Is work bad for health?  
The role of constraint vs choice

Andrea Bassanini
Eve Caroli

Janvier 2014
Is work bad for health?
The role of constraint vs choice*

Andrea Bassanini
(OECD and IZA)

Eve Caroli
(University Paris Dauphine, LEDa-LEGOS, Paris School of Economics and IZA)

Abstract
This paper reviews the literature on the impact of work on health. We consider work along two dimensions: (i) the intensive margin, i.e. how many hours an individual works and (ii) the extensive margin, i.e. whether an individual is in employment or not, independent of the number of hours worked. We show that most of the evidence on the negative health impact of work found in the literature is based on situations in which workers have essentially no control (no choice) over the amount of work they provide. In essence, what is detrimental to health is not so much work per se as much as the gap which may exist between the actual and the desired amount of work, both at the intensive and extensive margins.

Keywords: health, work, retirement, hours worked, job loss, individual choice

Résumé
Dans cet article, nous passons en revue la littérature sur l'impact du travail sur la santé. Nous considérons le travail selon deux dimensions : (i) la marge intensive, à savoir le nombre d'heures travaillées par un individu et (ii) la marge extensive, c'est-à-dire le fait qu'un individu soit ou non en emploi, indépendamment du nombre d'heures travaillées. Nous montrons que l'effet néfaste du travail sur la santé, mis en évidence dans la littérature, est basé sur des situations dans lesquelles les salariés n'ont, pour l'essentiel, aucun contrôle (aucun choix) sur la quantité de travail qu'ils fournissent. Dans les faits, l'effet néfaste sur la santé ne proviendrait pas tant du travail en soi que de l'écart éventuel entre la quantité de travail effectivement fournie et la quantité désirée, tant à la marge intensive qu'à la marge extensive.

Mots clés : santé, travail, retraite, heures travaillées, licenciement, choix individuel

JEL codes: I10, I31, J22, J28

* The views expressed here are those of the authors and cannot be attributed to the OECD or its member countries. We are grateful to Sylvie Blasco, Clémentine Garrouste, Mathilde Godard, Maarten Lindeboom, Claudio Lucifora, Pierre-Jean Messe, Lexane Weber and participants at the TEPP Conference on "Research in Health and Labour Economics for useful comments and suggestions. All remaining errors are our own. The authors gratefully acknowledge financial support by the Chaire Santé Dauphine.
Introduction

Whether and to what extent work may affect health is a widely debated issue. Cases of burnout or even suicide at work are regularly reported by the media, but conversely we all know that losing one's job may have severe health consequences.

The impact of work on health has long been investigated in the literature, in particular in epidemiology, occupational psychology and even sociology. More recently, health economists have started to be interested in this question. Surprisingly enough, we are still waiting for labour economists to join the group. This is all the more surprising that assessing the health effect of work is of major importance from a policy point of view. If excess, or alternatively, insufficient work is bad for health, this may call for public regulations and/or subsidies. Moreover, as the working-age population gets older, keeping people in employment may become harder and/or less socially desirable – and less advantageous for the government budget – if work turns out to be bad for health.

Assessing the impact of work on health is difficult, essentially for two reasons. First of all, work is a vague concept: when looking at the health impact of work, what is it that matters? Is it where people work? In which conditions? How much? etc. Work encompasses so many dimensions that it is extremely difficult to characterise it in an easy way. In this paper, we will consider work along two dimensions: (i) the intensive margin, i.e. how many hours an individual works and (ii) what we may call the extensive margin, i.e. whether an individual is in employment or not, independent of the number of hours worked. We will investigate how work affects health along both dimensions, i.e. whether working long hours and/or being in employment has any positive or negative effect on health.

Even once work has been defined, assessing its impact on health remains a hard task. This is mainly due to identification problems which plague the analysis both at the intensive and extensive margins. These identification problems mostly come down to selection issues. Static selection is known as the "healthy worker effect". It is essentially a problem of reverse causality: healthy workers are more likely to be in employment than unhealthy ones (extensive margin) and they are also more able to work longer hours (intensive margin). Solving this problem is tricky if only cross-sectional data are available. It is potentially easier using panel data which allow controlling for past health levels or, even better, estimating fixed-effect models. What makes things more difficult when trying to estimate the health effect of work is that dynamic selection may also be a concern. Changes in workers' health may generate changes in their employment status or in the number of hours they work. This
identification problem is not solved by fixed-effect estimates. One has to rely on more sophisticated empirical strategies.

In the rest of the paper, we will review the literature on the health impact of work. We will emphasise the quality of the empirical evidence provided by the various pieces of research and the extent to which they succeed in solving identification problems. We will show that at the intensive margin, working long hours turns out to be unambiguously bad for health. At the extensive margin too, work seems to be bad for health. The evidence on retirement mostly goes in the direction of a positive health impact of retiring, at least once addressed the most obvious endogeneity problems. The literature on unemployment and health is more mixed and suggests that losing one's job is bad for health. How can we account for this difference between the health effects of retirement and unemployment? One obvious explanation hinges on permanent-income losses. But, beyond the role of income, we argue that what matters is whether individuals are forced to work or stop working or whether they are free to choose. We will show that most of the evidence on the negative health impact of work found in the literature is based on situations in which workers have essentially no control (no choice) over the amount of work they provide. In essence, what is detrimental to health is not so much work per se as much as the gap which may exist between the actual and the desired amount of work, both at the intensive and extensive margins.

The remainder of the paper is organised as follows. Section 1 discusses the health impact of work at the intensive margin. Section 2 considers what occurs at the extensive margin. Section 3 revisits the evidence on both intensive and extensive margins emphasising the importance of choice versus constraint in determining the health effect of work.

1. The intensive margin: the health effects of working long hours

Defining long hours is not an easy task. It can be done on the basis of regulation: the 2003 EU Working Time Directive establishes that the average working week cannot exceed 48 hours, including over time. This limit can be used as a threshold beyond which hours worked are considered to be long, but it is clearly somewhat arbitrary. Another way to define long hours is with respect to the distribution of hours worked in the population under study. On a population of rural Chinese migrants, Frijters et al. (2009) consider that individuals work long hours if they work more than 60 hours per week which is very close to the median number of hours worked in their sample (56h). However, most papers in the literature simply define long
hours as overtime work, which is quite imprecise but has the advantage of being easily available in a large number of surveys.

The literature in epidemiology has long suggested that working long hours is correlated with lower health status. In their review of the literature, Spurgeon et al. (1997) conclude that "there is good evidence" that working time beyond 50 hours a week has adverse effects on health, in particular mental health and cardiovascular disorders. However, Beswick and White (2003) underline that most of the evidence on cardiovascular diseases is based on Japan. Karoshi – a syndrome of cardiovascular attack such as strokes, myocardial infarction or acute cardiac failure – is indeed more frequent among Japanese employees working long hours than among others. Consistent with Spurgeon et al. (1997), Sparks et al. (1997) conduct a meta-analysis of 19 studies and find that working long hours – whatever the definition – is negatively correlated with overall health as well as physiological and psychological health. These results are consistent with the historical evidence presented by Bourdieu and Reynaud (2006). They argue that the reduction in working time that took place in France in the 19th century was due to the combination of two factors: (i) the detailed evidence provided by doctors on the deterioration of health in the working age population at that time and (ii) the fact that employers realised that this was detrimental to labour productivity and that one of the main causes of this deterioration were excessive working hours.

As suggestive as this literature may be, it suffers either from lack of generality (historical evidence) or from major selection problems. The latter are widely acknowledged by epidemiologists. In particular, the "healthy worker effect" is mentioned in most papers as a potential source of bias towards zero in the results. However, the bias may also go in the opposite direction. For example, workers with lower mental health status due to anxiety may work longer hours if they are afraid they might lose their job. In an attempt to overcome these identification problems, a number of papers use fixed-effect panel-data models in order to control for time-invariant heterogeneity across workers. This is the case of Llena-Nozal (2009) who uses panel information for the United-Kingdom (BHPS\(^1\), 1991-2005), Australia (HILDA\(^2\), 2001-2005), Canada (NPHS\(^3\), 1994-2004) and Switzerland (SHP\(^4\), 1999-2005) to estimate the impact of changes in hours worked on an indicator of mental distress. Estimates are run separately for each country and results suggest that moving from standard work hours

---

\(^1\) British Household Panel Survey.
\(^2\) Household, Income, Labour Dynamics in Australia.
\(^3\) The National Population Health Survey.
\(^4\) Swiss Household Panel.
to overtime reduces mental health for men in Canada, the United-Kingdom and Australia. Symmetrically, moving to full-time regular hours is positive for mental health in all countries except for both men and women in Switzerland and for women in Canada, where results are not significant. Robone et al. (2011) find consistent results for self-assessed health. On 12 waves of the BHPS, they estimate a dynamic panel ordered-probit model with correlated random effects, controlling for prior health status. Their results suggest that having a part-time job (as compared to full-time job) has a positive and significant impact on health for both men and women who are satisfied with their number of hours worked. Overall, these papers not only suggest that working overtime is bad for (mental) health, but also that reducing to some extent the number of hours worked may be health improving.

To what extent are these results unbiased? They are so if there is no selection due to time-varying unobservable characteristics. This is a very strong assumption. Note however, that if negative health shocks induce individuals to reduce their number of hours worked, this generates a bias towards zero in the coefficients estimated in this literature. Of more concern would be situations in which a deterioration in mental health would turn employees into workaholics or, alternatively, changes in individuals' personal situation (such as children turning a certain age, for example) would force them to move from part-time to full-time work, while generating some depression symptoms. These confounding mechanisms would indeed generate an upward bias in the estimates so that the positive relationship between longer working hours and bad health would turn out to be spurious.

A couple of recent papers try to overcome these problems using instrumental variable strategies. Frijters et al. (2009) study Chinese workers who temporarily migrate from rural to urban areas. Given that the Chinese law forces them to go back to their home village after a while, most of these workers work very long hours (75% of them work more than 48 weekly hours) in order to save money for when they return to the rural countryside and start a family. Using the 2008 Urban Migrant Survey, Frijters et al. instrument the number of hours worked by the average wage in the home village and the proportion of inhabitants of the home village who have migrated. In the first stage, the former is negatively associated with long hours, while the latter attracts a positive coefficient. Both instruments are jointly significant and, in the second stage, the number of hours worked turns out to have a negative impact on mental health, as measured by a Likert scale based on a 12-item General Health Questionnaire (Goldberg and Williams, 1988). The same result holds if considering the impact of a dummy

---

5 This is measured as the wage that an unskilled person can earn in the home village.
variable equal to 1 when workers work more than 60 hours a week. All results are robust to controlling for demographic characteristics (gender, age, marital status, number of children, education, parents' health, number of friends and height), log weekly wages, occupation, industry and type of job contract. They are likely to be unbiased but it is not clear whether they are specific to very long working hours or may be generalised to lower numbers of hours worked. Berniell (2012) provides complementary evidence on this point using French data. Her research exploits the 1998 and 2002 waves of the ESPS^6 survey in order to study the impact of the number of hours worked on health behaviours. Her identification strategy is based on a quasi-experiment provided by the enactment of the Aubry I law on the reduction of the workweek in 1998. The law brought the legal number of weekly hours worked down from 39 to 35, but the reform was binding only for medium and large firms (more than 20 employees) and they had at least two years to abide by it. For health outcomes for which she has information at both dates, Berniell implements a difference-in-difference estimation in which she compares employees working more than 35 hours a week in 1998 and employed in 2002 in firms which had implemented the legal reduction in working time with a control group consisting of employees initially working less than 35 hours and employees working more than 35 hours but employed in firms which had not reduced the workweek yet by 2002. For health outcomes available only in 2002, she instruments the number of hours worked by a dummy variable equal to 1 if the individual worked in a firm which had already abided by the law in 2002 and the average number of hours worked in 2002 by individuals in the same occupational category. Both methods should yield consistent results as long as workers do not self-select themselves into firms more likely to implement the legal reduction of the workweek (more worker-friendly enterprises?) on the basis of characteristics correlated with expected changes in their health status. To the extent that this exclusion restriction is plausible, the results may be considered unbiased. Whatever the estimation method, the results suggest that the larger the number of hours worked, the worse the health behaviours: the reduction in working time enjoyed by employees in firms which complied with the requirements of the law between 1998 and 2002 reduced their probability of smoking as well as their BMI. Similarly, shorter hours worked proved to decrease alcohol consumption and increase physical activity.

^6 Enquête sur la Santé et la Protection Sociale.
Overall, at the intensive margin, the number of hours worked seems to have a negative effect on health. This is the case for long hours but it also seems that reducing "normal" hours to lower levels of hours worked is likely to be health-enhancing through better lifestyle habits.

The literature is not so clear, however, as to the reasons why longer working hours negatively affect health. According to Spurgeon et al. (1997), two channels may be considered. First, working long hours is a stressor per se. Long hours are a factor of fatigue. They have deleterious effects on sleep and on mental health. They also raise the risk of accident. This is particularly the case when people work for long (potentially night) shifts. Harrington (2001) recalls that major catastrophes such as Three Mile Island, Chernobyl or Exxon Valdez all occurred in the early hours of the morning because of errors made by employees who had been on duty for long hours. Second, and more problematic to us, working long hours increases exposure to other types of stressors such as, for example, bad working conditions or flexible work schedules. The respective roles of each channel are still to be assessed. This is a key point though since the second channel (exposure to other stressors) implies that confounding factors may be at work. Unfortunately, none of the IV strategies mentioned above can really rule out the role of such confounding variables. If Chinese migrants accept longer and more flexible working hours because they expect a lower wage when back to their home village, flexible work schedules may well be the "true" driving factor of mental health deterioration rather than long hours. Similarly if the legal reduction of the workweek was accompanied in France by less flexible work schedules and/or better working conditions, this may account for the observed improvement in health behaviours. The former effect may be ruled out since, if anything, the reduction of the workweek was accompanied by more flexible work patterns – Askenazy (2004a). The latter is more debatable although scattered empirical evidence suggests that work intensity has rather increased than decreased over the period of implementation of the reduction of the workweek – see Askenazy (2004b) and Mehaut et al. (2008).

Despite the limitations due to the identification problems persisting in the literature and the existence of confounding factors, the state of our knowledge goes in the direction of a negative impact of work on health at the intensive margin. Long hours worked seem to be detrimental to health and reducing working time comes out as a potential solution to improve

---

the health status in the working population. If working less is good for health, is it the case that not working at all would be even better?

2. The extensive margin: the positive (?) effects of not working

As far as individual employment is concerned, the notion of extensive margin is not quite standard. We define it here as the fact of working or not, independent of the number of hours worked – for a similar definition, see Lindeboom and Kerkhofs (2009). In the working age population, individuals who are not in employment may be either unemployed, retired, or out of the labour force for other reasons. Researchers have been much interested in the health consequences of retirement and unemployment. In contrast, the impact of being out of the labour force for reasons different from retirement has been little investigated, so that we will leave it aside in this review.

2.a The health impact of retirement

The literature on the health effects of retirement mostly suggests that, at the extensive margin too, working is bad for health. This literature faces a major identification problem due to what could be called the "unhealthy worker effect". Retirement is indeed a choice and there is widespread evidence in the literature that workers with poor health status tend to retire earlier (see Currie and Madrian, 1999). This generates a negative correlation between retirement and health which results in a downward bias when trying to estimate the causal effect of retirement on health.

A first strand in this literature has attempted to correct for time-invariant heterogeneity across workers using the time dimension of panel data. Using the Dutch CERRA survey\(^8\) of elderly workers, Kerkhofs and Lindeboom (1997) estimate a first-difference model in which changes in health – as measured by a summary index built out of responses to the 57-item Hopkins Symptom Checklist, HSCL – are explained by changes in labour market status between 1993 and 1995, controlling for health shocks. The results suggest that early retirement has no short-run effect on health conditions but that longer-run effects (two years later) tend to be positive. Dave et al. (2008) find different results on US data (the HRS survey\(^9\)). They estimate a fixed-effect model on the subsample of individuals who did not report any health problem before

\(^8\) Panel Survey of the Leiden University Center for Research on Retirement and Aging.

\(^9\) Health and Retirement Survey.
retiring. They find that retirement increases difficulties with mobility and other daily activities, but also illness conditions and mental distress. However, controlling for individual fixed effects does not guarantee that results are unbiased in particular if sources of endogeneity vary over time. Similarly, controlling for past health conditions does not ensure that very recent health deterioration is not the key factor driving retirement decisions.

In order to improve on these estimation strategies, Charles (2004) proposes an instrumental variable approach. Using the HRS survey too, he instruments retirement decisions with eligibility ages for pension benefits in the US Social Security system and the changes in these age thresholds over time. He finds that once the endogeneity of retirement decisions is accounted for, retirement appears to have a positive effect on mental health insofar as it reduces the probability of feeling depressed or lonely. Neuman (2008) uses the same data and instruments retirement decisions using several sets of instruments among which indicators for whether the individual is between 62 and 65 years old – i.e. eligible for early entitlement to Social Security benefit – and whether he is older than 70 and thus no longer must face the earnings test used to compute the pension replacement rate. Controlling for age, Neuman finds that retirement reduces the probability of subjective health decline for both men and women. For women, it also reduces the probability of decline in activities of daily life (ADL). Coe and Lindeboom (2008) use the same data to estimate the health impact of early retirement. They instrument the latter by a dummy variable indicating whether the individual has been offered an early retirement window by his employer. This is plausibly exogenous with respect to individual health to the extent that these windows are, in general, unexpected and targeted at entire units or divisions of the firm. The impact of early retirement on self-assessed health comes out as positive and marginally significant. However, this turns out to be a short-run effect which disappears in the longer run – four years later. Coe and Zamarro (2011) find stronger results using the 2004 wave of the SHARE\textsuperscript{10} survey, which is the European equivalent of the HRS. The instruments they use are very similar to those proposed by Charles (2004) and Neuman (2008): they consist in two dummy variables indicating whether the individual is above or below the full and early retirement eligibility ages in his country. They also condition on a polynomial in age and country fixed effects. Their results suggest that retiring has a positive effect on self-assessed health and on a health stock index obtained by regressing subjective health on a number of health symptoms. In contrast, the effect is insignificant for depression and cognitive abilities.

\textsuperscript{10} Survey of Health, Ageing and Retirement.
In contrast, other papers find rather negative effects of retirement on health using similar identification strategies. Godard (2013) uses the same instrument as Coe and Zamarro (2011) but on three different waves of SHARE which allows her to include individual fixed effects in her regressions. She finds that retirement increases the probability of being obese but only for men retiring from physically demanding jobs. Rohwedder and Willis (2010) pool 2004 cross-sectional data from SHARE, the US HRS and the British panel of elderly workers ELSA. They estimate the same model as Coe and Zamarro (2011). Their results suggest that there exists a negative effect of retiring on cognitive abilities, although they are difficult to interpret since their specification does not control for education and gender and does not include country fixed effects. The negative impact of retirement on cognitive abilities is confirmed by Mazzonna and Peracchi (2012) on SHARE data and by Bonsang et al. (2012) on the HRS who both use eligibility ages as instrumental variables for retirement. Eventually Behncke (2012) estimates the health impact of retirement on three waves of the ELSA survey. Using a matching estimator combined with an IV strategy relying on the same instrument as Coe and Zamarro (2011) and controlling for health before retirement, she finds a positive and significant effect of retirement on difficulties in ADL and walking, as well as a negative effect on a health stock index. However her empirical framework does not allow for differences in age trends before and after eligibility ages. This could be a confounding factor if the health impact of age accelerates at some point and this point is close to eligibility ages. This criticism applies, in principle, to all studies using eligibility ages as instrumental variables since their empirical strategy is similar to a regression discontinuity design which requires that patterns be stable around the discontinuity. However, this issue is potentially less of a concern in cross-country analyses where the variation in eligibility ages is much larger than in single-country studies. In the former case, the identifying assumption is that, even if the health impact of age varies over time, it does so in a similar way in all countries. Although there is no certainty that this is the case, this assumption looks plausible.

The empirical strategy used by Blake and Garrouste (2012) is immune from this problem since they do not rely on regression discontinuity design but rather on a pension reform to identify the causal impact of retirement on health. This reform was enacted in France in 1993. It raised the number of quarters of contribution required to be entitled to full pension benefit and extended the period over which the average wage used as a reference to compute pensions was calculated. Both changes were introduced in a progressive way so that earlier cohorts

---

11 English Longitudinal Study of Ageing.
were less affected than later ones, which generates the variation needed to instrument individual retirement decisions. Based on the 1993 and 2005 Baromètre Santé surveys, the authors use Duke health profiles and regress them on retirement decisions instrumented by the changes in the required contribution quarters and reference wage brought about by the 1993 reform. The results suggest that retiring has a positive and significant impact on physical health as measured through the ability to walk up stairs and run. In contrast, it has no significant effect on mental health. De Grip et al. (2012) also use a pension reform but look at the health impact of postponing the expected retirement age of cohorts that, before the reform, were too young to be eligible for retirement. They exploit the withdrawal of the favourable tax treatment for public-sector early-retirement pensions in the Netherlands in 2006. This change affected individuals born in 1950 or later but not those born earlier. It was quite massive: for example, in order to benefit from the same replacement rate as unaffected individuals (born in 1949), those born in 1950 had to delay retirement by about 13 months or increase their savings by 14% over 7 years. The authors compare the frequency of depression across these groups and find that individuals affected by the reform had a 40% higher risk of suffering from depression. These results suggest that postponing retirement – even if only in expected terms - has a negative impact on health.

Overall, most of the evidence regarding the health effects of retirement rather goes in the direction of a positive impact – except for cognitive abilities. This suggests that, at the extensive margin too, not working is good for health or, alternatively, that work can be a threat to health. This conclusion is further supported by Lindeboom and Kerkhofs (2009) who show that too long careers tend to be damaging to health. Using Dutch data (the CERRA survey) and controlling in a non-parametric way for age, they estimate the impact of the number of months worked by an individual over his career on a summary health index. Thus doing, they find evidence that working more than 25 years has a negative impact on health. Even though this research design is not necessarily immune from selectivity, let us underline that, should there be a selection effect here, it would most likely take the form of a healthy worker effect and hence bias the results towards zero.

In contrast with the findings in the retirement literature, the idea that working is bad for health, so that not working would be better, is however not fully supported by the literature on economic conditions and unemployment.
2.b Health and economic conditions

A recent strand of literature has extensively investigated the impact of local economic conditions on health and, here again, the results could be interpreted as suggesting that not working is good for health. The seminal paper in this research area was published by Ruhm in 2000 and investigates how health responds to transitory changes in economic conditions. Using US panel data from 1972 to 1991, economic conditions are proxied by state-level unemployment rates and the author estimates their impact on state mortality rates. His specification includes state-level time varying controls (age, education and ethnic structure of the population, average personal income) together with state and time fixed effects. The results suggest that in periods and areas where unemployment is higher, mortality rates are lower. This negative correlation between unemployment and mortality is particularly strong for young adults (20 to 44 years old) and to a smaller extent for senior citizens (above 65 years old) while it is not significant for prime-age workers (45 to 64 years old). Similar results are found for other countries than the USA – see Neumayer (2004) for Germany, Tapia-Granados (2005) for Spain, Buchmueller et al. (2007) for France and Gerdtham and Ruhm (2006) for OECD countries. In addition, Ruhm (2003) finds that the number of people affected by chronic conditions also increases in good times: a 1% fall in US state-level unemployment is associated with a 4.3% increase in the prevalence of ischemic heart diseases and a 8.7% increase in intervertebral disk disorders. The pro-cyclicality of mortality and morbidity may seem puzzling at first sight. However, it turns out to be consistent with evidence on the counter-cyclicality of health-enhancing behaviours. Using the same empirical framework as in Ruhm (2000) – although with individual-level dependent variables –, Ruhm (2005) shows that the probability of smoking, of being obese as well as physical inactivity decrease when state-level employment goes down. The key explanation for these findings put forward by Ruhm (2000; 2004) is that the opportunity cost of time goes down during downturns so that if health-producing activities are time intensive, people are likely to dedicate more time to them during bad times. In other words, people will spend more time cooking, doing sport etc. which will contribute to the improvement of their health status.

One interpretation of Ruhm’s results is that downturns are good for health because people who become unemployed have more time to dedicate to health-improving activities. This is consistent with the idea that work is bad for health at the extensive margin so that any reduction in the proportion of individuals who are out of employment is likely to raise the average level of health in the population.
However, this interpretation is challenged by various pieces of evidence. Miller et al. (2009) show that, adding to Ruhm's specification the unemployment rate of the demographic group to which the individual belongs – which is likely to be a good predictor of the own probability of becoming unemployed -, this variable turns out to be positively correlated with mortality - although often non statistically significant -, while state-level unemployment remains negatively signed and significant. If employment status were the major driver of the reduction in mortality during downturns, the unemployment rate in the individual's own demographic group should be a strong predictor of mortality and its correlation with it should be negative. Moreover, focusing on individuals with the lowest employment probabilities, Charles and De Cicca (2008) show that, in this subsample, the probability of being overweight or obese increases with state-level unemployment as do mental distress and the probability of smoking. This suggests that the pro-cyclicality of mortality and morbidity is not directly driven by a positive effect of unemployment (or unemployment prospects) on health and hence by work at the extensive margin.

An alternative explanation is that the action comes from the intensive margin: at times of high unemployment, individuals who have jobs work fewer hours (in particular due to the reduction in overtime work) and hence get more time to engage in health-improving activities. However, this interpretation is at odds with empirical evidence since Ruhm (2005) himself acknowledges that the positive effect of high employment levels on the probability of smoking, of becoming obese and on physical inactivity is no larger if restricting the sample to people in employment. Moreover, Miller et al. (2009) provide evidence that the largest part of the decrease in deaths during bad economic times is due to a reduction in fatalities for people aged 80 years old and above, whose labour market attachment is marginal or nil – see also Stevens et al. (2011). This suggests that other factors related to economic activity are at work in accounting for the pro-cyclicality of mortality. Ruhm (2004) underlines that pollution decreases during downturns as do fatalities due to motor vehicle accidents – see also Neumayer (2004) and Buchmueller et al. (2007). In addition, Ruhm (2006) and Jusot (2012) hypothesize that in bad times, adults of working age dedicate more time to caring for the elderly as suggested by Vistnes and Hamilton (1995). This suggests that the reduction in mortality and morbidity observed during downturns is not due to the direct benefit of working fewer hours or not working at all on individuals who have or used to have a job. The mechanisms at work are more indirect and may go through a safer or healthier environment.
which is likely to be key for the health status of the more fragile fraction of the population, among which the elderly, and/or more time and attention being dedicated to them.

The fact that becoming unemployed during downturns may not be considered as positive for health is actually consistent with findings in the unemployment and job loss literature.

2.c The health effects of unemployment and job loss

The impact of unemployment and/or job loss on health has been much investigated in the literature. The results are mixed, ranging from strong health damaging effects to insignificant ones. However, no article ever finds a positive health effect of becoming unemployed.

The traditional literature on unemployment and health was based on simple correlations (Theodossiou, 1998) and evidenced a strong negative association between unemployment and any measure of health. However, as underlined by Schmitz (2011), identifying a causal impact of unemployment on health is difficult because of selection effects: ill workers tend to be selected out of work into unemployment and poor health causes longer unemployment spells. Both points increase the probability of observing an ill individual in the pool of unemployed which, in turn, generates a lower average health status in this group.

In order to overcome these identification problems, three strategies have been used in the literature. The first one relies on estimating fixed-effect models while the second one compares the health outcomes of displaced and non-displaced workers in circumstances in which displacement is likely to be exogenous. The third one combines the first two.

Llena-Nozal (2009) estimates fixed-effect panel data models for men and women in Australia, Canada and the United-Kingdom. In all countries, changing from employment to unemployment increases mental distress as compared to individuals staying in employment. This increase is significant and large. It is larger for men than for women except in Australia where the impact for men is not significant at conventional levels. Similar estimates are run by Böckerman and Ilmakunnas (2009) for Finland. Their findings are somewhat different since, after controlling for individual fixed-effects, they do not find any significant effect of becoming unemployed on self-assessed health. This suggests that unemployment may affect in a different way mental and physiological health. However one cannot be conclusive on this point since dynamic selection problems cannot be ruled out here: a negative health shock may make workers less productive hence increase the probability that they be fired. Such
mechanism could account for the positive correlation between unemployment and mental distress found by Llena-Nozal (2009).

Given the difficulty of identifying any health effect of unemployment using fixed-effect estimates, a number of papers in the literature have focused on mass displacements in an attempt to find exogenous sources of job loss. The seminal paper in this literature is Sullivan and von Wachter (2009) who use US Social Security data on mortality matched with male workers' employment histories in Pennsylvania between 1980 and 2006. Workers are considered to be displaced if they have left their firm during the period 1980-1986 and the employment at their former firm was 30% or more below its peak level since 1974 the year after the employee left. The authors estimate the impact of displacement on mortality hazards. Given that displacement is likely to be non-random and hence endogenous, they control for the mean and standard deviation of workers' earnings over a period of several years prior to job loss. This is meant to correct for selectivity problems such as those due to the fact that firms are likely to layoff less productive workers who may turn out to be in poorer health. The results display a 10 to 15% increase in annual death hazards following displacement and the effect is still positive and significant more than 15 years after displacement. The identifying assumption underlying this research is that, in case of mass layoffs, there is no selection of displaced workers conditional on the mean and variance of wages.

Other papers in the literature have tried to improve on Sullivan and von Wachter by using data containing direct information on the reason for displacement and selecting displacements due to plant closure. The underlying assumption is that when plants close down, they have to fire all workers so that selection effects are potentially reduced. They may still exist at the plant level if plants which close down employ a large fraction of low-productive workers and if these tend to be in poorer health than average. In order to control at least for differences in observable characteristics of workers, notably the health status prior to displacement, Browning and Heinesen (2012) use a propensity score weighting estimator. Using Danish data from 1980 to 2006, they weight the observations in the non-displaced (control) group by their odds (calculated from their predicted propensity scores) so that the weighted number of control observations is equal to the number of displaced persons in each base year. The displacement groups of all base years are then pooled together and so are the control groups. The authors estimate non-parametric cumulative hazard functions (at different durations from the base year) for the displacement and control groups and calculate their ratio. The results suggest that the risk of overall mortality is much higher in the displacement group than in the
control group. The gap is maximum just after plant closure but it remains statistically significant even after 20 years. Most of the difference is due to deaths from circulatory diseases, in particular myocardial infarctions and strokes. Using a similar empirical strategy on Swedish data, Eliason and Storrie (2009a) find that overall mortality increases by 44% over the first four years following job loss with the main bulk of the increase being due to alcohol-related deaths and suicides. In a companion paper (Eliasson and Storrie, 2009b) they find that displacement significantly increases the probability of hospitalisation due to alcohol-related conditions in Sweden. In contrast, they find no evidence of increased risk of severe cardiovascular diseases.

Static propensity score estimators do correct for selection on observables but not for selection on unobservables. So, a third strategy used in the literature has been to investigate the health impact of job loss due to plant closure including individual fixed-effects. On Danish data, Browning et al. (2006) estimate a difference-in-difference model in which they compare hospitalisations of workers who have been displaced because of plant closure to those of non-displaced workers before and after displacement – using propensity score matching methods in order to control for selectivity into displacement. Thus doing, they do not find any significant health effect of job loss due to plant closure. Osthus (2012) estimates a similar model on Norwegian data and does not find either any significant effect of displacement on psychological distress, muscle-skeletal pain and chest pain. In contrast, using the same empirical framework on Austrian data, Kuhn et al. (2009) find a positive – although economically small – effect of job loss due to plant closure on public health expenditures generated by consumption of psychotropic drugs and hospitalisation of men for mental problems and stroke. Using German data (the GSOEP), Schmitz (2011) estimates fixed-effect models of the impact of unemployment due to plant closure on various health outcomes. He finds no significant effect whatsoever, be it on satisfaction with health, mental health or overnight stays in hospital. Using the US HRS, similar results are found by Salm (2009) who estimates a difference-in-difference model of the impact of job loss due to business closure on a long list of health outcomes (self-assessed health, limitations in ADL, longevity expectations and mental health). The underlying assumption is that business closure is more exogenous than plant closure since the latter may be due to unfitness or incompetence.

12 Note however that controlling for individual fixed-effect is unlikely to solve all identification problems since the health effects of displacement may still be heterogeneous across workers. If more dynamic workers with higher re-employment probabilities are more likely to quit establishments before closure, the estimated health impact of displacement is upward biased due to sample selection.

13 German Socio-Economic Panel.
of workers – which may be correlated to health outcomes – whereas business closure is more likely to be due to organisational decisions to restructure or relocate business units. Salm (2009) concludes that job loss due to business closure has no effect on health, whatever the specific outcome.\textsuperscript{14}

Overall, the literature on unemployment and job loss suggests that becoming unemployed has, at best, no health effect. This does not necessarily mean that losing one's job has no overall negative effect on health. Using the 2010 European Working Condition Survey, Caroli and Godard (2013) indeed find that, when properly instrumented,\textsuperscript{15} job insecurity – defined as the perceived risk of job loss – has a negative effect on a limited number of health outcomes including self-assessed health, headaches or eyestrain and stomach ache. If the health of workers who are more at risk of being fired has already deteriorated before layoff, due to perceived job insecurity, it may not decline further when layoff actually takes place. This would be consistent with the literature in occupational psychology which underlines that the anticipation of a stressful event represents an equally important or even greater source of anxiety than the event itself (Lazarus and Folkman, 1984). In this case, the overall impact of job loss on health would be negative with the greatest part of the health deterioration taking place before the actual layoff.\textsuperscript{16}

Overall, there is no evidence of a positive effect of unemployment or job loss on health. How can we account for this difference in health effects between not working due to retirement and not working due to job loss? The positive effect of retirement on health could easily be explained by a conceptual framework adapted from the intertemporal model of health capital proposed by Grossman (1972). Following Muurinen and Le Grand (1985) and Case and Deaton (2005), assume that health capital deteriorates over time because of both age and work but that this deterioration may be compensated by health-promoting time-consuming activities and/or purchases of medical care. In this case, when individuals retire, their permanent

\textsuperscript{14} Deb et al. (2011) also use business closure as a measure of "exogenous” job loss and find negative effects on health behaviours for small groups of individuals who were already at risk. They estimate a finite-mixture model controlling for the lagged dependent variable and show that job loss raises BMI for 16\% of the population whose BMI was already over average, and strongly increases alcohol consumption (by 42\%) for 10\% of the population who already consumed more alcohol than average.

\textsuperscript{15} They instrument perceived job insecurity by the stringency of the employment protection legislation in the country where the individual lives interacted with the natural rate of dismissals in the sector where she is employed. The intuition behind this instrument is that workers are likely to feel more secure with respect to their job if living in a country where employment is strongly protected by the law, and relatively more so if employed in sectors where employment protection legislation is more binding because of a higher natural rate of dismissal.

\textsuperscript{16} This interpretation is supported by results from Kuhn et al. (2009) on Austria who find that the number of days of sick leave starts increasing a few quarters before plant closure. Mandal et al. (2011) also find that job loss expectation has a stronger impact on mental distress as compared to actual job loss, at least for older workers.
income remains roughly constant thus leaving unchanged their capacity of purchasing medical care. At the same time, the depreciation of their health capital slows down and they have more time to dedicate to health-promoting activities. This should generate a clear positive effect on health in standard estimates.

In the case of involuntary job loss, the prediction of this model is more ambiguous. Health capital depreciation does slow down, because exposure to the health-damaging effects of work disappears, but so does permanent income – see Bonsang and Klein (2012). So, in countries where medical care is mostly paid for by the patient himself, the health consequences of job loss may be negative if the reduction in the amount of medical care purchased by displaced workers more than compensates the slower depreciation of health capital. In this case, the conceptual framework derived from the Grossman model may also explain the deterioration of health following job loss. However, a large part of the evidence on the negative health effect of job displacement comes from countries in which medical care is heavily subsidised if not freely available – e.g. Denmark, Sweden, Austria. In this case, investments in health taking the form of medical care have no reason to go down in case of job loss and the adapted Grossman model should unambiguously predict an improvement in health following job displacement. As mentioned above, this is not what is observed in the data.

In the next section, we propose a different - and potentially complementary – explanation of the different health effects of retirement and job loss. We argue that whether not working is a choice or a constraint is likely to be a key determinant of its health impact. We show that all the results reviewed so far on the negative health effect of work both at the intensive and extensive margins can be reinterpreted along these lines.

3. How does work affect your health? The key role of constraint vs choice

At the intensive margin, the literature in occupational psychology suggests that the extent to which workers have control over their work pattern determines how damaging long working hours may be for their health. Sparks et al. (1997) report that individual control over hours worked has been found to influence perceived stress levels (Hall and Savery, 1986) and tolerance of work schedules (Barton et al, 1993). Similarly, Barnett et al. (1999) provide evidence that long hours have a negative impact on home and family life but with a smaller effect when individuals have some control over the number of hours they work. Schmitt et al.
(1980) also find significant correlations between health symptoms and the inability to refuse overtime. Similarly, Barton and Folkard (1991) find that freedom to choose particular work patterns has implications for the degree to which shift workers subsequently experience health problems. More recently, Dockery (2006) estimates the health impact of working more than one would like to. Using four waves of the Australian HILDA survey, he estimates a random-effect model and finds that working full-time – or even part-time – while preferring to work less is associated with lower mental health. Interestingly, the effect is also negative and significant, although smaller, if individuals are forced to work less than they would like to. These results suggest that the gap between the actual and the desired length of the workweek is at least as important as the total number of hours worked. However, the specification estimated by Dockery does not properly control for unobserved heterogeneity across workers since the assumption underlying random effect models is that the individual component of the stochastic error term is uncorrelated with the covariates.

In order to tackle the potential endogeneity of hours worked, a number of articles rely on panel data. Friedland and Price (2003) use the 1989 and 1994 waves of the Americans' Changing Lives study to investigate the health impact of overemployment. Workers are defined as overemployed in terms of hours worked if they work more than 45 hours a week and would like to work fewer hours. Three health measures are regressed on the overemployment indicator controlling for age, gender, education, marital status, weekly working hours and health status in the first wave. The results suggest that overemployment has no significant impact on self-assessed health nor functional health, while being positively correlated with the number of chronic conditions. In contrast, underemployment – defined as working less than 35 hours a week while wishing to work more – has no significant health effect, whatever the indicator. Bell et al. (2012) further improve by estimating fixed-effect panel data models using the GSOEP and the BHPS in order to investigate the health effects of hour constraints. In the GSOEP, workers are considered to be unconstrained in terms of working hours if the gap between the actual and desired number of hours worked is between -4 and +4. If it is larger than +4, workers are considered as overemployed. In the BHPS, respondents are directly asked if they are happy with the number of hours they work or whether they would like to work longer or shorter hours. Estimating fixed-effect ordered logits separately for both countries and controlling for several time-varying covariates (job tenure, marital status, number of children, household income, whether overtime is paid, disability and occupation), the authors find that, in Germany, overemployed individuals are
less satisfied with their health and report lower self-assessed health than reference workers (working 35 to 40 hours and who are unconstrained). Results are similar for the United-Kingdom, where overemployment appears to have negative effects on both health measures in the sample of individuals working more than 20 hours a week. The authors acknowledge that they cannot rule out reverse causality if negative health shocks reduce desired working time thus giving rise to overemployment.

The IV strategy used by Berniell (2012) yields similar results. The French reduction of the legal workweek which is used to identify the effect of hours worked on health behaviours had the peculiarity of being income preserving, since the law imposed that it be implemented at constant monthly earnings. As a consequence, the resulting decrease in hours worked may be seen as a positive choice of workers. As a matter of fact the actual number of hours worked decreased by a smaller amount – on average 2 hours per week – than what was made possible by the law. The resulting improvement in health behaviours can thus be seen as driven by a voluntary decrease in working time. An interesting avenue for further research would be to assess the health impact of a truly chosen reduction in working time. Goux et al. (2013) suggest one way to do this by looking at interdependencies in spousal labour supply. Using the French Labour Force Survey matched with a survey indicating whether and when employers actually reduced the workweek in order to abide by the law, they estimate the impact of having a partner working in an enterprise who has reduced working time on the number of hours worked by the individual himself. This reduction is likely to result from pure cross-hour effects given the income preserving nature of the reform. The authors find that men worked about half an hour less per week when their wives became treated, while women's response to their husband's treatment was small and insignificant. Building on this methodology, one could instrument the number of hours worked by an individual, using information on whether or not his partner was affected by the reduction of the legal workweek. If this could be matched with data on health or medical consumption, it would allow estimating the causal health impact of a truly desired change in hours worked.

At the extensive margin, the positive impact of retirement on health can be re-interpreted as a negative health effect of being forced to stay in employment. Blake and Garrouste (2012) indeed find that the increase in the length of the contribution period required to be entitled to full pension benefit and the reduction in the average level of the reference wage brought about by the 1993 French reform negatively affected the probability of retiring. This increase in the length of careers was clearly undesired since it was driven by the deterioration in pension
levels brought about by the reform. So, workers affected by the reform found themselves "involuntarily" retained at work. The negative effect on physical health found by the authors can thus be seen as the consequence of being unexpectedly forced to stay in employment while one had planned to retire. As evidenced by De Grip et al. (2012), the same effect also applies if workers are forced to change their expectation on how long they will have to stay in employment. Moreover, their findings are robust to controlling for changes in the expected replacement rate at a given retirement age, which suggests that this evidence cannot be entirely attributed to a permanent-income effect. The authors insist that control over one's own retirement may partly account for their results. Such a view is consistent with evidence on the effect of transition to retirement on individual happiness provided by Calvo et al. (2009), who suggest that what matters is whether people perceive this transition as chosen or forced. What determines happiness of older workers appears to be the sense of control that they have over their own retirement decision.

A similar interpretation can be given to the results coming out of the literature using retirement eligibility ages as instruments for retiring decisions – see Section 2.a above. In these papers, the identification strategy relies on the fact that the probability that an individual retires discontinuously increases as he reaches the full and/or early retirement eligibility age in his country. This suggests that immediately below that age, individuals would like to retire but do not because of the pension loss that this would imply. As soon as they reach the full-benefit age threshold, a large proportion of them retire, thus closing the gap between the actual and the desired length of their career. To the extent that the main bulk of the evidence in this literature goes in the direction of a positive impact of retirement on health, this suggests that, symmetrically, being stuck in employment while preferring to retire has a health damaging effect.17

The view according to which work has a negative impact on health when it is a constraint sheds a new light on the results on the health effects of job loss. In this literature, the decision to stop working is clearly involuntary since the focus is on workers who have lost their job due to plant or business closure. Following our assumption according to which it is the constrained nature of work which generates the negative health effects, the results on job loss are consistent with those on retirement: being forced to stop working has negative health effects in the same way as being forced to keep on working when one does not want to.

17 Consistent results are found by Falba et al. (2009) who look at how depressive symptoms may be related to missed expectations about the time of retirement.
The importance of work-related constraints in generating health damaging effects is further confirmed by the literature on undesired organisational changes. Ferrie et al. (1998) use the British Whitehall sample of London-based civil servants to estimate the health impact of a major reorganisation in the public service. In 1986, a report recommended the examination of civil-service functions in the UK to determine whether they could be abolished or transferred to the private sector. Implementation of this recommendation started in 1988 through the "Next Step" programme which separated executive from policy functions of government and transferred the executive functions to units called agencies, thus generating major involuntary organisational changes. Phases 1 and 3 of the Whitehall sample collect information on civil servant's health as of 1985-1988 and 1993, respectively. In 1993 it also contains information on whether employees have been transferred to an executive agency or expect to be so. The authors use a difference-in-difference approach to compare the 1988-1993 changes in health outcomes between employees who have been either transferred or expect to be so shortly, and a control group of workers for whom transfer is not expected to take place. Men who have been transferred report significantly lower self-rated health and suffer from higher psychiatric morbidity as measured by a 30-item GHQ. Similar results are found for men expecting to be transferred who also report a higher probability of sleeping short hours (five or less). In contrast, women's health does not appear to be significantly affected by transfers, be they realised or expected. Rathelot and Romanello (2013) estimate the impact on mental health of a very similar reform in France. Following the transposition of the 1996 European Directive aiming at increasing competition in the energy market, a major reform of the two state-owned electricity and gas utilities – EDF and GDF – was implemented starting in 2000. Distribution and transport services were separated from the other activities and two new independent companies were created to take charge of them. Meanwhile, the other departments of both firms were getting prepared for privatisation. As a consequence, some services were created while others were downsized and the probability for workers to change unit actually increased by 50% over the period. Individuals employed at EDF and GDF were civil servants so they could not be fired. However, this large-scale restructuration had major consequences for their working conditions. The authors investigate the impact of such changes on depression. Health information is available for EDF and GDF workers from the GAZEL database. This is a longitudinal yearly survey covering a large sample of individuals employed in – or retired from – both firms. The authors use the subsample of 5,000 workers still in employment in 2002 and estimate a fixed-effect model for the period 1999-2002. Depression turns out to have increased by a substantial amount with the effect being larger for employees who are

22
further away from retirement and for those who change unit. This suggests that transfers between units are likely to have been involuntary and that these undesired changes in working conditions have had a negative effect on health. Overall, results from both Ferrie et al. (1998) and Rathelot and Romanello (2013) suggest that, when civil servants are subject to major reorganisations including undesired transfers to other units, negative health effects are to be expected. This evidence supports the idea that constraints imposed on work (and in particular on working conditions) are detrimental to health.

**Conclusion**

This review of the literature has uncovered the key role played by choice vs. constraint in shaping the health impact of work. At the intensive margin, working long hours appears to be unambiguously detrimental to health, in particular when employees have little control on the number of hours they work and/or on their work schedule. Symmetrically, a voluntary reduction in the length of the workweek seems to have a positive impact on health behaviours. The importance of choice in determining whether work has positive or negative effects on health is even stronger at the extensive margin. The literature on retirement indeed suggests that being forced to keep on working while one would like to retire tends to have adverse health effects. Symmetrically, being forced to stop working because of involuntary job loss is equally harmful to health. Eventually, changes in work organisation which are not welcome by workers also appear to have negative health effects. Overall, changes in work status or working conditions desired by workers are often health improving and, at least, do not seem to have major negative health effects. In contrast, changes which are imposed on workers either because of managerial decisions or changes in regulation are much more likely to have adverse consequences on health.

This idea that choice vs. constraint is a key determinant of the health impact of work has rarely been tested directly in the literature, except as regards constraints on work hours. Developing empirical frameworks which would permit to test this conjecture would be most valuable, both at the intensive and extensive margins. This is, of course, a hard task since selectivity, and more generally endogeneity, raise important identification problems here, as in most of the health and work literature. Instrumental variable approaches are potentially of great help but finding good instruments remains a challenge in a field in which natural experiments cannot be easily implemented.
If choice vs. constraint turns out to be an important determinant of the extent to which work negatively affects health, a major question immediately arises: why is it the case? How come that being forced to work or to stop working and/or to accept undesired working conditions may be harmful to health? Is this effect mediated by lower self-esteem and would it be weaker if giving rise to social action rather than being dealt with by individuals in isolation? If this were the case, it would suggest that the health impact of work may have increased as individualisation developed in advanced societies, in particular in the second half of the 20th century. Assessing the mechanisms through which work-related constraints may affect health and how these may have changed over time is a challenging avenue for further research.

References

Berniell Maria Ines, 2012, "The Effects of Working Hours on Health Status and Health Behaviors", CEMFI Mimeo.


Eliason Marcus and Donald Storrie, 2009b, "Job loss is bad for your health: Swedish evidence on cause-specific hospitalization following involuntary job loss", *Social Science & Medicine*, 68(8), pp. 1396-1406.


Godard Mathilde, 2013, "Gaining weight through retirement? Results from the SHARE survey", University of Paris Dauphine mimeo.


Rathelot Roland and Lucie Romanello, 2013, "Changes at the workplace and employee health", CREST mimeo.


