BACHELOR THESIS

MAJOR POLICY ECONOMICS

ERASMUS SCHOOL OF ECONOMICS

The relationship between working overtime and mental well-being: a longitudinal study

Author:
Dennis Boers - 566976

Supervisor:
dr. Aart Gerritsen

Second assessor:
dr. Josse Delfgaauw

July 13, 2023

ERASMUS UNIVERSITY ROTTERDAM

Abstract

In recent years, there has been a worldwide upward trend in the number of mental illness cases. It is interesting to know whether this is partly due to overworking. This thesis studies the causal relationship between overtime working and mental well-being. Past research has found a negative effect of overworking on mental well-being, but research on this relationship has been mainly concerned with finding associations rather than causality. This study focuses explicitly on causality and is thus more useful for policy setting. This study uses survey panel data covering the period 2009-2022 from the UK to study whether overtime working is one of many causes of lower mental well-being among UK employees. An individual fixed effects model is estimated using the variation over time in the number of overtime hours worked to uncover the causal impact of overworking on mental well-being. The results suggest that there is a negative impact of overworking on mental well-being. This effect is driven by the number of overtime hours, and not necessarily overworking per se. Moreover, the negative impact of overworking is only present among those that work in a non-manual job, which are often mentally more demanding, where no impact is found for those working in a manual job. Unpaid overtime hours seem to do more harm than paid overtime hours. Policy aimed at improving employee mental well-being should be aimed at decreasing the number of overtime hours worked by employees, and at forcing employers to pay their employees for overtime hours worked. There are many other factors influencing employee mental well-being, thus these policies should be complemented by other policies improving employee mental well-being, besides aiming at overtime hours.

1 Introduction

The prevalence of mental health problems has been rising over the past few years. The period 2007-2017 saw a worldwide rise in mental health and substance abuse problems by 13 percent (World Health Organization, 2022a). This led to an estimated 15 percent of working-age adults having a mental disorder in 2019. A proposed reason for this is a poor working environment, including excessive workloads and low job control (World Health Organization, 2022b).

These are already worrying numbers, but it can get more extreme if we continue having poor working environments. The most extreme case is the situation in Japan. In Japan, long working hours and many overtime hours are widespread. In 2021, Japanese workers on average worked 9.7 hours of overtime per month (Statista, 2022). This average seems to shadow large variations in overworking among workers, since there are several cases of death because of mental issues due to extreme overworking in Japan, which still is a major social issue in Japan. It even has its own word: ‘karoshi’, which translates to: ‘death by overwork’.

These issues are the consequence of the workplace culture in most workplaces in Japan. An article by the British Broadcasting Corporation (2020) (BBC) illustrates this perfectly. A citation from an interviewed Japanese worker makes this clear when being asked why he has not taken more days off: *“It’s difficult because the atmosphere in the workplace wouldn’t allow it”*. Death by overwork is however not limited to Japan, but is prevalent all over the world. In 2016, 745,194 individuals worldwide died from physical and mental health problems attributable to overwork (Pega et al., 2021)[1].

By now, it should be clear that overworking could lead to serious physical, but certainly also mental health problems, worsening overall mental well-being among employees worldwide. Therefore, this thesis studies whether overworking has an impact on mental well-being among a sample of UK employees.

There already exists some evidence on the impact of overworking on mental well-being, also in the UK, but most published papers focus more on associations rather than causality. This study utilises the within-individual variation over time in overtime hours worked to try to reach causality rather than correlation. What makes this study also unique is that it uses a different, more reliable measure of mental well-being: the 12-item General Health Questionnaire (GHQ-12), which has specifically been developed to measure psychological distress.

The data used in this study come from the UK Household Longitudinal Study (UKHLS) and covers the period 2009-2022. I estimate an individual fixed effects regression model to try to uncover the causal impact of working overtime on employee mental well-being, and I also conduct heterogeneity analyses to disentangle the differential impact of overworking in manual and non-manual jobs, and of paid and unpaid overtime hours.

The results suggest that overworking per se has no impact, but there is a small negative impact of the number of overtime hours on mental well-being, and that this is driven by those working in a non-manual job, which are most often more mentally demanding. Moreover, paid

[1]These deaths attributable to overwork are mainly caused by large burdens of ischematic heart diseases and strokes, as a consequence of overwork.

overworking seems to have much less impact on mental well-being than unpaid overtime hours, which suggests that employees value the acknowledgement of their extra effort put into work by their employer.

Several robustness tests are conducted. Instead of using an OLS individual fixed effects regression, I estimate a fixed effects ordered logit model developed by Baetschmann et al. (2015). As a second robustness test, I remove the covid-19 pandemic from the data. As a last robustness test I test the consistency of the GHQ-12 questionnaire in measuring mental well-being. The results of these tests indicate that the results found in this thesis are robust to those different approaches, which gives confidence in the validity of the results.

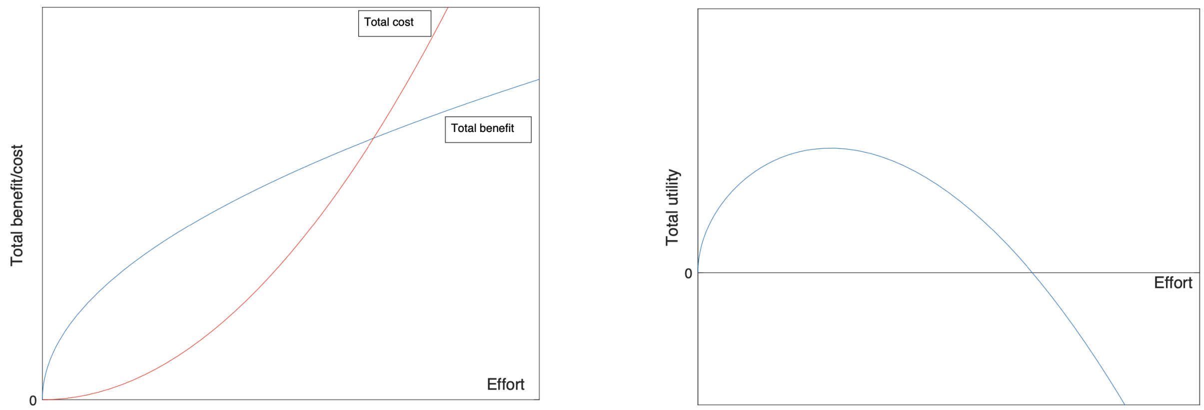
The remainder of this thesis is structured as follows: section 2 introduces the theoretical framework, section 3 describes the evidence already available on the relationship between work and (mental) well-being, especially overworking, section 4 describes the data used, section 5 describes the empirical strategy used in uncovering the causal impact of overworking on mental well-being, section 6 describes the results of various analyses related to this relationship, section 7 discusses the various robustness tests conducted to validate the results, section 8 discusses the potential explanations and policy prescriptions that follow from the results and the limitations of this study, and section 9 concludes.

2 Theoretical framework

In the classical marginalistic theory of labour, utility is defined as the sum of all *pleasures* and *pains*. The notion of utility maximization is described as follows: *"To satisfy our wants to the utmost with the least effort - to procure the greatest amount of what is desirable at the expense of the least that is undesirable - in other words, to maximize comfort and pleasure."*. Labour is defined as *"the painful exertion which we undergo to ward off pains and of greater amount, or to procure pleasures which leave a balance in our favour."* (Jevons, 1871). From the notion of utility maximization, it follows that one should exert effort until the pleasure of the extra income earned no longer outweighs the pain of exerting extra effort. The former is thus the marginal benefit of exerting an extra unit of effort, whereas the latter is the marginal cost, both in terms of utility. The assumption in the classical theory is that marginal benefit is decreasing in effort and marginal cost is increasing in effort.

This marginal analysis can best be shown graphically. The classical assumptions are best represented by the curves in figure 1. In both panels, effort is on the x-axis. Effort could be seen as the duration of work in hours, which is consistent with the measure for amount of work in this thesis. Total utility in figure 1b, pleasure minus pain, is given by the vertical distance between the total benefit and cost curve in figure 1a. Consequently, all points above the x-axis correspond to *'pleasure'* and the points below to *'pain'*, and utility maximization corresponds to the maximum of the total utility curve.

Modern micro-economic labour-leisure choice models based on classical marginalistic theory assume that a worker only can spend his time on work and leisure, is perfectly rational and consequently only cares about utility maximization, given his preferences between leisure and



(a) Total benefit and cost

(b) Total utility

Figure 1: Graphical analysis of the classical theory of labour

income, and given the market wage rate, net of tax. When overworking is a voluntary choice by the worker, the overworking should lead to utility maximization where his marginal value of leisure equals his net market wage rate, the former the marginal benefit of leisure, the latter the marginal cost of leisure. Equivalently, the former is the marginal cost of labour and the latter the marginal benefit of labour. When workers work more or less than these hours under utility maximization, utility would be less and thus well-being decreases.

These models however assume perfectly competitive markets, and thus no restrictions on working hours for workers. When there are restrictions on the number of hours one can choose to work, utility maximization often is impossible, and with concave preferences in both leisure and income (i.e. Cobb-Douglas preferences) and a linear budget constraint, working too much relative to a worker's 'bliss point' would imply decreased utility.

Mathematically, the model description is as follows. Worker i's utility over leisure L and net income Y is given by some utility function $U_i(Y_i, L_i)$, with utility increasing both in income and leisure, but at a decreasing rate[2]. The budget constraint of worker i given the gross wage rate w_i of worker i, total time available L_i^{max} and marginal tax rate t can be summarized by the following expression:

$$Y_i = (1-t)w_i(L_i^{max} - L_i) \tag{1}$$

Where $L_i^{max} - L_i$ is the total labour supply, and $(1-t)w_i(L_i^{max} - L_i)$ is net labour income for worker i. Workers choose their labour supply as to maximize their utility. Formally, they solve the following optimization problem:

$$\max_{Y_i \geq 0, L_i \geq 0} \quad U_i(Y_i, L_i) \quad \text{s.t.} \quad Y_i = (1-t)w_i(L_i^{max} - L_i) \tag{2}$$

Graphically, the model can be summarized by figure 2 below. The curve given by $U_i = U_i(Y_i, L_i)$ is the indifference curve, corresponding to utility level U_i. A higher curve corresponds to a higher utility level. When a worker works too many hours compared to his 'bliss point' (point L^* in figure 2), utility decreases. This model thus leads to the same prediction as the classical theory: when a worker involuntarily works too many hours (e.g. too little leisure), utility decreases.

[2]Mathematically, $\frac{\partial U_i(Y_i,L_i)}{\partial L_i}, \frac{\partial U_i(Y_i,L_i)}{\partial Y_i} > 0$, $\frac{\partial^2 U_i(Y_i,L_i)}{\partial L_i^2}, \frac{\partial^2 U_i(Y_i,L_i)}{\partial Y_i^2} < 0$.

The theory thus predicts that working a positive amount of hours is desirable and produces utility, or excess 'pleasure', but putting in too much hours of work could lead to disutility, or excess 'pain'. In the next section, I will review the empirical findings on the relationship between well-being, often used as measure for total utility, mental well-being, which could be seen as part of total well-being and effort as measured by work duration in hours.

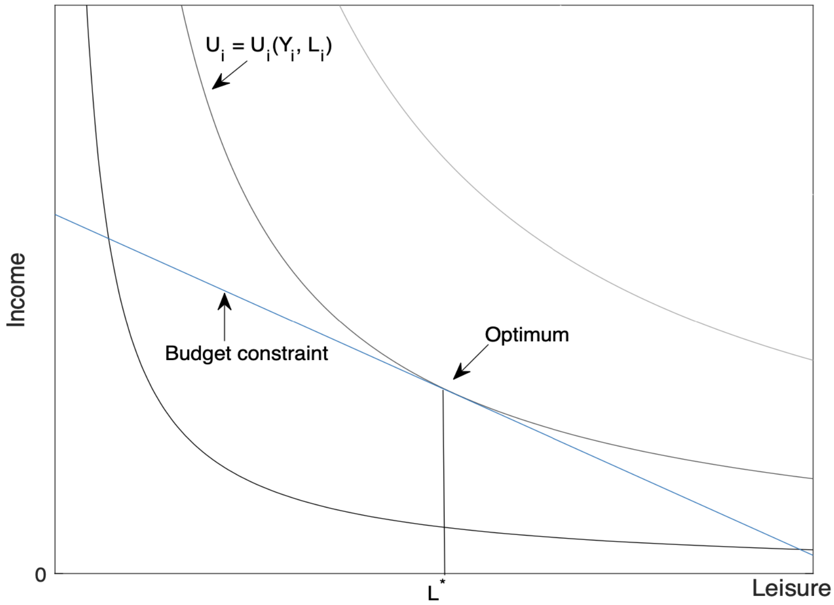


Figure 2: Modern micro-economic leisure-labour choice model

3 Related literature

My thesis is closely related to the broader literature on the determinants of mental well-being and more generally well-being, life satisfaction, working hours and happiness. This is a quite broad literature, with a lot of economic and also psychological research already existing on these topics. In the following subsections, I will give a brief overview of the most relevant work-related past research, bridging the gap between theory and empirics.

3.1 The relationship between work and well-being

The first strand of research within the broader related literature looks at the relationship between work and well-being, or utility. There are a lot of factors contributing to the well-being of individuals, and this certainly holds for the mental well-being of individuals. The first important economic determinant is unemployment. An early study on this relationship is the study by Clark and Oswald (1994). They use data from the BHPS for just one single wave (1991). This study uses the same questionnaire for measuring mental well-being as I do: the GHQ-12. The findings are that the jobless are on average more likely to be distressed, and thus have a lower mental well-being than the working people.

This is a common result in the literature. For example, Gerdtham and Johannesson (2001) find that unemployment is negatively related to subjective well-being, using a random sample of over 5,000 Swedish households, and Gerlach and Stephan (1996) find this result using GSOEP data on German households, and Frey and Stutzer (2000) find a strong negative relationship between unemployment and well-being for over 6,000 residents of Switzerland. A more recent study by Hetschko et al. (2014) confirms this, also concluding from their study that the unemployed report a lower life satisfaction than the employed. This finding is in accordance with the theory introduced in section 2. Even the experience of losing your job could result in lower subjective well-being. Clark et al. (2001) find evidence of 'scarring', using panel data on German households from the GSOEP.

These studies thus highlight that working is enhancing for well-being, as compared to not working, presumably because of social network building through colleagues and the extra income earned. On top of that, unemployment is found to be more harmful for individual happiness than non-participation (Winkelmann and Winkelmann, 1995). A second study by Winkelmann and Winkelmann (1998) tries to explain why the unemployed are experiencing a decreased happiness. Their findings are that the loss of income does have a relatively small effect on life satisfaction, as compared to the non-pecuniary costs of unemployment, such as a loss of social contact on the work floor.

Work characteristics also matter for well-being. For example, low levels of job strain impact well-being positively. A study by Stansfeld et al. (2013) looks at the well-being impact of personal social support and work characteristics. They specifically look at the association between level of job strain and well-being. A 'high strain job' is defined as a job with low decision latitude and high demands, and a 'low strain job' as one with high decision latitude and low demands. The results show a positive association between employees in a low strain job and employee well-being, compared to employees in a high strain job.

Also, job satisfaction seems an important factor in mediating the impact of certain work characteristics on well-being, such as work demands which likely influences job satisfaction. A recent study by Lorente et al. (2018) focuses on factors mediating the relationship between certain work characteristics and well-being[3]. Their results show that job meaningfulness and job satisfaction are the most important mediating factors for the relationship between various work characteristics and psychological well-being, and thus the impact of for example overworking may be greater or smaller between different jobs.

Besides these direct work characteristics, the work-family conflict is also found to be related to employee well-being, where the work-family conflict refers to the interference of work life in an employee's family life. Grant-Vallone and Donaldson (2001) study the effects of work-family conflicts on employee well-being using survey data on non-professional employees from the greater Los Angeles area, and find that the employees indicating experiencing work-family conflicts also reported a lower well-being[4]. This can be placed in the theoretical framework of

[3] The work characteristics included in the study by Lorente et al. (2018) are based on the Job Content Survey by Karasek et al. (1998), which includes the following items: decision latitude, psychological demands and mental workload, social support, physical demands and job insecurity.

[4] These mentioned studies all highlight important work related determinants of well-being. Other social and economic determinants influencing well-being include health (Shields and Price, 2005), age (Ferrer-i Carbonell and Gowdy, 2007), income (Pouwels et al., 2008), relative income (Clark et al., 2008), number of children and marital

section 2, as work could interfere with family life through employees having to work too many hours, and thus too little leisure, such that employees are way off their bliss point in figure 2.

However, the question remains whether hours worked and more importantly whether overtime hours worked also impact general well-being, especially mental well-being, which is the focus of this study. There is evidence that working is better than not working, and that lower job strain is positively associated with well-being. To shed some light on the possible impact of (over)working hours on mental well-being, some evidence is presented in the last two subsections.

3.2 The relationships between hours worked and well-being

The second strand of research within this literature studies the relationship between working hours and subjective well-being. Most studies approximate well-being by life satisfaction, and some others also look at hours and job satisfaction as extra measures of well-being. Gerritsen (2016) studies optimal income taxation, letting go of the typical assumption in public economics that utility maximization perfectly corresponds with well-being maximization. Subjective well-being is measured as the subjective life satisfaction reported by the respondent, using panel data on British households from the BHPS. Using an individual fixed effects approach, he finds that the subjective well-being function follows an inverted U-shape in the number of working hours. Working too little could thus result in lower well-being, but working too much also does, which perfectly corresponds with theory.

There is however no consensus on the relationship between hours worked and well-being. Booth and Van Ours (2008) study the causal relationship between type of job (part-time; 1-29 hours or full-time; 30+ hours) and partnered well-being, defined as the sum of the well-being of two partners and measured by hours, job and life satisfaction. They also use data from the BHPS, as Gerritsen (2016) does, but only for couples in which the female partner was aged between 25 and 50. They use a slightly different approach than in earlier research, using instead a fixed effects ordered logit model based on Ferrer-i Carbonell and Frijters (2004). They conclude that life satisfaction is not related to the number of hours worked for partnered men and women. They also find that having a job influences life satisfaction the most for men, and women with children, rather than the number of hours worked.

The article by Pouwels et al. (2008) studies the degree to which the positive effect of income on happiness is underestimated in earlier attempts to estimate this relationship. Besides their main results that this effect is underestimated by 25%, their results also indicate that weekly working hours is not significantly impacting happiness, and more broadly well-being.

These studies focused on the direct impact of working hours on well-being. However, Wooden et al. (2009) come up with a different explanation for the observed relationship between working hours and well-being: the mismatch between preferred and actual number of hours worked,

status (Shields and Price, 2005), job type (Bardasi and Francesconi, 2004), job changes (Björklund et al., 2013) and social interaction (Lelkes, 2006). Van Praag et al. (2003) confirm this. They study the associations between different aspects of life which they call domains, such as health, financial situation and job. They perform an ordered probit random-effects regression on GSOEP survey data, and conclude that general well-being is significantly related to job, financial, house, health, leisure and environment satisfaction. General well-being is thus related to all factors relating to these domains.

rather than the number of hours worked itself. They use panel data from the HILDA survey, an Australian panel survey, containing data on the lives of Australian households. They account for time-invariant individual heterogeneity and find that it is the mismatch between the preferred number of hours worked and the actual hours worked that is driving this relationship. For both men and women working more or less than their preferred hours, they document a negative relationship between working hours and well-being, though this relationship is stronger for the overemployed. They find no significant relationship between work hours and well-being for the workers working their preferred number of hours. These results do not comply with theory, since on the basis of the theory, one would expect to see a positive relationship between hours worked and well-being for the underemployed.

In addition, Ala-Mursula et al. (2002) find that among female workers, poor health and low well-being were more prevalent in the lowest quartile of worktime control than those in the highest quartile, where the worktime control measure is constructed from survey questions on a number of worktime aspects[5]. This association was not found for male workers.

3.3 The relationships between overtime hours and well-being

More closely related research on the relationship between overtime work and well-being also exists. For example, a study by Beckers et al. (2008) shows that workers working overtime involuntarily is associated negatively with mental well-being, proxied by fatigue and work satisfaction[6]. This negative association was even stronger for those getting no compensation for overtime work. Beckers et al. (2008) did use cross-sectional survey data for 2004 on 1,612 Dutch workers, and ran an ANOVA analysis. Since there is likely to be time-invariant heterogeneity between workers confounding the results, an ANOVA analysis is not sufficient to establish causality (Ferrer-i Carbonell and Frijters, 2004).

Nevertheless, there are some strong signs that working overtime causes lower mental well-being, especially when involuntary and unpaid, which is in line with the micro-economic leisure-labour choice model. Among Swedish business professionals, there also seems to be a negative association of working overtime on mental well-being (Broberg et al., 2020). Also, Golden and Wiens-Tuers (2006) find in their research that overworking has a negative impact on mental well-being. On top of that, the extra income of working overtime reduces the negative impact of overworking somewhat. It appears that the work-family conflict is the most important mechanism through which overtime causes lower mental well-being, which fits in the theoretical framework of section 2 (Golden and Wiens-Tuers, 2006).

These findings contrast earlier findings by Beckers et al. (2004), who also exploit Dutch survey data on 1,807 workers. They find that overtime workers do not report to be more fatigued than non-overtime workers. Not all studies thus conclude that overworking is negatively associated with mental well-being.

These studies look at the direct impact of overtime hours worked by the individuals them-

[5] Worktime control refers to the situation in which an employee has some power to set his working hours.

[6] The involuntary nature of overworking is measured by survey answers indicating whether a respondent wants to or does not want to work overtime.

selves, but spill-over effects could also be present. The study by Ishida et al. (2020) tries to examine the spillover effect of having an overtime-encouraging working environment. The results indicate that there is an positive association between having an overtime working environment (OWE) and psychological distress among a sample of Japanese workers, after controlling for the direct effect of overtime hours worked by the individuals themselves.

My thesis makes several contributions to the existing literature. First, the focus in my thesis is on uncovering the causal impact of overworking on mental well-being, which is quite different from the approaches taken in other articles and more relevant for policy. Second, I will use the 12-item General Health Questionnaire as a measure of mental well-being, which is a different measure used than in most of these earlier studies. This is one of the most reliable measures of mental well-being, originally developed as a measure for psychological distress. Broberg et al. (2020) and Robone et al. (2011) also use the 12-item GHQ as measure for mental well-being, but do not take into account fixed effects. My study provides evidence on mental well-being using this same measure, but taking into account fixed effects, allowing for causally more credible results. Third, my thesis provides evidence on the differential impact of overworking between manual and non-manual jobs, and between paid and unpaid overworking. Beckers et al. (2008) and Robone et al. (2011) find that uncompensated overworking is more harmful and that non-manual workers experience a negative effect of overworking on mental well-being, though they use methodologies only allowing for uncovering associations rather than causality. Last, my thesis will add to the emerging economic literature on the determinants of psychological health and mental well-being.

Based on the predictions made by the introduced theoretical framework and the associations found in earlier published work, I expect to find a negative effect of overworking on mental well-being, mainly driven by non-manual workers and uncompensated overworking hours.

4 Data

4.1 The dataset

I use data from the UK Household Longitudinal Study (UKHLS). This is a longitudinal survey dataset containing individual-level data on family life, wealth, education, employment and health for approximately 40,000 households in the UK. The data consists of 12 waves spanning the period 2009-2022. The respondents are surveyed annually, to produce a dataset of 1,063,224 individual-wave observations. In the remainder of this section, I will describe the variables and sample I will use in the analyses.

4.2 Measuring mental well-being

The dependent variable in my study is mental well-being. Earlier literature on well-being has mainly used life, job and hours satisfaction as measures for well-being (e.g. Gerritsen (2016); Booth and Van Ours (2008)). Since the focus in my thesis is more on mental well-being, I will use a different measure, namely the GHQ-12 questionnaire. This questionnaire is based on the 12-item General Health Questionnaire (GHQ-12)[7]. The GHQ-12 is a questionnaire on the mental well-being of an individual, and is widely used to measure mental well-being (e.g. Clark and Oswald (1994)), and to screen for mental health problems (Anjara et al., 2020). Another advantage of the GHQ-12 questionnaire is that it asks about a respondent's level of mental well-being over the past few weeks as opposed to some average level of mental well-being over a longer period, which makes it possible to capture mental well-being changes from working overtime in the past weeks.

The UKHLS survey contains a separate question on all 12 items. In table 1 below, the shortened 12-item General Health Questionnaire is given, together with the corresponding measured item. The possible answers to these questions are divided into 4 categories:"More so than usual","Same as usual", "Less so than usual" and "Much less than usual", which are then recoded into the numbers 1 to 4. The answers are ordered and thus the variables are ordinal categorical variables.

From the answers to these 12 questions, an aggregated score is derived. For the positively formulated questions (e.g. happiness), if a respondent reports "More so than usual", no point is added to their GHQ-12 score, if a respondent reports "Same as usual", 1 point is added, if a respondent reports "Less so than usual", 2 points are added, and if a respondent reports "Much less than usual", 3 points are added to their GHQ-12 score, which I will refer to with mental well-being score from now on. For the negatively formulated questions (e.g. constantly feeling under strain), the points are allocated in the opposite order. The mental well-being score thus ranges from 0 to 36, where 0 is least distressed, and thus corresponds with the best mental well-being, and a score of 36 is most distressed, and corresponds with the worst mental well-being.

The distribution of the mental well-being score is given in figure 3a and the distribution for the individuals working overtime and not working overtime separately is given in figure 3b. Most individuals (75.1%) report having a mental well-being score of 12 or less, which indicates that most have a fine mental well-being, where few (3.0%) report having a mental well-being score of 24 or higher, indicating a bad mental well-being. From figure 3b, we can see that the distributions of the mental well-being score does not differ greatly between those working and not working overtime, though the extreme values at the upper end of the distribution are a little more prevalent among the ones working overtime, which is already an indication that overworking potentially has a small negative impact on mental well-being.

4.2.1 The reliability of the GHQ-12 score as measure of mental well-being

As already mentioned, the mental well-being score is an aggregated score involving 12 separate items included in the questionnaire. To be able to reliably use this score as a measure of mental well-being, I need to know whether this questionnaire and the aggregated score does indeed measure a common underlying factor, in this case mental well-being, in the analysis sample, elaborately described in section 4.5. To test this, I use Cronbach's alpha to determine whether

[7]The 12-item General Health Questionnaire was introduced by Goldberg (1972), initially developed to measure psychological distress.

Table 1: GHQ-12 QUESTIONNAIRE

Question	**GHQ-12 item**
"Have you recently ...	
been able to concentrate on whatever you're doing?"	Concentration
lost much sleep over worry?"	Loss of sleep
felt that you were playing a useful part in things?"	Usefulness
felt capable of making decisions about things?"	Decision making
felt constantly under strain?"	Under strain
felt you couldn't overcome your difficulties?"	Overcoming difficulties
been able to enjoy your normal day-to-day activities?"	Enjoying day-to-day activities
been able to face up to problems?"	Facing problems
been feeling unhappy or depressed?"	Depressed
been losing confidence in yourself?"	Losing confidence
been thinking of yourself as a worthless person?"	Self-esteem
been feeling reasonably happy, all things considered?"	General happiness

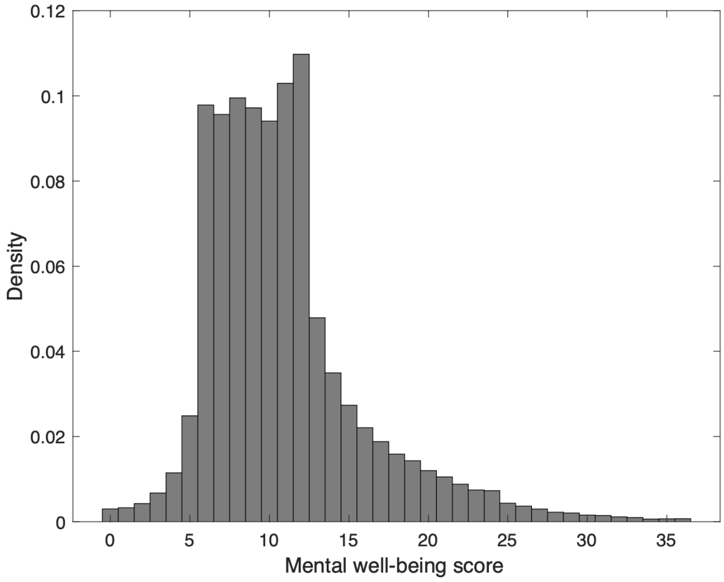
the 12 items are interrelated and indeed all measure mental well-being.

The results of this test are reported below in table 2. The first column shows the correlation between the corresponding item and the test scale formed by the other remaining items. The second and third column show the average inter-item correlation and value of Cronbach's alpha after dropping the corresponding item. The last row shows the overall average inter-item covariance and value of Cronbach's alpha for all 12 items.

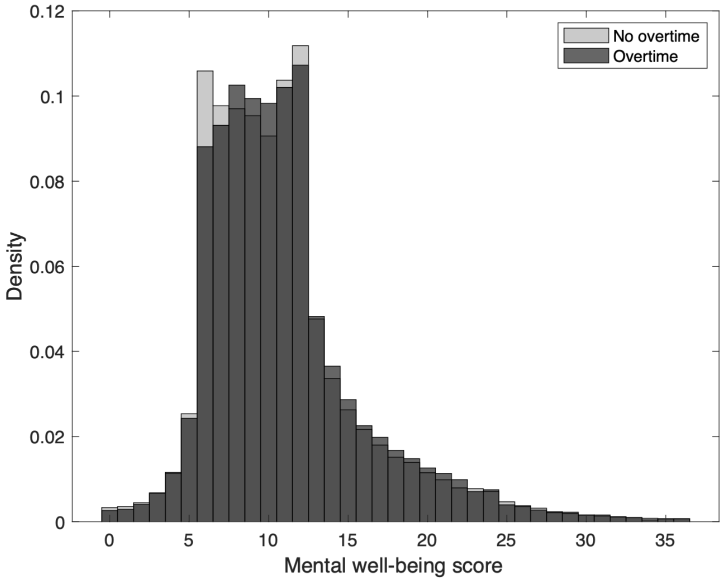
From the first column we can see that the items are all quite related to each other, with correlations bigger than 0.50 for almost all separate items. Column 2 reveals that by removing some items (e.g. concentration and usefulness), the average covariance between the items increases slightly from the overall average of 0.1525. However, we should look at the values of Cronbach's alpha in column 3. A quick look at this column leads to the conclusion that all items in the GHQ-12 questionnaire are valuable to measuring an underlying factor, since Cronbach's alpha does not increase by dropping an item from the average level of 0.8918[8]. More importantly, this value indicates that the items have a high internal consistency.

I conclude that the GHQ-12 questionnaire is a reliable measure of mental well-being, strengthening the case for using this as measure for mental well-being in my thesis.

[8]The value of Cronbach's alpha only tells us something about the *reliability* of the questionnaire. It tells us nothing about the *validity* of the questionnaire, meaning that it does not tell us anything about *what* common underlying factor it reliably measures. Since the items all relate to mental health, I will just assume that it is actually measuring mental well-being. Del Pilar Sánchez-López and Dresch (2008) for example show that the GHQ-12 questionnaire is valid for measuring overall mental well-being. The frequent use of the GHQ-12 as a measure of mental well-being in the literature strengthens the plausibility of this assumption (see for example Shields and Price (2005); Clark and Oswald (1994)).



(a) Full sample



(b) Overtime and no overtime

Figure 3: Distribution of the mental well-being score

Notes: figure 3 shows a histogram plot for the mental well-being score, derived from the respondent's answers to the GHQ-12 questionnaire, described in table 1, using the analysis sample described in section 4.5. Panel (a) does this for the full sample, panel (b) does this for the ones working overtime and those that do not work overtime separately. Each bar corresponds to the associated mental well-being score level.

4.3 Working hours and overtime

The main independent variable is the number of overtime hours worked in a normal workweek, as reported by the respondent. This includes paid and unpaid overtime. The second important variable is the number of hours worked in a normal week, excluding overtime hours. The questions regarding the normal hours worked and overtime hours worked are asked successively and formulated as follows: "Thinking about your (main) job, how many hours, excluding overtime

Table 2: Reliability test of the GHQ-12 questionnaire

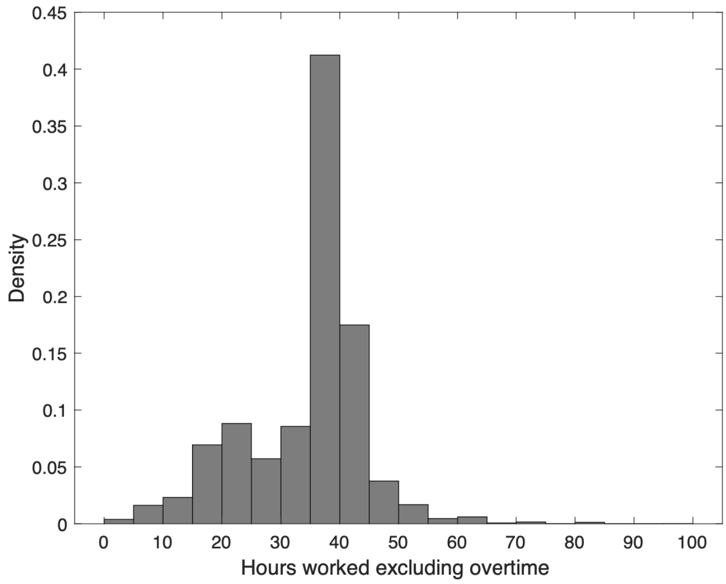
Item	Item-rest correlation	Average inter-item covariance	Cronbach's alpha
Concentration	.5473	.1610	.8862
Loss of sleep	.5899	.1470	.8848
Usefulness	.4892	.1628	.8885
Decision making	.4954	.1657	.8888
Under strain	.6298	.1457	.8819
Overcoming difficulties	.6772	.1447	.8787
Enjoy activities	.5833	.1588	.8845
Facing problems	.5462	.1635	.8869
Depressed	.7476	.1369	.8743
Losing confidence	.7157	.1400	.8764
Self-esteem	.6402	.1489	.8809
General happiness	.6178	.1555	.8826
Overall	$N = 89,913$	.1525	.8918

Notes: table 2 shows the results for the Cronbach's alpha test of survey reliability, using the full analysis sample from section 4.5. The first column report the correlation of the item with the test scale formed by the other items, the second column gives the average covariance between the items if the corresponding item is removed, and the third column gives the value for Cronbach's alpha if the corresponding item is removed.

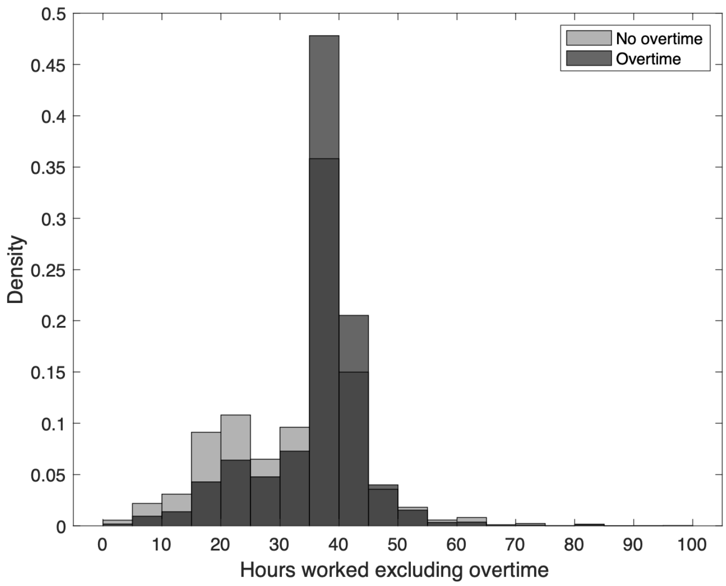
and meal breaks, are you expected to work in a normal week? And how many hours overtime do you usually work in a normal week?". These questions are only asked to employees, so I do not consider self-employed individuals or other forms of employment.

In the literature, there has already been found a relationship between well-being and hours worked (e.g. Gerritsen (2016)), thus individuals already working 40 hours per week excluding overtime already might have a different level of mental well-being than someone working 32 hours excluding overtime. The number of hours worked excluding overtime will also likely influence the overtime hours: an employee already working 40 hours a week may be more tired and thus less willing and therefore may have a lower chance of being asked to work overtime, where an employee working a lower amount of hours may agree on working overtime, or a workaholic working already a lot also wants to work a lot of overtime hours. If I do not control for these hours worked, the effect of hours worked excluding overtime would be picked up by the coefficient of overtime hours. Since I am focusing on the effect of overtime hours on mental well-being, it is necessary to control for these hours excluding overtime.

Figure 4a plots the distribution of hours worked excluding overtime in the full sample, figure 4b does this for the ones overworking and the ones that do not work overtime separately and figure 5 plots the distribution of overtime hours worked for the subsample working overtime. We can see from figure 4a that 90.4% of the respondents work 40 hours or less per week, excluding overtime, which is as expected. The distribution of normal hours worked excluding overtime does differ substantially between the ones overworking and the ones that do not overwork, as can be seen from figure 4b. It seems that employees overworking also tend to work more regular hours. We can also see from figure 5 that most respondents (79.6%) that report to overwork do work between 0 and 10 overtime hours.



(a) Full sample



(b) Overtime and no overtime

Figure 4: Distribution of hours worked excluding overtime

Notes: figure 4 shows a histogram plot of the number of hours worked excluding overtime hours in a normal week, as reported by the respondent, using the analysis sample described in section 4.5. Panel (a) does this for the full sample, panel (b) does this for the ones working overtime and those that do not work overtime separately. Each bar corresponds to an interval of width 5. For example: the first bar corresponds to 0-5 hours worked, the second bar to 5-10 hours worked and so on.

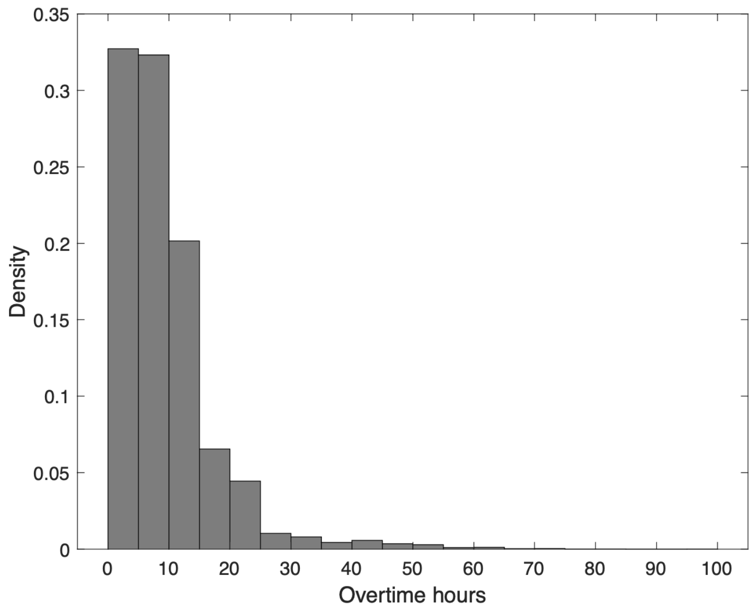


Figure 5: Distribution of overtime hours worked

Notes: figure 5 shows a histogram plot of the number of overtime hours worked in a normal week, as reported by the respondent, using the subsample of individuals that report to overwork of the analysis sample described in section 4.5. Each bar corresponds to an interval of width 5. For example: the first bar corresponds to 0-5 overtime hours worked, the second bar to 5-10 overtime hours worked and so on.

4.4 Confounders

I already introduced the first and most important confounder in section 4.3: hours worked excluding overtime hours. Leaving out this variable would lead to omitted variables bias in my estimates, prohibiting me from giving my estimates a causal interpretation.

To reach causality, I have to control for all time-varying confounders only, since the time-invariant confounders are accounted for by the fixed effects. I will thus also use data on confounders. Confounders are variables influencing the outcome and the treatment. In this case a confounder is a variable that simultaneously influences mental well-being and overtime hours worked.

Besides normal hours worked, I will control for age, job changes, job type, personal monthly net income, number of own dependent children in the household and marital status. I do not have a direct variable measuring job changes, but I use a variable indicating whether the respondent reports to work at the same workplace as the year before to control for workplace changes due to a change in employer. Changes in job functions within the same workplace are potentially also important, so to control for this I use a variable indicating the socio-economic group of the job (or simply job type), which is a categorical variable[9]. To control for marital status, I use cohabitation status as a proxy, since I do not observe the marital status of respondents directly.

To give some insights into why these variables are potential confounding my results, I will

[9]The classification of the socio-economic group of a job is based on the official UK NS-SEC coding, developed by the Office for National Statistics (2023).

give a brief overview on the relationships found in the literature between the control variables and the outcome and treatment variable in my study. The first is age. Steptoe et al. (2015) note in their study that well-being varies with age, and also that well-being follows a U-shape in age for high-income, English-speaking countries (e.g. the UK) with the lowest well-being in the age group 45-54 years. Besides, older workers are often more experienced and more often have a more managing function and thus are often the ones working longer hours, to lead the team on the work floor for example.

The second control variable is job changes. According to Björklund et al. (2013), an improvement in work motivation is significantly related to less exhaustion, and a decrease in work motivation is found to be significantly related to an increase in depressive feelings. Such a change in work motivation comes for example from changing jobs, and thus job changes potentially influence mental well-being. Changing job would in most cases also lead to a different number of overtime hours worked, since a different work arrangement would be made between employee and employer.

The third control variable included in my analyses is job type. It is straightforward that in some jobs or industries, overworking is more prevalent and more accepted compared to other jobs or industries. Less straightforward is the possibility that job type influences the mental well-being. One study by Bardasi and Francesconi (2004) finds evidence for lower job satisfaction among casual workers, which in turn impacts mental well-being.

To disentangle the partial effect of overworking on mental well-being, I control for personal monthly net income. It speaks for itself that more paid overtime hours lead to higher labour income and as a result higher net income. The relationship between income and mental well-being is maybe less obvious. The study by Kahneman and Deaton (2010) states that emotional well-being, which they define as *"the frequency and intensity of experiences of joy, stress, anger, and affection that make one's life pleasant or unpleasant"*, which is close to mental well-being, rises with income up until an annual income of $75,000[10]. Net income is thus an important mechanism and thus should be controlled for to derive the partial effect of overworking only on mental well-being.

The number of dependent children in the household is also a confounder. It seems quite natural that individuals with children, and especially those with a lot of children work in a job in which they have more control on the specific job arrangements, and therefore likely choose to work less overtime hours. Shields and Price (2005) find that number of children aged above 2 years have a negative association with psychological (or mental) well-being, also using the GHQ-12 as a measure of mental well-being. Furthermore, this research finds that marital status also is associated with mental well-being. It seems likely that married individuals work less hours, since they have someone around them to take care of and spend time with and thus attach a higher weight to leisure in their utility function, making marital status an important control.

[10]See also Pouwels et al. (2008).

4.5 Sample selection

The main sample I will use in my analyses is constructed as follows. I include all individual-wave observations for which I have data on all relevant variables mentioned in sections 4.2, 4.3 and 4.4. Individuals aged below 18 years are not allowed to work long overtime hours, so I will exclude these observations[11]. Since I focus on individuals in their prime working ages, I also exclude individuals aged between 18 and 24 years and individuals aged above 54 years. I also restrict my dataset to these individuals, because in this way I also exclude most likely all students[12]. I also exclude observations not working any hours, including overtime hours, because I focus on the working people in the survey. To control for workplace changes, I drop observations reporting to have changed workplace in some year[13]. In this way, there are no workplace changes in the sample.

4.5.1 Potential outliers and misreporting

In figure 4a and 5, we can see that the distributions of hours worked and overtime hours worked are right-skewed, so there might be potential outliers in the data. One approach is to set a maximum level of hours worked and overtime hours worked which seem reasonable from a practical perspective, and drop all observations from the sample which report a value above this threshold. I will not adopt this approach to potential outliers, since then I need to set some arbitrary threshold level of hours worked and overtime hours worked, which I want to avoid. Besides, there is reason to believe that these potential outliers just represent natural variation in the population, and thus should not be dropped.

The second potential reason for the observed extreme values documented in figure 4a and 5 is misreporting by respondents. It could be that some respondents misunderstood the question asking about normal hours worked per week, excluding overtime, and reported total number of hours worked including overtime, producing the extreme values in the distribution of normal hours worked per week. On the other hand, there are also some respondents reporting the same amount of normal hours and overtime hours worked, which could be possible but it could also be that respondents misunderstood the question and answered their normal weekly working hours twice. These concern 550 individual-wave observations. I just leave these in, since I do not know whether these double reported values represent their real values, or if they are misreported. These two potential threats are a limitation of my study, and more generally of survey data.

[11] According to ACAS (2022), teenagers in the UK aged below 18 are not allowed to work long hours.

[12] I do not have data on the current educational status of respondents. The best I thus can do in this case is assuming that all students are excluded with this restriction. It seems plausible that this is the case, since most full-time programs end when the student turns 21. Having students in the analysis sample is problematic, since students have less time to work, and also experience stress from studying. This would thus bias the results.

[13] I prefer to drop these observations rather than including an indicator for having changed workplace, since the construction of the variable on workplace changes does not allow me to perfectly control for these changes, while it does allow me to identify those that have changed workplace. I will include the job type indicator directly in my models. I also ran the same regressions without dropping these observations. The results remain qualitatively the same.

4.5.2 Accounting for the covid-19 pandemic

The worldwide covid-19 pandemic caused changes in demand for certain goods and services, and let to shutdowns of certain factories and other workplaces. The pandemic could thus influence the number of hours worked and more importantly, the number of overtime hours worked by the individuals in the analysis sample.

There have also been found indications that the pandemic caused lower levels of mental well-being. Research by Thorisdottir et al. (2021) for example found that among Icelandic adolescents, the pandemic caused an increase in depressive symptoms and worsened overall mental well-being. Failing to account for the pandemic could lead to misleading estimates. The waves corresponding to the covid-19 pandemic are 11 and 12, spanning the period 2020-2022.

A quick inspection of table 3 reveals no massive changes in the average mental well-being score, in the number of hours worked or in the overtime hours worked. The spread of these variables also did not change significantly, therefore the relative small changes in the mean for these variables is not caused by bigger variation at the top and bottom of the distribution. Therefore, I leave these waves in the sample. In section 7.2, I remove waves 11 and 12 to test the robustness of the results to the exclusion of these waves.

4.5.3 The sample

The resulting analysis sample consists of 89,913 individual-wave observations, containing 23,032 unique individuals, with an average of 3.9 observations per individual. The analysis sample spans the period 2009-2022. Descriptive statistics for the analysis sample are given in table 3.

We can see that in the analysis sample, the average reported mental well-being is 10.905, meaning that, on average, the respondent's self-reported mental well-being is quite good. We can also spot an upward trend in the mental well-being score, which indicates that the mental well-being of the respondents on average declined over the years. This downward trend in mental well-being is accompanied by an upward trend in the percentage of respondents working overtime and the average number of overtime hours worked in the sample, while the number of regular working hours has stayed relatively constant over time. This can be seen as a piece of evidence that the number of overtime hours worked could explain the decrease in the mental well-being, though this evidence is not sufficient to conclude that an increase in overtime hours worked cause lower mental well-being, as already noted in section 4.4. Personal monthly net income has been rising over the past years, standing at an average of £2,179 in 2022. The average age in the analysis sample also is relatively constant over time at around 41 years, as is the average number of own dependent children in the household at around 1 child. Over the years, only 18.6 percent lives together with a partner in the household, on average. Finally, the resulting analysis sample is 42.9 percent male, and 30.4 percent has a manual job. What is the most striking, is the significant downward trend in the proportion of individuals getting paid for their overtime hours worked. In wave 1, 42.5 percent of individuals are getting paid for their extra hours worked, while in wave 12 this decreases to 30.7 percent. Over the years, an average of 39 percent of individuals only were getting paid for their overwork.

Table 3: Descriptive statistics: main analysis sample

Variable	All	Wave 1	2	3	4	5	6	7	8	9	10	11	12
Outcome variable													
Mental well-being score	10.905	10.696	10.813	10.720	10.758	10.909	10.548	10.575	10.815	11.131	11.271	11.538	11.824
	(4.963)	(4.719)	(4.816)	(4.908)	(5.015)	(5.104)	(4.751)	(4.828)	(4.908)	(5.071)	(5.073)	(5.167)	(5.383)
Treatment variables													
Working overtime	.451	.423	.439	.433	.444	.459	.478	.462	.467	.469	.466	.461	.442
Overtime hours worked	8.292	8.185	7.968	7.806	7.979	8.191	7.996	8.023	8.560	8.670	9.102	8.476	9.381
	(7.772)	(7.325)	(6.905)	(6.359)	(6.801)	(7.266)	(6.625)	(7.190)	(7.992)	(8.753)	(9.776)	(8.907)	(10.579)
Control variables													
Hours worked	33.416	33.145	33.254	33.308	33.280	33.322	33.510	33.516	33.550	33.562	33.473	33.717	33.820
	(10.040)	(10.741)	(10.077)	(9.996)	(10.031)	(10.014)	(9.914)	(9.908)	(9.762)	(9.901)	(10.118)	(9.848)	(9.546)
personal monthly	1,873	1,651	1,712	1,763	1,790	1,838	1,905	1,935	1,955	1,994	2,045	2,109	2,167
net income	(1,286)	(1,097)	(1,061)	(1,159)	(1,094)	(1,281)	(1,376)	(1,569)	(1,282)	(1,319)	(1,351)	(1,459)	(1,466)
Age	41.088	39.590	40.903	40.970	41.123	41.175	41.406	41.303	41.437	41.504	41.593	41.657	41.739
	(8.178)	(8.287)	(8.097)	(8.040)	(8.007)	(8.042)	(8.091)	(8.118)	(8.144)	(8.184)	(8.305)	(8.364)	(8.316)
Number of dependent	1.091	1.056	1.077	1.087	1.103	1.110	1.121	1.122	1.118	1.090	1.085	1.072	1.056
children in the household	(1.067)	(1.069)	(1.049)	(1.053)	(1.068)	(1.072)	(1.084)	(1.084)	(1.075)	(1.072)	(1.069)	(1.054)	(1.059)
Cohabits	.186	.186	.177	.173	.178	.184	.188	.188	.190	.195	.196	.206	.183
Other characteristics													
Male	.429	.423	.422	.423	.427	.424	.430	.435	.442	.438	.430	.436	.435
Paid overtime	.390	.425	.412	.399	.394	.400	.422	.418	.385	.363	.349	.321	.307
Manual worker	.304	.320	.314	.315	.306	.304	.311	.314	.303	.301	.286	.274	.265
Observations	89,913	11,068	9,069	8,930	8,333	8,066	7,026	7,576	7,149	6,309	5,958	5,410	5,019

Notes: table 3 shows the mean of the variables listed for the full sample from section 4.5 in column 1, and for each wave separately in columns 2 until 13. Standard deviations of the continuous variables are shown in parentheses. The range of the continuous variables listed are as follows: the mental well-being score ranges from 0 to 36, overtime hours worked range from 0 to 96, normal hours worked range from 1 to 97.9, personal monthly net income ranges from 6 to 81,666.67, age ranges from 25 to 54 and number of children ranges from 0 to 8. The average overtime hours worked and proportion of paid overtime reported are conditional on overworking. The number of observations for the variable 'Overtime hours worked' and 'Paid overtime' is 40,572.

5 Methodology

5.1 Empirical strategy and theoretical model

The empirical strategy I will use to identify the impact of overworking on mental well-being follows the recommendations made by Ferrer-i Carbonell and Frijters (2004). They compare different studies using different methodologies to estimate a relationship between some explanatory variables and happiness, or more generally life satisfaction. They conclude that treating the outcome as cardinal or ordinal does not change the results significantly, but failing to account for time-invariant interpersonal heterogeneity does.

To take these time-invariant individual-specific factors into account, I will include individual fixed effects, and estimate a fixed effects model using OLS[14]. The inclusion of fixed effects in the regression to control for time-invariant individual heterogeneity is an approach often taken in the literature on well-being (e.g. Winkelmann and Winkelmann (1995); Booth and Van Ours (2008); Wooden et al. (2009). I will estimate the fixed effects model using OLS, since this allows for an easy interpretation of the coefficients.

I will conduct two distinct analyses: first I explore whether overworking in general leads to a different mental well-being outcome. Second, I explore the exact relationship between number of overtime hours worked and mental well-being. To this end, two distinct fixed effects regression models will be used. Theoretically, the regression specification of interest for the first analysis is given by equation 3:

$$MWS_{it} = \alpha_i + \beta * Overtime_{it} + \sum_j \delta_j * x_{j,it} + \gamma_t + \epsilon_{it} \tag{3}$$

Where MWS_{it} is the mental well-being score, as reported by the respondent, α_i is the individual fixed effect, absorbing the time-invariant individual-specific heterogeneity. The coefficient β is the coefficient of interest in my analysis, giving the average effect of working overtime on the mental well-being of employees. The term $Overtime_{it}$ indicates whether the respondent reports to work overtime ($Overtime_{it} = 1$) or not ($Overtime_{it} = 0$). The j included control variables as discussed in section 4.4 and their coefficients are given by the summation $\sum_j \delta_j * x_{j,it}$. The term γ_t is a time fixed effect, capturing the confounders which influence each individual in the sample in the same way, but differ over time, and ϵ_{it} is the error term. The subscript it refers to the observation for individual i in wave t.

For the second analysis, the theoretical regression specification is given by the following equation 4:

$$\begin{aligned} MWS_{it} = \alpha_i + \beta_1 * OvertimeHours_{it} + \beta_2 * OvertimeHours_{it}^2 \\ + \beta_3 * Overtime_{it} + \sum_j \delta_j * x_{j,it} + \gamma_t + \epsilon_{it} \end{aligned} \tag{4}$$

[14] Here, I am implicitly assuming cardinality of the mental well-being score, meaning that the relative difference between the points is perceived the same by everyone: a score of 6 is twice as bad as a score of 3. This is a quite restrictive assumption, so most economists assume therefore ordinality. In section 7.1, I will also let go of the cardinality assumption and assume the less restrictive ordinality assumption as a robustness check, by applying a fixed effects ordered logit model.

Where the meaning of the terms is the same as in equation 3, but now the term $OvertimeHours_{it}$ is the number of overtime hours worked, as reported by respondent i in wave t. To allow for possible nonlinearities in overtime hours, I added overtime hours squared[15]. It seems reasonable to assume that not every overtime hour has the same impact on mental well-being. For example, the tenth overtime hour could have a different impact on mental well-being than the third. I also included an indicator for overworking, as in equation 3 besides the number of overtime hours worked, to explicitly disentangle the effect of working an extra hour of overtime from the effect of working overtime itself.

In both equations 3 and 4, to allow for nonlinearities in the control variables, the term $\sum_j \delta_j * x_{j,it}$ also includes normal hours worked squared, age squared and the logarithm of personal monthly net income[16].

6 Results

This section presents the results of the main analyses in this thesis. First, I investigate whether the individuals reporting to work overtime also report a different level of mental well-being, compared to those not working overtime. Next, I present the results regarding the exact relationship between the number of overtime hours worked and the self-reported mental well-being. Then, I conduct two heterogeneity analyses: one exploring the heterogeneity of the treatment effect between manual workers and non-manual workers, and one exploring the heterogeneity of the treatment effect between paid and unpaid overtime.

6.1 Does working overtime affect mental well-being?

The first relevant question this subsection tries to answer is: does working overtime in general lead to a different mental well-being outcome? Table 4 below reports the results. As already mentioned, Ferrer-i Carbonell and Frijters (2004) stated that accounting for individual fixed effects changes the results significantly in most cases, when researching the relationships between certain variables and well-being. To see whether this also is the case in this study, I show the results of a pooled OLS regression model in columns 1 until 3. The results for the fixed effects model are shown in columns 4 until 6. The coefficient of the normal hours worked is also reported, because it is interesting to see whether these correspond to the findings in earlier research[17].

The results for this subsection are reported in columns 1 and 4. As a first observation, we can

[15] Higher-order polynomials are all not statistically significant either, so there need not be checked for different specifications.

[16] The number of own dependent children in the household is also a continuous variable, and thus there could also be a nonlinearity in the relationship between mental well-being and number of own dependent in the household. Since most individuals report having 1 or 2 children however, the relevant range of this variable is rather small and thus a linear term is preferred. Moreover, different nonlinear specifications in this variable are not statistically significant.

[17] Earlier research has found nonlinear relationships between hours worked and well-being. Therefore, I also control for and report the squared term of normal hours worked.

clearly see that the coefficient of interest in the first row changes significantly when accounting for individual fixed effects. The estimated coefficient in the pooled OLS model is almost triple the estimate in the fixed effects model. The estimated coefficient in column 4 is positive and statistically significant, which means that those working overtime, on average report a lower mental well-being than those not working overtime, though the estimated coefficient is rather small, standing at an average increase in the mental well-being score of 0.101 points for those working overtime.

Considering the scale of the mental well-being score from 0 to 36, this increase corresponds to 0.3 percent of the total range of the score. If we compare the estimate with the average within-individual standard deviation of the mental well-being score of 3.254, the estimate is 3.1 percent of the average within-individual variation in the mental well-being score. This first result thus shows that there is a statistically significant negative effect of working overtime on self-reported mental well-being, but the effect size and thus economic significance is limited.

Table 4: Results: working overtime

	Mental well-being					
	Pooled OLS			**Fixed effects**		
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Overtime	.294***	.112**	.077	.101**	-.038	-.115*
	(.046)	(.054)	(.067)	(.041)	(.048)	(.060)
Overtime hours		.023***	.030***		.020***	.036***
		(.004)	(.010)		(.004)	(.009)
Overtime hours squared			-.0002			-.0004**
			(.0002)			(.0002)
Hours worked	-.014	-.012	-.012	-.009	-.010	-.010
	(.008)	(.008)	(.008)	(.009)	(.009)	(.009)
Hours worked squared	-.0001	-.0001	-.0001	.0002**	.0003**	.0003**
	(.0001)	(.0001)	(.0001)	(.0001)	(.0001)	(.0001)
(Within) R^2	0.0110	0.0116	0.0116	0.0113	0.0118	0.0119
Individuals	23,032	23,032	23,032	23,032	23,032	23,032
Observations	89,913	89,913	89,913	89,913	89,913	89,913

Notes: table 4 shows the estimates for equation 3 in columns 1 and 4, and for equation 4 in columns 2, 3, 5, and 6, using the full analysis sample from section 4.5. Robust standard errors clustered at the individual level are given in parentheses. In all models, the dependent variable is the mental well-being score, which ranges from 0 to 36. A higher score means a lower mental well-being. The columns 1 until 3 show the results of the pooled OLS regression, the columns 4 until 6 show the results of the individual fixed effects regression. The controls included in all models are: normal hours worked, normal hours worked squared, age, age squared, logarithm of personal monthly net income, number of own dependent children in the household, cohabitation status, job type and a time fixed effect. $*: p < .10; ** : p < .05; *** : p < .01$.

6.2 The relationship between overtime hours worked and mental well-being

This section extends the previous analysis by looking at the exact relationship between overtime hours worked and mental well-being. The results for this analysis are reported in columns 2, 3, 5 and 6 of table 4. The coefficient of the quadratic term is statistically significant, so there seems to be a nonlinear, quadratic relationship between overtime hours worked and mental well-being. The right specification is thus the quadratic one, reported in columns 3 and 6.

A comparison between the pooled OLS model in column 3 and the fixed effects model in column 6 reveals that if individual time-invariant heterogeneity is accounted for, the coefficient of overtime hours increases by 0.006, or 20 percent, and the squared term becomes statistically significant. The estimated coefficient of overtime hours in column 6 is statistically significant and positive, while the point estimate for the quadratic term is statistically significant and negative, which means that the number of overtime hours worked negatively impacts the mental well-being of individuals, but at a decreasing rate.

The exact average change in the mental well-being score is $.036 - .0008 * OvertimeHours$, thus the average change is dependent on the initial level of overtime hours worked[18]. The estimated relationship between overtime hours and mental well-being is plotted in figure 6 below, together with the linear model in column 5 for comparison. Surprisingly, the first few hours of overtime worked lead to the greatest decrease in mental well-being, according to the estimates of column 6.

An explanation for this could be that the first few hours of overtime working are experienced as more unpleasant since they could be unexpected and be experienced as a 'loss', but after some hours of overtime the past workweeks, the employees adapt to the situation. The exact estimates in column 6 imply for example that on average, for an individual working the fifth hour of overtime, the mental well-being score increases by .033 points, which is 1 percent of the average within-individual standard deviation of the mental well-being score.

The estimated effect size is thus rather small, but it seems that overtime hours have a negative impact on mental well-being, and that this relationship is nonlinear. The most striking result however, is that overworking itself does not lead to lower mental well-being, but seems to impact mental well-being positively. Theoretically, this could be explained by the fact that overworking itself is voluntary and thus increases well-being (utility) via mental well-being in the leisure-labour choice model[19]. The first few hours of overtime would then in sum have a positive impact on mental well-being compared to not overworking, but when overtime hours become excessive, mental well-being decreases. Though, the coefficient is not statistically significant at the 5% level, and thus should be interpreted with caution.

The most important takeaway from this analysis is that overworking per se seems to not significantly lead to different mental well-being outcomes, but a greater number of hours overtime

[18] The change is simply the partial derivative of the estimated equation 4 with respect to overtime hours worked: $\frac{\partial MWS(.)}{\partial OvertimeHours}$.

[19] Of course, the assumption behind this reasoning is that an increase in mental well-being increases utility, and does not lead to a decrease in utility via other channels influencing total utility, or well-being. This seems reasonable, since mental well-being includes for example happiness, which is a rather big part of total well-being. In classical work, utility is often defined as happiness.

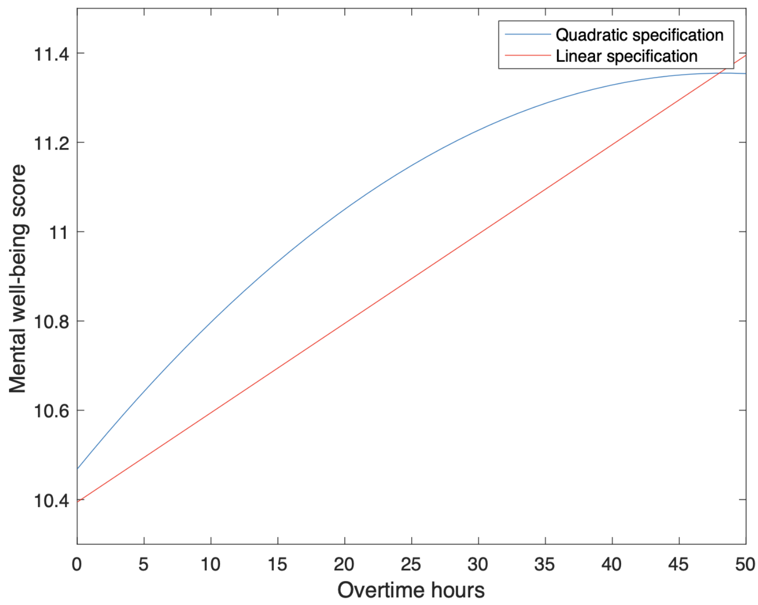


Figure 6: Estimated relationship between overtime hours and mental well-being

Notes: figure 6 plots the estimates for equation 4 from table 4, columns 5 and 6. The x-axis represents the number of overtime hours worked, and the y-axis represents the mental well-being score. A higher mental well-being score corresponds to a lower mental well-being.

working does lead on average to a lower mental well-being on the relevant range.[20]

As an interesting side result, every normal hour worked has a very small, negative impact on mental well-being. For example, 40 weekly hours worked leads to an increase in the mental well-being score of 0.48, on average, compared to not working.

As a final observation from the above analyses, we can see that the explanatory power of my models is quite low, namely only between 1.10% and 1.19% of the total within-individual variation in mental well-being is explained by overworking and the controls. This means that there are many other variables explaining the mental well-being of employees. This has some implications for predominantly policy, which are further discussed in section 8.1.

6.3 Exploring the heterogeneity of the relationship

In this section, I conduct two heterogeneity analyses, to explore possible heterogeneous effects of overworking on mental well-being, which gives more insight into the potential explanations of the documented impact found in the previous sections. The first heterogeneity analysis looks at the potentially differential impact of overworking between two job types: manual and non-manual jobs, and the second analysis is concerned with uncovering the differential impact of paid and unpaid overwork.

In conducting these analyses, one could split the sample or simply add an interaction term for

[20]The relevant range in this case corresponds to the range of number of overtime hours worked which seem reasonable from a practical perspective. Of course, this is not a waterproof definition, but it could be regarded as plausible that the estimated models all have an interpretation on the interval [0, 30] of overtime hours worked. There are only 844 observations working more than 30 hours of overtime, so the estimates become noisier on the interval $(30, \infty)$.

job type and paid overtime, respectively. For the first heterogeneity analysis, I split the sample between manual and non-manual workers, since all variables, including the controls, could impact individuals working in a manual and non-manual job differently, since these different job types involve different tasks and skills. For the second analysis, I simply add an interaction term, indicating paid overtime, since I do not expect the control variables to differently impact individuals working paid and unpaid overtime.

6.3.1 Heterogeneity analysis: manual versus non-manual workers

Until now, I have not distinguished between different job types, especially between manual and non-manual jobs. This distinction is a natural one, since the capabilities and tasks differ greatly between these two types of jobs. Manual jobs mostly involve (simple) repetitive tasks, where most non-manual jobs are more mentally demanding and could more frequently be experienced as stressful. A study by Robone et al. (2011) also using the 12-item GHQ as measure of mental well-being underlines this, finding an overall significant negative impact of overtime hours on mental well-being and a significant negative impact on mental well-being for non-manual workers, while finding no significant relationship between overtime hours worked and mental well-being for manual workers.

The results of this analysis are given in table 5 below. In panel A, the separate results are shown for the subsample working in a manual job, and in panel B for the ones working in a non-manual job. The following job types are categorized as manual: foreman manual, skilled manual worker, semi-skilled manual worker, unskilled manual worker, agricultural workers and personal service workers, and the following as non-manual: junior non-manual, intermediate non-manual workers, managers and professional employees.

In column 1, the same question as in subsection 6.1 is investigated. As can be seen from panel A in the table, the point estimate in column 1 is not statistically distinguishable from zero. Individuals in a manual job working overtime thus do not report a significantly different level of mental well-being, compared to those not working overtime. In contrast, individuals in a non-manual job working overtime report a significantly different level of mental well-being, compared to those not working overtime. The point estimate in column 1 of panel B is 0.184, which means that the ones working overtime on average have a 0.184 point lower mental well-being, compared to those not working overtime.

In columns 2 and 3, the same question as in subsection 6.2 is investigated for the two separate subsamples. In column 2 of panel A, the coefficient of overtime hours is not statistically significant, and also the quadratic term in column 3 is not statistically significant, so there also seems to be no impact of the number of overtime hours on mental well-being for those with a manual job. In column 3 of panel B, we can see that there seems to exist a nonlinear relationship between mental well-being and overtime hours worked for non-manual workers.

For non-manual workers, the estimates of the linear and quadratic terms in column 3 of panel B are both statistically significant, but the estimated coefficient of the linear term is positive and of the quadratic term negative. The change in the mental well-being of an individual equals $0.050 - 0.0012 * OvertimeHours$ and is thus dependent on the initial value of the number of overtime hours worked. Also, the coefficient of 'Overtime' in column 3 of panel B is not statistically significant, so it seems that overworking per se has no impact on mental well-being, as before.

To give a bit more insight into the relationship, the estimated parabola from column 3 is plotted in figure 7 below. As before documented in section 6.2, the first few hours result in the greatest decrease in mental well-being, whereas when more overtime hours are worked, the slope becomes less steep.

To sum up, there seems to be a negative nonlinear impact of overtime hours worked on mental well-being for those that have a non-manual job, where there does not seem to be an effect on mental well-being for those that have a manual job. This is in line with the findings by Robone et al. (2011).

In light of Lorente et al. (2018), this could be explained by the possibility that an employee's job satisfaction and/or job meaningfulness differ between the two job types. This difference could stem from differences in work characteristics between the two job types, such as the amount of job strain. As a consequence, manual workers could have a higher level of job satisfaction and/or job meaningfulness compared to non-manual workers and therefore these mediating factors mute the negative impact of overtime hours on mental well-being.

One could argue that the absence of an effect of overworking on mental well-being among manual workers could be due to the smaller sample size and as a consequence loss of statistical power. This could be true, since the point estimates in panel A are not precisely 0, but have a rather wide confidence interval. However, since the absolute values of the point estimates in panel A of table 5 are a lot smaller in magnitude than the same relevant estimates in panel B for non-manual workers, I can conclude that there at least is a relatively smaller effect of overworking on mental well-being for manual workers, if there is an effect of overworking for manual workers, which in itself already is an insightful result.

Table 5: Results manual vs non-manual job: working overtime

	Mental well-being		
	Fixed effects		
Variable	(1)	(2)	(3)
Panel A: manual job			
Overtime	-.073	-.143	-.086
	(.075)	(.094)	(.112)
Overtime hours		.010	-.002
		(.008)	(.015)
Overtime hours squared			.0003
			(.0004)
Hours worked	-.015	-.015	-.015
	(.016)	(.016)	(.016)
Hours worked squared	.0003	.0003	.0003
	(.0002)	(.0002)	(.0002)
(Within) R^2	.0074	.0075	.0075
Individuals	8,480	8,480	8,480
Observations	27,343	27,343	27,343
Panel B: non-manual job			
Overtime	.184***	.031	-.099
	(.050)	(.057)	(.071)
Overtime hours		.022***	.050***
		(.004)	(.010)
Overtime hours squared			-.0006***
			(.0002)
Hours worked	-.004	-.007	-.007
	(.012)	(.012)	(.012)
Hours worked squared	.0002	.0003	.0003*
	(.0002)	(.0002)	(.0002)
Within R^2	.0128	.0135	.0137
Individuals	16,016	16,016	16,016
Observations	62,570	62,570	62,570

Notes: table 5 shows the estimates for equation 3 in column 1 and for equation 4 in columns 2 and 3 using the full analysis sample from section 4.5. Panel A does this for the subsample working in a manual job, panel B does this for the subsample working in a non-manual job. Robust standard errors clustered at the individual level are given in parentheses. In all models, the dependent variable is the mental well-being score, which ranges from 0 to 36. A higher score means a lower mental well-being. The following job types are considered as manual: foreman manual, skilled manual worker, semi-skilled manual worker, unskilled manual worker, agricultural workers and personal service workers. The following job types are considered as non-manual: junior non-manual, intermediate non-manual workers, managers and professional employees. The controls included in all models are: normal hours worked, normal hours worked squared, age, age squared, logarithm of personal monthly net income, number of own dependent children in the household, cohabitation status, job type and a time fixed effect. $^* : p < .10; ^{**} : p < .05; ^{***} : p < .01$.

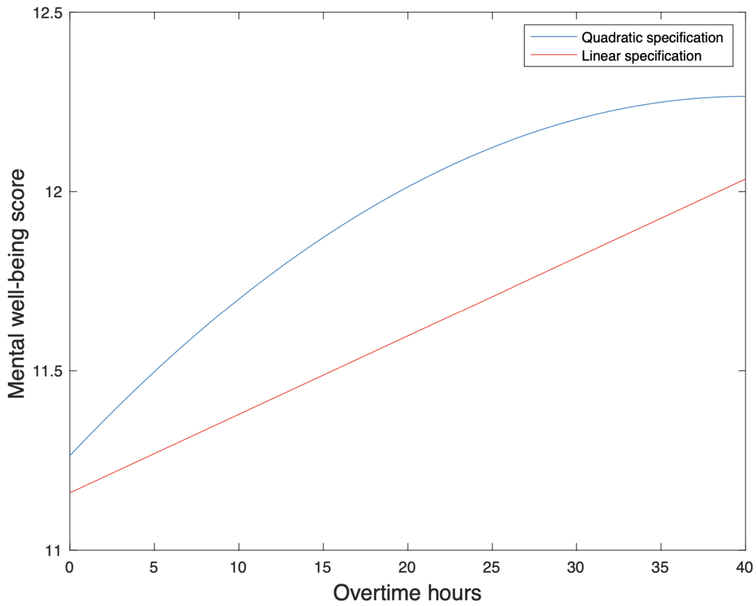


Figure 7: ESTIMATED RELATIONSHIP BETWEEN OVERTIME HOURS AND MENTAL WELL-BEING

Notes: figure 7 plots the estimates for equation 3 from table 5, columns 2 and 3 in panel B, using the subsample working in a non-manual job. The x-axis represents the number of overtime hours worked, and the y-axis represents the mental well-being score. A higher mental well-being score corresponds to a lower mental well-being.

6.3.2 Heterogeneity analysis: The differential impact of paid and unpaid overtime

I have already made a distinction between manual and non-manual jobs in the last section. The second logical distinction we can make is between paid and unpaid overtime. It could be the case that individuals experience the overtime hours as much less stressful when knowing they are rewarded for it. The study by Beckers et al. (2008) already presents evidence on the differential impact of compensated and uncompensated overworking.

To capture the potential differential impact of paid and unpaid overtime hours, I will include an interaction term in equation 4, indicating whether the overtime hours were paid. This interaction term is a binary variable, which equals 1 if all overtime hours are paid, and 0 if all overtime hours are unpaid[21].

The regression model I will estimate for this analysis is the following:

$$\begin{aligned} MWS_{it} = \alpha_i + \beta_1 * OvertimeHours_{it} + \beta_2 * OvertimeHours_{it}^2 + \beta_3 * Overtime_{it} \\ +\beta_4 * PaidOvertime_{it} + \beta_4 * OvertimeHours_{it} * PaidOvertime_{it} \\ +\beta_5 * OvertimeHours_{it}^2 * PaidOvertime_{it} + \sum_j \delta_j * x_{j,it} + \gamma_t + \epsilon_{it} \end{aligned} \tag{5}$$

Where the indicator $PaidOvertime_{it}$ is equal to 1 of individual i in wave t is working paid overtime, and equals 0 otherwise.

[21]There are 945 observations with a fraction of the total number of overtime hours being paid overtime hours. Because there are relatively little observations with just a fraction of paid overwork, I remove these observations and use a binary indicator instead.

The results of this analysis are given in table 6 below. The coefficients of interest in column 2 are all statistically significant, except for overtime, which again means that overworking itself has no significant impact on mental well-being. An interesting result from this table is that the effect of overtime hours seems to almost disappear when these hours are paid, but there is a clear negative effect of working an extra overtime hour on mental well-being when these are unpaid.

Table 6: Results paid versus unpaid overtime

	Mental well-being	
	Fixed effects	
Variable	(1)	(2)
Overtime	-.196***	-.121
	(.074)	(.083)
Overtime hours	.025***	.059***
	(.004)	(.012)
Paid overtime	-.320***	-.087
	(.090)	(.108)
Overtime hours ∗ Paid	-.015	-.067***
	(.009)	(.017)
Overtime hours squared		-.0007***
		(.0002)
Overtime hours squared ∗ Paid		.0013***
		(.0005)
Within R^2	.0125	.0128
Individuals	22,964	22,964
Observations	88,845	88,845

Notes: table 6 shows the estimates for equation 5, using the subsample from the full analysis sample from section 4.5 that reports to overwork. Robust standard errors clustered at the individual level are given in parentheses. In all models, the dependent variable is the mental well-being score, which ranges from 0 to 36. A higher score means a lower mental well-being. The controls included in all models are: normal hours worked, normal hours worked squared, age, age squared, logarithm of personal monthly net income, number of own dependent children in the household, cohabitation status, job type and a time fixed effect. $^{*}: p < .10; ^{**}: p < .05; ^{***}: p < .01$.

A striking result is that the sign of the estimated coefficient of the number of overtime hours worked and the squared term flips if these overtime hours are paid, compared to when these are unpaid. For example, from column 2 of table 6, we can see that on average, the quadratic relationship when the overwork is unpaid is summarized by the parabola $.059 * OvertimeHours - .0007 * OvertimeHours^2$, whereas this relationship when overwork is paid becomes $-.008 * OvertimeHours + .0006 * OvertimeHours^2$. The relationship thus changes in shape. The squared term in column 2 is significant, so it seems indeed there exists a nonlinear relationship, which is not surprising since we have already found this nonlinearity in the main analysis of section 6.2.

Figure 8 plots the estimated equation from column 2. Clearly, the effect size for paid overtime hours is much smaller on the relevant range, and positive for the first few hours. It looks like rewarding employees for their extra effort put in can for a big part help preventing a negative impact of the number of overtime hours on mental well-being, though too much compensated overwork would harm employees. This is a remarkable result, supporting the finding by Beckers et al. (2008).

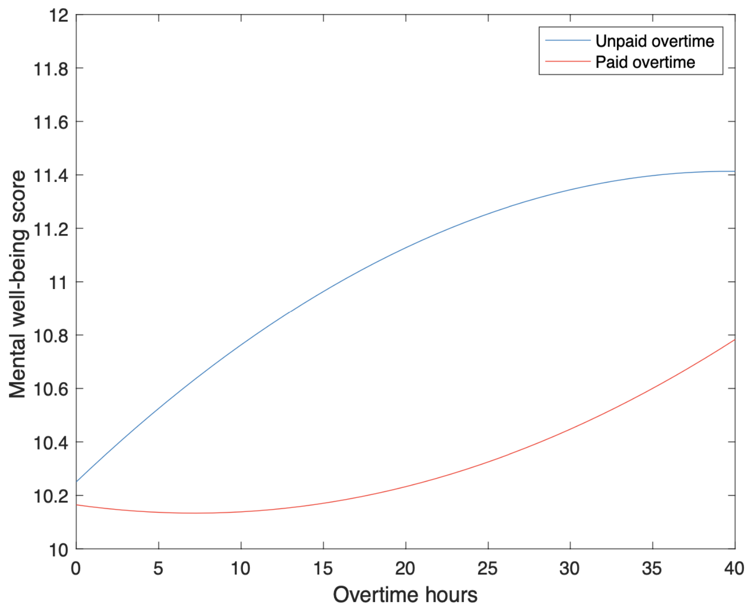


Figure 8: Estimated relationship between overtime hours and mental well-being

Notes: figure 8 plots the estimates for equation 3 from table 6, column 4, using the subsample working overtime. The x-axis represents the number of overtime hours worked, and the y-axis represents the mental well-being score. A higher mental well-being score corresponds to a lower mental well-being.

7 Robustness

In my main analyses from the previous sections, I have made some restrictive assumptions and drawn conclusions based on the results under those assumptions. One of these assumptions is the assumption of cardinality of the mental well-being score. As a first robustness test, I will let go of this assumption and instead assume ordinality of the mental well-being score. In my main analyses, I also made the assumption that the covid-19 pandemic did not influence the individuals in my study, at least not differently. As a second robustness test, I will relax this assumption. In section 4.2.1, I concluded based on Cronbach's alpha that the 12 items together measure one underlying latent factor: mental well-being. As a last test I will conduct the same analysis on a random subset of these items.

7.1 The ordinal approach

The first robustness test in this section lets go of the cardinality assumption, and instead assumes ordinality[22]. To this end, I will use a collapsed version of the mental well-being score used up until now: the GHQ-12 caseness score, which is constructed in a similar way, but ranges from 0 to 12[23].

The individual fixed effects model estimated using OLS assumes cardinality. The fixed effects ordered logit model proposed by Baetschmann et al. (2015) assumes ordinality, rather than cardinality[24]. The results for this robustness test are shown in table 7.

If my results are robust to the cardinality assumption, we should see the same relationships between overtime hours and mental well-being in the table. The interpretation of the coefficients is quite cumbersome, since the reported coefficients represent changes in the odds ratio of having a higher mental well-being score, as opposed to a lower mental well-being score. I deliberately omit the description of what a change to the odds ratio means, since that is not very important in understanding the coefficients. The most important is the sign, magnitude, and significance of the coefficients.

If we compare tables 4, 5 and 6 with table 7, we can see that the significance of the coefficients is largely unchanged, which is a good sign. There still seems to be an impact of working overtime on mental well-being, especially for those working in a non-manual job, and no effect for those working in a manual job. Moreover, the conclusion that overworking itself has no impact on mental well-being remains valid.

Though not one-for-one comparable, the magnitude of the coefficients also resemble the same order of impact of working overtime or working an extra hour of overtime on mental well-being. For example, the estimated coefficient of overtime hours in column 2 row 2 of 1.012 means that the odds ratio of having a higher mental well-being score (and thus a lower mental well-being) rather than a lower mental well-being score (and thus a higher mental well-being) increases by 1.2 percentage points on average if working an extra hour of overtime, which indicates that there is a small negative effect of an extra hour of working overtime on mental well-being, as I concluded in section 6.2. Similarly, a coefficient smaller than 1 means a decrease in the odds ratio. For example: the estimated coefficient of paid overtime hours in column 10 of .906 means that the odds ratio decreases by 9.4 percentage points on average when an individual works overtime, compared to those not working overtime.

We can also see that the relationship between overtime hours and mental well-being changes in shape in the same way as documented earlier when overtime hours are paid, compared to unpaid overtime hours. I conclude that my results in section 6 are robust to the cardinality

[22]The ordinality of the mental well-being score means that the mental well-being of an individual cannot be explicitly expressed as a number, but only the ordering is meaningful: a mental well-being score of 6 is perceived as worse than 3, but not necessarily twice as bad.

[23]I use the GHQ-12 caseness score instead of the mental well-being score earlier used, since fitting a fixed effects ordered logit model using a dependent variable with 37 categories is computationally very inefficient. Therefore, since the caseness score is constructed in nearly the same way using the same data and 'only' has 13 outcome levels, I use this variable for fitting this model.

[24]For technical details on this model, the odds ratio and the estimator, see Baetschmann et al. (2020, 2015)

assumption.

7.2 The potential impact of the covid pandemic

In section 4.5.2 I already discussed the potential impact of the covid-19 pandemic on both mental well-being and overtime hours worked. I have assumed that the pandemic did not impact the individuals in my study differently. If this would be the case, then my results would not be valid.

To test the robustness of my results to this assumption, I will drop waves 11 and 12 from the data and re-run the same analyses. The results of these analyses are shown below in table 8. If we compare the tables 4, 5, 6 with table 8, we can see that the coefficients virtually do not change, as well as their significance.

In conclusion, this means that my main results are not affected by the covid-19 pandemic, and therefore conclude that my results are robust to the potential influences of the covid-19 pandemic, and thus are not driven by the pandemic.

7.3 Testing the consistency of the GHQ-12 questionnaire

The last robustness check is concerned with the consistency of the measure for mental well-being I use in my thesis: the 12-item General Health Questionnaire. We have already seen that according to Cronbach's alpha, the items together are consistently measuring one underlying variable: mental well-being. On the basis of this result, we would expect to see the same relationships between overtime hours and mental well-being when running the main analyses on a subset of these items.

The items are listed in an arbitrary order in table 1, so I will simply pick the 6 items on top of the table: concentration, loss of sleep, usefulness, decision making, feeling under strain and ability to overcome difficulties. I will construct an aggregated score in the same way as the dependent variable in the main analyses described in section 4.2. To restore the same scale range as the 12-item aggregated score for comparability, an individuals score is multiplied by 2. This shortened 6-item score thus also ranges from 0 to 36.

The results of this robustness test are given in table 9. Reassuringly, the table mostly tells the same story as the results from the main analyses. Most conclusions from the main analyses still hold: there seems to be an effect of overtime hours on mental well-being, where this effect is mostly present under the ones with a non-manual job, and the ones who work unpaid overtime.

A significant difference with table 4 is that overworking itself seems to significantly impact mental well-being at the 5% level. Still, the conclusion I can draw from this is that most results are not driven by some items in the questionnaire, but rather all items measure mental well-being consistently, strengthening the assumption that the GHQ-12 questionnaire is a reliable measure of mental well-being.

Table 7: Results of robustness tests: ordinal approach

	Sample										
	All			Manual job			Non-manual job			Paid vs unpaid	
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Overtime	1.065**	.981	.942	.989	.952	.995	1.102***	1.007	.938	.906**	.935
	(.028)	(.030)	(.034)	(.053)	(.060)	(.073)	(.035)	(.036)	(.041)	(.045)	(.052)
Overtime hours		1.012***	1.021***		1.006	.996		1.013***	1.028***	1.014***	1.032***
		(.002)	(.005)		(.005)	(.009)		(.002)	(.006)	(.002)	(.007)
Paid overtime										.854***	.957
										(.049)	(.065)
Overtime hours			.9998**			1.0002			.9997***		.9996***
squared			(.0001)			(.0002)			(.0001)		(.0001)
Overtime hours*										.992	.968***
Paid										(.006)	(.010)
Overtime hours											1.0006**
squared*Paid											(.0002)
Pseudo R^2	.0077	.0084	.0085	.0071	.0072	.0073	.0085	.0094	.0097	.0091	.0093
Individuals	12,124	12,124	12,124	3,796	3,796	3,796	8,503	8,503	8,503	12,024	12,024
Observations	66,589	66,589	66,589	18,016	18,016	18,016	46,985	46,985	46,985	65,564	65,564

Notes: table 7 shows the odds ratios from the fixed effects ordered logit model for equation 3 in column 1, 4 and 7, for equation 4 in column 2, 3, 5, 6, 8, and 9, and for equation 5 in column 10 and 11, using the full analysis sample from section 4.5. Robust standard errors clustered at the individual level are given in parentheses. In all models, the dependent variable is the mental well-being score, which ranges from 0 to 36. A higher score means a lower mental well-being. The controls included in all models are: normal hours worked, normal hours worked squared, age, age squared, logarithm of personal monthly net income, number of own dependent children in the household, cohabitation status, job type and a time fixed effect. * : $p < .10$; ** : $p < .05$; *** : $p < .01$.

Table 8: Results of robustness tests: removing the pandemic

	Sample										
	All			Manual job			Non-manual job			Paid vs unpaid	
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Overtime	.104**	-.049	-.123*	-.070	-.164*	-.125	.196***	.031	-.096	-.180**	-.099
	(.044)	(.051)	(.063)	(.078)	(.096)	(.117)	(.054)	(.061)	(.076)	(.076)	(.089)
Overtime hours		.022***	.038***		.014	.005		.024***	.052***	.029***	.062***
		(.004)	(.009)		(.008)	(.017)		(.005)	(.011)	(.005)	(.013)
Paid overtime										-.272***	-.037
										(.093)	(.116)
Overtime hours squared			-.0004*			.0002			-.0006**		-.0007**
			(.0002)			(.0004)			(.0002)		(.0003)
Overtime hours*Paid										-.018*	-.072***
										(.010)	(.020)
Overtime hours squared*Paid											.0014***
											(.0006)
(Within) R^2	.0056	.0062	.0063	.0043	.0044	.0045	.0069	.0078	.0080	.0070	.0073
Individuals	22,155	22,155	22,155	8,103	8,103	8,103	15,300	15,300	15,300	22,085	22,085
Observations	79,484	79,484	79,484	24,532	24,532	24,532	54,952	54,952	54,952	78,603	78,603

Notes: table 8 shows the estimates for equation 3 in column 1, 4 and 7, for equation 4 in column 2, 3, 5, 6, 8, and 9, and for equation 5 in column 10 and 11, using the full analysis sample from section 4.5. Robust standard errors clustered at the individual level are given in parentheses. In all models, the dependent variable is the mental well-being score, which ranges from 0 to 36. A higher score means a lower mental well-being. The controls included in all models are: normal hours worked, normal hours worked squared, age, age squared, logarithm of personal monthly net income, number of own dependent children in the household, cohabitation status, job type and a time fixed effect. $* : p < .10; ** : p < .05; *** : p < .01$.

Table 9: RESULTS OF ROBUSTNESS TESTS: 6-ITEM MENTAL WELL-BEING SCORE

	Sample										
	All			**Manual job**			**Non-manual job**			**Paid vs unpaid**	
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Overtime	.119***	-.039	-.128**	-.110	-.196**	-.176	.224***	.050	-.086	-.203***	-.162*
	(.042)	(.049)	(.060)	(.077)	(.097)	(.114)	(.052)	(.059)	(.073)	(.076)	(.084)
Overtime hours		.022***	.041***		.012	.008		.025***	.054***	.028***	.065***
		(.004)	(.009)		(.009)	(.015)		(.004)	(.011)	(.004)	(.012)
Paid overtime										-.346***	-.132
										(.092)	(.110)
Overtime hours			-.0004**			.0001			-.0007***		-.0008***
squared			(.0002)			(.0003)			(.0002)		(.0003)
Overtime hours*										-.019**	-.066***
Paid										(.009)	(.017)
Overtime hours											.0012***
squared*Paid											(.0004)
(Within) R^2	.0076	.0082	.0083	.0047	.0048	.0049	.0092	.0101	.0103	.0091	.0094
Individuals	23,032	23,032	23,032	8,480	8,480	8,480	16,016	16,016	16,016	22,964	22,964
Observations	89,913	89,913	89,913	27,343	27,343	27,343	62,570	62,570	62,570	88,845	88,845

Notes: table 9 shows the estimates for equation 3 in column 1, 4 and 7, for equation 4 in column 2, 3, 5, 6, 8, and 9, and for equation 5 in column 10 and 11, using the full analysis sample from section 4.5. Robust standard errors clustered at the individual level are given in parentheses. In all models, the dependent variable is the 6-item mental well-being score, which ranges from 0 to 36. The items included are: concentration, loss of sleep, usefulness, decision making, feeling under strain and ability to overcome difficulties. A higher score means a lower mental well-being. The controls included in all models are: normal hours worked, normal hours worked squared, age, age squared, logarithm of personal monthly net income, number of own dependent children in the household, cohabitation status, job type and a time fixed effect. * : $p < .10$; ** : $p < .05$; *** : $p < .01$.

8 Discussion

So far, I have only interpreted the sign, significance and magnitude of the estimated coefficients and drawn some preliminary conclusions on the relationships between overtime hours worked and mental well-being. I have already discussed the relative small size of the point estimates and their sign. These findings are both in line with the expectations. The small effect size seems reasonable, since working some hours overtime does most likely not result in a mental illness, but nevertheless has a negative impact on mental well-being, especially when working a big amount of overtime hours. A mental illness often is the consequence of a simultaneous interplay of many factors, one of which is intensity of overworking in hours.

In the remainder of this section, I will discuss some possible (economic) mechanisms and their policy implications. I will end this section with a final note on the limitations of my study.

8.1 The negative effect of working overtime on mental well-being: the potential mechanisms

There is quite some concensus on the sign of the effect of overtime hours worked on mental well-being, as already outlined in section 3.3, and my study adds to this existing concensus with new empirical evidence. The question remains what could drive this negative relationship. There could be multiple, not mutually exclusive explanations for this evidence. A number of explanations follow from the results in this thesis.

A first explanation is the level of stress overtime working can cause to employees. This explanation is especially relevant, since my results indicate that non-manual workers experience worse mental well-being from working overtime. These non-manual jobs include most mentally demanding jobs. When putting in a lot of effort and constantly being under strain, this can cause larger levels of distress as research by Beckers et al. (2008) already stresses.

More overtime working also comes at the expense of less leisure time and thus less social activities with family and friends and work-family conflicts, which could lead to feelings of loneliness and depression, as Lelkes (2006) and Grant-Vallone and Donaldson (2001) also conclude. This explanation coincides with the leisure-labour choice model. If an employee works too many hours/too little leisure compared to his bliss point, this would decrease utility.

Also, acknowledgement by employers of the extra effort put in by employees via working overtime by paying them these overtime hours seem to weaken the negative impact of overworking on mental well-being. When getting pecuniary acknowledgement of their effort put in, and their willingness to sacrifice extra leisure for work, employees may feel satisfied with being helpful, suppressing their exhausted and stressed feeling. When sacrificing too much leisure, this could be detrimental for their feelings of stress and exhaustion. When not getting this pecuniary acknowledgement, this could be detrimental for employees' feelings of stress and exhaustion.

These potential explanations for the documented negative impact of overtime working on mental well-being are useful for policy, and some policy implications follow from my results and their potential explanations. Following from the result that non-manual workers experience a negative impact of working overtime on mental well-being, as well as those that do not get rewarded for their extra effort by their employer, a potential policy to enhance mental well-being under employees could be to tax overtime hours in some jobs more heavily, especially the mentally demanding jobs, and to reduce the number of overtime hours left unpaid by employers. This makes letting employees working overtime less attractive for employers, and thus likely helps reducing total overtime hours worked by their employees, and helps to reduce mental illnesses among employees.

I have implicitly assumed that every employee works overtime involuntarily. This is true from a theoretical perspective of the leisure-labour choice model, since every extra overtime hour negatively impacts mental well-being, and thus brings them further away from their optimum, on average. This need not be the case for everyone. Still, when an employee voluntarily chooses to work excessive amounts of overtime for future career gains, this is not desirable from a social perspective, since mental illnesses are a cost to society, and policy should be aimed at improving employee well-being, partly by aiming at making overworking less attractive for employers, regardless of employees wanting to work excessive amounts of overtime.

To significantly improve overall mental well-being, these policies should be accompanied by other mental well-being improving policies, since the impacts of working overtime are present but not very big, as my results suggest. The power of my models in explaining changes in mental well-being is quite low as already mentioned in section 6.2, which also confirms the conclusion from section 3.1 that there are many factors influencing mental well-being. These suggested policy interventions thus likely help reducing mental illnesses, though need to be complemented by a set of other policies to really make a great impact on improving employee mental well-being.

8.2 Threats to causality & limitations

Until now, I have interpreted the results as causal. Of course, this would be an ideal situation if this were the case, but it is likely that the estimated coefficients are not exactly reflecting the causal relationship between overtime hours worked and mental well-being in my sample. A possible threat to causality in this case is the possibility of *reverse causality* biasing my results. Reverse causality is present when in this case mental well-being also influences the number of overtime hours worked. The answer to these questions by a respondent is most likely based on the past few weeks, whereas the mental well-being questionnaire also asks about the mental well-being over the past weeks. The causal link therefore can run in two ways: from overtime hours to mental well-being, or vice versa. For example, a manager could prohibit his employees to work overtime if his employees suffer from a mental illness, or an employee could try to escape his private life by working more overtime. The former corresponds to a negative relationship between the mental well-being score used in my thesis and overtime hours worked, the latter to a positive relationship. The estimates could thus be biased upwards or downwards.

An earlier study sheds some light on this issue. De Lange et al. (2004) find in their paper that there actually most likely is a reverse relationship between mental well-being and work characteristics, such as job demands, but that work characteristics have a stronger effect on mental well-being than the effect in the other direction. This study thus makes it more plausible that there is a negative effect of working overtime on mental well-being.

The most important threat to causality is the possibility that there are other time-varying omitted variables, on which I do not have data. One such factor is already mentioned in section 3.1: conflicts between employee and employer. The exclusion of this variable for example would lead to downward biased estimates in my models, since conflicts would most likely lead to lower mental well-being, and working hours may be reduced by the employer, as a kind of retaliation. Ideally, one would control for all factors that vary over time and influence both overtime hours worked and mental well-being. This is often not possible, as is the case in my study. I can only hope that these factors do not significantly influence my results or nullify each other, which I cannot formally test. There are potentially many other time-varying confounders, so the sign of the total bias is not clear.

Another limitation which is also related to the *internal validity* of this study is the use of survey data. As already mentioned, survey data is prone to misunderstandings and as a consequence misreporting, and other behaviour that leads to measurement errors in the treatment variable. These cases of misreporting cannot be appropriately handled, since I do not know which observation is 'wrong' in some way. Another limitation of survey data is that respondents often can choose to refrain from answering the question. This does not lead to biased estimates, unless they do refrain from answering on the questions about working hours because of their mental well-being. This possibility cannot be neglected, and also biases the results ambiguously, though I cannot once again be sure whether this is the source of the missing data in my sample.

Taken altogether, the results in my thesis could be biased upward or downward. The estimates thus do not purely reflect the causal relationship between overtime hours worked and mental well-being, though one could believe that there actually is a negative impact and that the magnitude of the impact is limited, since only working some hours overtime does not immediately lead to mental illnesses, but is merely the consequence of the simultaneous interplay between many factors, one of which is overtime working.

9 Conclusion

This study presents evidence on the relationship between overworking and mental well-being, as well as evidence on the differential impact of overtime hours worked in manual and non-manual jobs, and the differential impact of paid and unpaid overtime hours worked. To the best of my knowledge, this is one of the first studies to present evidence on the differential impact of overtime hours worked in manual and non-manual jobs and the differential impact between paid and unpaid overtime hours worked.

The results suggest that there is a negative effect of the number of hours worked on mental well-being, and that this is driven by those working in a more mentally demanding non-manual job. Overworking per se is not significantly impacting mental well-being. Moreover, unpaid overtime hours do more harm than paid overtime hours. There are different explanations for the documented effect: work characteristics, stress, less time to socialise and the acknowledgement of effort by the employers.

Policy aiming at improving mental well-being among employees should make overworking by their employees less attractive for employers by raising taxes on overtime hours, especially in non-manual jobs, and make employers pay their employees for the overtime hours worked. These policies should be complemented with other policies aiming at improving mental well-being among employees. This is beyond the scope of this thesis.

Future research should focus on these other policies aiming at improving mental well-being of employees through other channels than overtime working, to establish an effective set of policy prescriptions for improving mental well-being. Caution is needed when literally interpreting the point-estimates in this study, since there might still be bias present in the estimates, though it is reasonable to think there is a modest negative effect of overtime hours worked on mental well-being, since mental well-being is influenced by many other unrelated factors.

References

ACAS (2022). Rules on working hours for people under 18: Young workers, apprentices and work experience. Accessed via https://www.acas.org.uk/young-workers-apprentices-and-work-experience.

Ala-Mursula, L., Vahtera, J., Kivimäki, M., Kevin, M. V., and Pentti, J. (2002). Employee control over working times: associations with subjective health and sickness absences. *Journal of Epidemiology & Community Health*, 56(4):272–278.

Anjara, S., Bonetto, C., Van Bortel, T., and Brayne, C. (2020). Using the ghq-12 to screen for mental health problems among primary care patients: psychometrics and practical considerations. *International journal of mental health systems*, 14:1–13.

Baetschmann, G., Ballantyne, A., Staub, K. E., and Winkelmann, R. (2020). feologit: A new command for fitting fixed-effects ordered logit models. *The Stata Journal*, 20(2):253–275.

Baetschmann, G., Staub, K. E., and Winkelmann, R. (2015). Consistent estimation of the fixed effects ordered logit model. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, pages 685–703.

Bardasi, E. and Francesconi, M. (2004). The impact of atypical employment on individual wellbeing: evidence from a panel of british workers. *Social science & medicine*, 58(9):1671–1688.

Beckers, D. G., Van der Linden, D., Smulders, P. G., Kompier, M. A., Taris, T. W., and Geurts, S. A. (2008). Voluntary or involuntary? control over overtime and rewards for overtime in relation to fatigue and work satisfaction. *Work & Stress*, 22(1):33–50.

Beckers, D. G., van der Linden, D., Smulders, P. G., Kompier, M. A., van Veldhoven, M. J., and van Yperen, N. W. (2004). Working overtime hours: relations with fatigue, work motivation, and the quality of work. *Journal of Occupational and Environmental Medicine*, pages 1282–1289.

Björklund, C., Jensen, I., and Lohela-Karlsson, M. (2013). Is a change in work motivation related to a change in mental well-being? *Journal of Vocational Behavior*, 83(3):571–580.

Booth, A. L. and Van Ours, J. C. (2008). Job satisfaction and family happiness: the part-time work puzzle. *The Economic Journal*, 118(526):F77–F99.

British Broadcasting Corporation (2020). How the japanese are putting an end to extreme work weeks. Accessed via https://www.bbc.com/worklife/article/20200114-how-the-japanese-are-putting-an-end-to-death-from-overwork.

Broberg, P., Tagesson, T., and Uman, T. (2020). Antecedents of psychological well-being among swedish audit firm employees. *International journal of environmental research and public health*, 17(10):3346.

Clark, A., Georgellis, Y., and Sanfey, P. (2001). Scarring: The psychological impact of past unemployment. *Economica*, 68(270):221–241.

Clark, A. E., Frijters, P., and Shields, M. A. (2008). Relative income, happiness, and utility: An explanation for the easterlin paradox and other puzzles. *Journal of Economic literature*, 46(1):95–144.

Clark, A. E. and Oswald, A. J. (1994). Unhappiness and unemployment. *The Economic Journal*, 104(424):648–659.

De Lange, A. H., Taris, T. W., Kompier, M. A., Houtman, I. L., and Bongers, P. M. (2004). The relationships between work characteristics and mental health: Examining normal, reversed and reciprocal relationships in a 4-wave study. *Work & Stress*, 18(2):149–166.

Del Pilar Sánchez-López, M. and Dresch, V. (2008). The 12-item general health questionnaire (ghq-12): reliability, external validity and factor structure in the spanish population. *Psicothema*, 20(4):839–843.

Ferrer-i Carbonell, A. and Frijters, P. (2004). How important is methodology for the estimates of the determinants of happiness? *The Economic Journal*, 114(497):641–659.

Ferrer-i Carbonell, A. and Gowdy, J. M. (2007). Environmental degradation and happiness. *Ecological economics*, 60(3):509–516.

Frey, B. S. and Stutzer, A. (2000). Happiness, economy and institutions. *The Economic Journal*, 110(466):918–938.

Gerdtham, U.-G. and Johannesson, M. (2001). The relationship between happiness, health, and socio-economic factors: results based on swedish microdata. *The Journal of Socio-Economics*, 30(6):553–557.

Gerlach, K. and Stephan, G. (1996). A paper on unhappiness and unemployment in germany. *Economics Letters*, 52(3):325–330.

Gerritsen, A. (2016). Optimal taxation when people do not maximize well-being. *Journal of Public Economics*, 144:122–139.

Goldberg, D. P. (1972). The detection of psychiatric illness by questionnaire: A technique for the identification and assessment of non-psychotic psychiatric illness. *Oxford University Press*.

Golden, L. and Wiens-Tuers, B. (2006). To your happiness? extra hours of labor supply and worker well-being. *The Journal of Socio-Economics*, 35(2):382–397.

Grant-Vallone, E. J. and Donaldson, S. I. (2001). Consequences of work-family conflict on employee well-being over time. *Work & stress*, 15(3):214–226.

Hetschko, C., Knabe, A., and Schöb, R. (2014). Changing identity: Retiring from unemployment. *The Economic Journal*, 124(575):149–166.

Ishida, Y., Murayama, H., and Fukuda, Y. (2020). Association between overtime-working environment and psychological distress among japanese workers: A multilevel analysis. *Journal of occupational and environmental medicine*, 62(8):641.

Jevons, W. S. (1871). *The theory of political economy*. Macmillan and Company.

Kahneman, D. and Deaton, A. (2010). High income improves evaluation of life but not emotional well-being. *Proceedings of the national academy of sciences*, 107(38):16489–16493.

Karasek, R., Brisson, C., Kawakami, N., Houtman, I., Bongers, P., and Amick, B. (1998). The job content questionnaire (jcq): an instrument for internationally comparative assessments of psychosocial job characteristics. *Journal of occupational health psychology*, 3(4):322.

Lelkes, O. (2006). Knowing what is good for you: Empirical analysis of personal preferences and the "objective good". *The Journal of Socio-Economics*, 35(2):285–307.

Lorente, L., Tordera, N., and Peiró, J. M. (2018). How work characteristics are related to european workers' psychological well-being. a comparison of two age groups. *International journal of environmental research and public health*, 15(1):127.

Office for National Statistics (2023). The national statistics socio-economic classification (ns-sec). Accessed via https://www.ons.gov.uk/methodology/classificationsandstandards/otherclassifications/thenationalstatisticssocioeconomicclassificationnssecrebasedonsoc2010.

Pega, F., Náfrádi, B., Momen, N. C., Ujita, Y., Streicher, K. N., Prüss-Üstün, A. M., Group, T. A., Descatha, A., Driscoll, T., Fischer, F. M., et al. (2021). Global, regional, and national burdens of ischemic heart disease and stroke attributable to exposure to long working hours for 194 countries, 2000–2016: A systematic analysis from the who/ilo joint estimates of the work-related burden of disease and injury. *Environment International*, 154:106595.

Pouwels, B., Siegers, J., and Vlasblom, J. D. (2008). Income, working hours, and happiness. *Economics letters*, 99(1):72–74.

Robone, S., Jones, A. M., and Rice, N. (2011). Contractual conditions, working conditions and their impact on health and well-being. *The European Journal of Health Economics*, 12:429–444.

Shields, M. A. and Price, S. W. (2005). Exploring the economic and social determinants of psychological well-being and perceived social support in england. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 168(3):513–537.

Stansfeld, S. A., Shipley, M. J., Head, J., Fuhrer, R., and Kivimaki, M. (2013). Work characteristics and personal social support as determinants of subjective well-being. *PloS one*, 8(11):e81115.

Statista (2022). Average number of non-scheduled hours worked per month in japan from 2012 to 2021. Accessed via https://www.statista.com/statistics/643896/japan-monthly-overtime-working-hours/.

Steptoe, A., Deaton, A., and Stone, A. A. (2015). Subjective wellbeing, health, and ageing. *The Lancet*, 385(9968):640–648.

Thorisdottir, I. E., Asgeirsdottir, B. B., Kristjansson, A. L., Valdimarsdottir, H. B., Tolgyes, E. M. J., Sigfusson, J., Allegrante, J. P., Sigfusdottir, I. D., and Halldorsdottir, T. (2021). Depressive symptoms, mental wellbeing, and substance use among adolescents before and during the covid-19 pandemic in iceland: a longitudinal, population-based study. *The Lancet Psychiatry*, 8(8):663–672.

Van Praag, B. M., Frijters, P., and Ferrer-i Carbonell, A. (2003). The anatomy of subjective well-being. *Journal of Economic Behavior & Organization*, 51(1):29–49.

Winkelmann, L. and Winkelmann, R. (1995). Happiness and unemployment: a panel data analysis for germany. *Applied Economics Quarterly*, 41(4):293–307.

Winkelmann, L. and Winkelmann, R. (1998). Why are the unemployed so unhappy? evidence from panel data. *Economica*, 65(257):1–15.

Wooden, M., Warren, D., and Drago, R. (2009). Working time mismatch and subjective well-being. *British Journal of Industrial Relations*, 47(1):147–179.

World Health Organization (2022a). Mental health. Accessed via https://www.who.int/health-topics/mental-health#tab=tab_2.

World Health Organization (2022b). Mental health at work. Accessed via https://www.who.int/news-room/fact-sheets/detail/mental-health-at-work.