	0.169	0.375	0	1	0.184	0.388	0	1
Fear of job future	0.188	0.391	0	1	0.207	0.405	0	1
Occupational changes	0.444	0.497	0	1	0.419	0.493	0	1

2 AIC and BIC value tables for every SA and SP duration models

In the following section, we present all the value tables for the AIC and BIC information criteria when we consider different distributions for the feeling of illness, i.e. logit, probit or cloglog. The three discrete distributions give very similar results. Moreover, for each of the models, the logit distribution appears to be the best distribution associated with the probability of feeling ill.

Table A3 – AIC values for the discrete marginal distributions and copulas with a logit in first step

	Sickness absence (SA)					Sickness presenteeism (SP)			
Copula	Poisson	PIG	NB I	NB II	Copula	Poisson	PIG	NB I	NB II
Normal	172,142	83,413	80,876	80,440	Normal	103,116	78,028	78,674	78,753
Frank	168,526	83,402	80,865	80,424	Frank	102,417	77,996	78,209	78,255
AMH	168,014	83,414	80,872	80,437	AMH	102,817	78,035	78,531	78,572
FGM	167,340	83,398	80,864	80,418	FGM	102,346	78,041	78,523	78,571
Student-t	167,450	83,503	80,878	80,461	Student-t	103,094	**77,958**	78,361	78,458
Clayton	168,810	83,429	80,884	80,453	Clayton 90	101,176	78,013	78,174	78,234
Clayton 180	168,201	83,149	80,825	80,359	Clayton 270	103,319	78,085	78,774	78,808
Joe	168,066	83,095	80,816	**80,348**	Joe 90	103,320	78,067	78,775	78,810
Joe 180	168,681	83,429	80,884	80,453	Joe 270	101,078	78,026	78,156	78,216
Gumbel	168,562	83,297	80,850	80,406	Gumbel 90	103,320	78,019	78,739	78,808
Gumbel 180	168,674	83,429	80,884	80,453	Gumbel 270	102,166	77,991	78,372	78,457

Note: logit in first step

Table A4 – BIC values for the discrete marginal distributions and copulas with a logit in first step

	Sickness absence (SA)					Sickness presenteeism (SP)			
Copula	Poisson	PIG	NB I	NB II	Copula	Poisson	PIG	NB I	NB II
Normal	173,244	84,522	81,985	81,549	Normal	104,217	79,137	79,783	79,862
Frank	169,627	84,512	81,974	81,533	Frank	103,519	79,106	79,318	79,365
AMH	169,116	84,524	81,981	81,546	AMH	103,918	79,145	79,640	79,681
FGM	168,442	84,507	81,973	81,528	FGM	103,448	79,151	79,632	79,681
Student-t	168,552	84,613	81,987	81,570	Student-t	104,196	**79,068**	79,470	79,567
Clayton	169,911	84,538	81,994	81,562	Clayton 90	102,277	79,123	79,283	79,343
Clayton 180	169,302	84,258	81,935	81,468	Clayton 270	104,421	79,194	79,883	79,917
Joe	169,167	84,204	81,925	**81,457**	Joe 90	104,421	79,177	79,884	79,919
Joe 180	169,782	84,538	81,994	81,562	Joe 270	102,179	79,135	79,265	79,325
Gumbel	169,664	84,406	81,959	81,515	Gumbel 90	104,422	79,128	79,849	79,917
Gumbel 180	169,776	84,538	81,994	81,562	Gumbel 270	103,267	79,100	79,482	79,566

Note: logit in first step

Table A5 – AIC values for the discrete marginal distributions and copulas with a probit in first step

	Sickness absence (SA)					Sickness presenteeism (SP)			
Copula	Poisson	PIG	NB I	NB II	Copula	Poisson	PIG	NB I	NB II
Normal	168,270	83,415	80,878	80,442	Normal	111,338	78,030	79,907	81,938
Frank	169,024	83,404	80,867	80,426	Frank	102,418	77,998	78,210	78,256
AMH	168,027	83,416	80,874	80,439	AMH	102,818	78,037	78,533	78,574
FGM	167,387	83,400	80,866	80,421	FGM	102,352	78,043	78,524	78,573
Student-t	167,456	83,505	80,880	80,463	Student-t	103,093	**77,960**	78,361	78,458
Clayton	168,829	83,431	80,886	80,455	Clayton 90	101,179	78,015	78,172	78,232
Clayton 180	168,191	83,151	80,827	80,361	Clayton 270	114,942	78,088	78,776	78,809
Joe	168,067	83,097	80,817	**80,350**	Joe 90	103,322	78,071	78,776	78,811
Joe 180	168,709	83,431	80,886	80,455	Joe 270	101,078	78,027	78,154	78,214
Gumbel	168,569	83,300	80,852	80,408	Gumbel 90	103,323	78,022	78,742	78,810
Gumbel 180	168,677	83,431	80,886	80,455	Gumbel 270	102,166	77,992	78,372	78,456

Note: probit in first step

Table A6 – BIC values for the discrete marginal distributions and copulas with a probit in first step

	Sickness absence (SA)					Sickness presenteeism (SP)			
Copula	Poisson	PIG	NB I	NB II	Copula	Poisson	PIG	NB I	NB II
Normal	169,372	84,524	81,987	81,551	Normal	112,439	79,139	81,016	83,047
Frank	170,125	84,514	81,976	81,535	Frank	103,519	79,108	79,319	79,366
AMH	169,128	84,526	81,983	81,548	AMH	103,919	79,147	79,642	79,683
FGM	168,488	84,509	81,975	81,530	FGM	103,453	79,153	79,634	79,682
Student-t	168,558	84,614	81,989	81,572	Student-t	104,195	**79,069**	79,470	79,567
Clayton	169,931	84,540	81,995	81,564	Clayton 90	102,280	79,124	79,281	79,342
Clayton 180	169,292	84,261	81,937	81,470	Clayton 270	116,043	79,197	79,885	79,918
Joe	169,169	84,206	81,927	**81,459**	Joe 90	104,423	79,180	79,886	79,920
Joe 180	169,811	84,540	81,995	81,564	Joe 270	102,180	79,136	79,263	79,323
Gumbel	169,670	84,409	81,961	81,518	Gumbel 90	104,424	79,131	79,851	79,919
Gumbel 180	169,779	84,540	81,995	81,564	Gumbel 270	103,268	79,101	79,481	79,565

Note: probit in first step

Table A7 – AIC values for the discrete marginal distributions and copulas with a cloglog in first step

	Sickness absence (SA)					Sickness presenteeism (SP)			
Copula	Poisson	PIG	NB I	NB II	Copula	Poisson	PIG	NB I	NB II
Normal	168,492	83,550	81,013	80,578	Normal	103,261	78,178	78,819	78,898
Frank	167,590	83,539	81,000	80,560	Frank	103,549	78,150	78,349	78,399
AMH	168,133	83,551	81,008	80,574	AMH	102,966	78,182	78,674	78,716
FGM	167,448	83,534	80,999	80,555	FGM	102,500	78,190	78,666	78,717
Student-t	167,533	83,642	81,017	80,597	Student-t	103,207	**78,112**	78,500	78,597
Clayton	169,182	83,565	81,020	80,589	Clayton 90	101,324	78,163	78,286	78,946
Clayton 180	168,307	83,282	80,957	80,485	Clayton 270	103,461	78,233	78,910	78,944
Joe	168,174	83,227	80,945	**80,472**	Joe 90	103,460	78,217	78,912	78,946
Joe 180	168,800	83,565	81,020	80,589	Joe 270	101,225	78,174	78,267	78,946
Gumbel	168,692	83,435	80,986	80,539	Gumbel 90	103,460	78,173	78,884	78,946
Gumbel 180	168,807	83,565	81,020	80,589	Gumbel 270	102,314	78,141	78,500	78,588

Note: cloglog in first step

Table A8 – BIC values for the discrete marginal distributions and copulas with a cloglog in first step

	Sickness absence (SA)					Sickness presenteeism (SP)			
Copula	Poisson	PIG	NB I	NB II	Copula	Poisson	PIG	NB I	NB II
Normal	169,585	84,652	82,114	81,679	Normal	104,355	79,279	79,920	79,999
Frank	168,684	84,640	82,102	81,662	Frank	104,643	79,252	79,451	79,500
AMH	169,227	84,652	82,109	81,675	AMH	104,059	79,284	79,775	79,818
FGM	168,542	84,636	82,100	81,656	FGM	103,593	79,291	79,767	79,818
Student-t	168,626	84,743	82,119	81,699	Student-t	104,301	**79,214**	79,601	79,698
Clayton	170,275	84,667	82,122	81,690	Clayton 90	102,417	79,264	79,387	80,047
Clayton 180	169,400	84,384	82,059	81,587	Clayton 270	104,554	79,335	80,012	80,045
Joe	169,267	84,328	82,047	**81,573**	Joe 90	104,553	79,319	80,012	80,047
Joe 180	169,893	84,667	82,122	81,690	Joe 270	102,318	79,275	79,369	80,047
Gumbel	169,786	84,537	82,087	81,641	Gumbel 90	104,554	79,274	79,986	80,047
Gumbel 180	169,901	84,667	82,122	81,690	Gumbel 270	103,408	79,243	79,601	79,689

Note: cloglog in first step

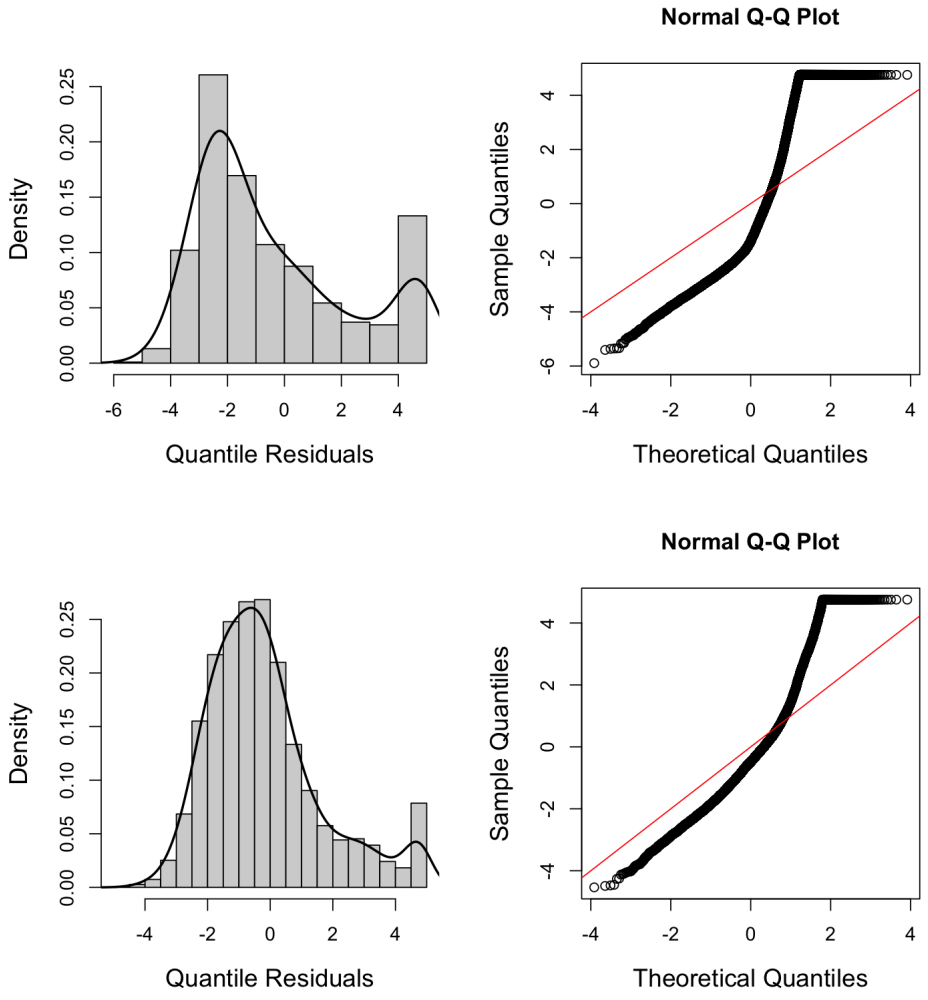

Figure A1 – Q-Q plots of the randomized and normalized residuals of the count regressions with endogenous participation. The first row corresponds to the regression of sickness absenteeism days with a Poisson margin and Gaussian copula, and the second row corresponds to the regression on sickness presenteeism days with a Poisson margin and Gaussian copula.

3 Estimates of the models with different sets of variables

This section presents the results of the models when we remove sets of variables that are likely to be endogenous, such as declared health. Firstly, Table 7 below presents the initial results of the chapter with confidence intervals instead of standard errors. This allows to determine whether the coefficients in all the robustness tables that follow remain within the confidence intervals of the coefficients estimated in the chapter. Then, Table 8 shows the results when health characteristics are excluded, Table 9 shows the results without the working conditions and finally Table 10 presents the results without the employment characteristics.

Table A9 – SA and SP models presented in the chapter with confidence intervals in square brackets

	Sickness absenteeism		Sickness presenteeism	
Equation	SA days	Illness	SP days	Illness
Distribution	(NB II)	(logit)	(PIG)	(logit)
Copula	Joe		Student-t	
Constant	**1.407**	−0.005	**1.167**	−0.044
	[0.687; 2.127]	[−0.724; 0.713]	[0.614; 1.721]	[−0.763; 0.675]
Health (ref: good)				
bad or very bad	**0.587**		**0.445**	
	[0.475; 0.700]		[0.348; 0.542]	
rather good	**0.179**		**0.286**	
	[0.112; 0.245]		[0.234; 0.339]	
very good	**−0.083**		**−0.150**	
	[−0.164; −0.001]		[−0.214; −0.086]	
Chronic disease	**0.199**		**0.063**	
	[0.137; 0.261]		[0.013; 0.113]	
Occupation (ref: employee)				
labourer	0.055	0.059	−0.025	0.076
	[−0.049; 0.159]	[−0.052; 0.169]	[−0.109; 0.058]	[−0.035; 0.186]
intermediate	0.069	−0.003	−0.005	−0.005
	[−0.016; 0.155]	[−0.096; 0.089]	[−0.072; 0.062]	[−0.097; 0.087]
executive	−0.029	−0.080	0.002	−0.064
	[−0.155; 0.096]	[−0.206; 0.046]	[−0.092; 0.096]	[−0.189; 0.061]
Seniority	0.004	**0.022**	−0.006	**0.024**
	[−0.008; 0.015]	[0.011; 0.034]	[−0.015; 0.003]	[0.012; 0.035]
Seniority^2	−0.000	**−0.001**	0.000	**−0.001**
	[−0.000; 0.000]	[−0.001; −0.000]	[−0.000; 0.000]	[−0.001; −0.000]
Work alone	−0.067	**0.091**	0.020	**0.100**
	[−0.137; 0.002]	[0.018; 0.164]	[−0.034; 0.073]	[0.027; 0.173]
Supervisor	**−0.194**	**−0.227**	−0.016	**−0.207**
	[−0.294; −0.094]	[−0.324; −0.131]	[−0.089; 0.058]	[−0.303; −0.111]
Public sector	**0.156**	**0.134**	−0.002	**0.112**
	[0.071; 0.240]	[0.042; 0.225]	[−0.068; 0.064]	[0.021; 0.204]
Temporary contract	**−0.466**	**−0.516**	0.036	**−0.506**
	[−0.621; −0.312]	[−0.652; −0.380]	[−0.077; 0.149]	[−0.641; −0.370]
working time	**0.067**	**0.041**	−0.014	**0.040**
	[0.045; 0.088]	[0.022; 0.060]	[−0.029; 0.000]	[0.021; 0.058]
working time^2	**−0.001**	**−0.001**	**0.000**	**−0.001**

	SA days	Illness	SP days	Illness
	[−0.001; −0.001]	[−0.001; −0.000]	[0.000; 0.000]	[−0.001; −0.000]
Atypical schedules	**−0.101**	**−0.075**	**0.056**	−0.064
	[−0.162; −0.040]	[−0.139; −0.010]	[0.008; 0.105]	[−0.128; 0.000]
Wage (ref:1,800-2,499)				
100-1,199	0.085	−0.023	−0.002	−0.021
	[−0.046; 0.216]	[−0.161; 0.116]	[−0.104; 0.100]	[−0.160; 0.117]
1,200-1,499	0.067	−0.080	−0.012	−0.051
	[−0.026; 0.160]	[−0.182; 0.022]	[−0.086; 0.063]	[−0.153; 0.051]
1,500-1,799	0.066	−0.072	0.051	−0.065
	[−0.019; 0.150]	[−0.166; 0.021]	[−0.017; 0.118]	[−0.158; 0.029]
2,500-2,999	0.028	−0.089	0.076	−0.089
	[−0.094; 0.149]	[−0.211; 0.033]	[−0.016; 0.168]	[−0.210; 0.033]
3,000-20,000	0.011	**−0.154**	−0.001	**−0.155**
	[−0.121; 0.144]	[−0.281; −0.027]	[−0.099; 0.097]	[−0.281; −0.029]
Quantified target	0.063	**0.120**	**0.086**	**0.107**
	[−0.006; 0.132]	[0.045; 0.196]	[0.032; 0.140]	[0.031; 0.183]
Work rhythm paced by:				
machine	0.067	**0.192**	0.014	**0.182**
	[−0.025; 0.159]	[0.081; 0.304]	[−0.061; 0.089]	[0.070; 0.294]
external demand	**−0.136**	−0.018	**0.056**	−0.009
	[−0.204; −0.068]	[−0.089; 0.053]	[0.002; 0.111]	[−0.080; 0.062]
norms	0.016	0.064	0.007	**0.077**
	[−0.049; 0.080]	[−0.005; 0.132]	[−0.043; 0.058]	[0.009; 0.145]
computing control	0.010	0.034	0.003	0.037
	[−0.054; 0.074]	[−0.036; 0.103]	[−0.047; 0.054]	[−0.032; 0.107]
monitoring	0.047	**0.151**	0.041	**0.142**
	[−0.020; 0.113]	[0.075; 0.226]	[−0.012; 0.093]	[0.066; 0.217]
colleagues	−0.009	**0.104**	0.015	**0.100**
	[−0.075; 0.056]	[0.031; 0.178]	[−0.037; 0.067]	[0.026; 0.173]
Help from Supervisors	−0.040	**−0.268**	**−0.075**	**−0.261**
	[−0.101; 0.022]	[−0.336; −0.201]	[−0.124; −0.026]	[−0.329; −0.193]
Help from colleagues	0.089	0.090	**−0.126**	**0.095**
	[−0.000; 0.179]	[−0.001; 0.182]	[−0.193; −0.060]	[0.003; 0.186]
Tensions at work	**0.132**	**0.473**	**0.111**	**0.473**
	[0.065; 0.199]	[0.408; 0.538]	[0.057; 0.165]	[0.408; 0.537]
Lack of time	0.006	**0.270**	**0.099**	**0.269**
	[−0.060; 0.072]	[0.193; 0.346]	[0.048; 0.151]	[0.193; 0.345]
Under pressure	0.013	**0.324**	**0.144**	**0.309**
	[−0.054; 0.081]	[0.247; 0.400]	[0.091; 0.198]	[0.232; 0.385]
Choice of methods	**−0.095**	−0.061	0.023	−0.046
	[−0.167; −0.024]	[−0.142; 0.020]	[−0.035; 0.081]	[−0.128; 0.036]
Initiative	−0.034	0.044	**0.094**	0.042
	[−0.095; 0.027]	[−0.020; 0.109]	[0.047; 0.142]	[−0.023; 0.106]
Work-to-life conflicts	−0.031	**0.317**	0.036	**0.321**
	[−0.107; 0.045]	[0.227; 0.407]	[−0.022; 0.095]	[0.231; 0.410]
Fear of job future	−0.062	**0.223**	**0.092**	**0.219**
	[−0.136; 0.011]	[0.140; 0.306]	[0.036; 0.149]	[0.136; 0.302]
Organizational changes	**0.131**	**0.270**	0.022	**0.263**
	[0.070; 0.191]	[0.206; 0.334]	[−0.026; 0.070]	[0.199; 0.327]
Family conflicts		**0.361**		**0.405**
		[0.292; 0.429]		[0.337; 0.474]
sigma	21.123	21.123	1.698	1.698
theta	1.820	1.820	−0.511	−0.511
tau	0.312	0.312	−0.341	−0.341
AIC	80, 348	80, 348	77, 958	77, 958
Log Likelihood	−40032.815	−40032.815	−38838.129	−38838.129
Num. obs.	11073	19285	11073	19285

Table A10 – Estimates of the effects of employment characteristics and working conditions on SA and SP days excluding health characteristics

	Sickness absenteeism		Sickness presenteeism	
Equation	SA days	Illness	SP days	Illness
Distribution	(NB II)	(logit)	(PIG)	(logit)
Copula	Joe		Student-t	
Constant	**1.289*****	−0.021	**1.046*****	−0.055
	(0.366)	(0.366)	(0.285)	(0.367)
Occupation (ref: employee)				
labourer	0.069	0.060	−0.000	0.076
	(0.053)	(0.056)	(0.043)	(0.056)
intermediate	0.067	−0.006	−0.004	−0.001
	(0.044)	(0.047)	(0.035)	(0.047)
executive	−0.027	−0.083	0.009	−0.059
	(0.064)	(0.064)	(0.049)	(0.064)
Supervisor	**−0.202*****	**−0.226*****	−0.033	**−0.209*****
	(0.051)	(0.049)	(0.038)	(0.049)
Public sector	**0.172*****	**0.136*****	0.010	**0.110****
	(0.043)	(0.047)	(0.034)	(0.047)
Seniority	0.004	**0.022*****	−0.006	**0.024*****
	(0.006)	(0.006)	(0.005)	(0.006)
Seniority^2	−0.000	**−0.001*****	0.000	**−0.001*****
	(0.000)	(0.000)	(0.000)	(0.000)
Temporary contract	**−0.486*****	**−0.519*****	0.029	**−0.505*****
	(0.079)	(0.069)	(0.059)	(0.069)
Working time	**0.065*****	**0.042*****	**−0.017****	**0.040*****
	(0.011)	(0.010)	(0.007)	(0.009)
Working time^2	**−0.001*****	**−0.001*****	**0.000*****	**−0.001*****
	(0.000)	(0.000)	(0.000)	(0.000)
Atypical schedules	**−0.107*****	**−0.077****	**0.057****	−0.063*
	(0.031)	(0.033)	(0.025)	(0.033)
Work alone	−0.066*	**0.091****	0.020	**0.099*****
	(0.035)	(0.037)	(0.028)	(0.037)
Wage (ref:1,800-2,499)				
100-1,199	0.116*	−0.024	0.028	−0.019
	(0.067)	(0.071)	(0.053)	(0.071)
1,200-1,499	0.086*	−0.080	0.009	−0.049
	(0.048)	(0.052)	(0.038)	(0.052)
1,500-1,799	0.075*	−0.075	0.061*	−0.060
	(0.043)	(0.048)	(0.035)	(0.048)
2,500-2,999	0.025	−0.088	0.080*	−0.090
	(0.062)	(0.062)	(0.048)	(0.062)
3,000-20,000	−0.013	**−0.156****	−0.015	**−0.155****
	(0.068)	(0.065)	(0.051)	(0.064)
Quantified target	**0.075****	**0.121*****	**0.100*****	**0.107*****
	(0.035)	(0.039)	(0.028)	(0.039)
Work rhythm paced by:				
machine	0.071	**0.193*****	0.011	**0.181*****
	(0.047)	(0.057)	(0.039)	(0.057)
external demand	**−0.132*****	−0.018	**0.056****	−0.008
	(0.035)	(0.036)	(0.028)	(0.036)
norms	0.021	0.064*	0.015	**0.074****
	(0.033)	(0.035)	(0.026)	(0.035)
computing control	0.005	0.034	−0.002	0.036
	(0.033)	(0.036)	(0.026)	(0.036)

	SA days	Illness	SP days	Illness
monitoring	0.057*	**0.151*****	**0.054****	**0.142*****
	(0.034)	(0.039)	(0.027)	(0.039)
colleagues relation	−0.000	**0.104*****	0.020	**0.104*****
	(0.034)	(0.038)	(0.027)	(0.038)
Help from supervisors	−0.049	**−0.269*****	**−0.085*****	**−0.261*****
	(0.031)	(0.035)	(0.025)	(0.035)
Help from colleagues	0.067	0.091*	**−0.146*****	**0.096****
	(0.046)	(0.047)	(0.034)	(0.047)
Tensions at work	**0.151*****	**0.473*****	**0.133*****	**0.470*****
	(0.034)	(0.033)	(0.028)	(0.033)
Choice of methods	**−0.103*****	−0.062	0.017	−0.047
	(0.037)	(0.041)	(0.030)	(0.042)
Lack of time	0.024	**0.268*****	**0.126*****	**0.267*****
	(0.034)	(0.039)	(0.027)	(0.039)
Under pressure	0.050	**0.327*****	**0.173*****	**0.310*****
	(0.034)	(0.039)	(0.028)	(0.039)
Initiative	−0.033	0.047	**0.086*****	0.043
	(0.031)	(0.033)	(0.025)	(0.033)
Work-to-life conflict	0.009	**0.317*****	**0.071****	**0.323*****
	(0.039)	(0.046)	(0.030)	(0.046)
Fear of job future	−0.024	**0.223*****	**0.121*****	**0.220*****
	(0.038)	(0.042)	(0.029)	(0.042)
Organizational changes	**0.148*****	**0.270*****	0.038	**0.262*****
	(0.031)	(0.033)	(0.025)	(0.033)
Family conflicts		**0.350*****		**0.418*****
		(0.035)		(0.035)
sigma	21.562	21.562	1.814	1.814
theta	1.790	1.790	−0.514	−0.514
tau	0.305	0.305	−0.344	−0.344
AIC	80561.374	80561.374	78220.290	78220.290
Log Likelihood	−40143.687	−40143.687	−38973.145	−38973.145
Num. obs.	11073	19285	11073	19285

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

We control for individual's characteristics (age, gender, marital status, having children and education level), firm sectors and firm size.

Table A11 – Estimates of the effects of employment characteristics on SA and SP days excluding working conditions

	Sickness absenteeism		Sickness presenteeism	
Equation	SA days	Illness	SP days	Illness
Distribution	(NB II)	(logit)	(PIG)	(logit)
Copula	Joe		Student-t	
Constant	**1.359*****	0.029	**1.277*****	0.027
	(0.365)	(0.354)	(0.285)	(0.352)
Occupation (ref: employee)				
labourer	0.093*	**0.126****	−0.041	**0.133****
	(0.052)	(0.054)	(0.042)	(0.053)
intermediate	0.074*	0.062	0.018	0.056
	(0.044)	(0.046)	(0.035)	(0.045)
executive	−0.040	−0.055	0.036	−0.055
	(0.063)	(0.062)	(0.048)	(0.062)

	SA days	Illness	SP days	Illness
Supervisor	**−0.206*****	**−0.148*****	0.028	**−0.138*****
	(0.051)	(0.047)	(0.038)	(0.047)
Public sector	**0.166*****	**0.091****	−0.030	0.077*
	(0.043)	(0.046)	(0.034)	(0.045)
Seniority	0.006	**0.029*****	−0.007	**0.031*****
	(0.006)	(0.006)	(0.005)	(0.006)
Seniority^2	−0.000	**−0.001*****	0.000	**−0.001*****
	(0.000)	(0.000)	(0.000)	(0.000)
Temporary contract	**−0.490*****	**−0.528*****	0.062	**−0.519*****
	(0.078)	(0.066)	(0.058)	(0.066)
Working time	**0.069*****	**0.036*****	**−0.018****	**0.034*****
	(0.011)	(0.009)	(0.007)	(0.009)
Working time^2	**−0.001*****	**−0.000*****	**0.000*****	**−0.000*****
	(0.000)	(0.000)	(0.000)	(0.000)
Atypical schedules	**−0.086*****	**0.100*****	**0.104*****	**0.103*****
	(0.030)	(0.031)	(0.024)	(0.031)
Work alone	**−0.097*****	0.052	0.050*	0.056
	(0.034)	(0.035)	(0.026)	(0.035)
Wage (ref:1,800-2,499)				
100-1,199	0.095	−0.017	−0.006	−0.013
	(0.067)	(0.069)	(0.053)	(0.069)
1,200-1,499	0.082*	−0.054	−0.018	−0.038
	(0.047)	(0.051)	(0.038)	(0.050)
1,500-1,799	0.073*	−0.050	0.055	−0.045
	(0.043)	(0.046)	(0.035)	(0.046)
2,500-2,999	0.028	−0.119*	0.072	**−0.122****
	(0.062)	(0.061)	(0.048)	(0.061)
3,000-20,000	0.000	**−0.224*****	0.004	**−0.220*****
	(0.068)	(0.063)	(0.051)	(0.063)
Quantified target	**0.091*****	**0.320*****	**0.129*****	**0.307*****
	(0.034)	(0.036)	(0.028)	(0.036)
Family conflicts		**0.509*****		**0.564*****
		(0.034)		(0.034)
sigma	20.723	20.723	1.891	1.891
theta	1.685	1.685	−0.539	−0.539
tau	0.276	0.276	−0.362	−0.362
AIC	81441.650	81441.650	79495.843	79495.843
Log Likelihood	−40611.825	−40611.825	−39638.922	−39638.922
Num. obs.	11073	19285	11073	19285

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

We control for individual's characteristics (age, gender, marital status, having children and education level), firm sectors and firm size.

Table A12 – Estimates of the effects of working conditions on SA and SP days excluding employment characteristics

	Sickness absenteeism		Sickness presenteeism	
Equation	SA days	Illness	SP days	Illness
Distribution	(NB II)	(logit)	(PIG)	(logit)
Copula	Joe		Student-t	
Constant	**2.152*****	−0.023	**1.130*****	−0.066
	(0.269)	(0.281)	(0.220)	(0.283)
Work rhythm paced by:				
machine	0.087*	**0.209*****	0.014	**0.206*****
	(0.047)	(0.056)	(0.038)	(0.056)
external demand	**−0.153*****	−0.034	**0.065****	−0.024
	(0.035)	(0.036)	(0.028)	(0.036)
norms	0.020	**0.068****	0.009	**0.083****
	(0.033)	(0.035)	(0.025)	(0.034)
computing control	0.023	0.062*	0.017	0.065*
	(0.032)	(0.035)	(0.025)	(0.035)
monitoring	**0.071****	**0.180*****	0.045*	**0.165*****
	(0.034)	(0.038)	(0.027)	(0.038)
colleagues relation	−0.002	**0.092****	0.011	**0.087****
	(0.033)	(0.037)	(0.026)	(0.037)
Help from supervisors	−0.044	**−0.276*****	**−0.075*****	**−0.270*****
	(0.031)	(0.034)	(0.025)	(0.034)
Help from colleagues	**0.131*****	0.073	**−0.134*****	0.079*
	(0.044)	(0.045)	(0.033)	(0.045)
Tensions at work	**0.149*****	**0.491*****	**0.116*****	**0.487*****
	(0.034)	(0.033)	(0.028)	(0.033)
Choice of methods	**−0.089****	−0.051	0.018	−0.036
	(0.037)	(0.041)	(0.030)	(0.041)
Initiative	**−0.091*****	0.010	**0.111*****	0.009
	(0.030)	(0.032)	(0.023)	(0.032)
Lack of time	−0.006	**0.265*****	**0.109*****	**0.266*****
	(0.034)	(0.038)	(0.026)	(0.038)
Under pressure	−0.010	**0.312*****	**0.158*****	**0.297*****
	(0.034)	(0.039)	(0.027)	(0.038)
Work-to-life conflict	**−0.097****	**0.269*****	**0.065****	**0.277*****
	(0.038)	(0.044)	(0.029)	(0.044)
Fear of job future	**−0.098*****	**0.155*****	**0.095*****	**0.150*****
	(0.037)	(0.041)	(0.028)	(0.041)
Organizational changes	**0.153*****	**0.294*****	0.022	**0.286*****
	(0.031)	(0.032)	(0.024)	(0.032)
Family conflicts		**0.359*****		**0.406*****
		(0.035)		(0.035)
sigma	21.423	21.423	1.692	1.692
theta	1.857	1.857	−0.502	−0.502
tau	0.322	0.322	−0.335	−0.335
AIC	80607.195	80607.195	78122.436	78122.436
Log Likelihood	−40202.598	−40202.598	−38960.218	−38960.218
Num. obs.	11073	19285	11073	19285

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

We control for individual's characteristics (age, gender, marital status, having children and education level).

4 Estimates of models with various restrictions of the sample of study

In this section, we present different regressions in which we relax some of the restrictions imposed on our sample. Then, Table 8 provides the results without the restriction on the age of the individuals and Table 9 presents the results without the restriction on the declared working time. Tables 10 and 11 give the results of the regressions when we change the total duration of the illness under consideration: either to 30, 120 or 365 days.

Table A13: SA and SP models without restriction on age

	Sickness absenteeism		Sickness presenteeism	
Equation	SA days	Illness	SP days	Illness
Distribution	(NB II)	(logit)	(PIG)	(logit)
Copula	Joe		Student-t	
Constant	**1.285*****	−0.266	**1.241*****	−0.286
	(0.364)	(0.338)	(0.281)	(0.338)
Health (ref: good)				
bad or very bad	**0.588*****		**0.446*****	
	(0.057)		(0.049)	
rather good	**0.180*****		**0.285*****	
	(0.034)		(0.027)	
very good	**−0.084****		**−0.149*****	
	(0.042)		(0.033)	
Chronic disease	**0.197*****		**0.064****	
	(0.032)		(0.025)	
Occupation (ref: employee)				
labourer	0.055	0.058	−0.026	0.074
	(0.053)	(0.056)	(0.043)	(0.056)
intermediate	0.069	−0.007	−0.006	−0.008
	(0.044)	(0.047)	(0.034)	(0.047)
executive	−0.033	−0.085	0.002	−0.068
	(0.064)	(0.064)	(0.048)	(0.064)
Seniority	0.004	**0.021*****	−0.006	**0.023*****
	(0.006)	(0.006)	(0.004)	(0.006)
Seniority^2	−0.000	**−0.000*****	0.000	**−0.000*****
	(0.000)	(0.000)	(0.000)	(0.000)
Work alone	−0.066*	**0.091****	0.019	**0.100*****
	(0.035)	(0.037)	(0.027)	(0.037)
Supervisor	**−0.195*****	**−0.226*****	−0.015	**−0.206*****
	(0.051)	(0.049)	(0.037)	(0.049)
Public	**0.158*****	**0.138*****	−0.003	**0.117****
	(0.043)	(0.047)	(0.034)	(0.047)
Temporary contract	**−0.467*****	**−0.520*****	0.036	**−0.509*****
	(0.079)	(0.069)	(0.058)	(0.069)

	SA days	Illness	SP days	Illness
Working time	**0.067*****	**0.041*****	**−0.014****	**0.040*****
	(0.011)	(0.009)	(0.007)	(0.009)
Working time^2	**−0.001*****	**−0.001*****	**0.000****	**−0.001*****
	(0.000)	(0.000)	(0.000)	(0.000)
Atypical schedules	**−0.101*****	**−0.077****	**0.056****	**−0.066****
	(0.031)	(0.033)	(0.025)	(0.033)
Wage (ref:1,800-2,499)				
100-1,199	0.085	−0.021	−0.003	−0.019
	(0.067)	(0.070)	(0.052)	(0.071)
1,200-1,499	0.066	−0.077	−0.011	−0.049
	(0.047)	(0.052)	(0.038)	(0.052)
1,500-1,799	0.066	−0.073	0.050	−0.065
	(0.043)	(0.048)	(0.034)	(0.048)
2,500-2,999	0.029	−0.090	0.076	−0.090
	(0.062)	(0.062)	(0.047)	(0.062)
3,000-20,000	0.012	**−0.148****	−0.001	**−0.150****
	(0.068)	(0.065)	(0.050)	(0.064)
Quantified target	0.062*	**0.119*****	**0.087*****	**0.107*****
	(0.035)	(0.039)	(0.028)	(0.039)
Work rhythm paced by:				
machine	0.066	**0.194*****	0.014	**0.183*****
	(0.047)	(0.057)	(0.038)	(0.057)
external demand	**−0.135*****	−0.016	**0.055****	−0.007
	(0.035)	(0.036)	(0.028)	(0.036)
norms	0.017	0.065*	0.006	**0.079****
	(0.033)	(0.035)	(0.026)	(0.035)
computing control	0.009	0.033	0.004	0.037
	(0.033)	(0.035)	(0.026)	(0.035)
monitoring	0.046	**0.153*****	0.040	**0.144*****
	(0.034)	(0.039)	(0.027)	(0.039)
colleagues	−0.010	**0.107*****	0.016	**0.102*****
	(0.034)	(0.037)	(0.026)	(0.037)
Help from Supervisors	−0.038	**−0.267*****	**−0.075*****	**−0.259*****
	(0.031)	(0.035)	(0.025)	(0.035)
Help from colleagues	**0.091****	**0.098****	**−0.127*****	**0.102****
	(0.046)	(0.047)	(0.034)	(0.047)
Tensions at work	**0.131*****	**0.475*****	**0.111*****	**0.474*****
	(0.034)	(0.033)	(0.028)	(0.033)
Lack of time	0.005	**0.268*****	**0.100*****	**0.268*****
	(0.034)	(0.039)	(0.026)	(0.039)
Under pressure	0.014	**0.321*****	**0.143*****	**0.307*****
	(0.034)	(0.039)	(0.027)	(0.039)
Choice of method	**−0.094****	−0.061	0.023	−0.046
	(0.037)	(0.041)	(0.030)	(0.042)
Initiative	−0.035	0.043	**0.094*****	0.041
	(0.031)	(0.033)	(0.024)	(0.033)
Work-to-life conflicts	−0.033	**0.318*****	0.037	**0.321*****
	(0.039)	(0.046)	(0.030)	(0.046)
Fear of job future	−0.061	**0.222*****	**0.092*****	**0.219*****
	(0.038)	(0.042)	(0.029)	(0.042)

	SA days	Illness	SP days	Illness
Organizational changes	**0.130*****	**0.268*****	0.022	**0.261*****
	(0.031)	(0.033)	(0.024)	(0.033)
Family conflicts		**0.359*****		**0.404*****
		(0.035)		(0.035)
sigma	21.110	21.110	1.695	1.695
theta	1.814	1.814	−0.511	−0.511
tau	0.311	0.311	−0.342	−0.342
AIC	80457.202	80457.202	78061.854	78061.854
Log Likelihood	−40087.601	−40087.601	−38889.927	−38889.927
Num. obs.	11084	19334	11084	19334

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$
We control for individual characteristics, firm sectors and firm size.

Table A14 – SA and SP models without restriction on working time

	Sickness absenteeism		Sickness presenteeism	
Equation	SA days	Illness	SP days	Illness
Distribution	(NB II)	(logit)	(PIG)	(logit)
Copula	Joe		Student-t	
Constant	**2.113*****	0.572*	**0.808*****	0.503
	(0.334)	(0.327)	(0.250)	(0.329)
Health (ref: good)				
bad or very bad	**0.585*****		**0.443*****	
	(0.057)		(0.049)	
rather good	**0.177*****		**0.284*****	
	(0.034)		(0.027)	
very good	**−0.086****		**−0.152*****	
	(0.042)		(0.032)	
Chronic disease	**0.203*****		**0.064****	
	(0.032)		(0.025)	
Occupation (ref: employee)				
labourer	0.057	0.072	−0.027	0.090
	(0.053)	(0.056)	(0.042)	(0.056)
intermediate	0.059	−0.015	−0.003	−0.016
	(0.044)	(0.047)	(0.034)	(0.047)
executive	−0.063	−0.111*	0.016	−0.094
	(0.064)	(0.064)	(0.047)	(0.063)
Supervisor	**−0.207*****	**−0.244*****	−0.011	**−0.221*****
	(0.051)	(0.049)	(0.037)	(0.049)
Public	**0.160*****	**0.137*****	−0.005	**0.117****
	(0.043)	(0.047)	(0.034)	(0.046)
Seniority	0.003	**0.022*****	−0.006	**0.023*****
	(0.006)	(0.006)	(0.004)	(0.006)
Seniority^2	−0.000	**−0.001*****	0.000	**−0.001*****
	(0.000)	(0.000)	(0.000)	(0.000)
Temporary contract	**−0.471*****	**−0.511*****	0.033	**−0.502*****

	SA days	Illness	SP days	Illness
	(0.078)	(0.068)	(0.057)	(0.068)
Working time	**0.026*****	0.006	0.003	0.006
	(0.008)	(0.005)	(0.003)	(0.005)
Working time^2	**−0.000*****	−0.000	0.000	−0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Atypical schedules	**−0.107*****	**−0.080****	**0.058****	**−0.070****
	(0.031)	(0.033)	(0.025)	(0.033)
Work alone	**−0.074****	**0.085****	0.023	**0.093****
	(0.035)	(0.037)	(0.027)	(0.037)
Wage (ref:1,800-2,499)				
100-1,199	0.026	−0.107	0.037	−0.103
	(0.066)	(0.068)	(0.050)	(0.068)
1,200-1,499	0.069	−0.082	−0.003	−0.054
	(0.047)	(0.052)	(0.038)	(0.052)
1,500-1,799	0.067	−0.072	0.050	−0.065
	(0.043)	(0.047)	(0.034)	(0.047)
2,500-2,999	0.017	−0.090	0.076	−0.088
	(0.062)	(0.062)	(0.047)	(0.062)
3,000-20,000	−0.021	**−0.187*****	0.009	**−0.186*****
	(0.068)	(0.064)	(0.050)	(0.064)
Quantified target	0.056	**0.116*****	**0.087*****	**0.102*****
	(0.035)	(0.038)	(0.027)	(0.039)
Work rhythm paced by:				
machine	0.069	**0.196*****	0.011	**0.184*****
	(0.047)	(0.057)	(0.038)	(0.057)
demand	**−0.135*****	−0.013	0.054*	−0.003
	(0.035)	(0.036)	(0.028)	(0.036)
norms	0.017	0.066*	0.010	**0.078****
	(0.033)	(0.035)	(0.025)	(0.035)
computing control	0.007	0.035	0.003	0.039
	(0.033)	(0.035)	(0.026)	(0.035)
monitoring	0.053	**0.152*****	0.041	**0.144*****
	(0.034)	(0.038)	(0.027)	(0.038)
colleagues	−0.003	**0.102*****	0.014	**0.098*****
	(0.034)	(0.037)	(0.026)	(0.037)
Help from Supervisor	−0.037	**−0.265*****	**−0.076*****	**−0.258*****
	(0.031)	(0.034)	(0.025)	(0.034)
Help from colleagues	**0.090****	**0.095****	**−0.122*****	**0.100****
	(0.046)	(0.046)	(0.034)	(0.046)
Tensions at work	**0.141*****	**0.477*****	**0.106*****	**0.476*****
	(0.034)	(0.033)	(0.027)	(0.033)
Lack of time	−0.000	**0.265*****	**0.100*****	**0.265*****
	(0.034)	(0.039)	(0.026)	(0.039)
Under pressure	0.008	**0.309*****	**0.146*****	**0.292*****
	(0.034)	(0.039)	(0.027)	(0.039)
Choice of methods	**−0.089****	−0.061	0.023	−0.047
	(0.037)	(0.041)	(0.029)	(0.041)
Initiative	−0.037	0.039	**0.096*****	0.037
	(0.031)	(0.033)	(0.024)	(0.033)
Work-to-life conflicts	−0.040	**0.305*****	0.042	**0.310*****

	SA days	Illness	SP days	Illness
	(0.038)	(0.045)	(0.030)	(0.045)
Fear of job future	−0.067*	**0.225*****	**0.094*****	**0.222*****
	(0.037)	(0.042)	(0.029)	(0.042)
Organizational changes	**0.130*****	**0.269*****	0.023	**0.262*****
	(0.031)	(0.033)	(0.024)	(0.033)
Family conflicts		**0.361*****		**0.407*****
		(0.035)		(0.035)
sigma	21.216	21.216	1.695	1.695
theta	1.818	1.818	−0.514	−0.514
tau	0.312	0.312	−0.343	−0.343
AIC	81109.588	81109.588	78756.515	78756.515
Log Likelihood	−40413.794	−40413.794	−39237.258	−39237.258
Num. obs.	11181	19510	11181	19510

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$
We control for individual characteristics, firm sectors and firm size.

Table A15 – SP days with different restrictions on the illness duration

	Sickness presenteeism					
	30 days of illness		120 days of illness		365 days of illness	
Equation	SP days	Illness	SP days	Illness	SP days	Illness
Distribution	(PIG)	(logit)	(PIG)	(logit)	(PIG)	(logit)
Copula	Student-t		Student-t		Student-t	
Constant	**0.788*****	−0.164	**1.408*****	−0.003	**1.887*****	0.110
	(0.269)	(0.373)	(0.294)	(0.362)	(0.314)	(0.357)
Health (ref: good)						
bad or very bad	**0.359*****		**0.522*****		**0.527*****	
	(0.051)		(0.049)		(0.048)	
rather good	**0.256*****		**0.309*****		**0.298*****	
	(0.026)		(0.028)		(0.029)	
very good	**−0.150*****		**−0.143*****		**−0.144*****	
	(0.031)		(0.034)		(0.035)	
Chronic disease	0.045*		**0.074*****		**0.053****	
	(0.025)		(0.026)		(0.027)	
Occupation (ref: employee)						
labourer	−0.039	0.080	−0.005	0.076	0.009	0.064
	(0.041)	(0.057)	(0.044)	(0.055)	(0.047)	(0.055)
intermediate	−0.027	−0.003	−0.001	−0.010	−0.022	−0.002
	(0.033)	(0.048)	(0.036)	(0.046)	(0.039)	(0.046)
executive	−0.007	−0.070	0.015	−0.080	0.036	−0.086
	(0.046)	(0.065)	(0.051)	(0.063)	(0.054)	(0.062)
Seniority	−0.005	**0.024*****	−0.007	**0.023*****	−0.007	**0.022*****
	(0.004)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)
Seniority^2	0.000	**−0.001*****	0.000*	**−0.001*****	**0.000****	**−0.001*****
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Work alone	0.019	**0.104*****	0.018	**0.101*****	0.031	**0.088****
	(0.026)	(0.038)	(0.028)	(0.037)	(0.030)	(0.036)
Supervisor	−0.015	**−0.210*****	−0.017	**−0.205*****	−0.013	**−0.195*****
	(0.035)	(0.050)	(0.039)	(0.048)	(0.041)	(0.048)
Public	−0.004	**0.131*****	0.031	**0.097****	0.007	**0.102****
	(0.032)	(0.048)	(0.035)	(0.046)	(0.038)	(0.045)
Temporary contract	0.037	**−0.481*****	0.063	**−0.526*****	**0.133****	**−0.543*****
	(0.054)	(0.070)	(0.060)	(0.068)	(0.063)	(0.068)

	SP days	Illness	SP days	Illness	SP days	Illness
Working time	**−0.017****	**0.040*****	**−0.022*****	**0.041*****	**−0.022*****	**0.039*****
	(0.007)	(0.010)	(0.007)	(0.009)	(0.008)	(0.009)
Working time^2	**0.000*****	**−0.000*****	**0.000*****	**−0.001*****	**0.000*****	**−0.001*****
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Atypical schedules	**0.059****	−0.058*	**0.056****	**−0.067****	**0.061****	**−0.070****
	(0.023)	(0.033)	(0.026)	(0.032)	(0.027)	(0.032)
Wage (ref:1,800-2,499)						
100-1,199	−0.002	−0.039	0.011	−0.008	−0.008	0.026
	(0.050)	(0.072)	(0.054)	(0.070)	(0.057)	(0.068)
1,200-1,499	−0.018	−0.060	−0.028	−0.042	−0.028	−0.017
	(0.036)	(0.053)	(0.040)	(0.051)	(0.042)	(0.050)
1,500-1,799	0.023	−0.084*	0.048	−0.056	0.049	−0.041
	(0.033)	(0.049)	(0.036)	(0.047)	(0.039)	(0.046)
2,500-2,999	0.078*	−0.070	0.092*	−0.095	**0.107****	−0.090
	(0.044)	(0.063)	(0.049)	(0.061)	(0.053)	(0.060)
3,000-20,000	−0.005	**−0.145****	0.012	**−0.166*****	0.034	**−0.187*****
	(0.047)	(0.065)	(0.053)	(0.064)	(0.056)	(0.063)
Quantified target	**0.071*****	**0.113*****	**0.072****	**0.108*****	**0.069****	**0.111*****
	(0.026)	(0.039)	(0.029)	(0.038)	(0.031)	(0.038)
Work rhythm paced by:						
machine	0.016	**0.183*****	0.017	**0.183*****	−0.011	**0.193*****
	(0.037)	(0.058)	(0.040)	(0.056)	(0.043)	(0.055)
external demand	0.037	−0.001	**0.075*****	−0.026	**0.092*****	−0.043
	(0.027)	(0.037)	(0.029)	(0.036)	(0.031)	(0.035)
norms	0.018	**0.074****	0.011	**0.078****	0.002	**0.078****
	(0.024)	(0.036)	(0.027)	(0.034)	(0.028)	(0.034)
computing control	0.023	0.033	0.003	0.031	−0.006	0.033
	(0.024)	(0.036)	(0.027)	(0.035)	(0.029)	(0.035)
monitoring	0.042	**0.123*****	0.043	**0.149*****	0.038	**0.145*****
	(0.026)	(0.039)	(0.028)	(0.038)	(0.030)	(0.037)
colleagues	−0.000	**0.091****	0.023	**0.096*****	0.004	**0.099*****
	(0.025)	(0.038)	(0.028)	(0.037)	(0.030)	(0.036)
help form Supervisors	**−0.063*****	**−0.249*****	**−0.063****	**−0.263*****	−0.045	**−0.268*****
	(0.024)	(0.035)	(0.026)	(0.034)	(0.028)	(0.034)
Help form colleagues	**−0.115*****	**0.103****	**−0.135*****	**0.093****	**−0.144*****	0.080*
	(0.033)	(0.048)	(0.035)	(0.046)	(0.038)	(0.045)
Tensions at work	**0.141*****	**0.470*****	**0.111*****	**0.466*****	**0.085*****	**0.454*****
	(0.026)	(0.034)	(0.029)	(0.033)	(0.030)	(0.032)
Lack of time	**0.106*****	**0.273*****	**0.108*****	**0.264*****	**0.106*****	**0.251*****
	(0.025)	(0.039)	(0.028)	(0.038)	(0.030)	(0.038)
Under pressure	**0.121*****	**0.307*****	**0.147*****	**0.319*****	**0.111*****	**0.328*****
	(0.026)	(0.040)	(0.029)	(0.039)	(0.031)	(0.038)
Choice of methods	0.020	−0.031	0.016	−0.058	0.018	−0.069*
	(0.028)	(0.043)	(0.031)	(0.041)	(0.033)	(0.040)
Initiative	**0.093*****	0.039	**0.094*****	0.046	**0.091*****	0.042
	(0.023)	(0.034)	(0.025)	(0.033)	(0.027)	(0.032)
Work-to-life conflicts	0.031	**0.327*****	0.007	**0.323*****	−0.010	**0.322*****
	(0.028)	(0.046)	(0.031)	(0.045)	(0.034)	(0.044)
Fear of job future	**0.088*****	**0.217*****	**0.098*****	**0.214*****	**0.077****	**0.231*****
	(0.028)	(0.043)	(0.030)	(0.042)	(0.032)	(0.041)
Organizational changes	0.038*	**0.261*****	0.023	**0.268*****	0.017	**0.262*****
	(0.023)	(0.033)	(0.026)	(0.032)	(0.027)	(0.032)
Family conflicts		**0.409*****		**0.399*****		**0.378*****
		(0.036)		(0.034)		(0.032)
sigma	1.113	1.113	2.400	2.400	4.349	4.349
theta	−0.391	−0.391	−0.585	−0.585	−0.726	−0.726
tau	−0.255	−0.255	−0.398	−0.398	−0.517	−0.517
AIC	70878.028	70878.028	82955.541	82955.541	86804.850	86804.850
Log Likelihood	−35298.014	−35298.014	−41336.771	−41336.771	−43261.425	−43261.425
Num. obs.	10213	18425	11692	19904	12131	20343

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$
We control for individual characteristics, firm sectors and firm size.

Table A16 – SA days with different restrictions on the illness duration

	Sickness absenteeism					
	30 days of illness		120 days of illness		365 days of illness	
Equation	SA days	Illness	SA days	Illness	SA days	Illness
Distribution	(NB II)	(logit)	(NB II)	(logit)	(NB II)	(logit)
Copula	Joe		Joe		Joe	
Constant	0.747*	−0.157	**1.754*****	0.017	**2.131*****	0.005
	(0.396)	(0.374)	(0.349)	(0.360)	(0.327)	(0.355)
Health (ref: good)						
bad or very bad	**0.506*****		**0.699*****		**0.829*****	
	(0.066)		(0.052)		(0.046)	
rather good	**0.167*****		**0.190*****		**0.183*****	
	(0.037)		(0.032)		(0.029)	
very good	−0.073*		**−0.087****		−0.067*	
	(0.044)		(0.039)		(0.036)	
Chronic disease	**0.178*****		**0.218*****		**0.243*****	
	(0.034)		(0.030)		(0.027)	
Occupation (ref: employee)						
labourer	0.056	0.052	0.073	0.068	0.066	0.072
	(0.057)	(0.058)	(0.051)	(0.055)	(0.048)	(0.054)
intermediate	0.065	−0.011	0.044	−0.012	0.040	−0.016
	(0.047)	(0.048)	(0.042)	(0.046)	(0.040)	(0.046)
executive	−0.051	−0.091	−0.091	−0.106*	−0.085	−0.110*
	(0.068)	(0.065)	(0.062)	(0.063)	(0.058)	(0.063)
Seniority	0.001	**0.022*****	0.006	**0.023*****	0.007	**0.023*****
	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.006)
Seniority^2	−0.000	**−0.001*****	−0.000	**−0.001*****	−0.000*	**−0.001*****
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Work alone	**−0.089****	**0.090****	−0.041	**0.099*****	−0.033	**0.102*****
	(0.038)	(0.038)	(0.034)	(0.037)	(0.032)	(0.036)
Supervisor	**−0.217*****	**−0.229*****	**−0.178*****	**−0.220*****	**−0.155*****	**−0.202*****
	(0.055)	(0.050)	(0.049)	(0.048)	(0.046)	(0.048)
Public	**0.196*****	**0.137*****	**0.125*****	**0.123*****	**0.126*****	**0.117*****
	(0.046)	(0.048)	(0.041)	(0.046)	(0.039)	(0.045)
Temporary contract	**−0.421*****	**−0.486*****	**−0.501*****	**−0.533*****	**−0.531*****	**−0.556*****
	(0.083)	(0.071)	(0.075)	(0.068)	(0.070)	(0.068)
Working time	**0.079*****	**0.041*****	**0.060*****	**0.040*****	**0.056*****	**0.038*****
	(0.012)	(0.010)	(0.010)	(0.009)	(0.010)	(0.009)
working time^2	**−0.001*****	**−0.001*****	**−0.001*****	**−0.001*****	**−0.001*****	**−0.000*****
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Atypical schedules	**−0.102*****	**−0.068****	**−0.097*****	**−0.073****	**−0.081*****	−0.062*
	(0.033)	(0.034)	(0.030)	(0.032)	(0.028)	(0.032)
Wage (ref:1,800-2,499)						
100-1,199	0.025	−0.049	0.101	0.004	**0.132****	0.048
	(0.072)	(0.072)	(0.064)	(0.069)	(0.060)	(0.068)
1,200-1,499	0.065	−0.089*	0.085*	−0.060	**0.121*****	−0.022
	(0.050)	(0.053)	(0.045)	(0.051)	(0.043)	(0.050)
1,500-1,799	0.026	**−0.097****	0.054	−0.063	0.058	−0.045
	(0.046)	(0.049)	(0.042)	(0.047)	(0.040)	(0.046)
2,500-2,999	0.098	−0.070	0.016	−0.092	0.015	−0.092
	(0.064)	(0.063)	(0.060)	(0.061)	(0.057)	(0.061)
3,000-20,000	0.023	**−0.147****	−0.013	**−0.158****	−0.057	**−0.178*****
	(0.072)	(0.066)	(0.065)	(0.064)	(0.062)	(0.063)
Quantified target	**0.082****	**0.116*****	0.053	**0.116*****	**0.073****	**0.122*****
	(0.038)	(0.039)	(0.034)	(0.038)	(0.032)	(0.037)
Work rhythm paced by:						
machine	0.052	**0.186*****	**0.088****	**0.202*****	**0.119*****	**0.216*****

	SA days	Illness	SA days	Illness	SA days	Illness
	(0.051)	(0.058)	(0.045)	(0.056)	(0.042)	(0.055)
external demand	**−0.150*****	−0.015	**−0.144*****	−0.024	**−0.139*****	−0.025
	(0.037)	(0.037)	(0.033)	(0.035)	(0.031)	(0.035)
norms	0.020	0.065*	0.025	0.064*	0.031	0.060*
	(0.035)	(0.036)	(0.032)	(0.034)	(0.030)	(0.034)
computing control	−0.010	0.032	0.006	0.030	0.012	0.030
	(0.035)	(0.036)	(0.031)	(0.035)	(0.030)	(0.034)
monitoring	0.015	**0.131*****	**0.072****	**0.163*****	**0.069****	**0.168*****
	(0.037)	(0.039)	(0.033)	(0.038)	(0.031)	(0.037)
colleagues	−0.048	**0.095****	−0.012	**0.100*****	0.001	**0.105*****
	(0.036)	(0.038)	(0.032)	(0.037)	(0.031)	(0.036)
Help form Supervisors	−0.008	**−0.258*****	−0.058*	**−0.268*****	**−0.096*****	**−0.279*****
	(0.034)	(0.035)	(0.030)	(0.034)	(0.028)	(0.033)
Help form colleagues	**0.107****	**0.103****	**0.090****	0.088*	**0.101****	0.088*
	(0.050)	(0.048)	(0.043)	(0.046)	(0.041)	(0.045)
Tensions at work	**0.109*****	**0.466*****	**0.159*****	**0.476*****	**0.181*****	**0.473*****
	(0.036)	(0.034)	(0.032)	(0.033)	(0.030)	(0.032)
Lack of time	−0.007	**0.266*****	0.013	**0.271*****	0.028	**0.275*****
	(0.036)	(0.040)	(0.032)	(0.038)	(0.031)	(0.038)
Under pressure	−0.010	**0.309*****	**0.068****	**0.342*****	**0.096*****	**0.353*****
	(0.037)	(0.040)	(0.033)	(0.039)	(0.031)	(0.038)
Choice of methods	−0.060	−0.045	**−0.111*****	−0.072*	**−0.095*****	−0.069*
	(0.040)	(0.042)	(0.035)	(0.041)	(0.033)	(0.040)
Initiative	−0.040	0.048	−0.012	0.050	−0.022	0.041
	(0.033)	(0.034)	(0.030)	(0.032)	(0.028)	(0.032)
Work-to-life conflicts	−0.040	**0.314*****	−0.012	**0.314*****	0.045	**0.335*****
	(0.042)	(0.046)	(0.037)	(0.045)	(0.035)	(0.044)
Fear of job future	**−0.085****	**0.221*****	−0.048	**0.226*****	0.010	**0.249*****
	(0.041)	(0.043)	(0.036)	(0.042)	(0.034)	(0.041)
Organizational changes	**0.132*****	**0.266*****	**0.146*****	**0.276*****	**0.145*****	**0.274*****
	(0.033)	(0.033)	(0.029)	(0.032)	(0.028)	(0.032)
Family conflicts		**0.365*****		**0.343*****		**0.321*****
		(0.036)		(0.034)		(0.032)
sigma	12.390	12.390	38.675	38.675	76.958	76.958
theta	1.686	1.686	2.259	2.259	3.158	3.158
tau	0.276	0.276	0.407	0.407	0.536	0.536
AIC	69470.049	69470.049	89897.785	89897.785	97941.076	97941.076
Log Likelihood	−34594.025	−34594.025	−44807.892	−44807.892	−48829.538	−48829.538
Num. obs.	10213	18425	11692	19904	12131	20343

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$
We control for individual characteristics, firm sectors and firm size.

5 Models with different choice of copula

In this section we investigate the sensitivity of our results to the choice of copula, the margins remaining unchanged. The first table below, Table 12, shows what happens if the error in the choice of copula is not very serious, meaning that we have chosen the copula with the second-best AIC. Table 13 reports the results with the worst choice of copula, where we selected the copula with the highest AIC.

Table A17 – SA and SP models with the second best choice of copula

	Sickness absenteeism		Sickness presenteeism	
Equation	SA days	Illness	SP days	Illness
Distribution	(NB II)	(logit)	(PIG)	(logit)
Copula	Clayton 180		Gumbel 270	
Constant	**1.412*****	−0.007	**1.134*****	−0.015
	(0.365)	(0.367)	(0.285)	(0.369)
Health (ref: good)				
bad or very bad	**0.585*****		**0.445*****	
	(0.057)		(0.050)	
rather good	**0.177*****		**0.285*****	
	(0.034)		(0.027)	
very good	−0.081*		**−0.149*****	
	(0.041)		(0.033)	
Chronic disease	**0.198*****		**0.058****	
	(0.032)		(0.025)	
Occupation (ref: employee)				
labourer	0.052	0.061	−0.026	0.076
	(0.053)	(0.056)	(0.043)	(0.057)
intermediate	0.070	−0.003	−0.003	−0.003
	(0.044)	(0.047)	(0.035)	(0.047)
executive	−0.027	−0.080	0.005	−0.064
	(0.063)	(0.064)	(0.049)	(0.064)
Seniority	0.004	**0.022*****	−0.007	**0.024*****
	(0.006)	(0.006)	(0.005)	(0.006)
Seniority^2	−0.000	**−0.001*****	0.000	**−0.001*****
	(0.000)	(0.000)	(0.000)	(0.000)
Work alone	**−0.070****	**0.091****	0.014	**0.096****
	(0.035)	(0.037)	(0.028)	(0.037)
Supervisor	**−0.194*****	**−0.227*****	−0.009	**−0.212*****
	(0.051)	(0.049)	(0.038)	(0.049)
Public	**0.156*****	**0.134*****	−0.008	**0.110****
	(0.043)	(0.047)	(0.034)	(0.047)
Temporary contract	**−0.460*****	**−0.516*****	0.062	**−0.502*****
	(0.078)	(0.070)	(0.058)	(0.070)
Working time	**0.067*****	**0.042*****	**−0.016****	**0.040*****
	(0.011)	(0.010)	(0.007)	(0.010)
Working time^2	**−0.001*****	**−0.001*****	**0.000*****	**−0.001*****
	(0.000)	(0.000)	(0.000)	(0.000)
Atypical schedules	**−0.101*****	**−0.076****	**0.058****	**−0.069****
	(0.031)	(0.033)	(0.025)	(0.033)
Wage (ref:1,800-2,499)				
100-1,199	0.084	−0.025	−0.005	−0.031
	(0.066)	(0.071)	(0.053)	(0.071)
1,200-1,499	0.068	−0.078	−0.013	−0.055
	(0.047)	(0.052)	(0.038)	(0.052)
1,500-1,799	0.066	−0.074	0.053	−0.064
	(0.043)	(0.048)	(0.035)	(0.048)
2,500-2,999	0.029	−0.087	0.076	−0.095

	SA days	Illness	SP days	Illness
	(0.061)	(0.062)	(0.048)	(0.062)
3,000-20,000	0.015	**−0.153****	0.005	**−0.161****
	(0.067)	(0.065)	(0.051)	(0.065)
Quantified target	0.062*	**0.120*****	**0.085*****	**0.111*****
	(0.035)	(0.039)	(0.028)	(0.039)
Work rhythm paced by:				
machine	0.066	**0.192*****	0.011	**0.181*****
	(0.047)	(0.057)	(0.039)	(0.057)
external demand	**−0.137*****	−0.019	**0.056****	−0.016
	(0.034)	(0.036)	(0.028)	(0.036)
norms	0.015	0.066*	0.002	**0.075****
	(0.033)	(0.035)	(0.026)	(0.035)
computing control	0.009	0.035	0.003	0.040
	(0.033)	(0.036)	(0.026)	(0.036)
monitoring	0.045	**0.151*****	0.039	**0.146*****
	(0.034)	(0.039)	(0.027)	(0.039)
colleagues	−0.011	**0.104*****	0.013	**0.107*****
	(0.033)	(0.038)	(0.027)	(0.038)
Help form Supervisors	−0.038	**−0.269*****	**−0.070*****	**−0.267*****
	(0.031)	(0.035)	(0.025)	(0.035)
Help from colleagues	**0.089****	0.091*	**−0.128*****	0.089*
	(0.045)	(0.047)	(0.034)	(0.047)
Tensions at work	**0.124*****	**0.473*****	**0.090*****	**0.471*****
	(0.034)	(0.033)	(0.028)	(0.033)
Lack of time	0.004	**0.269*****	**0.094*****	**0.270*****
	(0.034)	(0.039)	(0.027)	(0.039)
Under pressure	0.011	**0.324*****	**0.138*****	**0.317*****
	(0.034)	(0.039)	(0.028)	(0.039)
Choice of methods	**−0.095*****	−0.062	0.026	−0.052
	(0.036)	(0.041)	(0.030)	(0.042)
Initiative	−0.034	0.044	**0.093*****	0.046
	(0.031)	(0.033)	(0.025)	(0.033)
Work-to-life conflicts	−0.032	**0.319*****	0.032	**0.325*****
	(0.039)	(0.046)	(0.030)	(0.046)
Fear of job future	−0.064*	**0.222*****	**0.085*****	**0.218*****
	(0.037)	(0.042)	(0.029)	(0.042)
Organizational changes	**0.127*****	**0.270*****	0.011	**0.262*****
	(0.031)	(0.033)	(0.025)	(0.033)
Family conflicts		**0.364*****		**0.404*****
		(0.035)		(0.035)
sigma	21.160	21.160	1.463	1.463
theta	0.898	0.898	−1.519	−1.519
tau	0.310	0.310	−0.342	−0.342
AIC	80359.074	80359.074	77990.544	77990.544
Log Likelihood	−40038.537	−40038.537	−38854.272	−38854.272
Num. obs.	11073	19285	11073	19285

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$
We control for individual characteristics, firm sectors and firm size.

Table A18 – SA and SP models with the worst choice of copula

	Sickness absenteeism		Sickness presenteeism	
Equation	SA days	Illness	SP days	Illness
Distribution	(NB II)	(logit)	(PIG)	(logit)
Copula	Student-t		Clayton 270	
Constant	**1.725*****	0.003	**1.224*****	0.026
	(0.330)	(0.370)	(0.304)	(0.368)
Health (ref: good)				
bad or very bad	**0.541*****		**0.460*****	
	(0.053)		(0.052)	
rather good	**0.161*****		**0.296*****	
	(0.031)		(0.028)	
very good	−0.070*		**−0.152*****	
	(0.038)		(0.033)	
Chronic disease	**0.181*****		**0.065****	
	(0.029)		(0.026)	
Occupation (ref: employee)				
labourer	0.035	0.066	−0.018	0.071
	(0.047)	(0.057)	(0.043)	(0.057)
intermediate	0.068*	0.001	−0.004	−0.001
	(0.039)	(0.048)	(0.035)	(0.047)
executive	−0.011	−0.076	−0.001	−0.074
	(0.057)	(0.065)	(0.048)	(0.065)
Seniority	−0.000	**0.023*****	−0.004	**0.024*****
	(0.005)	(0.006)	(0.005)	(0.006)
Seniority^2	−0.000	**−0.001*****	0.000	**−0.001*****
	(0.000)	(0.000)	(0.000)	(0.000)
Work alone	**−0.084*****	**0.088****	0.021	**0.096****
	(0.031)	(0.038)	(0.028)	(0.038)
Supervisor	**−0.155*****	**−0.229*****	−0.035	**−0.207*****
	(0.046)	(0.050)	(0.038)	(0.049)
Public	**0.131*****	**0.136*****	0.002	**0.121*****
	(0.038)	(0.047)	(0.034)	(0.047)
Temporary contract	**−0.351*****	**−0.519*****	0.011	**−0.513*****
	(0.071)	(0.070)	(0.058)	(0.070)
Working time	**0.059*****	**0.044*****	−0.012	**0.041*****
	(0.010)	(0.010)	(0.007)	(0.010)
Working time^2	**−0.001*****	**−0.001*****	**0.000****	**−0.001*****
	(0.000)	(0.000)	(0.000)	(0.000)
Atypical schedules	**−0.088*****	**−0.080****	0.048*	**−0.080****
	(0.028)	(0.033)	(0.025)	(0.033)
Wage (ref:1,800-2,499)				
100-1,199	0.080	−0.039	−0.014	−0.035
	(0.060)	(0.072)	(0.053)	(0.071)
1,200-1,499	0.072*	−0.073	−0.022	−0.059
	(0.042)	(0.053)	(0.038)	(0.052)
1,500-1,799	0.073*	−0.078	0.047	−0.071
	(0.038)	(0.048)	(0.035)	(0.048)

	SA days	Illness	SP days	Illness
2,500-2,999	0.041	−0.085	0.065	−0.088
	(0.055)	(0.062)	(0.047)	(0.063)
3,000_20,000	0.046	**−0.145****	−0.026	**−0.153****
	(0.060)	(0.065)	(0.050)	(0.065)
Quantified target	0.042	**0.118*****	**0.099*****	**0.114*****
	(0.031)	(0.039)	(0.028)	(0.039)
Work rhythm paced by:				
machine	0.040	**0.189*****	0.022	**0.183*****
	(0.042)	(0.058)	(0.039)	(0.058)
external demand	**−0.126*****	−0.023	0.046	−0.023
	(0.031)	(0.036)	(0.028)	(0.036)
norms	0.002	**0.071****	0.011	**0.077****
	(0.029)	(0.035)	(0.026)	(0.035)
computing control	0.004	0.039	0.008	0.041
	(0.029)	(0.036)	(0.026)	(0.036)
monitoring	0.021	**0.147*****	0.051*	**0.149*****
	(0.030)	(0.039)	(0.028)	(0.039)
colleagues	−0.024	**0.105*****	0.030	**0.106*****
	(0.030)	(0.038)	(0.027)	(0.038)
Help from Supervisors	−0.005	**−0.272*****	**−0.097*****	**−0.268*****
	(0.028)	(0.035)	(0.026)	(0.035)
Help from colleagues	0.075*	**0.094****	**−0.131*****	**0.094****
	(0.041)	(0.047)	(0.034)	(0.047)
Tensions at work	0.039	**0.471*****	**0.148*****	**0.464*****
	(0.031)	(0.033)	(0.029)	(0.033)
Lack of time	−0.029	**0.268*****	**0.116*****	**0.257*****
	(0.030)	(0.039)	(0.028)	(0.039)
Under pressure	−0.031	**0.323*****	**0.172*****	**0.318*****
	(0.031)	(0.040)	(0.029)	(0.039)
Choice of methods	**−0.081****	−0.063	0.017	−0.074*
	(0.033)	(0.042)	(0.030)	(0.042)
Initiative	−0.036	0.045	**0.104*****	0.046
	(0.028)	(0.033)	(0.025)	(0.033)
Work-to-life conflicts	−0.061*	**0.330*****	0.056*	**0.317*****
	(0.035)	(0.046)	(0.031)	(0.046)
Fear of job future	**−0.089*****	**0.219*****	**0.105*****	**0.215*****
	(0.034)	(0.043)	(0.030)	(0.043)
Organizational changes	**0.079*****	**0.267*****	0.039	**0.261*****
	(0.028)	(0.033)	(0.025)	(0.033)
Family conflicts		**0.398*****		**0.406*****
		(0.036)		(0.036)
sigma	23.141	23.141	2.284	2.284
theta	0.249	0.249	−0.728	−0.728
tau	0.160	0.160	−0.267	−0.267
AIC	80460.959	80460.959	78085.045	78085.045
Log Likelihood	−40089.479	−40089.479	−38901.522	−38901.522
Num. obs.	11073	19285	11073	19285

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$
We control for individual characteristics, firm sectors and firm size.

6 Models with alternative exclusion restrictions

In this section, we present the results using the other two exclusion restrictions discussed in the chapter. These two exclusion variables are dummies indicating whether the individual had been subjected, before the age of 18, to violence at school or in the neighbourhood (Table 14) or to physical and/or psychological mistreatment (Table 15).

Table A19 – Determinants of sickness presenteeism and absence days with endogenous participation, without exclusion restriction in the first stage

	Sickness absence			Sickness presenteeism		
Equation	SA days		Illness	SP days		Illness
Distribution	(NB II)		(logit)	(PIG)		(logit)
	Coef.	AME[a]	Coef.	Coef.	AME[a]	Coef.
Constant	**1.374*****		−0.025	**1.170*****		−0.046
	(0.368)		(0.365)	(0.283)		(0.365)
Health (ref: good)						
bad or very bad	**0.580*****	**3.258**	−−	**0.453*****	**3.351**	−−
	(0.058)			(0.049)		
rather good	**0.175*****	**0.789**	−−	**0.290*****	**1.853**	−−
	(0.034)			(0.027)		
very good	**−0.081***	**−0.341**	−−	**−0.152*****	**−0.866**	−−
	(0.041)			(0.033)		
Chronic disease	**0.195*****	**0.865**	−−	**0.066*****	**0.403**	−−
	(0.032)			(0.025)		
Occupation (ref: employee)						
labourer	0.057	0.251	0.053	−0.025	−0.152	0.070
	(0.053)		(0.056)	(0.043)		(0.056)
intermediate	0.070	0.307	−0.002	−0.006	−0.037	−0.004
	(0.044)		(0.047)	(0.034)		(0.047)
executive	−0.032	−0.136	−0.081	0.000	0.002	−0.064
	(0.064)		(0.064)	(0.048)		(0.064)
Seniority	0.004	−0.002	**0.021*****	−0.006	−0.008	**0.023*****
	(0.006)		(0.006)	(0.005)		(0.006)
Seniority^2	−0.000		**−0.001*****	0.000		**−0.001*****
	(0.000)		(0.000)	(0.000)		(0.000)
Work alone	**−0.063***	**−0.272**	**0.100*****	0.020	0.118	**0.108*****
	(0.035)		(0.037)	(0.027)		(0.037)
Supervisor	**−0.197*****	**−0.790**	**−0.229*****	−0.015	−0.092	**−0.210*****
	(0.051)		(0.049)	(0.037)		(0.049)
Public sector	**0.156*****	**0.679**	**0.138*****	−0.001	−0.006	**0.118****
	(0.043)		(0.047)	(0.034)		(0.047)
Temporary contract	**−0.475*****	**−1.675**	**−0.507*****	0.032	0.196	**−0.498*****
	(0.079)		(0.069)	(0.058)		(0.069)
Working time	**0.067*****	**−0.035**	**0.041*****	**−0.014****	**0.018**	**0.039*****
	(0.011)		(0.009)	(0.007)		(0.009)
Working time^2	**−0.001*****		**−0.001*****	**0.000****		**−0.001*****
	(0.000)		(0.000)	(0.000)		(0.000)
Atypical schedules	**−0.101*****	**−0.442**	**−0.075****	**0.055****	**0.333**	**−0.064***
	(0.031)		(0.033)	(0.025)		(0.033)
Wage (ref:1,800-2,499)						
100-1,199	0.085	0.382	−0.017	−0.004	−0.022	−0.016
	(0.067)		(0.070)	(0.052)		(0.071)

	SA days		Illness	SP days		Illness
	Coef.	AME	Coef.	Coef.	AME	Coef.
1,200-1,499	0.067	0.294	−0.074	−0.013	−0.080	−0.046
	(0.047)		(0.052)	(0.038)		(0.052)
1,500-1,799	0.065	0.288	−0.073	0.049	0.301	−0.066
	(0.043)		(0.047)	(0.034)		(0.047)
2,500-2,999	0.027	0.120	−0.091	0.075	0.470	−0.086
	(0.062)		(0.062)	(0.047)		(0.062)
3,000-20,000	0.008	0.035	**−0.163****	−0.011	−0.011	**−0.166*****
	(0.068)		(0.065)	(0.050)		(0.064)
Quantified target	**0.063***	**0.278**	**0.126*****	**0.087*****	**0.531**	**0.117*****
	(0.035)		(0.038)	(0.028)		(0.039)
Work rhythm paced by:						
machine	0.070	0.311	**0.202*****	0.015	0.091	**0.191*****
	(0.047)		(0.057)	(0.038)		(0.057)
external demand	**−0.136*****	**−0.607**	−0.013	**0.055****	**0.330**	−0.004
	(0.035)		(0.036)	(0.028)		(0.036)
norms	0.016	0.068	**0.070****	0.009	0.051	**0.084****
	(0.033)		(0.035)	(0.026)		(0.035)
computing control	0.012	0.051	0.031	0.002	0.013	0.035
	(0.033)		(0.035)	(0.026)		(0.035)
monitoring	0.050	0.219	**0.157*****	0.040	0.246	**0.149*****
	(0.034)		(0.039)	(0.027)		(0.038)
colleagues relation	−0.009	−0.040	**0.109*****	0.017	0.102	**0.104*****
	(0.034)		(0.037)	(0.026)		(0.037)
Help from supervisors	−0.042	−0.185	**−0.269*****	**−0.076*****	**−0.461**	**−0.262*****
	(0.031)		(0.034)	(0.025)		(0.034)
Help from colleagues	**0.092****	**0.387**	0.090*	**−0.127*****	**−0.797**	**0.093****
	(0.046)		(0.047)	(0.034)		(0.046)
Tensions at work	**0.139*****	**0.589**	**0.491*****	**0.113*****	**0.668**	**0.492*****
	(0.034)		(0.033)	(0.028)		(0.033)
Lack of time	0.011	0.047	**0.268*****	**0.100*****	**0.609**	**0.269*****
	(0.034)		(0.039)	(0.026)		(0.039)
Under pressure	0.019	0.081	**0.350*****	**0.146*****	**0.890**	**0.339*****
	(0.034)		(0.039)	(0.027)		(0.039)
Choice of methods	**−0.097*****	**−0.434**	−0.051	0.023	0.139	−0.033
	(0.037)		(0.041)	(0.030)		(0.042)
Initiative	−0.033	−0.145	0.047	**0.094*****	**0.568**	0.044
	(0.031)		(0.033)	(0.024)		(0.033)
Work-to-life conflict	−0.026	−0.113	**0.326*****	0.037	0.226	**0.333*****
	(0.039)		(0.046)	(0.030)		(0.045)
Fear of job future	−0.058	−0.247	**0.231*****	**0.092*****	**0.571**	**0.230*****
	(0.038)		(0.042)	(0.029)		(0.042)
Organizational changes	**0.134*****	**0.580**	**0.284*****	0.022	0.134	**0.280*****
	(0.031)		(0.033)	(0.024)		(0.033)
Num. obs.	11,073		19,285	11,073		19,285
τ			0.329			−0.335
			(0.292, 0.37)			(−0.384, −0.29)
AIC			80454.396			78094.637
Log Likelihood			−40087.198			−38907.319

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

[a] Average marginal effects (AME) are calculated conditional on illness. We control for individual's characteristics (age, gender, marital status, having children and education level), firm sectors and firm size.

Table A20 – SA and SP models with violence at school as exclusion

	Sickness absenteeism		Sickness presenteeism	
Equation	SA days	Illness	SP days	Illness
Distribution	(NB II)	(logit)	(PIG)	(logit)
Copula	Joe		Student-t	
Constant	**1.369*****	−0.018	**1.171*****	−0.042
	(0.368)	(0.366)	(0.283)	(0.366)
Health (ref: good)				
bad or very bad	**0.581*****		**0.450*****	
	(0.058)		(0.049)	
rather good	**0.175*****		**0.290*****	
	(0.034)		(0.027)	
very good	**−0.082****		**−0.151*****	
	(0.041)		(0.033)	
Chronic disease	**0.196*****		**0.065****	
	(0.032)		(0.025)	
Occupation (ref: employee)				
labourer	0.056	0.053	−0.025	0.071
	(0.053)	(0.056)	(0.043)	(0.056)
intermediate	0.069	−0.001	−0.005	−0.002
	(0.044)	(0.047)	(0.034)	(0.047)
executive	−0.032	−0.081	0.001	−0.063
	(0.064)	(0.064)	(0.048)	(0.064)
Seniority	0.004	**0.022*****	−0.006	**0.023*****
	(0.006)	(0.006)	(0.005)	(0.006)
Seniority^2	−0.000	**−0.001*****	0.000	**−0.001*****
	(0.000)	(0.000)	(0.000)	(0.000)
Work alone	−0.064*	**0.096*****	0.020	**0.105*****
	(0.035)	(0.037)	(0.027)	(0.037)
Supervisor	**−0.197*****	**−0.225*****	−0.015	**−0.206*****
	(0.051)	(0.049)	(0.037)	(0.049)
Public	**0.156*****	**0.136*****	−0.001	**0.116****
	(0.043)	(0.047)	(0.034)	(0.047)
Temporary contract	**−0.474*****	**−0.513*****	0.032	**−0.505*****
	(0.079)	(0.069)	(0.058)	(0.069)
Working time	**0.068*****	**0.040*****	**−0.015****	**0.039*****
	(0.011)	(0.009)	(0.007)	(0.009)
Working time^2	**−0.001*****	**−0.001*****	**0.000****	**−0.000*****
	(0.000)	(0.000)	(0.000)	(0.000)
Atypical schedules	**−0.101*****	**−0.076****	**0.055****	**−0.065****
	(0.031)	(0.033)	(0.025)	(0.033)
Wage (ref:1,800-2,499)				
100-1,199	0.087	−0.024	−0.004	−0.021
	(0.067)	(0.070)	(0.052)	(0.071)
1,200-1,499	0.067	−0.076	−0.013	−0.047
	(0.047)	(0.052)	(0.038)	(0.052)
1,500-1,799	0.066	−0.074	0.049	−0.066
	(0.043)	(0.048)	(0.034)	(0.048)

	SA days	Illness	SP days	Illness
2,500-2,999	0.029	−0.089	0.076	−0.085
	(0.062)	(0.062)	(0.047)	(0.062)
3,000-20,000	0.008	**−0.162****	−0.002	**−0.165*****
	(0.068)	(0.065)	(0.050)	(0.064)
Quantified target	0.063*	**0.125*****	**0.087*****	**0.115*****
	(0.035)	(0.038)	(0.028)	(0.039)
Work rhythm paced by:				
machine	0.070	**0.196*****	0.015	**0.184*****
	(0.047)	(0.057)	(0.038)	(0.057)
external demand	**−0.136*****	−0.012	**0.056****	−0.002
	(0.035)	(0.036)	(0.028)	(0.036)
norms	0.015	0.067*	0.009	**0.081****
	(0.033)	(0.035)	(0.026)	(0.035)
computing control	0.011	0.030	0.002	0.033
	(0.033)	(0.035)	(0.026)	(0.035)
monitoring	0.050	**0.156*****	0.040	**0.149*****
	(0.034)	(0.039)	(0.027)	(0.038)
colleagues	−0.009	**0.108*****	0.016	**0.103*****
	(0.034)	(0.037)	(0.026)	(0.037)
Help from Supervisors	−0.042	**−0.270*****	**−0.076*****	**−0.262*****
	(0.031)	(0.034)	(0.025)	(0.034)
Help from colleagues	**0.091****	0.090*	**−0.126*****	**0.093****
	(0.046)	(0.047)	(0.034)	(0.047)
Tensions at work	**0.137*****	**0.485*****	**0.114*****	**0.487*****
	(0.034)	(0.033)	(0.028)	(0.033)
Lack of time	0.009	**0.266*****	**0.100*****	**0.267*****
	(0.034)	(0.039)	(0.026)	(0.039)
Under pressure	0.018	**0.341*****	**0.146*****	**0.330*****
	(0.034)	(0.039)	(0.027)	(0.039)
Choice of methods	**−0.097*****	−0.049	0.023	−0.030
	(0.037)	(0.041)	(0.030)	(0.042)
Initiative	−0.034	0.048	**0.094*****	0.044
	(0.031)	(0.033)	(0.024)	(0.033)
Work-to-life conflicts	−0.028	**0.321*****	0.037	**0.328*****
	(0.039)	(0.046)	(0.030)	(0.045)
Fear of job future	−0.058	**0.229*****	**0.092*****	**0.227*****
	(0.038)	(0.042)	(0.029)	(0.042)
Organizational changes	**0.134*****	**0.284*****	0.022	**0.279*****
	(0.031)	(0.033)	(0.024)	(0.033)
School violence		**0.340*****		**0.358*****
		(0.073)		(0.073)
sigma	21.325	21.325	1.689	1.689
theta	1.874	1.874	−0.502	−0.502
tau	0.326	0.326	−0.335	−0.335
AIC	80433.626	80433.626	78071.705	78071.705
Log Likelihood	−40075.813	−40075.813	−38894.852	−38894.852
Num. obs.	11073	19285	11073	19285

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$
We control for individual characteristics, firm sectors and firm size.

Table A21: SA and SP models with mistreatment as exclusion

	Sickness absenteeism		Sickness presenteeism	
Equation	SA days	Illness	SP days	Illness
Distribution	(NB II)	(logit)	(PIG)	(logit)
Copula	Joe		Student-t	
Constant	**1.376*****	0.038	**1.192*****	0.023
	(0.368)	(0.366)	(0.283)	(0.366)
Health (ref: good)				
bad or very bad	**0.582*****		**0.449*****	
	(0.058)		(0.049)	
rather good	**0.176*****		**0.288*****	
	(0.034)		(0.027)	
very good	**−0.082****		**−0.151*****	
	(0.041)		(0.033)	
Chronic disease	**0.198*****		**0.064****	
	(0.032)		(0.025)	
Occupation (ref: employee)				
labourer	0.057	0.051	−0.026	0.067
	(0.053)	(0.056)	(0.043)	(0.056)
intermediate	0.069	0.000	−0.006	−0.000
	(0.044)	(0.047)	(0.034)	(0.047)
executive	−0.030	−0.079	0.001	−0.061
	(0.064)	(0.064)	(0.048)	(0.064)
Seniority	0.004	**0.022*****	−0.006	**0.023*****
	(0.006)	(0.006)	(0.005)	(0.006)
Seniority^2	−0.000	**−0.001*****	0.000	**−0.001*****
	(0.000)	(0.000)	(0.000)	(0.000)
Work alone	−0.065*	**0.095****	0.019	**0.104*****
	(0.035)	(0.037)	(0.027)	(0.037)
Supervisor	**−0.197*****	**−0.226*****	−0.014	**−0.206*****
	(0.051)	(0.049)	(0.037)	(0.049)
Public	**0.155*****	**0.136*****	−0.002	**0.116****
	(0.043)	(0.047)	(0.034)	(0.047)
Temporary contract	**−0.472*****	**−0.510*****	0.035	**−0.503*****
	(0.079)	(0.069)	(0.058)	(0.069)
Working time	**0.068*****	**0.041*****	**−0.015****	**0.039*****
	(0.011)	(0.010)	(0.007)	(0.009)
Working time^2	**−0.001*****	**−0.001*****	**0.000*****	**−0.001*****
	(0.000)	(0.000)	(0.000)	(0.000)
Atypical schedules	**−0.102*****	**−0.076****	**0.056****	−0.064*
	(0.031)	(0.033)	(0.025)	(0.033)
Wage (ref:1,800-2,499)				
100-1,199	0.087	−0.021	−0.003	−0.021
	(0.067)	(0.070)	(0.052)	(0.071)
1,200-1,499	0.066	−0.076	−0.013	−0.048
	(0.047)	(0.052)	(0.038)	(0.052)
1,500-1,799	0.066	−0.075	0.050	−0.067

	SA days	Illness	SP days	Illness
	(0.043)	(0.048)	(0.034)	(0.048)
2,500-2,999	0.027	−0.087	0.077*	−0.083
	(0.062)	(0.062)	(0.047)	(0.062)
3,000-20,000	0.008	**−0.155****	0.001	**−0.157****
	(0.068)	(0.065)	(0.050)	(0.064)
Quantified target	0.063*	**0.123*****	**0.086*****	**0.112*****
	(0.035)	(0.039)	(0.028)	(0.039)
Work rhythm paced by:				
machine	0.070	**0.201*****	0.014	**0.190*****
	(0.047)	(0.057)	(0.038)	(0.057)
external demand	**−0.136*****	−0.015	0.054*	−0.008
	(0.035)	(0.036)	(0.028)	(0.036)
norms	0.015	0.067*	0.008	**0.081****
	(0.033)	(0.035)	(0.026)	(0.035)
computing control	0.011	0.032	0.002	0.034
	(0.033)	(0.035)	(0.026)	(0.035)
monitoring	0.049	**0.153*****	0.040	**0.146*****
	(0.034)	(0.039)	(0.027)	(0.038)
colleagues	−0.010	**0.106*****	0.015	**0.100*****
	(0.034)	(0.037)	(0.026)	(0.037)
Help from Supervisors	−0.041	**−0.270*****	**−0.074*****	**−0.263*****
	(0.031)	(0.034)	(0.025)	(0.034)
Help from colleagues	**0.091****	**0.094****	**−0.126*****	**0.097****
	(0.046)	(0.047)	(0.034)	(0.047)
Tensions at work	**0.137*****	**0.484*****	**0.109*****	**0.484*****
	(0.034)	(0.033)	(0.028)	(0.033)
Lack of time	0.009	**0.267*****	**0.099*****	**0.266*****
	(0.034)	(0.039)	(0.026)	(0.039)
Under pressure	0.017	**0.339*****	**0.143*****	**0.324*****
	(0.034)	(0.039)	(0.027)	(0.039)
Choice of methods	**−0.097*****	−0.051	0.025	−0.032
	(0.037)	(0.041)	(0.030)	(0.042)
Initiative	−0.033	0.044	**0.094*****	0.041
	(0.031)	(0.033)	(0.024)	(0.033)
Work-to-life conflicts	−0.028	**0.321*****	0.036	**0.327*****
	(0.039)	(0.046)	(0.030)	(0.046)
Fear of job future	−0.059	**0.228*****	**0.092*****	**0.226*****
	(0.038)	(0.042)	(0.029)	(0.042)
Organizational changes	**0.133*****	**0.282*****	0.021	**0.278*****
	(0.031)	(0.033)	(0.024)	(0.033)
Mistreatment		**0.349*****		**0.410*****
		(0.061)		(0.060)
sigma	21.273	21.273	1.706	1.706
theta	1.858	1.858	−0.515	−0.515
tau	0.322	0.322	−0.344	−0.344
AIC	80421.762	80421.762	78048.535	78048.535
Log Likelihood	−40069.881	−40069.881	−38883.267	−38883.267
Num. obs.	11073	19285	11073	19285

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$
We control for individual characteristics, firm sectors and firm size.

7 Additional controls for past health and employment history

Finally, in this last section, we present the results associated with models including additional control variables relating to past health and employment history in order to control for the possible self-selection of employees in their employment.

Table A22 – SA and SP models with additional controls about past health and employment history

	Sickness absenteeism		Sickness presenteeism	
Equation	SA days	Illness	SP days	Illness
Distribution	(NB II)	(logit)	(PIG)	(logit)
Copula	Joe		Student-t	
Constant	**1.575*****	0.236	**1.180*****	0.173
	(0.368)	(0.369)	(0.283)	(0.369)
Health (ref: good)				
bad or very bad	**0.584*****		**0.442*****	
	(0.057)		(0.050)	
rather good	**0.176*****		**0.285*****	
	(0.034)		(0.027)	
very good	-0.080^{*}		$\mathbf{-0.149}^{***}$	
	(0.042)		(0.033)	
Chronic disease	**0.194*****		**0.060****	
	(0.032)		(0.025)	
Unemployment 3-12 months	0.050	**0.092*****	−0.032	**0.082****
	(0.032)	(0.034)	(0.025)	(0.034)
Unemployment >1 year	**0.134*****	**0.236*****	0.017	**0.230*****
	(0.037)	(0.040)	(0.029)	(0.040)
Health problem >1 year	**0.166*****	**0.224*****	0.018	**0.245*****
	(0.059)	(0.074)	(0.049)	(0.075)
Other break >1 year	0.001	−0.019	0.039	−0.027
	(0.039)	(0.043)	(0.031)	(0.043)
Occupation (ref: employee)				
labourer	0.061	0.072	−0.024	0.088
	(0.053)	(0.057)	(0.043)	(0.057)
intermediate	0.084^{*}	0.020	−0.004	0.016
	(0.044)	(0.047)	(0.034)	(0.047)
executive	−0.018	−0.058	0.005	−0.042
	(0.064)	(0.064)	(0.048)	(0.064)
Seniority	0.006	**0.025*****	−0.006	**0.027*****
	(0.006)	(0.006)	(0.005)	(0.006)
Seniority^2	−0.000	$\mathbf{-0.001}^{***}$	0.000	$\mathbf{-0.001}^{***}$

	SA days	Illness	SP days	Illness
	(0.000)	(0.000)	(0.000)	(0.000)
Work alone	**−0.073****	**0.084****	0.019	**0.093****
	(0.035)	(0.037)	(0.027)	(0.037)
Supervisor	**−0.196*****	**−0.229*****	−0.015	**−0.207*****
	(0.051)	(0.049)	(0.037)	(0.049)
Public	**0.148*****	**0.125*****	−0.002	**0.106****
	(0.043)	(0.047)	(0.034)	(0.047)
Temporary contract	**−0.483*****	**−0.542*****	0.036	**−0.533*****
	(0.079)	(0.070)	(0.058)	(0.070)
Working time	**0.065*****	**0.039*****	−0.014*	**0.038*****
	(0.011)	(0.010)	(0.007)	(0.009)
Working time^2	**−0.001*****	**−0.001*****	**0.000****	**−0.000*****
	(0.000)	(0.000)	(0.000)	(0.000)
Atypical schedules	**−0.098*****	**−0.069****	**0.056****	−0.058*
	(0.031)	(0.033)	(0.025)	(0.033)
Wage (ref:1,800-2,499)				
100-1,199	0.058	−0.064	−0.005	−0.060
	(0.067)	(0.071)	(0.052)	(0.071)
1,200-1,499	0.047	**−0.112****	−0.011	−0.080
	(0.047)	(0.052)	(0.038)	(0.052)
1,500-1,799	0.056	−0.087*	0.049	−0.078
	(0.043)	(0.048)	(0.034)	(0.048)
2,500-2,999	0.035	−0.079	0.077*	−0.079
	(0.062)	(0.062)	(0.047)	(0.062)
3,000-20,000	0.029	**−0.131****	−0.000	**−0.139****
	(0.068)	(0.065)	(0.050)	(0.064)
Quantified target	0.060*	**0.116*****	**0.085*****	**0.103*****
	(0.035)	(0.039)	(0.028)	(0.039)
Work rhythm paced by:				
machine	0.058	**0.176*****	0.012	**0.168*****
	(0.047)	(0.057)	(0.038)	(0.057)
demand	**−0.137*****	−0.018	**0.055****	−0.010
	(0.035)	(0.036)	(0.028)	(0.036)
norms	0.015	0.064*	0.008	**0.077****
	(0.033)	(0.035)	(0.026)	(0.035)
computer control	0.010	0.033	0.003	0.037
	(0.033)	(0.036)	(0.026)	(0.036)
monitoring	0.048	**0.151*****	0.041	**0.140*****
	(0.034)	(0.039)	(0.027)	(0.039)
colleagues	−0.008	**0.107*****	0.014	**0.101*****
	(0.034)	(0.038)	(0.026)	(0.038)
Help from Supervisors	−0.038	**−0.266*****	**−0.074*****	**−0.260*****
	(0.031)	(0.035)	(0.025)	(0.035)
Help from colleagues	**0.093****	**0.095****	**−0.126*****	**0.099****
	(0.046)	(0.047)	(0.034)	(0.047)
Tensions at work	**0.128*****	**0.471*****	**0.110*****	**0.472*****
	(0.034)	(0.033)	(0.028)	(0.033)
Lack of time	0.008	**0.271*****	**0.100*****	**0.270*****
	(0.034)	(0.039)	(0.026)	(0.039)
Under pressure	0.016	**0.327*****	**0.144*****	**0.310*****

	SA days	Illness	SP days	Illness
	(0.034)	(0.039)	(0.027)	(0.039)
Choice of methods	**−0.092****	−0.055	0.023	−0.040
	(0.037)	(0.041)	(0.030)	(0.042)
Initiative	−0.031	0.051	**0.094*****	0.048
	(0.031)	(0.033)	(0.024)	(0.033)
Work-to-life conflicts	−0.034	**0.315*****	0.036	**0.318*****
	(0.039)	(0.046)	(0.030)	(0.046)
Fear of job future	−0.066*	**0.215*****	**0.094*****	**0.213*****
	(0.038)	(0.042)	(0.029)	(0.042)
Organizational changes	**0.130*****	**0.270*****	0.021	**0.262*****
	(0.031)	(0.033)	(0.024)	(0.033)
Family conflicts		**0.354*****		**0.398*****
		(0.035)		(0.035)
sigma	21.102	21.102	1.703	1.703
theta	1.820	1.820	−0.516	−0.516
tau	0.312	0.312	−0.345	−0.345
AIC	80293.115	80293.115	77907.737	77907.737
Log Likelihood	−39997.558	−39997.558	−38804.868	−38804.868
Num. obs.	11073	19285	11073	19285

Notes: $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$
We control for individual characteristics, firm sectors and firm size.

Chapter 2

All work and no rest makes Jack a sick boy: A causal analysis of sickness absence and presenteeism on health

1 Introduction

Sickness absences and work accidents have a significant cost for the healthcare system. In a context where the French Cour des Comptes (2024) underlines a huge increase in the cost of the social security system, that represents €12 billion in 2022, a 56% increase since 2017, reducing absenteeism and promoting workers' health would be a central concern. A report from IGF and IGAS by Triolle et al. (2024) indicates that sickness absenteeism rose by 63% between 2019 and 2022 in the French public sector, a far greater increase than in the private sector, which was not spared either, with an increase of 39%. For example, compensation for sick leave days in the public sector amounted to €15 billion in 2022. Since 2019, the rate for long-term sick leave (90 days or more) in the private sector has been rising steadily, with an increase of more than 30% between 2019 and 2023 . Further, according to the 2024 barometer published by WTW (Willis Towers Watson), the direct cost of absenteeism per employee in the French private sector would be €1,535 on average, an increase of 35% since 2019.

Therefore, being aware of the mechanisms that can help to limit this rise in absenteeism seems essential in this context. However, how can this goal be achieved? More importantly, is it always cost-effective to encourage attendance at work? With regard to sickness absenteeism, it is important to consider its antagonistic behaviour: sickness presenteeism. The introduction of strong incentives for attendance could lead workers to respond to these incentives inappropriately and come to work when sick rather than taking time off to rest (Aronsson et al., 2000). This behaviour can be expected to be detrimental to individuals in terms of health (Bergström et al., 2009a), as well as to the company, with numerous hidden costs (Hemp, 2004). This rises the importance of work sustainability that address the "regeneration of human and social resources" (Docherty et al., 2002). The aim of this chapter is thus to study the long-term effects of both sickness absenteeism and sickness presenteeism on individuals' health.

Report from the Inspection générale des finances and Inspection générale des affaires sociales.
Barometer Ayming/AG2R20 of 15 April 2024: "Ayming et AG2R LA MONDIALE dévoilent les premiers chiffres de leur 16è Baromètre de l'Absentéisme® et de l'Engagement 2024".

The study of sickness presenteeism (SP), defined as showing up to work while sick (Johns, 2010), continues to receive growing interest from scientific research in several areas of the social sciences. This is reflected in the very large number of literature reviews and meta-analyses on the subject over time (Patel et al., 2023; Ruhle et al., 2019; Lohaus & Habermann, 2019; Navarro et al., 2019; Kinman, 2019; Skagen & Collins, 2016; Miraglia & Johns, 2016; Garrow, 2016; Johns, 2010; Schultz & Edington, 2007). Many of these studies focus either on the determinants or the consequences of SP. Furthermore, these literature reviews often highlight the extensive use of cross-sectional data in the analysis of SP, as well as the issue of measuring this behaviour. Only a limited number of studies have examined in detail the dynamic relationship between SP and future health, sickness absence (SA) and well-being (Bergström et al., 2009b, 2009a; Hansen & Andersen, 2009; Gustafsson & Marklund, 2011; Dellve et al., 2011; Taloyan et al., 2012; Janssens et al., 2013; Gustafsson & Marklund, 2013; Collins et al., 2017). This chapter is part of this limited body of research. Additionally, studies focusing on the effects of SP rarely examine those of SA. However, both behaviours stem from a single decision by the worker whether or not to attend work in case of illness, constituting two exclusive alternatives (Johansson & Lundberg, 2004). In this sense, the effects of taking time off to recover from an illness and going to work while sick should be contrasted. Therefore, this chapter attempts to demonstrate this relation by analysing both the effects of SA and SP on future illness.

The literature review conducted by Skagen & Collins (2016) on the effects of SP on future health outcomes shows that previous findings are based on diverse data, sometimes drawn from very specific population samples, and use different definitions, measures, and statistical approaches, making comparisons highly difficult. More specifically, all the previous studies only explore the correlations between past SP or SA and the future outcomes. Omitting the effect of endogeneity, which has not been addressed in the previous research, has led the authors to a certain lack of understanding about the mechanisms associated with SA and SP and their relationship with future health outcomes. A debate has emerged on how to consider SA and SP in relation to each other. While intuition tends to consider these two behaviours as substitutes, and therefore negatively correlated, some results indicate that SA and

SP are positively correlated (Aronsson et al., 2000; Caverley et al., 2007; Hansen & Andersen, 2008; Johns, 2011; Aronsson et al., 2011) and that they share common determinants that influence them in the same direction (Gerich, 2016; Arnold & de Pinto, 2015). This leads the authors to argue that they are ultimately complementary. Further, the authors who have identified a positive association between both SA and SP and future health outcomes also corroborate this intuition.

The results of the past literature mainly underline that the two behaviours are negative predictors of future health. It is interesting to note that previous research has always argued in favour of the restorative aspect of SA (Aronsson et al., 2000), while their findings never confirm this intuition. It is assumed that these results were influenced by the endogeneity of SP and SA in relation to the intrinsic health of individuals, such as the well-known issue of the healthy worker effect. This is probably the effect of being in a poorer health, which is a prerequisite for doing SA and SP (Johansson & Lundberg, 2004), that explains the two positive associations with a future health deterioration. Research must therefore reassess previous findings by exploring larger databases using modern statistical approaches, that consider endogeneity, in order to draw reliable conclusions.

This chapter, that looks at the temporal dynamics between the duration of both SP and SA and future health and mental well-being, attempts to fill several gaps identified in the previous literature on this subject. First and foremost, probably the most important contribution, is the identification of causality in the relation of both SA and SP with future individuals' health outcomes. It is proposed to correct for endogeneity using a recent method based on a two stages copula endogeneity correction approach, called 2sCOPE, provided by Yang et al. (2024). This method is particularly interesting in the case of health studies using survey data where no valid instrumental variables are available to researcher. However, the use of copula methods is very rare in economics compared to the traditional IV approach: 4% in applied economics, 2% in human resources and 2% in econometrics (Qian et al., 2024). In doing so, it is argued that the two positive associations found between SP and SA and poorer future health in the past literature suggest that these two behaviours serve as predictors of employees' intrinsic propensity to be ill. Caverley et al. (2007) have already shown that they are complementary indicators to other

health indicators such as self-rated health. Thus, after controlling for endogeneity, which is supposed to capture the robustness or fragility of individuals with regard to illness, the results are supposed to better reflect the act of being absent or present when ill.

Secondly, past results has largely been based on the analysis of small and often specific samples. Only Taloyan et al. (2012) used a representative sample of the Swedish population. This chapter therefore contributes to the limited literature that explores dynamic relationships using longitudinal data on a representative sample of the working population. Using the two waves of the French "Working Conditions and Psychosocial Risks" survey from 2013 and 2016, involving 10,563 respondents in each wave, we can reliably explore the lagged effects of SA and SP.

Finally, it was often pointed out in the literature reviews that the measures used were not entirely satisfactory (Ruhle et al., 2019). Previous studies that analyse SP generally use frequency rather than duration. However, the use of cumulated number of SA and SP days is more economically relevant for companies when evaluating the consequences of each behaviour (Arnold & de Pinto, 2015). Productivity, for example, depends more on the duration of workers' episodes of SP (Pauly et al., 2008). The same applies to the expected effects on future health. In addition, even when the durations were analysed, qualitative indicators were often used in which the duration of SA and SP were dichotomised or categorised. The use of these variables leads to make choices on the thresholds that divide the categories. This may affect the results and do not allow comparisons between studies. Sometimes, even within the same study, comparisons between the effects of SA and SP are not possible due to the use of different variables (e.g. duration for SA and frequency for SP) or different thresholds for SA and SP categories. Therefore, this chapter uses the annual durations of SA and SP, defined in exactly the same way, to compare the relative impact of both behaviours on future outcomes. Further, contrary to the use of categories, this enables to determine the marginal effect of one additional day on future illness and mental well-being.

The results without controlling for endogeneity confirm the previous findings reported in the literature (Skagen & Collins, 2016). Both SA and SP are positively related to an increase in the future durations of SA, SP and total illness (calculated

as the sum of the SA and SP durations). Controlling for endogeneity, the 2sCOPE models provide very different results that have never been shown in previous literature. Lengthening the duration of SA at baseline significantly shortens the future duration of SA and total illness. These effects are monotonically decreasing and support the intuition of a restorative behaviour (Aronsson et al., 2000; Aboagye et al., 2019). At first glance, SP does not appear to have a significant effect on the future duration of SA. However, when looking at its effect on the total duration of illness, it describes a U-shaped curve. Consequently, even after a period of three years, long durations of SP would be related to an increase in the total duration of the illness.

The preliminary descriptive analysis highlights significant heterogeneity among individuals who report a positive duration of SP at baseline. This is confirmed by exploring the heterogeneity of the effects of SA and SP associated with perceived health in the benchmark models. While the effect of the baseline duration of SP is highly heterogeneous with regard to perceived health, this is not the case for the effect of baseline SA. Therefore, when this heterogeneity of perceived health at baseline is taken into account in the estimates of the effects of SP, in addition to controlling for endogeneity, the results ultimately confirm the intuition that longer duration of SP at baseline increases the duration of future illness.

Finally, the regressors generated by the 2sCOPE method to capture endogeneity show the expected signs. They are all positively correlated with future duration of SP, SA and illness and are expected to reflect the underlying health - or ill - state of individuals. Thus, the endogeneity results confirm the intuition that SP, on the contrary to SA, is a risk behaviour for future health and increases illness. This is the first study to jointly demonstrate the harmful effect of SP on the one hand and the restorative effect of SA on the other. This support the initial findings of Kivimäki et al. (2005) who conclude to the benefit of intermediate duration of SA. The control of endogeneity allows to a better rationalisation of the link between the observations and the theoretical intuition.

In addition to the main conclusion, although the correlations between SP duration and mental well-being may lead to ambiguous conclusions, the endogeneity correction models point to a significant deleterious effect of SP on the WHO-5 score,

with heterogeneous effects depending on self-rated health. These models also show a significant and important negative effect of the additional term that capture endogeneity with mental well-being, probably indicating that more fragile workers regarding health report lower score of mental well-being. On the contrary, the duration of SA is not related to future mental well-being.

The present results may have important implications for future research as well as for firm human resources and political policies. They confirm the harmful causal effects of SP on the lengthening of the future illness duration. Thus, SA has to be promoted over SP to protect future individual health. Further, the results confirm the great importance of considering endogeneity in health studies. Firms that want to prevent potential future absences due to illness must especially care about SP behaviour and favour SA instead. In addition to confirm the theoretical intuition, the results allow to obtain marginal effects to predict potential future costs associated to both the duration of SA and SP. For example, when a worker reports being in very good health, for an average SP duration of 2 days at baseline, a one-day increase in SP would result in an average increase in future illness duration of 9.4%. Conversely, a one-day increase in the duration of SA would be associated with a significant decrease in the duration of illness of 1.2% on average. As a result, companies have to balance the current cost of an absent worker against the future benefit of not letting that worker attend ill, thereby increasing the future cost associated with a poorer health and longer absences.

The paper is organized as follows. The section 2 reviews the existing literature that addresses the lagged effects of sickness presenteeism and absence. Section 3 presents the general intuition of the study. Then, section 4 presents the data, the measures and the methodology. Section 5 shows the results and finally, section 6 sets out the concluding remarks.

2 Review of the literature

2.1 The costs of sickness presenteeism and absenteeism

Both of these behaviours are known to be costly for companies, employees and society as a whole (Pauly et al., 2008). Over a given period, the more frequent and

longer the absences of workers, the longer it will take for them to return to their initial level of efficiency and productivity. From the strict point of view of human capital accumulation, the less experience individuals have, the less capital they accumulate to be more productive. For instance, Grinza & Rycx (2020) demonstrate a significant overall impact of SA on productivity in Belgian. For companies, the main cost of SA is that they have to continue paying workers while they are sick, at least in part (Allen, 1981). In addition, if production has to continue, companies have to pay overtime or hire temporary workers. These replacement workers, whose tasks are not their usual tasks, cannot be as productive as the absent worker (Herrmann & Rockoff, 2012). In a longer term, SA was also found to be associated to an increase in turnover intention (Čikeš et al., 2018). Furthermore, if there are complementarities of the absent worker with capital or team members for instance, an additional cost arises as the capital would not operate at its optimal level or the productivity of the team would be affected (Coles et al., 2007; Heywood et al., 2008).

For employees, depending on collective agreements, they may not receive their full wages during their period of illness, which encourages them not to be absent for too long. Chadi & Goerke (2018) show, for example, that SA is associated with a lower probability of being promoted but also with a greater probability of being laid off. Employees can develop strategies using their level of absence to avoid sending a damaging signal to their employer. For example, during a probation period, when job security is not guaranteed, workers appear to be less absent, and then return to a higher level of absence once the probation is over (Ichino & Riphahn, 2005). The same strategy may be applied in a context of downsizing. In these cases, SP would be a way of avoiding sending the wrong signal to the employer in order to protect oneself against the risk of dismissal (Johns, 2010). Furthermore, in the context of career competition, individuals may not want to reveal their vulnerabilities by being absent too often (Simpson, 1998).

The costs of SP are more hidden than those of SA. SP would be associated to a significant decrease in productivity of at least 30% compared to working under normal conditions (Hemp, 2004). More recently, Strömberg et al. (2017) evaluate the productivity loss of SP estimated by managers. They show that the costs for

organisations of both SA and SP may exceed the workers' wage. In comparison, the drop in productivity caused by SP would be largely higher than that only associated to the absent worker. Schultz & Edington (2007) already shown that working while sick leads to a greater drop in productivity than that caused by SA. Evans-Lacko & Knapp (2016) note that in eight countries, the costs associated with SP were five to ten times higher than those associated with SA. Considering short and long term costs of SP, Chou & Mach (2021) show that, although it is associated with an increase in productivity at 1 week, it is also well associated with a decrease in productivity 1 year later. This highlights the potential long-term effects of SP that are not always considered, either by individuals or by firms, at the time the behaviour is performed. Additionally, sick workers may contaminate coworkers leading to an higher drop of the overall team productivity (Barmby & Larguem, 2009), or even more to an increase of the absenteeism rate of the firm.

As a consequence, firms may want to reduce the number of days lost due to sickness. They can implement mechanisms that incentivise workers to be present rather than choosing to be absent. However, their policies can have harmful effect on employees behaviours and health. For example, Marie & Val Castello (2023) study the impact of a reduction in paid sick leave. The authors show a reduction in the frequency of sick leave but an increase in duration. More importantly, they observe a rise in the proportion of relapses in addition to an increase in work accidents. Taken together, SA and SP over time can have significant costs that may be underestimated at first. Finally, the dynamic between SP and SA, which could tend towards a vicious circle, needs to be explored further in order to identify the costs associated with each behaviour.

2.2 The effects of sickness presenteeism and absenteeism over time

To date, there is relatively few studies that explore the health consequences of both SP and SA over time, compared to the large body of the literature about each behaviour. Among the various literature reviews on SP and SA, Skagen & Collins (2016) focus on studies that used longitudinal data to explore the effects of SP on future health outcomes, including self-rated health, SA and well-being. They include

papers that adopted only a prospective study design with at least one follow-up after baseline.

The first article to highlight the deleterious effect of SP on health was that of Kivimäki et al. (2005). The authors show that men who not take sick leave during a 2 years period show a significant increase of future cardiovascular disease compared with whose reporting a moderate amount of sick leave. This first paper already showed the potential restorative effect of absence. SA would normally allow time to recover from physical or psychological problems by promoting health (Aronsson et al., 2000). On the opposite, SP has been found to heighten the risk of bad general health (Bergström et al., 2009a; Dellve et al., 2011; Taloyan et al., 2012; Gustafsson & Marklund, 2013) as well as the risk of future SA (Dellve et al., 2011; Janssens et al., 2013), especially long SA duration of more than 30 days (Bergström et al., 2009b) and of more than 90 days (Helgesson et al., 2021).

The articles cited above, with the exception of that by Gustafsson & Marklund (2013) and Helgesson et al. (2021) which study categories combining SP and SA, add a control for initial absenteeism in their models. Although the authors do not focus on interpreting the results associated with this control, when the results are presented, they show a positive relationship between past and future SA (Bergström et al., 2009b, 2009a; Gustafsson & Marklund, 2011; Janssens et al., 2013). Furthermore, Helgesson et al. (2021) show that individuals with high SA and SP together have higher risk to report long term SA during the 3 years follow-up period compared to individuals who only report SA or SP. A result that confirm previous findings by Gustafsson & Marklund (2013) who found higher risk of physical complaints for workers involved in high SA and SP at the same time. Interestingly, in line with the above comments, the results of Helgesson et al. (2021) highlight that individuals who report high SA with low SP show higher risk of future long term SA compared to those who report high SP and low SA. This confirms once again the conclusions of Gustafsson & Marklund (2013) showing that SA would represents a higher risk factor for future health than SP. These results reveal the counter-intuitive effect associated with resting from SA on health.

Concerning mental health, Demerouti et al. (2009) show the existence of a reciprocal relationship between SP and emotional exhaustion over time. SP favoured

by past emotional exhaustion would aggravate this exhaustion six months later. A result recently confirmed by Aboagye et al. (2019) who found an association between actual exhaustion and SP the previous year, but not with past SA. However, subsequently, this fatigue can translate into longer-term SA (Akerstedt et al., 2007). If SP is indeed a risk behaviour for exhaustion over time, pushed to its climax, it has also been shown that it was significantly associated with burnout two years later (Dellve et al., 2011). On the contrary, an efficient time off is associated with a better state of recovery (van der Pol & Irvine, 2018). Finally, mixed results were found by Collins et al. (2017) which indicate that, although people with SP in the previous three months showed lower psychological well-being, there was no significant association over time. In addition, they also found no associations over time between SA and any future outcomes, neither SA, reported health or well-being.

Even if we can conclude that a consensus has been reached regarding the fact that SP at baseline is indeed a risk factor for future sickness absence and lower self-rated health, the review of Skagen & Collins (2016) well exposes the various measures and methodological choices on which these findings are based. All the papers presented above have in common the fact of analysing the effects of predefined modalities, mostly based on the frequency or sometimes on the duration of SA and SP. They are therefore interested in the probability of seeing an effect on a certain threshold of frequency or duration of absence. For instance, Bergström et al. (2009b) found a risk increase for future SA of more than 30 days for individuals who show at least five episodes of SP at baseline, while Taloyan et al. (2012) found an effect for seven or more days of SP on reporting more than seven days of SA two years latter. Comparisons between studies regarding the effects of SA and SP are therefore difficult. Even within the same study, it is not always possible to compare the effects of SA with those of SP because of the different measures used. For example, many studies use the duration of SA, whereas they use the frequency of SP (Bergström et al., 2009b, 2009a; Hansen & Andersen, 2009; Gustafsson & Marklund, 2011; Dellve et al., 2011; Janssens et al., 2013; Gustafsson & Marklund, 2013). In addition, Taloyan et al. (2012), who use durations of SA and SP, do not construct the categories in the same way. The lowest modality for SA contains individuals who report up to seven days, including zero, whereas for SP, the lowest modality only includes individuals

with zero days. Only one paper, that of Collins et al. (2017), uses the durations of both behaviours and defines the variables modalities in the same way.

More importantly, all the reviewed studies base their conclusions solely on correlations and, as already noted, on the fact that SA and SP are just two predictors of individuals' health. Accordingly, researchers all agree that firms and policymakers should be concerned with workplace health in order to reduce both SA and SP, as they represent risk factors for long-term health. However, these studies do not allow us to highlight the mechanisms associated with the behaviours themselves or the effects associated with the underlying illness behind the behaviours. Further, there is no study that is able to promote SA over SP to favour rest behaviour.

3 Empirical intuition of the effects of SA and SP over time

Johansson & Lundberg (2004) rightly explain that poor health is a prerequisite for SA and SP behaviours and that these behaviours are the two exclusive alternatives for the worker. In this sense, the two behaviours are thus substitutes by definition. In addition, a common feature of the past literature is the intuition that the choice of individuals to be absent or present in case of illness is strongly linked to the type of illness encountered. Gosselin et al. (2013) show for instance that gastritis, insomnia or allergies are some of the illnesses for which people are willing to go to work instead of being absent. On the contrary, asthma, blood pressure problems and thyroid trouble often lead to SA. However, the authors also add that comorbidity and stress are associated to SP rather than to SA. Recently, Malmberg-Ceder et al. (2020) identified that recurrent headache is associated to SP but not to SA. In addition, mental disorders are shown to be related to high levels of SP while, on the contrary, people suffering from physical disorders are more likely to report a high level of SA (Helgesson et al., 2021). However, some studies have reached different conclusions. The most common reasons reported by workers for SA and SP would ultimately be very similar (Kinman, 2019), such as ordinary illness (flu, cold, gastroenteritis, etc.), chronic pain, headache, sleeping and mental disorders (Baptista et al., 2019) or depression (Maestas et al., 2020).

In fact, what is important in the choice of SA or SP is the individuals' own perception of the severity of illness for which they will decide to attend work (Lohaus & Habermann, 2019). Moreover, individuals' choice depends on the beliefs they have that they can achieve their objectives by taking SA or SP (Lohaus & Habermann, 2021). Therefore, the more severe the illness is perceived to be in relation to the objectives that the worker wishes to achieve, the greater the disutility of effort.

Health, and energy at large, may be considered as a resource that the employees manage over time. In this sense, individuals are producers of their own health (Grossman, 1972). Thus, health represents a stock that can be depreciated, but also in which individuals can invest. Furthermore, in line with the intuition regarding the determination of an optimal level of effort in a tournament context, the absence of rest can deteriorate the individual's energy resources (Amegashie et al., 2007). This idea is also supported by the Effort-Recovery Theory proposed by Meijman & Mulder (1998). By definition, SP is a demanding behaviour in terms of effort, and the lack of recovery would lead to depreciate the health stock. Consequently, the longer the duration of SP, the higher the probability of starting with a lower stock of health in the future. The consequences are firstly to increase the probability of falling ill and, secondly, that this illness will be more serious, leading to more frequent and/or longer absences. On the contrary, SA that allows workers to rest is supposed to restore workers' stock of health and energy. For instance, van Hooff et al. (2018) indicate that recovery efforts are needed in case of deficit in the energy level to restore individual's homeostasis in psychological well-being and abilities. This intuition has been observed in empirical literature, with SP shown to be associated with future exhaustion whereas SA is not (Aboagye et al., 2019). Thus, given the illness shock, the longer the duration of SA, the greater the restoration of the health stock. The intuition that will be tested in this chapter is detailed below.

It is thus assumed that individuals have an initial stock of health θ. When faced with a shock of illness δ, they will choose between taking time off work, SA, or going to work despite the illness, SP. Depending on two reservation health thresholds, δ^{sp} and δ^{sa}, the worker will choose SA or SP regarding the level of the illness shock δ (Brown & Sessions, 2004). If $\delta^{sp} < \delta < \delta^{sa}$ then they will choose SP because the perceived illness will not be felt to be serious enough to justify absence from work,

whereas if $\delta > \delta^{sa}$ then the individual will choose SA because of the perceived severity of the illness. The two thresholds, δ^{sp} and δ^{sa} are determined by the individual each time the choice arises, i.e. each day that has to be worked and when the individual is faced with a deteriorated state of health. We assume that we have $\delta^{sa}(\theta)$ and $\delta^{sp}(\theta)$ where both are increasing functions of θ. The greater the individual's stock of health, the higher these thresholds will be. In addition, they may vary according to individual characteristics, employment characteristics and the working conditions at the time the decision is made. Finally, in a dynamic context, these thresholds incorporate the past choices of SA and SP and the anticipated consequences of the current choice.

When the worker suffers a shock of illness, this damages his stock of health θ. Further, the more severe the illness, the longer the duration T of this illness. For the total duration of illness, the worker will have to choose between SA and SP T times. Moreover, it is assumed that the only way to protect and restore the stock of health is through rest, i.e. choosing SA. Otherwise, if the worker chooses SP, this will prevent him from restoring his health, leading to deteriorate his health stock. Therefore, we set $C(SA_k, SP_j, \delta_T)$ the total cost function of choosing SA k days and SP j days over the total duration T of the illness (with $j + k = T$). The first derivatives are given as $C'_{SA_k} < 0$, $C'_{SP_j} > 0$ and $C'_{\delta_T} > 0$. The longer the duration of SA, the higher the worker recovers from the cost of the illness shock. On the contrary, the longer the duration of SP, the more the worker damages his health capital. Indeed, the optimal choice of SA will depend of the costs associated with the illness shock and the sick pay awarded during the days of absence.

Suppose that the individual only chooses SP over the entire duration of the illness. His health stock at the end of the illness period is therefore $\theta - C(\delta_T) - \sum_{j=1}^{T} C(SP_j)$. It means that for a future illness shock, his health stock will be lower. Since $\delta^{sa}(\theta)$ is an increasing function of θ, when θ is lower, the threshold $\delta^{sa}(\theta)$ is also lower. According to this reasoning, the longer the duration of SP in the past, the lower the health stock and the more likely the future duration of SA will be longer. In other words, the more SP has been chosen in the past, the greater the chance that

In the economic literature, θ is generally seen as the worker's capacity. This assumes that the health shock will reduce the worker's capacity. Thus, the cost of effort, which is deceasing with capacity, would be greater given the illness shock (e.g. if the effort cost function is as follows: $c(e) = e/\theta$).

the sickness shock appearing above the threshold and leading to choose SA more regularly. It should be noted that this is also the case for the lower threshold δ^{sp}, which determines the feeling for which a sickness will be declared and leads to the choice of SP. In this way, if the two illness thresholds are affected in the same way by θ, the whole area between the two thresholds shifts downwards. Otherwise, it is quite possible that the functional form linking the health stock to each threshold is not the same. If, for example, $\delta^{sa}(\theta)$ is more sensitive than $\delta^{sp}(\theta)$ to a deterioration in θ, then the decision threshold to do SA will move closer to the decision threshold to do SP, thus reducing the likelihood of choosing SP.

Among the literature that deals with SA and SP, no paper to date has empirically demonstrated the relationship developed in the above intuition. Most papers have found negative effects of both SA and SP on future health. However, this does not lead to the conclusion that SA is detrimental to future health. This only reflects that SA is a predictor of illness. The longer the duration of SA, the poorer the health of the individual (Sanatkar et al., 2025). These empirical results indicate that it is the workers who are most vulnerable to illness who are likely to have longer SA both currently and in the future and to see their health deteriorate even further. It is this type of results that leads researchers to consider SA and SP as complementary behaviours. For instance, Gosselin et al. (2013) who observe a positive correlation between SP and SA argue that they corroborate the results of Johns (2011) and Hansen & Andersen (2009) who support a potential complementarity. The implied complementarity of the behaviours stems from their intrinsic link with illness (Johansson & Lundberg, 2004). It has been shown that in addition to SA, SP is also a good, and even better, predictor of illness, making these two behaviours complementary health indicators (Gustafsson & Marklund, 2013; Caverley et al., 2007).

This problem of the underlying state of health of the individual at the origin of SA and SP generates a traditional endogeneity problem known as the healthy worker effect. This effect indicates that workers self-select into jobs that are in line with their state of health and their resistance to effort (Haan & Myck, 2009). The most robust individuals self-select (or are selected) for the most demanding jobs in terms of working conditions. Conversely, workers who know they are more vulnerable to

illness are those who will self-select into jobs that allow them to be absent more easily. Thus, as mentioned above, an increase in SA over time would reflect more severe illnesses that would persist for the sickest individuals. With regard to SP, it also becomes impossible to dissociate the health effect associated with individual fragility from the deleterious effect of the behaviour.

It is therefore necessary to use methods that control for endogeneity when estimating the effects of SA and SP on future outcomes. In addition to properly estimating the effects of SA and SP, this would make it possible to compare their respective effects. Under the hypothesis that these two behaviours are not performed for the same degree of severity of illness, the effect of one day of SP cannot be compared with that of one day of SA if the population choosing SA is very different from that choosing SP in terms of perceived illness. Here again, controlling for the endogeneity associated with illness would make the results more comparable. According to the intuition and the empirical issues developed above, this chapter will answer to the following hypotheses:

Hypothesis 1. *Both SA and SP, as predictors of health, show a positive correlation with future health outcomes.*

Hypothesis 2. *Unobservable characteristics that generate endogeneity reflect individual propensity to cope with illness and explain the positive association between past SP or SA and future health outcomes.*

Hypothesis 3. *After controlling for endogeneity, a longer duration of SA would lead to better recovery and therefore a better future health, i.e. a shorter total duration of illness and duration of SA, a better self-rated health and greater mental well-being.*

Hypothesis 4. *After controlling for endogeneity, a longer duration of SP would lead to a worsening of future health, i.e. a lengthening of the total duration of illness, of the duration of SA, and a worsening of self-rated health and mental well-being.*

4 Data and methodology

4.1 Data and sample of study

To explore the link between past sickness absence and presenteeism (SA) on future health and absenteeism, two waves of the 'working conditions and psychosocial risks' survey from 2013 and 2016 are used. This sample allows for the examination of the effect of SP and SA at baseline on health outcomes after three years. This survey is provided by the French Ministry of Labour (Dares) and the French National Institute for Statistics and Economic Studies (Insee). The survey is conducted face-to-face in the respondent's home by an interviewer. Respondents are also asked to complete a self-assessment questionnaire. This second part of the survey deals with more sensitive subjects such as physical and mental health problems, relations with colleagues and superiors, tensions within the company, feelings about the content of the work, etc. Overall, the two surveys initially covered 17,144 respondents representative of the French population, thus constituting an exceptional data set covering a large part of the population over a long period of time. However, specific requests has been made by the Ministry to focus on the healthcare and public sectors. Given these specific requests, this focus has led to a substantial over-representation of these populations in the longitudinal data. In fact, 55% of the sample is made up of people from the healthcare sector. The following study will indeed control for sector and separated analyses will be provided in additional results.

Some restrictions were imposed in order to obtain a reliable sample free of outliers. Only active employees under contract for the entire period, aged over 18 in 2013 and under 60 in 2016, earning between 500 and 8,000 euros a month and working between 10 and 100 hours a week are included. Further, senior executives (CEOs) are also excluded. Finally, in order to avoid including people with serious chronic diseases or long-lasting illnesses (ALD) in the sample, the duration of sickness absenteeism has been limited to a maximum of 90 days (corresponding to 98% of the distribution) and sickness presenteeism to a maximum of 60 days (corresponding to more than 99% of the distribution). After removing the missing observations in the variables of interest, the final sample is made up of 10,563 individuals observed over the two periods, constituting a balanced panel.

4.2 Measures

The following section describes the different measures used to study SA and SP in relation to health, illness and mental well-being. As stated by Johns (2010), and more recently by Biron et al. (2022), the two behaviours must be measured by an open-ended fill-in-the-blank response to be completed by indicating the number of days that the respondents were present or absent when they were faced with a shock of illness. An important contribution of this research to the literature is also the use of count data in number of days for SA and SP. The latest literature reviews (Skagen & Collins, 2016; Ruhle et al., 2019) point out the complexity associated with comparing the results of previous research due to the measures employed. It is important to highlight two points on this subject. Firstly, although a number of studies have made available the duration of SA, it is common for the authors to transform their variable into a qualitative variable with two or more modalities. This means that choices have to be made regarding the thresholds defining the categories, which in most cases differ between studies, and also leads to a loss of information. The second important point concerns the measurement of SP. While it is common in the case of SA to report the duration, a number of studies have focused on the frequency of SP rather than its duration (Bergström et al., 2009b, 2009a; Hansen & Andersen, 2009; Gustafsson & Marklund, 2011; Dellve et al., 2011; Janssens et al., 2013; Gustafsson & Marklund, 2013; Chou & Mach, 2021). This choice has been strongly criticised as being tainted by a significant lack of information. The number of episodes alone does not indicate the duration, and therefore the importance, of each episode. Furthermore, when information on the duration of SP is available, as is the case for SA, the use of multi-modal qualitative variables is the most common (Taloyan et al., 2012; Conway et al., 2014; Collins et al., 2017; Aboagye et al., 2019; López-Bueno et al., 2021). According to Ruhle et al. (2019), the 'results regarding presenteeism propensity and count measures should be compared to draw valid conclusions'. However, previous studies have mainly used propensity measures, and few of them have adequately examined the duration of behaviours. Therefore, this chapter proposes to fill this gap by using the following measures.

Sickness presenteeism and absenteeism duration

In the present study, we focus on the effects of SA and SP, both measured in the same manner, namely the number of days over the past year. Sickness presenteeism is measured as the number of days an individual was going to work while sick instead of taking a sick leave. The related question is "*over the past twelve months, have you ever worked (gone to work) when you thought you shouldn't (should have stayed home) because you were sick?*". If respondents say they have experienced at least one episode of SP, they are asked to estimate the cumulative number of days this represents: "*How many times has this happened in the past twelve months?*", counted in days. Our SP measure is thus the aggregated annual duration of the phenomenon. The employed question is the same to that from the European Working Conditions Survey (Arnold & de Pinto, 2015). A similar question has been used in most previous reference publications (Bergström et al., 2009b, 2009a; Hansen & Andersen, 2009; Demerouti et al., 2009; Dellve et al., 2011; Taloyan et al., 2012).

Sickness absence is measured by the number of days individuals are absent from work because of sickness. Respondents are asked to report how many times they had taken sick leave over the past twelve months. If they experienced at least one episode of sick leave, they were also asked to evaluate their cumulative number of SA days in the past year: "*How many days of absence correspond to these sick leaves?*".

These measures will make it possible to evaluate the marginal effect of an additional day of SA or SP on our outcomes and to directly assess the relative importance of each behaviour. In addition, as stated by Arnold & de Pinto (2015), the use of the durations is supposed to be more informative from an economic point of view (Pauly et al., 2008). Knowing the unitary cost of one day of SA and SP, it allows companies to estimate their relative cost of SP and SA. Although the duration of the two behaviours has already been studied previously in the economic literature (Arnold & de Pinto, 2015; Gerich, 2016; Hirsch et al., 2017), these studies do not address the same question. Consequently, this analysis, in addition to improving our understanding of the link between the choice of absence or presence in the event of future illness and absence, is part of a body of literature restricted to the study of the duration of SA and SP in the same population.

Both SA and SP measures are self-reported and may be subject to measurement

bias, in particular to memory and perception bias. However, except for sick leave, all the studies about SP actually used the same auto-declared informations. However, Caverley et al. (2007) found a strong correlation (of 0.9) between self-reported and recorded absences. Several other studies that could check the correspondence between reported and recorded data have also shown good agreement between the two types of information, thereby validating self-reported data overall (Voss et al., 2008; Gustafsson & Marklund, 2011). In addition, concerning the data used in this chapter, Hillion (2024) matches the same data with administrative records and does not found significant difference between self-reported SA duration and administrative ones.

Illness, perceived health and mental well-being

Although SA and SP are the two behaviours of interest, in order to be consistent with existing literature, the following health indicators are also included as outcomes in the analysis. First, general self-rated health (SRH) is measured using the following question: "*How would you rate your general state of health?*". The levels are: "bad", "very bad", "rather good", "good" or "very good". The "bad" and "very bad" categories are combined because of the very low number of individuals reporting very poor health. For the purpose of the study and to follow previous literature, SRH will be dichotomized when considered as the dependent variable. The propensity of poor health is equal to 1 if SRH takes the modalities 'rather good" or "bad or very bad", zero otherwise.

Since the SRH variable is supposed to indicate the feeling of the individuals with their average health state, the WHO5 score is also considered in order to best reflect the different facets of individual mental well-being and health. The WHO5 score is measured by five questions with five modality and thus ranges from 0 to 25 where a low score indicate poor mental well-being and a high score a very good mental well-being in life .

Finally, it has been shown that SA, SP and self-reported health are complementary health indicators. Caverley et al. (2007) show that SA finally explains only 3%

The questions include in the WHO5 score are: I have felt cheerful in good spirits; I have felt calm and relaxed; I have felt active and vigorous; I woke up feeling fresh and rested; and my daily life has been filled with things that interest me. Respondents are asked to scaled from 0 (at no time) to 5 (all of the time).

of the total variability of the SRH variable and that the introduction of SP leads to increase the explained variability to 13%. Aronsson et al. (2011) confirm this conclusion, indicating that SA and SP must be considered together with SRH. However, considering the SA and SP variables separately always leads to confusion between the decision to choose SA or SP and the severity of the underlying illness. Therefore, as a complementary indicator to SRH, the total duration of illness is calculated as the sum of the durations of SA and SP. This is intended to reveal the burden of illness over the past year, regardless of whether these sick days were worked or not. This helps to complete the overall picture of an individual's health.

Control variables

As it has been widely identified in the literature, a large number of individual and occupational characteristics are associated with SP and SA. In particular, some of these characteristics, as shown in the first chapter, can have a significant impact on a worker's decision to engage in SP rather than SA, and vice versa. Thus, when studying the link between SA or SP and health outcomes, it is advisable to control for a set of characteristics that can affect both SA and SP but also these health outcomes. In other words, potential bias due to omitted variables should be controlled for to isolate as far as possible the effect associated with the behaviours themselves. The effect of SP should not, for example, reflect the effect of working conditions imposing an attendance constraint, or conversely, the existence of flexible absence conditions. To this end, a set of characteristics previously identified in the literature are introduced as control variables in this study. These characteristics can be classified into four groups: i) individual characteristics, ii) employment contract characteristics, iii) working conditions and iv) environmental conditions in the broad sense (work environment and the worker's social environment). These four groups of variables are described below.

The individual characteristics are age, gender, education, marital status and having dependent children under 18. The average age of the sample was 46 in 2016; females represented 55.6%, and 63.45% had a baccalaureate or higher degree. In addition, although the sample was restricted based on SA or SP duration, a binary indicator is included alongside health and mental well-being measures, to control for

whether individuals report a chronic disease, which accounts for 28% of the sample. The characteristics of the employment contract are seniority, hourly wage, sector of activity, occupational position, being a supervisor, working part-time, atypical working hours and being subject to quantitative targets. There seems to have been some improvement in careers over the period. Hourly wages rose by 4.38%, the number of supervisors rose from 12.41% to 13.62% and the proportions of higher professional categories increased, particularly in the private sector, from 9.26% to 10.53% for engineers and executives. These statistics are presented in Table A1 in appendix.

The controlled working conditions are: work paced by a machine, external demand, norms, colleagues, computer or hierarchical control, work under pressure and lack of time. The employee's environment is controlled by variables such as the size of the company, receiving help from colleagues or superiors, being uncertain about the future of one's job and being satisfied with promotion prospects. Finally, intrinsic motivation, which may influence the cost-benefit balance of SP over SA, will be captured by four variables: esteem and pride in one's work, pride in one's company and the feeling that one's work is useful to others. To consult the corresponding statistics, see Table A2 in the appendix section. In general, the working conditions to which employees are exposed do not vary greatly. Surprisingly, there has been a significant reduction in pressure at work and fear of losing one's job. In addition, more individuals say they are not concerned by promotion opportunities and there seems to be more of them who are satisfied by these opportunities. Consistent with these observations, the proportion of individuals who receive esteem from their work and are proud of their job increases significantly. These few descriptive findings may be explained by career progression over the period and therefore greater job security and satisfaction.

4.3 Empirical methodology

To broadly explore the effect of SA and SP on future health outcomes, the analysis is proposed to be conducted in two stages. First, the empirical methodology used in this chapter is a direct continuation of previous works on the subject (Bergström et al., 2009b, 2009a; Hansen & Andersen, 2009; Aronsson et al., 2011; Gustafsson &

Marklund, 2011; Taloyan et al., 2012). The longitudinal nature of the data makes it possible to test the long-term effect of the behaviours using lagged models. In addition, the observation of lagged effects of SA and SP on health initially secures the direction of causality that we wish to observe. However, these models in themselves do not solve the problem of endogeneity associated with the unobservability of initial conditions, capabilities and the robustness of individuals regarding illness. Therefore, in a second stage, endogeneity correction models using a copula approach are proposed to explore the causal relations between SP, SA and future health outcomes. Indeed, as previously explained, the objective of this study is to disentangle the impact of the individual's underlying health status from the influence of the decision to attend work or not when ill.

Benchmark models

The benchmark models presented below will, to a certain extent, enable us to compare our results with those of previous literature. The models specification is given by the following equation:

$$Y_{it} = \alpha + \delta_1 SP_{it-1} + \delta_2 SP_{it-1}^2 + \delta_3 SA_{it-1} + \delta_4 SA_{it-1}^2 + \delta_5 SRH_{it-1} + \mathbf{X}_{it}'\beta + \mu_{it} \quad (2.1)$$

When Y_{it} represents either SA, SP, or illness duration, we propose estimating ordinary least squares (OLS) regressions on the logarithm of these variables. This enables the estimation of the semi-elasticity of SP and SA on the outcomes. To correctly apply the logarithm to these variables, we calculate the logarithm of the duration plus one. Applying the ordinary least squares (OLS) method to log-transformed variables seems reasonable. For example, this strategy was previously used by Gosselin et al. (2013). Furthermore, the skewness and kurtosis values of each transformed variable are within the acceptable range (see Kline (2011) for a discussion of acceptable values for normal distribution). When Y_{it} is the WHO5 score, OLS are directly estimated without any transformation. When the dependent variable corresponds to the binary indicator of poor health, a logistic model is estimated.

The proposed model aims to study the quadratic effect of both sickness presenteeism (SP) and sickness absenteeism (SA). Given that no previous study has used the duration of the behaviours as a continuous variable, there are no empirical ex-

amples to justify this choice. Only Taloyan et al. (2012) could indicate a non-linear effect. However, as the highest category is "greater than 7 days", it is not possible to infer a non-linearity for long durations. This gap in the literature in itself justifies the study of quadratic effects. Intuitively, the longer the duration of SP, the greater the harmful effect. Nevertheless, it is possible that the expected positive relationship between SP and future illness may have decreasing returns. The same applies to the effect of the duration of SA. According to the intuition developed previously, the longer the duration of SA, the greater the recovery of the health stock. However, it is also possible to observe a decreasing effect in this case.

As already explained, SA, SP and self-reported health (SRH) would be complementary health indicators (Caverley et al., 2007; Aronsson et al., 2011). Furthermore, after calculating the VIF , we can observe that there is no significant effect of multicollinearity between the variables SA, SP and SRH in each model. This justifies the inclusion of SRH as a control in the models. In addition, it is assumed that this variable can capture the explanatory part associated with workers' health status that could pass through the effects of SA and SP behaviours.

$\mathbf{X}_{it}$ is the matrix that includes all the controls. In each model, the sets of controls will be successively included as defined above, i.e. i) individual characteristics, ii) i) plus the employment characteristics and iii) ii) plus the working conditions and the environmental characteristics. Finally, μ_{it} corresponds to the error term.

A copula endogeneity correction approach

In this section, we present the methodology used to take into account the potential endogeneity between the realisation of the SP or SA behaviours and the error term. To do this, we propose to explore an IV-free copula approach. According to the guidelines of Qian et al. (2024), we do not employ the original method proposed by Park & Gupta (2012) but rather the two stages copula endogeneity correction (2sCOPE) developed by Yang et al. (2024). The copula approach strategy is entirely appropriate in the context of occupational health research, where we are assessing the behaviours of individuals for whom a significant part of their variability can be attributed to unobservable and intrinsic characteristics such as their robustness to

Variance inflation factor.

illness, their abilities or their risk preferences over work. In addition, the use of a copula to model the link between the endogenous regressor and the error term is particularly useful because, in this context, the lack of appropriate information for estimating instrumental variable models prevents their use. Furthermore, unlike the IV approach, the copula one is less restrictive and will not suffer from a weak instrument problem or a wrong choice of instrument. This is particularly relevant in the context of the ongoing debate regarding the reliability of the choice of instrumental variables. Indeed, choosing the wrong instrument can introduce an even greater bias into the estimators than those obtained by the use of the ordinary least squares (OLS) estimator (Keane & Neal, 2023). Moreover, Eckert & Hohberger (2023) compared the results of the IV approach and the IV-free copula method proposed by Park & Gupta (2012). They show that the P&G method gives equivalent or even better results than the IV estimates.

Therefore, the 2sCOPE approach and the different assumptions required for its validity as recently detailed in Qian et al. (2024) is presented below. First of all, to better understand the concept, we will present the seminal work on endogeneity correction using copula proposed by Park & Gupta (2012). Then, since recent research has led to significant improvements in the methodology proposed by P&G (Qian et al., 2024), we present the advantages of the 2sCOPE approach (Yang et al., 2024).

Originally, the method proposed by Park & Gupta (2012) first involves estimating the empirical marginal distribution of the endogenous regressor P, denoted $H(P)$, where in our case P corresponds to either SP or SA. Secondly, using the marginal distribution of the error $G(\mu)$, one can derive the joint distribution function $F(P, \mu)$. So, according to Sklar's theorem (1959), there exists a copula function that links the two margins such as $F(P, \mu) = C(H(P), G(\mu), \theta) = C(U_P, U_\mu, \theta)$. For its desirable properties, Park & Gupta (2012) use the Gaussian copula. In this case, the dependence parameter θ can be converted into correlation coefficient ρ between the endogenous regressor and the error term and captures the endogeneity of P. Thus, the joint distribution function is given by:

$$F(P, \mu) = C(U_P, U_\mu, \rho) = \Psi_\rho \left(\Phi^{-1}(U_P), \Phi^{-1}(U_\mu) \right) \tag{2.2}$$

where Φ is the standard univariate normal distribution function and Ψ_ρ denotes the standard bivariate normal distribution with the correlation coefficient ρ. If ρ is significantly different from zero, this highlights that P is endogenous. According to the assumption that μ follows a normal distribution ($\mu \sim \mathcal{N}(0, \sigma_\mu)$), Park & Gupta (2012) demonstrate that the structural error is composed of $\mu = \sigma_\mu \mu^* = \sigma_\mu \rho P^* + \sigma_\mu \sqrt{1-\rho^2}\omega$ where $\mu^* = \Phi^{-1}(U_\mu)$ and $P^* = \Phi^{-1}(U_P)$. Therefore, the first part, $\sigma_\mu \rho P^*$, captures the endogeneity between P and μ and the second part, $\sigma_\mu \sqrt{1-\rho^2}\omega$, is an error term which is independent of P. Another practical point of this approach is the possibility to handle multiple endogenous regressors. In our case, where SP and SA are both endogenous, we have the following joint trivariate normal distribution:

$$\begin{pmatrix} SP^* \\ SA^* \\ \mu^* \end{pmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho_{12} & \rho_{\mu 1} \\ \rho_{12} & 1 & \rho_{\mu 2} \\ \rho_{\mu 1} & \rho_{\mu 2} & 1 \end{bmatrix} \right) \tag{2.3}$$

The interesting advantage in this model is the consideration of ρ_{12} which is the correlation between both endogenous regressors, in addition to the correlations of the regressors with the error term ($\rho_{\mu 1}$ and $\rho_{\mu 2}$). Finally, considering the specification with two endogenous regressors, SA and SP in our case, the method consists in the maximization of the log-likelihood with regards to the parameters of the following model:

$$\begin{aligned} Y_{it} = \alpha + \delta_1 SP_{it-1} + \delta_2 SP_{it-1}^2 + \delta_3 SA_{it-1} + \delta_4 SA_{it-1}^2 + \delta_5 SRH_{it-1} + \mathbf{X}'_{it}\beta + \\ \sigma_\mu \frac{\rho_{\mu 1} - \rho_{12}\rho_{\mu 2}}{1-\rho_{12}^2} SP^* + \sigma_\mu \frac{\rho_{\mu 2} - \rho_{12}\rho_{\mu 1}}{1-\rho_{12}^2} SA^* + \sigma_\mu \sqrt{1 - \rho_{\mu 1}^2 - \frac{(\rho_{\mu 2} - \rho_{12}\rho_{\mu 1})^2}{1-\rho_{12}^2}} \cdot \omega \end{aligned} \tag{2.4}$$

Park & Gupta (2012) suggest an augmented OLS estimation of this model with SP^* and SA^* as additional regressors to correct for endogeneity. This method is subject to several assumptions (Qian et al., 2024). Firstly, either the error term μ or the calculated regressor P^* has to be normally distributed. Secondly, either the couple (P, μ) or (P, P^*) follows a Gaussian copula. Thirdly, the regressor matrix must be of full rank. Fourthly, the endogenous regressor must have a non-normal distribution.

If the error and the endogenous regressor both follow a normal distribution, the endogenous regressor P and its empirical margin P^* are perfectly collinear, which violates the full rank assumption. Finally, the control variables must be exogenous and independent of the empirical margin of the endogenous regressor.

The aim of the present study is to model the effect of past SA and SP durations on different outcomes such as the log(SA days), log(SP days), log(Illness days) and the WHO5 score. Given the distribution of these outcomes, the estimated models do not necessarily lead to a normally distributed error. However, Qian et al. (2024) indicate that "the copula control function methods (including the P&G) do not require a normal error distribution or Gaussian copula regressor-error dependence and can be derived under substantially weaker conditions". This remark also holds for the following 2sCOPE method. The estimator can be derived as long as the endogenous part of the error (P^*) is normally distributed and jointly follows a Gaussian copula dependence structure with P and the control variables. This leads to robust inference even in the case that the error term is non-normal. The full rank condition and the exogeneity of the covariates are assumed to be true, just as they are in the OLS or IV models. In addition, our regressors of interest, the durations of SA and SP, are largely non-normally distributed and can, given their range, easily be considered as continuous variables. The assumption that is most likely to be violated in our case is the independence of the covariates with the endogenous regressors and their empirical margins. Research improving the P&G method shows that the correlation between the exogenous and endogenous regressors can negatively affect the effectiveness of the correction provided by the copula in eliminating the endogeneity bias (Qian & Xie, 2024; Yang et al., 2024). The 2sCOPE approach, by Yang et al. (2024), allows to relax the uncorrelatedness assumption between the endogenous regressor and the exogenous ones as well as the non-normality of the endogenous regressor. Given that our models take into account a wide range of covariates likely to affect both SA, SP and health outcomes, 2sCOPE is of particular interest in our case. For all of these reasons, the 2sCOPE methodology proposed by Yang et al. (2024) is chosen rather than the one initially developed by Park & Gupta (2012).

The idea behind 2sCOPE is to jointly model the endogenous regressor with the exogenous ones and the error term using a Gaussian copula model. They remove from the control function P^* the component which is correlated with the exogenous regressors, which leads to isolating the part of P^* that really controls the endogeneity, cleaned of its association with the control variables. Similarly to the original method, this implies that P^*, $\mathbf{X}^*$ and μ^* follows a multivariate normal distribution, where $\mathbf{X}^*$ corresponds to the matrix of all the empirical margins of the exogenous variables, i.e. all the $x_k^* = \Phi^{-1}(L(x_k))$ are calculated for each exogenous control k. Adding P^* and all the x_k^* in the initial model allow to capture endogeneity without requiring the exogeneity of the control variables and their independence with either SA or SP.

While Park & Gupta (2012) suggest integrating the kernel density estimate (IKDE) to obtain the marginal cumulative distribution function (CDF) of $H(P)$ and $G(\mu)$, the modern approaches mostly employ the non-parametric rank-based copula transformation based on the empirical marginal distributions of regressors. The copula transformation is adjusted for the maximum value of P since the inverse normal CDF of the 100th percentile is undefined. This transformation is thus carried out in two stages as follows:

$$P_i^* = \Phi^{-1}(H(P_i)) = \begin{cases} \Phi^{-1}(Rank(P_i)/n) & \text{if } P_i < max(P) \\ \Phi^{-1}(n/(n+1)) & \text{if } P_i = max(P) \end{cases} \tag{2.5}$$

The inclusion of many generated regressors associated with the control variables (all the x_k^*) can lead to significant multicollinearity problems. However, Yang et al. (2024) propose a two-stage strategy to overcome this issue. First, after obtaining SP^*, SA^* and all the x_k^*, the following models are regressed:

$$SP^* = \mathbf{X}^{*'}\rho_{\mathbf{x}1} + \varepsilon_{sp} \tag{2.6}$$

$$SA^* = \mathbf{X}^{*'}\rho_{\mathbf{x}2} + \varepsilon_{sa} \tag{2.7}$$

Then, by conditioning the initial model in equation 2.1 on ε_{sp} and ε_{sa}, μ becomes independent of SP and SA as well as on all the correlations with the controls. Thus, the second step consist in the estimation of the augmented model that includes both

ε_{sa} and ε_{sp}, instead of SA^*, SP^* and the x_k^* to address the endogeneity bias. The endogeneity-free model is defined as follows:

$$Y_{it} = \alpha + \delta_1 SP_{it-1} + \delta_2 SP_{it-1}^2 + \delta_3 SA_{it-1} + \delta_4 SA_{it-1}^2 + \delta_5 SRH_{it-1} + \mathbf{X}'_{it}\beta + \delta_{sa}\varepsilon_{sa} + \delta_{sp}\varepsilon_{sp} + \mu'_{it} \tag{2.8}$$

2sCOPE includes P&G as a special case, where $\rho_{\mathbf{x}1} = \rho_{\mathbf{x}2} = 0$. Thus, $\varepsilon_{sp} = SP^*$ and $\varepsilon_{sa} = SA^*$. In their guidelines, Qian et al. (2024) clearly explain that in the case of the inclusion of a quadratic effect, which is the case here, it is not necessary to consider this regressor as endogenous and apply the copula correction as if it were a second endogenous regressor. In fact, controlling the endogeneity with the main term is sufficient to capture and isolate the endogenous part of the regressor with the error. Moreover, the authors explain that computing an additional copula associated with the quadratic term is not a good way to identify the effect. They argue that "this substantially decreases precision of coefficient estimates, and makes copula correction methods perform worse than otherwise". Finally, 2sCOPE is also suitable when interaction terms with endogenous regressors are included.

5 Results

The results are presented as follows. First, the section below presents general descriptive statistics and provides some indications of the heterogeneity associated with the health of individuals who engage in SP. Second, the inferential analysis will first present the results of the standard models, without controlling for endogeneity, which reproduce the conclusions given in the previous literature. Then, the main results of the 2sCOPE models with endogeneity control are presented. Finally, as the descriptive statistics tend to highlight strong heterogeneity among the population with a positive duration of SP, the results of the effects interacting with self-rated health are presented.

5.1 Descriptive analysis

All the following statistics are reported in Table A3 in the appendix section . On average, 42.25%, respectively 45.52%, of the workforce reported SP at least once in 2013 and 2016. For these two periods, the average duration of SP was 2 and 2.52 days. For comparison, around 28.86% of the workers declared SA in 2013 with an average duration of 4.25 days while they are 33.26% in 2016 with an average duration of 5.33 days. Overall, there was an increase in the number of workers affected by illness over the two periods, from 54.61% to 59.04%. This is accompanied by an increase in the duration of illness from 6.26 days to 7.85 days. In general, self-reported health has also declined. We observe a transfer from the upper category of health to the lower ones. For instance, 26.04% of individuals declared themselves to be in very good health in 2013, compared with 21.27% in 2016. At the opposite, 18.84% said they were in rather good health in 2013, compared with 22.46% in 2016. The proportion of individuals reporting a chronic disease also increased significantly, from 25% to 28%.

Tables A4 and A5 in the appendix section present additional descriptive statistics providing information on the evolution of the link between self-rated health and both SA and SP behaviour between 2013 and 2016. For each health category, the proportion of individuals who engage in SP is systematically greater than the proportion who engage in SA. As expected, the duration of each behaviour increased significantly with the deterioration in self-rated health. However, the duration of SA increases more quickly than that of SP with the deterioration of perceived health, especially in 2016. Furthermore, individuals who reported poor or very poor health in 2013 showed a SA rate of 50.49% with a duration of 10.42 days, compared with a rate of 67.74% with a duration of 18.49 days. These figures clearly indicate a general deterioration in health over the three-year observation period.

What interests us now is to explore a potential difference in health status underlying the choice of SA and SP. To do that, we identify four groups with regard to the workers' choice of doing only SA, only SP or both alternately over the last year. These groups are expected to be very different from each other with regard to the

The descriptive statistics are not based on the 10,563 individuals used for the regressions but on the sample of 12,885 individuals, before removing individuals with missing data on hourly wage and WHO5 score.

underlying illness. For example, with regard to the statistics in Table 2.1, there is a significant difference between the average duration of SA and that of SP. Individuals who reported only SP during the past year had an average illness duration of 4.5 days, while those who reported only SA had an average duration of 16 days. This may indicate that, on average, the illnesses for which workers choose SP are less serious and resolve more quickly than those that lead them to choose SA. In addition, the total duration of illness for those who report both SA and SP is 21 days. While the average duration of SA for this group is not significantly different from that of individuals who only report SA, the average duration of SP is significantly higher than that of individuals who only take SP. Workers who report both SA and SP have certainly faced more serious illnesses that have led to a higher cumulative number of choices between SA and SP. In addition, there are significantly more individuals with at least two episodes of absence during the year among those who reported both SA and SP than among those who reported only SA. This could indicate a dynamic trade-off between SA and SP for individuals who experience multiple illness shocks during the year, a point we cannot examine using our data.

Table 2.1 – Descriptive statistics of health characteristics by behaviour group for SA and SP

Characteristic	Overall N = 25,770[1]	Healthy N = 11,126[1] (1)	SP only N = 6,641[1] (2)	SA only N = 3,335[1] (3)	p-value[2] (2)-(3)	SA&SP N = 4,668 (4)	p-value[2] (2)-(4)	p-value[2] (3)-(4)
SP days	2.26 (5.38)	0.00 (0.00)	4.56 (6.19)	0.00 (0.00)	<0.001	6.01 (8.24)	<0.001	<0.001
SP bracket					<0.001		<0.001	<0.001
No SP	14,461 (56.47%)	11,126 (100.00%)	0 (0.00%)	3,335 (100.00%)		0 (0.00%)		
1 to 3 days	6,887 (26.89%)	0 (0.00%)	4,325 (65.77%)	0 (0.00%)		2,562 (56.02%)		
4 to 7 days	2,449 (9.56%)	0 (0.00%)	1,370 (20.83%)	0 (0.00%)		1,079 (23.60%)		
1 to 2 weeks	1,431 (5.59%)	0 (0.00%)	702 (10.68%)	0 (0.00%)		729 (15.94%)		
2 to 4 weeks	382 (1.49%)	0 (0.00%)	179 (2.72%)	0 (0.00%)		203 (4.44%)		
SA days	4.79 (13.14)	0.00 (0.00)	0.00 (0.00)	15.94 (21.08)	<0.001	15.07 (18.82)	<0.001	0.4
SA bracket					<0.001		<0.001	0.059
No SA	17,767 (68.94%)	11,126 (100.00%)	6,641 (100.00%)	0 (0.00%)		0 (0.00%)		
1 to 3 days	2,223 (8.63%)	0 (0.00%)	0 (0.00%)	969 (29.06%)		1,254 (26.86%)		
4 to 7 days	2,041 (7.92%)	0 (0.00%)	0 (0.00%)	840 (25.19%)		1,201 (25.73%)		
1 to 2 weeks	1,598 (6.20%)	0 (0.00%)	0 (0.00%)	623 (18.68%)		975 (20.89%)		
2 to 4 weeks	1,129 (4.38%)	0 (0.00%)	0 (0.00%)	468 (14.03%)		661 (14.16%)		
1 to 3 months	1,012 (3.93%)	0 (0.00%)	0 (0.00%)	435 (13.04%)		577 (12.36%)		
SA episode					<0.001		<0.001	<0.001
No SA	17,767 (68.96%)	11,126 (100.00%)	6,641 (100.00%)	0 (0.00%)		0 (0.00%)		
1	5,834 (22.64%)	0 (0.00%)	0 (0.00%)	2,639 (79.18%)		3,195 (68.47%)		
2	1,424 (5.53%)	0 (0.00%)	0 (0.00%)	502 (15.06%)		922 (19.76%)		
3 or more	741 (2.88%)	0 (0.00%)	0 (0.00%)	192 (5.76%)		549 (11.77%)		
Illness days	7.05 (14.79)	0.00 (0.00)	4.56 (6.19)	15.94 (21.08)	<0.001	21.08 (21.35)	<0.001	<0.001
Health					<0.001		<0.001	<0.001
bad or very bad	877 (3.40%)	94 (0.84%)	236 (3.55%)	126 (3.78%)		421 (9.02%)		
rather good	5,322 (20.65%)	1,493 (13.42%)	1,592 (23.97%)	709 (21.26%)		1,528 (32.73%)		
good	13,476 (52.29%)	5,994 (53.87%)	3,566 (53.70%)	1,766 (52.95%)		2,150 (46.06%)		
very good	6,095 (23.65%)	3,545 (31.86%)	1,247 (18.78%)	734 (22.01%)		569 (12.19%)		
Chronic disease	6,810 (26.43%)	1,993 (17.91%)	1,906 (28.70%)	1,026 (30.76%)	0.033	1,885 (40.38%)	<0.001	<0.001
WHO5 score	17 (13, 19)	18 (15, 20)	15 (11, 18)	17 (14, 20)	<0.001	14 (10, 18)	<0.001	<0.001

[1]n (%); Mean (SD) or Median (IQR)
[2]Pearson's Chi-squared test; Wilcoxon rank sum test

With regard to self-rated health, when comparing the two groups with exclusive behaviour with the group of individuals reporting both SA and SP, it is clear that the latter group is in significantly poorer health than the others. The proportion of individuals reporting very poor or poor health is 9% in this group, compared to 3.55% and 3.78% in the other two groups. In contrast, only 12.2% of those who report both SA and SP declare themselves to be in very good health, compared to 18.8% of those who report only SP and 22% of those who report only SA.

From a descriptive point of view, we can therefore observe that the group of individuals who report both SA and SP is indeed a group of individuals in poorer health than those who only do SA and more than those who only do SP. These results highlight that looking at the duration of SP, workers may be very heterogenous regarding their health status. On the one side, there are healthy individuals who report short duration of illness for which, probably given its low severity, they decide to work during the entire period which represent 4.5 days in average. On the other side, the worst healthy individuals who cumulate both a longer average duration of SP of 6 days with an average duration of SA of 15 days. However, when studying the duration of SP only, it is not possible to capture this health heterogeneity.

To complete the descriptive analysis provided above, the Table 2.2 extends the study of being in a specific group on the future health outcomes of the following inferential analysis. This may answer to the question of persistent effect that can be attributed to the underlying bad individual's health. With regard to the results of the model in column (4), for which all individual characteristics, employment and working conditions are controlled, workers who have previously declared SA only or both SA and SP see their future duration of SA increase, contrary to the individuals who only did SP. On the contrary, in column (5), workers who have been involved in SP in the past are more likely to see their future duration of SP increase. Although this is also the case for workers who have only done SA in the past, the level of the coefficients is much lower (almost 5 to 6 times lower). Finally, the previous assumptions are confirmed here, with a certain graduation in the effect of the categories on the future total duration of illness in column (6). Individuals who had only done SP were associated with a smaller increase in the total duration of illness than those who had only done SA, who themselves were associated with a

smaller increase than those who had done both in the past.

An interesting result is that workers who have done either SP only or SP with SA in the past have a greater probability of declaring themselves to be in poorer health in the future and of seeing their mental well-being deteriorate, which is not the case for individuals who have only done SA. These results could be the first signs of a beneficial effect of SA on the recovery of health, showing no significant effect on any deterioration in future health and well-being.

Table 2.2 – Associations between group of baseline behaviour regarding SA and SP and outcomes

	Dependent variable:							
	log(SA days)$_t$				log(SP days)$_t$	log(Illness days)$_t$	low SRH$_t$	WHO5$_t$
	OLS				*OLS*	*OLS*	*probit*	*OLS*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SP only$_{t-1}$	**0.068*****	0.054*	0.038	0.031	**0.416*****	**0.366*****	**0.077****	**−0.647*****
	(0.026)	(0.028)	(0.028)	(0.028)	(0.021)	(0.028)	(0.038)	(0.108)
SA only$_{t-1}$	**0.428*****	**0.427*****	**0.417*****	**0.410*****	**0.079*****	**0.403*****	0.062	−0.023
	(0.037)	(0.040)	(0.040)	(0.040)	(0.025)	(0.040)	(0.049)	(0.134)
SA&SP$_{t-1}$	**0.596*****	**0.582*****	**0.557*****	**0.538*****	**0.499*****	**0.734*****	**0.093****	**−0.796*****
	(0.034)	(0.037)	(0.037)	(0.037)	(0.026)	(0.035)	(0.046)	(0.133)
rather good SRH$_{t-1}$				−0.069	−0.106*	−0.115	**−0.267*****	0.597*
				(0.083)	(0.059)	(0.076)	(0.084)	(0.305)
good SRH$_{t-1}$				−0.155*	**−0.154*****	**−0.205*****	**−0.921*****	**1.787*****
				(0.081)	(0.058)	(0.074)	(0.081)	(0.296)
very good SRH$_{t-1}$				−0.139*	**−0.161*****	**−0.211*****	**−1.213*****	**2.895*****
				(0.083)	(0.059)	(0.077)	(0.088)	(0.305)
Constant	**1.331*****	**1.377*****	**1.218*****	**1.344*****	**1.529*****	**2.243*****	**0.816*****	**14.709*****
	(0.088)	(0.111)	(0.140)	(0.159)	(0.108)	(0.152)	(0.202)	(0.562)
Controls								
Individual	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Employment	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Work conditions and environment	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,218	10,563	10,563	10,563	10,563	10,563	10,563	10,563
R^2	0.090	0.098	0.105	0.106	0.238	0.199		0.276
Adjusted R^2	0.089	0.095	0.101	0.101	0.234	0.194		0.272
Log Likelihood							−4,419	

Note: *p<0.1; **p<0.05; ***p<0.01

Robust standard errors in parentheses. Individual's controls include age, sex, marital status, children, education, WHO5 and chronic disease. Employment characteristics include seniority, supervisor, part-time, atypical schedules, hourly wage, quantitative targets, occupation, sectors and firm size. Working conditions and environment include work paced by machine, external demand, norms, computing control, hierarchical control or colleagues, work pressure, lack of time, help from colleagues or supervisors, fear of job future, organizational changes, promotion opportunities, work esteem, useful job, job and firm pride.

5.2 Inferential analysis

Results of the benchmark models

The first part of this study presents the estimates of the benchmark models, as presented in equation (2.1), which examine the effect of the duration of SA and SP at baseline on health outcomes three years later. The objective is to retrieve the results from previous literature that does not control for endogeneity between behaviours

and health outcomes. These results are reported in Table 2.3 below. Columns (1) to (4) present the results on the duration of SA, column (5) on the duration of SP and column (6) on the total duration of illness (the sum of the duration of SA and SP). Then, column (7) presents the results on the probability of declaring a health state lower than good and column (8) presents the results on mental well-being. The estimates in columns (1) to (3) control successively for individual characteristics, job characteristics, working conditions and the worker's environment. In column (4), the estimates are controlled for the reported health status of participants at baseline . The effects on SA, SP and illness durations are estimated by OLS on the log-transformed outcomes. However, the use of more complex models, such as the ZINB models, which make it possible to treat count data with zero inflation, leads to similar conclusions.

Table 2.3 – Estimates of the effects of SA and SP durations at baseline on health outcomes

	Dependent variable:							
	log(SA days)$_t$				log(SP days)$_t$	log(Illness days)$_t$	low SRH$_t$	WHO5$_t$
	OLS				*OLS*	*OLS*	*probit*	*OLS*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SP days$_{t-1}$	**0.026*****	**0.023*****	**0.021*****	**0.018*****	**0.072*****	**0.065*****	0.006	**−0.121*****
	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)	(0.005)	(0.006)	(0.019)
SP days$^2_{t-1}$	**−0.00045*****	**−0.00042*****	**−0.00038*****	**−0.00034****	**−0.00120*****	**−0.00111*****	−0.00011	**0.00238*****
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0005)
SA days^{t-1}	**0.032*****	**0.032*****	**0.031*****	**0.030*****	**0.006*****	**0.026*****	0.001	−0.007
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.009)
SA days$^2_{t-1}$	**−0.00038*****	**−0.00036*****	**−0.00035*****	**−0.00035*****	**−0.00008*****	**−0.00031*****	0.00001	0.00018
	(0.00004)	(0.00004)	(0.00004)	(0.00004)	(0.00003)	(0.00004)	(0.00004)	(0.0001)
rather good SRH$_{t-1}$				−0.035	−0.033	−0.037	**−0.263*****	0.556*
				(0.085)	(0.060)	(0.078)	(0.086)	(0.306)
good SRH$_{t-1}$				−0.133	−0.077	−0.132*	**−0.924*****	**1.748*****
				(0.084)	(0.059)	(0.077)	(0.084)	(0.299)
very good SRH$_{t-1}$				−0.131	−0.095	**−0.161****	**−1.219*****	**2.889*****
				(0.086)	(0.061)	(0.080)	(0.090)	(0.307)
Constant	**1.467*****	**1.416*****	**1.241*****	**1.350*****	**1.519*****	**2.251*****	**0.921*****	**14.348*****
	(0.087)	(0.118)	(0.147)	(0.167)	(0.114)	(0.161)	(0.212)	(0.584)
Controls								
Individual	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Employment	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Work conditions and environment	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,218	10,563	10,563	10,563	10,563	10,563	10,563	10,563
R^2	0.077	0.089	0.096	0.097	0.233	0.191		0.277
Adjusted R^2	0.076	0.085	0.090	0.091	0.228	0.185		0.272
Log Likelihood							−4,408	

Note: *p<0.1; **p<0.05; ***p<0.01

Robust standard errors in parentheses. Individual's controls include age, sex, marital status, children, education, WHO5 index (when is not the outcome) and chronic disease. Employment characteristics include seniority, supervisor, part-time, atypical schedules, hourly wage, quantitative targets, occupation, sectors and firm size. Working conditions and environment include work pace constrained by machine, external demand, norms, computing control, hierarchical control or colleagues, work pressure, lack of time, help from colleagues or supervisors, fear of job future, organizational changes, promotion opportunities, work esteem, useful job, job and firm pride.

The VIF has been calculated and does not show any increase of the indicator after the introduction of the past SRH.

Most of the coefficients in the models reported in Table 2.3 are highly significant. Only the parameters associated with both SA and SP on the probability of declaring a health state below good are not significant, revealing no significant effect of past SA and SP on perceived health. Finally, unlike SP, the duration of SA had no significant effect on future mental well-being. The introduction of the different sets of controls in the models on the duration of SA has relatively little effect on the results, especially for the past duration of SA, showing reliable estimates. In contrast, the coefficients associated with SP are slightly more sensitive to the introduction of the different controls, thus indicating a stronger association between SP and the control variables.

The results confirm hypothesis 1 by showing that the durations of both SA and SP behaviours are associated with an increase in the future duration of SA and SP as well as the total duration of future illness. The estimates show that the association between past and present SA is much greater than that between past SP and current SA. Further, the two behaviours appear to have a similar quadratic effect, showing a decreasing effect as the duration of the behaviours increases. This inverted U-shape is found in each model, on the duration of SA, SP and illness. This indicates that, beyond a certain turning point, located at 30 days for SP and 42 days for SA, an additional day of SP or SA leads to a decrease in the future duration of illness. Taking into account the quadratic term, the effect of an additional day of SP, for an average duration of 2 days of SP at baseline, would be to increase the future duration of SA by 1.68%. Similarly, the effect of an additional day of SA, for an average duration of 5 days at baseline, would be an increase of 2.7%.

In addition to the estimates on the future duration of SA, the estimates on the future duration of SP are presented. Although there is no clear theoretical support for the idea that past duration of SA and SP can influence the future choice of taking up SP, specifically over a duration as long as 3 years, this model reveals a correlation that is essential to an overall understanding of the mechanism. The estimates in column (5) show that an increase in the duration of past SP is associated with a much greater increase in future duration of SP than that associated with the past duration of SA. In contrast, the increase in SA duration at baseline has a much smaller effect on future SP than on future SA. The effect of an additional day of

SP, for a SP duration of 2 days at baseline, is associated with a 6.8% increase in the future duration of SP, while an increase of one day of SA, for an initial duration of 5 days of SA, is associated with a 0.5% increase.

Looking at the effect on the total duration of future illness, the results show a greater association between past SP, compared to past SA, and the future total duration of illness. In fact, an additional day of SP, for an average duration of SP of 2 days at baseline, would correspond to an increase of 6.1% in future illness. An increase of one day of SA, for an initial duration of 5 days of SA, would be associated with an increase of 2.3% in the total duration of future illness. The worker's duration of investment in SP behaviour would therefore be more deleterious to future health than the duration of SA. Finally, only the duration of SP is associated with a significant decline in mental well-being. Moreover, this effect decreases with the increase in the duration of SP. Thus, for a duration of 2 days of SP at baseline, one additional day of SP at baseline would be associated with a 0.113 reduction in this score, representing a decrease of 0.66%.

In summary, the results show that a longer duration of SA at baseline, compared to the effect of an increase in the duration of SP at baseline, is associated with a greater increase in the duration of SA three years later. Furthermore, the exact opposite occurs when examining the effects of longer initial durations of SA and SP on future SP duration. Therefore, as we do not control for potential endogeneity, these results simply seem to indicate that past behaviours are good predictors of future behaviours. This confirms the first hypothesis proposed in section 3. Furthermore, these positive correlations may be due to a deterioration in the underlying health status. Thus, whether it concerns SA or SP, it may be that the health status of individuals has worsened over time. If the intuition that individuals who engage in SP are on average healthier than those who engage in SA is correct, it means that, over time, the sickest individuals would choose SA more often, while the least sick individuals would continue to choose SP. Without controlling for endogeneity associated with the healthy worker effect, the results could reflect both the consequences of behaviours and the propensity of individuals to be ill.

Results of the two stages copula endogeneity correction models

Previous estimates suggest that both SA and SP are positively correlated with future illness. None of the above models identify a possible restorative effect of SA on future health. The positive effect of SA at baseline on the future duration of SA and illness is not consistent with the intuition of a restorative behaviour as expressed in section 3. Furthermore, although the positive effect of SP at baseline on future SA and illness is consistent, the issue of endogeneity remains and may explain much of this effect. Therefore, the subsequent models, which explicitly account for the endogeneity of SA and SP behaviours, seek to assess the extent to which the effects previously observed in the benchmark analysis are attributable to endogeneity.

Table 2.4 – Estimates of the 2sCOPE models

	Dependent variable:							
	log(SA days)$_t$				log(SP days)$_t$	log(Illness days)$_t$	low SRH$_t$	WHO5$_t$
	OLS				*OLS*	*OLS*	*OLS*	*OLS*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	**1.934*****	**1.770*****	**1.564*****	**1.999*****	**2.307*****	**3.343*****	**0.863*****	**13.092*****
	(0.114)	(0.141)	(0.162)	(0.208)	(0.137)	(0.187)	(0.058)	(0.707)
SP days$_{t-1}$	−0.014	−0.010	−0.006	−0.007	**−0.054*****	**−0.052*****	−0.002	0.062
	(0.015)	(0.014)	(0.014)	(0.015)	(0.012)	(0.014)	(0.005)	(0.054)
SP days$^2_{t-1}$	0.00009	0.00005	0.00001	0.00001	**0.0006*****	**0.00055*****	0.00003	−0.00023
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0001)	(0.001)
SA days$_{t-1}$	**−0.016****	**−0.016****	**−0.016****	**−0.016****	−0.001	**−0.013****	−0.001	0.020
	(0.007)	(0.007)	(0.007)	(0.007)	(0.004)	(0.006)	(0.002)	(0.022)
SA days$^2_{t-1}$	**−0.00016*****	**−0.00015*****	**−0.00015*****	**−0.00015*****	−0.00004	**−0.00013*****	0.00001	0.00005
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.000)	(0.000)	(0.000)	(0.0002)
SP^* 1st stage residuals	**0.194****	**0.167****	0.136*	0.128*	**0.690*****	**0.630*****	0.022	**−0.995*****
	(0.076)	(0.072)	(0.072)	(0.075)	(0.058)	(0.070)	(0.024)	(0.270)
SA^* 1st stage residuals	**1.029*****	**0.999*****	**0.984*****	**0.966*****	0.130*	**0.817*****	0.024	−0.542
	(0.128)	(0.122)	(0.123)	(0.124)	(0.073)	(0.115)	(0.035)	(0.401)
corr(SP^*, μ)	**0.124*****	**0.112*****	**0.095*****	**0.077*****	**0.389*****	**0.267*****	0.030	**−0.118*****
	(0.031)	(0.030)	(0.030)	(0.030)	(0.027)	(0.026)	(0.031)	(0.030)
corr(SA^*, μ)	**0.365*****	**0.353*****	**0.347*****	**0.331*****	**0.097*****	**0.296*****	0.030	−0.063
	(0.037)	(0.036)	(0.037)	(0.037)	(0.035)	(0.035)	(0.039)	(0.039)
Controls								
Individual	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Employment	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Work conditions and environment	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Health at baseline	No	No	No	Yes	Yes	Yes	Yes	Yes
Observations	10,563	10,563	10,563	10,563	10,563	10,563	10,563	10,563
sigma	1.271	1.257	1.248	1.242	0.875	1.255	0.368	4.298
R^2	0.345	0.352	0.356	0.357	0.538	0.592	0.482	0.935
Adjusted R^2	0.344	0.349	0.352	0.352	0.535	0.589	0.478	0.934

Note: *p<0.1; **p<0.05; ***p<0.01

Bootstrapped standard errors in parentheses. Individual's controls include age, sex, marital status, children, education, WHO5 index (when is not the outcome) and chronic disease. Employment characteristics include seniority, supervisor, part-time, atypical schedules, hourly wage, quantitative targets, occupation, sectors and firm size. Working conditions and environment include work pace constrained by machine, external demand, norms, computing control, hierarchical control or colleagues, work pressure, lack of time, help from colleagues or supervisors, fear of job future, organizational changes, promotion opportunities, work esteem, useful job, job and firm pride.

Table 2.4 above presents the results of the models that control for endogeneity as defined in equation (2.8). In addition to the usual coefficients, this table presents the coefficients associated with the first-stage residuals of SA^* and SP^* as well as the correlation coefficients between the endogenous part of SA and SP and the

model residuals. In this respect, regarding the results of the 2sCOPE models on the duration of SA, SP and illness, reported in columns (4) to (6), all the correlation coefficients between the unobservables are significantly positive. These correlations indicate that both behaviours share common unobservable characteristics that generate endogeneity, supporting the use of the 2sCOPE models. However, in column (4), when all the controls are introduced, the parameter associated to the first stage residuals from the SP^* equation is no longer significant at 5% but it is at the 10% level. The same applies to the parameter associated with the residuals from the first stage of the equation SA^* in column (5), which is only significant at 10%. According to the specification where the explained variable is a lagged explanatory variable, this introduces an endogeneity bias by construction. Thus, the terms associated with the residuals from the first stage of SA^* on the future duration of SA and that of SP^* on the future duration of SP have a significant and important effect. This is also the case for the model (6) where both the first stages residuals of SA and SP are significantly associated to future illness duration.

The parameters of the first stage residuals of SA^* and SP^* in columns (4) to (6) are all positive. This indicates that the unobservable characteristics associated with past and current behaviours that create endogeneity are positively related. This supports the intuition that these terms reflect the underlying poor health of the two behaviours. Specifically, the term capturing the endogeneity of SP appears to be weaker than that for SA on the total duration of illness in column (6), 0.63 versus 0.82. This could also support the idea that people who report SA are in poorer health, or are confronted to more severe illness, than those who report SP, which would lead to a stronger endogeneity effect for SA than for SP. The more vulnerable workers are to illness, the stronger the link between baseline SA, compared to baseline SP, and future illness.

These results therefore lead to the validation of the hypothesis 2, according to which the terms capturing endogeneity in the models appear to capture unobservable characteristics that increase both the risk of a longer duration of SA and SP at baseline and a longer future duration of illness three years later.

Even more important are the results concerning the effects of behaviours. When we control for endogeneity between SA and SP behaviours at baseline and the fu-

ture duration of SA, SP and illness, the results change radically from the previous findings. They reveal a significant negative effect of the lengthening of the duration of SA at baseline on the future duration of SA. Both terms, linear and quadratic, appear to be negative and significant. It should be noted that the results of models (1) to (4) associated with this estimation method are highly stable. Thus, for an average duration of 5 days, an additional day of SA at baseline would be associated with a 1.7% decrease in the future duration of SA, rather than a 2.7% increase as indicated by the estimates reported in Table 2.3. These findings lead to the validation of the hypothesis 3 according to which, once the endogeneity associated with individual intrinsic health is controlled for, a longer duration of SA in the past would be restorative for health and thus lead to a shorter duration of SA in the future. This conclusion is also validated by the estimates in column (6) on the total duration of illness. In this case, an additional day of SA would be associated with a 1.2% reduction in the total duration of illness, given an average SA duration of 5 days at baseline.

The results concerning the increase in the duration of SP at baseline on the future duration of both SA and illness are surprising. While the correlation between the unobservable factors that explain SP at baseline and the future SA duration is significantly positive, and the parameter of the residuals from the first stage of the SP^* equation also tends to be significant, the duration of SP has no significant effect on the lengthening of the future SA duration. This confirms that the correlation coefficients between SP, respectively SA, and the error term could well be explained by the same characteristics. Once these characteristics are controlled for, the behaviour of SP would have no direct effect on the future duration of SA. However, the intuition developed in Section 3 clearly assumes that SP behaviour reduces health capital and thus the worker's intrinsic ability to cope with a future illness shock. Therefore, while SP may not have a direct effect, the results tend to show that it may have an indirect effect through unobservable characteristics.

Furthermore, the estimates on the future duration of SP and illness in columns (5) and (6) indicate that SP at baseline would have a decreasing effect up to an inflection point at 45 and 47 days, respectively. Beyond this point, an additional day of SP would be associated with an increase in the duration of SP and future

illness. As with the findings on the future duration of SA, these results concern the direct effects of SP behaviour and should be considered in addition to the effects associated with the lengthening of both the future SP and illness durations through unobservable factors. However, the negative effect proper to the behaviour for the shorter SP durations should be justified.

Given that the estimated effect of the duration of SP at baseline is free of endogeneity, this effect would not be related to health status. However, SP behaviour may be related to work pressure or signalling issues, either to obtain a promotion or to avoid being laid off. Therefore, if workers invest in SP to cope with workload during a difficult period, once that period has passed, they may reduce their investment in SP and thus its duration. Similarly, if SP is used to signal motivation and investment in the job, once the period of dismissal has ended or the promotion has been obtained, it would no longer be necessary to invest in SP behaviour. However, the U-shaped curve leads to indicate a deleterious effect of SP on future illness when workers invest too much in this behaviour.

The linear probability model of reporting a poorer health state shows an unsignificant correlation between both behaviours and the model residuals (see column (7) in Table 2.4). Since the coefficients are also unsignificant, it confirms that there is no long-term association between past SA or SP and a lower perception of one's health, either through the observed measures or the unobservables. This would probably highlight that individuals assess their own health on the basis of recent events that may not be related to older illness shocks.

Furthermore, estimations reported in column (8) confirms that there is no correlation between the unobservables associated with the duration of SA, through SA^*, and the unobservables associated with future mental well-being. Since the coefficients associated with SA at baseline and its square are both unsignificant, we can conclude that there is no direct association between SA and mental well-being. On the contrary, the correlation coefficient between SP^* and the error term μ in the mental well-being equation is significant, indicating the presence of endogeneity. In addition, the first stage residuals of the SP equation also shows a significant and negative effect on future mental well-being. However, the parameters associated with SP and its square are not significant. This suggests that this is the unobserv-

able characteristics that explain the decline in mental well-being rather than the behaviour of SP itself.

Perceived health heterogeneity

Earlier, descriptive statistics showed that individuals who only did SP had a short duration of illness and were probably the least ill. On the contrary, individuals who reported having done both SP and SA in the past year were those with the poorest health and a much longer duration of illness. We can therefore suspect a certain heterogeneity in the effect of SP. Indeed, among individuals who engage in SP at baseline, it is possible that two radically different groups of individuals, in terms of health, coexist. Within the distribution of the duration of SP, the least ill individuals, who only report SP for relatively short episodes of illness, would be opposed to the most ill individuals, who report SP in addition to SA for longer periods of illness. Although previous models control for endogeneity at the mean, the presence of heterogeneity could also explain this apparent U-shaped curve in the effect of SP on future illness.

In order to explore the potential effects of this heterogeneity in more detail, interaction terms between SP and self-reported health categories at baseline are introduced into the 2sCOPE models. While health heterogeneity is strongly suspected with regard to the duration of SP, it is not suspected for the duration of SA, or at least not to the same extent. Furthermore, the results associated with self-reported health heterogeneity in the benchmark models, which are presented in Table A6 in the appendices, suggest that the effect of the duration of SA at baseline is not heterogeneous. Conversely, they show significant heterogeneity in the effect of the duration of SP. Finally, with regard to heterogeneous effects in the 2sCOPE models, when the durations of SA and SP are interacted with self-rated health, the effects of SA disappear overall (see Table A7 in the appendix section). As there does not appear to be any heterogeneity in the effect of the duration of SA, adding interaction terms may lead to over-identify the model and consequently to unsignificant results. Thus, Table 2.5 below presents the results of the 2sCOPE models only with the interaction between the duration of SP and the self-reported health categories.

Therefore, as might be expected, the results presented in Table 2.5 show sig-

Table 2.5 – Estimates of the 2sCOPE models with self-rated health interactions

	Dependent variable:						
	$\log(\text{SA days})_t$			$\log(\text{SP days})_t$	$\log(\text{Illness days})_t$	low SRH_t	WHO5_t
	OLS			*OLS*	*OLS*	*OLS*	*OLS*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Constant	**2.556*****	**2.341*****	**2.156*****	**2.529*****	**3.624*****	**0.908*****	**12.121*****
	(0.217)	(0.236)	(0.247)	(0.190)	(0.244)	(0.074)	(0.863)
rather good health	**−0.711*****	**−0.651*****	**−0.615*****	**−0.966*****	**−1.180*****	**−0.158*****	**2.594*****
	(0.196)	(0.185)	(0.173)	(0.146)	(0.179)	(0.058)	(0.677)
good health	**−0.872*****	**−0.811*****	**−0.769*****	**−1.073*****	**−1.369*****	**−0.386*****	**3.786*****
	(0.196)	(0.188)	(0.174)	(0.151)	(0.180)	(0.058)	(0.688)
very good health	**−0.931*****	**−0.862*****	**−0.806*****	**−1.111*****	**−1.443*****	**−0.418*****	**4.986*****
	(0.196)	(0.191)	(0.176)	(0.153)	(0.183)	(0.059)	(0.691)
SP days_{t-1}	−0.054*	−0.048	−0.043	**−0.122*****	**−0.127*****	−0.010	0.196*
	(0.032)	(0.030)	(0.030)	(0.028)	(0.032)	(0.010)	(0.111)
SP days^2_{t-1}	0.00063	0.00054	0.00046	**0.00151*****	**0.00149*****	0.00008	−0.00029
	(0.0005)	(0.0004)	(0.0004)	(0.0004)	(0.0005)	(0.0001)	(0.002)
$(\text{SP days}\times\text{rather good SRH})_{t-1}$	**0.094*****	**0.086*****	**0.080*****	**0.174*****	**0.188*****	0.010	**−0.307*****
	(0.032)	(0.031)	(0.030)	(0.027)	(0.032)	(0.010)	(0.114)
$(\text{SP days}^2\times\text{rather good SRH})_{t-1}$	**−0.00124****	**−0.00117****	**−0.00106****	**−0.00213*****	**−0.00226*****	−0.00006	0.00149
	(0.0005)	(0.0005)	(0.0005)	(0.0004)	(0.0005)	(0.0002)	(0.002)
$(\text{SP days}\times\text{good SRH})_{t-1}$	**0.091*****	**0.082****	**0.073****	**0.205*****	**0.210*****	0.014	**−0.297****
	(0.033)	(0.032)	(0.031)	(0.029)	(0.033)	(0.010)	(0.117)
$(\text{SP days}^2\times\text{good SRH})_{t-1}$	**−0.00118****	**−0.00103****	−0.00089*	**−0.00288*****	**−0.0028*****	−0.00011	0.00197
	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0002)	(0.002)
$(\text{SP days}\times\text{very good SRH})_{t-1}$	**0.089****	**0.078****	**0.068****	**0.233*****	**0.229*****	0.006	**−0.332*****
	(0.036)	(0.034)	(0.032)	(0.031)	(0.035)	(0.010)	(0.117)
$(\text{SP days}^2\times\text{very good SRH})_{t-1}$	−0.00121*	−0.00105	−0.00091	**−0.00383*****	**−0.00356*****	−0.00007	0.00298
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.0001)	(0.002)
SA days_{t-1}	**−0.016****	**−0.016****	**−0.016****	0.0001	−0.013*	−0.001	0.018
	(0.007)	(0.007)	(0.007)	(0.004)	(0.006)	(0.002)	(0.021)
SA days^2_{t-1}	**−0.00016*****	**−0.00015*****	**−0.00014*****	−0.00004	**−0.00013*****	0.00001	0.00006
	(0.0000)	(0.0001)	(0.0001)	(0.0000)	(0.0000)	(0.0000)	(0.0002)
SP 1st stage residuals	0.243	0.210	0.184	**1.158*****	**1.042*****	0.097	**−2.906*****
	(0.218)	(0.192)	(0.193)	(0.162)	(0.199)	(0.066)	(0.755)
SA 1st stage residuals	**1.014*****	**0.984*****	**0.972*****	0.115	**0.809*****	0.026	−0.529
	(0.133)	(0.123)	(0.133)	(0.079)	(0.117)	(0.037)	(0.399)
$\text{corr}(SP^*, \mu)$	**0.016****	**0.015****	0.013*	**0.083*****	**0.053*****	0.012	**−0.036*****
	(0.008)	(0.007)	(0.008)	(0.009)	(0.007)	(0.008)	(0.008)
$\text{corr}(SA^*, \mu)$	**0.342*****	**0.332*****	**0.328*****	0.066	**0.282*****	0.031	−0.055
	(0.039)	(0.036)	(0.040)	(0.040)	(0.036)	(0.042)	(0.039)
Controls							
Individual	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Employment	No	Yes	Yes	Yes	Yes	Yes	Yes
Work conditions and environment	No	No	Yes	Yes	Yes	Yes	Yes
Observations	10,563	10,563	10,563	10,563	10,563	10,563	10,563
sigma	1.256	1.245	1.239	0.819	1.216	0.368	4.287
R^2	0.345	0.352	0.357	0.534	0.590	0.483	0.935
Adjusted R^2	0.344	0.349	0.352	0.531	0.587	0.479	0.934

Note: *p<0.1; **p<0.05; ***p<0.01

Bootstrapped standard errors in parentheses. Individual's controls include age, sex, marital status, children, education, WHO5 index (when is not the outcome) and chronic disease. Employment characteristics include seniority, supervisor, part-time, atypical schedules, hourly wage, quantitative targets, occupation, sectors and firm size. Working conditions and environment include work pace constrained by machine, external demand, norms, computing control, hierarchical control or colleagues, work pressure, lack of time, help from colleagues or supervisors, fear of job future, organizational changes, promotion opportunities, work esteem, useful job, job and firm pride.

nificant heterogeneity in the effect of the duration of SP at baseline. While the estimates without interactions, presented in Table 2.4, show a unsignificant effect of SP on future SA, given the heterogeneity of self-rated health, the results confirm the adverse effect of SP on SA. A longer duration of SP at baseline would indeed be associated with an increase in future duration of SA for individuals who report being in rather good health at baseline. However, the better the individual's health status at baseline, the smaller the increase. For instance, given an average baseline duration of 2 days, a one-day increase in the duration of SP leads to a 7.2%

increase in the duration of SA when individuals report being in rather good health at baseline, compared with a 5% increase for individuals in very good health.

Controlling for heterogenous effects leads to a significant reduction in the correlation between the term capturing the endogeneity of SP and the error term, from 0.077 to 0.013. This correlation only becomes significant at the 10% level in column (3). Furthermore, the parameter associated with the first-stage residuals of the equation SP^* is no longer significant in columns (1) to (3).

In the same way, there is also a positive effect of lengthening the duration of SP at baseline on the future duration of SP and illness. Moreover, unlike the effects on the duration of SA, these effects increase with better health. The healthier the individual feels at baseline, the longer the future duration of SP and illness in the future. Given an average baseline duration of 2 days, a one-day increase in the duration of SP leads to a 11.2% increase in the total duration of illness for individuals in rather good health at baseline, whereas it corresponds to a 19.6% increase for individuals in very good health.

Finally, model (7) clearly shows a deleterious effect of the duration of SP on mental well-being. The healthier the individual, the greater the effect. Thus, in addition to a deterioration in health, leading to a longer duration of illness, SP leads to a deterioration in long-term mental well-being. This effect would be only linear.

6 Discussion and conclusion

The main contribution of this research is to study the causal relationship that may exist between SA or SP behaviours and future health outcomes. Ruhle et al. (2019) have highlighted this important lack in previous research. To date, no study has explored this causal relationship. It seems essential to consider this issue in order to better understand the fundamental relationship between the choice to attend or to be absent in case of illness and the consequences in terms of future health. Intuitively, if we agree to the fact that SP is detrimental to health, on the contrary, SA should be a restorative behaviour. However, past literature results do not identify this relationship and seem mainly to reflect the effect of health and

individual propensity to illness on future health. For instance, although the effect of SP was found to be detrimental to health and well-being (Skagen & Collins, 2016), the effects of SA and SP always seem to go in the same direction. The deleterious effect of SA on future health does not seem to be a rational result and contradicts the well-known medical prescriptions concerning rest in the event of illness. Similarly, if the results relating to SA do not seem to support intuition, what about the results associated with SP? Even if the previous literature does indeed conclude that SP has a deleterious effect on future health, how much of this effect is really due to the behaviour itself?

To solve this problem, it was proposed to explore an IV-free two-stage copula endogeneity correction model, called 2sCOPE, that takes into account the possible endogeneity of SA and SP on future health outcomes. This method makes it possible to control for the endogeneity assumed to be predominantly associated with the healthy worker effect issue (Haan & Myck, 2009). Correcting for endogeneity using copulas, making it possible to identify a causal effect without the use of additional information, is not a common approach in the economic literature compared to the marketing one. However, it is an entirely appropriate method in the case of microeconomic analysis of individual behaviour, and more specifically in the case of health studies. The high degree of endogeneity present in health studies remains a major challenge. The lack of relevant information in the survey data prevents the use of more traditional methods using instrumental variables. It is often difficult to validate all the assumptions required for their validity.

Therefore, if the estimates of SA and SP are cleared of the effects associated solely with health and individual's propensity to illness, it is assumed that they will adequately reflect the effect of the behaviours themselves. It is assumed that, without controlling for endogeneity, the results will reflect the fact that SA and SP are two good predictors of poor health (Caverley et al., 2007; Aronsson et al., 2011). In this sense, both behaviours will be positively associated with future illness, as has been shown in previous literature (Bergström et al., 2009b, 2009a; Gustafsson & Marklund, 2011; Janssens et al., 2013). On the other hand, once endogeneity is controlled for, it is assumed that the results better reflect the effect of choosing SA or SP.

To justify the intuition exposed above, this chapter first reproduces the results of the previous literature. Although the estimated models are different from the models generally used in past studies, the results allow us to recover the previous conclusions. Longer durations of SA and SP are positively associated with longer durations of SA, SP and illness 3 years later. Further, the duration of SA, respectively SP, at baseline is a better predictor of the lengthening of the future duration of SA, respectively SP. The duration of SP was also the strongest predictor of the total duration of illness (calculated as the sum of the reported durations of SA and SP). These results are consistent with the intuition that, without taking endogeneity into account, the behaviours are in fact predictors of health. Thus, the greater the duration of SA or SP reported by the worker, the greater the tendency to be ill, which would explain an increase in the total duration of illness, SA and SP in the future. However, these results are supposed to reflect the fact that both behaviours are intrinsically associated to illness propensity. In this sense, the results do not seem satisfactory to explain the consequences of the choice between SA and SP.

The 2sCOPE models with endogeneity correction bring to light some completely new results. Firstly, the models reveal the significant presence of endogeneity in the models. All the correlation coefficients that indicate the presence of endogeneity are significantly positive in the models associated with the durations of SA, SP and illness. As expected, they indicate that the unobservable factors that explain the two behaviours are positively and strongly associated with an increase in the duration of these behaviours and illness in the future. It is assumed that the endogeneity we are most likely dealing with is of the healthy worker effect type. In other words, this positive correlation would indicate that workers in poorer health would be more likely to fall ill at both observation periods. After controlling for this effect, it now appears that a longer duration of SA is indeed associated with a significant and continuous reduction in the duration of future SA and illness. This appears to be in line with the intuition that a longer duration of SA leads to a recovery of the health capital.

The 2sCOPE results related to the baseline duration of SP appear more ambiguous at first since they show no association with future SA, and describe a U-shape effect on the future duration of SP and illness. However, on the contrary to the

effect of SA which is decreasing and monotonic, that of SP appears at least positive for longer durations. The decreasing effect for shorter durations may be explained by several things.

Firstly, since the 2sCOPE models control for unobservable factors that are positively related to both the baseline duration of SP and the future duration of illness, the effects identified should be purely attributed to a behavioural effect. In this sense, SP may be used to signal the motivation and the investment of the worker in his job. For instance, in a period of uncertainty or dismissal, workers may invest in SP in order to avoid to be laid off. In addition, in a work environment where workers compete over a promotion to make a career, they may want to signal their propensity to effort by working while sick. Thus, once the period of uncertainty is over or the promotion has been obtained, workers do not perceived the utility of SP and reduce their duration.

Secondly, the individuals who report a positive duration of SP have been identified as the healthiest population. Given the three years gap between the two waves, workers who engaged in SP at baseline may have fully retrieved their health stock three years later. On the contrary, the longer the worker has been involved in SP, the lower the stock of health even three years later and therefore leads to a lengthening of the future duration of illness.

Finally, our results underline an important heterogeneity in individual's health for those who report a positive duration of SP. On the one side, there are individuals who are generally healthy and report few days of illness. While on the other side, there are individuals who are in poor health and show a much longer duration of illness. However, regarding the duration of SP, these two groups of individuals are confused. This finding justify, in addition to correct for endogeneity, the introduction of interaction terms between SP duration and self-rated health at baseline. Thus, after introducing these interaction terms in the 2sCOPE models, the results show significant heterogeneity in the effect of SP. Individuals who report at least a rather good health status will experience a longer duration of SA, SP and illness if they increase their baseline duration of SP. Overall, the results tend to conclude that SA is indeed a restorative behaviour while SP would deteriorate the health stock and leads to longer future duration of illness.

In addition to the main contribution associated with the consideration of endogeneity, allowing a better understanding of the long-term effects of SA and SP behaviours on future health, this chapter contributes to the literature in several ways. First of all, a large number of studies on SA and SP mainly explore cross-sections, while relatively few are interested in studying dynamic effects using longitudinal data. Moreover, the joint analysis of the effects of these behaviours on future illness had not received any interest since the latest work by Taloyan et al. (2012); Janssens et al. (2013) and Gustafsson & Marklund (2013). In addition, previous studies have generally been carried out on specific samples of individuals, with a relatively limited number of observations. Therefore, this chapter explores longitudinal data from the 2013 and 2016 "Working Conditions and Psychosocial Risks" survey, which covers a sample of 10,563 individuals representative of the employed working population in France. It is also the first study of its kind based on French data. Finally, the study of these behaviours in terms of cumulative number of days is relatively new (Arnold & de Pinto, 2015). In any case, previous studies that have used durations have not explored these data satisfactorily, generally grouping them into categories. This strategy, in contrast to the present study, does not make it possible to determine a marginal effect of SA and SP behaviours, but rather the probabilities of passing a certain threshold, set by the authors. These thresholds generally differ from one study to another and cannot be directly compared.

The results of this chapter also make an important operational contribution. As they provide the marginal effects of a day of SA and SP, it is possible to account for the costs and benefits associated with the cumulative duration of SA and SP. Encouraging the choice of SA will lead to reduce future duration of illness and SA. Firms may promote SA over SP in case of illness to limit their future cost. There is therefore a balance between the current costs of SA and SP and the future benefits of a healthier workforce. This rises the question of the sustainability of the workforce on the long run. In addition to increase the duration of illness, SP is also found to be detrimental to mental well-being, unlike SA. This should increase the potential costs associated with a decline in the workforce well-being, such as reduced motivation and productivity or increased employee turnover. In a context of developing the work life quality in order to promote workers' health and well-

being, human resources strategies should consider these results carefully. Finally, public policies should also consider these effects in the context of the substantial costs of paying days of absence. However, this represents a complex trade-off in order to favour SA over SP to invest in a workforce that will be in better health in the future. This raises important moral hazard issues. How to implement a policy that aims to promote absence only in the event of illness in order to discourage SP but also not to encourage shirking?

Finally, this analysis is not without limits. Firstly, the time gap of 3 years between the two waves is relatively large. Although this makes it possible to study long-term effects, it implies that several confounding factors may affect health, well-being and the possibility of being involved in SA and SP between these two observation periods. Future research should analyse surveys that are closer in time but also extended over a longer period in order to accurately test the duration of the effects of SP and SA over time. However, it is recommended that future research be inspired by the method used in this chapter to determine the existence of a causal relationship between SP and future SA over a more or less long period of time. It would be interesting to determine the duration of the effect of SP on health. It is not clear whether SP can have a causal impact on health and future SA at 6 months, 1 year, 3 years or even longer.

Furthermore, in addition to the time lapse between waves, this study uses only two waves of the survey. Although the 2019 wave is available, SP has unfortunately not been registered. This prevents the use of panel models to control for time-invariant individual fixed effects. However, the 2sCOPE method proposed in this chapter could be applied to time demeaned data. This allows, in addition to controlling for fixed effects, to explore the existence of endogeneity in the case of panel models. This would make it possible to identify in greater detail the unobservable determinants that are specific to individuals.

Finally, while it was proposed to study the cumulated duration of SA and SP, it should be emphasised that it is impossible in this case to identify episodes of SA and SP over the course of the year. A complete analysis of the two behaviours should cover both duration and frequency, while providing information on the periodicity of the behaviours in order to observe and measure the dynamics between the SA and SP

episodes with their respective duration. This would provide a better understanding of how workers arbitrate between SA and SP, taking into account their past episodes and their duration. Are there substitution strategies between SA and SP for the same illness episode, conditionally to its duration, or does the choice between SA and SP mainly concern several episodes? What is the effect of past behaviour choice, and duration, on the future choice between SA and SP? These questions require particular attention in order to gain a better understanding of the individual mechanisms that lead to the choice between SA or SP and then their respective durations. To do this, it would be extremely useful to study new data such as daily or weekly data.

Bibliography

Aboagye, E., Björklund, C., Gustafsson, K., Hagberg, J., Aronsson, G., Marklund, S., ... Bergström, G. (2019). Exhaustion and impaired work performance in the workplace: Associations with presenteeism and absenteeism. *Journal of Occupational and Environmental Medicine*, *61*, 1.

Akerstedt, T., Kecklund, G., Alfredsson, L., & Selen, J. (2007). Predicting long-term sickness absence from sleep and fatigue. *Journal of Sleep Research*, *16*(4), 341-345.

Allen, S. (1981). An empirical model of work attendance. *The Review of Economics and Statistics*, *63*(1), 77-87.

Amegashie, J. A., Cadsby, C. B., & Song, Y. (2007). Competitive burnout: Theory and experimental evidence. *Games and Economic Behavior*, *59*(2), 213-239.

Arnold, D., & de Pinto, M. (2015). How are work-related characteristics linked to sickness absence and presenteeism? theory and data. *Schmollers Jahrbuch*, *135*, 465-498.

Aronsson, G., Gustafsson, K., & Dallner, M. (2000). Sick but yet at work. an empirical study of sickness presenteeism. *J Epidemiol Community Health*, *54*, 502–509.

Aronsson, G., Gustafsson, K., & Mellner, C. (2011). Sickness presence, sickness absence, and self-reported health and symptoms. *International Journal of Workplace Health Management*, *4*, 228 - 243.

Baptista, M. C., Burton, W., Nahas, A., Wang, Y., Viana, M. C., & Andrade, L. H. (2019). Absenteeism and presenteeism associated with common health conditions in brazilian workers. *Journal of occupational and environmental medicine*, *61*, 303-313.

Barmby, T., & Larguem, M. (2009). Coughs and sneezes spread diseases: An empirical study of absenteeism and infectious illness. *Journal of Health Economics*, *28*(5), 1012-1017.

Bergström, G., Bodin, L., Hagberg, J., Aronsson, G., & Josephson, M. (2009b). Sickness presenteeism today, sickness absenteeism tomorrow? a prospective study on sickness presenteeism and future sickness absenteeism. *Journal of Occupational and Environmental Medicine*, *51*(6), 629–638.

Bergström, G., Bodin, L., Hagberg, J., Lindh, T., Aronsson, G., & Josephson, M. (2009a). Does sickness presenteeism have an impact on future general health? *International archives of occupational and environmental health*, *82*(10), 1179–1190.

Biron, C., Karanika-Murray, M., & Ivers, H. (2022). The health-performance framework of presenteeism: A proof-of-concept study. *Frontiers in Psychology*, *Volume 13 - 2022*.

Brown, S., & Sessions, J. G. (2004). Absenteeism, 'presenteeism', and shirking. *Economic Issues*, *9*(1), 15–23.

Caverley, N., Cunningham, J. B., & MacGregor, J. N. (2007). Sickness presenteeism, sickness absenteeism, and health following restructuring in a public service organization. *Journal of Management Studies*, *44*(2), 304–319.

Chadi, A., & Goerke, L. (2018). Missing at work – Sickness-related absence and subsequent career events. *Journal of Economic Behavior & Organization*, *153*(C), 153-176.

Chou, C.-Y., & Mach, M. (2021). Unlocking the contradictory outcomes of presenteeism through a temporal model: Effort exertion as a mediator. *Frontiers in Psychology*, *Volume 12 - 2021*.

Coles, M., Lanfranchi, J., Skalli, A., & Treble, J. (2007). Pay, technology, and the cost of worker absence. *Economic Inquiry*, *45*, 268–285.

Collins, A. M., Cartwright, S., & Cowlishaw, S. (2017). Sickness presenteeism and sickness absence over time: A uk employee perspective. *Work & Stress*, 1-16.

Conway, P. M., Hogh, A., Rugulies, R., & Åse Marie Hansen. (2014). Is sickness presenteeism a risk factor for depression?: A danish 2-year follow-up study. *Journal of Occupational and Environmental Medicine*, *56*(6), 595–603.

Cour des Comptes. (2024). *Communiqué de presse. arrêt de travail pour maladie : Mise au point du premier président de la cour des comptes* (Rapport technique). Consulté sur https://www.ccomptes.fr/fr/documents/70036

Dellve, L., Hadzibajramovic, E., & Ahlborg, G. (2011). Work attendance among healthcare workers: Prevalence, incentives, and long-term consequences for health and performance. *Journal of advanced nursing*, *67*, 1918-29.

Demerouti, E., Le Blanc, P. M., Bakker, A. B., Schaufeli, W. B., & Hox, J. (2009). Present but sick: a three-wave study on job demands, presenteeism and burnout. *Career Development International*, *14*(1), 50–68.

Docherty, P., Forslin, J., Shani, A. B., & Kira, M. (2002). Emerging work systems: Creating sustainable work systems: Emerging perspectives and practice. Routledge: London UK; New York, NY, USA.

Eckert, C., & Hohberger, J. (2023). Addressing endogeneity without instrumental variables: An evaluation of the gaussian copula approach for management research. *Journal of Management*, *49*(4), 1460-1495.

Evans-Lacko, S., & Knapp, M. (2016). Global patterns of workplace productivity for people with depression: absenteeism and presenteeism costs across eight diverse countries. *Social Psychiatry and Psychiatric Epidemiology Journal*, *51*, 1525–1537.

Garrow, V. (2016). Presenteeism: A review of current thinking. *Institute for Employment Studies, Report 507*.

Gerich, J. (2016). Determinants of presenteeism prevalence and propensity: Two sides of the same coin? *Archives of Environmental & Occupational Health*, *71*(4), 189–198.

Gosselin, E., Lemyre, L., & Corneil, W. (2013). Presenteeism and absenteeism: Differentiated understanding of related phenomena. *Journal of Occupational Health Psychology*, *18*(1), 75–86.

Grinza, E., & Rycx, F. (2020). The impact of sickness absenteeism on firm productivity: New evidence from belgian matched employer-employee panel data. *Industrial Relations*, *59*(1), 150–194.

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy*, *80*(2), 223–255.

Gustafsson, K., & Marklund, S. (2011). Consequences of sickness presence and sickness absence on health and work ability: A swedish prospective cohort study. *International Journal of Occupational Medicine and Environmental Health*, *24*, 153–165.

Gustafsson, K., & Marklund, S. (2013). Associations between health and combinations of sickness presence and absence. *Occupational medicine (Oxford, England)*, *64*.

Haan, P., & Myck, M. (2009). Dynamics of health and labor market risks. *Journal of Health Economics*, *28*(6), 1116-1125.

Hansen, C. D., & Andersen, J. H. (2008). Going ill to work – what personal circumstances, attitudes and work-related factors are associated with sickness presenteeism? *Social Science & Medicine*, *67*, 956–964.

Hansen, C. D., & Andersen, J. H. (2009). Sick at work–a risk factor for long-term sickness absence at a later date? *J Epidemiol Community Health*, *65*, 397–402.

Helgesson, M., Gustafsson, K., Marklund, S., Aronsson, G., & Leineweber, C. (2021). Sickness absence and sickness presence among health and care employees in sweden - health complaints, health behavior, and future long-term sickness absence. *Journal of occupational and environmental medicine*, *Publish Ahead of Print*.

Hemp, P. (2004). Presenteeism: At work – but out of it. *Harvard Business Review*, *82*(10), 49–58.

Herrmann, M. A., & Rockoff, J. E. (2012). Worker absence and productivity: Evidence from teaching. *Journal of Labor Economics*, *30*(4), 749-782.

Heywood, J. S., Jirjahn, U., & Wei, X. (2008). Teamwork, monitoring and absence. *Journal of Economic Behavior & Organization*, *68*, 676–690.

Hillion, M. (2024). L'effet des conditions de travail sur la santé et le recours aux soins. *DOCUMENT D'ÉTUDES N° 277, DARES*, 1–110.

Hirsch, B., Lechmann, D. S. J., & Schnabel, C. (2017). Coming to work while sick: An economic theory of presenteeism with an application to german data. *Oxford Economic Papers*, *69*(4), 1010–1031.

Ichino, A., & Riphahn, R. T. (2005). The effect of employment protection on worker effort: Absenteeism during and after probation. *Journal of the European Economic Association*, *3*(1), 120–143.

Janssens, H., Clays, E., De Clercq, B., De Backer, G., & Braeckman, L. (2013). The relation between presenteeism and different types of future sickness absence. *Journal of occupational health*, *55*.

Johansson, G., & Lundberg, I. (2004). Adjustment latitude and attendance requirements as determinants of sickness absence or attendance. empirical tests of the illness flexibility model. *Social Science & Medicine*, *58*, 1857–1868.

Johns, G. (2010). Presenteeism in the workplace: A review and research agenda. *Journal of Organizational Behavior*, *31*, 519–542.

Johns, G. (2011). Attendance dynamics at work: The antecedents and correlates of presenteeism, absenteeism, and productivity loss. *Journal of occupational health psychology*, *16*, 483-500.

Keane, M., & Neal, T. (2023). Instrument strength in iv estimation and inference: A guide to theory and practice. *Journal of Econometrics*, *235*(2), 1625-1653.

Kinman, G. (2019). Sickness presenteeism at work: prevalence, costs and management. *British Medical Bulletin*, 1–10.

Kivimäki, M., Head, J., Ferrie, J. E., Hemingway, H., Shipley, M. J., Vahtera, J., & Marmot, M. G. (2005). Working while ill as a risk factor for serious coronary events: The whitehall ii study. *American Journal of Public Health*, *95*(1), 98–102.

Kline, R. B. (2011). *Principles and practice of structural equation modeling.* (5th ed., pp. 3–427). New York: The Guilford Press.

Lohaus, D., & Habermann, W. (2019). Presenteeism: A review and research directions. *Human Ressource Management Review*, *29*, 43–58.

Lohaus, D., & Habermann, W. (2021). Understanding the decision-making process between presenteeism and absenteeism. *Frontiers in Psychology*, *12*, 716925.

López-Bueno, R., Clausen, T., Calatayud, J., Bláfoss, R., Vinstrup, J., & Andersen, L. L. (2021). Self-reported sickness absence and presenteeism as predictors of future disability pension: Cohort study with 11-year register follow-up. *Preventive Medicine*, *148*, 106565.

Maestas, N., Mullen, K., & Rennane, S. (2020). Absenteeism and presenteeism among american workers. *Journal of Disability Policy Studies*, *32*, 13-23.

Malmberg-Ceder, K., Vuorio, T., Korhonen, P., Kautiainen, H., Soinila, S., & Haanpää, M. (2020). The impact of self-reported recurrent headache on absenteeism and presenteeism at work among finnish municipal female employees. *Journal of Pain Research*, *13*, 2135-2142.

Marie, O., & Val Castello, J. (2023). Sick leave cuts and (unhealthy) returns to work. *Journal of Labor Economics*, *41*, 923–956.

Meijman, T. F., & Mulder, G. (1998). Psychological aspects of workload. In P. J. D. Drenth & H. Thierry (Eds.), *A handbook of work and organizational psychology* (Vol. 2, p. 5-33). Hove, United Kingdom: Psychology Press.

Miraglia, M., & Johns, G. (2016). Going to work ill: A meta-analysis of the correlates of presenteeism and a dual-path model. *Journal of Occupational Health Psychology*, *21*(3), 261–283.

Navarro, A., Salas-Nicás, S., Llorens, C., Moncada, S., Molinero-Ruíz, E., & na, D. M. (2019). Sickness presenteeism: Are we sure about what we are studying? a research based on a literature review and an empirical illustration. *American Journal of Industrial Medicine*, 1–10.

Park, S., & Gupta, S. (2012). Handling endogenous regressors by joint estimation using copulas. *Marketing Science*, *31*(4), 567–586.

Patel, C., Biron, M., Cooper, S. C., & Budhwar, P. S. (2023). Sick and working: Current challenges and emerging directions for future presenteeism research. *Journal of Organizational Behavior*, *44*(6), 839-852.

Pauly, M. V., Nicholson, S., Polsky, D., Berger, M. L., & Sharda, C. (2008). Valuating reductions in on-the-job illness: 'presenteeism' from managerial and economic perspectives. *Health Economics*, *17*, 469–485.

Qian, Y., Koschmann, A., & Xie, H. (2024). *A practical guide to endogeneity correction using copulas* (Working Paper N° 32231). National Bureau of Economic Research.

Qian, Y., & Xie, H. (2024). Correcting regressor-endogeneity bias via instrument-free joint estimation using semiparametric odds ratio models. *Journal of Marketing Research*, *61*(5), 914-936.

Ruhle, S., Breitsohl, H., Aboagye, E., Baba, V., Biron, C., Correia Leal, C., ... Muschalla, B. (2019). "to work, or not to work, that is the question" - recent trends and avenues for research on presenteeism. *European Journal of Work and Organizational Psychology*, *29*, 344-363.

Sanatkar, S., Stevelink, S., Aars, N., Bardal, I., Brandseth, O., Brinchmann, B., & Mykletun, A. (2025). Health consequences of graded, full, and no sickness absence among workers with musculoskeletal disorders: health profiles and six-months symptom changes of patients referred to norwegian outpatient clinics for chronic neck and back pain. *BMC Musculoskeletal Disorders*, *26*.

Schultz, A. B., & Edington, D. W. (2007). Employee health and presenteeism: A systematic review. *Journal of Occupational Rehabilitation*, *17*, 547–579.

Simpson, R. (1998). Presenteeism, power and organizational change: Long hours as a career barrier and the impact on the working lives of women managers. *British Journal of Management*, *9*(s1), 37–50.

Skagen, K., & Collins, A. M. (2016). The consequences of sickness presenteeism on health and wellbeing over time: A systematic review. *Social Sciences & Medicine*, *161*, 169–177.

Strömberg, C., Aboagye, E., Hagberg, J., Bergström, G., & Lohela-Karlsson, M. (2017). Estimating the effect and economic impact of absenteeism, presenteeism, and work environment–related problems on reductions in productivity from a managerial perspective. *Value in Health*, *20*, 1058–1064.

Taloyan, M., Aronsson, G., Leineweber, C., Magnusson Hanson, L., Alexanderson, K., & Westerlund, H. (2012). Sickness presenteeism predicts suboptimal self-rated health and sickness absence: A nationally representative study of the swedish working population. *PloS one*, *7*, e44721.

Triolle, A., Lefevre, N., Leclercq, M., Hairault, A., Lianova, R., & Pavis, I. (2024). *Revue de dépenses relative à la réduction des absences dans la fonction publique* (Rapport technique). Inspection générale des finances et Inspection générale des affaires sociales.

van der Pol, M., & Irvine, A. (2018). *Time preferences for health.* Oxford University Press.

van Hooff, M., Flaxman, P., Söderberg, M., Stride, C., & Geurts, S. (2018). Basic psychological need satisfaction, recovery state, and recovery timing. *Human Performance*, *31*, 125-143.

Voss, M., Stark, S., Alfredsson, L., Vingård, E., & Josephson, M. (2008). Comparisons of self-reported and register data on sickness absence among public employees in sweden. *Occupational and Environmental Medicine*, *65*(1), 61–67.

Yang, F., Qian, Y., & Xie, H. (2024). Addressing endogeneity using a two-stage copula generated regressor approach. *Journal of Marketing Research*, *0*(0), 00222437241296453.

Čikeš, V., Maškarin Ribarić, H., & Črnjar, K. (2018). The determinants and outcomes of absence behavior: A systematic literature review. *Social Sciences*, *7*(8).

Appendices

1 Descriptive statistics

Table A1 – Descriptive statistics of the individual and employment characteristics in 2013 and 2016

Characteristic	Overall N = 25,770[1]	2013 N = 12,885[1]	2016 N = 12,885[1]	p-value[2]
Age	44.31 (8.95)	42.81 (8.82)	45.81 (8.82)	<0.001
Female	14,328 (55.60%)	7,164 (55.60%)	7,164 (55.60%)	>0.9
Single	5,497 (21.33%)	2,750 (21.34%)	2,747 (21.32%)	>0.9
Child under 18	14,228 (55.21%)	7,344 (57.00%)	6,884 (53.43%)	<0.001
Education				>0.9
No diploma	1,388 (5.39%)	697 (5.41%)	691 (5.36%)	
less than Bac	8,051 (31.24%)	4,033 (31.30%)	4,018 (31.18%)	
Bac	4,835 (18.76%)	2,428 (18.84%)	2,407 (18.68%)	
Higher than Bac	11,496 (44.61%)	5,727 (44.45%)	5,769 (44.77%)	
Hourly wage	13.05 (5.01)	12.78 (4.92)	13.34 (5.09)	<0.001
Unknown	2,639	810	1,829	
Part-time	4,150 (16.10%)	2,110 (16.38%)	2,040 (15.83%)	0.2
supervisor	3,354 (13.02%)	1,599 (12.41%)	1,755 (13.62%)	0.004
Seniority y	15.23 (10.17)	14.31 (9.96)	16.15 (10.30)	<0.001
atypical schedules	14,408 (55.91%)	7,241 (56.20%)	7,167 (55.62%)	0.4
Quantitative targets	7,184 (27.88%)	3,650 (28.33%)	3,534 (27.43%)	0.11
Occupation				<0.001
A category or equiv	3,728 (14.47%)	1,822 (14.14%)	1,906 (14.79%)	
B category or equiv	2,806 (10.89%)	1,359 (10.55%)	1,447 (11.23%)	
CD category or equiv	4,530 (17.58%)	2,311 (17.94%)	2,219 (17.22%)	
Engineer executive	2,550 (9.90%)	1,193 (9.26%)	1,357 (10.53%)	
Foreman Sales rep	1,619 (6.28%)	789 (6.12%)	830 (6.44%)	
Technician	1,778 (6.90%)	937 (7.27%)	841 (6.53%)	
Office/Sales worker	4,823 (18.72%)	2,433 (18.88%)	2,390 (18.55%)	
Skilled worker	2,998 (11.63%)	1,502 (11.66%)	1,496 (11.61%)	
Specialized worker	938 (3.64%)	539 (4.18%)	399 (3.10%)	
Sectors				>0.9
care	13,321 (51.69%)	6,630 (51.46%)	6,691 (51.93%)	
comm	2,128 (8.26%)	1,054 (8.18%)	1,074 (8.34%)	
autind	1,704 (6.61%)	860 (6.67%)	844 (6.55%)	
admi	1,530 (5.94%)	764 (5.93%)	766 (5.94%)	
trans	1,073 (4.16%)	529 (4.11%)	544 (4.22%)	
cons	1,043 (4.05%)	511 (3.97%)	532 (4.13%)	
autser	996 (3.86%)	528 (4.10%)	468 (3.63%)	
finan	716 (2.78%)	357 (2.77%)	359 (2.79%)	
alim	601 (2.33%)	309 (2.40%)	292 (2.27%)	
info	586 (2.27%)	290 (2.25%)	296 (2.30%)	
mattr	520 (2.02%)	258 (2.00%)	262 (2.03%)	
extr	433 (1.68%)	221 (1.72%)	212 (1.65%)	
hebe	394 (1.53%)	199 (1.54%)	195 (1.51%)	
elec	287 (1.11%)	154 (1.20%)	133 (1.03%)	
agri	237 (0.92%)	118 (0.92%)	119 (0.92%)	
immo	201 (0.78%)	103 (0.80%)	98 (0.76%)	
Firm size				<0.001
less 10	3,455 (13.41%)	1,724 (13.38%)	1,731 (13.43%)	
10 50	6,197 (24.05%)	3,050 (23.67%)	3,147 (24.42%)	
50 200	5,863 (22.75%)	2,889 (22.42%)	2,974 (23.08%)	
200 500	3,530 (13.70%)	1,772 (13.75%)	1,758 (13.64%)	
more 500	5,676 (22.03%)	2,818 (21.87%)	2,858 (22.18%)	
No Info	1,049 (4.07%)	632 (4.90%)	417 (3.24%)	

[1]n (%); Mean (SD)

[2]Pearson's Chi-squared test; Wilcoxon rank sum test; Fisher's exact test

Table A2 – Descriptive statistics of the working conditions in 2013 and 2016

Characteristic	Overall N = 25,770[1]	2013 N = 12,885[1]	2016 N = 12,885[1]	p-value[2]
Work paced by:				
machine	2,776 (10.77%)	1,417 (11.00%)	1,359 (10.55%)	0.2
demand	18,044 (70.02%)	8,960 (69.54%)	9,084 (70.50%)	0.092
norms	11,555 (44.84%)	5,714 (44.35%)	5,841 (45.33%)	0.11
computer control	10,136 (39.33%)	4,995 (38.77%)	5,141 (39.90%)	0.063
hierarchical control	7,609 (29.53%)	3,913 (30.37%)	3,696 (28.68%)	0.003
colleagues	7,799 (30.26%)	3,968 (30.80%)	3,831 (29.73%)	0.063
Lack time	7,265 (28.19%)	3,603 (27.96%)	3,662 (28.42%)	0.4
Help from supervisor	16,907 (65.61%)	8,472 (65.75%)	8,435 (65.46%)	0.6
Help from colleagues	21,830 (84.71%)	10,894 (84.55%)	10,936 (84.87%)	0.5
Under pressure	8,831 (34.27%)	4,767 (37.00%)	4,064 (31.54%)	<0.001
Unsure	4,883 (18.95%)	2,592 (20.12%)	2,291 (17.78%)	<0.001
Changes	12,077 (46.86%)	6,155 (47.77%)	5,922 (45.96%)	0.004
Promo prospect				<0.001
Not concerned	2,518 (9.77%)	1,101 (8.54%)	1,417 (11.00%)	
Not satisfied at all	4,399 (17.07%)	2,219 (17.22%)	2,180 (16.92%)	
Not satisfied	7,918 (30.73%)	4,137 (32.11%)	3,781 (29.34%)	
Satisfied	8,650 (33.57%)	4,337 (33.66%)	4,313 (33.47%)	
Very satisfied	2,285 (8.87%)	1,091 (8.47%)	1,194 (9.27%)	
Work esteem	4,805 (18.65%)	2,280 (17.69%)	2,525 (19.60%)	<0.001
Firm pride	7,107 (27.58%)	3,610 (28.02%)	3,497 (27.14%)	0.12
Job pride	3,788 (14.70%)	1,666 (12.93%)	2,122 (16.47%)	<0.001
Useful job	6,264 (24.31%)	3,112 (24.15%)	3,152 (24.46%)	0.6

[1]n (%)

[2]Pearson's Chi-squared test; Wilcoxon rank sum test; Fisher's exact test

Table A3 – Descriptive statistics of the health characteristics in 2013 and 2016

Characteristic	Overall N = 25,770[1]	2013 N = 12,885[1]	2016 N = 12,885[1]	p-value[2]
SP	11,309 (43.88%)	5,444 (42.25%)	5,865 (45.52%)	<0.001
SP days	2.26 (5.38)	2.00 (4.96)	2.52 (5.75)	<0.001
SP bracket				<0.001
No SP	14,461 (56.47%)	7,441 (58.06%)	7,020 (54.87%)	
1 to 3 days	6,887 (26.89%)	3,528 (27.53%)	3,359 (26.26%)	
4 to 7 days	2,449 (9.56%)	1,089 (8.50%)	1,360 (10.63%)	
1 to 2 weeks	1,431 (5.59%)	606 (4.73%)	825 (6.45%)	
2 to 4 weeks	382 (1.49%)	153 (1.19%)	229 (1.79%)	
Unknown	160	68	92	
SA	8,003 (31.06%)	3,718 (28.86%)	4,285 (33.26%)	<0.001
SA days	4.79 (13.14)	4.25 (12.31)	5.33 (13.90)	<0.001
SA bracket				<0.001
No SA	17,767 (68.94%)	9,167 (71.14%)	8,600 (66.74%)	
1 to 3 days	2,223 (8.63%)	1,138 (8.83%)	1,085 (8.42%)	
4 to 7 days	2,041 (7.92%)	910 (7.06%)	1,131 (8.78%)	
1 to 2 weeks	1,598 (6.20%)	704 (5.46%)	894 (6.94%)	
2 to 4 weeks	1,129 (4.38%)	517 (4.01%)	612 (4.75%)	
1 to 3 months	1,012 (3.93%)	449 (3.48%)	563 (4.37%)	
SA episodes				<0.001
No SA	17,767 (68.96%)	9,167 (71.16%)	8,600 (66.75%)	
1	5,834 (22.64%)	2,734 (21.22%)	3,100 (24.06%)	
2	1,424 (5.53%)	661 (5.13%)	763 (5.92%)	
3 more	741 (2.88%)	320 (2.48%)	421 (3.27%)	
Unknown	4	3	1	
Illness	14,644 (56.83%)	7,037 (54.61%)	7,607 (59.04%)	<0.001
Illness days	7.05 (14.79)	6.26 (13.74)	7.85 (15.73)	<0.001
SA&SP groups				<0.001
No illness	11,126 (43.17%)	5,848 (45.39%)	5,278 (40.96%)	
SP only	6,641 (25.77%)	3,319 (25.76%)	3,322 (25.78%)	
SA only	3,335 (12.94%)	1,593 (12.36%)	1,742 (13.52%)	
SA&SP	4,668 (18.11%)	2,125 (16.49%)	2,543 (19.74%)	
health				<0.001
bad	877 (3.40%)	412 (3.20%)	465 (3.61%)	
rather good	5,322 (20.65%)	2,428 (18.84%)	2,894 (22.46%)	
good	13,476 (52.29%)	6,690 (51.92%)	6,786 (52.67%)	
very good	6,095 (23.65%)	3,355 (26.04%)	2,740 (21.27%)	
chronic disease	6,810 (26.43%)	3,212 (24.93%)	3,598 (27.92%)	<0.001
WHO5 score	17 (13, 19)	16 (13, 19)	17 (13, 20)	<0.001
Unknown	1,184	517	667	

[1]n (%); Mean (SD) or Median (IQR)

[2]Pearson's Chi-squared test; Wilcoxon rank sum test; Fisher's exact test

Table A4 – Descriptive statistics of SA and SP regarding health state in 2013

	self-rated health			
Characteristic	Bad or very bad N = 412[1]	rather good N = 2,428[1]	good N = 6,690[1]	very good N = 3,355[1]
SP	270 (65.53%)	1,368 (56.34%)	3,018 (45.11%)	1,209 (36.04%)
SP days	5.77 (9.39)	3.58 (6.83)	2.31 (5.26)	1.78 (4.98)
SA	208 (50.49%)	1,005 (41.39%)	2,165 (32.36%)	907 (27.03%)
SA days	10.42 (19.59)	7.30 (15.89)	5.07 (13.65)	3.81 (11.52)

[1]n (%); Mean (SD)

Table A5 – Descriptive statistics of SA and SP regarding health state in 2016

	self-rated health			
Characteristic	Bad or very bad N = 465[1]	rather good N = 2,894[1]	good N = 6,786[1]	very good N = 2,740[1]
SP	357 (76.77%)	1,738 (60.06%)	2,960 (43.62%)	810 (29.56%)
SP days	7.62 (10.74)	4.06 (7.59)	2.03 (4.64)	1.25 (3.61)
SA	315 (67.74%)	1,266 (43.75%)	2,077 (30.61%)	627 (22.88%)
SA days	18.49 (24.94)	7.66 (16.12)	4.42 (12.51)	2.88 (9.92)

[1]n (%); Mean (SD)

2 Interaction results between self-rated health and both SP and SA

Table A6 – Estimates of the effects of SA and SP interacted with self-rated health at baseline

	Dependent variable:						
	$\log(\text{SA days})_t$			$\log(\text{SP days})_t$	$\log(\text{Illness days})_t$	low SRH_t	WHO5_t
	OLS			*OLS*	*OLS*	*probit*	*OLS*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Constant	**1.729*****	**1.634*****	**1.441*****	**1.651*****	**2.434*****	**0.954*****	**14.317*****
	(0.142)	(0.168)	(0.189)	(0.129)	(0.181)	(0.232)	(0.674)
Rather good SRH_{t-1}	−0.209*	−0.157	−0.143	−0.159*	−0.219*	**−0.294****	0.579
	(0.122)	(0.131)	(0.130)	(0.088)	(0.121)	(0.134)	(0.478)
Good SRH_{t-1}	**−0.276****	−0.233*	−0.215*	**−0.235*****	**−0.329*****	**−1.006*****	**1.752*****
	(0.119)	(0.127)	(0.126)	(0.085)	(0.117)	(0.130)	(0.465)
Very good SRH_{t-1}	**−0.301****	−0.252*	−0.220*	**−0.266*****	**−0.379*****	**−1.295*****	**2.879*****
	(0.120)	(0.129)	(0.128)	(0.086)	(0.119)	(0.135)	(0.469)
SP days_{t-1}	−0.023	−0.017	−0.015	**0.034****	0.015	0.016	**−0.197*****
	(0.016)	(0.017)	(0.017)	(0.014)	(0.016)	(0.019)	(0.065)
SP days^2_{t-1}	0.00027	0.00017	0.00013	−0.00033	−0.00019	−0.00031	**0.00434*****
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0004)	(0.001)
(SP days×rather good $\text{SRH})_{t-1}$	**0.054*****	**0.046****	**0.043****	0.025	**0.045****	−0.017	0.069
	(0.019)	(0.020)	(0.020)	(0.016)	(0.019)	(0.021)	(0.074)
(SP days^2×rather good $\text{SRH})_{t-1}$	**−0.00087****	−0.00071*	−0.00063	−0.00066*	**−0.00080****	0.00032	−0.00218
	(0.0004)	(0.0004)	(0.0004)	(0.0003)	(0.0004)	(0.0004)	(0.002)
(SP days×good $\text{SRH})_{t-1}$	**0.050*****	**0.039****	0.034*	**0.047*****	**0.058*****	−0.002	0.109
	(0.017)	(0.019)	(0.019)	(0.015)	(0.018)	(0.021)	(0.071)
(SP days^2×good $\text{SRH})_{t-1}$	**−0.00072****	−0.00050	−0.00040	**−0.00103*****	**−0.00010*****	0.00019	−0.00282*
	(0.0004)	(0.0004)	(0.0004)	(0.0003)	(0.0003)	(0.0005)	(0.002)
(SP days×very good $\text{SRH})_{t-1}$	**0.045****	0.033	0.028	**0.079*****	**0.078*****	0.010	0.059
	(0.020)	(0.022)	(0.022)	(0.018)	(0.021)	(0.034)	(0.080)
(SP days^2×very good $\text{SRH})_{t-1}$	−0.00067	−0.00055	−0.00045	**−0.00215*****	**−0.00189*****	−0.00274	−0.00131
	(0.001)	(0.001)	(0.001)	(0.0005)	(0.001)	(0.002)	(0.002)
SA days_{t-1}	**0.037*****	**0.036*****	**0.035*****	0.006	**0.024****	−0.015	−0.001
	(0.010)	(0.010)	(0.010)	(0.007)	(0.009)	(0.011)	(0.035)
SA days^2_{t-1}	**−0.00038*****	**−0.00035*****	**−0.00034****	−0.00011	−0.00025*	0.00022	0.00036
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0005)
(SA days×rather good $\text{SRH})_{t-1}$	−0.006	−0.003	−0.003	0.004	0.006	0.014	0.005
	(0.011)	(0.012)	(0.012)	(0.008)	(0.011)	(0.012)	(0.040)
(SA days^2rather good $\text{SRH})_{t-1}$	−0.00000	−0.00006	−0.00007	−0.00002	−0.00014	−0.00020	−0.00037
	(0.0001)	(0.0002)	(0.0002)	(0.0001)	(0.0001)	(0.0002)	(0.001)
(SA days×good $\text{SRH})_{t-1}$	−0.009	−0.009	−0.009	−0.001	−0.001	0.020*	−0.017
	(0.010)	(0.011)	(0.011)	(0.008)	(0.010)	(0.011)	(0.037)
(SA days^2good $\text{SRH})_{t-1}$	0.00006	0.00005	0.00004	0.00005	−0.00001	−0.00027	−0.00011
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.001)
(SA days×very good $\text{SRH})_{t-1}$	0.0002	0.003	0.002	−0.006	0.008	0.020	0.007
	(0.012)	(0.012)	(0.012)	(0.009)	(0.012)	(0.014)	(0.041)
(SA days^2very good $\text{SRH})_{t-1}$	−0.00008	−0.00012	−0.00011	0.00013	−0.00013	−0.00016	−0.00018
	(0.0002)	(0.0002)	(0.0002)	(0.0001)	(0.0002)	(0.0002)	(0.001)
Controls							
Individual	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Employment	No	Yes	Yes	Yes	Yes	Yes	Yes
Work conditions and environment	No	No	Yes	Yes	Yes	Yes	Yes
Observations	12,218	10,563	10,563	10,563	10,563	10,563	10,563
R^2	0.080	0.091	0.098	0.237	0.193		0.278
Adjusted R^2	0.078	0.086	0.091	0.231	0.186		0.272
Log Likelihood						−4,400	

Note: *p<0.1; **p<0.05; ***p<0.01

Table A7 – Estimates of the 2sCOPE models with self-rated health interactions with SP and SA

	Dependent variable:						
	$\log(\text{SA days})_t$			$\log(\text{SP days})_t$	$\log(\text{Illness days})_t$	low SRH_t	WHO5_t
	OLS			*OLS*	*OLS*	*OLS*	*OLS*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Constant	**2.453*****	**2.294*****	**2.079*****	**2.539*****	**3.505*****	**0.827*****	**11.975*****
	(0.310)	(0.324)	(0.311)	(0.218)	(0.310)	(0.092)	(1.109)
Rather good SRH_{t-1}	**−0.888*****	**−0.809*****	**−0.770*****	**−1.019*****	**−1.261*****	−0.084	**2.850*****
	(0.304)	(0.293)	(0.279)	(0.198)	(0.275)	(0.080)	(0.993)
Good SRH_{t-1}	**−0.972*****	**−0.892*****	**−0.848*****	**−1.105*****	**−1.383*****	**−0.314*****	**4.051*****
	(0.301)	(0.298)	(0.282)	(0.199)	(0.277)	(0.081)	(1.001)
Very good SRH_{t-1}	**−0.999*****	**−0.913*****	**−0.856*****	**−1.135*****	**−1.433*****	**−0.343*****	**5.174*****
	(0.303)	(0.300)	(0.284)	(0.199)	(0.278)	(0.081)	(1.009)
SP days_{t-1}	−0.055*	−0.049	−0.044	**−0.120*****	**−0.127*****	−0.009	0.180
	(0.032)	(0.031)	(0.030)	(0.029)	(0.033)	(0.010)	(0.110)
SP days^2_{t-1}	0.001	0.0005	0.0005	**0.002*****	**0.002*****	0.0001	−0.0001
	(0.0005)	(0.0004)	(0.0004)	(0.0004)	(0.0005)	(0.0001)	(0.002)
$(\text{SP days}\times\text{rather good SRH})_{t-1}$	**0.083*****	**0.077****	**0.071****	**0.170*****	**0.179*****	0.009	**−0.288*****
	(0.032)	(0.031)	(0.031)	(0.028)	(0.033)	(0.010)	(0.112)
$(\text{SP days}^2\times\text{rather good SRH})_{t-1}$	**−0.001****	**−0.001****	−0.001*	**−0.002*****	**−0.002*****	−0.0001	0.001
	(0.0005)	(0.0005)	(0.0005)	(0.0004)	(0.0005)	(0.0002)	(0.002)
$(\text{SP days}\times\text{good SRH})_{t-1}$	**0.079****	**0.071****	**0.063****	**0.202*****	**0.200*****	0.013	**−0.270****
	(0.033)	(0.033)	(0.031)	(0.030)	(0.034)	(0.010)	(0.114)
$(\text{SP days}^2\times\text{good SRH})_{t-1}$	**−0.001****	−0.001*	−0.001	**−0.003*****	**−0.003*****	−0.0001	0.002
	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0002)	(0.002)
$(\text{SP days}\times\text{very good SRH})_{t-1}$	**0.074****	0.064*	0.056*	**0.231*****	**0.218*****	0.005	**−0.314*****
	(0.035)	(0.034)	(0.032)	(0.030)	(0.036)	(0.010)	(0.116)
$(\text{SP days}^2\times\text{very good SRH})_{t-1}$	−0.001	−0.001	−0.001	**−0.004*****	**−0.003*****	−0.0001	0.003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.0002)	(0.002)
SA days_{t-1}	−0.006	−0.004	−0.004	−0.0004	−0.002	0.003	0.028
	(0.021)	(0.020)	(0.019)	(0.013)	(0.018)	(0.005)	(0.066)
SA days^2_{t-1}	−0.0002*	−0.0002*	−0.0002	−0.0001	−0.0002	0.0000	0.0002
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0000)	(0.0005)
$(\text{SA days}\times\text{rather good SRH})_{t-1}$	0.037*	0.034*	0.034*	0.010	0.030*	−0.003	−0.021
	(0.020)	(0.020)	(0.019)	(0.013)	(0.018)	(0.005)	(0.066)
$(\text{SA days}^2\times\text{rather good SRH})_{t-1}$	−0.0001	−0.0001	−0.0001	0.0000	−0.0002	0.0000	−0.0003
	(0.0002)	(0.0002)	(0.0002)	(0.0001)	(0.0002)	(0.0000)	(0.001)
$(\text{SA days}\times\text{good SRH})_{t-1}$	0.035	0.032	0.031	0.005	0.025	−0.001	−0.046
	(0.021)	(0.021)	(0.020)	(0.013)	(0.019)	(0.005)	(0.069)
$(\text{SA days}^2\times\text{good SRH})_{t-1}$	−0.0001	−0.0001	−0.0001	0.0000	−0.0001	−0.0001	0.0000
	(0.0002)	(0.0002)	(0.0002)	(0.0001)	(0.0002)	(0.0000)	(0.0005)
$(\text{SA days}\times\text{very good SRH})_{t-1}$	**0.050****	**0.047****	**0.045****	0.001	0.036*	−0.003	−0.024
	(0.023)	(0.022)	(0.022)	(0.014)	(0.020)	(0.006)	(0.074)
$(\text{SA days}^2\times\text{very good SRH})_{t-1}$	−0.0003	−0.0003	−0.0003	0.0001	−0.0003	0.0000	0.0000
	(0.0002)	(0.0002)	(0.0002)	(0.0001)	(0.0002)	(0.0000)	(0.001)
1st stage SP	0.240	0.209	0.187	**1.144*****	**1.042*****	0.099	**−2.800*****
	(0.210)	(0.202)	(0.195)	(0.170)	(0.210)	(0.064)	(0.728)
1st stage SA	**1.078****	**1.022****	**1.003****	0.161	0.667*	−0.175	−0.728
	(0.441)	(0.429)	(0.408)	(0.269)	(0.377)	(0.116)	(1.381)
cor SP	0.016*	0.014*	0.013	**0.082*****	**0.054*****	0.011	**−0.035*****
	(0.008)	(0.008)	(0.008)	(0.009)	(0.008)	(0.008)	(0.008)
cor SA	**0.055*****	**0.053*****	**0.052*****	0.017	**0.042*****	−0.014	−0.012
	(0.015)	(0.015)	(0.014)	(0.014)	(0.013)	(0.013)	(0.013)
Controls							
Individual	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Employment	No	Yes	Yes	Yes	Yes	Yes	Yes
Work conditions and environment	No	No	Yes	Yes	Yes	Yes	Yes
Observations	10,563	10,563	10,563	10,563	10,563	10,563	10,563
sigma	1.189	1.182	1.177	0.817	1.170	0.368	4.279
Rsq	0.340	0.348	0.353	0.534	0.588	0.483	0.935
Adj Rsq	0.338	0.344	0.348	0.531	0.585	0.479	0.934

Note: *p<0.1; **p<0.05; ***p<0.01

Chapter 3

Is presenteeism such a valuable signal for employees? Analysis on wages, career, health and well-being

1 Introduction

In today's highly competitive world, is it always wise to demonstrate your motivation to your employer by continuously increasing your efforts at work? As many companies shape their corporate cultures to foster employee retention and enhance motivation, it is worth questioning what returns employees actually receive on these investments in terms of career advancement. Moreover, within the broader context of modern economies aiming to improve working conditions and define what constitutes a decent job, this chapter examines the links between individual outcomes — such as career progression, health, and well-being — and a behaviour that remains relatively unexplored or poorly understood: *competitive presenteeism*. This behaviour can be analysed within the broader research framework surrounding job intensification and workplace competition.

According to the definition of presenteeism given by Cooper (1996) , and previous literature, we can identify two types of presenteeism: i) sickness presenteeism (SP) as working while sick (Johns, 2010), and ii) competitive presenteeism as making an effort in excess of that required (by a norm or a contract). Contrary to a large body of literature which focuses exclusively on sickness presenteeism, we are rather interested in what is called competitive presenteeism. As noticed by Ashby & Mahdon (2010), presenteeism could be a behaviour not only related to illness. This behaviour is also linked to over-investment in the job in terms of effort and working time. In her paper, Simpson (1998) highlights that working long hours, leading to presenteeism defined as "the tendency to stay at work beyond the time needed for effective performance of the job", can emerge in competitive environment. She introduces the notion of competitive presenteeism defining the behaviour of individuals work harder than needed in the aim to compete over a promotion. Competitive presenteeism can be associated with working longer than the average working time, not taking time off, taking work home in the evenings and at weekends or working many overtime hours often uncompensated.

Cooper (1996) defines presenteeism such as "being at work when you should be at home either because you are ill or because you are working such long hours that you are no longer effective" (p.15)

See for example the special issue Presenteeism at Work: Developments and Avenues for Future Research, Volume 44, Issue 6 in the Journal of Organizational Behaviour (2023).

In addition, in contrast to the definitions widely considered in the literature to date, potential loss of productivity, due to illness or exhaustion, is not included in this definition. High involvement practices can actually be beneficial for employers as it can increase employee productivity (Cappelli & Neumark, 2001). Studying sickness presenteeism, Wang et al. (2022) show positive effects on performance and argue that this increase of productivity would be supported by affective commitment to the firm. Consequently, there is no reason to associate a loss of productivity to the definition but rather to focus on the behaviour itself.

The study of presenteeism behaviour conditional on illness shock requires an approach that controls for the fact that this shock is not exogenous to individual characteristics and working conditions. On the contrary, competitive presenteeism may concern all workers. This overcomes the problem of endogenous selection of workers related to illness. However, it does not overcome the more traditional endogenous selection of workers regarding their ability and capacity that will be considered in the following analysis. Thus, for the purpose of this chapter, we leave aside the analysis of sickness presenteeism, as both types of presenteeism need to be fully explored and cannot be properly studied in a single study.

Competitive presenteeism (CP) may be close to the concept of workaholism. However, workaholism refers to "an irresistible inner drive to work excessively hard" (Schaufeli et al., 2008). This concept refers more to the individual's psychological state in relation to work than to the description of a rational behaviour. The link between the two concepts would ultimately be that workaholism implies presenteeism. In other words, workaholism would be an individual trait that could explain the observation of presenteeism at work (both competitive and sickness, for that matter). Furthermore, an important distinction that can be made between these two concepts relates to the objective pursued. It is assumed that the main reason why employees engage in CP is to signal to their employer their ability or investment in their work, with a view to developing their career. This is not assumed to be the case with workaholism, which is said to be carried out for the purpose of personal fulfillment and not necessarily associated with a strategic objective. The signalling interpretation of presenteeism can be found in the research about long working hours and overtime (Avgoustaki & Frankort, 2019; Campbell & Green, 2002; Anger, 2008).

However, although this hypothesis was proposed to explain the decision to work harder, no empirical study, except that provided by Anger (2008), has explored the long term effects of CP as the signalling mechanism.

This chapter has therefore two objectives. The first is to determine the potential benefits of CP in terms of pay and career for the employee. The second is to assess the potential association between CP and a deleterious health and well-being. In other words, the aim is to enlighten employees and employers about the consequences of presenteeism. Can it be a valuable signal, in the sense that it can explain career enhancement? But also, are the potential career benefits to be balanced against its costs in terms of health? Thus, we presume that CP is a signal, and then seek to determine whether this signal is actually rewarded by the firms. This chapter therefore seek to validate the signalling hypothesis of presenteeism though its effect on wages and career.

The consequences on health can be considerable for workers: loneliness, social isolation, depression and burnout are some of the most serious consequences (F. Green et al., 2021; Barnay, 2020; Demerouti et al., 2009). These costs, additionally to that of fatigue and mistakes, are not to be neglected by the firm either. For example, due to burnout, the potential loss of an employee who has accumulated specific human capital could ultimately represent a significant cost for the company. Similarly, the negative externalities on the working environment and the tensions generated between colleagues by an employee's exacerbated competitive behaviour can be costly in the long term. The deterioration of the working environment, leading to a drop in job satisfaction and team motivation, can be associated with a reduction in overall performance, in the engagement of each worker in his or her tasks, and ultimately to turnover intention (Hofaidhllaoui & Chhinzer, 2014; Santhanam & Srinivas, 2019).

However, CP is not only costly for companies, but also for society as a whole and the healthcare system. The prevalence of this behaviour makes it a public health issue. For instance, almost 16% of the workers in our database say they often or always work unpaid overtime, and almost 14% say they often or always take work home. Finally, 40% declare they often or always think about their work outside working hours. For further information, according to the CNAM data, consumption

Caisse National d'Assurance Maladie

of antidepressants rose by 9.7% over the period studied in this research. In 2019, this represented a cost to the public health insurance system of 1,214 million euros. Shifting firms' management policies and individuals' attitudes towards more health-conscious behaviour could be a substantial source of savings.

Therefore, using data from the last three waves of the "Working Conditions and Psychosocial Risks" survey conducted between 2013 and 2019 by the French Ministry of Labour, this chapter examines the effects of CP on career, as well as its association with health and well-being. These data constitute a panel made up of 11,667 workers observed over the last three waves of the survey leading to 26,484 pooled observations. Further, they contain variables that provide measures of the intensive margin of the worker's effort. As the non-illness related behaviour of presenteeism has received few attention in previous literature, this chapter is the first to construct a CP indicator that addresses the multi-factorial aspect of the behaviour. The panel structure of the data allows to control for individual fixed effects. A powerful advantage when unobserved individual's characteristics may affect the studied behaviour. Finally, we attempt to provide causality insights between high investment in the job and either career or health outcomes using an IV-free approach based on copula and an instrumental variable strategy.

The rest of the chapter is organized as follows. Section 2 presents the theoretical framework of the chapter and the empirical results that underpin our theoretical intuition. Section 3 describes the data and the measures of the variables of interest, Section 4 explains the econometric methods and section 5 presents the results. Finally, section 6 draws the conclusion and provides discussion for future research.

2 Review of the literature

2.1 The career concern

While Böckerman (2018) still emphasises the non-existence of a theorisation of presenteeism, we can found related topics linked to other concepts and behaviours, such as over-work, over-commitment and workaholics culture for instance. Thus, several theoretical explanations would make it possible to explain this behaviour. From the individual's point-of-view, we can assume two fundamental mechanisms

to explain an investment in CP: the research of future extrinsic rewards, such as higher wages and hierarchical positions, or the intrinsic satisfaction perceived by doing the job. From an economic viewpoint, these two motivations for workers to exert more effort at work lead to the same end, an increase in their satisfaction. In this section, we aim to examine the nature of compensation that contributes to increase individual utility.

The extrinsic motivation of competitive presenteeism

We might firstly think that workers may react strategically to some market inflexibility. For example, Allen (1981) explains that individuals, for various reasons, may accept a contract with more hours than they are actually ready to work (due, for example, to the cost of finding a job). As a result, they reserve a rate of absence for themselves, taking into account the salary received, the potential cost associated with the increased probability of being laid off or the reduced probability of being promoted. According to this model, there is a relationship between the attendance and absence rates of individuals and their hidden cost of absence. Similarly, in the case of CP, it would be interesting to consider the opposite interpretation of this intuition. It would stipulate that a worker who perceives a higher utility, related to a hidden future gain, from his or her current attendance could then devote more hours or effort than stipulated in the contract.

According to this interpretation, the benefits of CP would be an increase in the probability of being promoted in the future. Ichino & Riphahn (2005) show in fact that workers reduce their level of absence during the probation period, during which their employment is not guaranteed, in order to signal their investment, and return to a normal level of absence once the probation period is finished. Further, Bratti & Staffolani (2007), with a tournament model, also show equilibria where individuals are encouraged to provide a level of effort beyond that contractualised with the employer. Indicating that, in the context of a competition (for promotion or not to be dismissed), individuals can multiply signalling actions such as reducing their absence and increasing their level of effort.

Several theoretical explanations can be thus exposed. Presenteeism could be used by workers to create a social image or a reputation within the company. Kreps

& Wilson (1982) show in finite repeated games that reputation effects appear. It could also be the firm's corporate culture that implicitly impose workers to be highly invest in their work. Further, this culture can be promoted by top managers if they think that such behavior may be profitable for the company (Graham et al., 2022). Therefore, employees have to adopt the corporate culture to be accepted by the team, the colleagues, for their self-esteem but also to climb the ladder or not being laid off (Zinko & Rubin, 2015). Moreover, Hoffmann et al. (2020) show that companies themselves can set up deferred compensation mechanisms in order to encourage employees to signal their level of investment and effort. Indeed, if the firm is facing an adverse selection issue with regard to the worker's ability and productivity, in the case where workers have present-biased preferences, deferring the payment encourages them to reveal their type as soon as possible, in order to reduce the time between two payments.

These interpretations have the same objective, namely that the employee should make an effort in excess of that defined by a standard or a contract in order to obtain a future compensation. This concept refers to the deferred compensation theory. For example, Campbell & Green (2002) show in their paper that one unpaid overtime hour increase future earnings by 4% while paid overtime only raise future pay by approximately 1%. Pannenberg (2005) find a long term increase of 10% in earnings for individuals working unpaid overtime. On the contrary, Meyer & Wallette (2005) do not find evidence that working overtime is associated to future benefits.

As a result, all of these explanations can be grouped into a single one: the signalling idea of workplace behaviours. The theoretical intuition is therefore based on the seminal article of Spence (1973). Indeed, the theories outlined above are all based on the fact that some characteristics are unknown to the employer, leading to strategic behaviour between workers and companies as well as incentive mechanisms to reveal workers' ability and motivations. For instance, as stated by Papagiannaki (2014), "unpaid overtime and employees' tolerance towards working unpaid long hours is the information set which is necessary to send messages to employers". Thus, the mechanism considered in this chapter is that of career concern (Holmstrom, 1982) where workers are in permanent competition. In this case, the over-effort linked to CP would not be devoid of interest and would expect a reward

from the firm in terms of future salary and career advancement. We therefore assume that workers invest in presenteeism in order to reveal their abilities to their employer. In fact, the employer wishes to select, i.e. promote, the worker who is the most efficient in the sense that he is the one for whom the cost of effort is the lowest. The worker who wishes to reveal his propensity for effort can therefore send a signal to his employer by working harder than expected, in other words, by engaging in CP. According to the points underpinned in the describe literature above, we derive the following hypothesis:

Hypothesis 1. *If competitive presenteeism is a valuable signal for employers, the more workers are involved in this behaviour, the higher their wages and professional status should be over time.*

The intrinsic motivation of competitive presenteeism

While future extrinsic rewards may explain a large part of the behaviour, we might also think that individuals derive greater utility from engaging in CP because they are motivated by the intrinsic characteristics of their job (Ma et al., 2018). The labor donation theory indicates that workers who value the prosocial characteristics of a job perceive an intrinsic utility due to this social value (Preston, 1989). This valuation can be reflected in a higher level of effort than if it did not exist. Thus, for equal wages, an intrinsically motivated employee will make an extra effort. For example, non-profit sector employees who perceive the utility of prosocial production would be more likely to work unpaid overtime than private sector employees (Gregg et al., 2011).

However, between two employees and for the same expected level of effort, it is not necessary for the employer to reward as much the intrinsically motivated employee as the one who is not (Besley & Ghatak, 2005). On the contrary, it could even be that financially rewarding an intrinsically motivated employee would crowd out the effect associated with his intrinsic motivation (Bénabou & Tirole, 2003). It could also be that these individuals are not looking for a career advancement given their attraction to their current job. Furthermore, this promotion could also have an effect on their reputation if it is not in line with the values that motivate them in their current job. These individuals may, for example, not have a competitive

mindset and would not aspire to more senior positions. According to this point of view, the mechanism that would explain over-effort is radically different from the assumed competitive objective of the behaviour. This implies to well differentiating both populations in the following analysis.

Therefore, the effects of over-investment in the job on future individual's career are supposed to be different between the extrinsically motivated individuals and the intrinsically motivated ones. Given that CP would not be linked to an idea of competitiveness for intrinsically motivated workers, we would not expect to observe significantly better career progression compared to the average population. therefore, the second hypothesis is derived:

Hypothesis 2. *If competitive presenteeism is sustained by the worker's intrinsic motivations in his or her job, its effects on career advancement would be unsignificant.*

2.2 The health consequences of over-work

Theoretical works also show the existence of equilibrium for which the worker's behaviour could be harmful for health. For instance, Amegashie et al. (2007) underline the impact of fatigue in the context of tournaments. They show that there exists an equilibrium composed of a group of individuals, called the "burnout group", who use all their resources during the first stage of the contest in order to win the tournament whatever the cost, while this equilibrium is Pareto-dominated by many other equilibria. The authors argue that burnout is in fact an intrinsic outcomes of competition. Ryvkin (2011) explains that tournaments need to be longer and more staggered in order to prevent the damaging effects of fatigue on productivity.

A large part of the empirical research that explore the work-health association focus on the work intensity through working conditions and employment characteristics. The deleterious effects that have been identified are often related to constrained characteristics and working conditions (Barnay, 2016). Few attempts have been made to understand the investment in voluntary over-effort at work as a behaviour. This aspect of work intensity, whether voluntary or involuntary, and the impact of this distinction on health, is rarely highlighted. However, the employment characteristics that create a competitive climate may have harmful effects on health and

well-being. For example, the recent empirical work of Avgoustaki & Frankort (2019) and Avgoustaki & Cañibano (2020) sheds light about the negative effects of overtime and work intensity, defines as the employment characteristics that push workers to exert more effort in a given amount of time, on health and psychological outcomes such as stress. Using objective health measures, Andelic et al. (2024) shows higher levels of blood pressure and fibrinogen, a protein associated with chronic stress, in workers with performance-related pay. In addition, a constant psychological commitment to work, such as CP, can affect individuals' daily lives since it may lead to sleep disorders (Garrow, 2016). Constantly being concerned about one's job can lead to major imbalances between private and professional life, which in turn leads to health problems (Kossek & Lee, 2017). For companies, this pressure on workers can result in a higher level and/or duration of absence over time. Havet & Plantier (2023) show that working conditions, and in particular psychosocial determinants such as stress, lead to more sickness spells and a longer duration of absence. In the psychological research, the deleterious effect of work intensity, time pressure but also engagement in work on health, and in particular on depressive symptoms, has long been explained by the Job Demand-Resource (JD-R) model (Bakker et al., 2014, 2023). More seriously, high work intensity would be a predictor of suicide (Younès et al., 2018).

Although this is natural, based on the intuition of effort recovery (van Hooff et al., 2018), to argue that CP can worsen workers' health status, the review of the literature by Ganster et al. (2018) argues that, contrary to the belief that long working hours are detrimental to health and well-being, there is no evidence of a "direct causal effect of work hours on either physical or mental well-being outcomes". However, this conclusion based on the review of existing papers may be due to a lack of longitudinal studies included in the analysis. In addition, an important point to emphasise is the fact that working time, while it may appear to be a good proxy for the level of effort, does not on its own reveal workers' investment. For instance, Ten Brummelhuis et al. (2017) show that working time alone is not a health risk factor. The authors emphasise that it is more the state of mind in which individuals place themselves in relation to their job that helps to explain an effect on health and well-being.

Finally, it has been shown previously that the relationship between CP and career outcomes may be different for intrinsically motivated workers. It is also expected that these workers may report significant differences in term of health and mental well-being. Avgoustaki & Frankort (2019) found that the effects of work intensity on well-being depend on the worker's motivations. Intrinsically motivated workers who are confronted to extensive work effort report a better well-being regarding the extrinsically motivated ones. Therefore, intrinsically motivated workers engaged in CP may not reveal lower mental well-being. However, there may exist health effects. Therefore, to sum up, we derive the last hypothesis:

Hypothesis 3. *As presenteeism is energy-intensive for the worker, increasing both physical fatigue and mental exhaustion, it should a) decrease self-reported health and b) deteriorate mental well-being. c) As it increases the likelihood of illness, it should also increase the duration of sickness absence. d) Intrinsically motivated workers should be able to cope better with overwork because of the interest they have in it, leading to lower the potential negative effects on health and mental well-being.*

3 Data and method

To analyse the effects of competitive presenteeism, we use the last three waves of the "Working conditions and psychosocial risks" survey from the French Ministry of Labour in 2013, 2016 and 2019. This national survey is conducted every three years on a set of individuals representative of the French working population, with a more in-depth focus on psychosocial risks every other time. The waves studied are the first in which the same individuals are re-entered, making it possible to create a panel. These data thus contain a wide range of socio-demographic information and individual's health situation, as well as employment contract characteristics, working conditions and employer practices. The data are collected in two ways. First, they are collected face-to-face by an interviewer. Then, a self-evaluation questionnaire is proposed to the individual. This second part of the survey tackles more detailed and sensitive topics on work relations, such as physical and mental health problems, relations with colleagues and supervisors, feeling about the job content, etc. Overall,

enquête "conditions de travail et risques psychosociaux"

the three waves initially covered 49,680 respondents, representing an exceptional data set covering a broad range of the French population over a long period.

However, according to specific requests from the ministry of Labor, the initial panel is no longer representative of the French population. A special focus on the healthcare sector in 2016 has resulted in an over-representation of this population in the longitudinal data. No less than 55% of the sample is made up of individuals from the healthcare sector. This sector is supposed to be relatively different from the others in many ways. Particularly in terms of intrinsic motivation, where pro-social investment plays an important role, but also in terms of the specific working conditions imposed by the sector. Although many of these characteristics are controlled for in the models, in order to keep the results more general, the main analysis will focus on the sub-sample of workers from non-healthcare sectors only. In addition, the variable used to identify career progression is undoubtedly too aggregated to observe mobility in the healthcare sector. However, we can take advantage of these two large samples to study the heterogeneity of the effects between the healthcare and non-healthcare sectors.

3.1 Sample restriction

For the purpose of the study we limit the working sample to a sub-population of individuals with the following characteristics. we only include salaried workers who are employed over the entire period. The sample is therefore made up solely of individuals aged over 18 in 2013 and under 60 in 2019, earning between 6 and 50 euros an hour or between 500 and 8,000 euros a month, and working between 10 and 100 hours a week. We also exclude self-employed, craft-persons, farmers, non-salaried workers and top managers. Finally, given the assumption that CP may represent an investment by workers to advance their careers within their company, we restrict the sample to individuals who have remained with the same employer. Otherwise, this would imply that CP behaviour could have effects on the labour market directly. This is not considered here, mainly because of the career objective underlying this behaviour, but also because it can be assumed that employees have changed company because they did not get the promotion they had hoped for. An employee who demonstrates a high level of commitment to his or her work but does

not perceive the corresponding returns may choose to return to the labour market in order to obtain a position offering greater responsibility and higher remuneration. Such a transition would not be associated with their level of commitment to their previous employer.

All in all, the analysed sub-sample is made up of 5,170 workers observed over the three waves, thus, a total sample of 10,168 observations. The study is based on the unbalanced panel. However, the unbalancedness measures as provided in Ahrens & Pincus (1981) does not indicate a strong panel attrition .

3.2 The outcome variables

Two principal issues related to presenteeism behaviours are addressed in this chapter. On the first hand, we are looking for to evaluate the potential benefits of CP on wages and occupation, and the potential costs bear by the workers regarding health and mental well-being on the second hand. Thus, we detail below the different variables that will be used as outcomes. The descriptive statistics of that variables are depicted in Table 3.1 below.

First, we are interested in wages to provide information about career status. Individuals were asked to give their last monthly wage. Thus, wages are on a declaration basis and are registered as a continuous variable. As the data provides information on the volume of hours worked by the individuals, we calculate the hourly wage based on the working time usually achieved during a week. In the following analysis we employ the logarithm of hourly wage. As stated by Card (1999), it is better for econometric purpose to use the log of hourly wages since it is normally distributed on the contrary to the distribution of the annual wages. An argument supported by the distribution of the *log(hourly wage)* variable in the CT-RPS data.

Secondly, we explain the worker's occupational position at the time of the interview in each wave. The *occupational grades* variable is a 5-point scale constructed as follows: 1-Specialized workers, 2-Skilled workers, 3-Office, sales workers and C category or equivalent, 4-Technician, foreman, sales representative and B category

These measures are equal to $\gamma = 0.823$ and $\nu = 0.86$

or equivalent, and 5-engineer, executive and A category or equivalent . The table 3.2 below shows this ranking as well as their public sector affiliation and the classification referring to the type of labour (either blue or white collars).

Table 3.1 – Descriptive statistics of the outcome variables

		Survey wave		
	Overall	2013	2016	2019
Characteristic	$N = 10,345$	$N = 4,282$	$N = 3,159$	$N = 2,904$
Monthly wage	1,957.10 (865.11)	1,867.11 (849.58)	1,949.95 (861.03)	2,097.58 (874.13)
Hourly wage	12.87 (4.78)	12.20 (4.54)	12.88 (4.74)	13.84 (5.01)
Profession:				
A category or equiv.	210 (2.03%)	83 (1.94%)	69 (2.18%)	58 (2.00%)
B category or equiv.	259 (2.50%)	97 (2.27%)	85 (2.69%)	77 (2.65%)
C category or equiv.	354 (3.42%)	136 (3.18%)	120 (3.80%)	98 (3.37%)
Engineer/executive	1,794 (17.34%)	713 (16.65%)	535 (16.94%)	546 (18.80%)
Foreman sales rep.	1,107 (10.70%)	446 (10.42%)	328 (10.38%)	333 (11.47%)
Office/Sales worker	2,759 (26.67%)	1,171 (27.35%)	876 (27.73%)	712 (24.52%)
Technician	1,002 (9.69%)	425 (9.93%)	296 (9.37%)	281 (9.68%)
Skilled worker	2,196 (21.23%)	881 (20.57%)	660 (20.89%)	655 (22.56%)
Specialized worker	664 (6.42%)	330 (7.71%)	190 (6.01%)	144 (4.96%)
Occupation (calculated):				
1	664 (6.42%)	330 (7.71%)	190 (6.01%)	144 (4.96%)
2	2,196 (21.23%)	881 (20.57%)	660 (20.89%)	655 (22.56%)
3	3,113 (30.09%)	1,307 (30.52%)	996 (31.53%)	810 (27.89%)
4	2,368 (22.89%)	968 (22.61%)	709 (22.44%)	691 (23.79%)
5	2,004 (19.37%)	796 (18.59%)	604 (19.12%)	604 (20.80%)
Health				
bad or very bad	419 (4.05%)	152 (3.55%)	121 (3.83%)	146 (5.03%)
rather good	2,120 (20.49%)	772 (18.03%)	717 (22.70%)	631 (21.73%)
good	5,362 (51.83%)	2,187 (51.07%)	1,643 (52.01%)	1,532 (52.75%)
very good	2,444 (23.62%)	1,171 (27.35%)	678 (21.46%)	595 (20.49%)
WHO5 score	15.78 (4.92)	15.52 (4.89)	16.10 (4.98)	15.81 (4.89)
SA	3,126 (30.22%)	1,172 (27.37%)	1,007 (31.88%)	947 (32.61%)
SA days	6.30 (19.98)	5.09 (17.24)	6.71 (20.45)	7.63 (22.92)
NA	22	10	6	6

n (%); Mean (SD)

Table 3.2 – Frequency of occupational hierarchical levels using professions

Initial professional categories:		Blue collars	occupational levels				
			1	2	3	4	5
Public	A category or equiv.	No	0	0	0	0	210
	B category or equiv.	No	0	0	0	259	0
	C category or equiv.	Yes	0	0	354	0	0
Private	Engineer executive	No	0	0	0	0	1,794
	Foreman sales rep.	No	0	0	0	1,107	0
	Technician	No	0	0	0	1,002	0
	Office/Sales worker	No	0	0	2,759	0	0
	Skilled worker	Yes	0	2,196	0	0	0
	Specialized worker	Yes	664	0	0	0	0

A, B, and C categories are used to rank the French civil servants

Regarding health outcomes, the respondents are asked to state their *general health state* on a 4-point scale (1-bad or very bad, 2-rather good, 3-good and 4-very good). Given that self-rated health can suffer from evaluation biases such as optimism or pessimism, the duration of sickness absence (SA) is used as a complementary indicator. Indeed, Caverley et al. (2007) show that the duration of SA and self-reported health are ultimately only weakly correlated. The duration of SA is a discrete count variable to which a logarithmic transformation has been applied . About mental well-being, we use the *WHO5 score*, as measured by the World Health Organisation, which ranges from 0 to 25. Zero represents the worst possible and 25 represents the best possible mental well-being in life .

3.3 The identification of competitive presenteeism

In order to identify competitive presenteeism in the data, we first construct an indicator of employee over-commitment based on variables that best reflect the defined concept. It can be seen as a measure of the intensive margin of effort made in addition to the contractual or average effort. Being engaged in CP would mean an excessive level of this indicator. The larger the value of this indicator the more likely the workers are invested in CP behaviour. For the purpose of the identification, we propose to categorise the continuous indicator into 4 modalities. This first allows to underline that the constructed indicator well reflect the intensity of the over-effort intuition and secondly, to highlight some specificities between the individuals with a high or a low score of this indicator. Indeed, as mentioned above, CP is described here as a signal sent by the worker. Therefore, the section below will explore the link between both intrinsic motivations and workload pressure with the levels of the CP indicator, which may explain an increase in investment not related to a signalling objective.

Since individuals can have zero days of absence, the log-transformation of the variable plus one was carried out.

The questions include in the WHO5 score are: I have felt cheerful in good spirits; I have felt calm and relaxed; I have felt active and vigorous; I woke up feeling fresh and rested; and my daily life has been filled with things that interest me. Respondents are asked to scaled from 0 (at no time) to 5 (all of the time).

The competitive presenteeism indicator

Regarding CP, we propose to analyse a set of variables that best reflect the different aspects of its definition. This set of variables is used to build an indicator that summarize the multidimensional aspects of CP using the first factorial axis of a multiple correspondence analysis (MCA). We therefore take into account the regularity of overtime, whether it is compensated in salary, in kind or not, the difference between the declared number of working hours and the average working time by occupation and industry , the frequency with which the workers take their work home, and the frequency with which they think about their job outside their working hours. One MCA is carried out for each wave of the survey .

Table 3.3 – Correlation matrix among the variables of interest

	Work hours gap	Overtime comp.	Work home	Think work
Overtime comp.	0.1517			
Work home	0.1268	0.1856		
Think work	0.1259	0.1975	0.2168	
SP days	0.0290	0.0636	0.0184	0.1045

For the purpose of the MCA, we modified some of the categories of the variables. Overtime and its compensation have been grouped together into a single 5-point scale variable called *compensated overtime* (1. no overtime, 2. sometimes compensated, 3. often or always compensated, 4. sometimes uncompensated and 5. often or always uncompensated). *Taking work home* is an ordinal variable on a 5-point scale (1. not concerned, 2. never, 3. sometimes, 4. often and 5. everyday) and *thinking about work* is also ordinal on a 4-point scale (1. never, 2. sometimes, 3. often and 4. always). Finally, the *work hours gap* from the mean by occupation and industry which was a continuous variable is transformed into a 5-point scale ordinal variable (1. less than the average, 2. the average, 3. 1 to 3 hours higher, 4. 4 to 6 hours higher and 5. more than 6 hours higher). Table 3.3 above shows the correlations among these variables. Finally, the cronbach's alpha for this set

This variable was calculated and previously used by Campbell & Green (2002) who study a possible response to long-term incentives by an increase in the number of hours worked.

The pooled MCA and thus the resulting pooled CP index does not change the results at all.

The initial variable was rounded to the hour. Individuals in the "mean" category have values between −0.5 and 0.5 compared with the occupation-industry average.

of variables is equal to 0.60, corresponding to an acceptable value for the internal consistency of this group of variables.

It may be appropriate to include information on sickness presenteeism (SP) in the CP indicator. SP is supposed to be one of the behaviours in which individuals can invest to signal their commitment to their employer. For instance, Hansen & Andersen (2008) found that being over-committed to work is associated to higher levels of SP. In this sense, it would be a component of the CP indicator. However, it is a behaviour in its own right, largely studied in the literature, and associated with illness. As a result, only individuals facing the shock of illness can seize this opportunity to signal their investment in their work using SP. On another hand, from a technical viewpoint, the omission of SP in the survey in 2019 would restrict the computation of the indicator on two periods, which is not desirable. From a statistical viewpoint also, the correlations between the variables used and the SP variable are fairly weak. Finally, the calculation of the cronbach's alpha clearly leads to the rejection of SP as a variable to include in the set, since it falls from 0.60 to 0.53, showing a deterioration of the internal consistency.

While the literature about effort has often focused on a single aspect of the intensive margin of effort, the idea of the CP indicator is to consider a set of job-related behaviours which, taken together, can send a strong signal to the employer. The advantage of using a synthetic indicator, in addition to best fitted the CP concept, is that we can measure the intensity of this signal. Furthermore, in order to understand CP behaviour among a group of heterogeneous individuals from different occupations, sectors and contracts, it is essential not to focus on overtime alone. For example, some executives are not subject to overtime but will still work a number of hours well above the average for executives in their sector. Thus, the use of an aggregate indicator seems more exhaustive. This method as already been used by F. Green et al. (2021) to construct a "required work intensity" index. Notice that the main difference between the Green's indicator and the one presented here is that the latter is not an indicator of work intensity based on job characteristics, but is composed of information on workers' behavioural choices.

Thus, the first MCA axis alone accounts for 13.3% of the total variability of the

information . The first plan of the MCA allows us to identify three distinct groups of individuals. The first factorial axis of this plan opposes individuals with strongly contrasting behaviours, while the second axis opposes individuals with either a very high or very low investment in CP against those with a moderate investment. The group of workers with strong positive coordinates on the first factorial axis corresponds to individuals who have all the characteristics of CP. These are individuals who work at least more than three hours longer than the average for their occupation and industry, who never stop thinking about work and who always take work home with them. In addition, they very often work unpaid overtime, if not always. On the contrary, workers with strongly negative coordinates on the axis are workers declaring all the lowest modalities of the CP variables. This results are supported by the observation of the Guttman effect in the Figure A3 in the appendices. As a result, the first axis provides a very good representation of the intensity with which individuals are involved in CP behaviour.

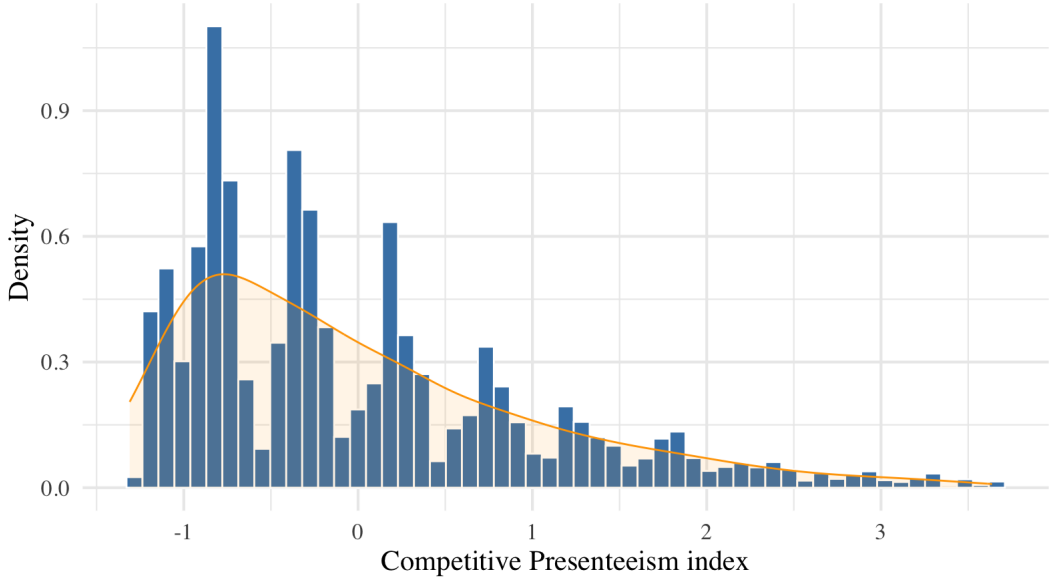

Figure 3.1 – Distribution of the competitive presenteeism index

The figure 3.1 above shows the distribution of the constructed index. After normalization, the CP indicator ranges from −1.31 to 3.64, with a median of −0.28, a mean of zero and a standard deviation of 1. As shown in this figure, its distribution has a skewness of 1.14, indicating a longer right-hand tail towards individuals with

The first two axes represent 21.8% of the variability, which is higher than the reference of 13.8%. These two axes therefore represent a significant proportion of the cloud variability.

higher values. This asymmetry to the left of the median reveals that it is not so common in the population to invest in CP behaviour and that a greater number of workers are below the average of this indicator.

Competitive presenteeism, intrinsic motivations and work pressure

The objective is to identify what we call competitive presenteeism in the population studied. It means, for example, frequently accumulating unpaid overtime, working well above the average working time for the sector and profession, often taking work home and often thinking about work outside working time. We also need to address two issues raised by the CP behaviour. On the one hand, we need to identify whether workers voluntarily engage in these various behaviours and not because they are constrained by work pressure and, on the other hand, whether they do so because they are intrinsically motivated. We present below some descriptive statistics to give a preliminary overview. These statistics are based on a discretization of the CP indicator. The CP indicator has been made discrete here for the simple purpose of carrying out descriptive statistics and testing the existence of significant differences between groups of workers. Furthermore, supplementary variables included in the MCA gave us the opportunity to identify some interesting associations among the information available in the data about these issues.

The table 3.4 below presents the descriptive statistics of the variables included in the CP indicator and the characteristics associated to the intrinsic motivations and pressure at work for each modality of the categorized CP indicator. These modalities have been made on the basis of the quantiles and adjusted with a graphical analysis of the Figure 3.1. The "Low" category corresponds to the workers who are strictly unengaged in over-investment (a CP score from the minimum to -0.49). The "Mean" category corresponds to the workers with a CP score from -0.50 to 0.50 which indicates a behaviour around the norm. The "Mean to 90%" category represents workers who are invested in CP but not excessively so (a CP score from 0.51 to 1.44, i.e. the tenth quantile) and finally the "Higher 10%" are the workers with a score of the CP indicator higher than the 90% of the CP distribution (from 1.45 to the maximum). As a first observation, we can see that high values of the CP indicator correctly reflect the idea of CP behaviour (i.e. working uncompensated

Table 3.4 – Summary statistics of CP characteristics by level of CP index

		Competitive presenteeism categories				
	Overall	**Low**	**Mean**	**Mean to 90%**	**Higher 10%**	
Characteristic	N = 10,345[1]	N = 3,853[1]	N = 3,842[1]	N = 1,613[1]	N = 1,037[1]	**p-value**[2]
CP index	0 (1.00)	-1 (0.17)	0 (0.28)	1 (0.26)	2 (0.57)	< 0.001
Work hours gap	0 (6.24)	-2 (4.65)	-1 (5.25)	2 (7.06)	7 (7.86)	< 0.001
Work hours gap cat.						< 0.001
Less than mean	4,671 (45.15%)	2,120 (55.02%)	1,829 (47.61%)	561 (34.78%)	161 (15.53%)	
Mean	887 (8.57%)	409 (10.62%)	362 (9.42%)	95 (5.89%)	21 (2.03%)	
1 to 3h	2,806 (27.12%)	1,222 (31.72%)	1,117 (29.07%)	342 (21.20%)	125 (12.05%)	
4 to 6h	958 (9.26%)	102 (2.65%)	385 (10.02%)	288 (17.85%)	183 (17.65%)	
More than 6h	1,023 (9.89%)	0 (0.00%)	149 (3.88%)	327 (20.27%)	547 (52.75%)	
Overtime						< 0.001
Never	2,050 (19.82%)	1,441 (37.40%)	538 (14.00%)	56 (3.47%)	15 (1.45%)	
Smts. Comp.	3,544 (34.26%)	2,180 (56.58%)	1,163 (30.27%)	181 (11.22%)	20 (1.93%)	
Oft./everyd. comp.	1,258 (12.16%)	78 (2.02%)	684 (17.80%)	363 (22.50%)	133 (12.83%)	
Smts. uncomp.	2,026 (19.58%)	154 (4.00%)	1,187 (30.90%)	489 (30.32%)	196 (18.90%)	
Oft./everyd. uncomp.	1,467 (14.18%)	0 (0.00%)	270 (7.03%)	524 (32.49%)	673 (64.90%)	
Work home						< 0.001
Not concerned	2,278 (22.02%)	1,206 (31.30%)	837 (21.79%)	188 (11.66%)	47 (4.53%)	
Never	5,773 (55.80%)	2,642 (68.57%)	2,395 (62.34%)	625 (38.75%)	111 (10.70%)	
Sometimes	1,448 (14.00%)	5 (0.13%)	568 (14.78%)	570 (35.34%)	305 (29.41%)	
Often	511 (4.94%)	0 (0.00%)	38 (0.99%)	173 (10.73%)	300 (28.93%)	
Everyday	335 (3.24%)	0 (0.00%)	4 (0.10%)	57 (3.53%)	274 (26.42%)	
Think work						< 0.001
Never	1,645 (15.90%)	1,403 (36.41%)	216 (5.62%)	24 (1.49%)	2 (0.19%)	
Sometimes	4,849 (46.87%)	2,446 (63.48%)	1,958 (50.96%)	384 (23.81%)	61 (5.88%)	
Often	2,826 (27.32%)	4 (0.10%)	1,425 (37.09%)	894 (55.42%)	503 (48.51%)	
Always	1,025 (9.91%)	0 (0.00%)	243 (6.32%)	311 (19.28%)	471 (45.42%)	
Work esteem	1,956 (18.91%)	794 (20.61%)	699 (18.19%)	277 (17.17%)	186 (17.94%)	0.4
Firm pride	2,101 (20.31%)	870 (22.58%)	704 (18.32%)	310 (19.22%)	217 (20.93%)	0.6
Job pride	1,561 (15.09%)	719 (18.66%)	548 (14.26%)	190 (11.78%)	104 (10.03%)	< 0.001
Useful job	1,651 (15.96%)	668 (17.34%)	627 (16.32%)	234 (14.51%)	122 (11.76%)	< 0.001
Lack of time	2,589 (25.03%)	466 (12.09%)	951 (24.75%)	630 (39.06%)	542 (52.27%)	< 0.001
Under pressure	3,566 (34.47%)	565 (14.66%)	1,369 (35.63%)	879 (54.49%)	753 (72.61%)	< 0.001
Ask excessive work	4,245 (41.03%)	1,060 (27.51%)	1,582 (41.18%)	881 (54.62%)	722 (69.62%)	< 0.001

[1] n (%); Mean (SD) ; [2] Pearson's Chi-squared test or Wilcoxon rank sum test to compare only the statistic of the higher group (Higher 10%) to that of all the other groups aggregated

overtime, working many more hours than their peers, regularly taking work home, etc.). The individuals among the higher 10% of the distribution have an average CP score of two, which means that on average they are two standard deviations away from the mean of the CP indicator. They correspond to workers who mostly declare a much higher than average working time, a high frequency of uncompensated overtime, regularly take work home and often think about their job. Thus, a competitively 'presentee' worker, in the sense of this study, would be a worker who accumulates a certain number of the highest modalities of the variables included in the indicator and has a CP score among the highest 10% of the population.

We would now like to identify whether workers who are invested in overcommitment behaviour do so for intrinsic reasons or because of work pressure. The descriptive statistics, reported in Table 3.4, about intrinsic characteristics declared by the workers does not clearly support a link between intrinsic motivations and CP. Firstly, there is no significant difference between workers in the highest CP category

and the rest of the population in terms of the proportion of individuals who are proud of their company or who receive adequate esteem and respect for their efforts. In addition, contrary to what might have been expected, the proportion of workers in the highest CP category who think their work is useful to others or who declare that they are proud of their job is much lower than in the rest of the population.

On the other hand, with regard to work pressure, it should be noted that workers in the highest CP category have a large and significantly higher proportion of individuals reporting that they work under pressure, lack time or are asked to do an excessive amount of work. Therefore, the econometric analysis will consider interaction effects to better identify the signalling effect of CP not associated to work pressure.

Lastly, the analysis of supplementary variables included in the MCA can help to understand who are the individuals with high score of CP. It is interesting to note that the additional variables most representative of this group indicate that these workers are motivated by quantitative objectives and that their working time is not controlled by their employer. In addition, although they have greater latitude in controlling their work and their time, these people state that they are not particularly proud of their work. The MCA results show, on the contrary, that the lower score group is characterized by individuals who do not control their choice of working methods and who have to clock in and out. Furthermore, these individuals have no deadlines and are not concerned by promotion opportunities either. However, this is the group whose workers are the most proud of the job they do. Finally, the intermediate group is composed of workers who have to clock in and out, but are free to adapt their working hours if necessary. They have no quantified targets and are satisfied with their opportunities for promotion within the company. To conclude, we can say that the 10% of individuals with the highest scores on the over-investment indicator are individuals who have more latitude in their jobs and can therefore more easily decide how to behave in the firm. Furthermore, the behaviour of these individuals does not seem to be driven by a pro-social motivation in their job. This group of workers seems to fit perfectly with the definition of presenteeism associated to a competitive objective.

3.4 Additional covariates

In order to isolate and estimate the effect of CP as accurately as possible, we need to control for all the characteristics that influence the utility and costs of effort for the individual, as well as the firm's incentive mechanisms in terms of effort and attendance. This should be done by controlling as far as possible for observable job characteristics and working conditions that differ from one worker to another. We detail below the variables that will be used as controls. A set of covariates are introduced in each models to control for several issues set out in the review of the literature in section 2. Descriptive statistics for these variables can be found in the tables A1 and A2 in the appendix section.

Firstly, we control for *promotion opportunities*, indicating whether the worker declared himself concerned by promotion opportunities and his level of satisfaction with these opportunities. Although the satisfaction of promotion opportunities may reveal barriers to career development, the aim is to control for individuals who declare that they are not concerned by promotion opportunities. Consequently, these individuals have no personal interest in competing for a promotion and therefore should not perceive the usefulness of investing in CP.

For the rest of the covariates, they can be sum up in four categories: the individual's characteristics, the health characteristics, the job characteristics, and the working conditions and environmental context in which the worker evolve. As individual characteristics, we control for *age*, *gender* indicating if the individual is a woman, *marital status*, either in a couple or single, having *children* under 18, the *education* levels in categories, the number of *active in the household* and the worker's *region* of employment. When they are not used as outcomes, the health characteristics, *self-reported health state* and the *WHO5 score*, are included in the models as lagged controls.

The studied employment characteristics are *seniority*, working in the *public sector*, *part time*, the *hourly wage*, working *atypical schedules*, being *supervisor* and having *quantitative targets*. In addition, we control for *sectors*, *occupations* and *firm size*. Controlling for these last three variables makes it possible to incorporate the unobserved characteristics of collective agreements negotiated by unions, relating to working hours and unpaid overtime for example, which may be important in specific

sectors and which affect employees' ability to invest in CP.

As working conditions, we control the work pace constraints imposed by a *machine*, a *demand*, *norms*, *computer* control, *hierarchy* control or *colleagues*. Further, pressure on job demand is controlled through several variables such as working *under pressure*, *lack time* to do the job properly, asking to *work in excess*, being *called out of the working hours* and *working time control* techniques. These variables make it possible to take into account workers who perceive an excessive demand for effort on the part of their employer. This demand would lead to the possibility of non-voluntary presenteeism, which would in fact respond to the pressure exerted on the job. In order to capture only the effects of presenteeism behaviour as a signal, it is essential to control the effects of job pressure as effectively as possible. We also check how much latitude the worker has in his job. To do this, we look at whether they can *vary deadlines* and whether they have to *follow strict instructions* or whether they have the freedom to organize their tasks. Additionally, we add the fact that the worker can receive *help from colleagues* and *help from superiors*. As stated in the literature review in section 2, we control for the worker's intrinsic motivations using information such as when the worker declares that his *job is useful to others*, where he is *proud of his job*, where he is *proud of his company* and finally, when the individual declares that he perceives sufficient *esteem for his work*.

Finally, the last piece of information considered in the models concerns the environment in which the worker operates. We control whether the workers experience major *organizational changes*, such as a redundancy plan, a merger or acquisition of their company, an outsourcing episode, etc. The workers also declare their perception of *job security* by indicating whether they are afraid for their professional future.

4 Method and econometric models

On the one hand, the proposed econometric strategy enables us to correctly reflect the intuition associated with career advancement and the future returns expected by the worker. On the other hand, it best accounts for the direction of causal effect between CP and both career and health outcomes. To do this, it is proposed to

take advantage of the panel structure of the data by studying the past effect of CP behaviour on the worker's present employment situation and health using lagged models.

Furthermore, the analysis of workers' behaviour and effort at work must take into account the issue of individual heterogeneity. This may be due to the worker's ability to withstand risk, effort, or illness to a greater or lesser extent. In the economic literature, this unobserved heterogeneity can be associated with the traditional issues of unobserved individual ability and the healthy worker effect. In this case, ability can refer to the worker's propensity to cope with effort in the broadest sense. The healthy worker effect stipulates that workers in better physical and/or mental condition are in fact those who would self-select into jobs with more difficult working conditions (Haan & Myck, 2009). Furthermore, since these workers are more resistant, they do not perceive the cost of effort associated with difficult working conditions in the same way. Thus, for equivalent working conditions, they would be more likely to be engaged in CP, to succeed in their career and would also report better physical and mental health. Consequently, the use of fixed-effects (FE) estimators that capture individual heterogeneity allows us to control more precisely the unobservable characteristics of workers, such as their resilience to difficult working conditions, but also their risk preferences and competitive spirit. In the following, we detail the models used.

4.1 The lagged models

To explore the relationship between over-effort and both career and health status, we take advantage of the panel structure of the data to control for individual heterogeneity. The principal model of interest is the individual fixed-effects model . The clustered standard-errors at the individual level are provided to address potential heteroskedasticity. Even though the random-effects (RE) estimator is more efficient than the FE estimator, it will not be used as it fails to meet the exogeneity assumptions required between the explanatory variables and the individual random

We do not present the first difference (FD) estimator, even though the exogeneity assumptions are less restrictive, because the models studied drop out the first wave. Thus, the FD estimator and the FE estimator lead to the same estimates since T=2. In addition, the Wooldridge specification test between FE and FD, based on autocorrelation, does not provide conclusive results.

effects. Furthermore, the Hausman specification test leads to conclude that the estimates are always significantly different between RE and FE estimators, leading to the rejection of the null hypothesis and to prefer the use of FE estimators. However, with a view to test the robustness of the results, we thus consider the Hausman-Taylor (HT) estimator into the analysis. The main drawback of the FE estimator is it wipes out time-invariant variables. This omission may aggravate the bias due to omitted variables. For instance, the FE models exclude controls such as *gender*, *education*, *public sector*, *region* and *sectors*. Thus, dropping the between variation would increase the influence of time-varying misspecification on the parameter estimates (Plümper & Troeger, 2019). Moreover, as the individual dimension is large, the fixed effects model leads to a great loss of degrees of freedom. Therefore, the HT estimator is a way to take into account these issues. It is based on instrumental variables (IV), without the need of additional variables. This estimator integrates time-invariant variables and correct for the endogeneity of the variables of interest with the unobserved individual effects. However, whether we are talking about the FE or HT estimator, they do not allow the explanatory variables to be correlated with the idiosyncratic error. This leads us, in any case, to analyse models that correct for endogeneity in order to identify a causal relationship. This question will be addressed later.

Thus, the estimated models are specified as follows:

$$y_{it} = \alpha_i + \beta_1 CP_{it-1} + \beta_2 CP_{it-1}^2 + \mathbf{X}'_{it}\delta_1 + \mathbf{H}'_{it-1}\delta_2 + \mathbf{E}'_{it}\delta_3 + \mathbf{W}'_{it}\delta_4 + \mathbf{G}'_{it}\delta_5 + \mu_{it} \quad (3.1)$$

where y_{it} stands for either the logarithm of hourly wage, occupational grades, health status or WHO5 score. As the logarithm of wage and the WHO5 score are continuous, a traditional estimation by OLS is provided (within and Hausman-Taylor), while the occupation and health outcomes are ordinal qualitative variables leading to the use of ordered logistic models. The α_i corresponds to the individual fixed effects and μ_{it} is assume to be an idiosyncratic error term. The use of the within estimator allows to suppress all the α_i. β_1 and β_2 are the main coefficients of interest in the models related to the CP indicator. The objective is to properly estimate these parameters. The δs are the vectors of coefficients associated to the exogenous explanatory variables. The $\mathbf{X}_{it}$ matrix corresponds to the

set of individual's characteristics, the $\mathbf{H}_{it-1}$ matrix to the lagged health characteristics, the $\mathbf{E}_{it}$ to the one composed of the employment characteristics, and $\mathbf{W}_{it}$ to the working conditions and worker's environmental characteristics. The $\mathbf{G}_{it}$ matrix corresponds to the general controls that will be included in each specifications. There are the *sectors*, the *occupational grades* , the *firm size* and the *survey wave*. To control for multi-collinearity between the explanatory variables and the CP indicator, each matrices, other than $\mathbf{G}_{it}$, will be included one after the other in the estimation. However, as an indication, the variance inflation factor (VIF) does not increase significantly after the inclusion of sets of controls in each model (around 2 at most with all controls). While the model specification is supposed to emphasise a sense of causality by examining lagged effects, potential endogeneity due to reverse causality or simultaneity must still be taken into account (Francesconi, 2001). In addition, time-varying endogeneity may also occur. The following section develops these issues.

4.2 Exploring causality

Although the FE estimator takes into account unobserved heterogeneous individual characteristics, including traditional assumption on the heterogeneity of individual's ability or the healthy worker effect, this method is not assumed to properly account for all the endogeneity of CP. It is important to note that there is a significant gap of three years between the surveys, leaving open the possibility that our results may have been influenced by external factors during this period. These unobserved confounders may affect CP and the outcomes at the same time. Also, it is necessary to take account potential endogeneity due to the non-random sorting of the jobs by workers (Böckerman et al., 2020). Further, it could remain possible dynamic endogeneity, such as specific human capital accumulation, time-varying health problems or the unobserved firm characteristics. For example, on the one hand, the accumulation of human capital can impact the worker's cost of investment in CP but also the probability of promotion associated to higher abilities (Booth et al., 2003). On the other hand, higher human capital would be linked to a better match between the employee and the job, probably leading to reducing stress, pressure,

Not included in the regressions on occupational grades indeed.

workload and thus increasing well-being. Finally, public policies or firm's incentive mechanisms, for which it is not possible to control in our data, may affect the results. It could be sick pay scheme, unions agreements or the employment legislation that constrain the possibility of overtime, and especially unpaid overtime. For example, unionized workers or workers in highly unionized countries work less unpaid overtime compared with the non-unionized (Bell et al., 2000).

Therefore, two strategies are presented to cope with endogeneity in addition to the individual heterogeneity using FE. Firstly, an IV-free approach based on the copula endogeneity correction initially proposed by Park & Gupta (2012) and recently extended by Yang et al. (2024) is presented. Their method is called 2sCOPE for two stages copula endogeneity correction and is directly applicable to time-demeaned variables. Secondly, traditional instrumental variable (IV) models are also proposed as a complementary analysis. While the copula approach has been proposed as a robustness of the IV (Eckert & Hohberger, 2023), there is an ongoing debate about the reliability of estimates using the IV, as well as their strength and validity (Dufour, 2003; Kiviet, 2020; Keane & Neal, 2023). Particularly in this case, in the absence of an instrument directly available in the data, a 'leave-one-out' type instrument is constructed. However, this type of instruments may be questionable (Betz et al., 2018). Furthermore, in the IV models, the controls are all assumed to be exogenous and uncorrelated with the endogenous variable. This assumption is quite strong since many controls are also potentially endogenous and may be correlated with the CP. Thus, in addition to correcting for endogeneity, the 2sCOPE approach also corrects for correlations between the endogenous variable and the exogenous controls. As a result, the copula approach is the preferred candidate to explore endogeneity.

The IV-free endogeneity copula correction approach

The use of a copula to model the link between the endogenous regressor and the error term is especially useful as it captures the link between unobservable characteristics and the error term without the use of an IV. This strategy was first developed by Park & Gupta (2012) and has since been largely extended. The

A blog post from the World Bank by David McKenzie recently sum up the different issues related to this kind of instruments, see the following link.

use of methods with copulas is mainly the appanage of marketing research and is relatively unused in the applied economics literature: 4% in applied economics, 2% in human resources and 2% in econometrics (Qian et al., 2024). However, this strategy is entirely appropriate in the context of microeconometric research, where we are assessing the behaviours of individuals for whom a significant proportion of their variability can be attributed to unobservable and intrinsic characteristics, and when no appropriate IV is available in the data.

According to the guidelines provided by Qian et al. (2024), we do not employ the original method proposed by Park & Gupta (2012) but rather the two stages copula endogeneity correction (2sCOPE) developed by Yang et al. (2024). Although Haschka (2022) proposes an extension of the Park and Gupta's strategy to fixed effects models, our two-time period is unfortunately too short to implement it. However, the 2sCOPE can be apply to the time-demeaned regressors to control for individual heterogeneity (Qian et al., 2024). Therefore, the chosen strategy is to apply the 2sCOPE approach proposed by Yang et al. (2024) on the transformed time-demeaned variables. The last enables to first control for fixed individual heterogeneity and then to apply the copula approach to control for potential endogeneity. Unlike the IV approach, this second approach is less restrictive and will not suffer from a weak instrument problem or a wrong choice of instrument. We present below the 2sCOPE approach and the different required assumptions.

Originally, the method provided by Park & Gupta (2012) consists first of estimating the empirical marginal distribution of the endogenous regressor CP, noted $H(CP)$. Second, using the marginal distribution of the error $G(\mu)$, one can derive the joint distribution function $F(CP, \mu)$. So, according to Sklar's theorem (1959), there exists a copula function that links the two margins such as $F(CP, \mu) = C(H(CP), G(\mu), \theta) = C(U_{CP}, U_{\mu}, \theta)$. For its desirable properties, Park & Gupta (2012) use the Gaussian copula. In this case, the dependence parameter θ can be converted into correlation coefficient ρ between the endogenous regressor and the error term and, in our case, capture the endogeneity of CP. Thus, the joint distri-

To simplify the notations, let $H(CP) = U_{CP}$ and $G(\mu) = U_{\mu}$.

bution function is given by:

$$F(CP, \mu) = C(U_{CP}, U_{\mu}, \rho) = \Psi_{\rho}\left(\Phi^{-1}(U_{CP}), \Phi^{-1}(U_{\mu})\right) \tag{3.2}$$

where Φ is the standard univariate normal distribution function and Ψ_{ρ} denotes the standard bivariate normal distribution with the correlation coefficient ρ. If ρ is significantly different from zero, this highlights that CP is endogenous. According to the assumption that μ follows a normal distribution ($\mu \sim \mathcal{N}(0, \sigma_{\mu})$), Park & Gupta (2012) demonstrate that the structural error is composed of $\mu = \sigma_{\mu}\mu^{*} = \sigma_{\mu}\rho CP^{*} + \sigma_{\mu}\sqrt{1-\rho^{2}}\omega$ where $\mu^{*} = \Phi^{-1}(U_{\mu})$ and $CP^{*} = \Phi^{-1}(U_{CP})$. Therefore, the first part, $\sigma_{\mu}\rho CP^{*}$, captures the endogeneity between CP and μ and the second part, $\sigma_{\mu}\sqrt{1-\rho^{2}}\omega$, where ω corresponds to a Gaussian idiosyncratic error term which is independent of CP. Finally, the method consists in the maximization of the log-likelihood with regards the parameters of the following model: $y_{it} = \alpha + \beta_1 CP_{it} + \mathbf{X}'_{it}\delta + \sigma_{\mu}\rho CP^{*} + \sigma_{\mu}\sqrt{1-\rho^{2}}\omega$. Park and Gupta suggest an augmented OLS estimation of this model with CP^{*} as an additional regressor to correct for endogeneity.

This method is subject to several assumptions. Firstly, the structural error term has to follow a normal distribution. The endogenous regressor must therefore have a non-normal distribution. Indeed, the regressor matrix must be of full rank. If the error and the endogenous regressor both follow a normal distribution, the endogenous regressor CP and its empirical margin CP^{*} are perfectly collinear, which violates the full rank assumption (Qian et al., 2024). Secondly, the control variables must be exogenous and independent of the endogenous regressor.

The 2sCOPE approach allows to relax the uncorrelatedness assumption between the endogenous regressor and the exogenous ones as well as the non-normality of the endogenous regressor. This last point is particularly interesting in our case, because we have first-differentiated our variables, which has led them to approach a normal distribution. In addition, even if 2sCOPE require the normality of the structural error term, it can be derived as long as the endogenous part of the structural error (CP^{*}) is normally distributed and jointly follows a Gaussian copula dependence

In this equation, $\mathbf{X}$ is the matrix that contains all the controls. It therefore includes $\mathbf{X}$, $\mathbf{H}$, $\mathbf{E}$, $\mathbf{W}$ and $\mathbf{G}$ detailed in equation (1).

structure with CP and the control variables. For all of these reasons, we choose to follow the 2sCOPE methodology proposed by Yang et al. (2024) rather than the one by Park & Gupta (2012).

The idea behind 2sCOPE is to jointly model the endogenous regressor with the exogenous ones and the error term using a Gaussian copula model. Similarly to the original method, this implies that CP^*, $\mathbf{X}^*$ and μ^* follows a multivariate normal distribution, where $\mathbf{X}^* = \Phi^{-1}(L(\mathbf{X}))$ is calculated for each exogenous control. Adding CP^* and all the $\mathbf{X}^*$ in the initial model allow to capture endogeneity without requiring the exogeneity of the control variables and their independence with CP. However, the inclusion of many generated regressors associated with the control variables (all the $\mathbf{X}^*$) can lead to significant multicollinearity problems. Therefore, the authors propose a two-stages strategy.

First, after obtaining CP^* and the $\mathbf{X}^*$, the following model is regressed:

$$CP^* = \gamma + \mathbf{X}^{*'}\rho_{\mathbf{x}} + \varepsilon \tag{3.3}$$

By conditioning the initial model on ε, μ becomes independent of CP and all the controls. Thus, the second step consist in the estimation of the augmented model that includes ε, instead of CP^* and $\mathbf{X}^*$ to address the endogeneity bias. This two-stages procedure is similar to the control function approach and the 2SRI already presented in the previous section. However, 2sCOPE does not require any instrument. Another advantage of 2sCOPE is that it includes Park & Gupta (2012) as a special case, where $\rho_{\mathbf{x}} = 0$, the generated regressor ε simply corresponds to CP^*.

Finally, in their guidelines, Qian et al. (2024) clearly explain that in the case of the inclusion of a quadratic effect, which is the case here, it is not necessary to consider this regressor as endogenous and to apply the copula correction as a second endogenous regressor. In fact, controlling the endogeneity with the main term is sufficient to capture and isolate the endogenous part of the regressor with the error. Moreover, the authors explain that computing an additional copula associated with the quadratic term is not a good way to identify the effect. They argue that "this substantially decreases precision of coefficient estimates, and makes copula correction methods perform worse than otherwise".

The instrumental variable strategy

Depending on the nature of the four variables explained, two IV estimation strategies will be used. Since the models used to estimate the effects on hourly wages and the WHO5 indicator are linear, the two stages Generalized Method of Moments (2SGMM) and fixed effects 2SGMM (FE2SGMM) estimators are presented. We use the 2SGMM since heteroskedasticity leads to inconsistent standard errors in IV models, preventing valid inference. The GMM uses orthogonality conditions to allow efficient estimation in the presence of heteroskedasticity of unknown form (Baum et al., 2003).

As the models used to estimate the effects on occupational status and self-reported health are non-linear, we follow Budidharma & Wahyuni (2017) who employ the two-stage residual inclusion (2SRI) method, corresponding to a control function method (Terza et al., 2008; Wooldridge, 2010). Like the 2SLS method, this involves estimating the two stages separately. First, the CP indicator is regressed on the instrument and the set of exogenous regressors. The residuals from the first step are then incorporated into the ordered logit regression of the occupational grades or health status. Since 2SRI is identical to 2SLS for non-linear models, the estimates of the parameters are thus consistent.

Therefore, for the two stages GMM models we have the following relation including the instrumental variables (IV):

$$y_{it} = \alpha_i + \beta_1 \widehat{CP}_{it-1} + \beta_2 \widehat{CP}^2_{it-1} + \mathbf{X}'_{it}\delta_1 + \mathbf{H}'_{it}\delta_2 + \mathbf{E}'_{it}\delta_3 + \mathbf{W}'_{it}\delta_4 + \mathbf{G}'_{it}\delta_5 + \mu_{it} \quad (3.4)$$

where $\widehat{CP}_{it-1}$ and $\widehat{CP}^2_{it-1}$ are instrumented by $\overline{CP}_{-it-1}$, the average of the CP indicator by industry and occupation for each wave not considering the individual and its square $\overline{CP}^2_{-it-1}$. We also include the lag of age and seniority and their square as additional instruments, as recommended by Wooldridge (2010). By adding these additional variables, we can over-identify our IV model and perform the standard Wu-Hausman test and the overidentification tests.

To properly estimate the model on occupational grades, the $\overline{CP}_{-it-1}$ is calculated as the average of the CP by industry and status (labourer, employee, intermediate and executive) rather occupation.

For the 2SRI method, the two stages are calculated as follows. The first step corresponds to :

$$\widehat{CP}_{it-1} = \gamma_{0i} + \mathbf{Z}'_{it-1}\gamma_1 + \mathbf{X}'_{it}\gamma_2 + \mathbf{H}'_{it}\gamma_3 + \mathbf{E}'_{it}\gamma_4 + \mathbf{W}'_{it}\gamma_5 + \mathbf{G}'_{it}\gamma_6 + \varepsilon_{it-1} \quad (3.5)$$

with $\mathbf{Z}_{it-1}$ which includes $\overline{CP}_{-it-1}$, $\overline{CP}^2_{-it-1}$, Age_{it-1}, Age^2_{it-1}, $Seniority_{it-1}$ and $Seniority^2_{it-1}$. We store the residuals ε_{it-1} to be incorporated in the second stage :

$$y_{it} = \alpha_i + \beta_1 CP_{it-1} + \beta_2 CP^2_{it-1} + \beta_3 \varepsilon_{it-1} + \mathbf{X}'_{it}\delta_1 + \mathbf{H}'_{it}\delta_2 + \mathbf{E}'_{it}\delta_3 + \mathbf{W}'_{it}\delta_4 + \mathbf{G}'_{it}\delta_5 + \mu_{it} \quad (3.6)$$

where ε_{it-1} captures the endogeneity that may exist between CP_{it-1} and μ_{it}, thus correcting the estimation bias on β_1 and β_2. The significance of β_3 therefore indicates whether or not there is an endogeneity issue.

The choice of the instrument

The use of an instrumental variable (IV) method lies in the choice of an appropriate instrument for the endogenous regressor. Unfortunately, it is not possible to obtain a valid instrument for the synthetic CP indicator directly from the CT-RPS data. This is why, following the literature, the choice of instrument was made by constructing a new variable. In their paper, Baktash et al. (2022) study the effect of performance pay on stress. To instrument performance pay, they use, for a given individual, the share of workers by occupation receiving performance pay excluding that worker as instrument. This type of instruments are called the 'leave-one-out' instruments. The idea is to identify an instrumental variable based on the aggregation of the information from the worker's environment. This strategy has been already used in the previous literature (Lai & Ng, 2014; Bilanakos et al., 2018; C. P. Green et al., 2021). They argue that a high share of workers receiving performance pay within an occupation increases the individual worker's probability of receiving performance pay.

Their argument is directly applicable to the case of CP. For reasons of peer effects, for example, the probability of an individual revealing a high level of CP is greater if the average level of CP is high in his or her industry and occupation. Indeed, previous literature about peer effects on the individual's productivity highlight

strong association between the average level of effort in a group and the individual level of effort (Falk & Ichino, 2006; Lindquist et al., 2022). Furthermore, Berlinski & Ramos (2020) notably shows, on a sample of teachers, that the promotion of a colleague increases the team's application rate for a professional recognition program. In addition, the closer the colleagues were to the promoted worker, the greater the participation in the competition. Therefore, a work team, or more broadly an entire profession, which is known to show on average a high level of investment and commitment to its work, will have an impact on the behaviour of each of its members. Following Baktash et al. (2022), for a given individual, we calculate the average of the CP indicator by industry and occupation for each wave without considering this individual in the calculation. This variable is supposed to have an impact on the decision to invest in CP, but does not directly explain the individual's career or health. This is because an employer promotes an employee on the basis of what he observes of that employee and not of his or her coworkers. Although the proposed instrument cannot perfectly reflect peer effects, as we cannot observe the direct colleagues of each individual in the database, it is nevertheless assumed that there is a strong identification of individuals with their reference group in terms of profession and sector of activity. Peers are considered here in a broad context.

Consequently, the effect of the IV on career outcomes is supposed to only be explained by its effect on the individual's behaviour in relation to that of his or her peers. This assumption is supported by Cornelissen et al. (2017) who found very weak peer effects in terms of wages. This precisely shows that the remuneration of individuals would indeed be based on individual assessment and not on consideration of the overall productivity of colleagues , thus validating the exclusion assumption of the instrumental variable used. In addition, the IV would also works for the health outcomes. Investment in a high level of effort by colleagues will not directly affect the worker's state of health. It is only through the incentive that the behaviour of the colleagues can have on the worker that the IV can affect his or her health. Once again, the literature on peer effects expressly states that the negative effects on an individual's health come from the adoption of deleterious behaviours. See for example the literature about peer effects among youth (Eisenberg et al., 2014;

An exception, however, would be a salary based on relative performance.

Bramoullé et al., 2020). Thus, the effects of peers on health would only be indirect, through the adoption of behaviour or norms in place within a work group that may be deleterious to health and well-being.

5 Results

In the following section, some descriptive statistics are first presented, including statistics for the variables of interest according to CP level categories as defined in section 3.3. Secondly, the inferential analysis presents the results of the lagged model as well as the models correcting for endogeneity. Then the potential moderating effects associated with work pressure and prosocial job interest are explored. A final part examines the heterogeneity of the effects of CP on the outcomes.

5.1 Preliminary analysis

In order to underline the importance of CP in career development and trajectory as well as in health and well-being, we can consult table 3.5 below. This table shows the average or proportion of individuals for each outcome variable in relation to the past CP categories. The *p-values* of the Chi-square tests or the Wilcoxon rank sum tests are given for the difference between the statistics of the higher category of CP ("Higher 10%") and the rest of the population.

In terms of career, there is a significant difference in wage and occupational grade between individuals whose CP score is in the 10th decile of the distribution and those whose CP score is below this decile. Indeed, regarding wage, there is a significant increase with the CP score, from 12 euros per hour to 16 euros per hour. Further, the wage difference between two successive waves is higher for individuals in the 10th decile of the distribution than for the others.

We can clearly see that an increase in the CP indicator is associated with a higher proportion of individuals with a high occupational grade. As we move from the "Low" CP category to the "Higher 10%" category, the proportion of workers in the first occupational categories (1 and 2) falls significantly, while that in the fifth category rises strongly. Surprisingly, when we look at career paths according to past CP level, the higher the CP level, the lower the change of occupational

Table 3.5 – Summary statistics of the outcomes and their change over time regarding CP index

		Lagged competitive presenteeism categories ($t-1$)				
Characteristic	**Overall** N = 5,170[1]	**Low** N = 2,027[1]	**Mean** N = 1,906[1]	**Mean to 90%** N = 779[1]	**Higher 10%** N = 458[1]	**p-value**[2]
Hourly wage	13 (4.76)	12 (3.80)	13 (4.75)	15 (5.23)	16 (6.10)	<0.001
Δ hourly wage	1 (2.75)	1 (2.47)	1 (2.67)	1 (2.98)	2 (3.64)	<0.001
Occupation						<0.001
1	280 (5.42%)	186 (9.18%)	78 (4.09%)	14 (1.80%)	2 (0.44%)	
2	1,182 (22.86%)	679 (33.50%)	414 (21.72%)	75 (9.63%)	14 (3.06%)	
3	1,465 (28.34%)	627 (30.93%)	579 (30.38%)	185 (23.75%)	74 (16.16%)	
4	1,197 (23.15%)	431 (21.26%)	460 (24.13%)	203 (26.06%)	103 (22.49%)	
5	1,046 (20.23%)	104 (5.13%)	375 (19.67%)	302 (38.77%)	265 (57.86%)	
Δ Occupation						<0.001
−2	122 (2.36%)	54 (2.66%)	53 (2.78%)	12 (1.54%)	3 (0.66%)	
−1	376 (7.27%)	171 (8.44%)	137 (7.19%)	47 (6.03%)	21 (4.59%)	
0	3,868 (74.82%)	1,442 (71.14%)	1,434 (75.24%)	609 (78.18%)	383 (83.62%)	
1	605 (11.70%)	263 (12.97%)	216 (11.33%)	87 (11.17%)	39 (8.52%)	
2	199 (3.85%)	97 (4.79%)	66 (3.46%)	24 (3.08%)	12 (2.62%)	
health						0.5
bad or very bad	228 (4.41%)	79 (3.90%)	98 (5.14%)	32 (4.11%)	19 (4.15%)	
rather good	1,164 (22.51%)	406 (20.03%)	452 (23.71%)	195 (25.03%)	111 (24.24%)	
good	2,723 (52.67%)	1,114 (54.96%)	983 (51.57%)	398 (51.09%)	228 (49.78%)	
very good	1,055 (20.41%)	428 (21.11%)	373 (19.57%)	154 (19.77%)	100 (21.83%)	
Δ Health						0.2
−3	12 (0.23%)	6 (0.30%)	5 (0.26%)	0 (0.00%)	1 (0.22%)	
−2	174 (3.37%)	57 (2.81%)	75 (3.93%)	27 (3.47%)	15 (3.28%)	
−1	1,128 (21.82%)	426 (21.02%)	404 (21.20%)	197 (25.29%)	101 (22.05%)	
0	2,885 (55.80%)	1,179 (58.16%)	1,058 (55.51%)	408 (52.37%)	240 (52.40%)	
1	871 (16.85%)	330 (16.28%)	324 (17.00%)	130 (16.69%)	87 (19.00%)	
2	93 (1.80%)	27 (1.33%)	39 (2.05%)	14 (1.80%)	13 (2.84%)	
3	7 (0.14%)	2 (0.10%)	1 (0.05%)	3 (0.39%)	1 (0.22%)	
health Dec	1,314 (25.42%)	489 (24.12%)	484 (25.39%)	224 (28.75%)	117 (25.55%)	0.095
health Inc	971 (18.78%)	359 (17.71%)	364 (19.10%)	147 (18.87%)	101 (22.05%)	0.2
WHO5 score	16 (4.93)	17 (4.67)	16 (4.92)	15 (4.97)	14 (5.21)	<0.001
Δ WHO5	0 (4.79)	0 (4.60)	0 (4.85)	1 (4.89)	1 (5.12)	0.024

[1] n (%); Mean (SD) ; [2] Pearson's Chi-squared test or Wilcoxon rank sum test to compare only the statistic of the higher group (Higher 10%) to that of all the other groups aggregated

status. There seem to be more transfers up and down for workers who are less involved in CP behaviour than for those who are more involved. For workers with a higher professional status, a high level of CP would be a signal of protection against redundancy.

With regard to declared health, the statistical tests do not show any significant association between a high level of CP and a deterioration in health. From a descriptive point of view, it can be seen that the higher the CP score, the lower the proportion of workers declaring themselves to be in good health. In addition, there was a slight increase in the proportion of individuals declaring themselves to be in rather good or poor health as the CP score increased. The link between the level of investment in work and health does not appear to be linear. It is thus difficult to establish a clear link between health and CP level. This is also true when we look at changes in self-reported health between two survey waves. No particular pattern

seems to emerge. Finally, although it is difficult to justify an association between health and CP, we can clearly see a significant drop in the mental well-being score from 17 for the "Low" CP category to 14 for the "Higher 10%" category. Furthermore, surprisingly, the difference between two waves of the mental well-being score appears to be positive for the two highest categories of the CP indicator, whereas it is zero for the two lowest categories. This difference is significant at the 5% level.

As we have already mentioned in section 3.3, we need to take into account the intrinsic motivations that may drive CP as well as the pressure exerted on the individual that could lead to a high CP score. The following table 3.6 presents the associations of two indicators of intrinsic motivations and two indicators of work pressure with the outcomes.

Table 3.6 – Descriptive statistics of the outcomes regarding intrinsic motivations and work pressure

	Useful job			Job pride			Under pressure			Lack of time		
	0	1		0	1		0	1		0	1	
N	8,694[1]	1,651[1]	p-value[2]	8,784[1]	1,561[1]	p-value[2]	6,779[1]	3,566[1]	p-value[2]	7,627[1]	2,541[1]	p-value[2]
Hourly wage	13.03	12.03	< 0.001	13.10	11.58	< 0.001	12.62	13.34	< 0.001	12.84	12.96	0.003
	(4.80)	(4.58)		(4.84)	(4.20)		(4.63)	(5.03)		(4.84)	(4.62)	
occupation			< 0.001			< 0.001			< 0.001			< 0.001
1	(6.18%)	(7.69%)		(5.75%)	(10.19%)		(7.32%)	(4.71%)		(6.84%)	(5.43%)	
2	(19.85%)	(28.47%)		(18.80%)	(34.91%)		(24.04%)	(15.87%)		(23.52%)	(15.11%)	
3	(28.96%)	(36.04%)		(29.63%)	(32.67%)		(32.14%)	(26.19%)		(31.09%)	(27.15%)	
4	(24.21%)	(15.93%)		(24.50%)	(13.84%)		(21.57%)	(25.41%)		(21.41%)	(26.37%)	
5	(20.80%)	(11.87%)		(21.32%)	(8.39%)		(14.93%)	(27.82%)		(17.14%)	(25.93%)	
health			< 0.001			< 0.001			< 0.001			< 0.001
bad/very bad	(4.05%)	(4.06%)		(4.00%)	(4.36%)		(2.37%)	(7.23%)		(3.04%)	(7.20%)	
rather good	(20.78%)	(18.96%)		(20.51%)	(20.37%)		(17.91%)	(25.41%)		(18.59%)	(26.21%)	
good	(52.55%)	(48.03%)		(52.74%)	(46.70%)		(53.02%)	(49.58%)		(52.54%)	(49.74%)	
very good	(22.61%)	(28.95%)		(22.75%)	(28.57%)		(26.70%)	(17.78%)		(25.83%)	(16.84%)	
WHO5 score	15.35	18.02	< 0.001	15.36	18.15	< 0.001	16.94	13.57	< 0.001	16.48	13.67	< 0.001
	(4.83)	(4.78)		(4.81)	(4.91)		(4.45)	(5.02)		(4.70)	(4.97)	

[1] n (%); Mean (SD); [2] Pearson's Chi-squared test; Wilcoxon rank sum test

On the first hand, We can see that workers who declare characteristics that emphasise their personal attachment to their job (having a useful job and being proud of their job) have slightly lower hourly wages on average. Also, individuals in the two highest occupational categories are less likely to say that they have a useful job and that they are proud of their job, unlike the other three occupational categories. On the second hand, even if there is a significant difference between the distributions of hourly wages of the workers who declared being under pressure or lacking time and the others, the averages are very close. Unsurprisingly, the two highest occupational categories are those with the highest proportion of individuals

working under pressure and lacking time.

Furthermore, in terms of health, we get the expected results: intrinsically motivated workers are on average healthier and happier. Workers declaring intrinsic motivations were less likely to report a good or rather good health status, but more likely to report a very good health. In addition, these workers also have a significantly higher mental well-being score in average. Not surprisingly, individuals who report working under pressure or who lack time to do their tasks are significantly more likely to report poorer health and have a lower well-being score.

These descriptive statistics provide some indication of the average trends in the population. The inferential analysis below seeks to determine whether, all other things being equal, CP can be, at the individual level, a good signal for a career and whether or not it is detrimental to health and mental well-being.

5.2 Inferential analysis

The results for each outcome are presented successively in two subsections. The Tables 3.7 and 3.8 present the results of the models on hourly wages and occupational grades in a first sub-section. Then, in a second sub-section, Tables 3.9, 3.10 and 3.11 present the results on self-rated health, duration of SA and mental well-being respectively. In these tables, columns (1) to (4) present the results of the fixed effects (FE) estimator, column (5) shows the results of the Hausman-Taylor (HT) estimator and columns (6) and (7) display the results of the 2sCOPE and IV models respectively, taking into account the fixed effects.

Since many covariates are included to the models, to check the robustness of the results to possible multicollinearity with the CP indicator, the columns (1) to (4) show the results after controlling for different sets of variables. These sets are indicated at the bottom of the table and corresponds to the matrix defined in section 4. Apart from the results for the effects of the CP on hourly wages and SA duration, which are largely robust to the introduction of controls, we can see that the addition of covariates has a fairly significant impact on the results for the effects on occupational grades and self-reported health. This seems to be particularly true after controlling for working conditions. For example, the coefficient associated with the CP square in the model on occupational status rises from 0.139 to 0.283 and

becomes significant. Similarly, in the models on self-reported health, the coefficient of the linear part of CP raises from -0.247, only significant at the 10% level, to -0.291, significant at the 5% level. These results indicate the importance to control for the characteristics identified so as not to increase the possibility of an omitted variable bias whose effects would pass through the CP indicator. According to that, in the following analysis, the focus will be on the results of the models in column (4) which control for a wide range of covariates. The 2sCOPE and IV models, presented in columns (6) and (7), also correspond to the specification that takes all controls into account.

In addition to the coefficients associated to the CP indicator, the 2sCOPE models report the first stage residuals and the correlation coefficient between the endogenous part of CP and the error term. Both provide information on the importance of endogeneity in the model. Finally, diagnostic statistics are reported at the end of the footnote in each table to determine the relevance of the IV models. The Kleibergen-Paap (KP) rank LM test is an underidentification test using Lagrange multiplier of whether the equation is identified. Under the null hypothesis the equation is underidentified. The Hansen's J statistic corresponds to the traditional Sargan-Hansen test of overidentifying restrictions. The joint null hypothesis is that the instruments are valid instruments. In addition, the Anderson-Rubin (AR) Wald test controls the validity of the instruments and tests the overidentifying restrictions. This test is robust to the presence of weak instruments. On the contrary to the Hansen's J statistic, the null hypothesis is that the coefficients are jointly equal to zero, thus that the instruments are not valid. Lastly, an endogeneity test is provided. Under conditional homoskedasticity, this endogeneity test is numerically equal to the Hausman test. However, the implemented test reports statistics that are robust to various violations of conditional homoskedasticity. Under the null hypothesis, the explanatory variables can be treated as exogenous.

The effect of competitive presenteeism on hourly wage and occupational status

The models in Table 3.7 allows to explain the relation between past investment in CP and current hourly wages. The results show that the higher the CP index the

previous period the higher the contemporaneous hourly wage. An increase of one standard deviation of the CP indicator leads to a wage increase of 1.8%. As seen in Table 3.4, workers in the highest decile are within two standard deviations from the average. Thus, for an increase of two standard deviations, hourly wages would rise by 3.6%. This increase seems roughly equivalent to the 4% increase associated with unpaid overtime over a four-year period on the following year's income (Campbell & Green, 2002). These results are strongly robust to the use of the Hausman-Taylor estimator. Further, the pooled models show very different results . This reflects the importance of considering individual heterogeneity as the pooled OLS would leads to different conclusions.

Table 3.7 – Estimates of the effects of competitive presenteeism on hourly wages

	Fixed effects				HT	FE2sCOPE	FE2sGMM[a]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable:	log(Hourly wage)$_t$						
CP index$_{t-1}$	**0.017****	**0.019*****	**0.018*****	**0.018****	**0.018****	0.023	0.089
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.015)	(0.087)
CP index$^2_{t-1}$	−0.001	−0.002	−0.002	−0.002	−0.001	−0.002	0.008
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.048)
1st stage resid.						−0.001	
						(0.004)	
Corr(CP, μ)						−0.012	
						(0.034)	
Log Lik.	6,544.03	6,553.15	6,565.26	6,620.97	.	.	3,306
N obs.	4,998	4,998	4,998	4,998	4,998	4,998	3,186
N group	3,405	3,405	3,405	3,405	3,405	.	1,593
R^2	0.15	0.15	0.16	0.17	.	0.168	0.07
Adj. R^2	0.15	0.15	0.15	0.17	.	0.162	
Individual's controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Health controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Employment controls	No	No	Yes	Yes	Yes	Yes	Yes
Working conditions	No	No	No	Yes	Yes	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; clustered standard errors at the individual level in parenthesis. HT: Hausman-Taylor estimator ; FE2sCOPE: two stages copula endogeneity correction on time-demeaned variable ; FE2sGMM: FE 2 stages GMM.
Individual and health controls include age, marital status, number of employed in the household, child, lagged declared health and lagged WHO5 score. Employment controls include seniority, supervisor, part time, atypical schedules, quantitative targets, occupations and firm size. Working conditions include work paced by machine, external demand, norms, computer control, hierarchical control or colleagues, work pressure, excess work demand, lack of time, working time control, called out of work, help from colleagues or supervisor, fear of job future, organizational changes, work esteem, job and firm pride, and useful job. HT also include gender, education, sectors and regions.
[a] IV diagnostic statistics: KP rk LM statistic $\chi^2(5)$ is 3.17 ; Hansen J statistic $\chi^2(4)$ is 3.85 ; AR Wald test $F(6)$ is 1.00 ; Endogeneity test $\chi^2(2)$ is 1.13.

The results in Table 3.8 show that past CP would increase the probability of being at higher occupational grades in the future. This effect passes through the quadratic

These results are reported in the tables presenting all the models in the appendix section

term while the linear term does not appear significant at 5%. Nonetheless, this effect seems to indicate a U-shape curve. In the case of the model (4), the turning point for which an increase of the CP indicator would lead to an increase in the probability of being in the higher occupational grades is 0.83. For instance, given an initial CP score of 1, a one standard deviation increase in the indicator leads to a 10% increase in the odds ratio for all occupational categories , all other things being equal. This shows that the positive effect of CP only appears for higher values of the indicator. According to the variable used as a proxy of career advancement, this effect well reflect the framework of tournaments (Bratti & Staffolani, 2007) in which only the most engaged worker would be promoted.

Table 3.8 – Estimates of the effects of competitive presenteeism on occupational grades

	Fixed effects Ordered Logit				HT	FE2sCOPE	FE2SRI[a]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable:				Occupation$_t$			
CP index$_{t-1}$	−0.143	−0.176	−0.230	−0.470*	−0.040	−0.044	2.185
	(0.173)	(0.182)	(0.195)	(0.247)	(0.028)	(0.054)	(2.312)
CP index$^2_{t-1}$	0.146	0.123	0.139	**0.283****	0.021*	0.020*	**0.276****
	(0.101)	(0.104)	(0.111)	(0.129)	(0.012)	(0.012)	(0.130)
1st stage resid.						0.000	−2.671
						(0.014)	(2.323)
Corr(CP, μ)						0.001	
						(0.036)	
Log Lik.	−317.82	−313.74	−304.14	−269.15	.		−268
N obs.	766	766	766	766	4,998	4,998	766
N group	383	383	383	383	3,405	.	383
R^2 (pseudo R^2)	0.06	0.07	0.10	0.20	.	0.034	0.20
Adjusted R^2						0.027	
Individual's controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Health controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Employment controls	No	No	Yes	Yes	Yes	Yes	Yes
Working conditions	No	No	No	Yes	Yes	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; clustered standard errors at the individual level in parenthesis. HT: Hausman-Taylor estimator ; FE2sCOPE: two stages copula endogeneity correction on time-demeaned variable ; FE2SRI: FE two stages residual inclusion.
Individual and health controls include age, marital status, number of employed in the household, child, lagged declared health and lagged WHO5 score. Employment controls include seniority, supervisor, part time, atypical schedules, quantitative targets, hourly wage and firm size. Working conditions include work paced by machine, external demand, norms, computer control, hierarchical control or colleagues, work pressure, excess work demand, lack of time, working time control, called out of work, help from colleagues or supervisor, fear of job future, organizational changes, work esteem, job and firm pride, and useful job. HT and pooled models also include gender, education, sectors and regions.
[a] IV diagnostic statistics: KP rk LM statistic $\chi^2(5)$ is 1.23 ; Hansen J statistic $\chi^2(4)$ is 1.69 ; AR Wald test $F(6)$ is 1.40 ; Endogeneity test $\chi^2(2)$ is 0.08

Furthermore, the two models which control for endogeneity, columns (6) and (7),

the calculation is $\partial Occupation/\partial CP = e^{-0.470}e^{2\times 0.283 CP}$, with $CP = 1$, $e^{-0.470}e^{0.566} = 1.1008$.

do not show significant results either for hourly wages or for occupational position. Only the coefficient associated with the quadratic term of the fixed effects IV model on hierarchical grades shows a positive and significant effect, which corroborates the previous results. On the other hand, the linear term of this model is strongly positive. Even if it is not significant, this may indicate limitations in the instrument used. Finally, in each method, 2sCOPE and IV, for both hourly wages and occupational categories, none of the indicators revealing the presence of endogeneity are significant (neither the residuals from the first stage, nor even the correlation between the endogenous part of the CP and the structural error, as well as the diagnostic statistics associated with the IV models). This leads to the conclusion that the fixed effects models already capture a significant proportion of the endogeneity associated with individual heterogeneity. This conclusion is also supported by the results of the 2sCOPE and IV models on the pooled models which, on the contrary, strongly show the presence of endogeneity affecting the coefficients.

These results therefore tend to validate the first hypothesis developed in section 2. Further, they are in line with past literature showing the effects of overtime on earnings and promotion in Britain and Germany for instance (Francesconi, 2001; Pannenberg, 2005; Anger, 2008). In the same vein, CP behaviour would indeed be associated with a better professional situation in the future in terms of wage and occupational status. It would therefore be a valuable signal for the employer who pays for this behaviour.

The effect of competitive presenteeism on self-rated health, sickness absence and mental well-being

Let's turn now to the results of the effects of CP on self-rated health, SA duration and mental well-being in the following Tables 3.9, 3.10 and 3.11. As previously indicated, the introduction of controls appears to have an impact on both the significance and the magnitude of the effect of CP on self-reported health. From -0.185 when only individual characteristics are introduced to -0.291 when all characteristics are controlled for. Following the results of model (5), the preferred one, CP reduces the probability of moving up to the higher health categories. An increase

Table 3.9 – Estimates of the effects of competitive presenteeism on self-rated health

	——— Fixed effects Ordered Logit ———				HT	FE2sCOPE	FE2SRI[a]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable:				Health$_t$			
CP index$_{t-1}$	−0.185	−0.228*	−0.247*	**−0.291****	**−0.073****	**−0.148****	2.031
	(0.127)	(0.125)	(0.126)	(0.136)	(0.032)	(0.063)	(2.285)
CP index$^2_{t-1}$	0.113	0.120*	0.115*	0.102	0.030*	**0.038****	0.102
	(0.070)	(0.067)	(0.067)	(0.070)	(0.017)	(0.015)	(0.071)
1st stage resid.						0.026	−2.336
						(0.016)	(2.294)
Corr(CP, μ)						0.063*	
						(0.035)	
Log Lik.	−507.25	−504.24	−499.29	−474.24	.		−473.70
N obs.	1,514	1,514	1,514	1,514	4,998	4,998	1,514
N group	678	678	678	678	3,405	.	678
R^2 (pseudo R^2)	0.03	0.04	0.05	0.10	.	0.041	0.10
Adjusted R^2						0.034	
Individual's controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Health controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Employment controls	No	No	Yes	Yes	Yes	Yes	Yes
Working conditions	No	No	No	Yes	Yes	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; clustered standard errors at the individual level in parenthesis. HT: Hausman-Taylor estimator ; FE2sCOPE: two stages copula endogeneity correction on time-demeaned variable ; FE2SRI: FE two stages residual inclusion.
Individual and health controls include age, marital status, number of employed in the household, child, lagged WHO5 score. Employment controls include seniority, supervisor, part time, atypical schedules, quantitative targets, occupations, hourly wage and firm size. Working conditions include work paced by machine, external demand, norms, computer control, hierarchical control or colleagues, work pressure, excess work demand, lack of time, working time control, called out of work, help from colleagues or supervisor, fear of job future, organizational changes, work esteem, job and firm pride, and useful job. HT and pooled models also include gender, education, sectors and regions.
IV diagnostic statistics: [a] KP rk LM statistic $\chi^2(5)$ is 1.52 ; Hansen J statistic $\chi^2(4)$ is 1.65 ; AR Wald test $F(6)$ is 1.82 ; Endogeneity test $\chi^2(2)$ is 0.07.

of one standard deviation in the CP indicator therefore leads to a 25% drop in this probability. Accordingly, increasing the CP score by two standard deviations, in order to reach the level of the last decile of the CP distribution, leads to a sharp drop of almost 56% in the probability of declaring a better health status. Models (2) and (3) tend to indicate a U-shaped curve for the effect of CP on health. Above a CP score of around 1, an increase in the level of investment in employment would be associated with an improvement in health. However, these results are only significant at the 10% level and do not take into account all the controls associated with working conditions.

With regard to endogeneity in the self-reported health models, the two-stage residuals inclusion model does not provide significant results and the coefficients appear to be greatly inflated. As with the model on occupational status, this probably reveals that the instrument used is too weak. On the contrary, the 2sCOPE model

allows to find and confirm this U-shape curve, even if the first stage residuals as well as the correlation coefficient are not significant. As a reminder, this model does not fit well for discrete outcomes. However, even if Breitung et al. (2024) indicate that OLS performs better than models with endogeneity correction in the case where there is no endogeneity, the 2sCOPE method also corrects for correlations between the endogenous variable and the controls, which nevertheless leads to interesting results.

The turning point at which an increase of one standard deviation rises the probability of reporting a better health is 1.95, close to the average CP of the top 10th decile. Thus, for a CP score below this threshold, a one standard deviation increase in this score is still associated with a significant drop in the probability of reporting better health. However, for workers with an initially high CP score, increasing their investment in CP behaviour would be associated with better future health. The previous results led us to conclude that workers who invest the most in CP behaviour are rewarded in terms of career and wages. An initial explanation for the positive effect of CP on self-rated health could therefore be the fact that rewarded workers no longer feel entitled to declare themselves to be in poor health given their hierarchical position and their wage level. Individuals would be grateful to their employer to see their efforts properly rewarded. This reward could also affect the worker's satisfaction in his job, corresponding to higher satisfaction with the financial and professional situation. In their article, Francesconi (2001) point out that an increase in wages and career advancement go hand in hand with greater job satisfaction. Furthermore, employees who have a high level of autonomy in determining their work pace or the intensity of their effort, characteristics mainly associated to higher hierarchical positions, would be less inclined to self-diagnose as sick (Johansson & Lundberg, 2004).

The results in Table 3.10 below show a positive effect of CP on the lengthening of SA duration. In each model, only the quadratic term is significantly positive. This would indicate that the increase in the duration of SA would be significantly greater for the workers most involved in CP behaviour. The values of the coefficients are also fairly stable in all the models controlling for individual heterogeneity. In addition, the FE and HT estimators give exactly the same results. Although not

Table 3.10 – Estimates of the effects of competitive presenteeism on sickness absence duration

	Fixed effects				HT	FE2sCOPE	FE2sGMM[a]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable:	log(SA days)$_t$						
CP index$_{t-1}$	0.031	0.035	0.035	0.038	0.039	−0.072	−0.082
	(0.073)	(0.073)	(0.073)	(0.072)	(0.072)	(0.248)	(1.129)
CP index$^2_{t-1}$	**0.078****	**0.080****	**0.077****	**0.075****	**0.075****	**0.100****	−0.085
	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.049)	(0.429)
1st stage resid.						−0.003	
						(0.059)	
Corr(CP, μ)						−0.025	
						(0.041)	
Log Lik.	−5,071	−5,046	−5,030	−4,975	.		−3,935
N obs.	4,988	4,988	4,988	4,988	4,988	4,998	3,174
N group	3,405	3,405	3,405	3,405	3,405	.	1,587
R^2	0.02	0.02	0.03	0.05	.	0.191	0.02
Adj. R^2	0.01	0.02	0.03	0.04	.	0.185	
Individual's controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Health controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Employment controls	No	No	Yes	Yes	Yes	Yes	Yes
Working conditions	No	No	No	Yes	Yes	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; clustered standard errors at the individual level in parenthesis. HT: Hausman-Taylor estimator ; FE2sCOPE: two stages copula endogeneity correction on time-demeaned variable ; FE2sGMM: FE 2 stages GMM.
Individual and health controls include age, marital status, number of employed in the household, child, lagged declared health and lagged WHO5 score. Employment controls include seniority, supervisor, part time, atypical schedules, quantitative targets, occupations, hourly wage and firm size. Working conditions include work paced by machine, external demand, norms, computer control, hierarchical control or colleagues, work pressure, excess work demand, lack of time, working time control, called out of work, help from colleagues or supervisor, fear of job future, organizational changes, work esteem, job and firm pride, and useful job. HT and pooled models also include gender, education, sectors and regions.
IV diagnostic statistics: [a] KP rk LM statistic $\chi^2(5)$ is 1.52 ; Hansen J statistic $\chi^2(4)$ is 1.50 ; AR Wald test $F(6)$ is 0.74 ; Endogeneity test $\chi^2(2)$ is 0.09.

significant, the linear term is also positive. Thus, according to the results of model (4), given an initial CP score of 1, a one standard deviation increase is associated with an 18.8% increase in the duration of SA. As before, the IV model does not show any significant results and the 2sCOPE model does not indicate the presence of endogeneity in the model controlling for individual heterogeneity. However, as stated here-above, the 2sCOPE allows to control for the correlations among the variables. Thus, according to this last model, an increase of one standard deviation in the CP indicator, given an initial score of 1, is related to an increase of 12.8%. The 2sCOPE model, contrary to the FE ones, indicates a U-shape curve. However, in all the models, the linear term is not significantly different from zero. Finally, with regard to the models on the WHO5 score, none of them showed a significant effect. The results seem to lead to a robust conclusion that CP behaviour would not be associated with mental well-being.

Table 3.11 – Estimates of the effects of competitive presenteeism on mental well-being

	Fixed effects				HT	FE2sCOPE	FE2sGMM[a]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable:	WHO5$_t$						
CP index$_{t-1}$	0.160	0.127	0.110	0.011	−0.016	−0.094	1.696
	(0.193)	(0.190)	(0.190)	(0.182)	(0.183)	(0.376)	(2.275)
CP index$^2_{t-1}$	−0.106	−0.108	−0.096	−0.110	−0.097	−0.073	0.592
	(0.114)	(0.112)	(0.113)	(0.106)	(0.108)	(0.102)	(1.068)
1st resid. stage						0.031	
						(0.095)	
Corr(CP, μ)						0.009	
						(0.037)	
Log Lik.	−10,124	−10,086	−10,074	−9,882	.		−7,290
N obs.	4,998	4,998	4,998	4,998	4,998	1,593	3,186
N group	3,405	3,405	3,405	3,405	3,405	.	1,593
R^2	0.02	0.03	0.04	0.11	.	0.085	−0.06
Adj. R^2	0.01	0.03	0.03	0.10	.	0.079	
Individual's controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Health controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Employment controls	No	No	Yes	Yes	Yes	Yes	Yes
Working conditions	No	No	No	Yes	Yes	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; clustered standard errors at the individual level in parenthesis. HT: Hausman-Taylor estimator ; FE2sCOPE: two stages copula endogeneity correction on time-demeaned variable ; FE2sGMM: FE 2 stages GMM.
Individual and health controls include age, marital status, number of employed in the household, child and lagged declared health. Employment controls include seniority, supervisor, part time, atypical schedules, quantitative targets, occupations, hourly wage and firm size. Working conditions include work paced by machine, external demand, norms, computer control, hierarchical control or colleagues, work pressure, excess work demand, lack of time, working time control, called out of work, help from colleagues or supervisor, fear of job future, organizational changes, work esteem, job and firm pride, and useful job. HT and pooled models also include gender, education, sectors and regions.
[a] IV diagnostic statistics: KP rk LM statistic $\chi^2(5)$ is 1.56 ; Hansen J statistic $\chi^2(4)$ is 1.51 ; AR Wald test $F(6)$ is 0.86 ; Endogeneity test $\chi^2(2)$ is 0.09.

To conclude, the results globally support the intuition about the effects of CP on health, leading to validate the third hypothesis a) and c) about self-rated health and SA duration. However, the results clearly reject the third hypothesis b) concerning the deleterious effects on mental well-being. These results partly support the previous findings associated to constrained working time and difficult working conditions (Barnay, 2016). An important point to highlight in our case is the voluntary aspect of CP behaviour. If CP is indeed a voluntary signalling behaviour, these results shed new light on the two facets of hard work, since a voluntary investment in CP would also lead to a deterioration in health. However, contrary to previous literature, the results presented do not confirm the existence of any effect on mental well-being. When work intensity is a choice, and not a pressure exerted by employment characteristics or working conditions, this over-investment has no effect on mental well-being. A psychological explanation may justified these results. The

most committed workers are those who find usefulness (value) in the spirit of competition. The most competitive workers are the ones who show the most creativity and adaptability, thus gaining greater self-esteem and having a more dynamic and positive spirit (Shimazu et al., 2012). Further, Shimazu et al. (2015) already found that work engagement is associated to better health and happiness while, on the contrary, workaholism is a predictor of long-term future unwell-being and ill-health. Therefore, the next section explores the moderating effects of work pressure in addition to those of intrinsic motivations to confirm the effects of a voluntary choice of CP.

5.3 The moderated effects of prosocial and work pressure characteristics

This section presents the results of the models where an interaction term has been included with the CP score. This was done in order to check whether the main results above are not influenced either by the pressure to which workers are subjected or by the their intrinsic motivations. The tables underlying these results are reported in the appendix section 3. The following comments are based on the results of the preferred model that take into account all the controls , corresponding to the specification already presented in the main results.

Work pressure

About the potential moderating effects associated with job pressure, the results do not reveal any significant difference on hourly wages between workers who declare themselves to be working under pressure or lacking time and all other workers. Moreover, in both cases, the effect of the CP on wages for workers who do not declare characteristics of pressure at work increases slightly, from 1.8% in the initial models to 2%. On the other hand, there is also no significant effect of the interaction terms in the occupational models. However, even if the coefficients becomes only significant at the 10% level, the results indicate that an increase of one standard deviation of the CP score, given an initial score of 1, is related to an increase between

These models correspond to the ones in column (5) in the appendix tables. The models using the HT estimator are not reported as HT does not support interaction effects.

21% and 25% in the probability of moving up the occupational grades, showing an important increase of the effect compared to the 10% in the initial results. In addition, the fact of declaring a lack of time to carry out tasks would be associated with a lower occupational situation in the future. These additional results tend to confirm and reinforce the signalling meaning of CP behaviour. The initial results would be borne by individuals who do not declare characteristics of pressure at work. Thus, the behaviour identified would indeed be a voluntary investment behaviour which is a valuable signal. These results seem to be in line with the intuition that it is not necessary to reward workers whose excess effort is constrained by the characteristics of the job since the firm already impose this high level of effort. It is thus not possible for them to signal over-effort given that it is required by the firm. On the contrary, employees who engage in CP behaviour even though nothing obliges them to do so send a signal that is correctly interpreted and rewarded by the firm.

The results on self-rated health show an increase in the negative effect of CP for workers who do not reveal any pressure at work. A one standard deviation increase in the CP score, which initially led to a 25% drop in the probability of being in better health, would lead to a 28% to 38% drop in the models interacting with lack of time and working under pressure respectively. Moreover, contrary to the initial results, the quadratic effect never tends to be significant in these models. It is interesting to note that, if people who do not declare characteristics linked to work pressure report poorer health in the future, the increase in the CP score for people who declare working under pressure would not be linked to poorer self-rated health in the future. In addition, an increase in CP for workers who declare a lack of time to perform their tasks is associated with a significant reduction in the duration of SA compared with other workers. Finally, an increase in the CP score would be related to a future lower mental well-being for the workers who do not declare lacking time. However, this effect is only significant at the 10% level.

Overall, these results tend to confirm the general intuition associated with a voluntary behaviour that signals one's ability to exert an additional effort. Furthermore, this voluntary over-investment seems to be detrimental to workers' health, all other things being equal.

Intrinsic motivations

As regards the moderating effect of intrinsic motivation in the job, there was no significant difference between workers declaring that they had a job that was useful to others or that they were proud of their work on the effect of CP on hourly wages. However, while the previous results on the effect of CP on the probability of increasing the occupational grade are confirmed, the interaction effects show that workers declaring both intrinsic motivation characteristics have a lower probability of moving towards higher hierarchical occupations. These results tend to confirm the second hypothesis and the idea that intrinsically motivated workers probably do not send the same signal to the employer and are not promoted in higher hierarchical levels.

With regard to the effects on the health outcomes, we can notice that the results of the fixed effects models on the WHO5 score remain insignificant. We also note the robustness of the initial results with regard to the duration of SA. No interaction term indicates a significant effect. The effect for non-intrinsically motivated workers is very close to the effect previously found. An increase of one standard deviation would be associated with a 7.6 to 8.2% increase in the future duration of SA. On the opposite, the results on self-reported health are quite impacted by the introduction of the interaction terms. Even if there is no significant difference between the intrinsically and non-intrinsically motivated workers in the effect of CP on self-rated health, when the CP is interacted with useful job, the negative linear effect is no longer significant while the positive quadratic effect becomes significant at the 10% level only. Further, when CP is interacted with job pride, we find again the U-shaped effect of CP on health we identified in the 2sCOPE model. The turning point for which an additional increase of the CP score lead to an increase in health stands at 1. Therefore, unlike the effect on occupational grades, there is no significant difference in the effects on the health outcomes between workers who are intrinsically motivated and those who are not. However, these results tend to confirm a U-shape curve of the effect of CP on self-rated health.

5.4 Additional results

This section depicts some additional results to explore the robustness of the results to some heterogeneity. First, we present the results associated to the interaction terms regarding gender, working in public sector and the type of occupation (blue or white collars). Also, we use the subset of the healthcare sector workers to explore the differences between the main results and those of this sector. It is assumed that the signalling intuition of the CP would not be the same between the healthcare sector and all the others. Some differences are expected according to the nature the tasks in the healthcare sector, that can be related to pro-social motivations as well as to the constraints on attendance, working time and overtime. In addition, we show the heterogeneous effect between public and private sectors among the healthcare industry. Finally, we present the results of the models where the variables used to construct the CP index are directly regress in the models. The tables showing these results are available in the appendix section 4.

In general, the results are largely robust to interaction effects, revealing no strong heterogeneity effects. Firstly, there is no significant difference between men and women, a results already shown by Francesconi (2001). Similarly, there is no change in the main effects after the interaction of the CP indicator with the public sector. The significant negative linear effect on health would be much greater for the public sector than for the private sector. Finally, there is no significant difference between blue and white collar workers.

The results of the models on the sub-sample of workers in the healthcare sector lead to rather different conclusions. Firstly, while the linear effect of CP on the hourly wage appears to be the same, the quadratic effect tends to be significantly positive. On the other hand, the results of the fixed effects models on occupation show a negative link between CP and the probability of increasing occupational grades. This surprising result may reflect either the limitation of the explained variable corresponding to occupations which can be not enough detailed, or the high heterogeneity that may exist in the healthcare sector. Indeed, in this sector, inter-professional mobility is relatively rare. The hierarchical structure of healthcare establishments is more conducive to career progression within a given profession.

The occupations among the blue and white collars are reported Table 3.2.

Progression to occupations at a higher hierarchical level depends on qualifications and degrees. However, it is important to note that, contrary to the previous heterogeneity results, there is much higher heterogeneity in the effects of CP on hourly wages between the public and private healthcare sectors. In fact, once this heterogeneity is taken into account, the effect of CP on hourly wage at increasing rates disappears to show a much greater linear effect in the private sector than in the public one. An increase of one standard deviation in the CP score leads to an increase of 3.6% in hourly wages in the private healthcare sector compared with the 1.1% in the public healthcare sector. Moreover, the results indicate no effect of CP on self-rated health, SA duration and mental well-being. This absence of significant results on health outcomes in the healthcare sector sub-sample should not lead to the conclusion that there is no deleterious effect, but rather that workers involved in CP behaviour in this sector do not declare a state of health, or a duration of absence, significantly different from the average in this sector. Furthermore, as stated earlier in this chapter, it is possible that the CP indicator does not reflect exactly the same thing in the healthcare sector than in other sectors. Finally, it is also possible to assume a selection bias relating to the health and/or psychology of workers in this sector, making it impossible to identify an effect on health.

Finally, instead of the CP indicator, we regress the outcomes directly on the variables that compose the indicator obtained from the MCA. Overall, the results of the models with fixed effects show that few factors have a significant effect on the outcomes. In terms of hourly wages, the greater the difference between the number of hours worked and the average in the occupation and sector, the higher the wages in the future. For example, a difference of more than 6 hours is associated with a 5.8% increase in pay. Moreover, the fact that you often take work home increases wages by 6.9%. This result particularly highlights the signalling nature of CP. Overtime, whether compensated or not, had no effect on earnings. In addition, the fact of regularly thinking about one's job is associated with a 1.6% fall in hourly wages, only in the HT model however. As regards occupational status, only working between 4 and 6 hours more than the sector and industry average is associated with a higher probability of moving into the higher occupational categories. There is no significant effect from any other variable. As regards health and well-being, here again there are

few significant effects. Nevertheless, working more than 6 hours above the average would have a negative effect on the self-reported health, significant at only 10%. This is also the case when the worker takes work home only sometimes. In accordance with the main results on mental well-being, no variable appears to have a significant effect on the WHO5 score. All of these results confirm the robustness of using the CP index, which better summaries the information associated with CP behaviour.

6 Conclusion and discussion

The aim of this chapter was to explore the effects of competitive presenteeism (CP) on the worker's career, through the hourly wage and occupational grades, but also on health through self-reported health, sickness absence (SA) and mental well-being. To this end, a synthetic CP indicator was created on the basis of variables that provide the best possible information on the nature of the worker's over-investment or over-effort in his or her job. It is assumed that this type of presenteeism behaviour can be used by employees to signal their ability and motivation to their employer(Holmstrom, 1982; Anger, 2008). This study is part of the limited literature on presenteeism behaviour not associated with illness. It is also the first to consider a multifactorial measure of this behaviour which synthesises different aspects of the effort intensity.

While the high level of investment in effort that characterises CP was supposed to be a valuable signal for employees career, it was also assumed to have a counterbalancing effect on workers' health. Estimating such effects remains a major challenge. Indeed, many studies on these subjects have focused on a single aspect of the intensive margin of effort. As a result, it has been shown that working time alone cannot constitute a health risk factor (Ten Brummelhuis et al., 2017), or that unpaid overtime has only a relatively small impact on a worker's wage and career. Consequently, using this multi-factorial indicator makes possible to measure the intensity of the signal an employee wishes to send to his employer.

In addition, we need to control for a set of covariates that can explain both workers' career and health. Indeed, it is necessary for instance to decorrelate the signalling effect of CP from the effect induced by intrinsic motivations or from

the effect of deleterious working conditions exerting pressure on the worker. This was made possible by the use of the last three waves of the "Working conditions and psychosocial risks" survey, which provide a very wide range of information on individual and professional characteristics, working conditions and the worker's environment. However, an additional difficulty that has not been much addressed in the literature is the endogeneity between CP and job characteristics or workers' health. It is important to establish the direction of causality between CP behaviour and the various possible outcomes (Böckerman, 2018). To this end, two approaches have been used in order to confirm or refute the main results. The first corresponds to a modern approach based on copula to correct for endogeneity and proposed by Yang et al. (2024). The second approach uses traditional instrumental variable models.

Looking at the lagged effects of the CP indicator on hourly wages and occupational grades, the results show that being involved in his or her job (e.g. working many more hours than the average of your industry and occupation, always working unpaid overtime, taking work home everyday, and always thinking to his or her job) would be well valued and rewarded by the firms. The results confirm both the effectiveness of the CP behaviour as a rewarded signal and the potential deferred compensation it may involve. Using individual fixed-effects models, we find that a one standard deviation increase in the indicator corresponds to a significant increase in the hourly wage from 1.8% to 2%, as well as an increase in the probability of moving up the career ladder by 10%. Furthermore, an increase of two standard deviations in the CP indicator, reflecting a shift to the highest decile of the population invested in CP, is associated with a 3.6% rise in the hourly wage.

With regard to the results associated with self-rated health, the duration of SA and mental well-being, it can be generally concluded that CP has a deleterious effect. Taking individual heterogeneity into account, a decrease of 25% in the probability of declaring a better health status has been found for a one standard deviation increase in the CP indicator. Furthermore, this increase of one standard deviation of the indicator also leads to an increase of nearly 19% in the duration of SA. While CP seems to lead to a robust increase in the duration of SA for the workers the most involved in CP, some results tend to highlight a possible U-shaped curve in

the effect on self-rated health. These two opposing results could be explained by the difference in the measurement used. Caverley et al. (2007) shown that there is ultimately little correlation between SA and self-rated health. The duration of SA, even if it is declarative, is expected to provide reliable information on the worker's health. Self-reported health, on the other hand, may suffer from psychological bias.

Consequently, the explanation put forward to justify this potential positive effect is that, despite a high level of effort and the fatigue that CP may generate, workers involved in CP are rewarded by the firm in terms of wage and career advancement. In turn, worker would not feel the right to declare themselves unhappy or in poor health. The better workers are rewarded and paid for their past effort, the less they feel entitled to call in sick. Moreover, the well balance between the costly effort in CP and the firm rewards may simply make workers satisfied. Previous literature already found positive effects of work engagement on job satisfaction, on the contrary to behaviour such as workaholism that could be deleterious for health and well-being (Shimazu et al., 2015). This suggests that the individual's state of mind towards their work would be a key factor that determine their feeling of health state. Johnston et al. (2016) also stress the importance of personality traits in assessing the effect of such behaviour.

Finally, CP does not appear to be associated with a decline in overall mental well-being. F. Green et al. (2016) highlighted in their study that high levels of effort and investment were linked to a decrease in job-related well-being during the 2006–2012 recession. While this may hold for job-related well-being, our findings indicate that, once individual heterogeneity is taken into account, there is no significant impact on general mental well-being. However, our pooled estimates do replicate the decline in well-being reported by F. Green et al. (2016), underlining the importance of accounting for individual heterogeneity in the analysis. Moreover, although the absence of significant effects may seem to be in contradiction with previous research on overcommitment and well-being (Hinsch et al., 2019), it can nevertheless be consistent within the framework of our study, where CP is regarded as a voluntary behaviour. Individuals who choose to engage in CP as a means of advancing their career may not experience a notable decline in mental well-being relative to the average population. In contrast, previous studies that report negative effects often

do so within frameworks of work intensification that are not voluntary but rather imposed on employees.

A large part of the variability of the CP behaviour would be explained by unobservable personal traits that enter the error term of the regressions. Therefore, it appears to be essential to employ methods that capture this endogeneity. While the instrumental variable approach is the most common method used in empirical econometrics, there is a constant debate about the use of appropriate instruments and their weakness (Keane & Neal, 2023). For instance, IVs are often associated with a significant decrease in statistical power (Semadeni et al., 2014). More complicated is the availability of adequate instruments in the original data. Therefore, this has led us to construct a leave-one-out instrument based on a sectorial peer effect justification. However, these instruments are widely criticised and their conditions of use are limited to a particular framework in which their validity is supposed to be fulfilled (Betz et al., 2018). Further, our IV results lead to the same conclusion as Semadeni et al. (2014), which raises the question of the weakness of the instrument, especially in the case of the fixed-effects estimates. Therefore, the use of recent IV-free correction methods based on copula seems especially interesting in the case of exploring microeconometric data when no reliable instrument are available.

However, the results of both methods lead to the same conclusion: fixed-effects models do not seem to suffer from endogeneity bias. All the endogeneity tests in the models using IV conclude to the absence of endogeneity in the fixed effects specifications. Moreover, the additional control that capture endogeneity in the copula method also lead to conclude that there is no endogeneity in these specifications. As previously stated, these models already partially control for individual heterogeneity. Thus, these models seem sufficiently robust to be able to carry out our statistical inference. On the contrary, this does not seem to be the case for the pooled models in the copula as well as in the IV models, justifying the correction of endogeneity in cross-sectional studies.

Some of the limitations of this analysis need to be taken into account for future research in order to increase our understanding of the mechanisms that help to explain the various facets presenteeism. Although this document is the first to construct a CP indicator, it was created on the basis of information available in the

CT-RPS data. It is certain that its explanatory power and its relevance could be increased by adding information such as: "not taking all your time off", "your family and friends say you are unavailable to them", "you prioritise work over your family and/or social commitments", "you regularly make calls outside working hours", etc. Future research should therefore agree on a common measure of CP behaviour unrelated to illness. What variables should be included in the indicator? What calculations should be used for this indicator? And so on.

For instance, one of the choices made in this chapter was not to include the behaviour of sickness presenteeism. This was justified by the fact that sickness presenteeism is a behaviour in its own right and deserves to be studied as such. Furthermore, under the assumption that CP can serve as a signal to the employee, engaging in sick presenteeism behaviour may also appear as a signal. For instance, Lu & Cooper (2022) show that working long hours and higher occupational grades are associated to higher sickness presenteeism. However, only individuals who have declared an illness would be able to claim this signal. This adds to the complexity of identification due to the endogenous selection of sick and non-sick workers. On the other hand, future research should focus more on the link between sickness presenteeism and CP associated with a signalling objective. Can sickness presenteeism, in the same vein as CP, serve as a valuable signal for workers? Are the effects on health the same ?

As regards the outcome variables, certain limits are imposed by the data available. If the wages and health or well-being variables correctly reflect the objective pursued, the use of occupations as a proxy for career improvement can be discussed. The variability of the variable that establishes the hierarchy between occupations is probably insufficient. Indeed, the categories of this variable may not correctly reflect all promotions within the individual's profession. Future research should identify more precise career evolution of individuals within their company and their occupation.

An important aspect of this chapter is the exploration of causality by correcting for endogeneity using an IV approach and a copula IV-free one. Unfortunately, our data do not directly provide a sufficiently robust instrument, which led us to construct a statistical leave-one-out instrument. Future research should focus on

building databases that are compatible with statistical methods for dealing with endogeneity. These databases should provide information on variables that are known to be used as IV. For example, the past characteristics and lifestyle habits of individuals. In the case of the present study, an interesting IV would have been physical activity and sport, which could explain individuals' personality traits. It could constitute a good instrument concerning the worker's investment in a behaviour such as CP without directly explaining his job characteristics such as wage or hierarchical position. Sporting activity could be a good indicator of a competitive state of mind and of surpassing oneself to achieve a set objective. In addition, recent methods based on copulas, which do not require additional information, make it possible to deal with endogeneity without the use of IV. Future research should also take a greater interest in these methods, as proposed in this chapter. However, we have to notice that the 2sCOPE method proposed by Yang et al. (2024) is suitable for linear model but not to non-linear ones. Thus, the results on occupational grades and health states should be treated with caution using this method. Indeed, the outcome variables are categorical with a small amount of modalities. However, the results provided give us important information when linked to the results of the fixed-effects models and the IV models.

Finally, although our results highlight the possibility of a strong response in terms of an individual's career from his employer when he is involved in CP, the incentive mechanisms of firms should be explored in more detail. In fact, the wage increase and the higher hierarchical position associated with individuals who are the most committed to their job could better reflect an equilibrium situation in our data. This would not allow us to identify whether, in the context of a tournament, the firm's optimal choice was indeed the most invested individual. In addition, we cannot rule out the possibility that it may not be the signal that is rewarded but the productivity of the workers. In a literature review by Kim et al. (2013), the authors show that 11 studies find a direct or indirect link between job commitment and performance. Furthermore, it is not possible for us to identify the visibility of the productivity involved in CP behaviour. As a result, future studies should address the question of the level of productivity of workers engaged in CP behaviour. This question is also at the heart of the very definition of the behaviour being studied.

Bibliography

Ahrens, H., & Pincus, R. (1981). On two measures of unbalancedness in a one-way model and their relation to efficiency. *Biometrical Journal*, *23*(3), 227-235.

Allen, S. (1981). An empirical model of work attendance. *The Review of Economics and Statistics*, *63*(1), 77-87.

Amegashie, J. A., Cadsby, C. B., & Song, Y. (2007). Competitive burnout: Theory and experimental evidence. *Games and Economic Behavior*, *59*(2), 213-239.

Andelic, N., Allan, J., Bender, K., Powell, D., & I.Theodossiou. (2024). Performance-related pay, mental and physiological health. *Industrial Relations: A Journal of Economy and Society*, *63*(1), 3–25.

Anger, S. (2008). Overtime work as a signaling device. *Scottish Journal of Political Economy*, *55*(2), 167–189.

Ashby, K., & Mahdon, M. (2010). Why do employees come to work when ill: An investigation into sickness presence in the workplace. *London: The Work Foundation*.

Avgoustaki, A., & Cañibano, A. (2020). Motivational drivers of extensive work effort: Are long hours always detrimental to well-being? *Industrial Relations: A Journal of Economy and Society*, *59*(3), 355-398.

Avgoustaki, A., & Frankort, H. T. W. (2019). Implications of work effort and discretion for employee well-being and career-related outcomes: An integrative assessment. *ILR Review*, *72*(3), 636-661.

Bakker, A. B., Demerouti, E., & Sanz-Vergel, A. (2023). Job demands–resources theory: Ten years later [Journal Article]. *Annual Review of Organizational Psychology and Organizational Behavior*, *10*(Volume 10, 2023), 25-53.

Bakker, A. B., Demerouti, E., & Sanz-Vergel, A. I. (2014). Burnout and work engagement: The jd–r approach. *The Annual Review of Organizational Psychology and Organizational Behavior*, *1*, 389–411.

Baktash, M. B., Heywood, J. S., & Jirjahn, U. (2022). Worker stress and performance pay: German survey evidence. *Journal of Economic Behavior & Organization*, *201*, 276-291.

Barnay, T. (2016). Health, work and working conditions: a review of the european economic literature. *The European Journal of Health Economics*, *17*, 693–709.

Barnay, T. (2020). Interactions of work and health: An economic perspective. In T. Theorell (Ed.), *Handbook of socioeconomic determinants of occupational health: From macro-level to micro-level evidence* (pp. 37–52). Cham : Springer International Publishing.

Baum, C. F., Schaffer, M. E., & Stillman, S. (2003). Instrumental variables and gmm: Estimation and testing. *The Stata Journal*, *3*(1), 1-31.

Bell, D., Hart, R., Hübler, O., & Schwerdt, W. (2000). Paid and unpaid overtime work in germany and the uk. *SSRN Electronic Journal*.

Berlinski, S., & Ramos, A. (2020). Peer effects in the decision to apply for a professional excellence award. *Labour Economics*, *67*, 101934.

Besley, T., & Ghatak, M. (2005). Competition and incentives with motivated agents. *American Economic Review*, *95*(3), 616-636.

Betz, T., Cook, S. J., & Hollenbach, F. M. (2018). On the use and abuse of spatial instruments. *Political Analysis*, *26*(4), 474–479.

Bilanakos, C., Heywood, J. S., Sessions, J. G., & Theodoropoulos, N. (2018). Does delegation increase worker training? *Economic Inquiry*, *56*(2), 1089-1115.

Booth, A. L., Francesconi, M., & Frank, J. (2003). A sticky floors model of promotion, pay, and gender. *European Economic Review*, *47*(2), 295-322.

Bramoullé, Y., Djebbari, H., & Fortin, B. (2020). Peer effects in networks: A survey. *Annual Review of Economics*, *12*(Volume 12, 2020), 603-629.

Bratti, M., & Staffolani, S. (2007). Effort-based career opportunities and working time. *International Journal of Manpower*, *28*, 489-512.

Breitung, J., Mayer, A., & Wied, D. (2024). Asymptotic properties of endogeneity corrections using nonlinear transformations. *The Econometrics Journal*, *27*(3), 362-383.

Budidharma, I., & Wahyuni, H. (2017). Does better physical function make people happier? *Asian Social Science*, *13*(3), 58–69.

Bénabou, R., & Tirole, J. (2003). Intrinsic and extrinsic motivation. *Review of Economic Studies*, *70*, 489–520.

Böckerman, P. (2018). Presenteeism at work. In C. L. Cooper & L. Lu (Eds.), (pp. 166–180). Cambridge University Press.

Böckerman, P., Bryson, A., Kauhanen, A., & Kangasniemi, M. (2020). Does job design make workers happy? *Scottish Journal of Political Economy*, *67*(1), 31-52.

Campbell, D., & Green, F. (2002). The long term pay-off from working longer hours. *Studies in Economics 0205, School of Economics, University of Kent*.

Cappelli, P., & Neumark, D. (2001). Do "high-performance" work practices improve establishment-level outcomes? *Industrial and Labor Relations Review*, *54*(4), 737–775.

Card, D. (1999). *The causal effect of education on earnings* (1re éd., Vols. 3, Part A; O. Ashenfelter & D. Card, Eds.). Elsevier.

Caverley, N., Cunningham, J. B., & MacGregor, J. N. (2007). Sickness presenteeism, sickness absenteeism, and health following restructuring in a public service organization. *Journal of Management Studies*, *44*(2), 304–319.

Cooper, C. L. (1996). Hot under the collar. *Times Higher Education Supplement*, 15.

Cornelissen, T., Dustmann, C., & Schönberg, U. (2017). Peer effects in the workplace. *The American Economic Review*, *107*(2), 425–456.

Demerouti, E., Le Blanc, P. M., Bakker, A. B., Schaufeli, W. B., & Hox, J. (2009). Present but sick: a three-wave study on job demands, presenteeism and burnout. *Career Development International*, *14*(1), 50–68.

Dufour, J.-M. (2003). Identification, weak instruments, and statistical inference in econometrics. *Canadian Journal of Economics/Revue canadienne d'économique*, *36*(4), 767-808.

Eckert, C., & Hohberger, J. (2023). Addressing endogeneity without instrumental variables: An evaluation of the gaussian copula approach for management research. *Journal of Management*, *49*(4), 1460-1495.

Eisenberg, D., Golberstein, E., & Whitlock, J. L. (2014). Peer effects on risky behaviors: New evidence from college roommate assignments. *Journal of Health Economics*, *33*, 126-138.

Falk, A., & Ichino, A. (2006). Clean evidence on peer effects. *Journal of Labor Economics*, *24*(1), 39–57.

Francesconi, M. (2001). Determinants and consequences of promotions in britain. *Oxford Bulletin of Economics and Statistics*, *63*(3), 279-310.

Ganster, D., Rosen, C., & Fisher, G. (2018). Long working hours and well-being: What we know, what we do not know, and what we need to know. *Journal of Business and Psychology*, *33*, 25–39.

Garrow, V. (2016). Presenteeism: A review of current thinking. *Institute for Employment Studies, Report 507*.

Graham, J. R., Grennan, J., Harvey, C. R., & Rajgopal, S. (2022). Corporate culture: Evidence from the field. *Journal of Financial Economics*, *146*(2), 552-593.

Green, C. P., Heywood, J. S., & Theodoropoulos, N. (2021). Hierarchy and the employer size effect on wages: Evidence from britain. *Economica*, *88*(351), 671-696.

Green, F., Felstead, A., Gallie, D., & Henseke, G. (2021). Working still harder. *ILR Review*, *75*(2), 458–487.

Green, F., Felstead, A., Gallie, D., & Inanc, H. (2016). Job-related well-being through the great recession. *Journal of Happiness Studies*, *17*, 389–411.

Gregg, P., Grout, P., Ratcliffe, A., Smith, S., & Windmeijer, F. (2011). How important is pro-social behaviour in the delivery of public services? *Journal of Public Economics*, *95*(7), 758–766.

Haan, P., & Myck, M. (2009). Dynamics of health and labor market risks. *Journal of Health Economics*, *28*(6), 1116-1125.

Hansen, C. D., & Andersen, J. H. (2008). Going ill to work – what personal circumstances, attitudes and work-related factors are associated with sickness presenteeism? *Social Science & Medicine*, *67*, 956–964.

Haschka, R. E. (2022). Handling endogenous regressors using copulas: A generalization to linear panel models with fixed effects and correlated regressors. *Journal of Marketing Research*, *59*(4), 860-881.

Havet, N., & Plantier, M. (2023). The links between difficult working conditions and sickness absences in the case of french workers. *Labour*, *37*, 160–195.

Hinsch, D. M., Spanier, K., Radoschewski, F. M., & Bethge, M. (2019). Associations between overcommitment, effort–reward imbalance and mental health: findings from a longitudinal study. *Int Arch Occup Environ Health*, *92*, 559–567.

Hofaidhllaoui, M., & Chhinzer, N. (2014). The relationship between satisfaction and turnover intentions for knowledge workers. *Engineering Management Journal*, *26*(2), 3–9.

Hoffmann, F., Inderst, R., & Opp, M. (2020). Only Time Will Tell: A Theory of Deferred Compensation. *The Review of Economic Studies*, *88*(3), 1253-1278.

Holmstrom, B. (1982). Moral hazard in teams. *The Bell Journal of Economics*, *13*(2), 324–340.

Ichino, A., & Riphahn, R. T. (2005). The effect of employment protection on worker effort: Absenteeism during and after probation. *Journal of the European Economic Association*, *3*(1), 120–143.

Johansson, G., & Lundberg, I. (2004). Adjustment latitude and attendance requirements as determinants of sickness absence or attendance. empirical tests of the illness flexibility model. *Social Science & Medicine*, *58*, 1857–1868.

Johns, G. (2010). Presenteeism in the workplace: A review and research agenda. *Journal of Organizational Behavior*, *31*, 519–542.

Johnston, C. S., Maggiori, C., & Rossier, J. (2016). Professional trajectories, individual characteristics, and staying satisfied and healthy. *Journal of Career Development*, *43*(1), 81-98.

Keane, M., & Neal, T. (2023). Instrument strength in iv estimation and inference: A guide to theory and practice. *Journal of Econometrics*, *235*(2), 1625-1653.

Kim, W., Kolb, J., & Kim, T. (2013). The relationship between work engagement and performance a review of empirical literature and a proposed research agenda. *Human Resource Development Review*, *12*, 248-276.

Kiviet, J. F. (2020). Testing the impossible: Identifying exclusion restrictions. *Journal of Econometrics*, *218*(2), 294-316.

Kossek, E. E., & Lee, K.-H. (2017). Work-family conflict and work-life conflict. *Oxford University Press*.

Kreps, D. M., & Wilson, R. (1982). Reputation and imperfect information. *Journal of Economic Theory*, *27*(2), 253–279.

Lai, T., & Ng, T. (2014). The impact of product market competition on training provision: Evidence from canada. *The Canadian Journal of Economics / Revue canadienne d'Economique*, *47*(3), 856–888.

Lindquist, M., Sauermann, J., & Zenou, Y. (2022). Peer effects in the workplace: A network approach. *SSRN Electronic Journal*.

Lu, L., & Cooper, C. L. (2022). Sickness presenteeism as a link between long working hours and employees' outcomes: Intrinsic and extrinsic motivators as resources. *International Journal of Environmental Research and Public Health*, *19*(4).

Ma, J., Meltzer, D. P., Yang, L.-Q., & Liu, C. (2018). Motivation and presenteeism: The whys and whats. In C. L. Cooper & L. Lu (Eds.), *Presenteeism at work* (p. 97–122). Cambridge University Press.

Meyer, A., & Wallette, M. (2005). Absence of absenteeism and overtime work – signaling factors for temporary workers? *Working Papers, Lund University, Department of Economics*(2005:15), 1–27.

Pannenberg, M. (2005). Long-term effects of unpaid overtime. *Scottish Journal of Political Economy*, *52*(2), 177-193.

Papagiannaki, E. (2014). Rising unpaid overtime: a critical approach to existing theories. *Int. J. Management Concepts and Philosophy*, *8*(1), 68–88.

Park, S., & Gupta, S. (2012). Handling endogenous regressors by joint estimation using copulas. *Marketing Science*, *31*(4), 567–586.

Plümper, T., & Troeger, V. E. (2019). Not so harmless after all: The fixed-effects model. *Political Analysis*, *27*(1), 21–45.

Preston, A. E. (1989). The nonprofit worker in a for-profit world. *Journal of Labor Economics*, *7*(4), 438–463.

Qian, Y., Koschmann, A., & Xie, H. (2024). *A practical guide to endogeneity correction using copulas* (Working Paper N° 32231). National Bureau of Economic Research.

Ryvkin, D. (2011). Fatigue in dynamic tournaments. *Journal of Economics & Management Strategy*, *20*(4), 1011–1041.

Santhanam, N., & Srinivas, S. (2019). Modeling the impact of employee engagement and happiness on burnout and turnover intention among blue-collar workers at a manufacturing company. *Benchmarking: An International Journal*, *ahead-of-print*.

Schaufeli, W. B., Taris, T. W., & Bakker, A. B. (2008). It takes two to tango: Workaholism is working excessively and working compulsively. *The long work hours culture: Causes, consequences and choices*, 203–226.

Semadeni, M., Withers, M. C., & Trevis Certo, S. (2014). The perils of endogeneity and instrumental variables in strategy research: Understanding through simulations. *Strategic Management Journal*, *35*(7), 1070-1079.

Shimazu, A., Schaufeli, W. B., Kamiyama, K., & Kawakami, N. (2015). Workaholism vs. work engagement: the two different predictors of future well-being and performance. *International journal of behavioral medicine*, *22*(1), 18-23.

Shimazu, A., Schaufeli, W. B., Kubota, K., & Kawakami, N. (2012). Do workaholism and work engagement predict employee well-being and performance in opposite directions? *Industrial Health*, *50*(4), 316-321.

Simpson, R. (1998). Presenteeism, power and organizational change: Long hours as a career barrier and the impact on the working lives of women managers. *British Journal of Management*, *9*(s1), 37–50.

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, *87*(3), 355–374.

Ten Brummelhuis, L. L., Rothbard, N. P., & Uhrich, B. (2017). Beyond nine to five: Is working to excess bad for health? *Academy of Management Discoveries*, *3*(3), 262–283.

Terza, J. V., Basu, A., & Rathouz, P. J. (2008). Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modeling. *Journal of Health Economics*, *27*(3), 531-543.

van Hooff, M., Flaxman, P., Söderberg, M., Stride, C., & Geurts, S. (2018). Basic psychological need satisfaction, recovery state, and recovery timing. *Human Performance*, *31*, 125-143.

Wang, M., Lu, C.-q., & Lu, L. (2022). The positive potential of presenteeism: An exploration of how presenteeism leads to good performance evaluation. *Journal of Organizational Behavior*, *44*.

Wooldridge, J. M. (2010). Econometric analysis of cross section and panel data, second edition. The MIT Press.

Yang, F., Qian, Y., & Xie, H. (2024). Addressing endogeneity using a two-stage copula generated regressor approach. *Journal of Marketing Research*, *0*(0), 00222437241296453.

Younès, N., Rivière, M., Plancke, L., Leroyer, A., Blanchon, T., Azevedo Da Silva, M., & Melchior, M. (2018). Work intensity in men and work-related emotional demands in women are associated with increased suicidality among persons attending primary care. *Journal of Affective Disorders*, *235*, 565-573.

Zinko, R., & Rubin, M. (2015). Personal reputation and the organization. *Journal of Management & Organization*, *21*(2), 217–236.

Appendices

1 Descriptive statistics

1.1 Tables of descriptive statistics of explanatory variables

Table A1 – Descriptive statistics of the individual characteristics

		Competitive presenteeism categories				
Characteristic	Overall N = 10,345[1]	Low than mean N = 3,853[1]	Mean N = 3,842[1]	Mean to 90% N = 1,613[1]	Higher 10% N = 1,037[1]	p-value[2]
Age annee	43 (8.68)	44 (8.81)	43 (8.84)	43 (8.31)	43 (8.12)	0.4
female	4,379 (42.33%)	1,630 (42.30%)	1,706 (44.40%)	681 (42.22%)	362 (34.91%)	< 0.001
single	2,115 (20.44%)	837 (21.72%)	803 (20.90%)	291 (18.04%)	184 (17.74%)	0.026
child under18	6,081 (58.78%)	2,192 (56.89%)	2,257 (58.75%)	997 (61.81%)	635 (61.23%)	0.093
education						< 0.001
Nodiploma	659 (6.37%)	364 (9.45%)	216 (5.62%)	57 (3.53%)	22 (2.12%)	
lessBac	3,559 (34.40%)	1,756 (45.57%)	1,291 (33.60%)	351 (21.76%)	161 (15.53%)	
Bac	2,077 (20.08%)	837 (21.72%)	789 (20.54%)	307 (19.03%)	144 (13.89%)	
SupBac	4,050 (39.15%)	896 (23.25%)	1,546 (40.24%)	898 (55.67%)	710 (68.47%)	
ActivMenage						< 0.001
1	2,993 (28.93%)	1,130 (29.33%)	1,164 (30.30%)	419 (25.98%)	280 (27.00%)	
2	6,562 (63.43%)	2,356 (61.15%)	2,399 (62.44%)	1,100 (68.20%)	707 (68.18%)	
3	790 (7.64%)	367 (9.53%)	279 (7.26%)	94 (5.83%)	50 (4.82%)	
sectors						< 0.001
admi	1,254 (12.12%)	330 (8.56%)	471 (12.26%)	261 (16.18%)	192 (18.51%)	
agri	192 (1.86%)	65 (1.69%)	91 (2.37%)	25 (1.55%)	11 (1.06%)	
alim	570 (5.51%)	273 (7.09%)	193 (5.02%)	68 (4.22%)	36 (3.47%)	
autind	1,511 (14.61%)	741 (19.23%)	477 (12.42%)	170 (10.54%)	123 (11.86%)	
autser	826 (7.98%)	265 (6.88%)	304 (7.91%)	167 (10.35%)	90 (8.68%)	
comm	1,847 (17.85%)	719 (18.66%)	702 (18.27%)	276 (17.11%)	150 (14.46%)	
cons	812 (7.85%)	301 (7.81%)	324 (8.43%)	105 (6.51%)	82 (7.91%)	
elec	265 (2.56%)	109 (2.83%)	93 (2.42%)	37 (2.29%)	26 (2.51%)	
extr	375 (3.62%)	158 (4.10%)	122 (3.18%)	61 (3.78%)	34 (3.28%)	
finan	567 (5.48%)	120 (3.11%)	267 (6.95%)	106 (6.57%)	74 (7.14%)	
hebe	326 (3.15%)	120 (3.11%)	127 (3.31%)	53 (3.29%)	26 (2.51%)	
info	485 (4.69%)	94 (2.44%)	183 (4.76%)	101 (6.26%)	107 (10.32%)	
mattr	426 (4.12%)	187 (4.85%)	150 (3.90%)	47 (2.91%)	42 (4.05%)	
trans	889 (8.59%)	371 (9.63%)	338 (8.80%)	136 (8.43%)	44 (4.24%)	
workhours	38 (7.05)	35 (5.04)	37 (5.49)	41 (7.22)	47 (8.42)	< 0.001
Public	823 (7.96%)	256 (6.64%)	288 (7.50%)	157 (9.73%)	122 (11.76%)	< 0.001
partime	1,257 (12.15%)	586 (15.21%)	506 (13.17%)	129 (8.00%)	36 (3.47%)	< 0.001
supervisor	1,598 (15.45%)	279 (7.24%)	546 (14.21%)	369 (22.88%)	404 (38.96%)	< 0.001
firmworkers						< 0.001
10-50	2,705 (26.15%)	975 (25.30%)	1,031 (26.83%)	439 (27.22%)	260 (25.07%)	
200-500	1,390 (13.44%)	584 (15.16%)	486 (12.65%)	179 (11.10%)	141 (13.60%)	
50-200	2,301 (22.24%)	930 (24.14%)	810 (21.08%)	344 (21.33%)	217 (20.93%)	
less 10	2,020 (19.53%)	690 (17.91%)	801 (20.85%)	340 (21.08%)	189 (18.23%)	
more 500	1,496 (14.46%)	501 (13.00%)	562 (14.63%)	231 (14.32%)	202 (19.48%)	
NoInfo	433 (4.19%)	173 (4.49%)	152 (3.96%)	80 (4.96%)	28 (2.70%)	
Seniority	14 (9.31)	14 (9.58)	14 (9.36)	13 (8.80)	13 (8.72)	0.015
atypsched	5,802 (56.09%)	2,047 (53.13%)	2,061 (53.64%)	943 (58.46%)	751 (72.42%)	< 0.001
quantarget	4,034 (38.99%)	1,186 (30.78%)	1,456 (37.90%)	752 (46.62%)	640 (61.72%)	< 0.001

[1] n (%); Mean (SD); [2] Pearson's Chi-squared test; Wilcoxon rank sum test

Table A2 – Descriptive statistics of the working conditions

		Competitive presenteeism categories				
Characteristic	Overall N = 10,345[1]	Low than mean N = 3,853[1]	Mean N = 3,842[1]	Mean to 90% N = 1,613[1]	Higher 10% N = 1,037[1]	p-value[2]
calloutwork	4,646 (44.91%)	1,185 (30.76%)	1,737 (45.21%)	950 (58.90%)	774 (74.64%)	< 0.001
privprolife						< 0.001
VeryDifficult	412 (3.98%)	95 (2.47%)	142 (3.70%)	94 (5.83%)	81 (7.81%)	
Difficult	1,375 (13.29%)	330 (8.56%)	465 (12.10%)	313 (19.40%)	267 (25.75%)	
Good	5,010 (48.43%)	1,861 (48.30%)	1,875 (48.80%)	783 (48.54%)	491 (47.35%)	
VeryGood	3,548 (34.30%)	1,567 (40.67%)	1,360 (35.40%)	423 (26.22%)	198 (19.09%)	
wrmachin	1,643 (15.88%)	887 (23.02%)	567 (14.76%)	129 (8.00%)	60 (5.79%)	< 0.001
wrdemand	7,392 (71.45%)	2,343 (60.81%)	2,846 (74.08%)	1,311 (81.28%)	892 (86.02%)	< 0.001
wrnorms	5,267 (50.91%)	2,013 (52.25%)	1,977 (51.46%)	790 (48.98%)	487 (46.96%)	0.007
wrinfo	4,272 (41.30%)	1,574 (40.85%)	1,561 (40.63%)	658 (40.79%)	479 (46.19%)	0.010
wrcontr	3,185 (30.79%)	1,195 (31.01%)	1,190 (30.97%)	499 (30.94%)	301 (29.03%)	0.6
wrcoll	3,252 (31.44%)	1,179 (30.60%)	1,255 (32.67%)	487 (30.19%)	331 (31.92%)	0.2
AskExcessWork						< 0.001
NotConcerned	510 (4.93%)	246 (6.38%)	193 (5.02%)	47 (2.91%)	24 (2.31%)	
No	1,209 (11.69%)	670 (17.39%)	409 (10.65%)	90 (5.58%)	40 (3.86%)	
Notreally	4,381 (42.35%)	1,877 (48.72%)	1,658 (43.15%)	595 (36.89%)	251 (24.20%)	
Yes	3,173 (30.67%)	874 (22.68%)	1,227 (31.94%)	644 (39.93%)	428 (41.27%)	
YesOften	1,072 (10.36%)	186 (4.83%)	355 (9.24%)	237 (14.69%)	294 (28.35%)	
wtime control						< 0.001
NoInfo	252 (2.44%)	72 (1.87%)	82 (2.13%)	67 (4.15%)	31 (2.99%)	
Nocontrol	5,421 (52.40%)	1,595 (41.40%)	2,018 (52.52%)	1,004 (62.24%)	804 (77.53%)	
colleagues	116 (1.12%)	56 (1.45%)	36 (0.94%)	17 (1.05%)	7 (0.68%)	
hierarchy	915 (8.84%)	407 (10.56%)	334 (8.69%)	118 (7.32%)	56 (5.40%)	
clock in out	3,641 (35.20%)	1,723 (44.72%)	1,372 (35.71%)	407 (25.23%)	139 (13.40%)	
deadlines						< 0.001
Nodeadlines	1,955 (18.90%)	929 (24.11%)	684 (17.80%)	246 (15.25%)	96 (9.26%)	
fixed	3,813 (36.86%)	1,501 (38.96%)	1,390 (36.18%)	545 (33.79%)	377 (36.35%)	
variable	4,577 (44.24%)	1,423 (36.93%)	1,768 (46.02%)	822 (50.96%)	564 (54.39%)	
lacktime	2,589 (25.03%)	466 (12.09%)	951 (24.75%)	630 (39.06%)	542 (52.27%)	< 0.001
instructions						< 0.001
NoInstruction	2,056 (19.87%)	610 (15.83%)	728 (18.95%)	417 (25.85%)	301 (29.03%)	
StridflyApply	3,203 (30.96%)	1,537 (39.89%)	1,098 (28.58%)	375 (23.25%)	193 (18.61%)	
DoneOtherwise	5,086 (49.16%)	1,706 (44.28%)	2,016 (52.47%)	821 (50.90%)	543 (52.36%)	
helpsup	3,212 (31.05%)	1,132 (29.38%)	1,149 (29.91%)	557 (34.53%)	374 (36.07%)	< 0.001
helpcol	1,820 (17.59%)	591 (15.34%)	690 (17.96%)	311 (19.28%)	228 (21.99%)	< 0.001
tensions firm						< 0.001
FirmTensions	5,367 (51.88%)	1,526 (39.61%)	2,099 (54.63%)	1,045 (64.79%)	697 (67.21%)	
NoTension	4,978 (48.12%)	2,327 (60.39%)	1,743 (45.37%)	568 (35.21%)	340 (32.79%)	
underpressure	3,566 (34.47%)	565 (14.66%)	1,369 (35.63%)	879 (54.49%)	753 (72.61%)	< 0.001
unsure	2,395 (23.15%)	872 (22.63%)	860 (22.38%)	411 (25.48%)	252 (24.30%)	0.056
changes	4,736 (45.78%)	1,494 (38.77%)	1,788 (46.54%)	822 (50.96%)	632 (60.95%)	< 0.001
promo prospect						< 0.001
NotConcerned	535 (5.17%)	236 (6.13%)	189 (4.92%)	71 (4.40%)	39 (3.76%)	
NotSatisAtall	1,830 (17.69%)	588 (15.26%)	767 (19.96%)	310 (19.22%)	165 (15.91%)	
NotSatisfied	3,374 (32.61%)	1,193 (30.96%)	1,256 (32.69%)	543 (33.66%)	382 (36.84%)	
Satisfied	3,637 (35.16%)	1,461 (37.92%)	1,291 (33.60%)	539 (33.42%)	346 (33.37%)	
VerySatisfied	969 (9.37%)	375 (9.73%)	339 (8.82%)	150 (9.30%)	105 (10.13%)	

[1] n (%); Mean (SD); [2] Pearson's Chi-squared test; Wilcoxon rank sum test

1.2 Graphics of the MCA to create the competitive presenteeism index

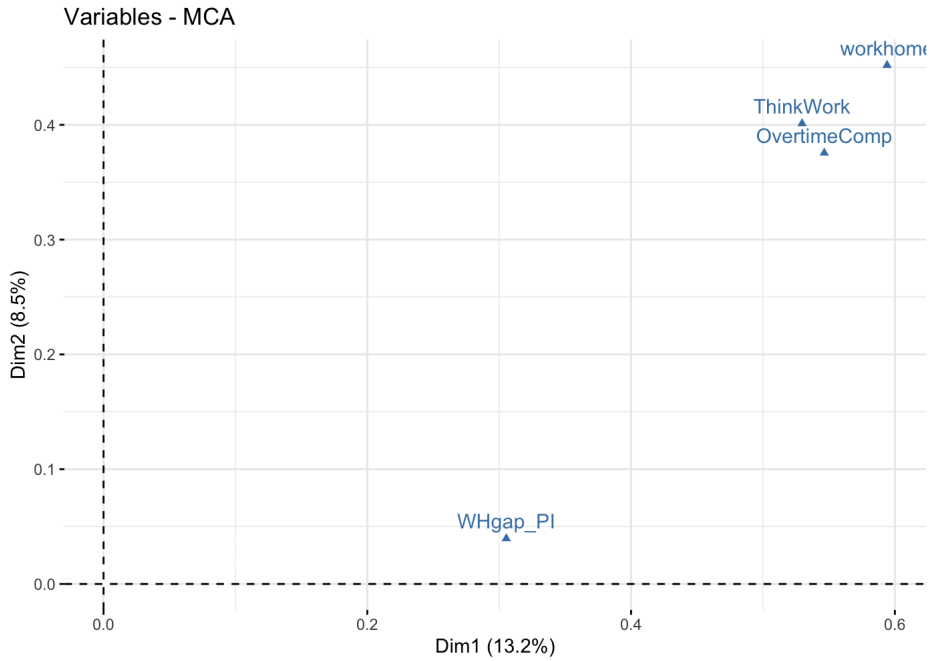

Figure A1 – Correlations of the variables with the first two dimensions

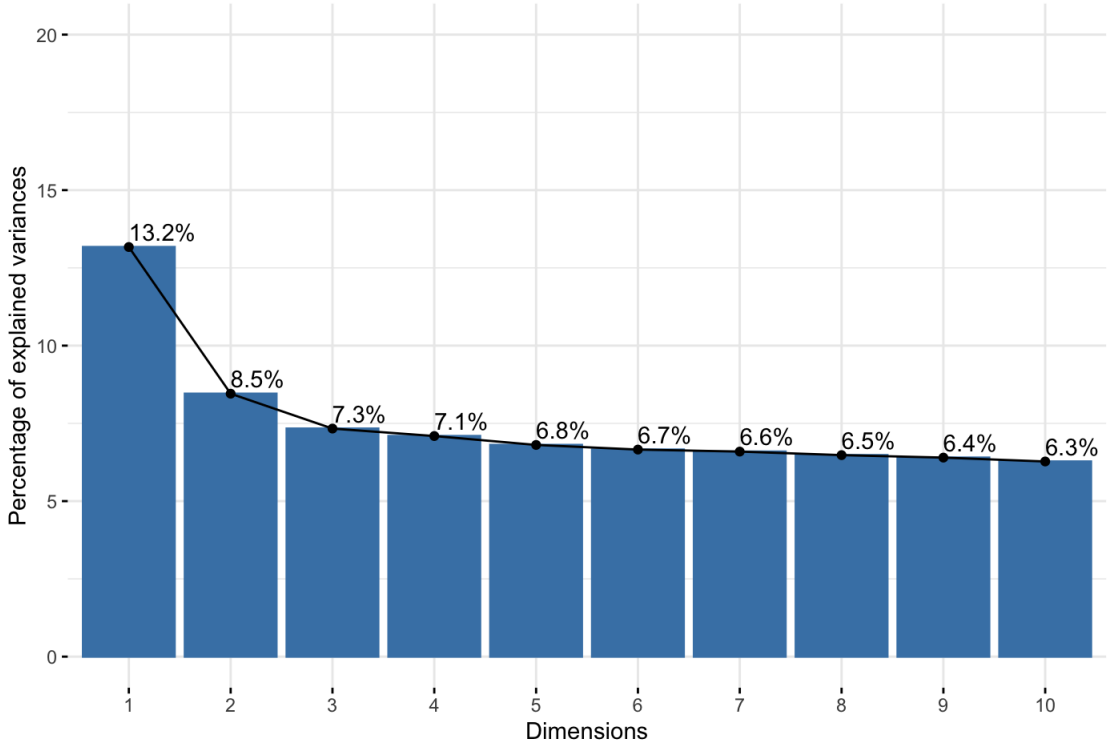

Figure A2 – Percentage of variance explained by dimension (eigenvalue)

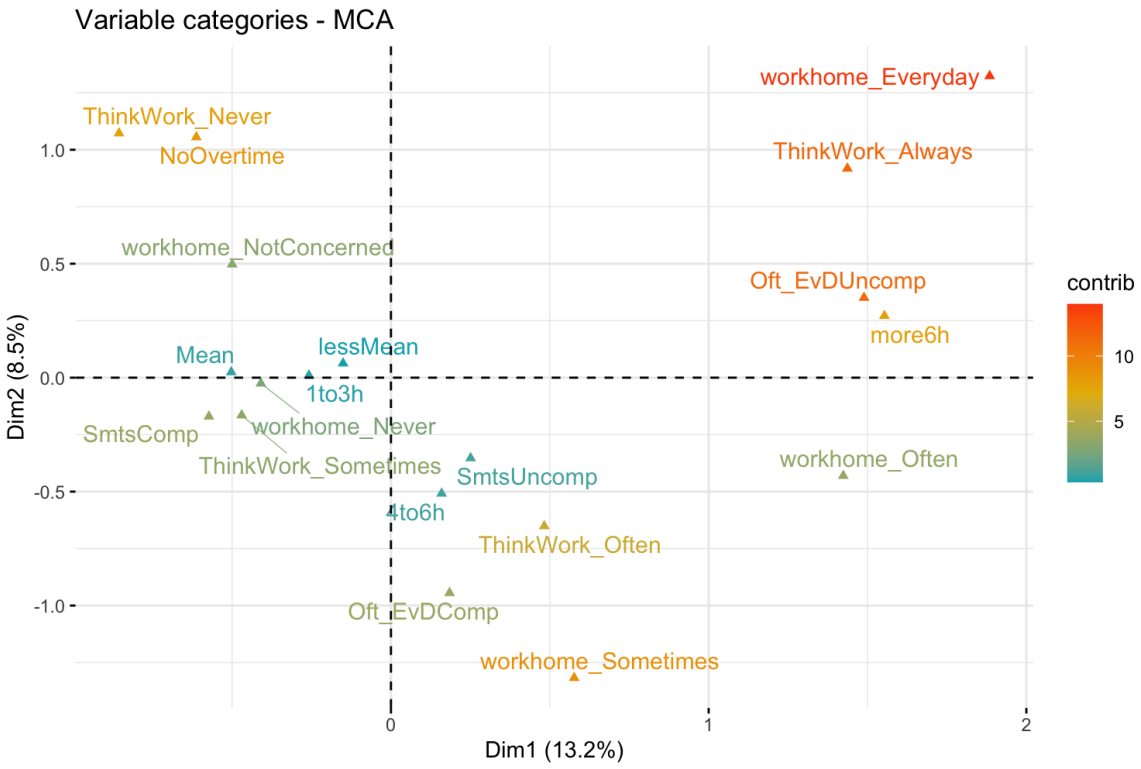

Figure A3 – Modalities position and contribution to the first two dimensions

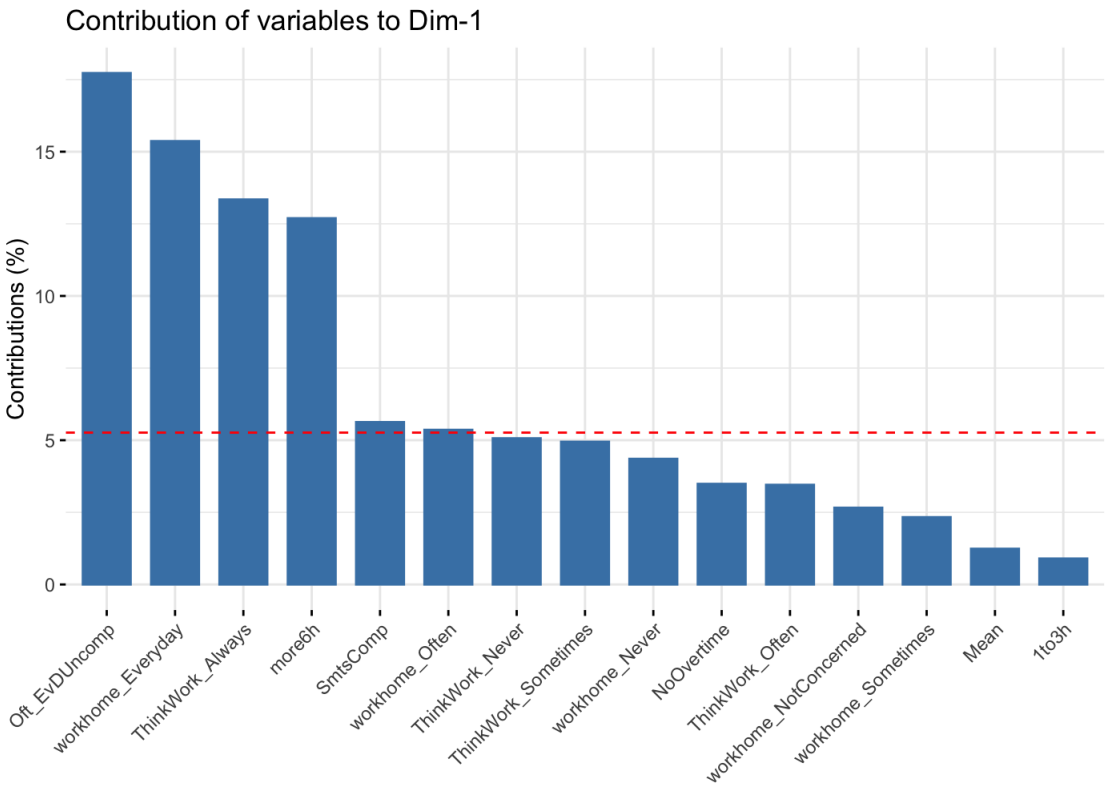

Figure A4 – Variables contribution to the first dimension

2 Complete tables of the main results

The models in this section correspond to the models presented in the main document. In addition, the model (1) in each table of this section reports the results without any controls and the last column of each tables reports the pooled regressions that are not shown in the main document. Further, the tables that report the 2sCOPE and the IV results present all the model with the different specifications.

2.1 Results of the pooled models

Table A3 – Pooled estimates of the effects of competitive presenteeism the outcomes

	Dependent variables:				
	$\log(\text{Hourly wage})_t$	Occupation_t	Health_t	$\log(\text{SA days})_t$	WHO5_t
	(OLS)	(Ordered logit)	(Ordered logit)	(OLS)	(OLS)
CP index_{t-1}	0.010*	**0.176*****	−0.087*	−0.011	**−0.294*****
	(0.005)	(0.046)	(0.047)	(0.030)	(0.099)
CP index^2_{t-1}	**−0.006****	0.043	**0.073*****	−0.004	0.062
	(0.003)	(0.033)	(0.027)	(0.016)	(0.053)
Log Lik.	503.35	−4,240	−5,131.38	−8,234.26	−14,199.47
N obs.	4,998	4,998	4,998	4,988	4,998
(pseudo) R^2	0.51	0.44	0.10	0.08	0.30
Adj. R^2	0.50			0.07	0.28
Individual's controls	Yes	Yes	Yes	Yes	Yes
Health controls	Yes	Yes	Yes	Yes	Yes
Employment controls	Yes	Yes	Yes	Yes	Yes
Working conditions	Yes	Yes	Yes	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; clustered standard errors at the individual level in parenthesis.

2.2 Results of the two stages Copula Endogeneity Correction models

Table A4 – 2sCOPE estimates of competitive presenteeism on career outcomes

	Fixed effects					Pooled
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	log(Hourly wage)$_t$					
CP$_{t-1}$	0.024*	0.025*	0.024	0.023	0.023	**0.120*****
	(0.015)	(0.015)	(0.015)	(0.014)	(0.015)	(0.017)
CP$^2_{t-1}$	−0.002	−0.002	−0.002	−0.002	−0.002	**−0.017*****
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Resid 1st stage	−0.002	−0.002	−0.001	−0.001	−0.001	**−0.091*****
	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)	(0.014)
ρ	−0.017	−0.015	−0.013	−0.012	−0.012	**−0.220*****
	(0.036)	(0.034)	(0.035)	(0.033)	(0.034)	(0.033)
σ_ε	0.101	0.101	0.101	0.101	0.100	0.246
Observations	4,998	4,998	4,998	4,998	4,998	4,998
R^2	0.159	0.162	0.163	0.167	0.168	0.992
Adjusted R^2	0.158	0.160	0.160	0.162	0.162	0.991
Dependent variable:	Occupation$_t$					
CP$_{t-1}$	−0.051	−0.053	−0.045	−0.044	−0.044	**0.221*****
	(0.052)	(0.049)	(0.053)	(0.053)	(0.054)	(0.056)
CP$^2_{t-1}$	0.018	0.019	0.018	0.020*	0.020*	−0.020*
	(0.011)	(0.012)	(0.012)	(0.012)	(0.012)	(0.011)
Resid 1st stage	0.004	0.003	0.001	0.000	0.000	0.001
	(0.013)	(0.013)	(0.013)	(0.014)	(0.014)	(0.047)
ρ	0.011	0.009	0.003	0.002	0.001	0.005
	(0.036)	(0.034)	(0.036)	(0.036)	(0.036)	(0.039)
σ_ε	0.370	0.369	0.369	0.368	0.368	0.796
Observations	4,998	4,998	4,998	4,998	4,998	4,998
R^2	0.022	0.025	0.029	0.033	0.034	0.948
Adjusted R^2	0.021	0.023	0.026	0.027	0.027	0.948
Individual's controls	No	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	Yes	Yes	Yes	Yes
Employment controls	No	No	No	Yes	Yes	Yes
Working conditions	No	No	No	No	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; clustered standard errors at the individual level in parenthesis.
Models controls are thus reported in the main tables

Table A5 – 2sCOPE estimates of competitive presenteeism on health outcomes

	Fixed effects					Pooled
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	Health$_t$					
CP$_{t-1}$	**−0.137****	**−0.132****	**−0.137****	**−0.148****	**−0.148****	0.011
	(0.063)	(0.061)	(0.059)	(0.063)	(0.062)	(0.049)
CP$^2_{t-1}$	**0.038****	**0.037****	**0.038****	**0.038****	**0.038****	0.018
	(0.016)	(0.015)	(0.016)	(0.017)	(0.015)	(0.012)
Resid 1st stage	0.026*	0.023	0.023	0.025*	0.026	−0.040
	(0.016)	(0.015)	(0.015)	(0.015)	(0.016)	(0.041)
ρ	0.063*	0.056*	0.057*	0.061*	0.063*	−0.038
	(0.036)	(0.034)	(0.033)	(0.034)	(0.036)	(0.037)
σ_ε	0.432	0.431	0.431	0.427	0.427	0.702
Observations	1,593	1,593	1,593	1,593	1,593	4,998
R^2	0.020	0.022	0.024	0.040	0.041	0.945
Adjusted R^2	0.019	0.020	0.021	0.034	0.034	0.945
Dependent variable:	SA days$_t$					
CP$_{t-1}$	−0.142	−0.172	−0.166	−0.145	−0.072	−0.034
	(0.272)	(0.264)	(0.251)	(0.251)	(0.247)	(0.087)
CP$^2_{t-1}$	**0.105***	**0.111****	**0.112****	**0.106****	**0.100****	0.004
	(0.054)	(0.050)	(0.049)	(0.048)	(0.050)	(0.020)
Resid 1st stage	0.015	0.018	0.019	0.016	-0.003	0.017
	(0.065)	(0.063)	(0.060)	(0.060)	(0.059)	(0.075)
ρ	−0.014	−0.011	−0.011	−0.013	−0.025	0.008
	(0.045)	(0.044)	(0.042)	(0.041)	(0.041)	(0.037)
σ_ε	1.368	1.365	1.363	1.360	1.359	1.273
Observations	1,593	1,593	1,593	1,593	1,593	4,998
R^2	0.173	0.179	0.182	0.186	0.191	0.315
Adjusted R^2	0.171	0.177	0.179	0.181	0.185	0.309
Dependent variable:	WHO5$_t$					
CP$_{t-1}$	−0.077	−0.058	−0.073	−0.171	−0.094	−0.026
	(0.377)	(0.386)	(0.381)	(0.372)	(0.374)	(0.290)
CP$^2_{t-1}$	−0.061	−0.070	−0.063	−0.070	−0.073	0.013
	(0.100)	(0.104)	(0.101)	(0.102)	(0.100)	(0.065)
Resid 1st stage	0.036	0.027	0.027	0.044	0.031	−0.259
	(0.095)	(0.096)	(0.099)	(0.094)	(0.095)	(0.239)
ρ	0.009	0.005	0.005	0.012	0.009	−0.042
	(0.036)	(0.036)	(0.037)	(0.036)	(0.037)	(0.036)
σ_ε	2.594	2.588	2.586	2.524	2.509	4.194
Observations	1,593	1,593	1,593	1,593	1,593	4,998
R^2	0.022	0.026	0.027	0.074	0.085	0.937
Adjusted R^2	0.020	0.024	0.024	0.068	0.079	0.936
Individual's controls	No	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	Yes	Yes	Yes	Yes
Employment controls	No	No	No	Yes	Yes	Yes
Working conditions	No	No	No	No	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; clustered standard errors at the individual level in parenthesis.
Models controls are thus reported in the main tables

2.3 Instrumental variable models

Table A6 – IV estimates of the effects of competitive presenteeism on hourly wages

Dependent variable:	log(Hourly wage)$_t$					
	Fixed effects 2SGMM					2SGMM
	(1)	(2)	(3)	(4)	(5)	(6)
CP index$_{t-1}$	**1.022*****	0.130	0.103	0.120	0.089	0.111
	(0.200)	(0.093)	(0.103)	(0.104)	(0.087)	(0.083)
CP index$^2_{t-1}$	**−0.292****	−0.004	0.005	−0.004	0.008	**0.209****
	(0.128)	(0.053)	(0.053)	(0.054)	(0.048)	(0.104)
Log Lik.	−893	3,118	3,222	3,189	3,306	−2,469
N obs.	3,186	3,186	3,186	3,186	3,186	4,358
N group	1,593	1,593	1,593	1,593	1,593	.
Uncentered R2	−12.04	−0.05	0.02	−0.00	0.07	0.97
KP rk LM statistic $\chi^2(5)$	1.75	3.84	3.46	3.18	3.17	26.30***
Hansen J statistic $\chi^2(4)$	1.56	3.76	4.26	4.00	3.85	5.78
AR Wald test $F(6)$	41.67	1.04	1.08	1.01	1.00	42.67***
Endogeneity test $\chi^2(2)$	0.03	1.01	0.85	0.81	1.13	140.75***
Individual's controls	No	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	Yes	Yes	Yes	Yes
Employment controls	No	No	No	Yes	Yes	Yes
Working conditions	No	No	No	No	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; clustered standard errors at the individual level. Models controls are the same as those reported in the main tables. KP rk LM statistic: Kleibergen-Paap rank LM test, AR: Anderson-Rubin.

Table A7 – IV estimates of the effects of competitive presenteeism on occupational grades

Dependent variable:	Occupation$_t$							
	Fixed effects 2SRI					2SRI	FE2SGMM	2SGMM
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
CP index$_{t-1}$	0.817	0.253	−0.283	0.411	2.185	**6.658*****	0.086	**2.628*****
	(0.611)	(0.651)	(0.817)	(0.992)	(2.312)	(0.263)	(0.524)	(0.227)
CP index$^2_{t-1}$	0.105	0.142	0.125	0.134	**0.276****	0.032	0.078	**−0.686*****
	(0.098)	(0.101)	(0.104)	(0.112)	(0.130)	(0.036)	(0.147)	(0.200)
resid 1st Stage	−0.994	−0.416	0.109	−0.649	−2.671	**−6.448*****		
	(0.632)	(0.674)	(0.821)	(0.982)	(2.323)	(0.266)		
Log Lik.	−331	−318	−314	−304	−268	−4,142	−967	−8,435
N obs.	972	972	972	972	972	4,358	3,186	4,358
N group	383	383	383	383	383	.	1,593	.
R2 (uncentered)	0.02	0.06	0.07	0.10	0.20	0.37	0.01	0.77
KP rk LM statistic $\chi^2(5)$							1.23	34.05***
Hansen J statistic $\chi^2(4)$							1.69	1.66
AR Wald test $F(6)$							1.40	173.82***
Endogeneity test $\chi^2(2)$							0.08	496.17***
Individual's controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Employment controls	No	No	No	Yes	Yes	Yes	Yes	Yes
Working conditions	No	No	No	No	Yes	Yes	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; clustered standard errors at the individual level in parenthesis.
Models controls are the same as those reported in the main tables.
KP rk LM statistic: Kleibergen-Paap rank LM test, AR: Anderson-Rubin

Table A8 – IV estimates of the effects of competitive presenteeism on self-rated health

Dependent variable:	Health$_t$							
	Fixed effects 2SRI					2SRI	FE2SGMM	2SGMM
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
CP index$_{t-1}$	**0.927****	**1.258*****	1.105*	0.975	2.031	0.227	−0.037	0.015
	(0.443)	(0.475)	(0.593)	(0.684)	(2.285)	(0.167)	(0.464)	(0.095)
CP index$^2_{t-1}$	0.108*	0.108	0.110*	0.109	0.102	0.063**	0.239	0.066
	(0.065)	(0.067)	(0.067)	(0.067)	(0.071)	(0.029)	(0.193)	(0.089)
resid 1st Stage	**−1.210*****	**−1.539*****	**−1.387****	−1.258*	−2.336	−0.321*		
	(0.460)	(0.492)	(0.608)	(0.698)	(2.294)	(0.173)		
Log Lik.	−509.01	−501.39	−501.29	−497.55	−473.70	−4,485.33	−1,597.23	−4,611.55
N obs.	1,514	1,514	1,514	1,514	1,514	4,358	3,186	4,358
N group	678	678	678	678	678	.	1,593	.
R2 (uncentered)	0.03	0.04	0.04	0.05	0.10	0.09	−0.10	0.95
KP rk LM statistic $\chi^2(5)$							1.52	24.88***
Hansen J statistic $\chi^2(4)$							1.65	6.07
AR Wald test $F(6)$							1.82	1.15
Endogeneity test $\chi^2(2)$							0.07	2.06
Individual's controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Employment controls	No	No	No	Yes	Yes	Yes	Yes	Yes
Working conditions	No	No	No	No	Yes	Yes	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; clustered standard errors at the individual level in parenthesis.
Models controls are the same as those reported in the main tables.
KP rk LM statistic: Kleibergen-Paap rank LM test, AR: Anderson-Rubin.

Table A9 – IV estimates of the effects of competitive presenteeism on SA days

Dependent variable:	$\log(\text{SA days})_t$					
	Fixed effects 2SLS					2SLS
	(1)	(2)	(3)	(4)	(5)	(6)
CP index_{t-1}	**1.225****	0.152	−0.096	−0.262	−0.082	**−0.411****
	(0.540)	(1.049)	(1.120)	(1.240)	(1.129)	(0.178)
CP index^2_{t-1}	−0.603*	−0.255	−0.142	−0.112	−0.085	0.246
	(0.360)	(0.445)	(0.448)	(0.497)	(0.429)	(0.170)
Log Lik.	−4,352.90	−4,052.97	−4,012.61	−4,021.44	−3,935.28	−7,301.31
N obs.	3,174	3,174	3,174	3,174	3,174	4,349
N group	1,587	1,587	1,587	1,587	1,587	.
Uncentered R2	−0.27	−0.05	−0.03	−0.03	0.02	0.29
KP rk LM statistic $\chi^2(5)$	1.74	1.60	1.57	1.54	1.52	24.72**
Hansen J statistic $\chi^2(4)$	0.67	1.41	1.48	1.52	1.50	8.27*
AR Wald test $F(6)$	0.73	0.71	0.81	0.87	0.74	2.37**
Endogeneity test $\chi^2(2)$	0.82	0.17	0.12	0.09	0.09	5.23*
Individual's controls	No	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	Yes	Yes	Yes	Yes
Employment controls	No	No	No	Yes	Yes	Yes
Working conditions	No	No	No	No	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; clustered standard errors at the individual level. Models controls are the same as those reported in the main tables. KP rk LM statistic: Kleibergen-Paap rank LM test, AR: Anderson-Rubin.

Table A10 – IV estimates of the effects of competitive presenteeism on mental well-being

Dependent variable:	WHO5_t					
	Fixed effects 2SLS					2SLS
	(1)	(2)	(3)	(4)	(5)	(6)
CP index_{t-1}	**−5.564*****	1.572	1.872	1.577	1.696	−0.779
	(1.932)	(2.247)	(2.073)	(2.071)	(2.275)	(0.577)
CP index^2_{t-1}	**2.825****	0.562	0.330	0.507	0.592	0.597
	(1.228)	(1.068)	(1.014)	(1.094)	(1.068)	(0.552)
Log Lik.	−8,091.98	−7,379.41	−7,346.97	−7,339.24	−7,289.62	−12,434.40
N obs.	3,186	3,186	3,186	3,186	3,186	4,358
N group	1,593	1,593	1,593	1,593	1,593	.
Uncentered R2	−0.75	−0.12	−0.10	−0.09	−0.06	0.94
KP rk LM statistic $\chi^2(5)$	1.75	1.59	1.59	1.55	1.56	24.78***
Hansen J statistic $\chi^2(4)$	1.69	1.53	1.50	1.54	1.51	1.74
AR Wald test $F(6)$	2.79	1.37	1.24	1.09	0.86	0.82
Endogeneity test $\chi^2(2)$	0.11	0.09	0.11	0.08	0.09	1.05
Individual's controls	No	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	Yes	Yes	Yes	Yes
Employment controls	No	No	No	Yes	Yes	Yes
Working conditions	No	No	No	No	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; clustered standard errors at the individual level. Models controls are the same as those reported in the main tables. KP rk LM statistic: Kleibergen-Paap rank LM test, AR: Anderson-Rubin.

3 Mediated effects of work pressure and intrinsic motivations

3.1 The mediated effects of work pressure

Table A11 – Estimates of the effects of competitive presenteeism on career moderated by work pressure

	Dependent variable:					
	Fixed effects					Pooled
	(1)	(2)	(3)	(4)	(5)	(6)
	log(Hourly wage)$_t$					
CP index$_{t-1}$	0.019**	0.019**	0.021**	0.020**	0.020**	0.003
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.42)
CP index^2	0.000	0.000	−0.001	−0.001	−0.000	−0.002
	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.34)
Under pressure	0.005	0.004	0.005	0.005	0.005	0.020**
	(0.010)	(0.010)	(0.010)	(0.010)	(0.011)	(2.11)
CP×Under pressure	−0.007	−0.007	−0.007	−0.008	−0.007	0.018*
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(1.70)
CP^2×Under pressure	−0.001	−0.001	−0.000	−0.000	−0.000	−0.010
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(1.54)
Log Lik.	6,531.26	6,545.49	6,554.73	6,566.90	6,622.46	506.75
N obs.	4,998	4,998	4,998	4,998	4,998	4,998
Rsq	0.14	0.15	0.15	0.16	0.18	0.51
R adj.	0.14	0.15	0.15	0.15	0.17	0.50
	Occupation$_t$					
CP index$_{t-1}$	−0.110	−0.160	−0.212	−0.281	−0.319	0.126**
	(0.205)	(0.212)	(0.215)	(0.224)	(0.235)	(2.23)
CP index^2	0.163	0.170	0.175	0.161	0.272*	0.029
	(0.138)	(0.149)	(0.153)	(0.152)	(0.150)	(0.59)
Under pressure	−0.315	−0.176	−0.114	−0.022	−0.052	0.061
	(0.256)	(0.258)	(0.266)	(0.265)	(0.296)	(0.68)
CP×Under pressure	−0.059	−0.084	−0.044	0.073	0.006	0.137
	(0.291)	(0.288)	(0.293)	(0.312)	(0.331)	(1.40)
CP^2×Under pressure	0.056	0.064	0.053	0.058	0.016	−0.011
	(0.177)	(0.182)	(0.186)	(0.190)	(0.191)	(0.15)
Log Lik.	−782.49	−758.13	−751.27	−730.94	−663.75	−4,238.68
N obs.	11,970	11,970	11,970	11,970	11,970	4,998
R p	0.81	0.82	0.82	0.82	0.84	0.44
Individual's controls	No	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	Yes	Yes	Yes	Yes
Employment controls	No	No	No	Yes	Yes	Yes
Working conditions	No	No	No	No	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; robust standard errors in parenthesis
Models controls are thus reported in the main tables

Table A12 – Estimates of the effects of competitive presenteeism on health moderated by work pressure

	Dependent variable:					
	Fixed effects					Pooled
	(1)	(2)	(3)	(4)	(5)	(6)
	Health_t					
CP index_{t-1}	−0.379**	−0.361**	−0.380**	−0.396**	−0.480***	−0.117**
	(0.156)	(0.157)	(0.156)	(0.156)	(0.161)	(0.057)
CP index^2_{t-1}	0.120	0.119	0.123	0.117	0.120	0.105**
	(0.088)	(0.088)	(0.088)	(0.087)	(0.092)	(0.045)
Under pressure	0.073	0.070	0.018	0.008	−0.162	0.074
	(0.202)	(0.206)	(0.206)	(0.209)	(0.229)	(0.085)
CP×Under pressure	0.468**	0.444**	0.435**	0.432**	0.475**	0.092
	(0.215)	(0.217)	(0.215)	(0.219)	(0.229)	(0.093)
CP^2×Under pressure	−0.045	−0.037	−0.038	−0.034	−0.030	−0.066
	(0.119)	(0.120)	(0.118)	(0.119)	(0.126)	(0.062)
Log Lik.	−1,437.00	−1,425.41	−1,419.71	−1,408.17	−1,329.09	−5,130.45
N obs.	14,570	14,570	14,570	14,570	14,570	4,998
R p	0.72	0.72	0.72	0.72	0.74	0.10
	$\log(\text{SA days})_t$					
CP index_{t-1}	0.082	0.085	0.077	0.076	0.071	0.020
	(0.081)	(0.081)	(0.081)	(0.081)	(0.081)	(0.56)
CP index^2	0.069	0.067	0.067	0.063	0.060	−0.020
	(0.042)	(0.043)	(0.043)	(0.042)	(0.044)	(0.81)
Under pressure	−0.020	−0.026	−0.007	−0.012	0.059	−0.040
	(0.097)	(0.097)	(0.096)	(0.096)	(0.106)	(0.75)
CP×Under pressure	−0.162	−0.165	−0.141	−0.154	−0.157	−0.098*
	(0.104)	(0.104)	(0.103)	(0.103)	(0.105)	(1.68)
CP^2×Under pressure	0.047	0.049	0.048	0.053	0.057	0.044
	(0.056)	(0.056)	(0.056)	(0.056)	(0.057)	(1.29)
Log Lik.	−5,080.54	−5,067.29	−5,043.69	−5,029.56	−4,972.56	−8,232.72
N obs.	4,988	4,988	4,988	4,988	4,988	4,988
Rsq	0.01	0.02	0.03	0.03	0.05	0.08
R adj.	0.01	0.01	0.02	0.03	0.04	0.07
	WHO5_t					
CP index_{t-1}	0.087	0.106	0.126	0.110	0.097	−0.272**
	(0.207)	(0.207)	(0.206)	(0.206)	(0.198)	(2.30)
CP index^2	−0.148	−0.157	−0.150	−0.137	−0.172	0.004
	(0.156)	(0.155)	(0.154)	(0.154)	(0.154)	(0.04)
Under pressure	0.299	0.272	0.168	0.162	−0.437	−0.574***
	(0.271)	(0.272)	(0.268)	(0.268)	(0.292)	(3.23)
CP×Under pressure	−0.010	−0.031	−0.128	−0.126	−0.130	0.184
	(0.302)	(0.303)	(0.300)	(0.299)	(0.291)	(0.93)
CP^2×Under pressure	0.089	0.096	0.097	0.096	0.114	0.024
	(0.203)	(0.202)	(0.199)	(0.200)	(0.193)	(0.20)
Log Lik.	−10,127.90	−10,120.28	−10,084.33	−10,071.53	−9,877.93	−14,191.91
N obs.	4,998	4,998	4,998	4,998	4,998	4,998
Rsq	0.02	0.02	0.03	0.04	0.11	0.30
R adj.	0.01	0.01	0.03	0.03	0.10	0.28
Individual's controls	No	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	Yes	Yes	Yes	Yes
Employment controls	No	No	No	Yes	Yes	Yes
Working conditions	No	No	No	No	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; robust standard errors in parenthesis
Models controls are thus reported in the main tables

Table A13 – Estimates of the effects of competitive presenteeism on career moderated by lack time

	Dependent variable:					
	Fixed effects					Pooled
	(1)	(2)	(3)	(4)	(5)	(6)
	log(Hourly wage)$_t$					
CP index$_{t-1}$	0.019***	0.019***	0.020***	0.020***	0.021***	0.007
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(1.05)
CP index$^2_{t-1}$	−0.005	−0.005	−0.005	−0.006	−0.006	−0.005
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(1.06)
Lack time	−0.002	−0.003	−0.003	−0.002	−0.012	−0.002
	(0.010)	(0.010)	(0.010)	(0.011)	(0.012)	(0.25)
CP × Lack time	−0.012	−0.013	−0.012	−0.013	−0.016	0.010
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.88)
CP^2 × Lack time	0.008	0.009	0.009	0.010	0.011	−0.005
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.66)
Log Lik.	6,532.63	6,547.84	6,556.70	6,569.44	6,618.13	492.58
N obs.	4,998	4,998	4,998	4,998	4,998	4,998
Rsq	0.15	0.15	0.15	0.16	0.17	0.51
R adj.	0.14	0.15	0.15	0.15	0.16	0.50
	Occupation$_t$					
CP index$_{t-1}$	−0.093	−0.141	−0.168	−0.184	−0.223	0.168***
	(0.201)	(0.199)	(0.201)	(0.209)	(0.220)	(3.16)
CP index^2	0.072	0.073	0.078	0.104	0.206*	0.036
	(0.126)	(0.120)	(0.123)	(0.124)	(0.121)	(0.78)
Lack time	−0.510*	−0.511**	−0.536**	−0.452*	−0.585**	0.066
	(0.263)	(0.260)	(0.263)	(0.265)	(0.296)	(0.71)
CP × Lack time	−0.228	−0.197	−0.252	−0.238	−0.333	0.011
	(0.306)	(0.290)	(0.296)	(0.298)	(0.312)	(0.10)
CP^2 × Lack time	0.242	0.258	0.264	0.245	0.206	0.011
	(0.164)	(0.157)	(0.161)	(0.162)	(0.172)	(0.16)
Log Lik.	−779.42	−753.49	−746.38	−727.62	−673.35	−4,240.23
N obs.	2,572	2,572	2,572	2,572	2,572	4,998
R p	0.81	0.82	0.82	0.82	0.84	0.44
Individual's controls	No	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	Yes	Yes	Yes	Yes
Employment controls	No	No	No	Yes	Yes	Yes
Working conditions	No	No	No	No	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; robust standard errors in parenthesis
Models controls are thus reported in the main tables

Table A14 – Estimates of the effects of competitive presenteeism on health moderated by lack time

	Dependent variable:					
	Fixed effects					Pooled
	(1)	(2)	(3)	(4)	(5)	(6)
	Health_t					
CP index_{t-1}	−0.258*	−0.246*	−0.268*	−0.287*	−0.334**	−0.107**
	(0.145)	(0.147)	(0.146)	(0.147)	(0.153)	(0.053)
CP index2	0.105	0.111	0.114	0.117	0.116	0.108***
	(0.080)	(0.082)	(0.081)	(0.082)	(0.087)	(0.036)
Lack time	0.176	0.187	0.155	0.159	−0.146	0.095
	(0.209)	(0.208)	(0.208)	(0.208)	(0.231)	(0.090)
CP × Lack time	0.129	0.127	0.109	0.120	0.107	0.041
	(0.248)	(0.255)	(0.253)	(0.259)	(0.252)	(0.106)
CP^2 × Lack time	0.048	0.038	0.042	0.026	0.045	−0.071
	(0.121)	(0.124)	(0.121)	(0.123)	(0.126)	(0.064)
Log Lik.	−1,441.75	−1,429.89	−1,424.36	−1,413.21	−1,346.19	−5,132.84
N obs.	2,800	2,800	2,800	2,800	2,800	4,998
R p	0.71	0.72	0.72	0.72	0.73	0.10
	$\log(\text{SA days})_t$					
CP index_{t-1}	0.095	0.099	0.094	0.091	0.080	0.000
	(0.078)	(0.078)	(0.077)	(0.077)	(0.076)	(0.01)
CP index2	0.065	0.064	0.065	0.062	0.072*	−0.002
	(0.040)	(0.040)	(0.040)	(0.040)	(0.039)	(0.12)
Lack time	−0.087	−0.093	−0.091	−0.094	0.062	−0.061
	(0.107)	(0.107)	(0.106)	(0.106)	(0.113)	(1.07)
CP × Lack time	−0.261**	−0.267**	−0.242*	−0.256**	−0.238*	−0.021
	(0.123)	(0.124)	(0.125)	(0.125)	(0.123)	(0.33)
CP^2 × Lack time	0.081	0.083	0.078	0.083	0.058	0.004
	(0.060)	(0.060)	(0.060)	(0.060)	(0.059)	(0.12)
Log Lik.	−5,075.35	−5,061.67	−5,038.79	−5,024.36	−4,971.04	−8,235.59
N obs.	4,988	4,988	4,988	4,988	4,988	4,988
Rsq	0.01	0.02	0.03	0.03	0.05	0.08
	WHO5_t					
CP index_{t-1}	0.028	0.044	0.040	0.028	−0.003	−0.355***
	(0.199)	(0.198)	(0.196)	(0.196)	(0.192)	(3.14)
CP index2	−0.227*	−0.232*	−0.231*	−0.222*	−0.242*	0.069
	(0.128)	(0.127)	(0.124)	(0.125)	(0.124)	(0.94)
Lack time	0.158	0.143	0.128	0.102	−0.415	−0.207
	(0.279)	(0.278)	(0.276)	(0.273)	(0.292)	(1.11)
CP × Lack time	0.307	0.282	0.170	0.163	0.208	0.278
	(0.331)	(0.331)	(0.330)	(0.332)	(0.318)	(1.28)
CP^2 × Lack time	0.152	0.156	0.176	0.181	0.183	−0.086
	(0.197)	(0.197)	(0.194)	(0.194)	(0.189)	(0.70)
Log Lik.	−10,121.21	−10,114.01	−10,078.05	−10,065.43	−9,903.77	−14,212.86
N obs.	4,998	4,998	4,998	4,998	4,998	4,998
Rsq	0.02	0.02	0.04	0.04	0.10	0.29
R adj.	0.02	0.02	0.03	0.03	0.09	0.28
Individual's controls	No	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	Yes	Yes	Yes	Yes
Employment controls	No	No	No	Yes	Yes	Yes
Working conditions	No	No	No	No	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; robust standard errors in parenthesis
Models controls are thus reported in the main tables

3.2 The mediated effects of intrinsic motivations

Table A15 – Estimates of the effects of competitive presenteeism on career moderated by useful job

	Dependent variable:					
	Fixed effects					Pooled
	(1)	(2)	(3)	(4)	(5)	(6)
	log(Hourly wage)$_t$					
CP index$_{t-1}$	0.019***	0.020***	0.021***	0.021***	0.019***	0.014**
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(2.19)
CP index$^2_{t-1}$	0.000	−0.000	−0.001	−0.001	−0.000	−0.007*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(1.89)
Useful job	0.004	0.002	0.002	0.003	0.002	−0.016
	(0.012)	(0.012)	(0.012)	(0.012)	(0.015)	(1.19)
CP × Useful job	−0.013	−0.015	−0.016	−0.015	−0.010	−0.021
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(1.59)
CP^2 × Useful job	−0.013	−0.011	−0.011	−0.011	−0.013	−0.001
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.09)
Log Lik.	6,542.38	6,556.35	6,566.19	6,577.66	6,632.61	507.01
N obs.	4,998	4,998	4,998	4,998	4,998	4,998
Rsq	0.15	0.15	0.16	0.16	0.18	0.51
R adj.	0.15	0.15	0.15	0.16	0.17	0.50
	Occupation$_t$					
CP index$_{t-1}$	0.011	0.002	−0.027	−0.036	−0.136	0.188***
	(0.196)	(0.188)	(0.194)	(0.197)	(0.205)	(3.58)
CP index$^2_{t-1}$	0.149	0.147	0.158*	0.186*	0.284***	0.036
	(0.099)	(0.096)	(0.096)	(0.100)	(0.108)	(0.90)
Useful job	−0.549	−0.643*	−0.566	−0.492	−0.016	−0.104
	(0.354)	(0.356)	(0.366)	(0.382)	(0.403)	(0.88)
CP × Useful job	−0.954**	−1.056***	−1.079***	−1.085***	−1.138***	−0.060
	(0.372)	(0.371)	(0.380)	(0.388)	(0.382)	(0.55)
CP^2 × Useful job	0.343	0.411*	0.371	0.270	0.188	0.038
	(0.223)	(0.225)	(0.228)	(0.236)	(0.241)	(0.41)
Log Lik.	−773.04	−744.86	−737.77	−717.26	−647.55	−4,239.71
N obs.	11,970	11,970	11,970	11,970	11,970	4,998
R p	0.81	0.82	0.82	0.83	0.84	0.44
Individual's controls	No	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	Yes	Yes	Yes	Yes
Employment controls	No	No	No	Yes	Yes	Yes
Working conditions	No	No	No	No	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; robust standard errors in parenthesis
Models controls are thus reported in the main tables

Table A16 – Estimates of the effects of competitive presenteeism on health moderated by useful job

	Dependent variable:					
	Fixed effects					Pooled
	(1)	(2)	(3)	(4)	(5)	(6)
	Health_t					
CP index_{t-1}	−0.120	−0.107	−0.137	−0.154	−0.230	−0.093*
	(0.148)	(0.149)	(0.147)	(0.150)	(0.152)	(0.053)
CP index^2_{t-1}	0.126*	0.126*	0.130*	0.125*	0.145*	0.068**
	(0.072)	(0.073)	(0.071)	(0.071)	(0.075)	(0.031)
Useful job	−0.599*	−0.606*	−0.572*	−0.610*	−0.721**	−0.049
	(0.317)	(0.318)	(0.317)	(0.319)	(0.349)	(0.123)
CP × Useful job	−0.430	−0.433	−0.427	−0.443	−0.467	0.046
	(0.296)	(0.297)	(0.298)	(0.294)	(0.313)	(0.113)
CP^2 × Useful job	−0.045	−0.036	−0.039	−0.013	−0.098	0.050
	(0.177)	(0.176)	(0.177)	(0.173)	(0.176)	(0.086)
Log Lik.	−1,432.22	−1,419.56	−1,414.59	−1,402.75	−1,321.66	−5,130.57
N obs.	14,570	14,570	14,570	14,570	14,570	4,998
R p	0.72	0.72	0.72	0.72	0.74	0.10
	$\log(\text{SA days})_t$					
CP index_{t-1}	0.023	0.023	0.026	0.022	0.036	−0.006
	(0.076)	(0.075)	(0.075)	(0.075)	(0.075)	(0.18)
CP index^2_{t-1}	0.084**	0.083**	0.085**	0.081**	0.076**	−0.007
	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.44)
Useful job	0.025	0.027	0.030	0.024	−0.100	−0.031
	(0.135)	(0.136)	(0.138)	(0.138)	(0.155)	(0.42)
CP × Useful job	0.039	0.044	0.047	0.040	−0.004	−0.027
	(0.136)	(0.136)	(0.138)	(0.139)	(0.139)	(0.40)
CP^2 × Useful job	−0.036	−0.040	−0.034	−0.026	−0.009	0.025
	(0.079)	(0.080)	(0.080)	(0.080)	(0.080)	(0.53)
Log Lik.	−5,083.68	−5,070.49	−5,046.10	−5,032.57	−4,976.11	−8,234.11
N obs.	4,988	4,988	4,988	4,988	4,988	4,988
Rsq	0.01	0.02	0.02	0.03	0.05	0.08
R adj.	0.01	0.01	0.02	0.02	0.04	0.07
	WHO5_t					
CP index_{t-1}	0.200	0.208	0.177	0.164	0.032	−0.269**
	(0.203)	(0.204)	(0.201)	(0.201)	(0.193)	(2.43)
CP index^2_{t-1}	−0.109	−0.117	−0.117	−0.107	−0.104	0.054
	(0.123)	(0.122)	(0.121)	(0.121)	(0.114)	(0.88)
Useful job	−0.159	−0.170	−0.148	−0.172	0.436	0.847***
	(0.409)	(0.407)	(0.399)	(0.402)	(0.437)	(3.61)
CP × Useful job	−0.272	−0.265	−0.281	−0.300	−0.149	−0.116
	(0.366)	(0.365)	(0.358)	(0.357)	(0.365)	(0.50)
CP^2 × Useful job	0.054	0.055	0.036	0.052	−0.045	0.038
	(0.291)	(0.289)	(0.279)	(0.283)	(0.292)	(0.25)
Log Lik.	−10,131.24	−10,123.20	−10,085.19	−10,072.21	−9,878.48	−14,186.46
N obs.	4,998	4,998	4,998	4,998	4,998	4,998
Rsq	0.01	0.02	0.03	0.04	0.11	0.30
R adj.	0.01	0.01	0.03	0.03	0.10	0.29
Individual's controls	No	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	Yes	Yes	Yes	Yes
Employment controls	No	No	No	Yes	Yes	Yes
Working conditions	No	No	No	No	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; robust standard errors in parenthesis
Models controls are thus reported in the main tables

Table A17 – Estimates of the effects of competitive presenteeism on career moderated by job pride

	Dependent variable:					
	Fixed effects					Pooled
	(1)	(2)	(3)	(4)	(5)	(6)
	log(Hourly wage)$_t$					
CP index$_{t-1}$	0.017**	0.017**	0.018**	0.018**	0.017**	0.011*
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(1.82)
CP index$^2_{t-1}$	−0.000	−0.000	−0.001	−0.001	−0.001	−0.006*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(1.71)
Job pride	0.017	0.016	0.017	0.016	0.008	−0.018
	(0.015)	(0.015)	(0.015)	(0.015)	(0.016)	(1.23)
CP × Job pride	0.002	0.002	0.001	0.002	0.003	−0.006
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.46)
CP^2 × Job pride	−0.011	−0.011	−0.012	−0.012	−0.012	−0.008
	(0.014)	(0.014)	(0.015)	(0.015)	(0.015)	(0.70)
Log Lik.	6,533.36	6,547.71	6,557.36	6,569.39	6,624.07	507.27
N obs.	4,998	4,998	4,998	4,998	4,998	4,998
Rsq	0.15	0.15	0.15	0.16	0.18	0.51
R adj.	0.14	0.15	0.15	0.15	0.17	0.50
	Occupation$_t$					
CP index$_{t-1}$	−0.046	−0.077	−0.115	−0.118	−0.190	0.175***
	(0.204)	(0.194)	(0.200)	(0.204)	(0.215)	(3.33)
CP index$^2_{t-1}$	0.222**	0.234**	0.238***	0.257***	0.328***	0.047
	(0.092)	(0.091)	(0.092)	(0.093)	(0.102)	(1.20)
Job pride	0.206	0.128	0.200	0.190	0.360	0.089
	(0.308)	(0.306)	(0.310)	(0.319)	(0.393)	(0.72)
CP × Job pride	−0.746**	−0.743**	−0.753**	−0.781**	−0.730**	−0.002
	(0.322)	(0.313)	(0.313)	(0.329)	(0.360)	(0.01)
CP^2 × Job pride	−0.387*	−0.385*	−0.392*	−0.391*	−0.384	−0.038
	(0.224)	(0.223)	(0.224)	(0.230)	(0.276)	(0.36)
Log Lik.	−766.61	−740.63	−732.92	−711.60	−648.44	−4,239.76
N obs.	11,970	11,970	11,970	11,970	11,970	4,998
R p	0.82	0.82	0.82	0.83	0.84	0.44
Individual's controls	No	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	Yes	Yes	Yes	Yes
Employment controls	No	No	No	Yes	Yes	Yes
Working conditions	No	No	No	No	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; robust standard errors in parenthesis
Models controls are thus reported in the main tables

Table A18 – Estimates of the effects of competitive presenteeism on health moderated by job pride

	Dependent variable:					
	Fixed effects					Pooled
	(1)	(2)	(3)	(4)	(5)	(6)
	Health_t					
CP index_{t-1}	−0.233	−0.221	−0.250*	−0.261*	−0.350**	−0.081
	(0.145)	(0.146)	(0.144)	(0.146)	(0.147)	(0.052)
CP index^2_{t-1}	0.162**	0.162**	0.165**	0.157**	0.174**	0.062**
	(0.075)	(0.076)	(0.073)	(0.073)	(0.079)	(0.031)
Job pride	0.233	0.235	0.249	0.192	0.420	−0.178
	(0.294)	(0.300)	(0.304)	(0.311)	(0.355)	(0.120)
CP × Job pride	0.226	0.216	0.193	0.159	0.181	−0.023
	(0.258)	(0.269)	(0.277)	(0.276)	(0.297)	(0.107)
CP^2 × Job pride	−0.161	−0.173	−0.173	−0.139	−0.184	0.119
	(0.185)	(0.190)	(0.194)	(0.192)	(0.202)	(0.081)
Log Lik.	−1,445.83	−1,433.39	−1,426.84	−1,415.78	−1,336.81	−5,129.54
N obs.	14,570	14,570	14,570	14,570	14,570	4,998
R p	0.71	0.72	0.72	0.72	0.74	0.10
	$\log(\text{SA days})_t$					
CP index_{t-1}	0.027	0.029	0.032	0.028	0.033	−0.001
	(0.074)	(0.073)	(0.074)	(0.073)	(0.073)	(0.03)
CP index^2_{t-1}	0.087**	0.086**	0.087**	0.085**	0.082**	−0.001
	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.07)
Job pride	0.032	0.042	0.041	0.035	−0.015	0.033
	(0.136)	(0.136)	(0.136)	(0.136)	(0.138)	(0.43)
CP × Job pride	0.008	0.008	0.013	0.001	−0.011	−0.063
	(0.134)	(0.134)	(0.134)	(0.136)	(0.131)	(0.94)
CP^2 × Job pride	−0.123	−0.127	−0.117	−0.116	−0.116	−0.052
	(0.078)	(0.078)	(0.077)	(0.076)	(0.076)	(1.05)
Log Lik.	−5,079.64	−5,066.40	−5,042.72	−5,028.81	−4,972.28	−8,232.10
N obs.	4,988	4,988	4,988	4,988	4,988	4,988
Rsq	0.01	0.02	0.03	0.03	0.05	0.08
R adj.	0.01	0.01	0.02	0.03	0.04	0.07
	WHO5_t					
CP index_{t-1}	0.098	0.108	0.075	0.062	−0.040	−0.267**
	(0.198)	(0.197)	(0.196)	(0.195)	(0.190)	(2.46)
CP index^2_{t-1}	−0.082	−0.089	−0.088	−0.077	−0.089	0.039
	(0.117)	(0.117)	(0.115)	(0.116)	(0.109)	(0.65)
Job pride	0.308	0.324	0.336	0.327	0.790*	0.612**
	(0.414)	(0.416)	(0.416)	(0.414)	(0.441)	(2.39)
CP × Job pride	0.348	0.339	0.334	0.312	0.390	−0.135
	(0.351)	(0.351)	(0.348)	(0.347)	(0.339)	(0.57)
CP^2 × Job pride	−0.063	−0.081	−0.119	−0.113	−0.103	0.184
	(0.223)	(0.224)	(0.222)	(0.225)	(0.228)	(1.16)
Log Lik.	−10,130.00	−10,122.07	−10,084.51	−10,071.77	−9,873.79	−14,188.87
N obs.	4,998	4,998	4,998	4,998	4,998	4,998
Rsq	0.01	0.02	0.03	0.04	0.11	0.30
R adj.	0.01	0.01	0.03	0.03	0.10	0.29
Individual's controls	No	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	Yes	Yes	Yes	Yes
Employment controls	No	No	No	Yes	Yes	Yes
Working conditions	No	No	No	No	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; robust standard errors in parenthesis
Models controls are thus reported in the main tables

4 Heterogenous effects

4.1 Gender

Table A19 – Gender heterogeneity in the estimates of the effects of competitive presenteeism on career

	Dependent variable:					
	Fixed effects					Pooled
	(1)	(2)	(3)	(4)	(5)	(6)
	log(Hourly wage)$_t$					
CP index	0.017*	0.018*	0.020**	0.019**	0.022**	−0.006
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.007)
CP index^2	−0.003	−0.004	−0.004	−0.005	−0.005	−0.006
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.008)
CP × Female	−0.000	−0.000	−0.002	−0.002	−0.010	0.028***
	(0.015)	(0.014)	(0.014)	(0.014)	(0.014)	(0.009)
CP^2 × Female	0.006	0.006	0.007	0.007	0.008	−0.017***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)
Female						−0.062***
						(0.009)
Log Lik.	6,531.04	6,545.73	6,555.11	6,567.33	6,623.02	509.89
N obs.	4,998	4,998	4,998	4,998	4,998	4,998
Rsq	0.14	0.15	0.15	0.16	0.18	0.51
R adj.	0.14	0.15	0.15	0.15	0.17	0.50
	Occupation$_t$					
CP index	−0.315	−0.385*	−0.436*	−0.528**	−0.654**	0.248***
	(0.232)	(0.230)	(0.230)	(0.241)	(0.263)	(0.061)
CP index^2	0.182	0.211*	0.233*	0.295**	0.420***	0.016
	(0.122)	(0.122)	(0.120)	(0.131)	(0.133)	(0.039)
CP × Female	0.344	0.419	0.457	0.628*	0.692*	-0.161**
	(0.383)	(0.368)	(0.373)	(0.379)	(0.388)	(0.081)
CP^2 × Female	0.022	−0.019	−0.059	−0.187	−0.291	0.057
	(0.179)	(0.178)	(0.179)	(0.192)	(0.195)	(0.070)
Female						−0.405***
						(0.095)
Log Lik.	−782.59	−756.30	−749.29	−727.10	−659.02	−4,238.31
N obs.	2,572	2,572	2,572	2,572	2,572	4,998
R p	0.81	0.82	0.82	0.82	0.84	0.44
Individual's controls	No	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	Yes	Yes	Yes	Yes
Employment controls	No	No	No	Yes	Yes	Yes
Working conditions	No	No	No	No	Yes	Yes
Notes:	*p<0.1; **p<0.05; ***p<0.01 ; robust standard errors in parenthesis Models controls are thus reported in the main tables					

Table A20 – Gender heterogeneity in the estimates of the effects of competitive presenteeism on health

	Dependent variable:					
	Fixed effects					Pooled
	(1)	(2)	(3)	(4)	(5)	(6)
	Health_t					
CP index	−0.195	−0.179	−0.208	−0.231	−0.249	−0.128**
	(0.176)	(0.177)	(0.174)	(0.176)	(0.186)	(0.058)
CP index^2	0.148*	0.153*	0.159*	0.161*	0.199**	0.129***
	(0.083)	(0.085)	(0.083)	(0.083)	(0.091)	(0.032)
CP × Female	0.011	−0.008	−0.022	−0.012	−0.228	0.091
	(0.268)	(0.270)	(0.269)	(0.272)	(0.276)	(0.078)
CP^2 × Female	−0.014	−0.024	−0.029	−0.042	−0.077	−0.144***
	(0.141)	(0.142)	(0.136)	(0.136)	(0.140)	(0.052)
Female						0.048
						(0.088)
Log Lik.	−1,447.58	−1,435.05	−1,428.38	−1,416.56	−1,336.65	−5,126.48
N obs.	14,570	14,570	14,570	14,570	14,570	4,998
R p	0.71	0.72	0.72	0.72	0.74	0.10
	$\log(\text{SA days})_t$					
CP index	0.016	0.014	0.013	0.019	0.011	−0.011
	(0.101)	(0.101)	(0.100)	(0.101)	(0.099)	(0.037)
CP index^2	0.089**	0.087*	0.089**	0.082*	0.077*	−0.011
	(0.045)	(0.045)	(0.044)	(0.044)	(0.044)	(0.019)
CP × Female	0.032	0.040	0.050	0.037	0.067	0.002
	(0.146)	(0.146)	(0.146)	(0.146)	(0.146)	(0.050)
CP^2 × Female	−0.023	−0.021	−0.021	−0.010	−0.001	0.020
	(0.069)	(0.069)	(0.069)	(0.069)	(0.069)	(0.033)
Female						0.088
						(0.054)
Log Lik.	−5,083.83	−5,070.71	−5,046.16	−5,029.87	−4,974.93	−8,233.88
N obs.	4,988	4,988	4,988	4,988	4,988	4,988
Rsq	0.01	0.02	0.02	0.03	0.05	0.08
R adj.	0.01	0.01	0.02	0.03	0.04	0.07
	WHO5_t					
CP index	0.144	0.212	0.187	0.175	0.176	−0.340***
	(0.234)	(0.236)	(0.234)	(0.233)	(0.239)	(0.119)
CP index^2	−0.098	−0.125	−0.127	−0.117	−0.133	0.071
	(0.143)	(0.143)	(0.143)	(0.143)	(0.135)	(0.062)
CP × Female	−0.128	−0.123	−0.143	−0.152	−0.398	0.109
	(0.393)	(0.396)	(0.391)	(0.387)	(0.374)	(0.170)
CP^2 × Female	0.029	0.044	0.045	0.049	0.039	−0.021
	(0.237)	(0.235)	(0.230)	(0.230)	(0.218)	(0.110)
Female						−0.934***
						(0.189)
Log Lik.	−10,153.04	−10,124.03	−10,086.18	−10,073.25	−9,879.81	−14,199.21
N obs.	4,998	4,998	4,998	4,998	4,998	4,998
Rsq	0.01	0.02	0.03	0.04	0.11	0.30
R adj.	0.00	0.01	0.03	0.03	0.10	0.28
Individual's controls	No	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	Yes	Yes	Yes	Yes
Employment controls	No	No	No	Yes	Yes	Yes
Working conditions	No	No	No	No	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; robust standard errors in parenthesis
Models controls are thus reported in the main tables

4.2 Public sector

Table A21 – Public sector heterogeneity in the estimates of the effects of competitive presenteeism on career

	Dependent variable:					
	Fixed effects					Pooled
	(1)	(2)	(3)	(4)	(5)	(6)
	log(Hourly wage)$_t$					
CP index	0.017**	0.017**	0.019***	0.018***	0.018***	0.010*
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)
CP index^2	−0.001	−0.001	−0.001	−0.002	−0.002	−0.003
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)
CP × Public	−0.004	−0.008	−0.008	−0.007	−0.027	−0.016
	(0.040)	(0.041)	(0.041)	(0.041)	(0.040)	(0.019)
CP^2 × Public	−0.005	−0.003	−0.005	−0.005	0.003	−0.014
	(0.015)	(0.015)	(0.015)	(0.015)	(0.014)	(0.010)
Public						−0.015
						(0.017)
Log Lik.	6,529.78	6,544.30	6,553.52	6,565.68	6,621.77	510.14
N obs.	4,998	4,998	4,998	4,998	4,998	4,998
Rsq	0.14	0.15	0.15	0.16	0.18	0.51
R adj.	0.14	0.15	0.15	0.15	0.17	0.50
	Occupation$_t$					
CP index	−0.176	−0.213	−0.251	−0.264	−0.343*	0.195***
	(0.185)	(0.181)	(0.187)	(0.195)	(0.208)	(0.047)
CP index^2	0.183*	0.197**	0.200**	0.203**	0.281***	0.066*
	(0.093)	(0.091)	(0.091)	(0.094)	(0.098)	(0.035)
CP × Public	0.150	0.341	0.475	0.626	0.623	−0.469**
	(0.664)	(0.586)	(0.600)	(0.661)	(0.749)	(0.189)
CP^2 × Public	0.682*	0.472	0.372	0.354	0.488	−0.058
	(0.375)	(0.363)	(0.367)	(0.372)	(0.367)	(0.082)
Public						0.668***
						(0.158)
Log Lik.	−783.52	−758.06	−751.13	−730.35	−663.05	−4,232.48
N obs.	2,572	2,572	2,572	2,572	2,572	4,998
R p	0.81	0.82	0.82	0.82	0.84	0.44
Individual's controls	No	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	Yes	Yes	Yes	Yes
Employment controls	No	No	No	Yes	Yes	Yes
Working conditions	No	No	No	No	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; robust standard errors in parenthesis
Models controls are thus reported in the main tables

Table A22 – Public sector heterogeneity in the estimates of the effects of competitive presenteeism on health

	Dependent variable:					
	Fixed effects					Pooled
	(1)	(2)	(3)	(4)	(5)	(6)
	Health_t					
CP index	−0.175	−0.163	−0.196	−0.213	−0.296**	−0.084*
	(0.136)	(0.138)	(0.136)	(0.138)	(0.141)	(0.048)
CP index^2	0.135*	0.136*	0.138*	0.132*	0.150**	0.076***
	(0.072)	(0.073)	(0.071)	(0.070)	(0.076)	(0.027)
CP × Public	−0.886	−1.059	−1.028	−1.145	−1.517*	−0.079
	(0.839)	(0.812)	(0.815)	(0.821)	(0.832)	(0.173)
CP^2 × Public	0.208	0.224	0.247	0.315	0.310	−0.004
	(0.336)	(0.334)	(0.318)	(0.324)	(0.333)	(0.104)
Public						−0.115
						(0.156)
Log Lik.	−1,446.40	−1,433.50	−1,426.97	−1,414.77	−1,336.40	−5,131.10
N obs.	14,570	14,570	14,570	14,570	14,570	4,998
R p	0.71	0.72	0.72	0.72	0.74	0.10
	$\log(\text{SA days})_t$					
CP index	0.019	0.019	0.025	0.026	0.027	−0.010
	(0.074)	(0.074)	(0.074)	(0.074)	(0.073)	(0.031)
CP index^2	0.078**	0.077**	0.079**	0.076**	0.076**	0.002
	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.016)
CP × Public	0.441	0.472	0.406	0.342	0.473	−0.055
	(0.399)	(0.398)	(0.372)	(0.376)	(0.384)	(0.133)
CP^2 × Public	−0.026	−0.032	−0.036	−0.025	−0.070	−0.022
	(0.149)	(0.147)	(0.144)	(0.144)	(0.147)	(0.061)
Public						0.376***
						(0.109)
Log Lik.	−5,081.79	−5,068.42	−5,044.65	−5,028.69	−4,973.31	−8,233.34
N obs.	4,988	4,988	4,988	4,988	4,988	4,988
Rsq	0.01	0.02	0.03	0.03	0.05	0.08
R adj.	0.01	0.01	0.02	0.03	0.04	0.07
	WHO5_t					
CP index	0.134	0.198	0.159	0.140	0.052	−0.262***
	(0.195)	(0.196)	(0.194)	(0.193)	(0.185)	(0.101)
CP index^2	−0.100	−0.118	−0.122	−0.110	−0.131	0.065
	(0.116)	(0.115)	(0.114)	(0.114)	(0.107)	(0.054)
CP × Public	−1.732*	−1.541	−1.279	−1.197	−1.659	−0.678*
	(1.021)	(1.008)	(1.035)	(1.047)	(1.049)	(0.411)
CP^2 × Public	0.552	0.453	0.479	0.476	0.681	0.132
	(0.685)	(0.686)	(0.667)	(0.669)	(0.673)	(0.198)
Public						−1.170***
						(0.361)
Log Lik.	−10,149.86	−10,121.57	−10,084.39	−10,071.67	−9,878.01	−14,197.44
N obs.	4,998	4,998	4,998	4,998	4,998	4,998
Rsq	0.01	0.02	0.03	0.04	0.11	0.30
R adj.	0.00	0.01	0.03	0.03	0.10	0.28
Individual's controls	No	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	Yes	Yes	Yes	Yes
Employment controls	No	No	No	Yes	Yes	Yes
Working conditions	No	No	No	No	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; robust standard errors in parenthesis
Models controls are thus reported in the main tables

4.3 Occupation (blue or white collars)

Table A23 – Blue collars heterogeneity in the estimates of the effects of competitive presenteeism on career

	Dependent variable:					
	Fixed effects					Pooled
	(1)	(2)	(3)	(4)	(5)	(6)
	log(Hourly wage)$_t$					
CP index	0.023***	0.022**	0.023***	0.022***	0.021**	0.014**
	(0.009)	(0.009)	(0.009)	(0.009)	(0.008)	(0.006)
CP index^2	−0.003	−0.003	−0.003	−0.004	−0.003	−0.007**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)
CP × BlueCol	−0.014	−0.013	−0.011	−0.011	−0.010	−0.019*
	(0.015)	(0.015)	(0.015)	(0.014)	(0.014)	(0.012)
CP^2 × BlueCol	0.009	0.008	0.008	0.009	0.007	−0.013
	(0.013)	(0.013)	(0.013)	(0.013)	(0.012)	(0.010)
BlueCol						0.019
						(0.029)
Log Lik.	6,533.48	6,546.92	6,555.53	6,567.67	6,622.80	506.07
N obs.	4,998	4,998	4,998	4,998	4,998	4,998
Rsq	0.15	0.15	0.15	0.16	0.18	0.51
R adj.	0.14	0.15	0.15	0.15	0.17	0.50
	Occupation$_t$					
CP index	−0.087	−0.132	−0.209	−0.240	−0.758*	0.008
	(0.272)	(0.265)	(0.278)	(0.277)	(0.422)	(0.069)
CP index^2	0.174	0.225*	0.196	0.193	0.495**	0.108**
	(0.124)	(0.129)	(0.136)	(0.146)	(0.198)	(0.048)
CP × BlueCol	−0.541	−0.548	−0.535	−0.013	−0.693	0.226*
	(0.585)	(0.603)	(0.587)	(0.642)	(0.969)	(0.120)
CP^2 × BlueCol	−0.870	−0.838	−0.818	−0.414	−1.559**	−0.208**
	(0.537)	(0.582)	(0.554)	(0.603)	(0.786)	(0.096)
BlueCol						−23.831***
						(0.148)
Log Lik.	−183.59	−170.85	−165.82	−158.82	−124.10	−2,727.15
N obs.	766	766	766	766	766	4,998
R p	0.46	0.49	0.51	0.53	0.63	0.64
Individual's controls	No	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	Yes	Yes	Yes	Yes
Employment controls	No	No	No	Yes	Yes	Yes
Working conditions	No	No	No	No	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; robust standard errors in parenthesis
Models controls are thus reported in the main tables

Table A24 – Blue collars heterogeneity in the estimates of the effects of competitive presenteeism on health

	Dependent variable:					
	Fixed effects					Pooled
	(1)	(2)	(3)	(4)	(5)	(6)
	Health_t					
CP index	−0.351**	−0.335**	−0.376**	−0.398**	−0.554***	−0.015
	(0.163)	(0.164)	(0.163)	(0.166)	(0.162)	(0.056)
CP index^2	0.199**	0.198**	0.204***	0.199**	0.243***	0.050
	(0.080)	(0.081)	(0.078)	(0.078)	(0.084)	(0.031)
CP × BlueCol	0.390	0.374	0.396	0.396	0.548*	−0.278***
	(0.291)	(0.292)	(0.285)	(0.290)	(0.302)	(0.097)
CP^2 × BlueCol	−0.146	−0.152	−0.148	−0.146	−0.248	−0.030
	(0.235)	(0.232)	(0.224)	(0.229)	(0.230)	(0.092)
BlueCol						0.127
						(0.261)
Log Lik.	−1,441.77	−1,430.34	−1,423.43	−1,411.77	−1,331.93	−5,126.93
N obs.	14,570	14,570	14,570	14,570	14,570	4,998
R p	0.71	0.72	0.72	0.72	0.74	0.10
	$\log(\text{SA days})_t$					
CP index	−0.004	−0.003	0.003	0.008	0.027	−0.019
	(0.080)	(0.079)	(0.080)	(0.080)	(0.079)	(0.034)
CP index^2	0.097***	0.095***	0.095***	0.090**	0.084**	0.007
	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)	(0.017)
CP × BlueCol	0.051	0.052	0.054	0.049	−0.009	−0.023
	(0.157)	(0.158)	(0.158)	(0.160)	(0.156)	(0.066)
CP^2 × BlueCol	−0.147	−0.149	−0.127	−0.106	−0.109	−0.128***
	(0.132)	(0.132)	(0.132)	(0.132)	(0.131)	(0.048)
BlueCol						0.327*
						(0.193)
Log Lik.	−5,077.24	−5,063.70	−5,040.49	−5,024.70	−4,969.89	−8,230.34
N obs.	4,988	4,988	4,988	4,988	4,988	4,988
Rsq	0.01	0.02	0.03	0.03	0.05	0.09
R adj.	0.01	0.01	0.02	0.03	0.04	0.07
	WHO5_t					
CP index	−0.151	−0.102	−0.149	−0.178	−0.297	−0.236**
	(0.221)	(0.221)	(0.220)	(0.219)	(0.208)	(0.113)
CP index^2	0.011	−0.002	0.003	0.020	0.008	0.028
	(0.126)	(0.125)	(0.124)	(0.124)	(0.116)	(0.058)
CP × BlueCol	0.645	0.705	0.717*	0.755*	0.869**	−0.093
	(0.431)	(0.430)	(0.415)	(0.416)	(0.427)	(0.224)
CP^2 × BlueCol	−0.434	−0.470	−0.543	−0.558	−0.486	0.273
	(0.417)	(0.411)	(0.396)	(0.399)	(0.415)	(0.176)
BlueCol						0.283
						(0.649)
Log Lik.	−10,145.40	−10,115.61	−10,076.79	−10,063.07	−9,870.10	−14,197.75
N obs.	4,998	4,998	4,998	4,998	4,998	4,998
Rsq	0.01	0.02	0.04	0.04	0.11	0.30
R adj.	0.01	0.02	0.03	0.04	0.10	0.28
Individual's controls	No	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	Yes	Yes	Yes	Yes
Employment controls	No	No	No	Yes	Yes	Yes
Working conditions	No	No	No	No	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; robust standard errors in parenthesis
Models controls are thus reported in the main tables

4.4 Healthcare sector subset

Table A25 – Estimates of the effects of competitive presenteeism on career (Healthcare subset)

	Dependent variable:						
	Fixed effects					HT	Pooled
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	log(Hourly wage)$_t$						
CP index	0.020***	0.015**	0.017***	0.019***	0.017***	0.019***	0.020***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.004)
CP index^2	0.010**	0.009**	0.009**	0.009**	0.008*	0.008*	−0.016***
	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)
Log Lik.	7,481.63	7,729.39	7,738.42	7,812.15	7,877.79	.	782.80
N obs.	6,859	6,859	6,859	6,859	6,859	6,859	6,859
Rsq	0.01	0.08	0.09	0.11	0.12	.	0.50
R adj.	0.01	0.08	0.08	0.10	0.12	.	0.49
	Occupation$_t$						
CP index	−0.448**	−0.422**	−0.428**	−0.446**	−0.527**	−0.047***	0.862***
	(0.203)	(0.206)	(0.208)	(0.208)	(0.214)	(0.018)	(0.045)
CP index^2	0.178	0.165	0.165	0.164	0.161	0.018**	0.053
	(0.114)	(0.112)	(0.111)	(0.116)	(0.114)	(0.009)	(0.033)
Log Lik.	−625.42	−615.85	−611.56	−603.32	−562.93	.	−4,871.72
N obs.	3,536	3,536	3,536	3,536	3,536	6,859	6,859
R p	0.89	0.89	0.89	0.90	0.90	.	0.39
	Health$_t$						
CP index	−0.098	−0.048	−0.081	−0.079	−0.119	−0.049	−0.020
	(0.118)	(0.121)	(0.121)	(0.121)	(0.123)	(0.037)	(0.040)
CP index^2	0.032	0.026	0.027	0.027	0.017	0.011	0.006
	(0.062)	(0.063)	(0.063)	(0.063)	(0.064)	(0.020)	(0.022)
Log Lik.	−3,795.54	−3,735.21	−3,724.33	−3,719.12	−3,578.53	.	−7,840.63
N obs.	24,802	24,802	24,802	24,802	24,802	6,859	6,859
R p	0.56	0.57	0.57	0.57	0.58	.	0.02
	log(SA days)$_t$						
CP index	0.003	0.004	0.039	0.035	0.051	0.051	−0.029
	(0.067)	(0.067)	(0.066)	(0.066)	(0.066)	(0.066)	(0.028)
CP index^2	0.006	0.006	−0.003	−0.008	−0.015	−0.016	−0.010
	(0.031)	(0.031)	(0.031)	(0.031)	(0.031)	(0.030)	(0.015)
Log Lik.	−7,535.61	−7,514.72	−7,462.31	−7,452.50	−7,399.64	.	−11,750.28
N obs.	6,841	6,841	6,841	6,841	6,841	6,841	6,841
Rsq	0.01	0.01	0.03	0.03	0.04	.	0.07
R adj.	0.00	0.01	0.02	0.03	0.04	.	0.06
	WHO5$_t$						
CP index	0.042	0.072	−0.040	−0.024	−0.105	−0.112	−0.318***
	(0.164)	(0.164)	(0.159)	(0.159)	(0.157)	(0.158)	(0.089)
CP index^2	0.054	0.069	0.096	0.094	0.101	0.102	0.091*
	(0.085)	(0.085)	(0.084)	(0.084)	(0.083)	(0.083)	(0.050)
Log Lik.	−14,095.41	−14,074.14	−14,005.76	−13,995.74	−13,758.06	.	−19,575.93
N obs.	6,859	6,859	6,859	6,859	6,859	6,859	6,859
Rsq	0.00	0.01	0.03	0.03	0.10	.	0.25
R adj.	0.00	0.01	0.03	0.03	0.09	.	0.25
Individual's controls	No	Yes	Yes	Yes	Yes	Yes	
Health controls	No	No	Yes	Yes	Yes	Yes	
Employment controls	No	No	No	Yes	Yes	Yes	
Working conditions	No	No	No	No	Yes	Yes	

Notes: *p<0.1; **p<0.05; ***p<0.01 ; robust standard errors in parenthesis
Models controls are thus reported in the main tables

4.5 Public healthcare sector

Table A26 – Public sector heterogeneity in the estimates of the effects of competitive presenteeism on career (Healthcare subset)

	Dependent variable:					
	Fixed effects					Pooled
	(1)	(2)	(3)	(4)	(5)	(6)
	log(Hourly wage)$_t$					
CP index	0.036***	0.036***	0.038***	0.039***	0.036***	0.027***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.010)
CP index^2	0.008	0.007	0.008	0.009	0.010	−0.021**
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)
CP × Public	−0.027**	−0.027**	−0.027**	−0.026**	−0.025**	−0.009
	(0.013)	(0.012)	(0.012)	(0.012)	(0.012)	(0.011)
CP^2 × Public	0.002	0.002	0.002	0.000	−0.001	0.006
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Public						−0.014
						(0.011)
Log Lik.	7,734.14	7,737.83	7,746.65	7,820.19	7,885.07	783.29
N obs.	6,859	6,859	6,859	6,859	6,859	6,859
Rsq	0.08	0.09	0.09	0.11	0.12	0.50
R adj.	0.08	0.08	0.08	0.10	0.12	0.50
	Occupation$_t$					
CP index	−0.242	−0.249	−0.250	−0.303	−0.421	0.634***
	(0.348)	(0.347)	(0.350)	(0.347)	(0.340)	(0.118)
CP index^2	0.189	0.193	0.177	0.171	0.168	0.295***
	(0.274)	(0.270)	(0.277)	(0.284)	(0.280)	(0.096)
CP × Public	−0.224	−0.215	−0.228	−0.179	−0.129	0.287**
	(0.402)	(0.416)	(0.416)	(0.411)	(0.394)	(0.122)
CP^2 × Public	0.015	0.011	0.031	0.024	0.041	−0.286***
	(0.307)	(0.301)	(0.306)	(0.314)	(0.304)	(0.101)
Public						1.494***
						(0.115)
Log Lik.	−605.90	−594.24	−591.74	−585.69	−548.46	−4,865.22
N obs.	3,536	3,536	3,536	3,536	3,536	6,859
R p	0.90	0.90	0.90	0.90	0.91	0.39
Individual's controls	No	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	Yes	Yes	Yes	Yes
Employment controls	No	No	No	Yes	Yes	Yes
Working conditions	No	No	No	No	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; robust standard errors in parenthesis
Models controls are thus reported in the main tables

Table A27 – Public sector heterogeneity in the estimates of the effects of competitive presenteeism on health (Healthcare subset)

	Dependent variable:					
	Fixed effects					Pooled
	(1)	(2)	(3)	(4)	(5)	(6)
	Health_t					
CP index	0.219	0.307	0.270	0.264	0.171	−0.159**
	(0.228)	(0.232)	(0.230)	(0.230)	(0.222)	(0.078)
CP index^2	0.092	0.084	0.089	0.069	0.037	0.103
	(0.168)	(0.171)	(0.165)	(0.159)	(0.157)	(0.063)
CP × Public	−0.412	−0.460*	−0.454*	−0.441*	−0.373	0.177**
	(0.256)	(0.259)	(0.257)	(0.257)	(0.251)	(0.085)
CP^2 × Public	−0.058	−0.054	−0.058	−0.035	−0.010	−0.115*
	(0.181)	(0.184)	(0.178)	(0.174)	(0.171)	(0.067)
Public						0.045
						(0.087)
Log Lik.	−3,785.14	−3,723.87	−3,713.19	−3,708.80	−3,570.36	−7,838.07
N obs.	24,802	24,802	24,802	24,802	24,802	6,859
R p	0.56	0.57	0.57	0.57	0.58	0.02
	$\log(\text{SA days})_t$					
CP index	0.066	0.059	0.095	0.081	0.103	0.006
	(0.134)	(0.133)	(0.132)	(0.132)	(0.132)	(0.058)
CP index^2	−0.042	−0.030	−0.032	−0.033	−0.038	−0.104**
	(0.076)	(0.075)	(0.075)	(0.075)	(0.074)	(0.043)
CP × Public	−0.087	−0.075	−0.076	−0.064	−0.071	−0.047
	(0.146)	(0.145)	(0.144)	(0.144)	(0.144)	(0.062)
CP^2 × Public	0.059	0.045	0.037	0.032	0.030	0.104**
	(0.083)	(0.082)	(0.082)	(0.081)	(0.081)	(0.046)
Public						-0.036
						(0.065)
Log Lik.	−7,532.85	−7,512.22	−7,459.61	−7,449.83	−7,397.34	−11,748.17
N obs.	6,841	6,841	6,841	6,841	6,841	6,841
Rsq	0.01	0.01	0.03	0.03	0.05	0.07
R adj.	0.00	0.01	0.02	0.03	0.04	0.06
	WHO5_t					
CP index	−0.117	−0.063	−0.159	−0.129	−0.252	−0.259
	(0.293)	(0.294)	(0.282)	(0.283)	(0.284)	(0.175)
CP index^2	0.076	0.078	0.094	0.112	0.033	0.063
	(0.212)	(0.204)	(0.187)	(0.188)	(0.182)	(0.145)
CP × Public	0.210	0.179	0.158	0.142	0.192	−0.074
	(0.333)	(0.334)	(0.322)	(0.322)	(0.320)	(0.190)
CP^2 × Public	−0.032	−0.017	−0.003	−0.027	0.069	0.035
	(0.232)	(0.226)	(0.210)	(0.210)	(0.204)	(0.154)
Public						−0.206
						(0.192)
Log Lik.	−14,094.72	−14,073.40	−14,004.86	−13,994.69	−13,756.96	−19,575.85
N obs.	6,859	6,859	6,859	6,859	6,859	6,859
Rsq	0.00	0.01	0.03	0.03	0.10	0.25
R adj.	0.00	0.01	0.03	0.03	0.09	0.24
Individual's controls	No	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	Yes	Yes	Yes	Yes
Employment controls	No	No	No	Yes	Yes	Yes
Working conditions	No	No	No	No	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; robust standard errors in parenthesis
Models controls are thus reported in the main tables

5 Direct effects of the variables included in the CP indicator

Table A28 – Estimates of the variables included in the CP index$_{t-1}$ on hourly wage

	log(hourly wage)$_t$						
	Fixed effects					HT	Pooled
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
SmtsComp	−0.005	−0.006	−0.005	−0.006	−0.004	−0.004	−0.004
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
SmtsUncomp	0.010	0.008	0.010	0.009	0.011	0.011	−0.005
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Oft_EvDComp	0.014	0.015	0.017	0.016	0.016	0.015	0.002
	(0.015)	(0.015)	(0.015)	(0.015)	(0.014)	(0.015)	(0.013)
Oft_EvDUncomp	0.001	−0.000	0.001	0.000	0.001	0.002	−0.012
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Mean	0.014	0.013	0.014	0.013	0.011	0.011	−0.008
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.012)
1to3h	0.023**	0.024**	0.023**	0.022**	0.023**	0.023**	−0.030***
	(0.010)	(0.010)	(0.010)	(0.009)	(0.009)	(0.009)	(0.008)
4to6h	0.022	0.023*	0.023*	0.021	0.023*	0.024*	−0.014
	(0.013)	(0.013)	(0.013)	(0.013)	(0.014)	(0.014)	(0.013)
more6h	0.056***	0.057***	0.058***	0.056***	0.058***	0.058***	−0.040**
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.016)
NotConcerned	0.001	0.001	0.001	0.002	0.000	0.000	−0.011
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Sometimes	0.014	0.014	0.014	0.015	0.018	0.019	0.019
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.012)	(0.012)
Often	0.065***	0.067***	0.068***	0.068***	0.067***	0.068***	0.035*
	(0.023)	(0.023)	(0.023)	(0.022)	(0.022)	(0.022)	(0.020)
Everyday	0.011	0.012	0.011	0.009	0.008	0.011	0.016
	(0.036)	(0.037)	(0.037)	(0.037)	(0.036)	(0.036)	(0.026)
Sometimes	−0.013	−0.012	−0.011	−0.012	−0.012	−0.013	0.015
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Often	−0.013	−0.013	−0.012	−0.013	−0.016	−0.018	0.017
	(0.013)	(0.013)	(0.013)	(0.012)	(0.013)	(0.013)	(0.012)
Always	−0.002	−0.002	0.001	0.000	−0.001	−0.002	0.034**
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.016)
Log Lik.	6,569.37	6,585.87	6,595.70	6,607.13	6,664.69	.	517.99
N obs.	4,998	4,998	4,998	4,998	4,998	4,998	4,998
Rsq	0.16	0.16	0.17	0.17	0.19	.	0.51
R adj.	0.15	0.16	0.16	0.16	0.18	.	0.50
Individual's controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	Yes	Yes	Yes	Yes	Yes
Employment controls	No	No	No	Yes	Yes	Yes	Yes
Working conditions	No	No	No	No	Yes	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; robust standard errors in parenthesis
Models controls are thus reported in the main tables

Table A29 – Estimates of the variables included in the CP index$_{t-1}$ on occupation

	Occupation$_t$						
	Fixed effects					HT	Pooled
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
SmtsComp	0.143	0.134	0.093	0.019	0.019	0.031	0.076
	(0.282)	(0.287)	(0.291)	(0.291)	(0.272)	(0.046)	(0.085)
SmtsUncomp	−0.051	−0.094	−0.144	−0.218	−0.350	−0.012	0.242**
	(0.305)	(0.325)	(0.321)	(0.320)	(0.329)	(0.050)	(0.099)
Oft_EvDComp	−0.010	−0.091	−0.106	−0.085	−0.286	0.033	0.152
	(0.395)	(0.417)	(0.430)	(0.447)	(0.466)	(0.062)	(0.113)
Oft_EvDUncomp	−0.150	−0.103	−0.176	−0.203	−0.272	−0.039	0.516***
	(0.411)	(0.409)	(0.409)	(0.412)	(0.425)	(0.056)	(0.127)
Mean	−0.010	−0.008	0.011	0.022	0.107	0.048	−0.482***
	(0.287)	(0.305)	(0.306)	(0.306)	(0.323)	(0.055)	(0.112)
1to3h	0.196	0.244	0.292	0.310	0.412	0.013	−0.192**
	(0.259)	(0.274)	(0.277)	(0.284)	(0.319)	(0.040)	(0.077)
4to6h	0.741**	0.748**	0.805**	0.796**	0.784**	0.110*	−0.345***
	(0.324)	(0.335)	(0.330)	(0.337)	(0.345)	(0.058)	(0.116)
more6h	0.304	0.287	0.288	0.234	0.339	0.065	−0.178
	(0.428)	(0.425)	(0.425)	(0.440)	(0.481)	(0.065)	(0.143)
NotConcerned	−0.124	−0.113	−0.133	−0.096	−0.133	−0.021	−0.070
	(0.239)	(0.238)	(0.238)	(0.244)	(0.247)	(0.041)	(0.081)
Sometimes	−0.306	−0.346	−0.384	−0.196	0.055	−0.030	0.261**
	(0.419)	(0.420)	(0.423)	(0.444)	(0.456)	(0.041)	(0.116)
Often	0.352	0.167	0.079	−0.293	0.030	−0.009	0.574***
	(0.706)	(0.733)	(0.740)	(0.750)	(0.643)	(0.064)	(0.206)
Everyday	−0.097	−0.399	−0.523	−0.673	−0.370	−0.041	0.408*
	(0.929)	(0.884)	(0.850)	(0.856)	(0.828)	(0.064)	(0.231)
Sometimes	−0.487	−0.435	−0.420	−0.393	−0.406	−0.101**	0.036
	(0.297)	(0.292)	(0.295)	(0.303)	(0.318)	(0.049)	(0.088)
Often	−0.762**	−0.668*	−0.674*	−0.644*	−0.719*	−0.156***	0.061
	(0.356)	(0.350)	(0.359)	(0.372)	(0.395)	(0.054)	(0.104)
Always	−0.077	0.122	0.097	0.196	0.137	−0.040	0.158
	(0.496)	(0.488)	(0.507)	(0.491)	(0.510)	(0.070)	(0.135)
Log Lik.	−765.69	−741.48	−733.59	−714.66	−652.94	.	−4,220.82
N obs.	2,572	2,572	2,572	2,572	2,572	4,998	4,998
R p	0.82	0.82	0.82	0.83	0.84	.	0.44
Individual's controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	Yes	Yes	Yes	Yes	Yes
Employment controls	No	No	No	Yes	Yes	Yes	Yes
Working conditions	No	No	No	No	Yes	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; robust standard errors in parenthesis
Models controls are thus reported in the main tables

Table A30 – Estimates of the variables included in the CP index$_{t-1}$ on health

	Health$_t$						
	Fixed effects					HT	Pooled
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
SmtsComp	−0.027	−0.018	0.001	0.044	0.056	0.003	−0.091
	(0.225)	(0.226)	(0.225)	(0.227)	(0.223)	(0.048)	(0.083)
SmtsUncomp	−0.192	−0.145	−0.160	−0.137	−0.116	−0.049	−0.109
	(0.234)	(0.228)	(0.227)	(0.228)	(0.235)	(0.052)	(0.099)
Oft_EvDComp	−0.289	−0.279	−0.259	−0.249	−0.344	−0.091	−0.110
	(0.309)	(0.309)	(0.307)	(0.312)	(0.325)	(0.070)	(0.118)
Oft_EvDUncomp	−0.005	0.050	0.039	0.049	0.001	−0.015	0.037
	(0.281)	(0.277)	(0.274)	(0.277)	(0.279)	(0.067)	(0.115)
Mean	0.129	0.122	0.136	0.199	0.214	0.009	−0.052
	(0.255)	(0.255)	(0.256)	(0.256)	(0.259)	(0.052)	(0.108)
1to3h	−0.003	−0.037	−0.046	−0.030	0.001	−0.028	−0.106
	(0.189)	(0.189)	(0.190)	(0.192)	(0.195)	(0.043)	(0.075)
4to6h	−0.309	−0.330	−0.323	−0.273	−0.327	−0.116*	−0.044
	(0.273)	(0.273)	(0.273)	(0.281)	(0.288)	(0.059)	(0.106)
more6h	−0.607*	−0.598*	−0.612*	−0.640*	−0.805**	−0.166**	0.013
	(0.334)	(0.337)	(0.337)	(0.351)	(0.359)	(0.078)	(0.124)
NotConcerned	−0.029	−0.035	−0.043	−0.085	−0.145	0.007	−0.086
	(0.211)	(0.210)	(0.210)	(0.212)	(0.214)	(0.043)	(0.079)
Sometimes	−0.412	−0.424	−0.447	−0.456	−0.529*	−0.074	0.083
	(0.282)	(0.277)	(0.278)	(0.282)	(0.281)	(0.066)	(0.097)
Often	−0.057	−0.109	−0.077	−0.204	−0.269	−0.078	0.050
	(0.394)	(0.384)	(0.378)	(0.382)	(0.407)	(0.097)	(0.145)
Everyday	−0.444	−0.458	−0.412	−0.420	−0.500	−0.119	0.160
	(0.447)	(0.451)	(0.436)	(0.431)	(0.453)	(0.120)	(0.195)
Sometimes	0.009	0.040	0.023	0.041	0.096	−0.002	−0.305***
	(0.213)	(0.213)	(0.217)	(0.216)	(0.227)	(0.053)	(0.086)
Often	0.246	0.267	0.231	0.221	0.199	0.030	−0.332***
	(0.265)	(0.264)	(0.265)	(0.266)	(0.270)	(0.064)	(0.104)
Always	0.607*	0.577	0.473	0.474	0.521	0.068	−0.378***
	(0.361)	(0.360)	(0.363)	(0.365)	(0.376)	(0.090)	(0.141)
Log Lik.	−1,430.42	−1,420.01	−1,415.04	−1,402.87	−1,323.60	.	−5,123.32
N obs.	14,570	14,570	14,570	14,570	14,570	4,998	4,998
R p	0.72	0.72	0.72	0.72	0.74	.	0.10
Individual's controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	Yes	Yes	Yes	Yes	Yes
Employment controls	No	No	No	Yes	Yes	Yes	Yes
Working conditions	No	No	No	No	Yes	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; robust standard errors in parenthesis
Models controls are thus reported in the main tables

Table A31 – Estimates of the variables included in the CP index$_{t-1}$ on WHO5 score

	log(SA days)$_t$						
	Fixed effects					HT	Pooled
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
SmtsComp	0.019	0.018	0.007	0.011	0.025	0.027	0.012
	(0.111)	(0.110)	(0.110)	(0.110)	(0.110)	(0.111)	(0.054)
SmtsUncomp	0.084	0.080	0.073	0.088	0.097	0.097	−0.009
	(0.117)	(0.117)	(0.117)	(0.118)	(0.118)	(0.119)	(0.062)
Oft_EvDComp	0.042	0.051	0.038	0.036	0.048	0.033	0.045
	(0.159)	(0.160)	(0.158)	(0.158)	(0.158)	(0.158)	(0.075)
Oft_EvDUncomp	0.236*	0.236*	0.242*	0.230	0.278**	0.279**	−0.050
	(0.142)	(0.142)	(0.142)	(0.142)	(0.141)	(0.142)	(0.073)
Mean	0.111	0.111	0.112	0.110	0.123	0.120	0.064
	(0.125)	(0.125)	(0.126)	(0.125)	(0.126)	(0.127)	(0.071)
1to3h	0.099	0.101	0.111	0.107	0.113	0.116	0.048
	(0.100)	(0.099)	(0.099)	(0.100)	(0.100)	(0.101)	(0.047)
4to6h	0.166	0.171	0.176	0.174	0.188	0.187	0.010
	(0.132)	(0.132)	(0.133)	(0.133)	(0.132)	(0.133)	(0.068)
more6h	0.513***	0.505***	0.520***	0.535***	0.526***	0.525***	0.007
	(0.166)	(0.166)	(0.167)	(0.171)	(0.171)	(0.172)	(0.074)
NotConcerned	0.175*	0.181*	0.169*	0.168*	0.173*	0.170*	0.126**
	(0.103)	(0.103)	(0.101)	(0.101)	(0.101)	(0.102)	(0.055)
Sometimes	−0.031	−0.032	−0.053	−0.049	−0.088	−0.087	0.042
	(0.125)	(0.124)	(0.123)	(0.122)	(0.124)	(0.126)	(0.060)
Often	0.068	0.072	0.033	0.058	0.023	0.019	−0.004
	(0.205)	(0.205)	(0.206)	(0.206)	(0.211)	(0.212)	(0.088)
Everyday	0.037	0.048	−0.013	−0.013	−0.041	−0.034	−0.150
	(0.228)	(0.225)	(0.220)	(0.219)	(0.224)	(0.225)	(0.105)
Sometimes	−0.027	−0.027	−0.024	−0.038	−0.017	−0.011	−0.010
	(0.109)	(0.108)	(0.109)	(0.110)	(0.111)	(0.112)	(0.055)
Often	−0.113	−0.113	−0.104	−0.121	−0.109	−0.104	−0.008
	(0.139)	(0.138)	(0.138)	(0.139)	(0.140)	(0.142)	(0.066)
Always	−0.126	−0.128	−0.082	−0.096	−0.079	−0.081	0.057
	(0.182)	(0.181)	(0.181)	(0.181)	(0.183)	(0.183)	(0.089)
Log Lik.	−5,068.34	−5,054.92	−5,030.38	−5,012.75	−4,955.30	.	−8,227.62
N obs.	4,988	4,988	4,988	4,988	4,988	4,988	4,988
Rsq	0.02	0.02	0.03	0.04	0.06	.	0.09
R adj.	0.01	0.02	0.02	0.03	0.05	.	0.07
Individual's controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	Yes	Yes	Yes	Yes	Yes
Employment controls	No	No	No	Yes	Yes	Yes	Yes
Working conditions	No	No	No	No	Yes	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; robust standard errors in parenthesis
Models controls are thus reported in the main tables

Table A32 – Estimates of the variables included in the CP index$_{t-1}$ on WHO5 score

	WHO5 score$_t$						
	Fixed effects					HT	Pooled
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
SmtsComp	0.024	0.015	0.078	0.083	0.244	0.027	0.299*
	(0.290)	(0.290)	(0.288)	(0.287)	(0.272)	(0.159)	(0.157)
SmtsUncomp	0.181	0.178	0.194	0.170	0.203	0.074	0.159
	(0.337)	(0.337)	(0.334)	(0.330)	(0.319)	(0.191)	(0.196)
Oft_EvDComp	−0.054	−0.017	0.046	0.063	0.106	0.203	0.686***
	(0.431)	(0.432)	(0.432)	(0.428)	(0.407)	(0.244)	(0.228)
Oft_EvDUncomp	−0.254	−0.255	−0.260	−0.273	−0.407	−0.149	0.110
	(0.413)	(0.412)	(0.408)	(0.406)	(0.380)	(0.238)	(0.228)
Mean	0.078	0.071	0.050	0.020	−0.007	0.027	0.122
	(0.328)	(0.326)	(0.322)	(0.322)	(0.314)	(0.222)	(0.220)
1to3h	−0.323	−0.333	−0.386	−0.385	−0.387	−0.014	0.219
	(0.276)	(0.276)	(0.274)	(0.276)	(0.265)	(0.155)	(0.148)
4to6h	0.156	0.185	0.164	0.168	0.159	0.266	0.317
	(0.374)	(0.373)	(0.370)	(0.373)	(0.366)	(0.231)	(0.230)
more6h	−0.037	−0.057	−0.134	−0.120	−0.081	0.310	0.632***
	(0.433)	(0.430)	(0.432)	(0.437)	(0.435)	(0.237)	(0.229)
NotConcerned	0.201	0.196	0.236	0.236	0.097	0.276*	0.269*
	(0.243)	(0.243)	(0.246)	(0.246)	(0.237)	(0.160)	(0.156)
Sometimes	0.094	0.087	0.137	0.082	0.200	0.235	0.266
	(0.377)	(0.374)	(0.365)	(0.363)	(0.341)	(0.191)	(0.191)
Often	−0.143	−0.149	0.007	0.014	−0.062	0.038	0.693**
	(0.566)	(0.569)	(0.555)	(0.553)	(0.559)	(0.320)	(0.308)
Everyday	0.311	0.345	0.569	0.660	0.708	−0.129	0.235
	(0.893)	(0.893)	(0.870)	(0.866)	(0.854)	(0.398)	(0.371)
Sometimes	−0.099	−0.075	−0.104	−0.132	−0.188	0.152	−0.056
	(0.291)	(0.294)	(0.294)	(0.294)	(0.278)	(0.155)	(0.151)
Often	0.118	0.134	0.064	0.042	−0.106	0.274	−0.675***
	(0.352)	(0.352)	(0.349)	(0.348)	(0.339)	(0.197)	(0.188)
Always	0.264	0.265	0.035	−0.031	−0.396	0.504*	−1.490***
	(0.540)	(0.539)	(0.528)	(0.525)	(0.505)	(0.300)	(0.286)
Log Lik.	−10,122.82	−10,114.85	−10,075.43	−10,062.06	−9,866.54	.	−28,927.94
N obs.	4,998	4,998	4,998	4,998	4,998	10,168	10,168
Rsq	0.02	0.02	0.04	0.04	0.11	.	0.29
R adj.	0.01	0.01	0.03	0.03	0.10	.	0.28
Individual's controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Health controls	No	No	Yes	Yes	Yes	Yes	Yes
Employment controls	No	No	No	Yes	Yes	Yes	Yes
Working conditions	No	No	No	No	Yes	Yes	Yes

Notes: *p<0.1; **p<0.05; ***p<0.01 ; robust standard errors in parenthesis
Models controls are thus reported in the main tables

General conclusion

1 Main contributions

Although research on presenteeism first began more than twenty years ago, it's still characterised by the difficulty of coming up with a single, adequate definition of this behaviour, as well as a clear research framework (Karanika-Murray et al., 2021). This thesis therefore sets out to examine two forms of presenteeism: sickness presenteeism on the one hand, and competitive presenteeism on the other. Presenteeism behaviour is also considered here as an effort behaviour and not as a passive presence at work. The latter ultimately associates presenteeism with a cheating behaviour, linked to the moral hazard issue related to the unobservability of effort, whose empirical evaluation remains complex.

This proposal to widen the scope of the study by redefining presenteeism behaviour represents a new perspective that has not been addressed in existing literature reviews or meta-analyses. The proposal for a general definition based on several types of presenteeism therefore already contributes to the current debate on the scope of the behaviour to be studied. For example, the latest literature review by Patel et al. (2023) highlights the fact that sickness presenteeism is not always a harmful behaviour, both for the employee in terms of health and for the employer in terms of productivity. Thus, while for several years the decrease in productivity was an inherent part of the definition of presenteeism, current research tends to broaden its analytical perspective, at least in this regard.

This thesis contributes to empirical literature in several ways. First, from an academic perspective, this thesis attempts to estimate the effects studied as accurately as possible by investigating new methods based on the use of copulas. The first chapter, for example, shows the importance of considering non-random sample selection bias when studying behaviours associated with illness. Indeed, our results highlight that the conclusions we would have reached without taking endogenous participation into account would have been very different. Some factors that were identified as being associated with sickness absenteeism and sickness presenteeism would ultimately only explain the deterioration in health. Thus, addressing these factors would limit their harmful effects on health and reduce the likelihood of workers having to take sick leave or come to work while ill.

The second chapter focused on the issue of endogeneity of both sickness absence and sickness presenteeism in estimating their lagged effects on future health. To our knowledge, the effects of endogeneity in this literature had never been explored before. The intuition that sickness absence, unlike sickness presenteeism, would be beneficial to health had never been observed in longitudinal studies dealing with both behaviours at the same time. Taking into account unobservable factors, which explain both absenteeism and presenteeism behaviours and future health variables, is supposed to capture individual sensitivity to illness. In other words, the results of previous literature do not allow us to observe the specific effects of sickness absence and sickness presenteeism, but rather estimate the extent to which these behaviours are indicators of poor health. On the contrary, controlling for endogeneity in the models allows us to identify the specific effects of choosing sickness absence or presenteeism. Indeed, our results, based on copula model estimates, show that the additional regressors capturing the endogeneity associated with the duration of sickness absence and presenteeism are still significant and have the same sign. This tends to indicate that the unobservable factors are probably of the same nature. Given all the controls and the lack of information on the type and severity of illness episodes, it is very likely that these additional terms accurately capture individuals' underlying health. Furthermore, our results confirm the intuition that choosing to take time off when ill has a long-term restorative effect on individuals. On the other hand, a longer duration of sickness presenteeism tends to deteriorate health and increase the total duration of future illness.

Finally, the last chapter, which deals with competitive presenteeism, takes individual heterogeneity into account by using panel data to estimate fixed effects models. Here again, the conclusions based on pooled estimates would have been very different from those associated with estimates controlling for individual heterogeneity. Furthermore, this chapter highlights the robustness of the results with respect to endogeneity when the models take individual fixed effects into account. Indeed, unlike pooled models, the additional regressor capturing endogeneity in the fixed-effects models with copulas never appears to be significant. Furthermore, statistical tests for endogeneity associated with the instrumental variable approach are also insignificant in fixed-effects models, unlike in the pooled ones. In the case of

the latter chapter, several factors may explain the presence of endogeneity. This would be associated with intrinsic characteristics of individuals, such as individual ability and preferences for risk and competition. On the other hand, in addition to the factors already mentioned, a form of endogeneity that would explain both investment in over-effort and health could also come from an effect similar to that of the healthy worker effect. These unobservable factors, which are assumed to identify individuals in terms of their propensity to effort and resistance to illness, would be largely fixed over time. Thus, controlling for fixed effects would be sufficient to capture the endogeneity associated with these unobservable determinants. Of course, there remains the possibility that unobservable factors that vary over time could bias the estimates. These factors could, for example, correspond to an accumulation of human capital that would make over-effort easier over time. However, no model indicating the presence of endogeneity after controlling for individual heterogeneity supports this hypothesis.

The results regarding endogeneity in Chapters 2 and 3 tend to confirm the advantages of panel data and fixed-effects models, while also highlighting the need for endogeneity correction methods when working with cross-sectional data.

This thesis has therefore aimed to employ statistical methods that remain relatively unexplored in labour and health economics research. However, the use of copulas offers several significant advantages. On the one hand, copula methods allow for great flexibility in modelling the distributions of the variables studied. Secondly, they allow several distributions to be modelled jointly and a correlation coefficient for their relationship to be obtained. These first two characteristics, specific to copulas, appear to be a major asset for studies in health and labour. Data relating to these fields of study are very often summarised by discrete variables. For example, with regard to the measurement of absenteeism and presenteeism in this thesis, the use of durations in number of days leads to the estimation of distributions of count variables that follow specific statistical distributions such as Poisson, negative binomial or Poisson inverse Gaussian distributions. Furthermore, these distributions are generally characterised by an excess of zeros corresponding to two different data generation processes. Once again, copulas allow for great flexibility in estimating the parameters of these distributions. An additional advantage of using copulas is

the possibility of specifying continuous explanatory variables as splines, which allows for a more flexible functional form to be taken into account in order to capture the effect of one variable on another. Although the use of these splines carries a risk of overfitting, making it difficult to identify a trend in the estimated effect, it nevertheless provides an indication of the approximation achieved by introducing a quadratic parameter into the model. Last but not least, copulas are of particular interest when estimating a causal relationship between two variables when no instrumental variables are available in the data. The robustness and validity of the instruments used are always a matter of debate that is rarely resolved. As a result, copulas, which do not require any additional information in the data and are relatively flexible in their application, appear to be the preferred estimation method for exploring causal relationships or modelling relatively complex phenomena.

Finally, from an operational perspective, the findings of this thesis could be useful to human resources managers in developing more enlightened and appropriate policies. In this respect, the empirical results fall within the field of human resources economics, one of the objectives of which is to inform personnel management practices through the use of economic science tools. For example, as indicated above, our results highlight the characteristics that should be targeted in order to limit the probability of illness. In addition, with regard to absence and attendance incentive policies, our results indicate the different employment characteristics and working conditions that are associated with both sickness absence and sickness presenteeism, and those that are associated with only one of these behaviours. In this way, it is possible, for example, taking into account the objectives and costs, to reduce the duration of sickness absence without encouraging sickness presenteeism.

Furthermore, while the first chapter provides the keys to taking joint or exclusive action on sickness absenteeism and presenteeism, the second chapter highlights the need for intertemporal trade-offs. More specifically, in a context where firms generally seek to reduce their level of absence, the results tend to show that, on the contrary, from a long-term perspective, it could be beneficial to encourage absence in order to avoid encouraging sickness presenteeism and lengthening the future duration of absence. Although an incentive mechanism to encourage attendance may be beneficial in the short term, the costs of past presenteeism may offset the short-

term benefits of this incentive mechanism in the long-term. The findings of these two chapters thus provide complementary results that contribute to a more comprehensive understanding of the mechanisms associated with sickness presenteeism and sickness absenteeism behaviours.

The last chapter provides guidance on the need to strike the balance between incentives to work hard and the deterioration of workers' health. Here again, an intertemporal trade-off between the short-term benefits of incentive mechanisms and their long-term costs is necessary. While the use of deferred compensation mechanisms would take into account the benefits of incentives and the costs associated with promotion, these must also take into account the fact that they can be detrimental to workers' health in the medium to long-term. In this way, firms must put in place the means to regulate the workers level of investment, in order to prevent it from becoming excessive and detrimental to their health.

2 Limitations and future research

This thesis also faced a number of limitations. First, although the data available for the analyses carried out is rich in information, it does not allow for the study of the dynamics between decisions to be absent or present on a day-to-day basis when workers are faced with an illness shock. Future research should focus more on these dynamics, which require specific survey data that is difficult to obtain but would enable more in-depth and detailed studies. Obtaining daily data on illness and the choices made by workers would probably require a field experiment in a private company or public institution. Useful data for studying the decision between absence and attendance should also include the nature of the illness that led the worker to choose between the two. Such information, in addition to that included in existing surveys, would be a significant contribution to research.

Furthermore, a important gap in the current data is the lack of variables that can be used as instruments in studies. The results of this thesis focus exclusively on data from the French survey on working conditions and psychosocial risks. Unfortunately, this survey does not allow us to use the information provided as an instrument for measuring sickness presenteeism or competitive presenteeism. Therefore, in order

for research to use instrumental variable methods to identify causal relationships, information that can be used as instruments would need to be included. This could be related to past health and employment events, but also survey questions that would reveal individuals' risk preferences, such as those used in experimental studies. Finally, as indicated in the last chapter, information on activities outside employment would make it possible to identify individuals' preferences more accurately, such as their sporting or creative activities. Finally, with regard to the data used, the various results obtained should be validated by examining surveys from other countries. For example, it would be interesting to validate the findings of this thesis using the European Working Conditions Surveys (EWCS) in order to explore potential heterogeneity between countries due to different legislation or different work cultures.

On the other hand, the models estimated in chapters 2 and 3 aim to test the past effects of both sickness and competitive presenteeism behaviours on health or career variables. However, the three-year interval between each wave of observations is likely to influence the results presented. While the career-related results are assumed to be robust given the time required to build a career, both in terms of wage growth and promotion, the health-related results may be impacted by a large number of factors during these three years. Given the relative heterogeneity in individuals' vulnerability to illness, the types of illnesses individuals face, and organisational changes in employment, it is possible that sickness presenteeism and absenteeism may not have similar long-term effects. For example, it is possible that absenteeism has short-term effects, while presenteeism has longer-term effects. Furthermore, this depends on the duration of investment in each behaviour. Future research should therefore focus on studying the short, medium and long-term effects of absence, sickness presenteeism and competitive presenteeism on future health.

In chapters 1 and 2, the study of sickness absence and sickness presenteeism reveals different behavioural profiles. Some workers, when they are ill, will exclusively choose to be absent, while others will exclusively choose to be present. Finally, a third group of individuals will sometimes choose to be absent and sometimes to be present when ill. However, these two chapters do not explore in detail the characteristics specific to these different behavioural profiles. It might be relevant to

conduct a comparative study of these different behavioural profiles in terms of the effect of an additional day of absence or presence while ill. In addition, it would be useful to study whether belonging to these groups of individuals can be explained more by their employment characteristics, which are more conducive to absence or, on the contrary, too restrictive, or by individual and health characteristics and the duration of the illness.

While the first chapter controls for the endogenous participation bias associated with illness, it does not take into account the endogeneity between all the explanatory variables of employment or working conditions and the duration of sickness absence and presenteeism. Unfortunately, our methodology does not allow us to do this. However, as copula-based methods are still in development, it would be beneficial to use methods such as that of Yang et al. (2024) in conjunction with the issue of non-random sample selection raised by Marra & Wyszynski (2016). For instance, taking sample selection into account would make it possible to assess the signalling effect of presenteeism on career progression, as in the analysis conducted in the third chapter on competitive presenteeism. Without taking this into account, it is likely that the estimated effect primarily reflects the impact of poor health on career trajectory. While it can be assumed that sickness presenteeism is potentially beneficial to one's career, the signal associated with poor health would, on the contrary, be detrimental to it (Chadi & Goerke, 2018).

Although this thesis provides guidance for human resources policies, these policies are mainly based on cost-benefit analysis. Future research should therefore focus more on the costs of presenteeism, both illness-related and competitive. From this perspective, the two types of behaviour are known to have relatively different costs. It would be useful to quantify the cost of a day of presenteeism and that of a day of absence. It would then be necessary to take into account a certain degree of heterogeneity among firms, particularly in terms of their production function. In order to examine the costs of absence and presence in case of illness, production technology makes it possible to identify the cost structure more concretely. To this end, the use of linked employer-employee data would be a significant asset and would make a real contribution to the literature on presenteeism. This would provide information on the organisation of the firm, the incentives and policies put in place, as well as

declarative information from employees, such as that used to date. In addition, it could enable managers and company executives to be questioned about their perception of the costs associated with absence and presence. This would provide an overview of the mechanisms involved between the incentives put in place by the employer and the response of employees in terms of behaviour. The surveys used here were originally designed to enable employer-employee matching, but the matching rate observed between employees and employers was found to be unsatisfactory.

Finally, Lohaus & Habermann (2019) explain that researchers have "generally failed to explicitly describe the decision-making process that determines how absenteeism or presenteeism is chosen". Although each chapter presents a theoretical insight, no theoretical model with testable hypotheses is proposed. The problem of theorising presenteeism behaviour is not recent. Böckerman (2018) already mentioned this gap in economics. Indeed, attempts to theorise this behaviour are more commonly made in the field of social psychology. Furthermore, the theoretical models proposed do not constitute a theorisation of presenteeism. They most commonly include sickness presenteeism behaviour within the framework of models on absenteeism. It is therefore essential to be able to propose a framework based on a theoretical approach that describes the mechanisms of presenteeism in its own right. This can be done, for example, within the framework of contract theory, tournament theory or signalling theory. The contractual approach is the most widely used in the context of employment relationships, in theoretical analyses of absenteeism (Coles & Treble, 1993; Treble & Barmby, 2011) and also of sickness presenteeism (Chatterji & Tilley, 2002; Brown & Sessions, 2004; Hirsch et al., 2017). Furthermore, Johns (2010), who focuses specifically on sickness presenteeism, explains that the theory underpinning this behaviour should take into account its relationship with sickness absenteeism. However, as presenteeism is not necessarily associated with illness, the theoretical framework should cover presenteeism behaviour in a broad sense and then in a specific one in relation to illness. Finally, like shirking related to absenteeism, cheating on presenteeism appears difficult to identify empirically in the data. Therefore, this type of behaviour should be studied more theoretically, especially since it is a common moral hazard issue in personnel economics (Kahana & Weiss, 1992).

Bibliography

Aboagye, E., Björklund, C., Gustafsson, K., Hagberg, J., Aronsson, G., Marklund, S., ... Bergström, G. (2019). Exhaustion and impaired work performance in the workplace: Associations with presenteeism and absenteeism. *Journal of Occupational and Environmental Medicine*, *61*, 1.

Ahrens, H., & Pincus, R. (1981). On two measures of unbalancedness in a one-way model and their relation to efficiency. *Biometrical Journal*, *23*(3), 227-235.

Akerstedt, T., Kecklund, G., Alfredsson, L., & Selen, J. (2007). Predicting long-term sickness absence from sleep and fatigue. *Journal of Sleep Research*, *16*(4), 341-345.

Allen, S. (1981). An empirical model of work attendance. *The Review of Economics and Statistics*, *63*(1), 77-87.

Amegashie, J. A., Cadsby, C. B., & Song, Y. (2007). Competitive burnout: Theory and experimental evidence. *Games and Economic Behavior*, *59*(2), 213-239.

Amilon, A., & Wallette, M. (2009). Work absence—a signalling factor for temporary workers? *Labour*, *23*, 171-194.

Andelic, N., Allan, J., Bender, K., Powell, D., & I.Theodossiou. (2024). Performance-related pay, mental and physiological health. *Industrial Relations: A Journal of Economy and Society*, *63*(1), 3–25.

Anger, S. (2008). Overtime work as a signaling device. *Scottish Journal of Political Economy*, *55*(2), 167–189.

Arnold, D. (2016). Determinants of the annual duration of sickness presenteeism: Empirical evidence from european data. *Labour*, *30*(2), 198–212.

Arnold, D., & de Pinto, M. (2015). How are work-related characteristics linked to sickness absence and presenteeism? theory and data. *Schmollers Jahrbuch*, *135*, 465-498.

Aronsson, G., Gustafsson, K., & Dallner, M. (2000). Sick but yet at work. an empirical study of sickness presenteeism. *J Epidemiol Community Health*, *54*, 502–509.

Aronsson, G., Gustafsson, K., & Mellner, C. (2011). Sickness presence, sickness absence, and self-reported health and symptoms. *International Journal of Workplace Health Management*, *4*, 228 - 243.

Ashby, K., & Mahdon, M. (2010). Why do employees come to work when ill: An investigation into sickness presence in the workplace. *London: The Work Foundation*.

Avgoustaki, A., & Cañibano, A. (2020). Motivational drivers of extensive work effort: Are long hours always detrimental to well-being? *Industrial Relations: A Journal of Economy and Society*, *59*(3), 355-398.

Avgoustaki, A., & Frankort, H. T. W. (2019). Implications of work effort and discretion for employee well-being and career-related outcomes: An integrative assessment. *ILR Review*, *72*(3), 636-661.

Bakker, A. B., & Demerouti, E. (2007). The job demands-resources model: state of the art. *Journal of Managerial Psychology*, *22*(3), 309–328.

Bakker, A. B., Demerouti, E., & Sanz-Vergel, A. (2023). Job demands–resources theory: Ten years later [Journal Article]. *Annual Review of Organizational Psychology and Organizational Behavior*, *10*(Volume 10, 2023), 25-53.

Bakker, A. B., Demerouti, E., & Sanz-Vergel, A. I. (2014). Burnout and work engagement: The jd–r approach. *The Annual Review of Organizational Psychology and Organizational Behavior*, *1*, 389–411.

Baktash, M. B., Heywood, J. S., & Jirjahn, U. (2022). Worker stress and performance pay: German survey evidence. *Journal of Economic Behavior & Organization*, *201*, 276-291.

Baptista, M. C., Burton, W., Nahas, A., Wang, Y., Viana, M. C., & Andrade, L. H. (2019). Absenteeism and presenteeism associated with common health conditions in brazilian workers. *Journal of occupational and environmental medicine*, *61*, 303-313.

Barmby, T., & Larguem, M. (2009). Coughs and sneezes spread diseases: An empirical study of absenteeism and infectious illness. *Journal of Health Economics*, *28*(5), 1012-1017.

Barmby, T., Orme, C., & Treble, J. (1991). Worker absenteeism: An analysis using microdata. *The Economic Journal*, *101*, 214-229.

Barnay, T. (2016). Health, work and working conditions: a review of the european economic literature. *The European Journal of Health Economics*, *17*, 693–709.

Barnay, T. (2020). Interactions of work and health: An economic perspective. In T. Theorell (Ed.), *Handbook of socioeconomic determinants of occupational health: From macro-level to micro-level evidence* (pp. 37–52). Cham : Springer International Publishing.

Baum, C. F., Schaffer, M. E., & Stillman, S. (2003). Instrumental variables and gmm: Estimation and testing. *The Stata Journal*, *3*(1), 1-31.

Beemsterboer, W., Stewart, R., Groothoff, J., & Nijhuis, F. (2009). A literature review on sick leave determinants (1984–2004). *International Journal of Occupational Medicine and Environmental Health*, *22*(2), 169–179.

Bell, D., Hart, R., Hübler, O., & Schwerdt, W. (2000). Paid and unpaid overtime work in germany and the uk. *SSRN Electronic Journal*.

Ben Halima, M. A., Greenan, N., & Lanfranchi, J. (2023). Getting sick for profit? the impact of cumulative ict and management changes on long term sickness absence. *Journal of Economic Behavior and Organization*, *212*, 659-688.

Berg, N., Kiviruusu, O., Karvonen, S., Rahkonen, O., & Huurre, T. (2017). Pathways from problems in adolescent family relationships to midlife mental health via early adulthood disadvantages–a 26-year longitudinal study. *PLos One*, *12*(5), 460-468.

Bergström, G., Bodin, L., Hagberg, J., Aronsson, G., & Josephson, M. (2009b). Sickness presenteeism today, sickness absenteeism tomorrow? a prospective study on sickness presenteeism and future sickness absenteeism. *Journal of Occupational and Environmental Medicine*, *51*(6), 629–638.

Bergström, G., Bodin, L., Hagberg, J., Lindh, T., Aronsson, G., & Josephson, M. (2009a). Does sickness presenteeism have an impact on future general health? *International archives of occupational and environmental health*, *82*(10), 1179–1190.

Berlinski, S., & Ramos, A. (2020). Peer effects in the decision to apply for a professional excellence award. *Labour Economics*, *67*, 101934.

Besley, T., & Ghatak, M. (2005). Competition and incentives with motivated agents. *American Economic Review*, *95*(3), 616-636.

Betz, T., Cook, S. J., & Hollenbach, F. M. (2018). On the use and abuse of spatial instruments. *Political Analysis*, *26*(4), 474–479.

Bilanakos, C., Heywood, J. S., Sessions, J. G., & Theodoropoulos, N. (2018). Does delegation increase worker training? *Economic Inquiry*, *56*(2), 1089-1115.

Biron, C., Karanika-Murray, M., & Ivers, H. (2022). The health-performance framework of presenteeism: A proof-of-concept study. *Frontiers in Psychology*, *Volume 13 - 2022*.

Booth, A. L., Francesconi, M., & Frank, J. (2003). A sticky floors model of promotion, pay, and gender. *European Economic Review*, *47*(2), 295-322.

Bramoullé, Y., Djebbari, H., & Fortin, B. (2020). Peer effects in networks: A survey. *Annual Review of Economics*, *12*(Volume 12, 2020), 603-629.

Bratti, M., & Miranda, A. (2011). Endogenous treatment effects for count data models with endogenous participation or sample selection. *Health Economics*, *20*, 1090-1109.

Bratti, M., & Staffolani, S. (2007). Effort-based career opportunities and working time. *International Journal of Manpower*, *28*, 489-512.

Breitung, J., Mayer, A., & Wied, D. (2024). Asymptotic properties of endogeneity corrections using nonlinear transformations. *The Econometrics Journal*, *27*(3), 362-383.

Brown, S., & Sessions, J. G. (2004). Absenteeism, 'presenteeism', and shirking. *Economic Issues*, *9*(1), 15–23.

Budidharma, I., & Wahyuni, H. (2017). Does better physical function make people happier? *Asian Social Science*, *13*(3), 58–69.

Bénabou, R., & Tirole, J. (2003). Intrinsic and extrinsic motivation. *Review of Economic Studies*, *70*, 489–520.

Böckerman, P. (2018). Presenteeism at work. In C. L. Cooper & L. Lu (Eds.), (pp. 166–180). Cambridge University Press.

Böckerman, P., Bryson, A., & Ilmakunnas, P. (2012). Does high involvement management improve worker wellbeing? *Journal of Economic Behavior & Organization*, *84*(2), 660–680.

Böckerman, P., Bryson, A., Kauhanen, A., & Kangasniemi, M. (2020). Does job design make workers happy? *Scottish Journal of Political Economy*, *67*(1), 31-52.

Cameron, A. C., Li, T., Trivedi, P. K., & Zimmer, D. M. (2004). Modelling the differences in counted outcomes using bivariate copula models with application to mismeasured counts. *The Econometrics Journal*, *7*(2), 566–584.

Campbell, D., & Green, F. (2002). The long term pay-off from working longer hours. *Studies in Economics 0205, School of Economics, University of Kent*.

Cappelli, P., & Neumark, D. (2001). Do "high-performance" work practices improve establishment-level outcomes? *Industrial and Labor Relations Review*, *54*(4), 737–775.

Card, D. (1999). *The causal effect of education on earnings* (1re éd., Vols. 3, Part A; O. Ashenfelter & D. Card, Eds.). Elsevier.

Caverley, N., Cunningham, J. B., & MacGregor, J. N. (2007). Sickness presenteeism, sickness absenteeism, and health following restructuring in a public service organization. *Journal of Management Studies*, *44*(2), 304–319.

Chadi, A., & Goerke, L. (2018). Missing at work – Sickness-related absence and subsequent career events. *Journal of Economic Behavior & Organization*, *153*(C), 153-176.

Chatterji, M., & Tilley, C. J. (2002). Sickness, absenteeism, presenteeism, and sick pay. *Oxford Economic Papers*, *54*, 669–687.

Chou, C.-Y., & Mach, M. (2021). Unlocking the contradictory outcomes of presenteeism through a temporal model: Effort exertion as a mediator. *Frontiers in Psychology*, *Volume 12 - 2021*.

Clark, C., Caldwell, T., Power, C., & Stansfeld, S. A. (2010). Does the influence of childhood adversity on psychopathology persist across the lifecourse? a 45-year prospective epidemiologic study. *Annals of epidemiology*, *20*(5), 385-394.

Coles, M., Lanfranchi, J., Skalli, A., & Treble, J. (2007). Pay, technology, and the cost of worker absence. *Economic Inquiry*, *45*, 268–285.

Coles, M., & Treble, J. (1993). The price of worker reliability. *Economics Letters*, *41*, 149–155.

Collins, A. M., Cartwright, S., & Cowlishaw, S. (2017). Sickness presenteeism and sickness absence over time: A uk employee perspective. *Work & Stress*, 1-16.

Conway, P. M., Hogh, A., Rugulies, R., & Åse Marie Hansen. (2014). Is sickness presenteeism a risk factor for depression?: A danish 2-year follow-up study. *Journal of Occupational and Environmental Medicine*, *56*(6), 595–603.

Cooper, C. L. (1996). Hot under the collar. *Times Higher Education Supplement*, 15.

Cornelissen, T., Dustmann, C., & Schönberg, U. (2017). Peer effects in the workplace. *The American Economic Review*, *107*(2), 425–456.

Cour des Comptes. (2024). *Communiqué de presse. arrêt de travail pour maladie : Mise au point du premier président de la cour des comptes* (Rapport technique). Consulté sur https://www.ccomptes.fr/fr/documents/70036

Dellve, L., Hadzibajramovic, E., & Ahlborg, G. (2011). Work attendance among healthcare workers: Prevalence, incentives, and long-term consequences for health and performance. *Journal of advanced nursing*, *67*, 1918-29.

Demerouti, E., Le Blanc, P. M., Bakker, A. B., Schaufeli, W. B., & Hox, J. (2009). Present but sick: a three-wave study on job demands, presenteeism and burnout. *Career Development International*, *14*(1), 50–68.

Docherty, P., Forslin, J., Shani, A. B., & Kira, M. (2002). Emerging work systems: Creating sustainable work systems: Emerging perspectives and practice. Routledge: London UK; New York, NY, USA.

Dufour, J.-M. (2003). Identification, weak instruments, and statistical inference in econometrics. *Canadian Journal of Economics/Revue canadienne d'économique*, *36*(4), 767-808.

Eckert, C., & Hohberger, J. (2023). Addressing endogeneity without instrumental variables: An evaluation of the gaussian copula approach for management research. *Journal of Management*, *49*(4), 1460-1495.

Ehlert, A., & García-Morán, E. (2022). Workers' self-selection into public sector employment: A tale of absenteeism. *Kyklos*, *75*, 394-409.

Eisenberg, D., Golberstein, E., & Whitlock, J. L. (2014). Peer effects on risky behaviors: New evidence from college roommate assignments. *Journal of Health Economics*, *33*, 126-138.

Evans-Lacko, S., & Knapp, M. (2016). Global patterns of workplace productivity for people with depression: absenteeism and presenteeism costs across eight diverse countries. *Social Psychiatry and Psychiatric Epidemiology Journal*, *51*, 1525–1537.

Falk, A., & Ichino, A. (2006). Clean evidence on peer effects. *Journal of Labor Economics*, *24*(1), 39–57.

Francesconi, M. (2001). Determinants and consequences of promotions in britain. *Oxford Bulletin of Economics and Statistics*, *63*(3), 279-310.

Ganster, D., Rosen, C., & Fisher, G. (2018). Long working hours and well-being: What we know, what we do not know, and what we need to know. *Journal of Business and Psychology*, *33*, 25–39.

Garrow, V. (2016). Presenteeism: A review of current thinking. *Institute for Employment Studies, Report 507*.

Gerich, J. (2016). Determinants of presenteeism prevalence and propensity: Two sides of the same coin? *Archives of Environmental & Occupational Health*, *71*(4), 189–198.

Gosselin, E., Lemyre, L., & Corneil, W. (2013). Presenteeism and absenteeism: Differentiated understanding of related phenomena. *Journal of Occupational Health Psychology*, *18*(1), 75–86.

Graham, J. R., Grennan, J., Harvey, C. R., & Rajgopal, S. (2022). Corporate culture: Evidence from the field. *Journal of Financial Economics*, *146*(2), 552-593.

Green, C. P., Heywood, J. S., & Theodoropoulos, N. (2021). Hierarchy and the employer size effect on wages: Evidence from britain. *Economica*, *88*(351), 671-696.

Green, F., Felstead, A., Gallie, D., & Henseke, G. (2021). Working still harder. *ILR Review*, *75*(2), 458–487.

Green, F., Felstead, A., Gallie, D., & Inanc, H. (2016). Job-related well-being through the great recession. *Journal of Happiness Studies*, *17*, 389–411.

Greene, W. H. (1997). *FIML estimation of sample selection models for count data.* (Discussion Paper EC-97-02, Department of Economics, Stern School of Business, New York University)

Gregg, P., Grout, P., Ratcliffe, A., Smith, S., & Windmeijer, F. (2011). How important is pro-social behaviour in the delivery of public services? *Journal of Public Economics*, *95*(7), 758–766.

Grinza, E., & Rycx, F. (2020). The impact of sickness absenteeism on firm productivity: New evidence from belgian matched employer-employee panel data. *Industrial Relations*, *59*(1), 150–194.

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy*, *80*(2), 223–255.

Gustafsson, K., & Marklund, S. (2011). Consequences of sickness presence and sickness absence on health and work ability: A swedish prospective cohort study. *International Journal of Occupational Medicine and Environmental Health*, *24*, 153–165.

Gustafsson, K., & Marklund, S. (2013). Associations between health and combinations of sickness presence and absence. *Occupational medicine (Oxford, England)*, *64*.

Haan, P., & Myck, M. (2009). Dynamics of health and labor market risks. *Journal of Health Economics*, *28*(6), 1116-1125.

Hansen, C. D., & Andersen, J. H. (2008). Going ill to work – what personal circumstances, attitudes and work-related factors are associated with sickness presenteeism? *Social Science & Medicine*, *67*, 956–964.

Hansen, C. D., & Andersen, J. H. (2009). Sick at work–a risk factor for long-term sickness absence at a later date? *J Epidemiol Community Health*, *65*, 397–402.

Harbaugh, W. T., & van Den Nouweland, A. (2005). Demonstrating worker quality through strategic absenteeism. *International Game Theory Review*, *23*, 189-209.

Haschka, R. E. (2022). Handling endogenous regressors using copulas: A generalization to linear panel models with fixed effects and correlated regressors. *Journal of Marketing Research*, *59*(4), 860-881.

Hasebe, T., & Vijverberg, W. P. (2012). *A Flexible Sample Selection Model: A GTL-Copula Approach* (IZA Discussion Papers N° 7003). Institute of Labor Economics (IZA).

Havet, N., & Plantier, M. (2023). The links between difficult working conditions and sickness absences in the case of french workers. *Labour*, *37*, 160–195.

Helgesson, M., Gustafsson, K., Marklund, S., Aronsson, G., & Leineweber, C. (2021). Sickness absence and sickness presence among health and care employees in sweden - health complaints, health behavior, and future long-term sickness absence. *Journal of occupational and environmental medicine*, *Publish Ahead of Print*.

Hemp, P. (2004). Presenteeism: At work – but out of it. *Harvard Business Review*, *82*(10), 49–58.

Herrmann, M. A., & Rockoff, J. E. (2012). Worker absence and productivity: Evidence from teaching. *Journal of Labor Economics*, *30*(4), 749-782.

Heywood, J. S., Jirjahn, U., & Wei, X. (2008). Teamwork, monitoring and absence. *Journal of Economic Behavior & Organization*, *68*, 676–690.

Hilbe, J. (2014). *Modelling count data* (1st éd.). New York: Cambridge University Press.

Hillion, M. (2024). L'effet des conditions de travail sur la santé et le recours aux soins. *DOCUMENT D'ÉTUDES N° 277, DARES*, 1–110.

Hillion, M., Mette, C., & Pisarik, J. (2021). L'effet des conditions de travail sur la santé et le recours aux soins. *Preliminary version*, 1–98.

Hinsch, D. M., Spanier, K., Radoschewski, F. M., & Bethge, M. (2019). Associations between overcommitment, effort–reward imbalance and mental health: findings from a longitudinal study. *Int Arch Occup Environ Health*, *92*, 559–567.

Hirsch, B., Lechmann, D. S. J., & Schnabel, C. (2017). Coming to work while sick: An economic theory of presenteeism with an application to german data. *Oxford Economic Papers*, *69*(4), 1010–1031.

Hofaidhllaoui, M., & Chhinzer, N. (2014). The relationship between satisfaction and turnover intentions for knowledge workers. *Engineering Management Journal*, *26*(2), 3–9.

Hoffmann, F., Inderst, R., & Opp, M. (2020). Only Time Will Tell: A Theory of Deferred Compensation. *The Review of Economic Studies*, *88*(3), 1253-1278.

Holmstrom, B. (1982). Moral hazard in teams. *The Bell Journal of Economics*, *13*(2), 324–340.

Hudson, H. L., Nigam, J. A., Sauter, S. L., Chosewood, L., Schill, A. L., & Howard, J. E. (2019). *Total worker health: Integrative approaches to safety, health, and well-being* (Rapport technique). American Psychological Association.

Ichino, A., & Riphahn, R. T. (2005). The effect of employment protection on worker effort: Absenteeism during and after probation. *Journal of the European Economic Association*, *3*(1), 120–143.

Janssens, H., Clays, E., De Clercq, B., De Backer, G., & Braeckman, L. (2013). The relation between presenteeism and different types of future sickness absence. *Journal of occupational health*, *55*.

Jensen, U. T., Andersen, L. B., & Holten, A.-L. (2019). Explaining a dark side: Public service motivation, presenteeism, and absenteeism. *Review of Public Personnel Administration*, *39*(4), 487-510.

Johansson, G., & Lundberg, I. (2004). Adjustment latitude and attendance requirements as determinants of sickness absence or attendance. empirical tests of the illness flexibility model. *Social Science & Medicine*, *58*, 1857–1868.

Johns, G. (2010). Presenteeism in the workplace: A review and research agenda. *Journal of Organizational Behavior*, *31*, 519–542.

Johns, G. (2011). Attendance dynamics at work: The antecedents and correlates of presenteeism, absenteeism, and productivity loss. *Journal of occupational health psychology*, *16*, 483-500.

Johnston, C. S., Maggiori, C., & Rossier, J. (2016). Professional trajectories, individual characteristics, and staying satisfied and healthy. *Journal of Career Development*, *43*(1), 81-98.

Kahana, N., & Weiss, A. (1992). Absenteeism: A comparison of incentives in alternative organizations. *Journal of Comparative Economics*, *16*, 573-595.

Karanika-Murray, M., Biron, C., Hervieux, V., Whysall, Z., & Chen, H. (2021). Managing presenteeism to optimize health and performance. In T. Wall, C. Cooper, & P. Brough (Eds.), *The sage handbook of organizational wellbeing* (pp. 232–247). London : SAGE.

Karanika-Murray, M., & Cooper, C. L. (2018). Presenteeism: an introduction to prevailing global phenomenon. *Cambridge Core*, 1–34.

Keane, M., & Neal, T. (2023). Instrument strength in iv estimation and inference: A guide to theory and practice. *Journal of Econometrics*, *235*(2), 1625-1653.

Kelly, E. L., & Moen, P. (2020). *Review of overload: How good jobs went bad and what we can do about it* (Rapport technique). Princeton University Press.

Kim, W., Kolb, J., & Kim, T. (2013). The relationship between work engagement and performance a review of empirical literature and a proposed research agenda. *Human Resource Development Review*, *12*, 248-276.

Kinman, G. (2019). Sickness presenteeism at work: prevalence, costs and management. *British Medical Bulletin*, 1–10.

Kiviet, J. F. (2020). Testing the impossible: Identifying exclusion restrictions. *Journal of Econometrics*, *218*(2), 294-316.

Kivimäki, M., Head, J., Ferrie, J. E., Hemingway, H., Shipley, M. J., Vahtera, J., & Marmot, M. G. (2005). Working while ill as a risk factor for serious coronary events: The whitehall ii study. *American Journal of Public Health*, *95*(1), 98–102.

Kline, R. B. (2011). *Principles and practice of structural equation modeling.* (5th ed., pp. 3–427). New York: The Guilford Press.

Kossek, E. E., & Lee, K.-H. (2017). Work-family conflict and work-life conflict. *Oxford University Press*.

Kreps, D. M., & Wilson, R. (1982). Reputation and imperfect information. *Journal of Economic Theory*, *27*(2), 253–279.

Lai, T., & Ng, T. (2014). The impact of product market competition on training provision: Evidence from canada. *The Canadian Journal of Economics / Revue canadienne d'Economique*, *47*(3), 856–888.

Landstedt, E., Hammarström, A., & Winefield, H. (2015). How well do parental and peer relationships in adolescence predict health in adulthood? *Scandinavian journal of public health*, *43*(5), e0178136.

Lanfranchi, J., & Treble, J. (2010). Just-in-time production, work organization and absence control. *The Manchester School*, *78*(5), 460–483.

Leineweber, C., Westerlund, H., Hagberg, J., Svedberg, P., & Alexanderson, K. (2012). Sickness presenteeism is more than an alternative to sickness absence: results from the population-based slosh study. *International Archives of Occupational and Environmental Health*.

Lepinteur, A. (2021). The asymmetric experience of gains and losses in job security on health. *Health Economics*, *30*, 2217–2229.

Lindquist, M., Sauermann, J., & Zenou, Y. (2022). Peer effects in the workplace: A network approach. *SSRN Electronic Journal*.

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

Lohaus, D., & Habermann, W. (2019). Presenteeism: A review and research directions. *Human Ressource Management Review*, *29*, 43–58.

Lohaus, D., & Habermann, W. (2021). Understanding the decision-making process between presenteeism and absenteeism. *Frontiers in Psychology*, *12*, 716925.

Lu, L., & Cooper, C. L. (2022). Sickness presenteeism as a link between long working hours and employees' outcomes: Intrinsic and extrinsic motivators as resources. *International Journal of Environmental Research and Public Health*, *19*(4).

Lu, L., Lin, H. Y., & Cooper, C. L. (2013). Unhealthy and present: Motives and consequences of the act of presenteeism among taiwanese employees. *Journal of Occupational Health Psychology*, *18*(4), 406–416.

López-Bueno, R., Clausen, T., Calatayud, J., Bláfoss, R., Vinstrup, J., & Andersen, L. L. (2021). Self-reported sickness absence and presenteeism as predictors of future disability pension: Cohort study with 11-year register follow-up. *Preventive Medicine*, *148*, 106565.

Ma, J., Meltzer, D. P., Yang, L.-Q., & Liu, C. (2018). Motivation and presenteeism: The whys and whats. In C. L. Cooper & L. Lu (Eds.), *Presenteeism at work* (p. 97–122). Cambridge University Press.

Maestas, N., Mullen, K., & Rennane, S. (2020). Absenteeism and presenteeism among american workers. *Journal of Disability Policy Studies*, *32*, 13-23.

Malmberg-Ceder, K., Vuorio, T., Korhonen, P., Kautiainen, H., Soinila, S., & Haanpää, M. (2020). The impact of self-reported recurrent headache on absenteeism and presenteeism at work among finnish municipal female employees. *Journal of Pain Research*, *13*, 2135-2142.

Marie, O., & Val Castello, J. (2023). Sick leave cuts and (unhealthy) returns to work. *Journal of Labor Economics*, *41*, 923–956.

Marra, G., Radice, R., & Zimmer, D. (2024). A unifying switching regime regression framework with applications in health economics. *Econometric Reviews*, *43*(1), 52-70.

Marra, G., & Wyszynski, K. (2016). Semi-parametric copula sample selection models for count responses. *Computational Statistics and Data Analysis*, *104*, 110-129.

Mastekaasa, A. (2020). Absenteeism in the public and the private sector: Does the public sector attract high absence employees? *Journal of Public Administration Research and Theory*, *30*(1), 60–76.

Meijman, T. F., & Mulder, G. (1998). Psychological aspects of workload. In P. J. D. Drenth & H. Thierry (Eds.), *A handbook of work and organizational psychology* (Vol. 2, p. 5-33). Hove, United Kingdom: Psychology Press.

Meyer, A., & Wallette, M. (2005). Absence of absenteeism and overtime work – signaling factors for temporary workers? *Working Papers, Lund University, Department of Economics*(2005:15), 1–27.

Miraglia, M., & Johns, G. (2016). Going to work ill: A meta-analysis of the correlates of presenteeism and a dual-path model. *Journal of Occupational Health Psychology*, *21*(3), 261–283.

Navarro, A., Salas-Nicás, S., Llorens, C., Moncada, S., Molinero-Ruíz, E., & na, D. M. (2019). Sickness presenteeism: Are we sure about what we are studying? a research based on a literature review and an empirical illustration. *American Journal of Industrial Medicine*, 1–10.

Niedhammer, I., Chastang, J. F., Sultan-Taïeb, H., Vermeylen, G., & Parent-Thirion, A. (2013). Psychosocial work factors and sickness absence in 31 countries in europe. *The European Journal of Public Health*, *23*, 622–629.

Nikolova, M., & Cnossen, F. (2020). What makes work meaningful and why economists should care about it. *Labour Economics*, *65*, Aricle 101847.

Norman, R. E., Byambaa, M., De, R., Butchart, A., Scott, J., & Vos, T. (2012). The long-term health consequences of child physical abuse, emotional abuse, and neglect: a systematic review and meta-analysis. *PLoS medicine*, *9*(11), e1001349.

Pannenberg, M. (2005). Long-term effects of unpaid overtime. *Scottish Journal of Political Economy*, *52*(2), 177-193.

Papagiannaki, E. (2014). Rising unpaid overtime: a critical approach to existing theories. *Int. J. Management Concepts and Philosophy*, *8*(1), 68–88.

Park, S., & Gupta, S. (2012). Handling endogenous regressors by joint estimation using copulas. *Marketing Science*, *31*(4), 567–586.

Patel, C., Biron, M., Cooper, S. C., & Budhwar, P. S. (2023). Sick and working: Current challenges and emerging directions for future presenteeism research. *Journal of Organizational Behavior*, *44*(6), 839-852.

Pauly, M. V., Nicholson, S., Polsky, D., Berger, M. L., & Sharda, C. (2008). Valuating reductions in on-the-job illness: 'presenteeism' from managerial and economic perspectives. *Health Economics*, *17*, 469–485.

Pichler, S., Wen, K., & Ziebarth, N. (2021). Positive health externalities of mandating paid sick leave. *Journal of Policy Analysis and Management*, *40*, 715–743.

Pichler, S., & Ziebarth, N. (2015). The pros and cons of sick pay schemes: Testing for contagious presenteeism and noncontagious absenteeism behavior. *Journal of Public Economics*, *156*, 14–33.

Plümper, T., & Troeger, V. E. (2019). Not so harmless after all: The fixed-effects model. *Political Analysis*, *27*(1), 21–45.

Preston, A. E. (1989). The nonprofit worker in a for-profit world. *Journal of Labor Economics*, *7*(4), 438–463.

Qian, Y., Koschmann, A., & Xie, H. (2024). *A practical guide to endogeneity correction using copulas* (Working Paper Nº 32231). National Bureau of Economic Research.

Qian, Y., & Xie, H. (2024). Correcting regressor-endogeneity bias via instrument-free joint estimation using semiparametric odds ratio models. *Journal of Marketing Research*, *61*(5), 914-936.

Ruhle, S., Breitsohl, H., Aboagye, E., Baba, V., Biron, C., Correia Leal, C., ... Muschalla, B. (2019). "to work, or not to work, that is the question" - recent trends and avenues for research on presenteeism. *European Journal of Work and Organizational Psychology*, *29*, 344-363.

Ryvkin, D. (2011). Fatigue in dynamic tournaments. *Journal of Economics & Management Strategy*, *20*(4), 1011–1041.

Sanatkar, S., Stevelink, S., Aars, N., Bardal, I., Brandseth, O., Brinchmann, B., & Mykletun, A. (2025). Health consequences of graded, full, and no sickness absence among workers with musculoskeletal disorders: health profiles and six-months symptom changes of patients referred to norwegian outpatient clinics for chronic neck and back pain. *BMC Musculoskeletal Disorders*, *26*.

Santhanam, N., & Srinivas, S. (2019). Modeling the impact of employee engagement and happiness on burnout and turnover intention among blue-collar workers at a manufacturing company. *Benchmarking: An International Journal*, *ahead-of-print*.

Schaufeli, W. B., Taris, T. W., & Bakker, A. B. (2008). It takes two to tango: Workaholism is working excessively and working compulsively. *The long work hours culture: Causes, consequences and choices*, 203–226.

Schultz, A. B., & Edington, D. W. (2007). Employee health and presenteeism: A systematic review. *Journal of Occupational Rehabilitation*, *17*, 547–579.

Semadeni, M., Withers, M. C., & Trevis Certo, S. (2014). The perils of endogeneity and instrumental variables in strategy research: Understanding through simulations. *Strategic Management Journal*, *35*(7), 1070-1079.

Shifrin, N. V., & Michel, J. (2022). Flexible work arrangements and employee health: A meta-analytic review. *Work & Stress*, *36*, 60–85.

Shimazu, A., Schaufeli, W. B., Kamiyama, K., & Kawakami, N. (2015). Workaholism vs. work engagement: the two different predictors of future well-being and performance. *International journal of behavioral medicine*, *22*(1), 18-23.

Shimazu, A., Schaufeli, W. B., Kubota, K., & Kawakami, N. (2012). Do workaholism and work engagement predict employee well-being and performance in opposite directions? *Industrial Health*, *50*(4), 316-321.

Simpson, R. (1998). Presenteeism, power and organizational change: Long hours as a career barrier and the impact on the working lives of women managers. *British Journal of Management*, *9*(s1), 37–50.

Skagen, K., & Collins, A. M. (2016). The consequences of sickness presenteeism on health and wellbeing over time: A systematic review. *Social Sciences & Medicine*, *161*, 169–177.

Sorensen, G., Dennerlein, T., Peters, S. E., Sabbath, E. L., Kelly, E. L., & Wagner, G. R. (2021). The future of research on work, safety, health and wellbeing: A guiding conceptual framework. *Social Science and Medicine*, *269*, 113593.

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, *87*(3), 355–374.

Strömberg, C., Aboagye, E., Hagberg, J., Bergström, G., & Lohela-Karlsson, M. (2017). Estimating the effect and economic impact of absenteeism, presenteeism, and work environment–related problems on reductions in productivity from a managerial perspective. *Value in Health*, *20*, 1058–1064.

Taloyan, M., Aronsson, G., Leineweber, C., Magnusson Hanson, L., Alexanderson, K., & Westerlund, H. (2012). Sickness presenteeism predicts suboptimal self-rated health and sickness absence: A nationally representative study of the swedish working population. *PloS one*, *7*, e44721.

Ten Brummelhuis, L. L., Rothbard, N. P., & Uhrich, B. (2017). Beyond nine to five: Is working to excess bad for health? *Academy of Management Discoveries*, *3*(3), 262–283.

Terza, J. V. (1998). Estimating count data models with endogenous switching: Sample selection and endogenous treatment effects. *Journal of Econometrics*, *84*, 129–154.

Terza, J. V., Basu, A., & Rathouz, P. J. (2008). Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modeling. *Journal of Health Economics*, *27*(3), 531-543.

Treble, J., & Barmby, T. (2011). *Worker absenteeism and sick pay.* Cambridge University Press.

Triolle, A., Lefevre, N., Leclercq, M., Hairault, A., Lianova, R., & Pavis, I. (2024). *Revue de dépenses relative à la réduction des absences dans la fonction publique* (Rapport technique). Inspection générale des finances et Inspection générale des affaires sociales.

Turanovic, J. J. (2022). Exposure to violence and victimization: Reflections on 25 years of research from the national longitudinal study of adolescent to adult health. *Journal of Adolescent Health*, *71*(6), s14-s23.

van der Pol, M., & Irvine, A. (2018). *Time preferences for health.* Oxford University Press.

van Hooff, M., Flaxman, P., Söderberg, M., Stride, C., & Geurts, S. (2018). Basic psychological need satisfaction, recovery state, and recovery timing. *Human Performance*, *31*, 125-143.

Voss, M., Stark, S., Alfredsson, L., Vingård, E., & Josephson, M. (2008). Comparisons of self-reported and register data on sickness absence among public employees in sweden. *Occupational and Environmental Medicine*, *65*(1), 61–67.

Wang, M., Lu, C.-q., & Lu, L. (2022). The positive potential of presenteeism: An exploration of how presenteeism leads to good performance evaluation. *Journal of Organizational Behavior*, *44*.

Wooldridge, J. M. (2010). Econometric analysis of cross section and panel data, second edition. The MIT Press.

Yang, F., Qian, Y., & Xie, H. (2024). Addressing endogeneity using a two-stage copula generated regressor approach. *Journal of Marketing Research*, *0*(0), 00222437241296453.

Younès, N., Rivière, M., Plancke, L., Leroyer, A., Blanchon, T., Azevedo Da Silva, M., & Melchior, M. (2018). Work intensity in men and work-related emotional demands in women are associated with increased suicidality among persons attending primary care. *Journal of Affective Disorders*, *235*, 565-573.

Zinko, R., & Rubin, M. (2015). Personal reputation and the organization. *Journal of Management & Organization*, *21*(2), 217–236.

Čikeš, V., Maškarin Ribarić, H., & Črnjar, K. (2018). The determinants and outcomes of absence behavior: A systematic literature review. *Social Sciences*, *7*(8).

List of Tables

1.1 Descriptive statistics of dependent variables . . . 55
1.2 Contingency tables comparing sickness presenteeism and sickness absence . . . 56
1.3 AIC values for the discrete marginal distributions and copulae . . . 65
1.4 Determinants of sickness presenteeism and absence days with endogenous participation . . . 67
1.5 Average marginal effects comparison between univariate and bivariate models . . . 69
A1 Variables description . . . 97
A2 Descriptive statistics of covariables . . . 98
A3 AIC values for the discrete marginal distributions and copulas with a logit in first step . . . 99
A4 BIC values for the discrete marginal distributions and copulas with a logit in first step . . . 99
A5 AIC values for the discrete marginal distributions and copulas with a probit in first step . . . 100
A6 BIC values for the discrete marginal distributions and copulas with a probit in first step . . . 100
A7 AIC values for the discrete marginal distributions and copulas with a cloglog in first step . . . 100
A8 BIC values for the discrete marginal distributions and copulas with a cloglog in first step . . . 101
A9 SA and SP models presented in the chapter with confidence intervals in square brackets . . . 102
A10 Estimates of the effects of employment characteristics and working conditions on SA and SP days excluding health characteristics . . . 104

A11 Estimates of the effects of employment characteristics on SA and SP days excluding working conditions . . . 105
A12 Estimates of the effects of working conditions on SA and SP days excluding employment characteristics . . . 107
A13 SA and SP models without restriction on age . . . 108
A14 SA and SP models without restriction on working time . . . 110
A15 SP days with different restrictions on the illness duration . . . 112
A16 SA days with different restrictions on the illness duration . . . 114
A17 SA and SP models with the second best choice of copula . . . 116
A18 SA and SP models with the worst choice of copula . . . 118
A19 Determinants of sickness presenteeism and absence days with endogenous participation, without exclusion restriction in the first stage . . 120
A20 SA and SP models with violence at school as exclusion . . . 122
A21 SA and SP models with mistreatment as exclusion . . . 124
A22 SA and SP models with additional controls about past health and employment history . . . 126

2.1 Descriptive statistics of health characteristics by behaviour group for SA and SP . . . 159
2.2 Associations between group of baseline behaviour regarding SA and SP and outcomes . . . 161
2.3 Estimates of the effects of SA and SP durations at baseline on health outcomes . . . 162
2.4 Estimates of the 2sCOPE models . . . 165
2.5 Estimates of the 2sCOPE models with self-rated health interactions . 170
A1 Descriptive statistics of the individual and employment characteristics in 2013 and 2016 . . . 187
A2 Descriptive statistics of the working conditions in 2013 and 2016 . . . 188
A3 Descriptive statistics of the health characteristics in 2013 and 2016 . . 189
A4 Descriptive statistics of SA and SP regarding health state in 2013 . . 190
A5 Descriptive statistics of SA and SP regarding health state in 2016 . . 190
A6 Estimates of the effects of SA and SP interacted with self-rated health at baseline . . . 191

A7 Estimates of the 2sCOPE models with self-rated health interactions with SP and SA . . . 192

3.1 Descriptive statistics of the outcome variables . . . 206
3.2 Frequency of occupational hierarchical levels using professions . . . 206
3.3 Correlation matrix among the variables of interest . . . 208
3.4 Summary statistics of CP characteristics by level of CP index . . . 212
3.5 Summary statistics of the outcomes and their change over time regarding CP index . . . 227
3.6 Descriptive statistics of the outcomes regarding intrinsic motivations and work pressure . . . 228
3.7 Estimates of the effects of competitive presenteeism on hourly wages . 231
3.8 Estimates of the effects of competitive presenteeism on occupational grades . . . 232
3.9 Estimates of the effects of competitive presenteeism on self-rated health 234
3.10 Estimates of the effects of competitive presenteeism on sickness absence duration . . . 236
3.11 Estimates of the effects of competitive presenteeism on mental well-being . . . 237
A1 Descriptive statistics of the individual characteristics . . . 257
A2 Descriptive statistics of the working conditions . . . 258
A3 Pooled estimates of the effects of competitive presenteeism the outcomes 261
A4 2sCOPE estimates of competitive presenteeism on career outcomes . . 262
A5 2sCOPE estimates of competitive presenteeism on health outcomes . 263
A6 IV estimates of the effects of competitive presenteeism on hourly wages 264
A7 IV estimates of the effects of competitive presenteeism on occupational grades . . . 265
A8 IV estimates of the effects of competitive presenteeism on self-rated health . . . 265
A9 IV estimates of the effects of competitive presenteeism on SA days . . 266
A10 IV estimates of the effects of competitive presenteeism on mental well-being . . . 266

A11 Estimates of the effects of competitive presenteeism on career moderated by work pressure . . . 267
A12 Estimates of the effects of competitive presenteeism on health moderated by work pressure . . . 268
A13 Estimates of the effects of competitive presenteeism on career moderated by lack time . . . 269
A14 Estimates of the effects of competitive presenteeism on health moderated by lack time . . . 270
A15 Estimates of the effects of competitive presenteeism on career moderated by useful job . . . 271
A16 Estimates of the effects of competitive presenteeism on health moderated by useful job . . . 272
A17 Estimates of the effects of competitive presenteeism on career moderated by job pride . . . 273
A18 Estimates of the effects of competitive presenteeism on health moderated by job pride . . . 274
A19 Gender heterogeneity in the estimates of the effects of competitive presenteeism on career . . . 275
A20 Gender heterogeneity in the estimates of the effects of competitive presenteeism on health . . . 276
A21 Public sector heterogeneity in the estimates of the effects of competitive presenteeism on career . . . 277
A22 Public sector heterogeneity in the estimates of the effects of competitive presenteeism on health . . . 278
A23 Blue collars heterogeneity in the estimates of the effects of competitive presenteeism on career . . . 279
A24 Blue collars heterogeneity in the estimates of the effects of competitive presenteeism on health . . . 280
A25 Estimates of the effects of competitive presenteeism on career (Healthcare subset) . . . 281
A26 Public sector heterogeneity in the estimates of the effects of competitive presenteeism on career (Healthcare subset) . . . 282

A27 Public sector heterogeneity in the estimates of the effects of competitive presenteeism on health (Healthcare subset) 283
A28 Estimates of the variables included in the CP index$_{t-1}$ on hourly wage 284
A29 Estimates of the variables included in the CP index$_{t-1}$ on occupation 285
A30 Estimates of the variables included in the CP index$_{t-1}$ on health . . . 286
A31 Estimates of the variables included in the CP index$_{t-1}$ on WHO5 score 287
A32 Estimates of the variables included in the CP index$_{t-1}$ on WHO5 score 288

List of Figures

1 Empirical research framework 30

1.1 Distribution of sickness absence and presenteeism days of ill individuals 55

1.2 Illness thresholds for which workers decide to be present sick (SP) or absent (SA) 58

1.3 Q-Q plots of the randomized and normalized residuals of the count regressions with endogenous participation. The first row corresponds to the regression of sickness absenteeism days with a negative binomial type II margin, and the second row corresponds to the regression of sickness presenteeism days with a Poisson inverse Gaussian margin. 66

1.4 Smooth functional forms of work hours on sickness absence and presenteeism 77

A1 Q-Q plots of the randomized and normalized residuals of the count regressions with endogenous participation. The first row corresponds to the regression of sickness absenteeism days with a Poisson margin and Gaussian copula, and the second row corresponds to the regression on sickness presenteeism days with a Poisson margin and Gaussian copula. 101

3.1 Distribution of the competitive presenteeism index 210

A1 Correlations of the variables with the first two dimensions 259

A2 Percentage of variance explained by dimension (eigenvalue) 259

A3 Modalities position and contribution to the first two dimensions . . . 260

A4 Variables contribution to the first dimension 260

Titre et résumé :

Trois essais sur les déterminants et les conséquences du présentéisme des salariés en France

Dans un contexte de hausse des charges de sécurité sociale en France, comprendre les mécanismes expliquant l'augmentation de l'absence est essentiel. Au-delà du coût important de l'absence, il convient de considérer celui du présentéisme, c'est-à-dire le fait de fournir un effort alors qu'on ne devrait pas. Cette thèse étudie deux formes de présentéisme : le présentéisme maladie (travailler en étant malade) et le présentéisme compétitif (travailler plus que nécessaire). En outre, elle propose une analyse empirique conjointe des comportements d'absence et de présentéisme. Ainsi, le premier chapitre explore les effets des caractéristiques de l'emploi sur la survenue de la maladie, ainsi que sur les durées cumulées d'absence et de présence en cas de maladie. Il montre que certains facteurs préalablement identifiés comme affectant le présentéisme et l'absence concernent en réalité la maladie elle-même. Par ailleurs, la durée d'absence dépend surtout des caractéristiques du contrat de travail, tandis que le présentéisme maladie est davantage lié aux conditions de travail. Le deuxième chapitre étudie les effets à long terme des choix entre absence et présence en cas de maladie sur la santé future. Celui-ci confirme que le présentéisme maladie nuit à la santé future, contrairement à l'absence, qui, bien que coûteuse à court terme, est bénéfique pour le rétablissement de la santé à long terme. Enfin, le troisième chapitre examine quant à lui les effets du présentéisme compétitif sur la carrière et la santé. Si ce comportement est effectivement associé à une progression de carrière en termes de salaire et de niveau professionnel, celui-ci peut néanmoins dégrader la santé et conduire à une augmentation de la durée d'absence future.

Descripteurs :

Présentéisme, Absentéisme, Santé, Bien-être, Motivation, Carrière, Signal, Endogénéité, Copules.

Abstract:

In a context of rising social security costs in France, it is essential to understand the mechanisms behind the increase in absence. Beyond the significant cost of absence, it is important to consider the cost of presenteeism, i.e. the act of working when one should not. This thesis examines two forms of presenteeism: sickness presenteeism (working while sick) and competitive presenteeism (working more than necessary). It also offers a joint empirical analysis of absenteeism and presenteeism behaviours. The first chapter explores the effects of employment characteristics on illness, as well as on the cumulative duration of absenteeism and presenteeism in the event of illness. It shows that some factors previously identified as affecting presenteeism and absenteeism are in fact related to the illness itself. Furthermore, the duration of absence depends mainly on the characteristics of the employment contract, while sickness presenteeism is more closely related to working conditions. The second chapter examines the long-term effects of choices between absence and attendance in the event of illness on future health. It confirms that sickness presenteeism is detrimental to future health, unlike absence, which, although costly in the short term, is beneficial to recovery in the long run. Finally, the third chapter examines the effects of competitive presenteeism on career and health. While this behaviour is indeed associated with career advancement in terms of wages and occupational status, it can nevertheless damage health and lead to an increase in the future duration of sickness absence.

Keywords:

Presenteeism, Absenteeism, Health, Well-being, Motivation, Career, Signal, Endogeneity, Copula.