CEDLAS

Facultad de Ciencias Económicas

UNIVERSIDAD NACIONAL DE LA PLATA

DOCUMENTOS DE TRABAJO

The Effect of Working Hours on Health

Inés Berniell y Jan Bietenbeck

Documento de Trabajo Nro. 237
Noviembre 2018
ISSN 1853-0168

www.cedlas.econo.unlp.edu.ar

The Effect of Working Hours on Health[☆]

Inés Berniell

Universidad Nacional de la Plata and CEDLAS

Jan Bietenbeck

Lund University and IZA

May 2018

Abstract

Does working time affect workers' health? We study this question in the context of a French reform which reduced the standard workweek from 39 to 35 hours, at constant earnings. Our empirical analysis exploits variation in the reduction of working time across employers, which was driven by the institutional features of the reform and thus exogenous to workers' health. We find that longer working hours increase smoking and decrease self-reported health, and that these impacts are concentrated among blue-collar workers. In contrast, white-collar workers' body mass index increases with hours worked.

Keywords: working hours, health, smoking, BMI
JEL codes: I10, I12, J22

[☆]We thank Lian Allub, Manuel Arellano, Manuel Bagues, Lucila Berniell, Guillermo Caruana, Laura Crespo, Dolores de la Mata, Paul Dourgnon, Romain Fantin, Daniel Hamermesh, Matilde Machado, Kaveh Majlesi, Alessandro Martinello, Claudio Michelacci, Barbara Petrongolo, Ansgar Wübker, anonymous referees, and numerous seminar and conference audiences for helpful comments. We thank IRDES for providing us with the ESPS data. An earlier version of this paper was released under the title "The Effects of Working Hours on Health Status and Health Behaviors." Author email addresses: ines.berniell@econo.unlp.edu.ar and jan.bietenbeck@nek.lu.se.

1. Introduction

Does working time affect workers' health? Data from employee surveys suggest so: for example, in a recent study of European workers, the share of respondents who stated that their work negatively affects their health rose monotonically from 19% for those working less than 30 hours per week to 30% for those working at least 40 hours per week.[1] Perceived negative health impacts from work also motivated the change to a 6-hour workday, at constant earnings, by some Swedish employers, a decision that received extensive international media coverage.[2] From a theoretical point of view, working time may affect health because of potential direct health impacts of work, such as physically strenuous work leading to exhaustion, or because of its impact on the time available for health production at home, such as longer working hours reducing the time for physical exercise.

Empirical studies of the effect of working time on health face two fundamental challenges. First, working hours are not randomly assigned, introducing bias into any naive regression estimate of the impact of hours. This bias may be due to omitted unobserved factors that influence both hours and health, or due to reverse causality, whereby health affects hours rather than the other way around. Second, estimates of the impact of working time are usually confounded by the influence of hours on income, which has an important independent effect on health (e.g. Frijters, Haisken-DeNew, and Shields, 2005; Lindahl, 2005). Both for determining the importance of working time as an input into health production and from a policy perspective, however, the effect of working hours on health keeping income constant is particularly relevant.

In this paper, we study the impact of working hours on health in the context of a French workweek reform which allows us to address both of

[1]These figures are for EU-27 respondents in the 2015 European Survey of Working Conditions. Shares of respondents who perceived negative health impacts from their work were: 19% (respondents working <30 hours per week), 26% (30-34 hours per week), 28% (35-39 hours per week), and 30% (40+ hours per week).

[2]For example, the switch to a 6-hour workday by a Gothenburg retirement home in 2015 was covered in *The New York Times*, *The Guardian*, and *Die Zeit*, among many other media outlets. Other Swedish employers who reduced or plan to reduce weekly working time at constant earnings include a Toyota production plant, several technology start-ups, and the municipal administration of Malmö, Sweden's third largest city.

these challenges. Introduced by the socialist government in 1998, the reform reduced the standard workweek from 39 to 35 hours, at constant earnings. Importantly, the laws mandating this reduction included different deadlines for implementation for firms of different sizes, which led to substantial employer-level variation in working time in subsequent years. These policy-driven, exogenous changes in working time, together with the absence of income effects, make the French context uniquely suited to study the impact of working hours on health.[3]

Our empirical analysis draws on data from a longitudinal health survey, which allows us to follow a sample of male workers from the pre-reform to the post-reform period, namely from 1998 to 2002. For each worker, we observe whether his employer had implemented the shorter workweek by the year 2002, and we use this information to create our binary treatment variable. Our main outcome variables are measures of smoking behavior, body mass index (BMI), and self-reported health. Notably, smoking and high BMI are among the leading preventable causes of death, and both outcomes have been widely studied in the medical literature on the impacts of working time, yielding mixed results.

We first estimate the impacts of the workweek reform in a difference-in-differences framework, comparing the evolution of health outcomes of workers in treated and control firms. As a complementary strategy, we also present results from lagged dependent variable models, which directly exploit the longitudinal dimension of the data. Whereas the difference-in-differences specifications assume that any gaps in outcomes between treated and control workers would have remained stable absent the reform, the lagged dependent variable models instead rely on unconfoundedness given past outcomes for identification. Thus, these two models are not nested, and we can gain some confidence in our results if they yield similar estimates. Finally, for both identification strategies, we also run regressions in which we instrument actual hours worked with our treatment variable; under the

[3] Estevão and Sá (2008) and Chemin and Wasmer (2009) study the labor market impacts of this workweek reduction. Goux, Maurin, and Petrongolo (2014) exploit within-household variation in working hours induced by this reform to examine inter-dependencies in spousal labor supply. Saffer and Lamiraud (2012) study the impact of working time on social interaction in the context of this reform.

additional assumption that the reform affected health only via its impact on working time, these specifications identify the causal effect of working hours on health.

The results consistently indicate that working time has negative effects on workers' health behavior and health. In particular, instrumental variable regressions show that one additional hour of work increases smoking by 1.4-2.4 percentage points and reduces self-reported health by 0.05-0.08 points on a scale from 0-10. Working time moreover appears to raise body mass index, but this effect is small and imprecisely estimated in the overall sample. A heterogeneity analysis reveals that while the impacts on smoking and self-reported health are concentrated among blue-collar workers, hours raise BMI only among white-collar workers.

All these estimates are very similar across our different identification strategies, and they survive a variety of robustness checks aimed at mitigating any remaining concerns about selection effects driving our results. For example, we show that effects are unchanged if we focus on a sample of likely job stayers, thus effectively ruling out that they are due to endogenous switching by healthier workers to firms which implemented the shorter workweek early on. Similarly, using the method developed by Oster (2017), we show that selection based on unobserved factors would need to be about ten times as large as selection based on observed control variables to explain away the impacts on smoking and self-reported health.

Our paper is related to a large medical literature on the health impacts of working time. Studies in that literature have focused predominantly on overtime hours and have generally found negative effects on health behaviors and health (e.g. Sparks and Cooper, 1997; van der Hulst, 2003; Kivimäki et al., 2015). However, most of those studies have failed to adequately address the empirical challenges described above.[4] Our work further connects to two strands of literature within economics that examine the health impacts of job displacement (e.g. Sullivan and von Wachter, 2009; Marcus, 2014; Black, Devereux, and Salvanes, 2015; Schaller and Stevens,

[4]One exception is the study by Åkerstedt et al. (2001), which experimentally varied workweek length, at constant earnings, among a group of female health care and day care workers in Sweden. That study found positive effects of a shorter workweek on sleep quality, mental fatigue, and heart/respiratory symptoms.

2015) and retirement (e.g. Coe and Zamarro, 2011). Whereas those papers estimate the combined effect of reduced hours and everything else changing with job loss or retirement, our study focuses on the pure working time impact. Finally, our work also relates to studies showing that health tends to improve during recessions (e.g. Ruhm, 2000, 2005). While those impacts could theoretically be driven by reductions in hours, more recent evidence has identified business cycle externalities as their probable main driver (Miller et al., 2009). To conclude, this paper's main contribution is to provide the first credibly causal estimates of the impact of working hours on health at a policy-relevant margin.[5]

2. Institutional background

Until the late 1990s, the standard workweek in France was set at 39 hours, with a legal maximum of 130 overtime hours per year and a 25% overtime wage premium. This situation changed considerably in 1998, when the newly elected left-wing government launched the reform that provides the backdrop for our study. The coalition of socialists and several smaller parties had campaigned on a program of reducing unemployment via work-sharing; in particular, the standard workweek was to be shortened from 39 to 35 hours, at constant earnings. Once in government, the coalition implemented this reduction via two distinct laws, known as Aubry I and Aubry II after the then Minister of Labor Martine Aubry. We now describe the provisions of these laws which are relevant for our analysis.[6]

Aubry I was passed in June 1998 and set the standard workweek at 35 hours in the private sector, with deadlines for implementation in January 2000 for large firms with more than 20 employees and in January 2002 for smaller firms. The reduction in hours was to be achieved through bargained agreements between employers and employee representatives at the firm level. Employers' incentives to sign such 35-hours agreements were threefold. First, after the relevant deadline, hours worked beyond the

[5] Our work is also related to a recent study by Hamermesh, Kawaguchi, and Lee (2017), who show that life satisfaction improved in Korea and Japan after an exogenous reduction in the standard workweek.

[6] This section draws heavily on Estevão and Sá (2008), Askenazy (2013), and Goux, Maurin, and Petrongolo (2014).

thirty-fifth hour were subject to the overtime wage premium, increasing labor costs. Second, the law introduced generous payroll tax cuts for firms which implemented the shorter workweek before these deadlines. Third, the negotiated agreements could allow for more flexible work schedules, the possibility of which had been very limited until then. Importantly, because workers should not bear the full costs of the reform, Aubry I required all agreements to keep the earnings of minimum-wage workers constant. In practice, previous studies have found near-zero effects of the reform on earnings also for higher-wage workers (Estevão and Sá, 2008; Goux, Maurin, and Petrongolo, 2014), a result that we further corroborate in the empirical analysis below.

Aubry II was passed in January 2000 and amended some of the rules regarding the implementation of the 35-hour workweek. For example, it introduced a transitional period with reduced overtime payments for small firms, allowing them to employ workers for 39 hours per week at almost no additional cost until 2005. The law also made it possible to achieve some nominal reduction in hours by simply re-defining working time to exclude 'unproductive breaks' (Askenazy, 2013). Moreover, firms could now implement the shorter hours on an annual basis, with a cap of 1,600 hours per worker and year. Finally, both Aubry I and Aubry II included special provisions for managers and other professionals with 'genuine autonomy' in their work: depending on their rank, these workers either could sign agreements restricting the number of days (but not hours) worked, or even were fully exempt from the new working time regulations.

In the general elections of June 2002, the conservative parties came back to power and almost immediately started to remove the incentives for employers to sign 35-hours agreements, meaning that the reform was discontinued in practice. By that time, however, many firms had already switched to the shorter workweek. As could be expected, this group disproportionately included large firms, which faced the earlier deadline for implementation (see Estevão and Sá, 2008). But it also encompassed the majority of public sector institutions, which reduced their employees' working time even though they were not formally bound by the Aubry laws. Taken together, the different deadlines for implementation and the abrupt discontinuation of the reform led to substantial employer-level variation in working time in the year 2002. Below, we exploit this variation to estimate the impact of working hours on health.

3. Data

We draw on data from the Enquête sur la Santé et la Protection Sociale (ESPS), a longitudinal survey of health, health insurance, and health care utilization. Around the time of the workweek reduction, the survey followed a representative sample of individuals in Metropolitan France, who were interviewed every four years. An important feature of ESPS is that it allows us to identify which workers were actually affected by the reform. In particular, the 2002 wave of the survey asked respondents whether the 35-hours workweek had been implemented at their current workplace, and we construct our treatment variable based on the answers to this question. In the remainder of this section, we summarize our data construction and measurement, with many more details provided in the Data Appendix.

Our analysis uses individual-level data from the 1998 and 2002 waves of ESPS. Specifically, we focus on the subsample of employees interviewed in both 2002, when information on treatment was collected, and 1998, giving us one pre- and one post-treatment observation per individual.[7] To ensure that we concentrate on workers whose hours were indeed reduced if treated, we impose some additional sample restrictions. In particular, we select individuals aged 18-61 and working more than 35 hours in 1998 (but any number of hours in 2002), and we exclude managers and high-level professionals who either were not covered by the Aubry laws or were subject to a different treatment (see Section 2). Although the remaining sample includes 744 men and 460 women, the main empirical analysis focuses exclusively on male workers. The reason is that the first-stage effect

[7]Due to sample attrition and sample refreshments, not all individuals surveyed in 1998 were also surveyed in 2002 and vice versa. The year 1998 can reasonably be assumed to belong to the pre-treatment period since virtually no employer signed a 35-hours agreement before 1999 (see Goux, Maurin, and Petrongolo, 2014). While we also obtained data from the 1994 and 2006 waves of ESPS, we did not augment our sample with these years because (1) the sampling method of the survey changed in 1998, such that only a small and unrepresentative sub-sample of 27% of workers is observed also in 1994, and (2) various counter-reforms by the conservative government after 2002 affected treated and untreated individuals in different ways, confounding any impacts of the original reform measured in 2006 (see Askenazy, 2013).

of treatment on hours is close to zero for women in this particular sample, mainly because treated women are less likely to switch from full-time to part-time work; we discuss this issue in more detail in Section 5.

We extract three health-related outcome variables from the data: an indicator for current smoking, body mass index (BMI), and self-reported health, which ranges from 0 to 10. The effect of working time on smoking has been widely studied in the medical literature and has yielded mixed results (e.g. Lallukka et al., 2008; Angrave, Charlwood, and Wooden, 2014). The proposed mechanism tying hours to smoking in these studies is usually job-related stress. Working time may also influence BMI via changes in diet or (the time spent on) physical exercise. Notably, both smoking and high BMI – in particular, a BMI higher than 25 – are among the leading preventable causes of death. Finally, working hours may affect self-reported health via a large number of physical and psychological channels.[8]

The treatment variable in our regressions is an indicator for working for an employer who had implemented the 35-hours workweek. While the exact dates that these hours reductions were carried out are not observed in the data, Goux, Maurin, and Petrongolo (2014) show that only very few firms switched to the shorter hours before the year 2000. Thus, the treatment captures exposure to the 35-hours workweek for at most 2–3 years.

Table 1 reports means and standard deviations of key variables in 1998 separately for the 588 treated and 156 control workers in the sample. While the two groups appear similar regarding age, marital status, and household income, treated workers tend to have higher levels of education. Interestingly, treated workers also work fewer hours on average already before the introduction of the 35-hours week, and they are more likely to be employed in the public sector. In contrast, there are no statistically significant dif-

[8]ESPS also asks respondents which health conditions they are currently suffering from, with answers coded according to the International Classification of Diseases. Unfortunately, due to the small sample size, estimates of the impact of the shortened workweek on even broad groups of diseases were always very imprecise and thus little informative. This motivates our focus on smoking, BMI, and self-reported health, which have relatively high incidence or variation in the sample, see Table 1. Furthermore, while the 2002 wave of ESPS contains information on other health behaviors with high incidence such as frequency of drinking and exercising, the lack of data for 1998 means that we cannot use these behaviors as outcomes in our analysis.

ferences in terms of smoking, body mass index, and self-reported health between the two groups. Below, we explain in detail how our regressions account for these observable as well as for unobservable differences between treated and control workers.

4. Empirical strategy

Two fundamental challenges arise when trying to estimate the effect of working hours on health. First, working time is not randomly assigned, introducing bias into any naive regression estimate of the impact of hours. This bias may be due to omitted unobserved factors that influence both hours and health, or due to reverse causality, whereby health affects hours rather than the other way around. Second, even if working time were randomly assigned, the estimate would still be confounded by the usual impact of hours on income, which has an important independent effect on health (e.g. Frijters, Haisken-DeNew, and Shields, 2005; Lindahl, 2005). For determining the importance of working time as an input into health production, however, the pure hours effect is the actual quantity of interest.

The French workweek reform provides us with the unique opportunity to address both of these empirical challenges. In particular, it generated policy-driven, employer-level variation in working time which was arguably exogenous from an individual worker's perspective. Moreover, since income was unaffected by the reform, the pure hours effect can be disentangled from the income effect under some additional assumptions set out below. Our first identification strategy leverages these features in a difference-in-differences framework similar to the one used by Goux, Maurin, and Petrongolo (2014). We estimate:

$$Y_{it} = \alpha_{\mathbf{i}} + \beta_1 Post_t + \beta_2 Treated_i * Post_t + \varepsilon_{it}, \qquad (1)$$

where Y_{it} is a health-related outcome for individual i at time t, $\alpha_{\mathbf{i}}$ is a vector of individual fixed effects, $Post_t$ is an indicator taking value 1 for $t = 2002$ and value 0 for $t = 1998$, and $Treated_i$ is an indicator for whether i's employer in 2002 adopted the 35-hours workweek. Note that the individual fixed effects absorb all time-invariant individual characteristics, including treatment status $Treated_i$. In other words, the specification in equation (1) controls for any constant differences between treated and control workers.

Equation 1 is a classical difference-in-differences specification with two groups and two periods. Under the assumption that differences in health between treated and untreated individuals would have been stable in absence of the workweek reform ("parallel trends"), it identifies the causal effect of switching to the 35-hours workweek. A drawback of having only a single pre-treatment period is that we cannot provide evidence in support of this assumption, for example by showing that trends in health for the two groups were parallel before the reform. To lend additional credibility to our results, we therefore also present estimates of the following lagged dependent variable specification:

$$Y_{i,2002} = \gamma_1 Treated_i + \gamma_2 Y_{i,1998} + \mathbf{X'}_{i,1998}\gamma_3 + \varepsilon_{i,2002}, \tag{2}$$

where $\mathbf{X}_{i,1998}$ is a vector of individual-level control variables measured in 1998 and the other variables are defined as above. Unlike the difference-in-differences model, which accounts for selection into treatment based on fixed group and worker characteristics, the specification in equation 2 relies on the assumption of unconfoundedness given past outcomes for identification. Thus, the two specifications are not nested, and we can gain some confidence in our results if they yield similar estimates.

The regression models considered so far aim at identifying the overall, reduced-form effect of the workweek reform on workers' health. In contrast, the policy-relevant question that this paper intends to address is how working hours affect workers' health. As described in Section 2, the Aubry laws mainly mandated a shortening of the standard workweek from 39 to 35 hours, but also introduced some other changes such as flexible work schedules. Under the assumption that the reform influenced health only via its effect on working time, we can use the treatment variable as an instrument for hours to provide a direct estimate of the impact of working hours on health. Accordingly, Section 5 below presents estimates from both the reduced-form specifications in equations 1 and 2 and the corresponding instrumental-variable regressions.

Finally, we note that from the description of the workweek reform in Section 2, one could devise at least two alternative identification strategies which are not used here. First, one may want to directly exploit variation in firm size in conjunction with the different deadlines for small and large firms. Unfortunately, this strategy is not feasible here because firm size is not observed in the ESPS data.[9] Second, one may be tempted to use part-time workers as an alternative control group. However, Oliveira and Ulrich (2002) show that part-time workers in treated firms actually *increased* their hours slightly in response to the reform, a result which we confirmed in our data. Thus, part-time workers were also affected by the reform, rendering them a bad control group.[10] In contrast, we present results from two complementary specifications which rely on distinct (untestable) assumptions for identification. Comparing the estimates from these models allows us to assess the robustness of our results.

5. Results

5.1. *Effects on hours and income*

Figure 1 shows the distributions of hours in 1998 and 2002 separately for the treatment and control groups. In both groups, the distribution peaks at 39 hours in 1998, with about half the workers reporting this amount of weekly working time. In the treatment group, this peak shifts to 35 hours in 2002, whereas the mode stays at 39 hours in the control group, pointing to a strong negative impact of the reform on working time.

Column 1 of Table 2 quantifies this first-stage effect. Panel A reports an estimate of a 2.5-hour decrease for treated workers based on the difference-in-differences specification. In comparison, the estimate based on the lagged dependent variable model in Panel B is 3.4 hours. The two regressions thus yield roughly similar results; however, both estimates fall short of the nominal 4-hour reduction in the standard workweek. Potential reasons for this difference include re-definitions of working time, implementation of the shorter hours at the annual rather than weekly level (see Section 2), or

[9] We are not aware of any dataset which contains relevant information on both firm size and health outcomes for the period before and after the workweek reform.

[10] Similarly, managers are unlikely to be a valid control group, as they were also partly affected by the reform. Moreover, because the Aubry laws were vague on who actually could be considered a manager, it is impossible to cleanly identify this group in the data.

simply an increased use of overtime work by employers who implemented the 35-hours workweek.[11]

Column 2 of Table 2 reports estimates of the effect of the reform on monthly household income. This outcome serves as a rough proxy for individual earnings, which unfortunately are not observed in the ESPS data (see the Data Appendix for details). In line with the findings from previous studies of the French workweek reduction (Estevão and Sá, 2008; Goux, Maurin, and Petrongolo, 2014), the results indicate an economically and statistically insignificant effect of the shorter workweek on income. Overall, the estimates in Table 2 thus confirm the expected impacts of the reform: it reduced weekly working hours at constant earnings.[12]

5.2. *Effects on smoking, BMI, and self-reported health*

Table 3 presents our main results for smoking, BMI, and self-reported health. Column 1 shows that working for a treated firm leads to a 6 percentage point decrease in smoking, independently of the identification strategy used (panels A and B). Under the assumption that this effect is driven only by the reduction in hours, this translates into a 1.4-2.4 percentage point increase in smoking per additional hour worked (panels C and D). Columns 2 and 3 show impacts on smoking separately for individuals who did versus did not smoke in 1998. The estimates reveal that the negative effect in the overall sample is driven primarily by quitting of baseline smokers, rather than non-initiation of baseline non-smokers.[13]

Column 4 reports a small negative impact of the workweek reform on BMI, with instrumental variable regressions suggesting a 0.03-0.04 increase for each additional hour of work. Qualitatively similar results are obtained when rather than a continuous outcome measure, indicators for being

[11] Previous studies have also found that workers who were affected by the reform reduced their labor supply by less than 4 hours; see Estevão and Sá (2008), Saffer and Lamiraud (2012), and Goux, Maurin, and Petrongolo (2014).

[12] Throughout the paper, we report results from regressions which weight observations using the sampling weights provided by ESPS, although in practice this makes little difference. Furthermore, in order to maximize sample size, we always report results for the full set of workers observed with a particular outcome; restricting the sample to workers who are observed with all outcomes gives very similar estimates.

[13] Table 3 reports estimates for smoking based on linear probability models. Results from probit specifications are very similar and are available on request.

overweight or obese are used (results available on request). However, none of these estimates is statistically significant at conventional levels. Finally, column 5 shows a negative effect of working time on self-reported health: for each additional hour worked, health decreases by 0.05-0.08 on a scale from 0-10.

Taken together, the results in Table 3 consistently indicate that workers' health improves as working hours decline. Across all outcomes, the estimates from the difference-in-differences and lagged dependent variable specifications are quite similar, which should give us some confidence that they reflect causal effects. While we are unable to provide direct evidence on the mechanisms behind these health improvements, a decrease in work-related health damage and stress and an increase in leisure time spent on health-promoting activities appear natural explanations for our findings.

5.3. Heterogeneity

An interesting question is whether the impact of the shorter workweek differs by workers' occupation or age. In Table 4, we separate workers into blue-collar and white-collar occupations and report estimates of the effect of treatment on hours and health for each of the two groups. Even though both types of workers experience the same reduction in hours, there are striking differences in the impacts of the shorter workweek on their health. In particular, whereas treatment decreases smoking by 10 percentage points and increases self-reported health by 0.2-0.4 for blue-collar workers, the estimated effects for white-collar workers are close to zero and not statistically significant at conventional levels. In contrast, BMI decreases among white-collar workers but, if anything, increases among blue-collar workers. A potential explanation for this last result is that blue-collar workers burn more calories on the job, and that they do not use the additional free time for a correspondingly larger increase in physical exercise.

Table 5 reports estimated impacts of the workweek reform separately for workers who were aged 18-39 versus 40-61 at baseline. Columns 1 and 5 show that the size of the hours reduction was about twice as large for older workers at 3.4-4.4 hours. Analogously, these workers experienced a substantially larger improvement in their health. For example, treatment increased older workers' self-reported health by a significant 0.3-0.4, whereas the estimated impact on younger workers is only about one third of that size and not statistically significant at conventional levels.

5.4. Results for women

As discussed in Section 3, the empirical analysis focuses on male workers because the first-stage effect of treatment on hours is close to zero for female workers in our particular sample. Importantly, this is not due to treated women not reducing their working time; rather, women in the control group are more likely to switch to part-time work. This pattern is clearly visible in Appendix Figure 1, which replicates Figure 1 for the sample of female workers, and is confirmed by the finding of small and statistically insignificant coefficients in first-stage regressions of hours on treatment.[14] One potential explanation for this pattern is that women find it easier to combine a 35-hours workweek with caring for their children, which might be why there is less switching to part-time work among treated women. This intuition is also shared by other researchers investigating the French workweek reform (Askenazy, 2013; Estevão and Sá, 2008).[15]

6. Robustness

6.1. Accounting for endogenous employer switching

One potential worry with the results presented above is that they are due to endogenous mobility between the treatment and control groups. For example, healthy workers might value their leisure time more and decide to switch to employers with a reduced workweek. Such endogenous sorting would not compromise our difference-in-differences estimates as long as workers' preferences are fixed over time. But it might jeopardize our lagged dependent variable estimates if sorting is not fully accounted for by

[14]The first-stage coefficient estimates are –0.21 and –1.08 in the difference-in-differences specification and lagged dependent variable specification, respectively. Separate regressions moreover indicate that treatment raises women's likelihood of working full-time by 8 percentage points. These results are available upon request.

[15]For example, Askenazy (2013) states that "a large number of women who work four days per week (i.e. women who do not work Wednesdays, when there is no school for young children) can more easily supply 35 hours of full-time work than 39 hours of full-time work." While childbearing cannot be perfectly observed in the ESPS data, we confirmed that the first stage is stronger for women in households without children.

differences in lagged outcomes. One way to rule this mechanism out is to focus on a subsample of job-stayers, that is, workers who did not change employer between 1998 and 2002. Unfortunately, because the ESPS data do not include firm identifiers, we are unable to unambiguously identify job-stayers. Instead, we present results for increasingly stringent subsamples of *likely* job-stayers.

The first part of Table 6 shows results for a subsample of workers who report having a permanent contract in both 1998 and 2002 and who intuitively are less likely to switch jobs than workers on temporary contracts. The second part of the table further restricts this subsample to workers who did not switch between the public and private sector. Finally, the last part of the table additionally excludes workers who changed occupation type or profession between 1998 and 2002.[16] Across all these subsamples and specifications, the impacts of the shorter workweek are very similar to our main estimates, even though the precision of the estimates is naturally reduced. Therefore, we can be reasonably confident that endogenous employer switching is not driving our results.

6.2. Differences between treated and control firms

As described in detail in Section 2, firms of different sizes were incentivized to implement the 35-hours workweek at different points of time. Therefore, the bulk of the variation in treatment status observed in 2002 is likely coming from differences in firm size (see also Estevão and Sá (2008), who directly exploit differences in firm size for identification). One might nevertheless be concerned that employers who did versus did not operate on a 35-hour schedule differ in ways related to workers' health, and that these differences are not constant over time (and thus not accounted for by the difference-in-differences models) and not fully captured by observable differences in baseline health (which are accounted for by the lagged dependent variable models). Here, we present two pieces of evidence that

[16] The subsample in the last part of Table 6 almost certainly excludes some workers who actually did *not* change jobs. This is because the questions eliciting occupation type and profession are ambiguous in the ESPS survey. For example, workers are asked to name the *perceived* type of their occupation, with possible answers including the very similar "qualified worker" and "specialized worker." See the Data Appendix for further details on these variables.

this is not the case.

First, we show results for a matched sample of workers with comparable socio-demographic and job characteristics. Intuitively, if workers are very similar on these characteristics, they are also unlikely to be on differential trends in health-related variables. Therefore, following the suggestion by Crump et al. (2009), we estimated workers' propensity to be treated using a logit regression, and restricted the sample to individuals with estimated propensity scores between 0.1 and 0.9.[17] As Appendix Table 1 shows, workers in this sample appear much more similar in terms of their socio-demographic and job characteristics compared to the unrestricted sample. Importantly, the regression results for the matched sample, which are shown in Appendix Table 2, are generally very similar to the ones reported above. This suggests that differential trends are not behind the health improvements of treated workers.

Second, we address the specific concern that employers who operate on a 35-hour schedule might be disproportionately located in areas where the local economy is trending upwards, a trend that might itself be related to improvements in health. To rule this explanation out, we estimated difference-in-differences specifications in which we controlled for the local unemployment rate as a proxy for economic activity. The results from these regressions were again very similar to those reported above, and are available upon request. Overall, there is thus no evidence that endogenous implementation of the shorter workweek is driving our results.

6.3. Judging the importance of selection on unobservables

As an alternative to ruling out specific ways in which selection on unobservables could drive our results, we now ask how large such selection would need to be to explain away our main effects. Our analysis builds on the methodology presented in Altonji, Elder, and Taber (2005) and recently refined by Oster (2017), which relies on comparing the coefficient of interest and the R-squared between regressions with and without control variables to gain insights into the importance of omitted variable bias. Here, we focus on the calculation of δ, which is the ratio of the impact of unobservables

[17]The characteristics used to predict treatment are the ones used in the lagged dependent variable specifications.

to the impact of observable controls that would drive the coefficient on the treatment variable to zero. As a point of reference, Oster (2017) suggests that effects for which $\delta > 1$ can be considered robust.

Table 7 shows the results from our analysis. We concentrate on the lagged dependent variable specification, which explicitly relies on the assumption that selection effects can be captured by observable control variables, and present estimates only for smoking and self-reported health, for which we find (marginally) significant effects in the overall sample. Columns 1 and 2 show that in a regression of smoking on the treatment dummy, adding controls reduces the coefficient in absolute value from -0.073 to -0.057, while increasing the R-squared from 0.004 to 0.596. The corresponding δ indicates that selection on unobservables would have to be more than nine times as large as the selection on observed controls to make the effect in column 2 go to zero, a value well beyond the threshold of one.[18]

Columns 3 and 4 present the results for self-reported health. In contrast to smoking, the inclusion of controls moves the coefficient on the treatment dummy away from zero. The corresponding δ implies that to explain away the impact in column 4, unobservables would have to move the coefficient in the opposite direction as observables, and their influence would have to be ten times as large. Taken together, the results in Table 7 thus strongly suggests that omitted variable bias is not driving our results.

7. Conclusion

In this paper, we study whether working time causally affects workers' health, a question that is important both for learning about the health production function and for informing labor market policy. To overcome problems of non-random assignment of hours and confounding income effects, our empirical analysis exploits a French reform that shortened the standard workweek from 39 to 35 hours, at constant earnings. Our difference-in-differences and lagged dependent variable models use variation in the adoption of this shorter workweek across workplaces, which is mostly dri-

[18]For our calculation of δ, we use the Stata command `-psacalc-`. Following the recommendation in Oster (2017), we assume that the inclusion of unobservables would increase the R-squared to 1.3 times the value in the regression with controls.

ven by institutional features of the reform and thus arguably exogenous from an individual worker's perspective.

Our estimates show that working time negatively affects health behaviors and health: four years after the reform was initiated, treated workers who saw their hours reduced are 6 percentage points less likely to smoke and report 0.2 units higher self-reported health on a scale from 0-10. Results are always very similar across our different identification strategies, and they survive a series of robustness checks which address potential concerns about time-varying differences between treated and control workers as well as sorting of workers across firms. This consistency across specifications makes us confident that our estimates reflect causal effects.

Our paper provides the first credibly causal evidence on the impact of working hours on health at a policy-relevant margin. As such, our results inform the current debate in many firms and countries about the potential benefits of shorter working days.

References

Åkerstedt, T., B. Olsson, M. Ingre, M. Holmgren, and G. Kecklund. 2001. "A 6-hour Working Day - Effects on Health and Well-Being." *Journal of Human Ergology* 30:197–202.

Altonji, J., T. Elder, and C. Taber. 2005. "Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools." *Journal of Political Economy* 113:151–184.

Angrave, D., A. Charlwood, and M. Wooden. 2014. "Working time and cigarette smoking: Evidence from Australia and the United Kingdom." *Social Science and Medicine* 112:72–79.

Askenazy, P. 2013. "Working Time Regulation in France from 1996 to 2012." *Cambridge Journal of Economics* 37:323–347.

Black, S.E., P.J. Devereux, and K.G. Salvanes. 2015. "Losing heart? the effect of job displacement on health." *Industrial and Labor Relations Review* 68:833–861.

Chemin, M., and E. Wasmer. 2009. "Using Alsace Moselle Local Laws to Build a Difference in Differences Estimation Strategy of the Employment Effects of the 35 Hour Workweek Regulation in France." *Journal of Labor Economics* 27:487–524.

Coe, N.B., and G. Zamarro. 2011. "Retirement effects on health in Europe." *Journal of Health Economics* 30:77 – 86.

Crump, R.K., V.J. Hotz, G.W. Imbens, and O.A. Mitnik. 2009. "Dealing with limited overlap in estimation of average treatment effects." *Biometrika* 96:187–199.

Estevão, M., and F. Sá. 2008. "The 35-hour workweek in France: Straightjacket or welfare improvement?" *Economic Policy* 23:417–463.

Frijters, P., J.P. Haisken-DeNew, and M.A. Shields. 2005. "The causal effect of income on health: Evidence from German reunification." *Journal of Health Economics* 24:997–1017.

Goux, D., E. Maurin, and B. Petrongolo. 2014. "Worktime regulations and spousal labor supply." *American Economic Review* 104:252–276.

Hamermesh, D.S., D. Kawaguchi, and J. Lee. 2017. "Does labor legislation benefit workers? Well-being after an hours reduction." *Journal of the Japanese and International Economies* 44:1 – 12.

Kivimäki, M., M. Jokela, S.T. Nyberg, A. Singh-Manoux, E.I. Fransson, L. Alfredsson, J.B. Bjorner, M. Borritz, H. Burr, A. Casini, E. Clays, D. De Bacquer, N. Dragano, R. Erbel, G.A. Geuskens, M. Hamer, W.E. Hooftman, I.L. Houtman, K.H. Jöckel, F. Kittel, A. Knutsson, M. Koskenvuo, T. Lunau, I.E.H. Madsen, M.L. Nielsen, M. Nordin, T. Oksanen, J.H. Pejtersen, J. Pentti, R. Rugulies, P. Salo, M.J. Shipley, J. Siegrist, A. Steptoe, S.B. Suominen, T. Theorell, J. Vahtera, P.J.M. Westerholm, H. Westerlund, D. O'Reilly, M. Kumari, G.D. Batty, J.E. Ferrie, and M. Virtanen. 2015. "Long working hours and risk of coronary heart disease and stroke: A systematic review and meta-analysis of published and unpublished data for 603 838 individuals." *The Lancet* 386:1739–1746.

Lallukka, T., E. Lahelma, O. Rahkonen, E. Roos, E. Laaksonen, P. Martikainen, J. Head, E. Brunner, A. Mosdol, M. Marmot, M. Sekine, A. Nasermoaddeli, and S. Kagamimori. 2008. "Associations of job strain and working overtime with adverse health behaviors and obesity: Evidence from the Whitehall II Study, Helsinki Health Study, and the Japanese Civil Servants Study." *Social Science and Medicine* 66:1681–1698.

Lindahl, M. 2005. "Estimating the effect of income on health and mortality using lottery prizes as an exogenous source of variation in income." *Journal of Human Resources* 40:144–168.

Marcus, J. 2014. "Does Job Loss Make You Smoke and Gain Weight?" *Economica*, pp. 626–648.

Miller, D.L., M.E. Page, A.H. Stevens, and M. Filipski. 2009. "Why are recessions good for your health?" *American Economic Review* 99:122–127.

Oliveira, A., and V. Ulrich. 2002. "L'incidence des 35 heures sur le temps partiel." Premières synthèses No. 07.1.

Oster, E. 2017. "Unobservable Selection and Coefficient Stability: Theory and Evidence." *Journal of Business and Economic Statistics*, pp. 1–18.

Ruhm, C.J. 2000. "Are Recessions Good for Your Health?" *The Quarterly Journal of Economics* 115:617–650.

—. 2005. "Healthy living in hard times." *Journal of Health Economics* 24:341–363.

Saffer, H., and K. Lamiraud. 2012. "The effect of hours of work on social interaction." *Review of Economics of the Household* 10:237–258.

Schaller, J., and A.H. Stevens. 2015. "Short-run effects of job loss on health conditions, health insurance, and health care utilization." *Journal of Health Economics* 43:190–203.

Sparks, K., and C. Cooper. 1997. "The effects of hours of work on health: A meta-analytic review." *Journal of Occupational and Organizational Psychology* 391:391–408.

Sullivan, D., and T. von Wachter. 2009. "Job Displacement and Mortality: An Analysis Using Administrative Data." *Quarterly Journal of Economics* 124:1265–1306.

van der Hulst, M. 2003. "Long workhours and health." *Scandinavian Journal of Work, Environment and Health* 29:171–188.

Data Appendix

Merging the 1998 and 2002 waves of ESPS

The empirical analysis is based on the 1998 and 2002 waves of the Enquête sur la Santé et la Protection Sociale (ESPS). The survey draws a random sample of individuals from an administrative database of the three main public health insurance funds in France. The selected individuals, who are referred to as "assurés principaux" (APs, "main insured"), as well as all members of their households are then interviewed for the survey. APs interviewed in 1998 were contacted again to participate in the 2002 wave of ESPS, and also in that wave, the current (i.e. 2002) members of their households were asked to participate. As usual, there was some attrition such that not all APs surveyed in 1998 are observed also in 2002; moreover, the sample was refreshed with some individuals not surveyed in the earlier years. The resulting sample is representative of 95% of the households in Metropolitan France. In our analysis, we weight observations using the sampling weights provided with the 1998 data.[19]

The data contain unique household identifiers that are consistent across all waves of ESPS. Moreover, there is an indicator for whether an individual is an AP. Together, these variables let us uniquely identify APs across the two waves of our sample. In order to identify non-AP household members across the two waves, we matched individuals on their relationship to the AP (partner, child, father or mother, brother or sister), gender, and age within households, keeping only unique matches. In principal, these matches could still be "false positives," e.g. when the AP changes partner between 1998 and 2002 and the new partner has the same gender and age as the old partner. To get a sense of the magnitude of this problem, we exploited the fact that in 1994 and 1998 (but not in 2002), the first five letters of individuals' first names are available in the data. In our final sample of males used in the empirical analysis, only two out of the 220 individuals who are observed also in 1994 did not have the same first name in 1994 and 1998 (and results are robust to excluding them from the sample).[20] This suggests that our within-household matching procedure works very well.

Construction of variables

The data contain information on individuals' age, gender, and education. For the latter variable, we collapse the available six categories into three education levels: lower secondary or less, upper secondary, and tertiary. We also use information on household size and household income.

[19]Results are qualitatively and quantitatively very similar if no sampling weights are used. For detailed information on ESPS sampling procedures, questionnaires, etc. (in French), see the ESPS website: http://www.irdes.fr/recherche/enquetes/esps-enquete-sur-la-sante-et-la-protection-sociale/questionnaires.html.

[20]We allowed for some differences in the spelling of names; for example, we would not count "JJacq" (which likely stands for Jean-Jacques) and "Jean-" as different names.

The latter is only available as a categorical variable, with different intervals in 1998 and 2002. For our analysis, we construct a continuous variable by imputing household income at the midpoint of each interval and converting the values to 1998 euros.[21] Finally, we use information on the region of residence (eight different regions) of the respondent.

We construct our hours variable from the answers to the question "Combien d'heures travaille-t-elle par semaine hors trajet?," which translates as "How many hours do you work per week, not counting commuting time?" We discard the top 1% of values, corresponding to working more than 70 hours, as many of these values are likely misreported (e.g., some individuals report working 160 hours per week).

Regarding occupation type, the data contain information on whether an employee works in the public or private sector as well as information about her occupation from two questions. The first of these questions asks employees about their perceived occupation type, with possible answers "unskilled worker / specialized worker," "qualified worker," "employee," "technician, foreman," and "engineer, professional" ("cadre" in French). The second question asks about employees' profession, with answers coded into 19 different categories. As described in the main text, managers and high-level professionals were subject to special rules under the Aubry laws and are therefore excluded from our analysis. Unfortunately, the laws were not very specific regarding the definition of these managers. In our analysis, we consider employees with the following profession to be managers or high-level professionals: artists, traders, business and executive managers, and liberal and intellectual professionals.[22] We experimented with a host of alternative definitions of managers and found that our results were robust to using any of them (details are available upon request). Finally, we considered employees with perceived occupation "unskilled worker / specialized worker" or "qualified worker" as blue-collar workers, and all other employees as white-collar workers. Again, we experimented with using alternative

[21]The highest income intervals in 1998 and 2002 are not bounded from above. In our newly-constructed variable, we set household income to missing for these intervals.

[22]In French, the categories are: "artisan," "commerçant et assimilé," "chef d'entreprise de 10 salariés et plus," "profession libérale," "profession intellectuelle, artiste, cadre fonction publique," and "cadre d'entreprise."

definitions and found that our results were robust to this.

Our three main outcome variables are an indicator for whether an individual is a current smoker, self-reported health on a scale from 0 to 10, and body mass index (BMI). For the latter variable, we exclude extreme values above 65 which are likely misreported (a BMI of 65 corresponds, e.g., to a person measuring 175cm and weighing 200kg).

Sample restrictions

As described in the main text, we focus on a sample of male workers who are aged 18-61 in 1998 and who are employed in both 1998 and 2002. We drop individuals without information on treatment status or on the health-related outcomes used in our analysis. We further drop individuals working less than 35 hours in 1998 as well as managers and professionals, who received special treatment under the Aubry laws.

Figure 1
Weekly working hours by treatment status and year

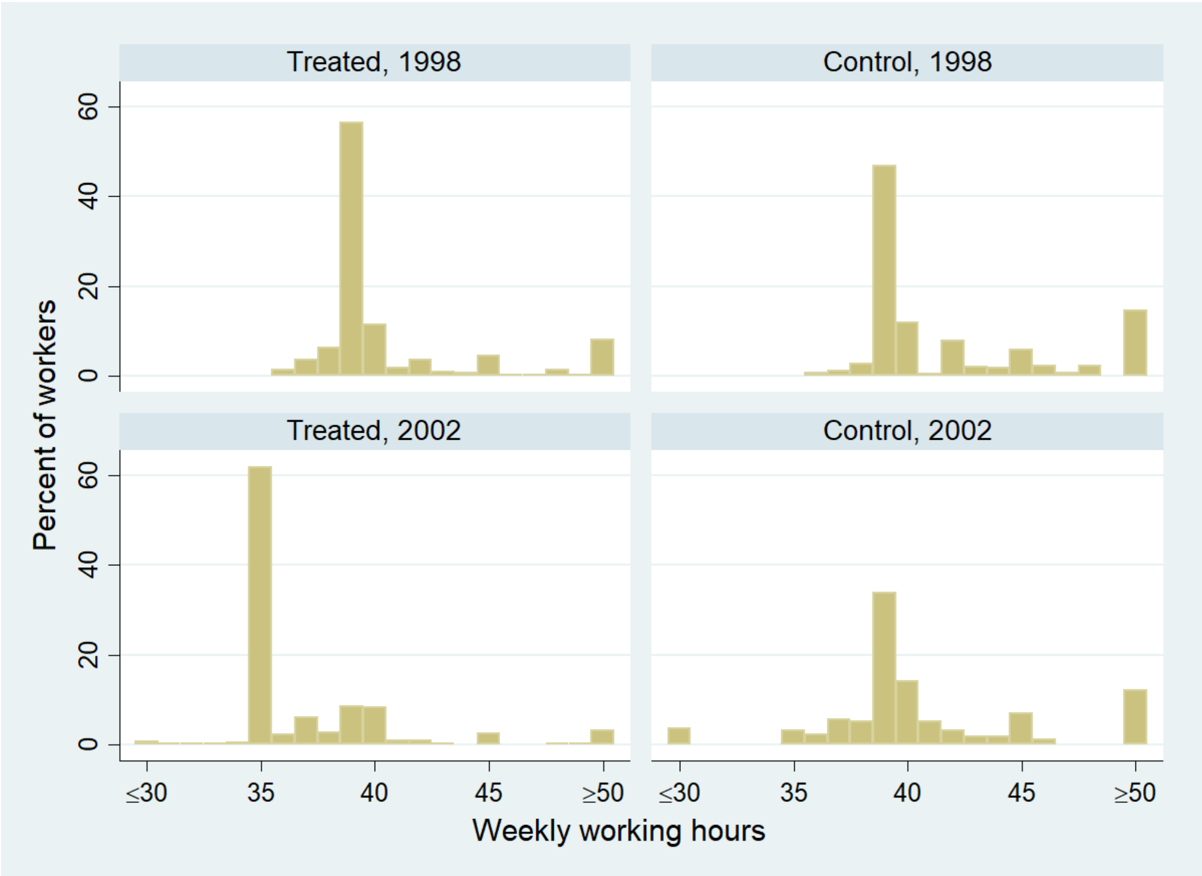


Table 1
Means and standard deviations in 1998 by treatment status

	Treated	Control	Difference [p-value]
Socio-demographic characteristics			
Age	38.16	37.23	0.93
	(8.21)	(8.32)	[0.21]
Education			
Lower secondary	0.66	0.79	-0.13
	(0.47)	(0.41)	[<0.01]
Upper secondary	0.17	0.12	0.05
	(0.38)	(0.33)	[0.14]
Tertiary	0.17	0.09	0.08
	(0.37)	(0.29)	[0.02]
Married	0.84	0.87	-0.02
	(0.36)	(0.34)	[0.45]
Household size	3.32	3.52	-0.19
	(1.31)	(1.31)	[0.10]
Household income	2033	1932	101.35
	(790)	(763)	[0.16]
Job characteristics			
Hours	40.76	42.45	-1.69
	(4.62)	(5.97)	[<0.01]
Blue collar	0.44	0.64	-0.19
	(0.50)	(0.48)	[<0.01]
Public sector	0.21	0.15	0.06
	(0.41)	(0.35)	[0.08]
Health-related outcomes			
Current smoker	0.36	0.37	-0.02
	(0.48)	(0.48)	[0.71]
Body mass index	24.81	25.18	-0.31
	(3.17)	(4.03)	[0.33]
Self-reported health	8.53	8.60	-0.08
	(1.35)	(1.24)	[0.50]
No. of workers	588	156	

Notes: The table reports means and standard deviations (in parentheses) of key variables separately for the 588 treated and the 156 control workers in the sample. Household income measures monthly income in euros. Self-reported health ranges from 0–10, with higher values indicating better health. For further details regarding all variables used in the empirical analysis, see the Data Appendix.

Table 2
Effects on hours and household income

	Hours (1)	Household income (2)
Panel A: difference-in-differences estimates		
Treated × post	–2.516***	–22.333
	(0.516)	(75.355)
Panel B: lagged dependent variable estimates		
Treated	–3.439***	–2.983
	(0.506)	(70.352)
No. of workers	744	613

Notes: The table reports estimates of the effect of workplace implementation of the 35-hours workweek on working hours and household income. Specifications in panel A control for individual fixed effects and a dummy for post. Specifications in panel B control for the dependent variable measured in 1998 as well as for age, age squared, education, marital status, household size, five occupation-type dummies, eleven profession dummies, a public-sector dummy, and eight region dummies, all measured in 1998. Standard errors in parentheses are clustered at the individual level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table 3

Effects on smoking, BMI, and self-reported health

	Current smoker			BMI	Self-rep. health
	All workers	1998=yes	1998=no		
	(1)	(2)	(3)	(4)	(5)
Panel A: difference-in-differences estimates					
Treated × post	–0.059**	–0.115**	–0.033	–0.106	0.202*
	(0.029)	(0.045)	(0.036)	(0.154)	(0.120)
Panel B: lagged dependent variable estimates					
Treated	–0.057**	–0.102**	–0.029	–0.123	0.198*
	(0.029)	(0.048)	(0.038)	(0.156)	(0.114)
Panel C: difference-in-differences instrumental-variable estimates					
Hours	0.024*	0.031**	0.018	0.041	–0.078*
	(0.013)	(0.013)	(0.021)	(0.061)	(0.047)
Panel D: lagged dependent variable instrumental-variable estimates					
Hours	0.014*	0.025**	0.007	0.029	–0.047*
	(0.007)	(0.012)	(0.009)	(0.038)	(0.028)
No. of workers	734	265	469	725	705

Notes: The table reports estimates of the effect of workplace implementation of the 35-hours workweek on smoking behavior, BMI, and self-reported health. Specifications in panel A control for individual fixed effects and a dummy for post. Specifications in panel B control for the dependent variable measured in 1998 as well as for age, age squared, education, marital status, household size, five occupation-type dummies, eleven profession dummies, a public-sector dummy, and eight region dummies, all measured in 1998. Specifications in panels C and D mirror those in panels A and B, respectively, with the difference that the treatment dummy is used as an instrumental variable for actual hours worked. First-stage F statistics are always above 10, except for in column 3 of panel C. Standard errors in parentheses are clustered at the individual level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table 4
Heterogeneity by occupation type

	Blue-collar workers				White-collar workers			
	Hours	Current smoker	BMI	Self-rep. health	Hours	Current smoker	BMI	Self-rep. health
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: difference-in-differences estimates								
Treated × post	–2.632***	–0.097***	0.144	0.388**	–2.469***	–0.015	–0.421*	–0.077
	(0.638)	(0.035)	(0.209)	(0.165)	(0.869)	(0.055)	(0.237)	(0.166)
Panel B: lagged dependent variable estimates								
Treated	–3.691***	–0.078**	0.091	0.246	–3.385***	–0.013	–0.545**	0.068
	(0.592)	(0.034)	(0.219)	(0.166)	(0.896)	(0.053)	(0.238)	(0.164)
Panel C: difference-in-differences IV estimates								
Hours		0.037**	–0.052	–0.142**		0.006	0.174	0.030
		(0.016)	(0.078)	(0.063)		(0.023)	(0.122)	(0.067)
Panel D: lagged dependent variable IV estimates								
Hours		0.019**	–0.023	–0.062		0.003	0.117*	–0.015
		(0.009)	(0.055)	(0.041)		(0.011)	(0.064)	(0.037)
No. of workers	370	365	360	350	374	369	365	355

Notes: The table reports estimates of the effect of workplace implementation of the 35-hours workweek on working hours, smoking behavior, BMI, and self-reported health, separately for workers in blue-collar occupations and workers in white-collar occupations in 1998. For information on the categorization of occupations into these two groups, see the Data Appendix. For details on specifications, see the notes to Tables 2 and 3. First-stage F statistics are always above 10, except for in columns 6, 7, and 8 of panel C. Standard errors in parentheses are clustered at the individual level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table 5

Heterogeneity by age

	Workers aged 18-39				Workers aged 40-61			
	Hours	Current smoker	BMI	Self-rep. health	Hours	Current smoker	BMI	Self-rep. health
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: difference-in-differences estimates								
Treated × post	–1.867***	–0.043	–0.005	0.115	–3.387***	–0.096**	–0.103	0.424*
	(0.674)	(0.039)	(0.206)	(0.145)	(0.862)	(0.046)	(0.239)	(0.218)
Panel B: lagged dependent variable estimates								
Treated	–2.638***	–0.027	–0.023	0.105	–4.356***	–0.107**	–0.202	0.327*
	(0.611)	(0.038)	(0.211)	(0.138)	(0.806)	(0.044)	(0.242)	(0.188)
Panel C: difference-in-differences IV estimates								
Hours		0.024	0.002	–0.060		0.028*	0.030	–0.124*
		(0.023)	(0.109)	(0.077)		(0.016)	(0.071)	(0.063)
Panel D: lagged dependent variable IV estimates								
Hours		0.008	0.007	–0.033		0.021**	0.039	–0.065*
		(0.012)	(0.065)	(0.045)		(0.009)	(0.048)	(0.037)
No. of workers	399	394	386	379	345	340	339	326

Notes: The table reports estimates of the effect of workplace implementation of the 35-hours workweek on working hours, smoking behavior, BMI, and self-reported health, separately for workers aged 18-39 and workers aged 40-61 in 1998. For details on specifications, see the notes to Tables 2 and 3. First-stage F statistics are always above 10, except for in columns 2, 3, and 4 of panel C. Standard errors in parentheses are clustered at the individual level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table 6
Accounting for endogenous employer switching

	Hours (1)	Current smoker (2)	BMI (3)	Self-rep. health (4)
Subsample A: workers with permanent contracts in 1998 and 2002				
Panel A: difference-in-differences estimates				
Treated × post	–2.754***	–0.054*	–0.118	0.225*
	(0.544)	(0.031)	(0.165)	(0.132)
Panel B: lagged dependent variable estimates				
Treated	–3.738***	–0.041	–0.167	0.169
	(0.561)	(0.031)	(0.170)	(0.128)
No. of workers	658	648	643	622
Subsample B: within subsample A, workers who did not change public-sector status				
Panel A: difference-in-differences estimates				
Treated × post	–2.736***	–0.063*	–0.118	0.235*
	(0.566)	(0.032)	(0.168)	(0.137)
Panel B: lagged dependent variable estimates				
Treated	–3.824***	–0.049	–0.178	0.185
	(0.578)	(0.032)	(0.178)	(0.133)
No. of workers	603	593	590	570
Subsample C: within subsample B, workers who did not change occupation				
Panel A: difference-in-differences estimates				
Treated × post	–3.017***	–0.046	–0.021	0.349*
	(0.984)	(0.055)	(0.275)	(0.206)
Panel B: lagged dependent variable estimates				
Treated	–3.910***	–0.066	–0.094	0.350
	(1.069)	(0.054)	(0.292)	(0.224)
No. of workers	274	268	268	255

Notes: The table reports estimates from regressions which probe the robustness of the main results in Table 3 to endogenous employer switching. In the upper part of the table ("subsample A"), the sample is restricted to workers who report having a permanent work contract in both 1998 and 2002. In the middle part of the table ("subsample B"), this sample is further restricted to only include workers who did not change from the public to the private sector or vice versa between 1998 and 2002. The lower part of the table ("subsample C") additionally restricts this sample to workers who report the same occupation type and profession in 1998 and 2002; see the Data Appendix for details on these variables. For details on specifications, see the notes to Tables 2 and 3. Standard errors in parentheses are clustered at the individual level. * p<0.10, ** p<0.05, *** p<0.01.

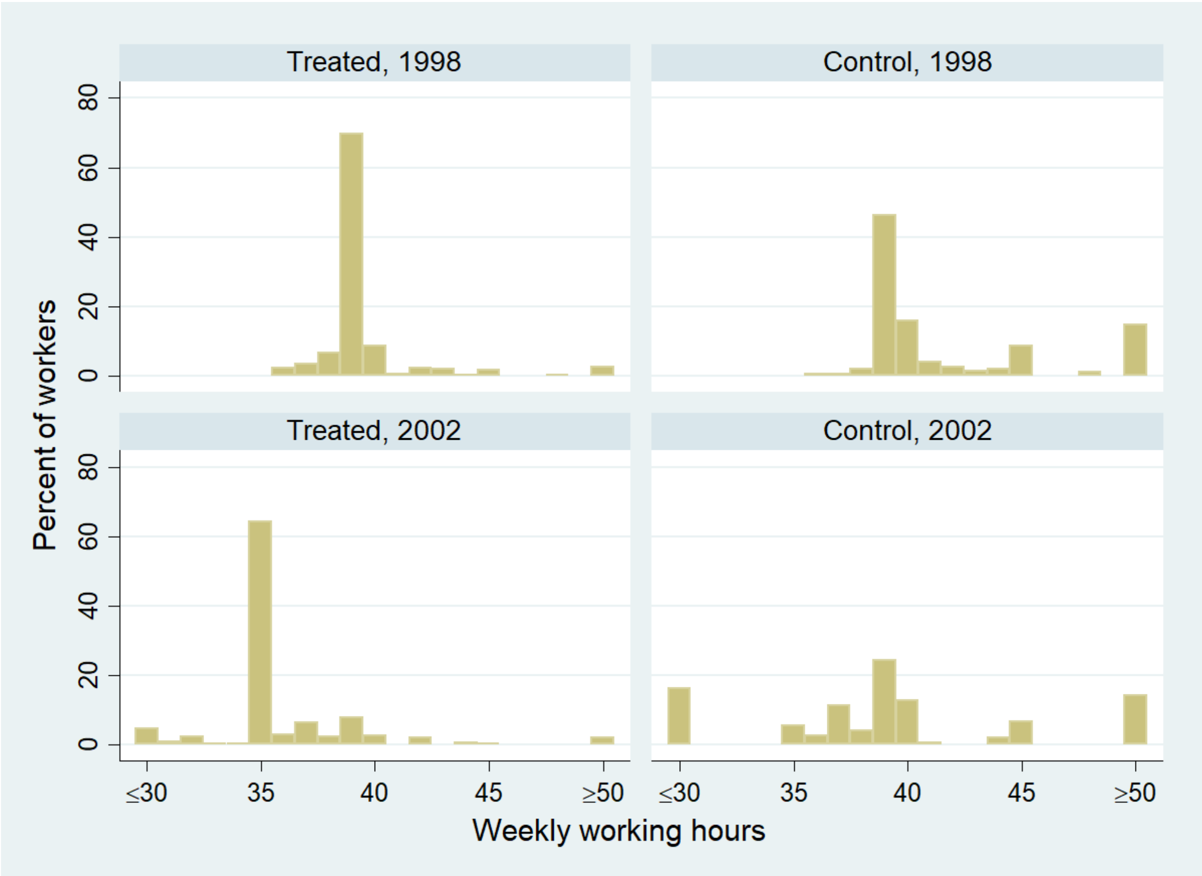
Table 7
Judging the importance of selection on unobservables

	Current smoker		Self-reported health	
	no controls (1)	with controls (2)	no controls (3)	with controls (4)
Treated	–0.073	–0.057**	0.126	0.198*
	(0.047)	(0.029)	(0.146)	(0.114)
No. of workers	734	734	705	705
R^2	0.004	0.596	0.001	0.393
δ		9.107		–10.222

Notes: Estimates based on the lagged dependent variable specification. Columns 1 and 3 report estimates from regressions of smoking behavior and self-reported health, respectively, on the treatment dummy without further controls. Columns 2 and 4 add controls as in panel B of Table 3. The final row shows the amount of selection on unobservables necessary, relative to the amount of selection on observable controls, to explain away the coefficient in the respective column. For the calculation of this δ, we use the Stata command `-psacalc-`, setting $Rmax$ to 1.3 times the R^2 in the respective column; for details, see text and Oster (2017). * p<0.10, ** p<0.05, *** p<0.01.

Appendix Figure 1

Weekly working hours by treatment status and year for female workers



Appendix Table 1

Means and standard deviations in 1998 by treatment status (matched sample)

	Treated	Control	Difference [p-value]
Socio-demographic characteristics			
Age	37.64	37.03	0.61
	(7.88)	(8.29)	[0.42]
Education			
Lower secondary	0.75	0.80	-0.05
	(0.44)	(0.40)	[0.18]
Upper secondary	0.13	0.11	0.03
	(0.34)	(0.31)	[0.38]
Tertiary	0.12	0.09	0.03
	(0.32)	(0.29)	[0.38]
Married	0.86	0.87	0
	(0.34)	(0.34)	[0.92]
Household size	3.46	3.54	-0.08
	(1.29)	(1.33)	[0.50]
Household income	1967	1900	67.22
	(753)	(747)	[0.36]
Job characteristics			
Hours	40.79	42.24	-1.45
	(4.63)	(5.24)	[$<$0.01]
Blue collar	0.56	0.68	-0.11
	(0.50)	(0.47)	[0.01]
Public sector	0.19	0.13	0.05
	(0.39)	(0.34)	[0.14]
Health-related outcomes			
Current smoker	0.38	0.37	0.01
	(0.49)	(0.48)	[0.80]
Body mass index	24.80	25.27	-0.39
	(3.10)	(4.10)	[0.24]
Self-reported health	8.49	8.62	-0.14
	(1.39)	(1.23)	[0.29]
No. of workers	464	148	

Notes: For details on the variables, see the Notes to Table 1 and the Data Appendix. For details on the construction of the matched sample, see text.

Appendix Table 2
Regression results for the matched sample

	Hours (1)	Current smoker (2)	BMI (3)	Self-rep. health (4)
Panel A: difference-in-differences estimates				
Treated × post	–2.470***	–0.080**	–0.084	0.217*
	(0.548)	(0.031)	(0.163)	(0.128)
Panel B: lagged dependent variable estimates				
Treated	–3.984***	–0.069**	–0.114	0.155
	(0.527)	(0.030)	(0.159)	(0.119)
Panel C: difference-in-differences instrumental-variable estimates				
Hours		0.033**	0.033	–0.086*
		(0.015)	(0.065)	(0.052)
Panel D: lagged dependent variable instrumental-variable estimates				
Hours		0.017**	0.029	–0.039
		(0.008)	(0.041)	(0.030)
No. of workers	612	604	594	578

Notes: The table reports estimates of the effect of workplace implementation of the 35-hours workweek on working hours, smoking behavior, BMI, and self-reported health for the matched sample. For details on the specifications, see the notes to Tables 2 and 3. For details on the construction of the matched sample, see text. First-stage F statistics are always above 10. Standard errors in parentheses are clustered at the individual level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.