STOCKHOLM SCHOOL OF ECONOMICS

Department of Economics

5350 Master's Thesis in Economics

Academic Year 2021-2022

# Does Work Sharing Work? A Two-Step Analysis on the French Work Time Reduction Experience Using Macro and Micro Data

Shannon Felländer-Tsai (23990) and Valter Frankenberg (24093)

**Abstract:** Hours per worker have steadily declined since the Industrial Revolution, which can partly be attributed to work time reduction policies. In recent years the debate and proposals of such policies have become increasingly popular. The purpose of this master's thesis is to give empirical evidence on the employment effects of one such policy, namely the French work time reduction that took place between 1998 and 2002. With the aim of reducing unemployment through work sharing, the policy lowered the standard work week from 39 to 35 hours. This thesis contributes to the literature evaluating work time reductions by studying effects on the extensive margin for workers across skill levels, economic sectors, age and type of contract. The analysis is conducted in two steps. Firstly, a synthetic control method using variation across countries is used to study aggregate employment effects for low-skilled industry and service workers, young and senior workers, as well as the prevalence of temporary contracts. Secondly, this is supplemented with an event study that uses variation within France from the French Labor Force survey to investigate potential heterogeneity in effects for workers with different skill levels. Results show that the policy had limited effects on employment for the studied groups, and that effects on workers' employment probabilities across skill levels were not heterogeneous after the reform. Reducing the standard work week may bring several benefits, but the results from this thesis indicate that such policies have limited effects on employment.

**Keywords:** Work Sharing, Work Time Reduction, Employment, Counterfactual Prediction, Synthetic Control, Event Study

**JEL:** E24, J60, J68, J80, J81

| | |
|---|---|
| Supervisor: | Abhijeet Singh |
| Date submitted: | May 15, 2022 |
| Date examined: | May 23, 2022 |
| Discussants: | Lan Yao & Mara Balasa |
| Examiner: | Magnus Johannesson |

## Acknowledgements

We would like to thank our supervisor Associate Professor Abhijeet Singh for his valuable comments and insights regarding methodology and scope of the analysis, as well as for his general support during the writing process. In addition, we want to thank Gwenaelle Moubarak for her help with our French research proposal, Professor Søren Overgaard for providing insights on the Danish working time reduction, as well as Isak Kupersmidt and Emil Bustos for their advice on methodology. We also extend our appreciation to Olof Tydén, Jacob Lange, David Huber and Elias Höckerfelt for their valuable feedback.

# Contents

# List of Figures



# List of Tables

# List of Abbreviations

| | |
|---|---|
| **ADISP** | *Archives de Données Issues de la Statistique Publique* [Archives of Data Issued from Public Statistics] |
| **COT** | Capital operating time |
| **DARES** | *Direction de l'animation de la recherche, des études et des statistiques* [Direction of the animation of research, studies and statistics] |
| **DiD** | Difference-in-differences |
| **FLFS** | French Labor Force survey (*Enquête Emploi*) |
| **GDP** | Gross domestic product |
| **ILO** | International Labor Organization |
| **ISIC** | Standard Industrial Classification of All Economic Activities |
| **ISCO** | International Standard Classification of Occupations |
| **LPM** | Linear probability model |
| **MSPE** | Mean squared prediction error |
| **OECD** | Organization for Economic Cooperation and Development |
| **RMSPE** | Root mean squared prediction error |
| **RTT** | *Réduction du temps de travail* [Reduction of the working time] |
| **SC** | Synthetic control |
| **SCM** | Synthetic control method |
| **SMIC** | *Salaire minimum de croissance* [Minimum growth wage] |
| **WTR** | Work time reduction |

# 1 Introduction

Hours per worker have since the Industrial Revolution been steadily declining across Western countries (Boppart & Krusell, 2020). This can be attributable both to efforts by the labor rights movement, as well as increasing productivity gains that have enabled higher wages, profits and more leisure time (Estevão & Sá, 2008). A motivation for reducing the standard work week has at times been to reduce unemployment through so-called work sharing, meaning sharing the labor among a larger pool of workers. In light of the long-term downward trend in weekly worked hours, further reductions are perhaps to be expected. This is reflected in recent support for further working time reductions in several industrialized countries, with trials of a four-day workweek taking place in for example Iceland, Spain and New Zealand (Haraldsson & Kellam, 2021; Kassam, 2021). Given the more recent debate regarding working time reduction (WTR) policies, studying past implementations are of interest to draw conclusions regarding their efficacy.

In this master's thesis we evaluate the employment effects of one such WTR policy in France, the *Réduction du temps de travail* (RTT), using both variation across countries and within France. The RTT was introduced between 1998 and 2002 and reduced the standard working week from 39 to 35 hours. Previous evaluations of the RTT have focused on the aggregate employment effects of the policy, but there are few studies that investigate effects for particular skill and sector groups. Our aim is to contribute to the literature studying the extensive margin employment effects of WTR policies, with a particular focus on the potential heterogeneous effects for certain groups on the labor market. Specifically, we believe effects of the policy may differ across skill levels, economic sectors, age and contract type. We therefore identify five groups believed to be especially affected by the policy. These are low-skilled labor, particularly in the services and industry[1] sector, the young and old labor force, as well as workers with temporary contracts. We begin by making use of Hunt's (1999) static labor demand model, used to study a WTR policy in Germany, to formulate priors regarding potential policy effects. Our analysis is then performed in two steps. Firstly, we employ a synthetic control method (SCM), using variation across countries through macro-level data from France and other European OECD countries, to study aggregate employment effects. Secondly, using the findings from the SCM, we supplement the analysis with an event study, employing variation within France through micro data from the *Enquête Emploi*, the French Labor Force survey, to study individual employment outcomes.

Our synthetic control (SC) analysis fails to find any significant effect of the RTT policy on employment in the groups studied. However, this null result is not due to poor pre-treatment fits, which rather show synthetic controls that mimic the development of actual France rather well. Instead, treatment effects of the policy are largely absent in the outcome variables analyzed, which indicates that the policy had little effects on employment in the groups studied. The only

[1] Industry is defined according to the International Standard Industrial Classification of All Economic Activities (ISIC)

significant treatment effect is found for the group low-skilled service workers. However, upon further investigation this appears to be the result of including Great Britain in the donor pool, where a time series break accounts for the negative treatment effect. Overall, this part of our analysis finds no results indicating any employment effects of the policy.

For the supplementary event study analysis, we are unable to make any causal claims, due to the lack of a convincing control group. Instead, we look for signs of heterogeneity in trend developments of employment probability across different skill groups, which we believe would suggest that they were affected differently by the reform. However, the event study design indicates no heterogeneous developments in employment probability for high-skilled workers compared to medium and low-skilled workers in the years directly after policy enforcement. Three years after enforcement, the employment probability for high-skilled workers remains on a higher level than that of the other two skill groups, where the probability decreases three years after policy announcement. This heterogeneity is statistically significant, but takes place too late after enforcement to be attributed to the policy.

In conclusion, our analysis finds that the RTT policy in France did not have any substantial effects on employment in any of the labor market groups thought to be most affected. There was also no indication of heterogeneity in employment effects across skill levels directly after the policy. Our study adds to the literature assessing the impact of WTR policies, and challenges work sharing motivations for implementing such reforms. Although reducing the standard hours worked per week may bring other benefits, doing so in search of raising employment might be unproductive.

This thesis starts by presenting a background in Section 2 of WTRs in general, and the French RTT in particular. Section 3 presents a short literature review of previous research on WTRs. This is followed by a presentation of a theoretical framework in Section 4. Section 5 moves on to present our research question and hypotheses. Section 6 describes the model for the SCM, and is followed by Section 7, which presents the SC results. The supplementary event study design is presented in Section 8, and its results follow in Section 9. The thesis rounds off with a final discussion and conclusion in Section 10 and Section 11.

# 2 Background

## 2.1 The Debate and History of Work Time Reduction

An example of the increasing popularity of WTRs is the Icelandic case. Iceland is often seen as a country with an extensive social safety net, an advanced economy and a well-developed healthcare system. However, the nation suffered from poor nationwide work-life balance in comparison to its Nordic neighbors in the beginning of the 21th century. According to a report from the Icelandic pressure group Alda and the British think-tank Autonomy (Haraldsson & Kellam, 2021), Iceland natives had few weekly hours for leisure and personal care due to long working hours. As a result, the Icelandic civil society and trade unions campaigned together for a reduction in working hours. Two major trials shortening the working week were launched in 2015 and 2017. Eventually, these came to involve around 2,500 workers, corresponding to around one percent of the country's labor force. Participants ranged from a number of different workplaces, such as offices, playschools, social service providers and hospitals, for whom the working week was reduced from 40 hours to 35 or 36 hours. The reductions were introduced without a decrease in salaries. The trials aimed to improve the problem of work-life balance among Icelanders, but were also an attempt to maintain, or even increase, productivity.

The Icelandic trials have gained worldwide attention due to their results – the majority of the workplaces in the trial did not experience changes in productivity or service provision. Moreover, workers' well-being improved for several indicators, such as stress, burnout and health. The success and popularity of the trials have resulted in permanent work time reductions for tens of thousands of union members all over Iceland. In June 2021, around 86 percent of the working population had reduced their weekly working week, or had gained the right to switch to a shorter one.

The Icelandic trial is one of many in a growing movement advocating for WTRs. Following Iceland, countries like Scotland and Spain have announced plans to launch a nationwide trial of a four-day workweek (Kassam, 2021). WTRs have also taken place on the firm level, with the multinational company Unilever engaging in trials of, and tech startup Bolt permanenting (Mehta, 2022), the four-day workweek. Moreover, several other voices have been raised in favor of WTRs. For instance, in 2017, the European Trade Union Institute published a report, authored by Spiegelaere and Piasna (2017), that examined recent trends in WTRs. The report outlined several motivations for such reductions, for example improved health and safety, increased gender equality, and the redistribution of work, and proposed ways through which a reduction could be implemented.

In spite of the increasing advocacy for a shortened work week, reduction in work time is far from a new idea. Already in the 1930s, John Maynard Keynes argued that increasing productivity and efficiency would lead to more leisure time in advanced economies. In a famous essay entitled *Economic possibilities for our grandchildren* (1930/2016), Keynes predicted that future generations would benefit from a fifteen-hour workweek, as productivity gains would lead to more leisure time.

On a policy level, reasons for supporting WTRs vary. Estevão and Sá (2008) show that WTRs historically have had two main objectives: improving overall well-being by increasing leisure time for workers and work sharing, which refers to sharing available work among workers more equitably. The latter is also touched upon by Keynes (1930/2016) in his essay from 1930, where he discusses the term technological unemployment, and defines it as "unemployment due to our discovery of means of economising the use of labour outrunning the pace at which we can find new uses for labour" (Keynes, 1930/2016, p. 364). This continues to be a topic of debate among policymakers across the world, through discussions on how productivity improvements and automation have affected the prevalence of certain job types (Brynjolfsson & McAfee, 2016).

Work sharing in a wider sense is also a common policy tool in times of economic crisis through short-time working or job retention schemes. These are designed to help firms maintain human capital and labor during a recession or times of economic instability. Within these schemes, employees' working time is partly cut for a limited period, which is fully or partly funded by the government (European Foundation for the Improvement of Living and Working Conditions, 2010). Short-time working schemes were for instance employed in Sweden (Regeringskansliet, 2020), the UK, Australia and New Zealand during the COVID-19 pandemic (Andrews et al., 2021).

The concept of work sharing has been the explicit aim of several national reductions in standard working hours, such as in Germany in the 1980s (Hunt, 1999) and in Canada in the late 1990s (Skuterud, 2007). It is also present in the more recent debate on WTRs, and is proposed as one motivation for such policies by Spiegelaere and Piasna (2017).

In spite of its popularity in the public debate, the idea of work sharing has been subject to criticism. The most important one concerns the lump of labor fallacy, which means that the labor supply in an economy is not a fixed amount that can be perfectly redistributed. According to Spiegelaere and Piasna (2017) there are several reasons for this. For instance, fixed costs of employment raise hiring costs relatively for firms in the case of a WTR. Moreover, there is also imperfect substitution between workers, and it is unlikely that new employees are perfect substitutes for existing ones. Finally, not all jobs have working tasks which can be easily divided without causing productivity losses. In other words, the phenomenon is not as simple as it may seem.

Recent trials and adoptions of WTRs begs the question what the economic effects of such a policy tool are. This thesis aims to cast light upon one such French adoption, and study its employment effects for different groups in the labor market. Investigating such employment effects is of interest, since it helps answer the question whether a WTR is capable of fulfilling its often intended goal – reducing unemployment.

## 2.2 The Réduction du Temps de Travail

The length of the work week has historically been subject to much discussion in France, which led to a WTR in 1982. Already at this point the aim was to move from a 40 to a 35-hour work week. However, due to budgetary deficits and difficult negotiations between unions and employers, the policy only constituted a one-hour reduction, setting the work week to 39 hours (Askenazy, 2013). The idea resurged in the 1990s, when work sharing was proposed as a policy tool to combat high unemployment. Although the evidence of work sharing being successful was contested, the concept gained traction in France, resulting in the Robien law, which was introduced in 1996 (Askenazy, 2013). It gave reductions in social security contributions to firms who voluntarily reduced hours worked by 10-15 percent. Obtaining the reductions were conditional on either proportional job creation, or a promise to not lay off workers. According to Fiole and Roger (2002), around 280,000 French employees had their working hours reduced within the Robien law, corresponding to approximately one percent of the workforce.

Despite intense debate and vocal opposition from many economists, the Robien law was followed by two more extensive laws, Aubry I and Aubry II, which together set a road map for the RTT policy. With these laws, France transitioned nationally from a 39 to a 35-hour workweek (Askenazy, 2013). By 2000, firms in the private sector with more than 20 employees had to introduce a 4-hour reduction through the Aubry I law. Smaller firms in the private sector could defer implementation until 2002, dictated by the Aubry II law, which regulated the RTT in more detail. The announcement of policies for both small and large firms occurred in June 1998 (Askenazy, 2013). According to Bilous (2000) the consequence of the laws was in practice that the hours-threshold where firms needed to start paying workers overtime was lowered from 39 to 35 hours, and stimulated implementation through negotiations between unions and employers, both on the sector and firm level.

The principal aim of the policy was to lower high unemployment levels of around 12 percent through work sharing, as well as to introduce more leisure time to workers and a higher level of flexibility in labor contracting (Askenazy, 2013). The announcement of the policy was accompanied by financial incentives for immediate reductions in working hours. Firms setting the working time to a maximum of 35 hours weekly within a collective agreement, and that committed to either not abolishing or creating jobs, received progressive subsidies for social security contributions. These subsidies were the highest for workers earning the minimum wage, in France referred to as the SMIC.

Full-time workers who were paid the SMIC did not experience a drop in their nominal monthly earnings, despite the RTT. Instead, they received a bonus equal to four paid hours from their employer, which was to be subsidized by the government through the social security cuts (Bilous, 2000). In the case of working time being reduced to less than 35 hours, SMIC earners were still to be compensated through this mechanism, but in proportion to the WTR below 35 hours. The same was to be applied to part-time SMIC earners whose working time was reduced after the policy. The intended legislation around part-time work at the SMIC level also stirred legal controversies. Aubry II stated that part-time SMIC workers whose working time had not been reduced would not receive the nominal bonus targeted to full-time workers. However, the French Constitutional Council (*Conseil constitutionnel*) censured this measure as it was seen as discriminatory. As a result, the measure was proportionally extended to part-time SMIC workers whose working hours were unaffected by the policy (Bilous, 2000). In contrast, the policy did not contain any wage compensation mechanisms for other wage groups.

The introduction of the RTT also implied changes to the regulation of part-time work. According to Bilous (2000), part-time work was classified as that with a working time less than 20 percent of the standard working hours before policy implementation. Afterwards, part-time work was defined as all work with contracted working hours below 35. Moreover, preceding the policy, there had been a 30 percent reduction in social security contributions for employing part-time workers. This was abolished in 2001. However, cuts in social security contributions for part-time work could still be obtained, but was contingent on the previously mentioned scheme.

An additional element of the reform was the introduction of a more flexible work schedule, which was in line with an additional policy goal of company modernization (Bilous, 2000). According to Askenazy (2013), there was limited flexibility of work schedules before the reform. After the reform, firms that introduced the reduction through signed agreements between employers and trade unions could employ greater time schedule flexibility. The flexibility was formalized within the Aubry II law, where hours went from being counted on a weekly to an annual basis. There was also special consideration paid to managers, for which a dramatic reduction in hours was predicted to have especially adverse effects. Here, the law allowed for working time to be calculated by days per year, meaning managers could work much longer hours without it being classified as overtime (Askenazy, 2013).

The Aubry laws only applied to the private sector. The public sector implemented a shorter working week to varying degrees. For instance, 25 percent of public sector employees were completely exempted from the policy change, among them teachers in all levels of schooling (Askenazy, 2013). Several groups of public administration loosely applied the RTT, mostly in the form of allowing employees to use the reduction through extra vacation at a later stage (Askenazy, 2013).

As stated by Bilous (2000), industrial relations in France are subject to heavy regulation, which meant that the RTT negotiated between employers and unions caused a shock to a detailed and tight regulatory framework. This led to effects for several dimensions of the French labor market. For instance, according to Askenazy (2013) the policy affected aspects such as workplace organization, working conditions and wages. Before investigating these effects further, we will present a short literature review of previous studies on WTR.

# 3 Literature Review

## 3.1 Previous Research on Work Time Reduction Policies

Several studies have investigated the labor market effects of the French RTT policy. Many of them use the French Labor Force survey (FLFS) as their main data source. One effort to empirically investigate the employment effects of the RTT was done by Chemin and Wasmer (2009). They recognize the lack of obvious identification strategies in investigating the policy, since it was introduced nationally. However, they propose the region Alsace-Moselle as a viable control group, due to variation in national holidays in this region. The discrepancy can, according to the authors, be used in an identification strategy that compares employment effects in Alsace-Moselle with the rest of France. Using the FLFS, the authors first establish that hours worked were affected differently in Alsace-Moselle in comparison to the rest of France. However, they later fail to find significant effects of the reform on employment outcomes. Although the mean effect of job creation is positive, it is not found to be significant.

A second study making use of the FLFS is that of Estevão and Sá (2008). The authors employ the difference in enforcement dates for small and large firms for identification. Since small firms did not have to enforce the policy until 2002, they use small firms as a control group in order to study policy effects on wages and employment within a difference-in-differences (DiD) framework. The authors find that the policy had heterogeneous effects for men and women, where the hourly wages for men earning around the minimum wage increased. Results also suggest an increased job turnover as a result of the policy. The authors find no effect on the general employment level.

Other European WTRs have also been subject to evaluation. One example is West Germany, where unions started to reduce standard working hours across industries in the 1980s. Like the French RTT, the aim of this reform was to reduce unemployment. Hunt (1999) uses the German Socio-Economic Panel to study effects from these reductions on employment, wages and hours worked between 1984-1994 using a fixed effects model. Hunt finds indications that the reductions led to negative employment effects among men. However this finding was not significant.

The phenomenon of work sharing from a macro perspective has more recently been studied by Cárdenas and Villanueva (2021). In their paper, they use micro data on the economically active population and wages to simulate a hypothetical Spanish WTR from 40 to 35 hours, in order to study changes to employment, hours worked and wages. According to their estimation, the WTR would have led to the creation of 560,000 full-time equivalent jobs. They then proceed to study macroeconomic effects of such a WTR using the fact that the created jobs would lead to an increase in the labor share. Through an extension of the Bhaduri–Marglin model (1990) the authors find that the WTR would lead to a 1.55 percentage point increase in GDP growth.

A common criticism to work sharing stems from the assumption of profit-maximizing firms. This is presented by, for instance, Calmfors and Hoel (1988). Under the assumption of profit-maximizing firms, a WTR will lead to a negative scale effect on output, due to the policy reducing the marginal product of employment (workers) more than the cost per worker (wages). According to this framework, even if a WTR has the possibility of producing positive employment effects, these will be counteracted by the negative output effect. This output assumption has been studied empirically for the French WTR. Gilles (2015) studies the capital-operating time (COT) and shift work in the French manufacturing sector, using survey data on COT and firms' WTR agreements. Employing conditional DiD estimators, the study finds that for French firms with more than 20 employees the WTR policy did not lead to a reduction in COT, contrary to the theoretical negative scale effect on output.

### 3.2 Heterogeneous Labor Market Effects for Low-Skilled Labor

Previous research on WTRs is concerned with its aggregate labor market and employment effects. Heterogeneity in the labor market is an established and much studied phenomenon among economists. One particular focus has been devoted to heterogeneity across skill level, building on the human capital theory by Becker (1962) and Rosen (1976). It is therefore of interest to study how policies and other shocks impact groups in the labor market differently. Low-skilled workers face the most adversity in the labor market, and has therefore been singled out as a particular group of study. One example is Card and Krueger's (1993) seminal paper, which shows that the employment effects of a policy may contradict economic theories. Their DiD study on a minimum wage increase in fast food restaurants from $4.25 to $5.05 in New Jersey shows that employment did not decrease. If anything, results indicate that employment increased. This went against the mainstream economic theory at the time.

Additionally, the low-skilled labor market has been studied in relation to supply shocks, such as increased immigration. Card (1990) studies the Mariel Boatlift in 1980, which caused an inflow of Cuban immigrants to Miami, increasing the labor force by 7 percent. This increase particularly affected the low-skilled labor force, as a majority of the immigrants had lower educational attainment and were relatively low-skilled (Card, 1990). Despite the influx of labor, the study fails to find the predicted adverse effects for low-skilled workers that were already in Miami. Instead the wages and unemployment level remained relatively stable. Peri and Yasenov (2017) return to Card's heavily influential study and attempt to improve on the analysis by using a SCM. In general, their findings support those of Card and show that if any labor market effects of the immigration shock occurred, they were relatively small.

# 4 Theoretical framework

### 4.1 Hunt's Static Labor Demand Model

Several economic theories have been formulated to predict the effects of different WTR policies, such as those by Calmfors and Hoel (1988, 1989). These models tend to predict negative employment effects from a reduction in working hours, except for some specific cases. One model that illustrates potential effects of a WTR is Hunt's (1999) simple static labor demand model, used to study the WTR within German unions in the 1980s. After presenting the model and its predictions, we will use it as a foundation on which we will formulate hypotheses regarding potential effects of the RTT.

Hunt considers a firm that takes as given standard hours ($h_s$), hourly wages ($w$) and the rental rate of capital ($r$). The firm chooses workers' actual hours ($h$), employment, or number of employees, ($N$), and capital ($K$). In addition, the firm faces a fixed cost of employment ($f$) and an overtime premium ($p$). The setup gives the following profit maximization problem:

$$\max_{h,N,K} g(h,N,K) - whN - fN - pw(h-h_s)N - rK \tag{1}$$

The firm has optimal hours per worker equal to $h^*$. If the fixed cost of employment, $f$, is high, the considered firm hires workers on overtime ($h^* > h_s$). A WTR policy that lowers standard hours, $h_s^0 > h_s^1$, will increase labor costs due a larger share of the wage having to be paid at the overtime premium, $p$. Within this framework, Hunt identifies one scale effect and two substitution effects from a WTR. The scale effect reduces employment and hours per week of the firm, causing the number of workers and hours worked to fall. There will also be a substitution effect from labor to capital. The second substitution effect is from workers to hours per worker, with firms hiring existing workers on overtime. This effect is contingent on firms already hiring workers on overtime before the policy. This can be seen by looking at the marginal cost of hiring an additional worker for $h^*$ hours ($MC_N$) and the marginal cost of hiring existing workers on $h^*$ additional hours ($MC_h$).

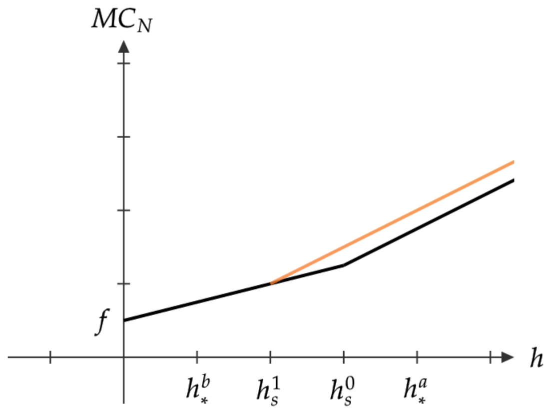
$$MC_N = wh^* + f + pw(h^* - h_s) \tag{2}$$

$$MC_h = (1+p)wh^* \tag{3}$$

From Equation 2 and Equation 3, one observes that the marginal cost of overtime ($MC_h$) is unaffected by any policy change in standard hours. It is instead determined only by the wage and overtime premium. On the other hand, the marginal cost of hiring an additional worker ($MC_N$) increases, since a larger share of the wage is paid at the overtime premium. The firm will thus substitute from workers to hours in this framework, which will decrease employment. This is illustrated graphically in Figure 1 and Figure 2, where a firm before the policy has optimal hours above standard hours at $h_*^a$. The orange and black lines represent the marginal cost after and before the policy respectively. We see in this case that the marginal cost of an additional worker ($MC_N$) has increased for this firm, whereas the marginal cost of an additional hour ($MC_h$) has remained constant.

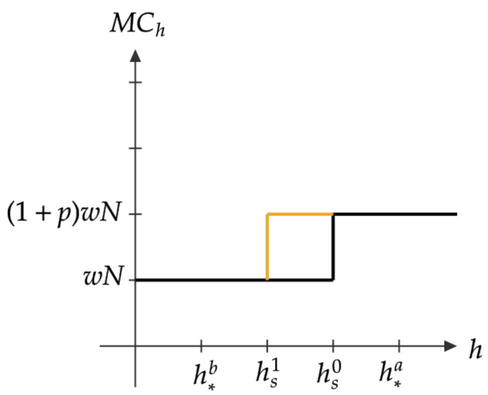
From Figure 1 and Figure 2 one can also consider the case where optimal hours before the policy is below the old and new standard hours at ($h_*^b$). Under the assumption that full-time workers must work at least standard hours, firms will substitute from hours to workers, and employment will rise. This is because the marginal cost of employment ($MC_N$) falls, while the marginal cost of hiring workers on additional hours ($MC_h$) remains constant. The scale and substitution effects thus depend on the optimal hours of the firm before a WTR. To conclude, this model tells us that if positive employment effects are to arise from a WTR, firms must substitute from hours to workers.

**Figure 1:** Marginal cost of an additional worker ($MC_N$) plotted against working hours ($h$)




Note: Black line represents the marginal cost before, and the orange line the marginal cost after, a reduction in standard working hours.
Source: Hunt (1999).

**Figure 2:** Marginal cost of an additional hour of work ($MC_h$) plotted against working hours ($h$)




Note: Black line represents the marginal cost before, and the orange line the marginal cost after, a reduction in standard working hours.
Source: Hunt (1999).

## 4.2 The French WTR Design

The simple model described above opens up for several possible scenarios. Some French deviations from the model that motivate further investigation of the topic will now be considered. According to Askenazy (2013), a key concern for the French policymakers designing the RTT was that unit labor costs were to remain unchanged after the reform. They believed that any positive employment effects from the reform might be offset if hourly wages rose proportionally, as suggested by Hunt (1999). However, unit labor costs were expected to decrease due to hourly productivity gains from the WTR, as a result of the flexibility in work schedules that accompanied the reform. Furthermore, social security contributions were cut to further decrease unit labor costs. Therefore, policymakers did not believe the reform would lead to greater unit labor costs.

An important policy element was the condition to obtain the social security cuts. For a firm to be able to benefit from the cuts, it had to commit to either creating or at least not shedding jobs (Bilous, 2000). It is possible that this policy element impacted the effects that the policy had on employment. According to Bosch and Lehndorff (2001), if the state financially supports WTR policies adopted by unions and employers, it can considerably improve the employment effectiveness of such policies, by helping firms with the organizational challenges that WTRs entail.

As previously mentioned, the RTT's goal to offset a proportional wage rise was handled differently for the minimum wage. The social security contribution cut was highest for minimum wage workers, whose nominal monthly wage also stayed the same after the policy through a bonus equal to the four 'lost' hours. Askenazy (2013) writes that this bonus was meant to decline in real terms as the minimum wage increased over time. The minimum wage in France is similar to the one modelled by Hunt, as it is exogenous. Since social contributions were greatly reduced for this group of workers, it is possible that it affected the substitution effect between hours and workers in Hunt's framework. If firms adopting the policy received a lower marginal cost for hiring new workers, whereas the marginal cost of overtime remained the same, it is possible that the reform had positive employment effects for this group of workers.

Moreover, the policy was accompanied with changes to the overtime premium. As can be seen in Table 1, the highest overtime pay had to be paid earlier after the reform. In Hunt's model, this means that the slopes of the marginal cost of additional hours in Figure 2 change. Since this is the driving factor behind the substitution effect between hours and workers, it is probable that the changes to the overtime policy had an impact on this substitution effect in the French case.

**Table 1:** Overtime premium in France before and after the RTT

| Before policy | | Transition (2000-2001) | | After policy | |
|---|---|---|---|---|---|
| Hours | Premium | Hours | Premium | Hours | Premium |
| 40-47 hours | 25% | 36-39 | 10% | 36-39 | $\geq$ 25% or time off |
| >47 hours | 50% | 40-43 | 25% | 40-43 | $\geq$ 25% or time off |
| | | >44 hours | 50% | >44 hours | 50% |

Source: Bilous (2000).

Firing and hiring costs will also impact the outcome predicted by the model. Kramarz and Michaud (2010) study these using panel data on French firms with measures on hiring and terminating costs, as well as measures of employment flows, for years 1992 and 1996. They find that collective terminations are more expensive than individual terminations, and that the former lead to large fixed costs. Further, they find that hiring costs are small, and are only present when hiring on permanent contracts. In other words, the marginal cost for hiring an additional worker in France was low before the policy, which would further impact the substitution effect in Hunt's framework.

Additionally, the model predicts that in the case of optimal hours being above standard hours, firms will substitute from workers to overtime hours. This would in practice mean that working hours should increase, and not decrease, as a result of the reform. However, according to Askenazy (2013) annual working time of French employees decreased by around 7 percent between 1995 and 2003, which suggests that the policy did not lead to a higher use of overtime. In addition, there is also a limit to how many overtimes hours a worker can be hired to perform. Employment protection legislation, as well as the number of hours in a day, limits the number of possible overtime hours, something that is not incorporated in Hunt's model.

In conclusion, there are several elements of the policy which call into question the predicted negative employment effects of a WTR from Hunt's model. We will now move on to discuss the importance of heterogeneity of the labor market in the context of the model.

## 4.3 Potential Heterogeneity in WTR Effects Across Workers

Hunt's model illustrates the general effects of a WTR policy. As previously presented, it is not adapted to a specific type of WTR. In addition, it does not take into account heterogeneity among firms or workers. Yet, labor market policies of any kind are likely to affect different employment groups heterogeneously. We therefore believe it is relevant to investigate policy effects on a more granular level. Specifically, a WTR reform will impact those workers that have little flexibility in their working hours. This indicates that the French case might lead to different employment outcomes, compared to Hunt's model.

The nature of the policy supports the idea of varying effects across labor market groups. Firstly, one can imagine that the reduction in working hours will be more strictly enforced for sectors where workers work in shifts. Such work is often performed on-site, with the wage strictly related to the number of hours worked. It is therefore likely that shift work required an especially large adjustment to the RTT, with the introduction of new shifts. One can therefore imagine that the RTT policy led to positive employment effects for shift workers. Shift work occurs on all skill levels on the labor market. Examples of common low-skilled shift work include work within the services sector, such as within hospitality or cleaning. It is also present in the industry sector through, for instance, certain occupations within manufacturing. Around 45 percent of French manufacturing workers were employed on shift work in the mid 1990s (Gilles, 2015). On the other hand, the tasks of most health care workers and pilots are examples of high-skilled work that is done in shifts.

Conversely, there are occupations that are likely to be less affected by the policy, for example employees with discrepancies between actual and reported hours. The reasons for flexibility around actual working hours vary. Whether due to the culture or nature of the work, there are several examples of occupations where hours are not carefully logged or the standard work week is not always strictly followed, such as white collar work. We therefore believe that individuals with such occupations will be less affected by the policy.

Moreover, one can imagine that another group especially affected by the RTT policy were those earning the minimum wage. The cuts in social security costs for firms that reduced working hours were the highest for SMIC workers. These workers tend to have a lower skill level (Cahuc & Michel, 1996). One can thus hypothesize that this group will experience stronger employment effects compared to other wage groups, since the social security cuts substantially lowered the marginal cost of hiring new workers at the minimum wage. Previous literature has already established ways in which this group can be especially sensitive to certain labor market policies, such as Card and Krueger's (1993) study on workers in the fast-food sector. With this reasoning in mind, we formulate a hypothesis regarding potential policy effects for minimum wage workers in shifts. We firstly proxy minimum wage with low skill, and shift work by work in the services and industry sectors. As such, we believe that the RTT led to positive employment effects for these low-skilled workers.

An additional characteristic of low-skilled shift work, is that the use of part-time is frequent, especially in the services sector. According to a *Direction de l'animation de la recherche, des études et des statistiques* (DARES) report by Ulrich and Zilberman (2007) using data from the 2005 FLFS, 21.7 percent of workers in the services sector were employed on part-time contracts. Even though this data point is from after the policy implementation, we note that it is the highest among the studied sectors in the report. According to an additional DARES report by Pak (2013), the use of part-time work in France reflects varying labor-needs throughout the working day. As exemplified by the services sector, the prevalence of part-time work is explained by the need to compensate for peaks in the working day where more labor is needed, in for instance distribution, cleaning or hospitality. Such work is thus of a different nature compared to white collar work. Moreover, according to the Ulrich and Zilberman report (2007), half of all French part-time workers had an education level inferior to a high school degree in 2005. We thus conclude that part-time work in France is common for low-skilled service workers.

The notion of optimal hours is of central importance in determining the direction of employment effects from a WTR within Hunt's framework. However, it seems that the optimal hours of firms employing low-skilled workers vary, due to the frequent use of part-time work. Such characteristics are not taken into account in Hunt's framework, since the model requires firms to hire only full-time workers. The employment effects of a WTR on sectors with low-skilled labor and frequent use of part-time in Hunt's framework are thus not clear. We therefore believe it is reasonable to hypothesize positive employment effects for this group.

The labor market does not only exhibit heterogeneity across skill and sector levels, but also for workers of different ages. The policy may have affected the young and senior workers in different ways. The labor market entry among the young is one such example. Before the policy, France experienced high rates of youth unemployment, amounting to around 28 percent (OECD, 2022c). If firms increase employment after the policy, it is reasonable that a large part of this expansion comes from the available pool of unemployed workers within this age group. More generally, we make the assumption that firms' policy adjustment to the RTT will be observed in their hiring behaviour of junior and inexperienced workers. As labor rights protect existing employees from volatile changes, potential new hires will to a larger extent be affected. In addition, young people entering the labor force tend to disproportionately take low-skilled work (Vanvitelli & Pastore, 2018), a skill group that we hypothesize will experience positive employment effects after the policy. We therefore believe low-skilled youth employment will be positively impacted by the RTT.

Similarly, it is possible that the policy brought particular effects on the other end of the age spectrum through labor market exit and retirement. On one hand, workers may want to retire earlier as a result of the policy if the monthly salary, and thereby the alternative cost to retirement, decreases. On the other hand, it is also possible that workers will decide to retire later, as working will be less burdensome with a shorter working week. When formulating a hypothesis regarding the behavior of this group post-RTT, we believe it is important to take into account the historical debate on pension conditions and retirement in France. Policies altering the retirement conditions have historically been met with strong public opposition in France. For instance, in December 1995 a general strike in Paris overturned a proposal to reduce pension benefits for public sector workers (Marier, 2005). General strikes against increasing the retirement age have also taken place more recently in 2010 (France24, 2010) and 2019 (BBC, 2019). This strong vocal opposition, especially against an increased pension age, can be interpreted as an advocacy of leisure time for older workers. The RTT gave increased leisure time to workers, and therefore addressed some of the concerns raised by the French public regarding retirement. We therefore view it as unlikely that workers will choose to retire earlier as a result of the reform. Rather, we hypothesize the policy will result in senior workers choosing to stay longer in their job. We believe this will be the case for senior workers of all skill levels.

Apart from heterogeneity across skill-sector groups and ages, differences among workers also exist through contract types. Hiring a worker on a temporary contract is often cheaper compared to a permanent contract, as shown by Kramarz and Michaud (2010). This dimension is not accounted for in Hunt's (1999) model. In addition, it is possible that the RTT leads to uncertainty regarding firms' demand for labor, due to potential productivity gains from the policy, which were hypothesized by policymakers. Firms might pass on this uncertainty to employees and in a transition phase employ more people through temporary contracts directly after the policy. It is therefore possible that positive employment effects from the policy arose through the use of temporary contracts.

Since Hunt's model differs from the French scenario, and does not take heterogeneity of workers into account, we hypothesize positive employment effects of the policy for the discussed groups on the labor market. These are presented in the next section.

# 5 Research Question and Hypotheses

Bearing in mind that previous research has mostly studied aggregate effects of the RTT reform, we wish to focus on particular groups in the labor market we believe were especially affected by the policy. Our research question is therefore:

*What were the employment effects of the French RTT reform for low-skilled workers, particularly in the industry and services sectors, young workers, workers close to retirement and the prevalence of temporary contracts?*

Building on Hunt's (1999) framework, as well as our considerations regarding heterogeneity, we develop some hypotheses with regards to our research question. Firstly, we believe that the policy will affect workers across skill levels differently, and give rise to positive employment effects for low-skilled workers compared to workers of other skill levels, especially within the industry and services sector. Secondly, we believe that the policy will positively impact low-skilled youth employment. Thirdly, we believe that the senior labor force will choose to retire later as a result of the reform. Lastly, we believe that the policy will give rise to an increase in the use of temporary contracts.

In order to answer the research question and shed light on these hypotheses, two methods with different data sources are used to give deeper insights into the policy effects. We begin with studying the aggregate effects of the policy using macro data from different countries within an SCM. In addition, micro-level data from the FLFS are used within a supplementary event study to further shed light on the policy's employment effects, and provide statistical power to the analysis.

# 6 Synthetic Control Method

## 6.1 Conceptual Framework

The SCM has in the past 20 years become an increasingly popular way to investigate causality. According to Abadie (2021) the method rests on the idea that a combination of untreated units provides a better counterfactual when studying causal effects of a treatment, rather than one single control unit. It was first introduced by Abadie and Gardeazabal (2003) in order to study the economic costs of the conflict in the Basque country in Spain. By creating a 'synthetic' Basque country, consisting of a combination of other, untreated, Spanish regions, the authors were able to obtain a counterfactual that could mimic what would have happened to the Basque country in the absence of the conflict.

Following Abadie (2021), the setup of the synthetic control method can briefly be described as follows.[2] Consider $J+1$ units, where $j = 1, 2, ..., J+1$, and where the the first unit ($j = 1$) experiences some form of treatment, the effects of which we want to investigate. The remaining units, $j = 2, ..., J+1$ consist of what is referred to as the donor pool, and are used to construct the counterfactual outcome in the absence of treatment. The treated unit and the donors are observed for $T$ periods, divided into $T_0$ pre-treatment periods and subsequent post-treatment periods. We define $Y_{jt}$ as the outcome variable of interest. For each unit $j$ we also collect data on a set of $k$ predictors $X_{1t}, ..., X_{kt}$, which are variables thought to predict $Y_{jt}$. These may include pre-intervention values of the outcome, referred to as outcome lags, as well as other economic predictors of the outcome. We thus have $k \times 1$ vectors $\mathbf{X_1}, ..., \mathbf{X_{J+1}}$, one for each of the units in the donor pool and the treated unit. The vector for the treated unit is denoted as $\mathbf{X_1}$ and the vector of the donors $j = 2, 3, ..., J+1$ is stacked together and denoted as $\mathbf{X_0} = [\mathbf{X_2}\ \mathbf{X_3}\ \ldots\ \mathbf{X_{J+1}}]$.

The fundamental issue that the SCM, tries to solve is the unobserved counterfactual of the treated unit after treatment, the untreated $Y_{1t}$ for $t = T_0+1, T_0+2, ..., T$. The SCM posits that a valid counterfactual can be constructed using a weighted average of the donor pool, $j = 2, ..., J+1$. The weights are summarized in the vector $\mathbf{W} = (w_2, w_3, ..., w_{J+1})$, giving the synthetic control estimate:

$$\hat{Y}_{1t}^N = \sum_{j=2}^{J+1} w_j Y_{j_t} \tag{4}$$

where $\hat{Y}_{1t}^N$ is the counterfactual estimate for the outcome variable of interest in time $t$, and is given by a weighted average of the donor pool. For this thesis, the weights are restricted to be non-negative and sum to one, in order to avoid extrapolation. However, this restriction is not always necessary as proposed by for instance Doudchenko and Imbens (2016).

The synthetic control is then used to estimate the treatment effect for the unit of interest, which is given by:

$$\hat{\tau}_{1t} = Y_{1t}^I - \hat{Y}_{1t}^N \tag{5}$$

where $t > T_0$. The estimate reflects the difference between the observed outcome for the unit of interest, denoted as $Y_{1t}^I$, and the unit of interest in absence of treatment, $\hat{Y}_{1t}^N$.

The next step is to determine what weights to assign to the units in the donor pool. Theoretically, this could be done by the researcher subjectively assigning weights based on previous knowledge. However, other methods are often preferable. Abadie and Gardeazabal (2003) and Echevarría and García-Enríquez (2019) propose to select weights that minimize the difference between the synthetic control and treatment unit in terms of pre-treatment values of predictors of the outcome variable. This is formally described as minimizing the following equation:

[2] This conceptual framework is largely based on Abadie (2021), for additional methodological details we refer to this paper.

$$\|\mathbf{X_1} - \mathbf{X_0}\mathbf{W}\| = \left( \sum_{h=1}^{k} v_h (X_{h1} - w_2 X_{h2} - ... - w_{J+1} X_{hJ+1}) \right)^{1/2} \quad (6)$$

where $v_1, ..., v_k$ are positive weights for the selected predictors. The choice of these $v$ weights are in turn determined by the minimization of the pre-treatment mean squared prediction error (MSPE):

$$Pre - MSPE = \frac{\sum_{t=1}^{T_0} \left( Y_{1t} - \sum_{j=2}^{J+1} Y_{Jt} w_j(V) \right)^2}{T_0} \quad (7)$$

The synthetic control weights in vector **W** are determined by a nested, double minimization problem. The **V** and **W** vectors are jointly calculated, in order to obtain the optimal solution $\mathbf{w}^*(\mathbf{V})$, which minimizes the pre-treatment MSPE of the SC. The inclusion of other predictors than the outcome lags ensures that the estimated synthetic control is in fact a proper synthetic control group in the economic sense, as opposed to being the result of a spurious relationship from having matched a variable on a trend line. Abadie (2021) notes that the fewer pre-treatment periods that are available for the analysis, the more important the set of included predictors becomes in order to obtain a valid SC estimate. As Kaul et al. (2015) describe, the choice of the predictor vector is important. Specifically, the inclusion of pre-treatment values for the outcome variable needs to be considered with care. Specifically, including all pre-treatment outcome lags will remove any explanatory value of the economic predictors, and should therefore be avoided.

## 6.2 Inference

According to Abadie et al. (2015) there are several difficulties with statistical inference within the SCM. For instance, since the data sample is small and there is no randomization in the assignment of treatment or of the control units, it is difficult to draw conclusions regarding the robustness of the obtained results. However, the authors argue that there are falsification exercises, or placebo tests, that can be used to draw conclusions whether SC results are statistically significant.

Firstly, the in-time placebo, or backdating, test aims to see whether a treatment effect can be attributed to the policy in question. It was first implemented by Abadie et al. (2015) in their study on the economic effects of the German reunification in 1990. In their study, they estimate a SC for Germany with treatment set to 1975 before reunification, to make sure no treatment effect was observed before the studied intervention. No such effect was observed, which indicates that the model was in fact picking up a causal treatment effect from the reunification. If the SC had picked up a treatment effect in 1975, the authors would not have been able to draw the conclusion that the observed treatment effect in 1990 was caused by the reunification. Rather, it would have indicated that the SCM was unable to produce a good counterfactual of Germany. Using this logic, the in-time placebo test can also work as a way to track any potential anticipation effects of an intervention.

Secondly, the in-space, or unit, placebo test, first developed by Abadie et al. (2010), assigns treatment to, and estimates a synthetic control for, all donors in the donor pool. The estimated placebo treatment effects are then pooled together for comparison. This way, one observes whether the treatment effect for the treatment unit is extreme in magnitude compared to the untreated units. If this is the case, it indicates that the obtained effect is significant. If treatment effects for untreated units are similar or larger in magnitude to the treatment effect of the treated unit, any evidence of a causal effect from the studied intervention is undermined. This causes the obtained results to become less significant.

Results from this test can be used to obtain a distribution of effects, that can be used to compute p-values. This procedure is formally described by Abadie (2021). The first step is to compute the RMSPE, the root mean-squared prediction error (as opposed to the MSPE described above), of the synthetic control for donors and the treated unit, $j = \{1, ..., J+1\}$, from period $t_1$ to period $t_2$:

$$R_j(t_1, t_2) = \left( \frac{1}{t_2 - t_1} \sum_{t=t_1}^{t_2} ((Y_{jt} - \hat{Y}_{jt}^N)^2)^{1/2} \right) \tag{8}$$

This is then used to compute the ratio for unit $j$ between the post-intervention RMSPE, period $T_0 + 1$ to $T$, and the pre-intervention RMSPE, period $T_0 + 1$ to $T$:

$$r_j = \frac{R_j(T_0 + 1, T)}{R_j(1, T_0)} \tag{9}$$

The ratio measures the goodness-of-fit for the SC in the post-treatment period, relative to the pre-treatment period. In the case of a large observed treatment effect, the ratio will be large. The ratios for the treated unit and the donors are then used to compute a p-value:

$$p = \frac{1}{J+1} \sum_{j=1}^{J+1} I_+(r_j - r_1) \tag{10}$$

For this computation, $I_+$ is an indicator which takes on one for nonnegative values and zero otherwise. In other words, the indicator will be one for all donors that have an RMSPE ratio larger than the treated unit. The p-value can be seen as the fraction of treatment effects greater than or equal to that of the treated unit in the distribution, and is therefore largely dependent on the number of donors.

Finally, the leave-one-out placebo test, also developed by Abadie et al. (2015) is an additional robustness check for the SCM. The test aims to shed light on whether any obtained results can be attributed to a particular donor. It can also provide a range for the estimated treatment effect. The test iteratively estimates a SC, each time excluding one donor from the pool which had a positive weight in the initial SC. If results are insensitive to the exclusion of particular donors, it indicates that the results are robust. If results change greatly when excluding a particular donor, it suggests that they are driven by the excluded donor, and calls into question whether they capture a true causal effect. Estimated effects when leaving one donor out also gives a range of effects, that in the case of a large donor pool can be compared to a confidence interval.

## 6.3 Contextual Requirements

There are several contextual requirements needed for the SC to be an appropriate evaluation tool, as presented by Abadie (2021). Firstly, the size of the effect must be large enough in magnitude compared to other shocks to the outcome variable, in order for the SC to be able to pick up an effect. If treatment effects are small in comparison to idiosyncratic shocks to the outcome variable, they will not be picked up by the SCM.

Secondly, a comparison group must be available for the SCM to be able to estimate a treatment effect. This means that none of the donor pool units can adopt similar interventions as the unit of interest during the studied period. If that is the case, it is not a valid control unit, and must be excluded from the donor pool. Moreover, if any potential donors have experienced large shocks to the studied outcome during the time period of interest, and it is believed that such shocks would not have affected the unit of interest in the absence of treatment, these should also be excluded from the final donor pool. This since the synthetic control might contain effects from these shocks, and will thus not be a valid counterfactual. In addition, all donors must be relatively similar to the treated unit, in order to avoid interpolation bias. Interpolation bias can stem from the SC creating a counterfactual of donors that are extreme in the outcome variable, but whose discrepancies compared to the unit of interest are averaged away in the SC estimate. The counterfactual will thus not be valid.

Thirdly, as is similar to any research design aiming to study effects from an intervention, there can be no anticipation of the studied policy. If actors anticipate the policy under study, and adjust their behavior accordingly, any obtained treatment effects will be biased. Time placebo tests can serve to check this requirement formally.

An additional requirement for the SC is the stable unit treatment value assumption (SUTVA), elaborated by Rubin (1974). It requires the policy to have no spillover effects on the control group. If this assumption is violated, it is not possible to establish causal treatment effects from the policy, due to the counterfactual becoming invalid.

Lastly, a vital requirement for the SCM is the convex hull condition. This means that a convex combination of the donors should be able to approximate the predictors of the unit of interest for the pre-treatment period. In practice, this means that if the treated unit is extreme in terms of the outcome variable, or other predictors, the method would not be able to create a synthetic treatment unit. The fulfillment of these contextual requirements within this thesis are described in the following section.

## 6.4 Data and Model

In order to test our hypotheses, we construct five different synthetic controls, using as outcome variables low-skilled labor in the industry and services sectors, low-skilled youth employment, the senior share of the labor force as well as the prevalence of temporary contracts. The obtained synthetic controls are then compared to the actual French outcome, in an attempt to measure employment effects for these groups.

To proceed with our analysis, we use panel data from the International Labor Organization's (ILO) database ILOSTAT as well as from the OECD.Stat database. ILOSTAT compiles labor statistics from several sources for a range of countries, including labor force surveys, establishment surveys and administrative records. For our SCM, statistics from ILOSTAT's Labor force statistics database is used. The database includes statistics on indicators such as employment, unemployment, working time, and monthly earnings. The indicators are estimated by ILOSTAT based on survey representative samples, such as national household surveys. OECD.Stat contains data for OECD countries sorted by selected themes, such as labor, growth and other economic indicators.

We collect data on each of our respective outcome variables and relevant economic predictors to create our $k \times 1$ vectors. This is done for the longest time period possible around the policy announcement for France, and for a donor pool which is described in more detail below. Selecting the predictors of a SC model is an important step in the analysis. Kaul et al. (2015) describe why one should avoid choosing all or too many of the lagged outcomes as predictors, since it removes the predictive power of the economic predictors. However, as stated by Ferman et al. (2020) there is no established way to do this in the literature, and much responsibility is left to the author. We therefore start by creating a simple SC for each outcome variable using only lagged outcome variables as predictors. For this part, we only studied the time period before treatment, in order to avoid specification searching. We limited the number of lagged outcomes to maximum half of the total pre-treatment periods available, taking Kaul et al. (2015) into account, and tested different combinations of outcome lags in order to obtain the smallest pre-treatment RMSPE. We then added economic predictors in two steps, and evaluated whether they improved the model's fit. First we included averaged labor market and macroeconomic indicators. Second, indexed metrics, like for example employment protection, were added to the models. When the model is optimized according to Equation 6, only the predictors that improve the model's fit are included in the final specification. None of the models were improved by this second set of predictors. Thus, only the first type are included in the final models.

Since the countries of our donor pool differ in population size, using level measurements provided by ILOSTAT would cause France to potentially fall outside of the convex hull of our synthetic control. We therefore transform all outcome variables by dividing them with the total working age population for each respective country each year. The outcome variable is divided by working age population, as opposed to for instance the labor force, since we want to capture changes on the extensive margin, and not only changes within the composition of the labor force. This way, employment effects can be attributed to structural changes in the labor force, and not simply indicate transitions in and out of employment. A detailed description of the coding, calculation and considerations for each outcome variable and their predictors can be found in Appendix A.

### 6.4.1 Donor pool and time period

The donor pool has been chosen based on data availability, as well as Abadie's (2021) second contextual requirement regarding the availability of a comparison group. In order to avoid interpolation bias, we want the labor markets of our donor pool to be adequately similar to the French labor market. We therefore restrict our initial donor pool search to European OECD countries. Restricting the donor pool to Europe ensures that geographical, historical and cultural factors remain closely related in all countries. Furthermore, restricting the search to OECD members means that the economies and labor markets of the prospective donors will be relatively similar. Data on the outcome variables and relevant predictors for a sufficient number of pre-treatment periods were available for 12 European OECD countries. ILOSTAT's data source for these countries is the European Labor Force survey, which ensures measurement is somewhat harmonized for our variables.

In this initial selection, three donors experienced policy changes similar to the RTT during the time period of interest, namely Germany (Hunt, 1999), Belgium (SPF, n.d.) and Portugal (Varejao, 2005). This goes against the comparison group requirement of the SCM, and the three countries were consequently dropped from the donor pool. Additionally, Denmark introduced a policy reducing working hours in 1987 (Søndergård, 2007). As we will see, this is during the studied period for two outcome variables - the senior labor force and the prevalence of temporary contracts, and Denmark is thus removed from the donor pool for these variables. We expect potential effects from the Danish WTR to occur within the first five years following the policy, and therefore include it in the three remaining outcome variables, where the first included time period is in 1992. The final donor pool for the synthetic control method thus consists of the countries in Table 2.

In order to avoid anticipation effects, we set treatment in the SCM to the year of the announcement of the policy, $T_0 = 1998$, as opposed to the first year of enforcement, 2000. Due to data availability, the first year of our analysis varies for our five outcome variables, starting in 1986, 1987 or 1992. We believe that potential policy effects will primarily occur within the first number of years after implementation.

We do not want any donors to have experienced large idiosyncratic shocks to their labor markets during the studied period (Abadie, 2021) as this risks misrepresenting the development of the counterfactual unit that we are trying to create. Our final post-treatment period is 2007 for all outcome variables. This avoids including the Great Recession in 2008, which had severe employment effects and would distort our results if we included periods after 2007. The time period available for each outcome variable can be found in Table 2. There are other events during the time period that could potentially have caused idiosyncratic shocks to our outcome variables. To ensure the validity of the results all outcome variables are therefore visually inspected before proceeding with the SC analysis.

**Table 2:** Donor pool and time period for all outcome variables

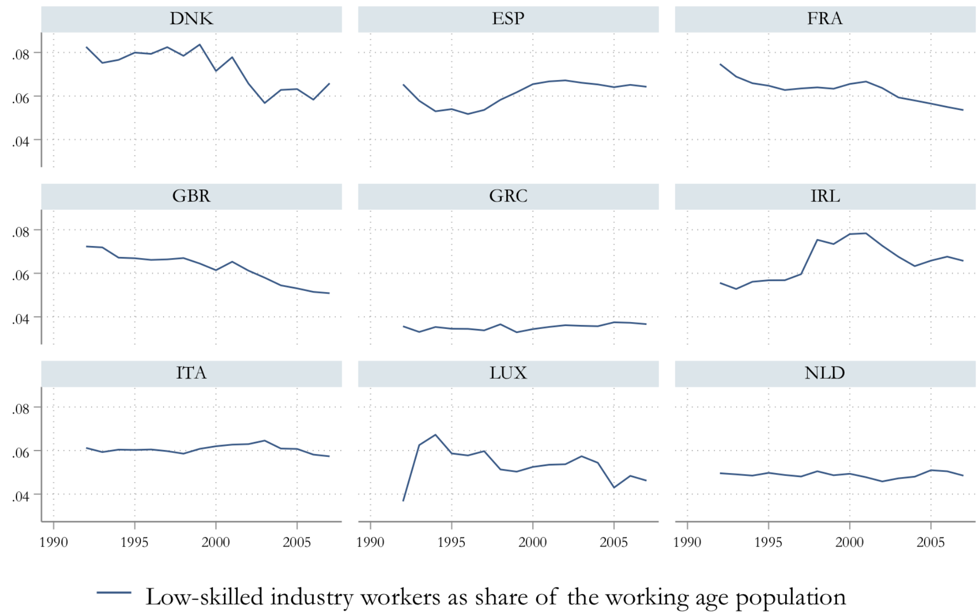
| Industry | Services | Youth | Senior | Temporary |
|---|---|---|---|---|
| 1992-2007 | 1992-2007 | 1992-2007 | 1986-2007 | 1987-2007 |
| Denmark | Denmark | Denmark | Great Britain | Great Britain |
| Great Britain | Great Britain | Great Britain | Greece | Greece |
| Greece | Greece | Greece | Ireland | Ireland |
| Ireland | Ireland | Ireland | Italy | Italy |
| Italy | Italy | Italy | Luxembourg | Luxembourg |
| Luxembourg | Luxembourg | Luxembourg | Spain | Spain |
| Spain | Spain | Spain | | |

### 6.4.2 Low-skilled workers in the industry and services sectors

The two outcome variables for the industry and services sectors are measured in employees by economic activity and occupation divided by the working age population. The two variables include low-skilled occupations within the ISCO classification (International Standard Classification of Occupations), separated by economic sector: industry or services. For more details on the chosen ISCO categories and economic sectors, as well as economic predictors for these two variables, see Appendix A. Before computing SCs of these outcome variables, we visually inspect their time series for both France and the donor pool. If we observe a shock or trend break in the outcome variable for the treated unit at the time of the policy, moving on with the SCM is motivated. Even though it is not possible to attribute an observed shock to the policy itself upon visual inspection, it gives motivation for further investigation. We also want to inspect the outcome variable for the donor pool, in order to make sure that the convex hull condition is fulfilled.

Figure 3 plots the outcome variable for low-skilled industry workers in France and the donor countries. We observe a slight positive trend in the variable for France between years 1998-2002. However, the change only corresponds to roughly half a percentage point. For the remaining time period, the trend is negative or constant. We note that France is contained within the convex hull of the donor pool. The donor pool exhibits quite heterogeneous characteristics. The time series for Denmark, Luxembourg and Ireland are rather volatile. The Danish volatility suggests industry employment is susceptible to shocks, which perhaps can be attributed to the Danish flexicurity system, where weak employment protection is combined with a strong social security system (Mailand, 2005). Luxembourg exhibits a sharp two percentage point upward shift in the variable between 1992 and 1993, which is attributable to missing data for some included ISCO categories in 1992. We still choose to include Luxembourg in the donor pool, since the rest of the time series does not contain the same volatility and measurement error and we want to keep the donor pool as large as possible. Ireland's episode of higher levels of industry employment in the late 1990s and early 2000s coincides with the country's period of economic expansion, where for instance lowered corporate taxes and favorable demographic conditions led to an Irish economic boom (**the˙economist˙model˙2004**). The overall downward trend observed for the countries Denmark, France, Great Britain and Luxembourg reflects the decline in number of industry workers, especially in manufacturing, that has taken place across OECD countries in the last half century (Pilat et al., 2006).

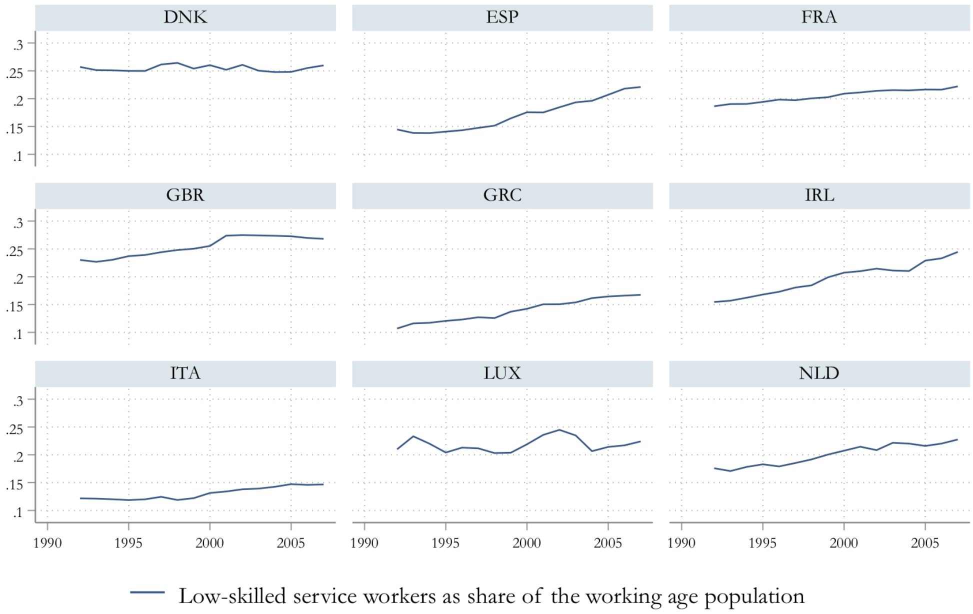
**Figure 3:** Annual time series of low-skilled industry workers




Source: ILOSTAT Labor Force statistics database.

While the industry variable exhibited a downward trend, the same is not reflected in the services variable. We see in Figure 4 that the share of service workers in the working age population has a positive trend in most donor countries. France does not exhibit any large shocks to the outcome variable during the time period of study, which would have given a stronger motivation for studying the policy with an SCM. Rather, France experiences a small positive trend in the variable during the studied period. France is not extreme among the donors, which indicates that the convex hull condition is fulfilled. Furthermore, we note that Great Britain experiences a slight upward shift in the variable in year 2001. The overall positive trend for this variable in Spain, France, Greece, Ireland and the Netherlands is in line with the employment growth in the services sector during the last half century (Schettkat & Yocarini, 2006).

**Figure 4:** Annual time series of low-skilled service workers

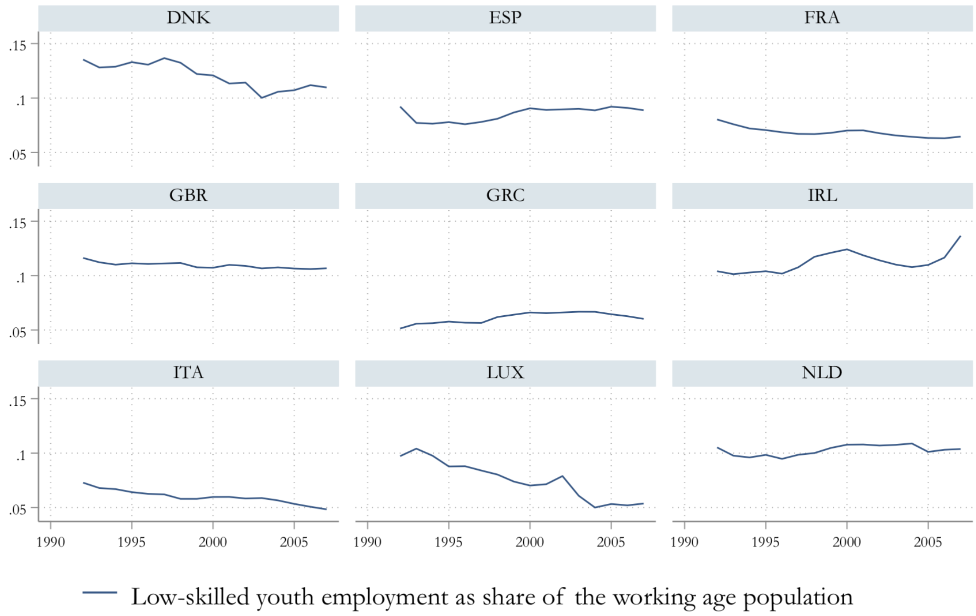



Source: ILOSTAT Labor Force statistics database.

### 6.4.3 Low-skilled youth employment

The outcome variable for young workers is measured by low-skilled youth employment divided by the working age population. The variable uses the same ISCO categories to measure skill as the previous variables. Unfortunately, this measurement is of employment, and not employees, which means that it includes the self-employed. This must be taken into account when interpreting the results. For this variable, we use the age band 15-29 years. More details regarding this variable transformation and its predictors can be found in Appendix A. Figure 5 plots this variable for France and the donors. We note that the French low-skilled youth employment has a quite stable development during the time period, starting with a slightly negative trend. We note that the convex hull condition is fulfilled for this variable as well. The development for the donors is quite stable, with slight upward and downward trends during the period of study. Denmark exhibits the same volatility as for industry workers in Figure 3.

**Figure 5:** Annual time series of low-skilled youth employment




Source: ILOSTAT Labor Force statistics database.

### 6.4.4 Senior labor force

This outcome variable is measured as the labor force aged 55-64 divided by the working age population. Ideally, we would want to study the policy's effect on the extensive margin by looking at the number of individuals that retire each year. However, we have not encountered a good data source for this measurement, but believe the labor force aged 55-64 serves as a good proxy. If workers in this age band exit the labor force, we assume it is because they are retiring. More detail regarding the construction and predictors of this variable can be found in Appendix A. Studying the development of this variable for France in Figure 6, we observe a downward trend until around the time policy announcement, after which we see a sharp increase. Moreover, all donors seem to experience a positive trend in this variable starting in the late 1990s. We note a common positive trend among almost all countries after the treatment year. This is especially the case for the Netherlands, whose upward trend is the strongest of donors. Greece experiences a slight downward shift in the mid 1990s.

**Figure 6:** Annual time series of senior labor force

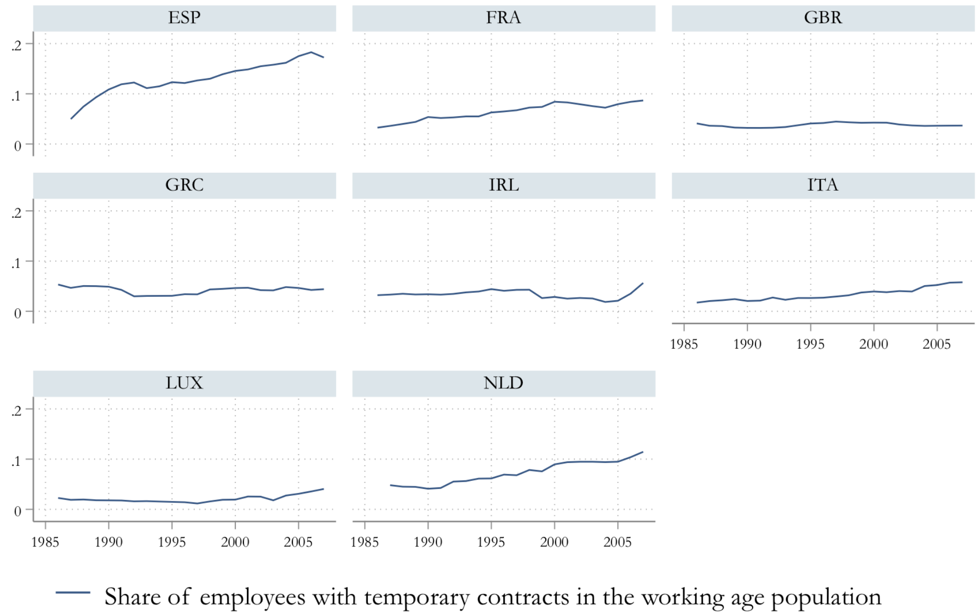

Source: ILOSTAT Labor Force statistics database.

#### 6.4.5 Prevalence of temporary contracts

This variable is measured as the share of employees with temporary contracts in the working age population. Appendix A explains the details of this variable transformation, as well as our collected predictors. The time series for this variable for France and the donors is seen in Figure 7. Upon visually inspecting the series for France, we see a positive trend until the year of policy enforcement in 2000, where we see a slight decline. In 2004, the trend once again becomes positive. Thus, we observe a negative change in the variable around policy enforcement. The donors exhibit constant or upwards trends, without any major volatility. We note that there is variation in the use of temporary contracts in the donor labor markets. Spain exhibits an increase in, and a high level of, temporary contracts, compared to Luxembourg or Italy.

We conclude that we are unable to spot any causal effects from the WTR policy by simply inspecting the time series of our constructed outcome variables. Changes to the time series of these around the time of the policy could just as well stem from other confounding factors. We therefore proceed with constructing synthetic counterfactuals of what would have happened in France in absence of the WTR policy using the SCM. This way, we will be better equipped to draw possible conclusions regarding causal effects of the policy.

**Figure 7:** Annual time series of temporary contracts




Source: ILOSTAT Labor Force statistics database.

# 7 Synthetic Control Results

In this section we present the results from the SCM. The selected weights for donors and predictors, the **W** and **V** vectors, are presented in Table 3 and Table 4 respectively. In order to present aggregated tables, we use a point (.) in the tables to show that a given country or predictor was not included in a final model, following our model selection explained in the previous section. As previously mentioned the models have been obtained by selecting a set of predictors, for which the pre-treatment MSPE has been minimized. A balance table with the comparison of actual and synthetic predictor outcomes is found in Table 5. As in the previous tables, a point for a given predictor shows that it is not included in the model. In this section, when referring to the fit of the model, we refer to the size of the pre-treatment MSPE.

**Table 3:** Country weights (W)

| | Industry | Services | Youth | Old | Temporary |
|---|---|---|---|---|---|
| DNK | 0.094 | . | 0.002 | . | . |
| ESP | 0.278 | 0.196 | 0.278 | 0.162 | 0.238 |
| GBR | 0.628 | 0.509 | 0.004 | . | 0.003 |
| GRC | . | 0.185 | . | . | . |
| IRL | . | . | 0.005 | . | . |
| ITA | . | . | 0.644 | 0.517 | 0.405 |
| LUX | . | 0.110 | 0.057 | 0.114 | . |
| NLD | . | . | 0.009 | 0.207 | 0.354 |

Note: A null weight for a predictor is represented by a point (.)

**Table 4:** Predictor weights (V)

| | Industry | Services | Youth | Senior | Temporary |
|---|---|---|---|---|---|
| Outcome lag (1987) | . | . | . | . | 0.022 |
| Outcome lag (1989) | . | . | . | . | 0.061 |
| Outcome lag (1991) | . | . | . | . | 0.168 |
| Outcome lag (1992) | 0.255 | 0.439 | 0.067 | . | . |
| Outcome lag (1993) | 0.195 | . | . | . | 0.277 |
| Outcome lag (1994) | . | 0.000 | 0.288 | . | . |
| Outcome lag (1995) | . | . | . | . | 0.258 |
| Outcome lag (1996) | . | 0.424 | 0.645 | . | . |
| Outcome lag (1997) | 0.429 | . | . | . | 0.213 |
| Outcome lag (1986-1988) | . | . | . | 0.317 | . |
| Outcome lag (1989-1991) | . | . | . | 0.157 | . |
| Outcome lag (1992-1994) | . | . | . | 0.381 | . |
| Outcome lag (1995-1997) | . | . | . | 0.140 | . |
| GDP growth | 0.039 | 0.009 | . | . | . |
| GDP/capita | 0.025 | 0.020 | . | 0.005 | . |
| Investment(1997) | 0.020 | 0.004 | . | . | . |
| Low skilled share of LF | 0.030 | 0.003 | . | . | . |
| Total LF participation rate | 0.003 | 0.092 | . | . | . |
| Unemployment rate | 0.004 | 0.009 | . | . | . |

Sources: Employment data are collected from the EU Labor Force Survey through ILOSTAT. GDP data are from OECD.
Notes: Outcome lag refers to the pre-treatment values for each repsective outcome variable of interest. When a year range is specified the predictor is averaged over these years. If not otherwise specified, all predictors are averaged over the whole pre-treatment period.
A null weight for a predictor is represented by a point (.).

**Table 5:** Balance table for all outcome variables

| | Industry | | Services | | Youth | | Senior | | Temporary | |
|---|---|---|---|---|---|---|---|---|---|---|
| Predictors | Actual | Synthetic | Actual | Synthetic | Actual | Synthetic | Actual | Synthetic | Actual | Synthetic |
| Outcome lag (1987) | . | . | . | . | . | . | . | . | 0.036 | 0.037 |
| Outcome lag (1989) | . | . | . | . | . | . | . | . | 0.044 | 0.048 |
| Outcome lag (1991) | . | . | . | . | . | . | . | . | 0.052 | 0.052 |
| Outcome lag (1992) | 0.075 | 0.071 | 0.186 | 0.188 | 0.080 | 0.080 | . | . | . | . |
| Outcome lag (1993) | 0.069 | 0.068 | . | . | . | . | . | . | 0.055 | 0.056 |
| Outcome lag (1994) | . | . | 0.191 | 0.190 | 0.072 | 0.072 | . | . | . | . |
| Outcome lag (1995) | . | . | . | . | . | . | . | . | 0.063 | 0.062 |
| Outcome lag (1996) | . | . | 0.198 | 0.196 | 0.069 | 0.069 | . | . | . | . |
| Outcome lag (1997) | 0.064 | 0.064 | . | . | . | . | . | . | 0.067 | 0.066 |
| Outcome lag (1986-1988) | . | . | . | . | . | . | 0.058 | 0.057 | . | . |
| Outcome lag (1989-1991) | . | . | . | . | . | . | 0.054 | 0.055 | . | . |
| Outcome lag (1992-1994) | . | . | . | . | . | . | 0.050 | 0.051 | . | . |
| Outcome lag (1995-1997) | . | . | . | . | . | . | 0.048 | 0.049 | . | . |
| GDP growth | 1.531 | 2.491 | 1.531 | 2.448 | . | . | . | . | . | . |
| GDP/capita | 29779.025 | 27648.121 | 29779.025 | 29356.910 | . | . | 28568.797 | 32097.354 | . | . |
| Investment(1997) | 0.789 | 1.515 | 0.789 | 2.042 | . | . | . | . | . | . |
| Low skilled share of LF | 0.500 | 0.491 | 0.500 | 0.479 | . | . | . | . | . | . |
| Total LF participation rate | 67.637 | 71.554 | 67.637 | 67.956 | . | . | . | . | . | . |
| Unemployment rate | 11.873 | 12.465 | 11.873 | 10.868 | . | . | . | . | . | . |

[a] Employment data are collected from the EU Labor Force Survey through ILOSTAT. GDP data are from OECD.
[b] Outcome lag refers to the pre-treatment values for each repsective outcome variable of interest. When a year range is specified the predictor is averaged over these years. If not otherwise specified, all predictors are averaged over the whole pre-treatment period.
A null weight for a predictor is represented by a point (.).

## 7.1 Pre-Treatment Analysis

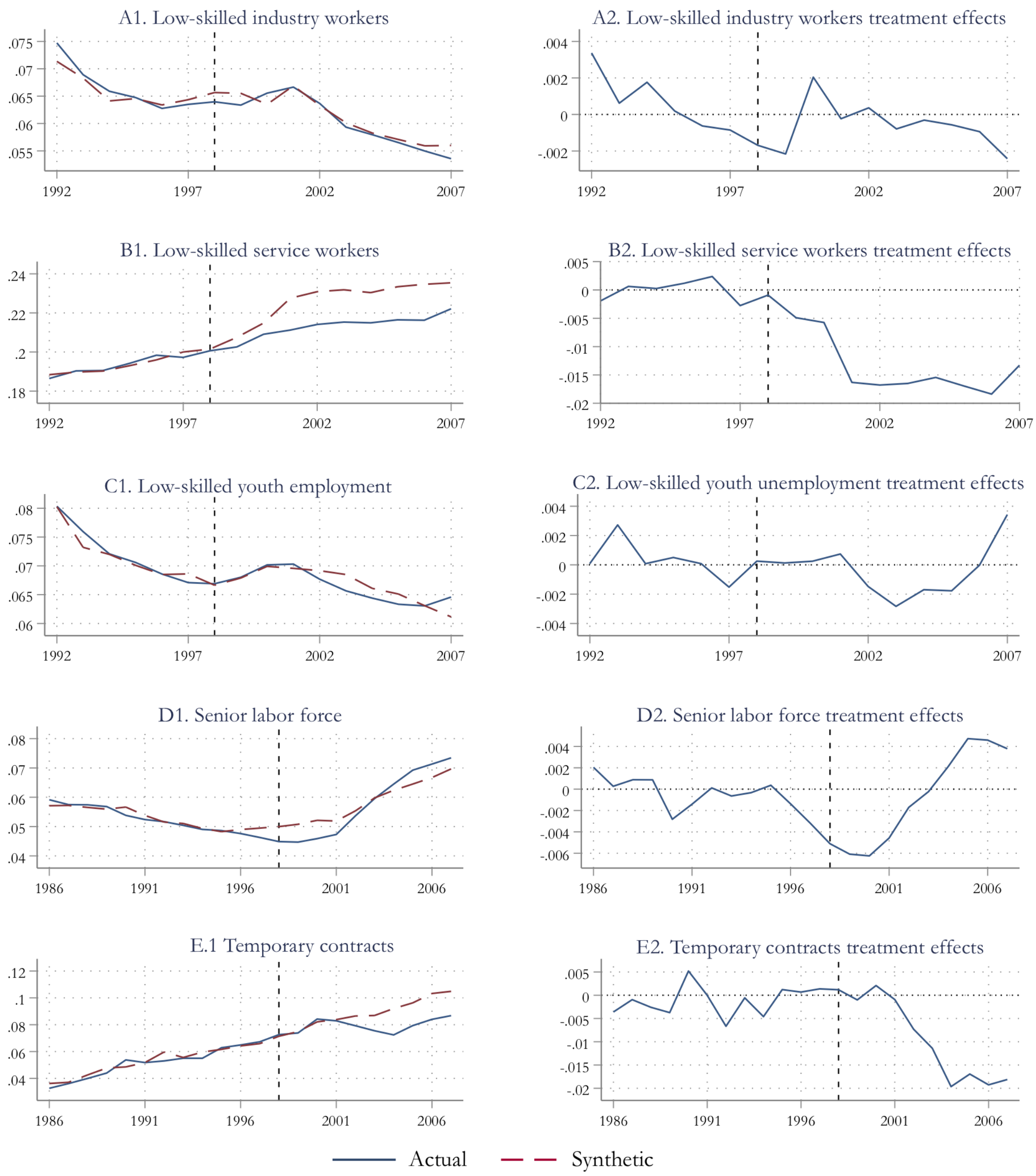
The main results are shown in Figure 8, where the left-hand panels display actual (solid blue line) and synthetic (dashed red line) France. The right-hand panels show treatment effects for the outcome variables, which is the difference between the synthetic and actual France. The donors and the treated unit sum to eight or nine, which gives the dimension $N = 8$ or $N = 9$ in Abadie's (2021) framework. All outcome variables are measured in shares of the working age population. The number of time periods differ across outcome variables between $15 < T < 22$, where the treated periods are the last nine, with treatment in 1998.

We start by studying the **W** vector, which is seen in Table 3. We note that each synthetic control is quite sparse with few donors. The number of donors ranges from three to four for almost all outcome variables. This is not the case for the youth employment variable which assigns weight to all donors in the pool apart from Greece. However, the donor weights are distributed unevenly with two major donors, and small weights on the remaining countries. Similarly, four out of five synthetic controls have the majority of their weight on one donor. For the SCs of low-skilled industry and service workers, Great Britain has 62.8 and 50.9 percent of the weight respectively. For the youth employment and the senior labor force, Italy accounts for 64.4 and 51.7 percent of the SC. As for the prevalence of temporary contracts, the weights are more balanced among the three major donors. However, Italy still has the largest weight with 40.5 percent. The fact that the synthetic controls have their weights concentrated to one donor increases the risk that the obtained results are driven by their majority donor. This will be tested further along in Section 7.3.1.

A more detailed overview of predictor weights can be found in Table 4, which presents the **V** vector. In general the pre-treatment fit has been optimized using primarily lagged outcome values, as opposed to economic predictors, for all specifications. There are multiple explanations for this. One is that the best predictors of future outcome values are historical data, and placing a high weight on lagged outcomes is a consequence of a causal relationship between the past and the future. As a consequence the other predictors are given less predictive power.

Table 5 shows the actual and synthetic outcomes for all predictors and lagged outcome variables. If synthetic France is able to closely replicate the pre-treatment outcomes and predictors of actual France, it indicates that the SC is a valid counterfactual for what would have happened in France in absence of the WTR policy. The relatively good replication of the pre-treatment outcome variables is shown in this table and in Figure 8, where we observe little difference between synthetic and real France in all pre-treatment values. Economic predictors are included in the models for industry, services and senior labor force outcome variables. We observe replication of mixed quality for these predictors. The replication of the predictors which relate more closely to the labor market, and consequently to our outcome variables, are relatively good. Although these are only included in the first two specifications, the predictors low-skilled share of the labor force, total labor force participation rate and total unemployment rate are similar for actual and synthetic France in both the industry and services models. In contrast, the predictors that are more generally related to the economy, GDP per capita, GDP growth and investment, are notably less well-replicated. This indicates that the outcome variables' relationship with more detailed labor market indicators is stronger, compared to more aggregate macroeconomic measures. An alternative explanation is that our donor pool and data fail to create a synthetic France of sufficient quality.

**Figure 8:** Synthetic and actual outcome with treatment effects



The close pre-treatment fit is also seen in panels A1-E1 in Figure 8. We start by noting that the scale of the figures is overall small. This is especially the case for panels A1 and C1, where each unit interval corresponds to a 0.5 percentage change. For panels B1 and E1 as well as D1, each interval corresponds to a two and one percentage point change respectively. As for the pre-treatment fit, we see in panels A1-E1, that actual France is fairly well replicated by the synthetic controls, even though the fit in panels A1, C1 and E1 is noisy. However, the noise becomes less important considering the small scale of the figures, and the pre-treatment fit can thus be considered to be tight.

## 7.2 Post-Treatment Analysis

The right side panels A2-E2 in Figure 8 present the treatment effects from the policy, ($\hat{\tau}_{1t} = Y_{1t}^{I} - \hat{Y}_{1t}^{N}$), obtained using the synthetic Frances as control groups. As in the case of the the pre-treatment analysis, we note that the magnitude of all treatment effects is very small, amounting at most to two percentage points. For low-skilled industry workers, in panel A2, we observe a slightly negative treatment effect starting already before the policy. However the magnitude of this effect is very small, of only 0.2 percentage points, and it is indistinguishable from noise. Given the close pre and post-treatment fit, this synthetic control indicates no treatment effect for this variable. In Panel B2, we observe the treatment effect for low-skilled service workers. Here, we see a negative treatment effect, starting at the time of the policy. The effect is magnified in 2000, the year when the policy was enforced, and corresponds to around 1.5 percentage points of the working age population, which roughly amounts to 550,000 service workers. The observed negative employment effect goes in the opposite direction of our hypothesis. However, as we will see in the leave-one-out placebo test, this result is largely driven by including Great Britain in the donor pool.

Panel C2 in Figure 8 shows the treatment effect for low-skilled youth employment. This SC shows a slightly negative treatment effect in the year of enforcement for small firms in 2002. However the magnitude is of around 0.3 percentage points, and is too small to differentiate from noise. In 2006, the synthetic and actual France develop in different directions. However, this is too far from the treatment date for us to interpret as a treatment effect from the policy. We thus find no treatment effect for this variable.

The senior labor force treatment effect is seen in Panel D2 in Figure 8. We note that a divergence between actual and synthetic France starts to appear before treatment in 1996. At its minimum, it amounts to a negative 0.6 percentage point change of the working age population, corresponding to around 200,000 seniors. In 2003, the difference becomes positive. Since the divergence starts to emerge in 1996, before treatment, it indicates that it is not a result of the RTT. We therefore conclude that the policy had no treatment effect for the senior labor force in France.

Finally, we observe the treatment effect for the prevalence of temporary contracts in panel E2 in Figure 8. A negative treatment effect starts to emerge in 2002, the year of policy enforcement for small firms. At its lowest point, the treatment effect corresponds to a two percentage points negative change, which roughly equals 740,000 workers. Before this divergence, synthetic France closely replicates actual France, even though the fit between 1990 and 1995 is noisy. As the result for low-skilled service workers, the observed treatment effect goes in the opposite direction of our hypothesis. However, the in-space placebo test will later show that this effect is not significant.

To summarize, we observe negative treatment effects on employment for outcome variables low-skilled service workers and share of the working age population with temporary contracts. In the following section with placebo tests of these results, we find that they are not robust to falsification exercises. For the other outcome variables, we do not observe any treatment effects.

## 7.3 Placebo Tests

This section will present results from the leave-one-out, in-space and and in-time placebo tests. These three tests are standard in the SC literature, and aim to test whether any obtained treatment effects are statistically significant and economically meaningful. Since we obtained treatment effects for low-skilled service workers and prevalence of temporary contracts, we are most interested in conducting placebo tests for these outcome variables. Placebo tests for the variables where no relevant treatment effect was found will mostly serve to see how strong the null result is.

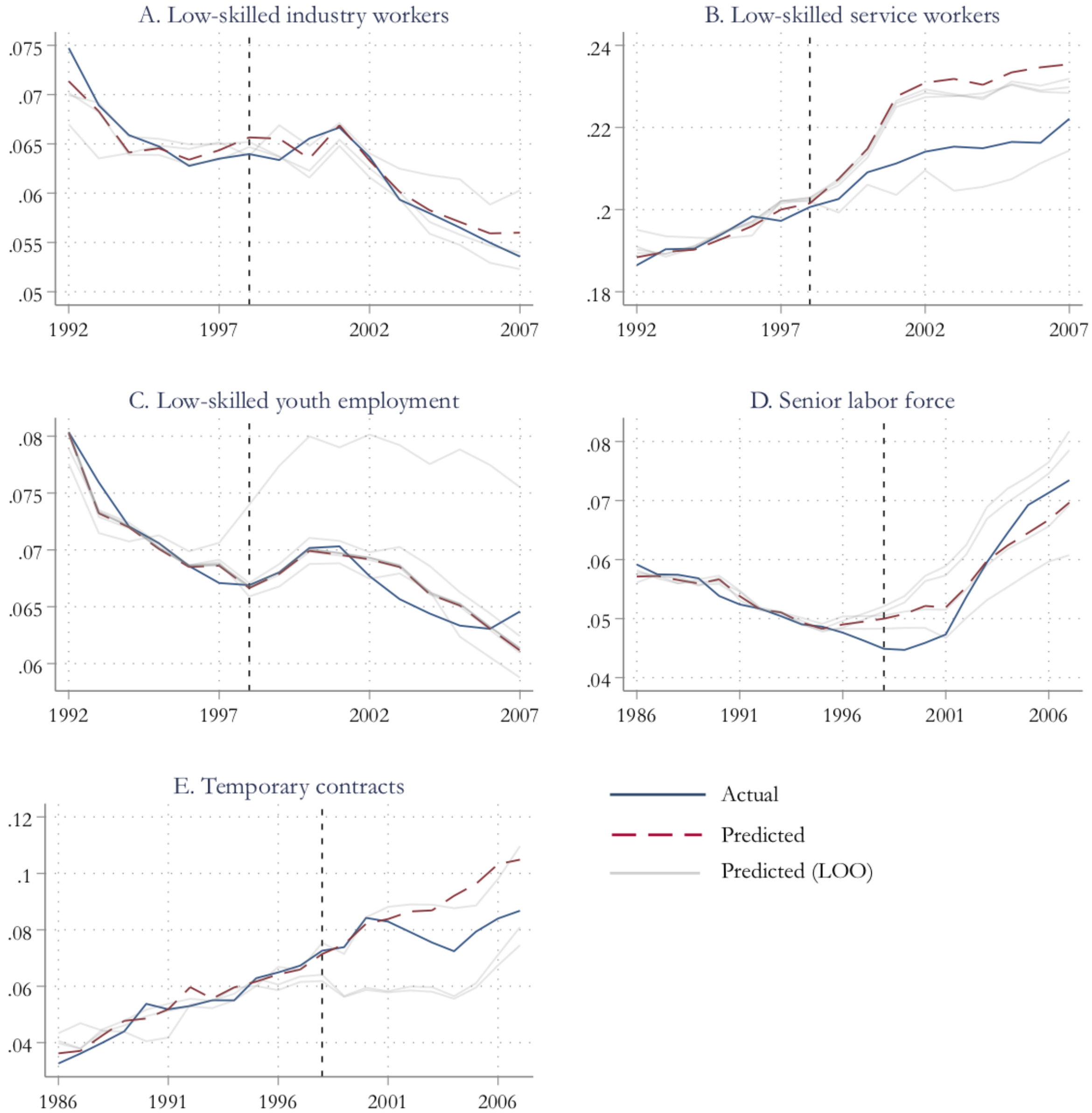
### 7.3.1 Leave-one-out placebo tests

Since our estimated synthetic controls assigned large weights to single donors, the leave-one-out test is of particular interest. If results do not significantly change when excluding the largest donor in the initial synthetic control, it indicates that the results are robust.

Figure 9 shows results of the leave-one-out placebo tests. The blue solid line represents actual France, the dashed red line synthetic France, and the light gray lines the SC for when each donor with a positive weight is excluded. We observe in Panel A that the low-skilled industry workers outcome is quite insensitive to excluding the obtained SC donors. Even though the pre-treatment fit becomes noisier, we still find no clear treatment effect when donors are excluded. This strengthens the finding that the policy had no employment effects for low-skilled industry workers.

In addition, for low-skilled service workers, we see in Panel B that the pre-treatment fit remains reasonably tight even when all donors but one are dropped from the pool. The treatment effect becomes positive when Great Britain is dropped from the donor pool, which received an initial SC weight of 50.9 percent. In contrast, when Great Britain is included in the donor pool, the treatment effect is always negative. We observe an indication of this finding in the visual inspection of this variable in Figure 4. There we see that the share of low-skilled service workers in the British working age population experienced an upward shift of around two percentage points in year 2000. When studying this further, we note from our data source that it is attributed to a break in the

**Figure 9:** Leave-one-out tests




time series of this variable. This effect is also seen in Figure 9 Panel B, where the synthetic control replicates this upward shift in 2000. The difference in treatment effect magnitude between the synthetic controls with and without Great Britain is around two percentage points. This means that Great Britain is the main driver of our negative treatment effect. It also shows that the observed treatment effect for low-skilled service workers probably does not stem from the WTR, but rather from using Great Britain as a donor country. Moreover, since the time series break in the data gives rise to an idiosyncratic shock that is large enough in magnitude to show up as a negative treatment effect, it is not a proper control unit for France. This in spite of it being given considerable weight in the synthetic control. Thus, with our data and method, we cannot observe any effects of the WTR policy for the outcome variable low-skilled service workers.

For low-skilled youth employment, we also observe that our obtained results are sensitive to the exclusion of a particular donor in Figure 9 Panel C. Results for this variable are driven by Italy, which was given 64.4 percent weight in the initial synthetic control. In the time series of this variable in Figure 5, we see that Italy, similarly to France, also experienced a downward trend for the variable during the studied period. As a result, when Italy is excluded from the SC, we observe a negative treatment effect. This is due to the SC giving a larger weight to Spain, which experienced a rise in low-skilled youth employment after the RTT, as seen in Figure 5. Using Italy as a large share of the counterfactual is thus the driver of the null result. We conclude that for the youth employment variable, there is a lack of appropriate donor countries, which makes the SC unable to produce robust results. Moreover, the SC of this outcome variable includes several small donor weights ranging from 0.2 to 5 percent. As expected, we see that excluding these does not affect the SC, illustrated by the gray lines closely following the dashed line of the initial SC. In other words, the null result for this variable is driven by Italy. This makes it more difficult to draw any conclusions on the effects of the policy for low-skilled youth employment.

In Figure 9 Panel D we see that results are not as sensitive to excluding a particular donor for the senior labor force as for the youth and services measurements. We note that between 1996 and 2001 the actual outcome is consistently lower than all of the counterfactuals, showing a negative treatment effect. However, since this effect started before the policy announcement, we are again unable to attribute the effect to the policy. For the years after 2001 the outcome for actual France lies inside the range of the leave-one-out counterfactual outcomes. We cannot explicitly interpret this observed range as a confidence interval, due to the small number of donors in the test, but it indicates that the employment effect after 2001 is not significant. We thus receive further indications of a null result for this variable.

Lastly, Panel E shows the same test for temporary contracts. The actual outcome lies inside the interval of the effects that exclude one donor, which indicates that the effect is not significant. We also see that the treatment effect remains negative when excluding Spain, but becomes positive when excluding Italy or the Netherlands. The treatment effect is thus once again sensitive to the choice of donors. This gives the same conclusion as for the senior labor outcome variable - we are unable to draw any conclusions of a significant causal effect from the policy.

#### 7.3.2 In-space placebo tests

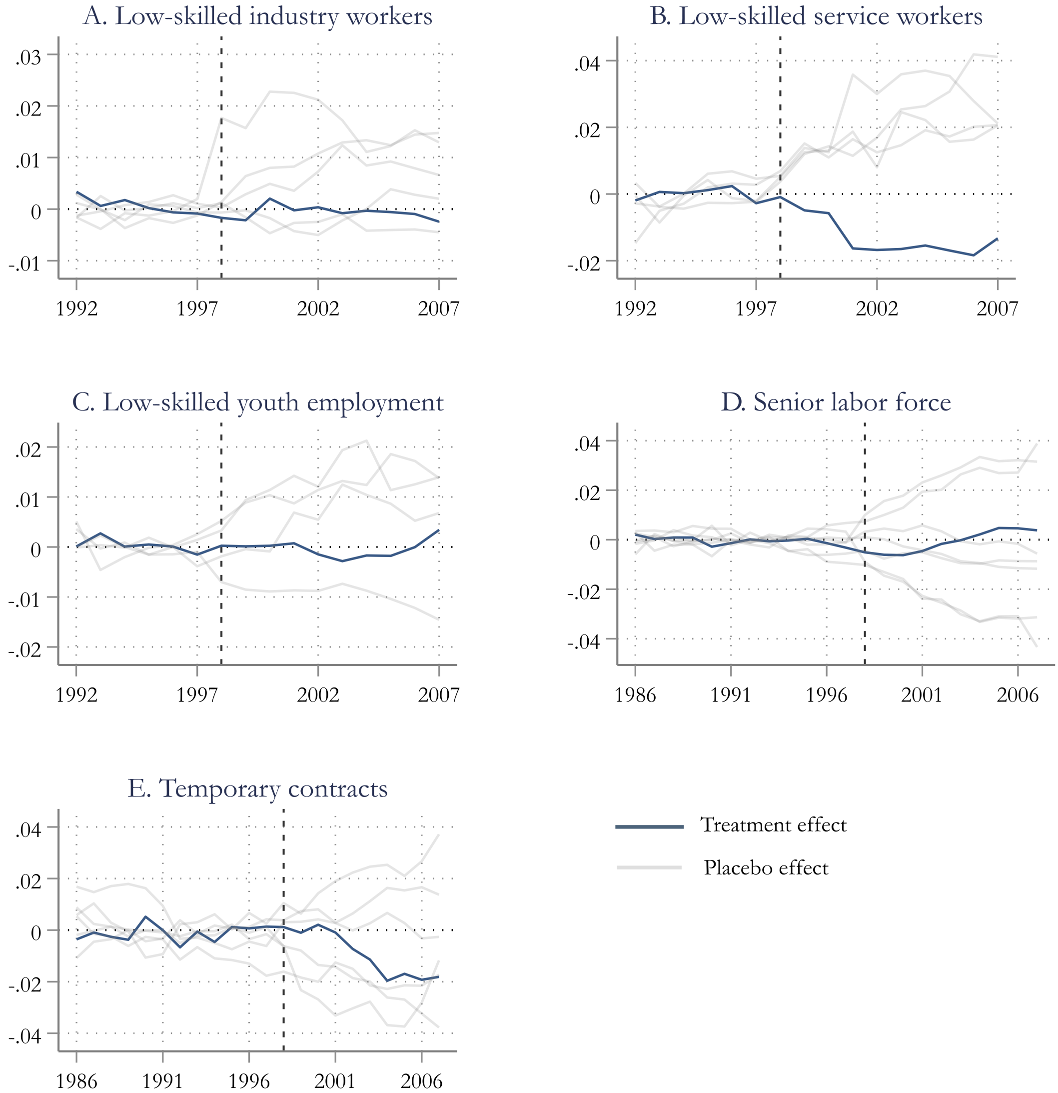
The in-space placebo test compares treatment effects when treatment is assigned to all donors in the pool. This helps to determine if the obtained effects are large in magnitude and thereby significant compared to those obtained for untreated donors. When conducting the in-space placebo tests, we remove any treatment effects from donors for which we do not obtain a good pre-treatment fit, in line with Abadie et al. (2010). We only want to compare donors with good pre-treatment fits, since treatment effects from donors with poor pre-treatment fits could just as well be the result of an invalid synthetic control. However, since our donor pool only consists of seven or eight countries depending on the outcome variable, this should be done with caution. We still want to be able to compare the treatment effect to a large enough distribution of placebo effects. We therefore remove any SCs of donors that have a pre-intervention MSPE 20 times larger than that of France. A method proposed in the SC literature to set this threshold is to follow Abadie et al. (2010) and only remove units whose pre-treatment MSPE is more than twice the size of the treated unit. However, due to the small magnitude of our estimates following this rule would have excluded a majority of our donor placebos. We therefore increased this upper limit to 20.

Results from the unit placebo test can be seen in Figure 10. The blue line shows the treatment effect for France, and the gray lines, shows treatment effects for donors. In Panel A, we see that the treatment effect for low-skilled industry workers is very small compared to the unit placebo treatment effects. Since we did not observe a treatment effect for this outcome, the result is not surprising, and further strengthens the conclusion that the policy did not affect the employment of low-skilled industry workers. As for low-skilled service workers, we see in Panel B that our observed treatment effect is the most extreme among the donors. The observed treatment effect is negative, compared to the treatment effects of other donors, which are all positive. This could have indicated that the observed effect was significant, but given that the leave-one-out-test showed that Great Britain was the main driver of the services result, this test loses its importance.

Furthermore, we did not observe any treatment effect for the youth employment variable. This finding is further strengthened by the in-space placebo test in Figure 10 Panel C. Here, we see that the unit placebo pre-treatment fits are a less noisy than for our industry and services measurements. The fact that we find no treatment effect is strengthened, since the treatment effect for France is the smallest among all donors. Looking at the senior labor force, we see in Figure 10 Panel D that the pre-treatment fit is tight for all unit placebos. We also note that our observed treatment effect for France is small in magnitude, compared to large effects observed among the donors, that also seem start before 1998. This is not surprising, as upon visually inspecting the data in Figure 6 we already noticed increases in this variable among the donors in the late 1990s. Even though we were already cautious of the observed change for the senior labor force, since it occurred before the policy, this test indicates that the treatment effect for senior labor is not significant.

The results indicated a treatment effect for the prevalence of temporary contracts among the working age population. However, we see in Figure 10 Panel E that the magnitude of this effect is not extreme in relation to the treatment effects for other donors. We can thus not regard the observed effect as significant. Moreover, p-values corresponding to these unit placebo tests are presented in Figure 13 in the Appendix B. Since most treatment effects for France are very small in magnitude compared to those of the donors, the p-values are high.

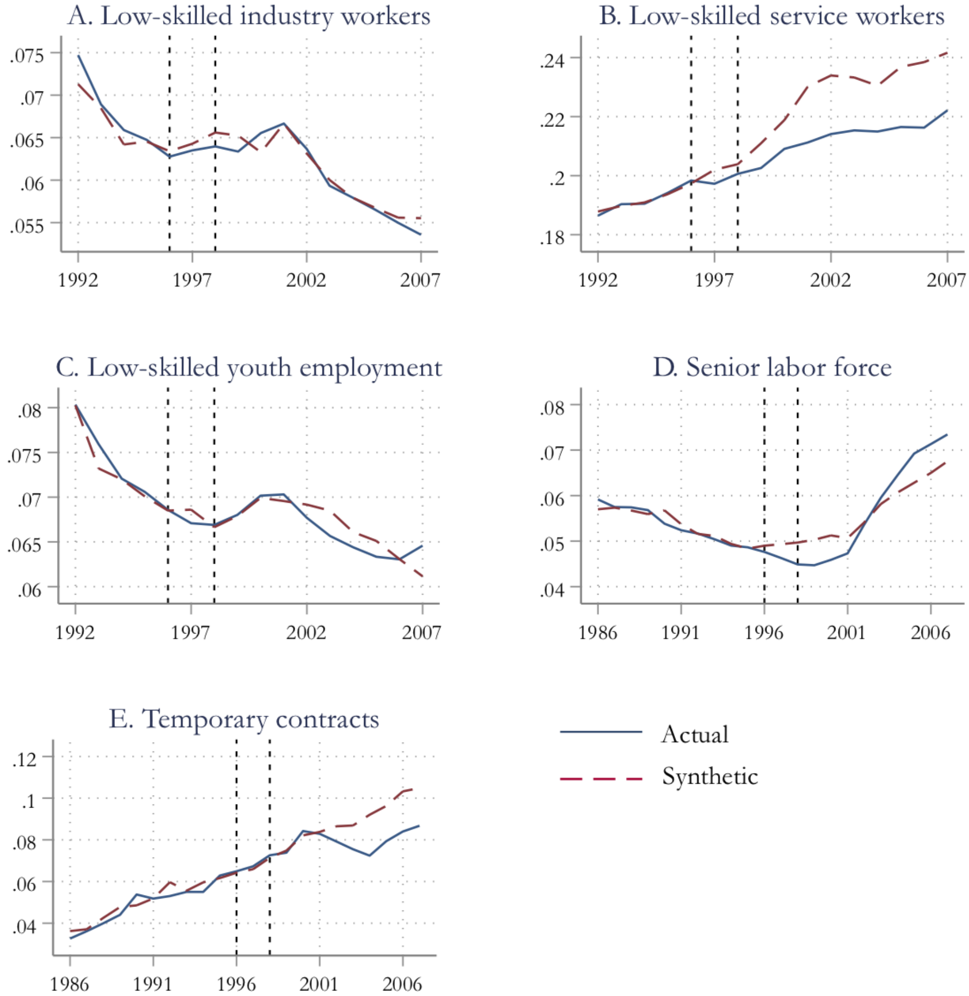
**Figure 10:** In-space placebo tests




### 7.3.3 In-time placebo tests

Figure 11 shows time-placebo tests for our outcomes, where treatment year is set to 1996, as opposed to 1998. In our case, this corresponds to the year of the announcement of the Robien Law. We see that our results are not sensitive to a change in the treatment date. Some differences in the treatment effects emerge between the placebo and actual results, such as in the treatment effects for the industry outcome in Panel A, where the placebo graph is shifted slightly upwards. However, these are very small in magnitude. These results indicate that there were no anticipation effects of the policy. In theory, it also indicates that we can attribute treatment effects to the RTT. However, since the in-space and leave-one-out placebo tests showed that the SC results were not significant, this is not relevant.

**Figure 11:** In-time placebo tests




In conclusion, the observed effects for temporary contracts and low-skilled service workers were not robust to placebo tests. The effect for low-skilled service workers was heavily reliant on the inclusion of Great Britain as a donor. In addition, the observed effect for temporary contracts was small in magnitude compared to that of other donors when conducting the unit placebo test. It was also largely driven by the inclusion of Italy as a donor. This indicates that the obtained causal effects are not a result of the RTT, but are rather attributed to the inclusion of specific donors in the SC. The placebo tests also further strengthen the finding that there were no employment effects from the policy on low-skilled industry workers, low-skilled youth employment and the senior labor force. We are thus unable to conclude any significant employment effects of the RTT on our studied labor market outcomes with the SCM.

# 8 Event Study Design

As a final supplementary analysis we take inspiration from an event study design to investigate the research question using French micro data.[3] This data can provide a more detailed analysis, as the heterogeneity we are interested in across skill, sector, age and contract type can be observed on the individual level. Furthermore, there is no clear method of measuring statistical power within the SCM. As such, results from a micro data analysis can complement our SC by providing statistical power.

The results from the SCM showed a treatment effect for low-skilled service workers. Even though this effect is shown to be less stable in the leave-one-out placebo test, we believe it warrants further investigation. In our micro data analysis we therefore choose to focus on the potential heterogeneous effects across skill groups in the labor market. If we in this analysis find a similar result as the one within the SCM, it can validate the result, because of the statistical power provided by the micro data. This section proceeds by presenting a general conceptual framework for an event study, followed by a description of our data source and subsequent considerations made for our event study model. It concludes with descriptive statistics relating to our specification.

## 8.1 Conceptual Framework

The event study design has in recent years grown in use for policy analysis. The design aims to estimate the impact of an event, or a treatment, where treatment differs across units and time. For instance, a policy intervention might be introduced at different times across states or regions within a country. This method can be applied to both panel and cross sectional data. Clarke and Tapia-Schythe (2021) present a generalized event study specification as:

$$Y_{gt} = \alpha + \sum_{j=2}^{J} \beta_j (\text{Lead } j)_{gt} + \sum_{k=1}^{K} \gamma_k (\text{Lag } k)_{gt} + \mu_g + \lambda_t + \mathbf{X}'_{gt}\mathbf{\Gamma} + \epsilon_{gt} \tag{11}$$

where $Y_{gt}$ denotes the outcome variable for a particular unit or group at a particular time, $\mu_g$ and $\lambda_t$ are group and time fixed effects, $\mathbf{X}_{gt}$ are time-varying controls and $\epsilon_{gt}$ is an unobserved error term. Lead $j$ and Lag $k$ are dummy variables that indicate how many periods away from the intervention each group is in each time period. The omitted category in this specification is the year before treatment, $j = 1$. The coefficients on the lead and lag variables are thus compared to the year preceding the intervention. Within the SCM, a lag is referred to as a year preceding treatment. However, in this framework, a lag should be interpreted as the estimate for a year after treatment.

We note that an event study specification can just as well be written as DiD specification. The event study design can be seen as an extension of the DiD method, where treatment is allowed to

[3] We applied for micro-level data from the French Labor Force Survey (FLFS) in anticipation of results from our SCM, which were then to be investigated further. Our aim was that potentially significant treatment effects were to be further analyzed and contrasted with variation within France.

vary across units and time. The control group, or counterfactual, in an event study design thus consists of units or groups in periods where they are untreated, as well as, if existing, groups or units that are never treated.

### 8.1.1 Assumptions for identification

The assumptions for identification within the event study design are similar to those for a DiD design (Clarke & Tapia-Schythe, 2021). Both methods are used to study the causal effect of a intervention, by comparing the outcome for the studied unit to a control unit that did not receive the treatment. If one observes a change, or a difference, in the differences between the treated and control group after the policy, it can under the right assumptions be interpreted as a causal effect. In Equation 11 such effects are captured in the lag estimates.

In order for the model to be able to estimate any causal effects, the parallel trends assumption must hold (Cunningham, 2018). The assumption states that in the absence of treatment, differences between the treatment and control groups would be the same over time – with trends parallel. If a trend break is observed after intervention for the treated group, it indicates a causal effect. In practice, it is not possible to fully verify the parallel trends assumption, since the outcome in the absence of treatment is never observed. However, the assumption can be partly checked by studying if trends between treatment and control groups are parallel in the periods preceding the policy. In an event study specification this is done by by studying the coefficients of the event leads. If the trends are not parallel for the treatment and control group before intervention, it is not likely that they they will be parallel after the intervention. Any trend break will thus not capture a causal effect.

## 8.2 The French Labor Force Survey

Our data source for the event study consists of the FLFS, in French called *Enquête Emploi* for years 1993-2007. The survey is collected and produced by the National Institute of Statistics and Economic Studies (*Institut national de la statistique et des études économiques*, INSEE), and distributed by Archives of Data Issued from Public Statistics (*Archives de Données Issues de la Statistique Publique*, ADISP). It uses a stratified systematic sampling method to collect data on labor market outcomes, such as employment status, salary and contract type, as well as demographic variables on the household and individual level. In 2003, the survey underwent a large transformation. Sampling went from being performed in March every year to being performed continuously, with results reported every quarter. The sampling strategy, a rotational panel design, also changed. Before 2003 the annual sample size was approximately 100,000 and a third of the sample was renewed each year, with each household being interviewed for three consecutive years in total. After the change, a sixth of the sample was renewed each quarter, and households were followed for six quarters in total, corresponding to 1.5 years. The frequency thus increased, but the sample size for each quarterly data reporting was reduced to approximately 50,000. Due to the change in sampling strategy during the time period of study, we choose to regard the data as a repeated cross section, as opposed to rotating panels. Our final model specification thus differs from the one in Equation 11, as we will not employ fixed effects but instead OLS, and therefore not make use of the panel nature of the data. In order to avoid potential fluctuations from seasonality, after 2003 we use the sample reported for the first quarter each year. We thus know that observations are measured during the same time of each year.

#### 8.2.1 Obstacles to causal inference

There are key characteristics of the data and the policy that limit the scope and design of our micro data analysis compared to the general framework presented in Section 8.1. This section aims to clarify these differences, explain their implications and establish what our model can investigate instead.

The first difference to the generalized event study outlined above is the lack of time variation in treatment. As opposed to many other examples where the event study method is employed, the WTR policy was announced in 1998 for all of France. Enforcement differed on the firm level, but as we will return to, we do not believe this heterogeneity can be leveraged in our specification. Furthermore, the timeline for enforcement was announced publicly, meaning there are likely anticipation effects from the policy starting in 1998. Thus, the treatment period is the same for all individuals.

Secondly, due to the sweeping nature of the policy, the data do not contain an appropriate control group that can be seen as untreated and comparable to a treated unit. Although much of the public sector was only mildly affected by the RTT, and for instance teachers were left entirely untreated, we do not view the public sector as a valid control group. This is because the employment protection legislation of the public and private sector are distinctly different in France. For example, a teacher is much less likely to lose their job compared to a French employee in a private firm, due to the French *fonctionnaire* status, which is a special status with favorable employment conditions within the French public sector (Ministère de la Transformation et de la Fonction publiques [Ministry of Transformation and Public Service] 2021). Since we are studying employment effects, there is little reason to believe that in absence of the policy the employment for public sector employees and low-skilled private sector workers would have had a similar development over time. In other words, the parallel trends assumption is unlikely to hold. An alternative control group is proposed by Estevão and Sá (2008), who utilize the variation in enforcement across firms. As enforcement of the policy was delayed by two years for small firms ($\leq$19 employees) they are viewed as untreated in this period and used as a control group for large firms with earlier implementation. Similarly to the reasoning for teachers and the public sector, we view small and large firms to be fundamentally different. A policy that aims to alter the working hours of employees will likely have different effects on a large company with more employees and a complex organizational structure, relative to a small and more flexible company with fewer employees. Thus, we find it unsuitable to use small firms as a control group in our specification.

In spite of the aforementioned obstacles to causal inference, we believe that investigating relative changes among skill groups is still of interest. Although far from perfect comparison groups, we choose to compare the employment probabilities of low-skilled with that of high and medium-skilled groups. The simultaneous treatment of these groups will limit the possible causal inference that can be drawn from the analysis, but enables a comparison of how employment developed after the policy in each of the groups. The approach serves as a way around the identification issues. We believe that if we observe parallel trends across skill groups before the RTT, heterogeneous employment changes across skill levels directly after the policy, this will give strong indications of heterogeneous policy effects, in spite of the obstacles to causal inference. We therefore proceed with constructing an event study model.

## 8.3 Model

This section describes the model specification used for our event study, as well as some methodological considerations due to the nature of the data. In order to study employment effects for low-skilled workers from the French RTT, we study the employment probability of workers across three different skill levels in three specifications, similar to the one shown in Equation 11. We then compare the estimates for the lags of the low-skilled group with the remaining skill groups, in order to establish whether the employment probability for the low-skilled group changed relative to the others after the reform. Using the same logic as within the SCM, we set 2007 as our last time period of study.

The outcome variable for the event study model is a binary variable measuring employment, partly following Estevao and Sá's (2008) DiD specification. We employ a linear probability model (LPM). In an LPM, all included covariates need to be categorical, in order to account for the binary nature of the outcome variable. Moreover, estimates should be interpreted as probabilities.

The FLFS codes individuals into three activity types, employed, unemployed or inactive. One approach would be to restrict the sample to only study the unemployed and the employed. However, transition from unemployment to employment, and vice versa, is a naturally occurring process in the economy among workers. Studying such transitions would potentially not capture structural changes to the labor force as individuals go from being inactive to participating in the labor market. For example, an important share of the individuals in their late teens and early twenties will be university students, and will therefore be coded as inactive in the FLFS. The ease of labor market entry for these new job seekers may reflect some of the structural effects of this type of policy, as firms adjust by hiring more or less junior employees. Moreover, classifications of unemployment in labor force surveys is often restrictive, meaning that the share of unemployed in the sample is potentially underestimated. Restricting our sample to only the employed and unemployed would mean a large part of the sample, which we believe would be impacted by the policy, would be excluded. We therefore widen the scope and include inactive individuals in our sample. Any observed employment changes will therefore be indicative of structural changes on the extensive margin in the labor market.

However, because we include inactive individuals, as opposed to only studying the employed and unemployed, we do not observe the economic sector they have, or would have been, employed in. This since the FLFS only contains sector measurements of the last held job for the unemployed. The division of the sample into skill-sector groups, as performed within the SC model, is thus no longer possible when including the inactive individuals. We therefore restrict the scope of the analysis for the event study, and choose to only study differences in employment outcomes across skill levels, as opposed to skill-sector levels.

Skill is proxied by education level in our model. We classify education into three skill levels: low, medium and high. The detailed coding of education levels into skill levels can be found in Appendix C. We could opt for a more fine-grained analysis and include more skill levels. However, this would lead to a lower number of observations for each specification, and thus a loss of statistical power. Therefore, we opt for a three-level skill classification.

The announcement of the policies for both small and large firms occurred in June 1998. Since the FLFS at this time took place in March each year, our treatment date is set to 1999, as opposed to 1998 for the SCM, as it is the first period where any effects will be observed. Several control variables are included in our final event study specification. We control for the same set of covariates in each model. Taking inspiration from Estevão and Sá (2008) we control for sex, dummy variables of being married and having at least one child under the age of six, and age in five year bands. Metropolitan France (which excludes overseas regions) is divided into 13 *régions*, henceforth regions, which are further subdivided into 96 *départements*, henceforth departments. We also include department as a control variable. Further, we also interact the covariates of sex, being married and having at least one child under six years old. In line with previous research regarding child penalties, such as the event study by Kleven et al. (2019), we believe women's employment probabilities covary differently with being married and having a young child compared to men.

We cluster standard errors on the department level, as opposed to regions. This way, we make sure we do not have too few clusters, following the rule of thumb by Angrist and Pischke (2009) that the number of clusters should ideally be equal to or greater than 42. We choose to cluster standard errors as we believe observations will be affected differently across departments. Within these departments, or clusters, there will be potential for covariance across observations, due to differing regional economic conditions across departments. The standard errors in our specifications are therefore cluster-robust.

In the final sample we choose to exclude some observations. Individuals employed within the public sector are excluded, since they were not impacted by the policy in the same way as other employees, and have special status in France with strict employment protection as explained above. Individuals that are self-employed are also excluded from the final sample, since their employment status differs from employees and we imagine that they will not be impacted in the same way by the policy. We choose to only include individuals aged 18-64. The upper limit of 64 follows the one chosen within the SCM. We choose to exclude individuals aged 15-17, since a majority of these will be in school and will therefore not be affected by the RTT. Moreover, as opposed to Estevão and Sá (2008), we include part-time workers in our final sample. Even though only full-time workers were affected by the WTR, we believe that some of its potential employment effects will be observed through part-time employment.

Our final model specification thus becomes:

$$Y_{igt} = \alpha_g + \sum_{j=2}^{J} \beta_j (\text{Lead } j)_g + \sum_{k=1}^{K} \gamma_k (\text{Lag } k)_g + \mathbf{X}'_{ig}\mathbf{\Gamma} + \epsilon_{igt} \tag{12}$$

where $Y_{igt}$ denotes employment for an individual $i$ in skill group $g$ in time $t$, and $\alpha$ is a group-varying intercept. Lead $j$ and Lag $k$ are coefficients of employment probability for being on the labor market a certain year for a certain group, relative to the omitted year 1998, $\mathbf{X}_{ig}$ is a vector of the control variables for an individual $i$ in group $g$ and $\epsilon_{igt}$ is an unobserved error term.

### 8.3.1 Methodological considerations

Several methodological considerations have been taken into account before opting for an LPM. A plausible alternative would have been to study employment probabilities through a logit or probit model, which calculates the log odds ratio of probabilities. There is a clear advantage to using such a model compared to an LPM. In a logit or probit model, the sum of the probability coefficients are constrained to lie in the range between 0 and 1. On the other hand, within an LPM framework it is possible that the sum of the probability coefficients and intercept sum to more than one, which is an inconsistent result. In spite of this, we have moved forward with an LPM, since we believe its results are more intuitive to interpret. Also, for the type of marginal effects that we are attempting to study the difference between the two alternatives are often negligible (Angrist & Pischke, 2009).

In addition, there are alternative ways to study employment effects as opposed to looking at employment probability. In line with Estevão and Sá (2008) one could make use of the three year panel nature of the FLFS, and study transitions in and out of employment. Results from such an analysis would not specifically study the employment level, they would rather shed light on how workers were more directly affected by the policy. However, this approach is complicated by the change in sampling strategy of the FLFS in 2003, where the frequency of the panels changed. Since we wish to study a longer post-treatment period, exploiting the panel nature of the data is therefore not an option.

The absence of a credible control group in our data necessitates a discussion on what the lack of counterfactuals implies for the analysis of our results. We run our specification for each skill group, in which the estimates for the leads and lags are set relative to the respective employment probability of each group in the year preceding treatment. The baseline for each group, which in a general event study specification would be the counterfactual, is thus the probability of employment in the year before treatment, in this case 1998. The inferences that can be made from our analysis will therefore be (a) whether there is a change in employment probability for each group relative to the year preceding treatment and (b) if the relative development after the policy is heterogeneous across the three skill groups. Heterogeneous effects will be observed if the relative development of groups is different after the policy introduction. However, the lack of a control group and our relative measures mean that we will be unable to infer any causality of the policy, as treatment is simultaneous for all skill groups. Before presenting and interpreting the results, the descriptive statistics for all variables in our specification are presented in the next section.

### 8.4 Descriptive Statistics

Table 6 shows descriptive statistics for our outcome and control variables across the three skill groups. We note that the total number of observations between 1993 and 2007 amounts to around 1.2 million. They are concentrated to the high-skilled group, which accounts for around 65 percent, or 800,000 observations, of the total sample. We also see that the share of inactive individuals is concentrated to the medium and high-skilled groups. Moreover, we observe that the two lower skill levels have a slight majority of females. In addition, we note that marriage is more common for the low and high-skilled groups, and having children under the age of six is more common in the low-skilled group.

**Table 6:** Descriptive statistics

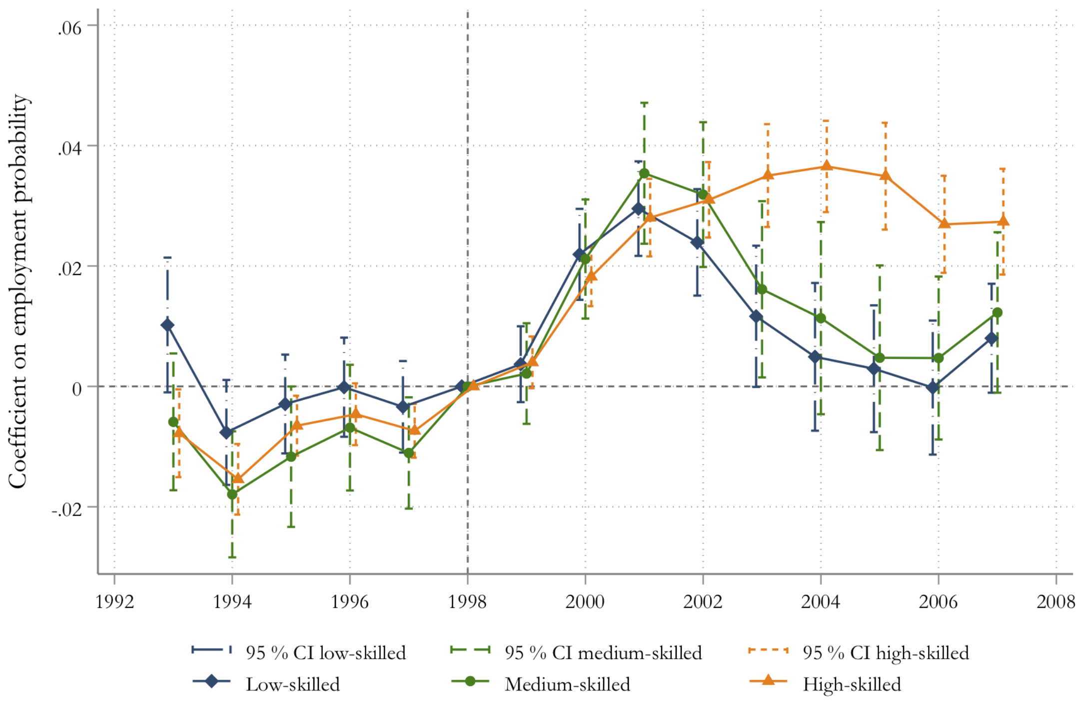
| | | Low | Medium | High | Total |
|---|---|---|---|---|---|
| Employed | N | 178,436 | 92,631 | 443,536 | 714,603 |
| | % | 75.01 | 54.41 | 55.55 | 59.23 |
| Unemployed | N | 13,423 | 11,790 | 73,554 | 98,767 |
| | % | 5.64 | 6.92 | 9.21 | 8.19 |
| Inactive | N | 46,023 | 65,832 | 281,338 | 393,193 |
| | % | 19.35 | 38.67 | 35.24 | 32.59 |
| Male | N | 110,057 | 77,897 | 412,096 | 600,050 |
| | % | 46.27 | 45.75 | 51.61 | 49.73 |
| Female | N | 12,7825 | 92,356 | 386,332 | 606,513 |
| | % | 53.73 | 54.25 | 48.39 | 50.27 |
| Married | N | 122,602 | 64,772 | 469,444 | 656,818 |
| | % | 51.54 | 38.04 | 58.80 | 54.44 |
| Children | N | 48,756 | 26,664 | 126,874 | 202,294 |
| | % | 20.50 | 15.66 | 15.89 | 16.77 |
| Mean age | | 37.81 | 32.91 | 41.50 | 39.56 |
| Total | | 237,882 | 170,253 | 798,428 | 1,206,563 |

Source: *Enquête Emploi* [French Labour Force Survey] 1993-2007. Producer: INSEE. Distributor: ADISP.
Note: Variable children is an indicator variable for having one or more children under six years of age.

# 9 Event Study Results

We will now proceed to present the results from the event study. Figure 12 plots the coefficients of the leads and lags from Equation 11 for the three respective skill groups in our studied time period. Each estimate includes a 95 percent confidence band. The baseline year is 1998 for all skill groups. The outcome variable is a binary indicator for employment, and can be interpreted as the probability of employment. Estimates should thus be seen as what change in employment probability an individual experiences from being on the labor market in a specific year compared to 1998. Appendix D contains the full set of estimates from the three regressions. As previously mentioned, our final sample includes employed, unemployed and inactive individuals. This means that all results should be interpreted as structural changes to the labor force. If the employment probability increases, it indicates an increased transition to employment for both the unemployed and inactive.

**Figure 12:** Event study model




For all groups, we see a significant positive change in employment probability starting in 2000, the year of policy enforcement. This change corresponds to a two percentage points positive increase in employment probability. For the low and medium-skilled, this change starts to decrease in 2002, and in 2003 it is no longer significant. The high-skilled group on the other hand remains on a statistically significant higher level in employment probability throughout the studied period, and increases to around three percentage points. We also note that the confidence intervals for the high-skilled group are tighter, which is probably due to the larger sample size for this group. We are unable to attribute these positive changes in employment probability to the policy since we lack a control group. However, the results indicate that workers of all skill levels on the French labor market experienced positive employment effects in the years directly after policy enforcement.

As previously mentioned, the parallel trends assumption is not essential to our model. However, if we do observe parallel trends, and a trend break after treatment, this implies that there has been a change in the employment probability with heterogeneous effects across skill groups around the same time as the policy.

When studying Figure 12 we see that all skill groups developed similarly before the RTT, with no statistically significant differences between groups. In other words, we observe parallel trends in the pre-intervention period. However, since this period is only five years we do not know if this relationship is temporary or constant in nature. We note that after 2002, the year of enforcement for small firms, the change in employment probability is statistically different between the high and low-skilled group. Furthermore, in 2004, the change in employment probability for the high-skilled group is significantly higher than both the low and medium-skilled groups. Our results therefore indicate that there were heterogeneous changes in employment probabilities across skill groups starting in 2003, as opposed to the pre-treatment period where development was relatively homogeneous across skill groups. However, since the heterogeneous changes start in 2003, we cannot be sure that this heterogeneity is attributable to the policy.

Moreover, Figure 12 shows that the low and medium-skilled groups have similar trajectories of their employment probability changes both before and after the policy. We hypothesized that low-skilled groups would be more affected by the RTT, due to the nature of both the policy and certain low-skilled work. Our results show no indication that this group developed differently compared to the medium-skilled group.

After specifying and running the event study model, we also perform an F-test in order to test the joint significance of the respective leads and lags. We see in Table 7 that these leads and lags are jointly significant for all three models, with p-values ranging from 0.2 percent to being infinitely small. This means that we reject that the estimates for our leads and lags are zero at any conventional significance level. In other words, we know that the employment probability for our respective skill groups changed across our studied time period.

**Table 7:** Joint significance test for leads and lags

| Low-skilled | | Medium-skilled | | High-skilled | |
|---|---|---|---|---|---|
| Leads (1992-1997) | | | | | |
| F-stat | 5.5418 | F-stat | 4.1260 | F-stat | 7.2265 |
| P-value | 0.0002 | P-value | 0.0020 | P-value | 0.0000 |
| Degrees of freedom | (5,94) | Degrees of freedom | (5,94) | Degrees of freedom | (5,94) |
| Lags (1999-2007) | | | | | |
| F-stat | 8.8059 | F-stat | 6.1174 | F-stat | 14.4641 |
| P-value | 0.0000 | P-value | 0.0000 | P-value | 0.0000 |
| Degrees of freedom | (9,94) | Degrees of freedom | (9,94) | Degrees of freedom | (9,94) |

To summarize, we find no indication of heterogeneous effects across skill groups in the years directly after the policy. However, we find heterogeneity across skill starting in 2003. The direction of this heterogeneity goes in the opposite direction of our hypothesis, which states that employment effects for low-skilled workers would be more positive compared to other groups. However, since this occurs after policy enforcement, and due to the lack of control group, we cannot attribute it to the RTT.

# 10 Discussion

After this two-step analysis, using an SCM as well as an event study design, we conclude that we are unable to find employment effects or heterogeneity in employment probability across skill after the French RTT. This section will now relate our findings to previous literature, shed light on some methodological concerns, as well as discuss the study's implications for policymaking.

Within the SC framework, we cannot observe any effects from the policy for low-skilled industry workers and low-skilled youth employment. These null results are strengthened by the results from the placebo tests. We observe a negative treatment effect for the senior labor force that cannot be attributed to the RTT, and is not robust to placebo tests, which we therefore interpret as a null result. Moreover, the observed negative treatment effects for the variables low-skilled service workers and the prevalence of temporary contracts are not robust to placebo tests either. The results for these outcome variables are largely driven by specific donors, making it difficult to draw clear conclusions regarding null results. The results from the event study show that all skill groups experienced significant positive changes to employment probability after the reform. Since our F-test concluded that the lags and leads were jointly significant, we can conclude that there were significant changes to employment probability during the studied period. However, we do not find any heterogeneity in employment probability across skill levels in the years directly following the policy. These results are in line with previous evaluations of the RTT by Chemin and Wasmer (2009) and Estevão and Sá (2008), who also are unable to find significant employment effects from the RTT.

Data limitations could explain the lack of reliable treatment effects within our SCM. Firstly, we note that the composition of the donor pool may negatively impact the validity of the results. Even though any donors that introduced a similar WTR during the studied period were excluded from the pool, other issues remain. For instance, the members of the donor pool may be heterogeneous with respect to labor markets in a way that makes them poor candidates for a counterfactual, which risks leading to interpolation bias. This may not only be due to geographical and cultural factors, but also due to different economic structures across European countries. For instance, the use of temporary contracts was much more frequent in Spain compared to the rest of the donor pool during our studied period, as seen in Figure 7. This implies that even though a donor country was not subject to a WTR, it might still not be a suitable control unit. However, after having visually inspected the variables we do not believe this was a major problem with our SCM. We instead believe that a more serious issue stems from studying labor markets within an SCM. As a result of their political importance, labor markets across Europe tend to be subject to frequent policy changes (Kluve et al., 2007). The outcome variables of interest may be indirectly affected by such interventions. These shocks can potentially show up in our SC and, if the SC is sparse, make the counterfactual reflect a specific shock in a donor country.

Proponents of the SCM would argue that problems with heterogeneity of donors can be alleviated by expanding the donor pool or the pre-treatment period. This was however not possible in our case due to limited data availability. This limitation can be attributed to two factors. Firstly, the data was collected from the European Labour Force survey through ILOSTAT, to ensure harmonized and comparable data. However, this made the data availability dependent on accession to the EU, and thereby limited the size of the potential donor pool. In order to have enough pre-treatment data we could thus only include countries that were EU members before 1992, which prevented the use of other potential donors, such as the Scandinavian countries. Secondly, in order to meet Abadie's (2021) second contextual SC requirement, three to four potential donors were excluded from the SCs, since they introduced WTRs during the period of study. Although we observe reasonable pre-treatment fits for several of our variables of interest, the limited data availability increases the risk of obtaining spurious fits within the SCM. A spurious fit would mean the obtained treatment effects are inconsequential, since they are not the result of an economic relationship. Thus, the SCM may have more merit if studying our labor market outcomes with more and better data.

The results from the event study cannot be attributed to the policy, due to the lack of a control group. Our idea was that if we received heterogeneous changes after the policy implementation, it would indicate that the policy affected skill groups differently. However, this was not the result we obtained within the event study. Instead, all skill groups experienced positive employment changes after the policy implementation. There are several other variables that could explain this change. Since all groups were affected to relatively the same extent in the years directly after the policy, it could instead be the result of an economic boom in France during this period. Indeed, the country underwent a period of growth rates of around 3.5 to four percent in the years 1998-2000 (OECD, 2022b). It is therefore highly likely that it is rather the effect of this economic boom that is displayed in the event study results, as opposed to positive employment effects from the RTT.

The problem of confounders in the analysis could potentially be solved by more detailed data. For instance, with data on what particular sectors were early adopters of the policy, a more fine-grained analysis on adoption effects could be performed. Moreover, if observations could be matched to employer firms, and if these were early adopters of the policy, this could be used to construct a control group. However, such analyses are not within the scope of this thesis, but could be the topic of further research.

We note in the results that the low and medium-skilled groups exhibit similar trajectories over the studied time period, compared to the high-skilled group. This calls into question both the hypothesis that the reform would give rise to positive employment effects for low-skilled workers compared to workers of other skill levels, as well as our skill measurements. For instance, it could indicate that our hypothesis, that workers of different skill levels are affected differently by the policy, is incorrect. On the other hand, the similar trajectories could indicate that low-skilled labor is not as different to what we classify as medium-skilled labor and that our measurement of low-skilled labor is thus too narrow. What we observe in our results is rather a dichotomous development, which suggests a dual separation of the labor market according to skill.

Moreover, we note that the employment probability for the low and medium-skilled groups started to decrease in 2002 relative to the 1998 level. This is around the time that France started to experience decreasing growth rates. It could therefore indicate that the obtained employment probabilities followed the business cycle for these skill groups. In contrast, the high-skilled group remains on the higher employment probability level, in spite of the decreasing French growth rates. Since the heterogeneity across skill levels is observed after policy enforcement and starts in 2003, it is unlikely that it is a result of workers across different skill levels being affected differently by the RTT. Instead, one can imagine that it is a result of the education premium.

As previously mentioned, there is an inherent issue in studying employment effects across economic sector levels using micro data. The source of variation across sectors means that all individuals with a sector measurement are employed, disabling any meaningful analysis on employment. Our way around this is to only study employment effects across skill levels. There are, however, possible extensions to this method that could alleviate the sector variation problem. For instance, one could specify a multinomial logit model that could be used to compute probabilities of individuals with certain characteristics entering certain sectors. The probabilities could then be extrapolated to assign sectors to unemployed or inactive individuals. This is could be incorporated in an extension to this thesis.

In spite of the mentioned limitations to both methodology and data, the obtained results indicate that the policy did not produce any significant employment effects, and did not affect workers across skill levels differently. This is in contrast to the theoretical predictions by Calmfors and Hoel (1988) and Hunt (1999), which state that a WTR policy will lead to a reduction in employment. There are several possible economic explanations behind this result. For instance, reducing the work week by four hours may not constitute a large enough structural change for clear employment effects to take place. One could imagine that a policy with a larger hour reduction, would have led to clearer policy effects. Moreover, since the four-hour change was accompanied by increased flexibility among companies with regards to implementation and organization, it is possible that potential positive employment effects were weakened by this policy element.

The lack of aggregate employment effects can be studied through Hunt's (1999) framework. Within Hunt's model, employment effects from a WTR occurs through a negative scale effect on both labor and capital, a substitution effect from labor to capital, as well as a substitution from workers to hours, due to changes in the marginal costs of workers and hours. Previous literature by Gilles (2015) has shown that the use of capital was not affected by the RTT, which indicates that there was no scale effect as a result of the policy, nor a substitution from labor to capital. Our results can thus indirectly be seen to support those of Gilles (2015). In addition, the policy may not have affected the marginal cost of either employment or additional hours enough for a substitution from workers to hours to take place, which would explain the lack of observed employment effects. For instance, one can imagine that negative employment effects resulting from a substitution from workers to hours were offset by the cuts in social security contributions, which lowered the fixed costs of employment. To sum up, the mechanisms behind employment effects from a reduction in working time in Hunt's framework were perhaps not strong enough to produce employment effects in the case of the RTT.

When formulating our hypotheses, we believed that heterogeneity across skill and sector would be important when studying the effects of a WTR policy, something that was not taken into account in Hunt's (1999) labor demand model. Since we found no indication of heterogeneity in effects across workers with different skill levels within the event study, this indicates that heterogeneity across workers is not as consequential as we theorized. Rather, low-skilled workers do not seem more sensitive to WTRs compared to other skill groups. Moreover, the lack of employment effects within the industry and services sector suggests that the work in these sectors does not have the hypothesized characteristics of being especially sensitive to reductions in standard working hours. In other words, we receive no evidence to support our claim of the importance of heterogeneity in effects across workers of different skill and sector groups.

Results from the SC strengthen the claim that the policy had little effects on employment. A contextual requirement of the SCM is that treatment effects must be large enough in magnitude compared to other shocks to the outcome variable (Abadie, 2021). Overall, the obtained results are small in magnitude, and there is therefore a risk that other idiosyncratic shocks influence these more strongly than the RTT. Specifically, we note that a time series break in the data from Great Britain led to a negative treatment effect for our low-skilled service workers outcome variable, which disrupted our analysis. The time series break was large in magnitude in comparison to the other obtained, although non-robust, treatment effects (ranging from 0.5 to two percentage points). This indicates that effects from the policy might not be possible to distinguish from other shocks to the outcome variable, such as time series breaks, and this contextual requirement breaks down. The insufficient effect magnitude gives further indication that the policy overall had little effect on employment.

Several limitations prevent this study from establishing employment effects from the RTT. However, if the French WTR policy was in fact relatively toothless for employment, this could lead to relevant policy implications. Considering the enduring popularity and debate of this type of reform, the arguments for implementation should perhaps not come from benefits to employment. Instead its raison d'être should be considered through other benefits, such as increased leisure time, potential productivity gains or positive environmental effects. Policymakers might need to reframe the goals of a WTR policy or, if unemployment is indeed the goal, look at alternatives with stronger empirical support than the French RTT.

We thus conclude that studying any causal effects of the French WTR policy is associated with several methodological obstacles, regardless of whether micro or macro data are used. Some of these issues can be solved through improved data availability. However, much of it stems from the lack of available control units.

# 11 Conclusion

In 1998 the French government announced two laws shortening the standard workweek from 39 to 35 hours, with the explicit goal of reducing unemployment through work sharing. This master's thesis aimed to investigate the employment effects of this policy, focusing on specific groups on the labor market that were hypothesized to be especially affected. Namely, employment in the industry and services sector were believed to be positively impacted by the policy, due to the presence of shift work in these sectors. Moreover, since the policy introduced cuts in social security contributions concentrated towards workers earning the minimum wage, low-skilled workers were also hypothesized to reap positive employment effects. In addition, young and senior workers, as well as workers employed on temporary contracts, were believed to be especially affected.

The analysis was performed in two steps, and was framed within Hunt's simple static labor demand model. Firstly, aggregate data with variation across countries was collected and, using the synthetic control method, counterfactuals of what would have happened in France absent the policy were constructed for five worker groups. Overall, results from this analysis indicated that the policy did not lead to any employment effects for any studied groups. However, methodological concerns as well as data limitations should be taken into account when interpreting these. The results from the synthetic controls were then used as a starting point for supplementary analysis, using micro data and variation within France through an event study design to study possible heterogeneity in employment effects across workers of different skill. The study found no indication of such skill heterogeneity, but instead positive changes to employment for all skill groups in the years directly after the policy. Due to the lack of control group, these changes were not attributable to the policy.

Studying the effects of nationwide reductions in working hours is of importance, since such policies are being increasingly advocated for by social actors ranging from individual firms to policy makers and trade unions. This master's thesis gives several contributions to the empirical literature of WTRs. In line with previous evaluations of the French RTT, we provide further evidence that it led to little employment effects. Moreover, we also test the hypothesis of potential heterogeneous effects across certain labor market groups. In spite of concerns regarding heterogeneity in effects for workers across different skill levels and sectors, ages and contract types, we find that these are not especially impacted by reductions in standard working hours. A final conclusion to be drawn from this study is thus that the goal of WTR policies should by guided by other motives, such as increased leisure time or productivity gains, rather than work sharing.

## 12 References


Abadie, A. (2021). Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects. *Journal of Economic Literature*, *59*(2), 391–425. https://doi.org/10.1257/jel.20191450

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association*, *105*(490), 493–505. https://doi.org/10.1198/jasa.2009.ap08746

Abadie, A., Diamond, A., & Hainmueller, J. (2015). Comparative Politics and the Synthetic Control Method. *American Journal of Political Science*, *59*(2), 495–510. https://doi.org/10.1111/ajps.12116

Abadie, A., & Gardeazabal, J. (2003). The Economic Costs of Conflict: A Case Study of the Basque Country. *American Economic Review*, *93*(1), 113–132. https://doi.org/10.1257/000282803321455188

Abrassart, A. (2015). Low-skill jobs or jobs for low-skilled workers? An analysis of the institutional determinants of the employment rates of low-educated workers in 19 OECD countries, 1997–2010. *Journal of European Social Policy*, *25*(2), 225–241. https://doi.org/10.1177/0958928715573485

Andrews, D., Charlton, A., & Moore, A. (2021). *COVID 19, Productivity and Reallocation: Timely evidence from three OECD countries* (OECD Economics Department Working Papers No. 1676). Retrieved May 3, 2022, from https://www.oecd-ilibrary.org/economics/covid-19-productivity-and-reallocation-timely-evidence-from-three-oecd-countries_d2c4b89c-en

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics: an empiricist's companion* [ISBN: 978-0-691-12034-8]. Princeton University Press.

Askenazy, P. (2013). Working time regulation in France from 1996 to 2012. *Cambridge Journal of Economics*, *37*(2), 323–347. https://doi.org/10.1093/cje/bes084

BBC. (2019). Macron pension reform: Strike continues for second day. *BBC News*. Retrieved May 11, 2022, from https://www.bbc.com/news/world-europe-50682071

Becker, G. S. (1962). Investment in Human Capital: A Theoretical Analysis. *Journal of Political Economy*, *70*(5, Part 2), 9–49. https://doi.org/10.1086/258724

Bhaduri, A., & Marglin, S. (1990). Unemployment and the real wage: the economic basis for contesting political ideologies. *Cambridge Journal of Economics*, *14*(4), 375–393. Retrieved May 1, 2022, from http://www.jstor.org/stable/23598376

Bilous, A. (2000). *Law on the 35-hour week is in force* (tech. rep.). Eurofound. Retrieved May 2, 2022, from https://www.eurofound.europa.eu/publications/article/2000/law-on-the-35-hour-week-is-in-force

Boppart, T., & Krusell, P. (2020). Labor Supply in the Past, Present, and Future: A Balanced-Growth Perspective. *Journal of Political Economy*, *128*(1), 118–157. https://doi.org/10.1086/704071

Bosch, G., & Lehndorff, S. (2001). Working-time reduction and employment: experiences in Europe and economic policy recommendations. *Cambridge Journal of Economics*, *25*(2), 209–243. https://doi.org/10.1093/cje/25.2.209

Brynjolfsson, E., & McAfee, A. (2016). *The second machine age: work, progress, and prosperity in a time of brilliant technologies*. W. W. Norton & Company.

Cahuc, P., & Michel, P. (1996). Minimum wage unemployment and growth. *European Economic Review*, *40*(7), 1463–1482. https://doi.org/10.1016/0014-2921(95)00035-6

Calmfors, L., & Hoel, M. (1988). Work Sharing and Overtime. *The Scandinavian Journal of Economics*, *90*(1), 45. https://doi.org/10.2307/3440149

Calmfors, L., & Hoel, M. (1989). Work Sharing Employment and Shiftwork. *Oxford Economic Papers*, *41*(1), 758–773. https://doi.org/10.1093/oxfordjournals.oep.a041926

Campus France, French degrees, LMD system and equivalences [The French Agency for the promotion of higher education, international student services, and international mobility.]. (n.d.). Retrieved May 3, 2022, from https://www.campusfrance.org/en/French-degrees-LMD-equivalences

Card, D. (1990). The Impact of the Mariel Boatlift on the Miami Labor Market. *Industrial and Labor Relations Review*, *43*(2), 245. https://doi.org/10.2307/2523702

Card, D., & Krueger, A. (1993). *Minimum Wages and Employment: A Case Study of the Fast Food Industry in New Jersey and Pennsylvania* (tech. rep. w4509). National Bureau of Economic Research. Cambridge, MA. https://doi.org/10.3386/w4509

Cárdenas, L., & Villanueva, P. (2021). Challenging the working time reduction and wages trade-off: a simulation for the Spanish economy. *Cambridge Journal of Economics*, *45*(2), 333–351. https://doi.org/10.1093/cje/beaa055

Chemin, M., & Wasmer, E. (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*(4), 487–524. https://doi.org/10.1086/605426

Clarke, D., & Tapia-Schythe, K. (2021). Implementing the panel event study. *The Stata Journal: Promoting communications on statistics and Stata*, *21*(4), 853–884. https://doi.org/10.1177/1536867X211063144

Cunningham, S. (2018). *Causal Inference: The Mixtape.*

Doudchenko, N., & Imbens, G. (2016). *Balancing, Regression, Difference-In-Differences and Synthetic Control Methods: A Synthesis* (tech. rep.). National Bureau of Economic Research. Cambridge, MA. https://doi.org/10.3386/w22791

Echevarría, C. A., & García-Enríquez, J. (2019). The economic cost of the Arab Spring: the case of the Egyptian revolution. *Empirical Economics*, *59*(3), 1453–1477. https://doi.org/10.1007/s00181-019-01684-7

Estevão, M., & Sá, F. (2008). The 35-hour workweek in France: Straightjacket or welfare improvement? *Economic Policy*, *23*(55), 417–463. https://doi.org/10.1111/j.1468-0327.2008.00204.x

European Foundation for the Improvement of Living and Working Conditions. (2010). *ERM report 2010 : Extending flexicurity – The potential of short-time working schemes* (tech. rep.). Publications Office. LU. Retrieved May 3, 2022, from https://data.europa.eu/doi/10.2806/21361

Ferman, B., Pinto, C., & Possebom, V. (2020). Cherry Picking with Synthetic Controls (B. Barnow, Ed.). *Journal of Policy Analysis and Management*, *39*(2), 510–532. https://doi.org/10.1002/pam.22206

Fiole, M., & Roger, M. (2002). Les effets sur l'emploi de la loi du 11 juin 1996 sur la réduction du temps de travail. *ÉCONOMIE ET STATISTIQUE*, *357-358*, 3–22.

France24. (2010). More disruptions expected after weekend protests. *France 24*. Retrieved May 11, 2022, from https://www.france24.com/en/20101017-france-faces-further-disruption-after-weekend-protests-strikes-pension-reform

Gilles, F. (2015). Evaluating the Impact of a Working Time Regulation on Capital Operating Time: The French 35-hour Work Week Experience. *Scottish Journal of Political Economy*, *62*(2), 117–148. https://doi.org/10.1111/sjpe.12067

Haraldsson, G. D., & Kellam, J. (2021). *Going Public: Iceland's Journey to a Shorter Working Week* (tech. rep.). Alda, Association for Democracy and Sustainability.

Hunt, J. (1999). Has Work-Sharing Worked in Germany? *The Quarterly Journal of Economics*, *114*(1), 117–148. https://doi.org/10.1162/003355399555963

ILO. (2012). *International Standard Classification of Occupations: ISCO-08.* [OCLC: 794974321]. International Labour Office. https://ilostat.ilo.org/resources/concepts-and-definitions/classification-occupation/

ILOSTAT. (2022a). Employees by economic activity and occupation (thousands) [Type: dataset]. Retrieved February 23, 2022, from https://www.ilo.org/shinyapps/bulkexplorer40/?lang=en&segment=indicator&id=EES_TEES_ECO_OCU_NB_A

ILOSTAT. (2022b). Employees by sex, age and temporary or permanent contract (thousands) [Type: dataset]. Retrieved February 23, 2022, from https://www.ilo.org/shinyapps/bulkexplorer11/?lang=en&segment=indicator&id=EES_TEES_SEX_AGE_JOB_NB_A

ILOSTAT. (2022c). International Standard Industrial Classification of All Economic Activities (ISIC) [Section: methods]. Retrieved May 11, 2022, from https://ilostat.ilo.org/resources/concepts-and-definitions/classification-economic-activities/

ILOSTAT. (2022d). Labour force by sex and age (thousands) [Type: dataset]. Retrieved February 23, 2022, from https://www.ilo.org/shinyapps/bulkexplorer48/?lang=en&segment=indicator&id=EAP_TEAP_SEX_AGE_NB_A

ILOSTAT. (2022e). Labour force participation rate by sex and age (%) (Annual indicators) [Type: dataset]. Retrieved February 23, 2022, from https://www.ilo.org/shinyapps/bulkexplorer52/?lang=en&segment=indicator&id=EAP_DWAP_SEX_AGE_RT_A

ILOSTAT. (2022f). Unemployment rate by sex and age (%) [Type: dataset]. Retrieved February 23, 2022, from https://www.ilo.org/shinyapps/bulkexplorer8/?lang=en&segment=indicator&id=UNE_DEAP_SEX_AGE_RT_A

ILOSTAT. (2022g). Working-age population by sex and age (thousands) [Type: dataset]. Retrieved February 23, 2022, from https://www.ilo.org/shinyapps/bulkexplorer27/?lang=en&segment=indicator&id=POP_XWAP_SEX_AGE_NB_A

ILOSTAT. (2022h). Youth employment by sex, age and occupation (thousands) [Type: dataset]. Retrieved February 23, 2022, from https://www.ilo.org/shinyapps/bulkexplorer19/?lang=en&segment=indicator&id=EMP_3EMP_SEX_AGE_OCU_NB_A

ILOSTAT. (2022i). Youth labour force participation rate by sex, age and rural / urban areas (%) [Type: dataset]. Retrieved February 23, 2022, from https://www.ilo.org/shinyapps/bulkexplorer18/?lang=en&segment=indicator&id=EAP_3WAP_SEX_AGE_GEO_RT_A

ILOSTAT. (2022j). Youth unemployment rate by sex, age and rural / urban areas (%) [Type: dataset]. Retrieved February 23, 2022, from https://www.ilo.org/shinyapps/bulkexplorer18/?lang=en&segment=indicator&id=UNE_3EAP_SEX_AGE_GEO_RT_A

INSEE. (1993). Enquête Emploi - 1993 [Producer: INSEE, Distributor ADISP, Type: dataset].

INSEE. (1994). Enquête Emploi - 1994 [Producer: INSEE, Distributor ADISP, Type: dataset].

INSEE. (1995). Enquête Emploi - 1995 [Producer: INSEE, Distributor ADISP, Type: dataset].

INSEE. (1996). Enquête Emploi - 1996 [Producer: INSEE, Distributor ADISP, Type: dataset].

INSEE. (1997). Enquête Emploi - 1997 [Producer: INSEE, Distributor ADISP, Type: dataset].

INSEE. (1998). Enquête Emploi - 1998 [Producer: INSEE, Distributor ADISP, Type: dataset].

INSEE. (1999). Enquête Emploi - 1999 [Producer: INSEE, Distributor ADISP, Type: dataset].

INSEE. (2000). Enquête Emploi - 2000 [Producer: INSEE, Distributor ADISP, Type: dataset].

INSEE. (2001). Enquête Emploi - 2001 [Producer: INSEE, Distributor ADISP, Type: dataset].

INSEE. (2002). Enquête Emploi - 2002 [Producer: INSEE, Distributor ADISP, Type: dataset].

INSEE. (2003). Enquête Emploi en continu (version FPR) - 2003 [Producer: INSEE, Distributor ADISP, Type: dataset].

INSEE. (2004). Enquête Emploi en continu (version FPR) - 2004 [Producer: INSEE, Distributor ADISP, Type: dataset].

INSEE. (2005). Enquête Emploi en continu (version FPR) - 2005 [Producer: INSEE, Distributor ADISP, Type: dataset].

INSEE. (2006). Enquête Emploi en continu (version FPR) - 2006 [Producer: INSEE, Distributor ADISP, Type: dataset].

INSEE. (2007). Enquête Emploi en continu (version FPR) - 2007 [Producer: INSEE, Distributor ADISP, Type: dataset].

Kassam, A. (2021). Spain to launch trial of four-day working week. *The Guardian*. Retrieved April 28, 2022, from https://www.theguardian.com/world/2021/mar/15/spain-to-launch-trial-of-four-day-working-week

Kaul, A., Kloßner, S., Pfeifer, G., & Schieler, M. (2015). Synthetic Control Methods: Never Use All Pre-Intervention Outcomes Together With Covariates.

Keynes, J. (2016). *Essays in persuasion* [Original work published 1919-1931]. Springer.

Kieffer, A. (2011). *Niveau d'éducation : expérience d'une collaboration entre chercheurs et statisticiens [Level of education: experiences of collaboration between researchers and statisticians]* (tech. rep.). CNRS, Centre Maurice-Halbwachs (CMH).

Kleven, H., Landais, C., & Søgaard, J. E. (2019). Children and Gender Inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*, *11*(4), 181–209. https://doi.org/10.1257/app.20180010

Kluve, J., Card, D., Fertig, M., Góra, M., Jacobi, L., Jensen, P., Leetmaa, R., Nima, L., Patacchini, E., Schaffner, S., Schmidt, C. M., van der Klaauw, B., & Weber, A. (2007). *Active Labor Market Policies in Europe*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-48558-2

Kramarz, F., & Michaud, M.-L. (2010). The shape of hiring and separation costs in France. *Labour Economics*, *17*(1), 27–37. https://doi.org/10.1016/j.labeco.2009.07.005

Mailand, M. (2005). *The Danish flexicurity model: the role of the collective bargaining system* [Compiled for the Danish Ministry of Employment]. FAOS, Forskningscenter for Arbejdsmarkeds- og Organisationsstudier.

Marier, P. (2005). Where Did the Bureaucrats Go? Role and Influence of the Public Bureaucracy in the Swedish and French Pension Reform Debate. *Governance*, *18*(4), 521–544. https://doi.org/10.1111/j.1468-0491.2005.00290.x

Mehta, S. (2022). Tech company Bolt is permanently embracing a 4-day workweek. Retrieved April 28, 2022, from https://www.fastcompany.com/90710084/tech-company-bolt-is-permanently-embracing-a-4-day-workweek

Ministère de l'Education Nationale de la Jeunesse et des Sports [Ministry of National Education, Youth and Sports], Le brevet d'études professionnelles (BEP) [The vocational studies certificate (BEP)]. (2021). Retrieved May 3, 2022, from https://www.education.gouv.fr/reussir-au-lycee/le-brevet-d-etudes-professionnelles-bep-7631

Ministère de l'Education Nationale de la Jeunesse et des Sports [Ministry of National Education, Youth and Sports], Le certificat d'aptitude professionnelle (CAP) [The certificate of professional aptitude (CAP)]. (2022). Retrieved May 3, 2022, from https://www.education.gouv.fr/reussir-au-lycee/le-certificat-d-aptitude-professionnelle-cap-9959

Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation [The Ministry of Higher Education, Research and Innovation], Nomenclature relative au niveau de diplôme [Nomenclature relative to the level of education]. (2021). Retrieved May 3, 2022, from https://www.enseignementsup-recherche.gouv.fr/fr/nomenclature-relative-au-niveau-de-diplome-45785

Ministère de la Transformation et de la Fonction publiques [Ministry of Transformation and Public Service] - Portail de la Fonction publique [The civil service portal]. (2021). Retrieved May 2, 2022, from https://www.fonction-publique.gouv.fr/fonction-publique-france

OECD. (2019). Gross domestic product (GDP), 2019 archive: GDP per head, US $, constant prices, constant PPPs, reference year 2010, 2019 archive [Type: dataset]. Retrieved February 25, 2022, from https://stats.oecd.org/Index.aspx?datasetcode=SNA_TABLE1_ARCHIVE

OECD. (2021a). Strictness of employment protection – individual and collective dismissals (regular contracts) [Type: dataset]. Retrieved February 25, 2022, from https://stats.oecd.org/Index.aspx?DataSetCode=EPL_OV

OECD. (2021b). Strictness of employment protection – temporary contracts [Type: dataset]. Retrieved February 25, 2022, from https://stats.oecd.org/Index.aspx?DataSetCode=EPL_T

OECD. (2021c). Trade Union Dataset [Type: dataset]. Retrieved February 25, 2022, from https://stats.oecd.org/Index.aspx?DataSetCode=TUD

OECD. (2022a). Investment (GFCF) (indicator) [Type: dataset]. https://doi.org/10.1787/b6793677-en

OECD. (2022b). Quarterly GDP (indicator) [Type: dataset]. https://doi.org/10.1787/b86d1fc8-en

OECD. (2022c). Unemployment rate [Type: dataset]. https://doi.org/10.1787/997c8750-en

Pak, M. (2013). Le travail à temps partiel. *Dares, Synthèses Stat*, (4), 13–14.

Peri, G., & Yasenov, V. (2017). The Labor Market Effects of a Refugee Wave: Synthetic Control Method Meets the Mariel Boatlift. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2940595

Pilat, D., Cimper, A., Olsen, K. B., & Webb, C. (2006). The changing nature of manufacturing in OECD economies [Publisher: OECD].

Rabaté, S., & Rochut, J. (2020). Employment and substitution effects of raising the statutory retirement age in France. *Journal of Pension Economics and Finance*, *19*(3), 293–308. https://doi.org/10.1017/S1474747218000392

Regeringskansliet. (2020). Korttidsarbete [Stort-Time Work] [[Government Offices of Sweden]]. Retrieved May 3, 2022, from https://regeringen.se/regeringens-politik/regeringens-arbete-med-coronapandemin/foretag/om-forslaget-korttidspermittering/

Rosen, S. (1976). A Theory of Life Earnings. *Journal of Political Economy*, *84*(4, Part 2), S45–S67. https://doi.org/10.1086/260532

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, *66*(5), 688–701. https://doi.org/10.1037/h0037350

Schettkat, R., & Yocarini, L. (2006). The shift to services employment: A review of the literature. *Structural Change and Economic Dynamics*, *17*(2), 127–147. https://doi.org/10.1016/j.strueco.2005.04.002

Skuterud, M. (2007). Identifying the Potential of Work-Sharing as a Job-Creation Strategy. *Journal of Labor Economics*, *25*(2), 265–287. https://doi.org/10.1086/511379

Søndergård, K. (2007). Strategier for flexicurity [Strategies for flexicurity], 81. https://faos.ku.dk/pdf/forskningsnotater/forskningsnotater_2007/fnotat79.pdf

SPF. (n.d.). Durée du travail et diminution de la durée de travail [Duration of work and the reduction of working time]. Retrieved May 11, 2022, from https://emploi.belgique.be/fr/themes/reglementation-du-travail/duree-du-travail-et-temps-de-repos/duree-du-travail-et-diminution

Spiegelaere, S. d., & Piasna, A. (2017). *The why and the how of working time reduction*. European Trade Union Institute.

Ulrich, V., & Zilberman, S. (2007). De plus en plus d'emplois à temps partiel au cours des vingt-cinq dernières années. *Premières synthèses*, (39.3).

Vanvitelli, L., & Pastore, F. (2018). Why is youth unemployment so high and different across countries? *IZA World of Labor*. https://doi.org/10.15185/izawol.420

Varejao, J. (2005). *Work-sharing revisited-lessons from a natural experiment* (tech. rep.). Mimeo, Universidade do Porto.

# Appendices

## A Variable Transformation for the SCM

### Low-Skilled Workers in the Industry and Services Sectors

To construct the outcome variables for low-skilled workers in the industry and services sectors, ILOSTAT's variable of employees by economic activity and occupation is used (ILOSTAT, 2022a). The variable consists of total number of employees in thousands by occupation and sector for each country in the donor pool, excluding the self-employed. The outcome variable is divided by the working age population (ILOSTAT, 2022g). Economic activity is coded according to the International Standard Industrial Classification of All Economic Activities (ISIC) into four broad categories – agriculture, industry, services and not elsewhere classified (ILOSTAT, 2022c). Industry in this classification refers to activities such as manufacturing, construction or mining. Services refers to activities such as trade or transportation. Occupation is coded according to the ten major occupation groups of the International Classification of Occupations (ISCO-88) which are sorted on skill level. This classification is seen in Table 8. Out of these ten categories, we choose four low-skilled occupation categories that we believe will be most affected by the policy, namely workers who tend to work in strict hour schedules.

**Table 8:** Relevant ISCO-88 major groups

| Group | Broad skill level | Included |
|---|---|---|
| 1. Legislators, senior officials and managers | 3-4 | NO |
| 2. Professionals | 3-4 | NO |
| 3. Technicians and associate professionals | 3-4 | NO |
| 4. Clerks | 2 | YES |
| 5. Service workers and shop and market sales workers | 2 | YES |
| 6. Skilled agricultural and fishery workers | 2 | NO |
| 7. Craft and related trades workers | 2 | NO |
| 8. Plant and machine operators and assemblers | 2 | YES |
| 9. Elementary occupations | 1 | YES |
| 10. Armed forces | NA | NO |

Source: International Labor Organization (ILO, 2012).

For each respective variable, the occupation categories clerks, service workers and shop and market sales workers, plant and machine operators and assemblers, as well as elementary occupations, are included respectively, as seen in Table 8.

Data for the donor pool on these two variables are available from 1992 and onwards, meaning we have six pre-treatment periods. We collect predictors for France and all donors of the outcome variable. Firstly, GDP per capita, in US dollars, constant 2010 prices and constant PPPs and GDP growth is compiled from the OECD (2019, 2022b), since this will capture general economic development and the business cycle, which in turn impacts the labor market. We also include the investment rate in the year preceding the policy, as we believe this will have explanatory power for employment especially in the industry sector (OECD, 2022a). Following Abrassart's (2015) analysis of the institutional determinants of low-educated workers' employment rates, we also include OECD indeces of employment protection and trade union density (OECD, 2021a, 2021c), as well as the share of low-skilled workers in the labor force (using groups 4-9 from Table 8) compiled from ILOSTAT (2022d). In addition, we collect Peri and Yasenov's (2017) alternative predictor for low-skilled labor market outcomes, the labor force participation rate, as well as the overall unemployment rate, both measurements compiled from ILOSTAT (2022e, 2022f).

### Low-Skilled Youth Employment

For the youth employment variable we use ILOSTAT's measurement of youth employment by occupation in thousands in the age band 15-29 (ILOSTAT, 2022h), divided by the working age population. As such, it does not contain youth in school or university, who are not classified as employed. The ILOSTAT measurement is split into the same ten major ISCO-88 occupation groups as for the industry and services variables, and we use the same low-skilled categories for the youth variable, which are seen in Table 8. These are clerks, service workers and shop and market sales workers, plant and machine operators and assemblers, and elementary occupations.

Data is available from 1992 and onwards, which is the same period as our first two outcome variables. We again collect data on the predictors GDP per capita, GDP growth, employment protection legislation and trade union density, and also include some youth-specific predictors. These are the youth labor force participation rate for individuals aged 15-29, the youth unemployment rate for individuals aged 15-29 as well as the share of 15-29 year olds in the labor force, all compiled from ILOSTAT (2022d, 2022i, 2022j). The reason for including the last predictor is that the labor market could have difficulties adjusting quickly to demographic developments. A new generation entering the labor market will face the labor supply of the generation before them. If there are major changes in generation sizes, it may therefore affect the youth employment.

### The Senior Labor Force

To study the senior labor force we use ILOSTAT's measurement labor force by sex and age in thousands in the age band 55-64 (2022d), divided by the working age population. These are the workers closest to retirement within the French pension system. The legal retirement age in France was lowered from 65 to 60 in 1983. In addition, as laid out by Rabaté and Rochut (2020), the minimum pension age in France became individual-specific in 2003, meaning that individuals with a long career duration could retire at age 56.

Similar to our other outcome variables, we want these to come from the same labor force survey, due to possible variation in sampling and measurement strategies across countries. Therefore, we set our first pre-treatment period to 1986, which is the first period for which data is available for all donors from the EU Labor Force Survey, meaning we have 12 pre-treatment periods. Collected predictors for this variable, apart from pre-treatment values are GDP per capita, GDP growth and unemployment rate. We also include the share of the labor force aged 55-64, compiled from ILOSTAT, to account for structural labor force changes. Ideally, we would like to use the unemployment rate for our studied age band, 55-64, as opposed to the general unemployment rate. However, this data was not available for Luxembourg, which is why we instead use the overall unemployment as a predictor.

### Prevalence of Temporary Contracts

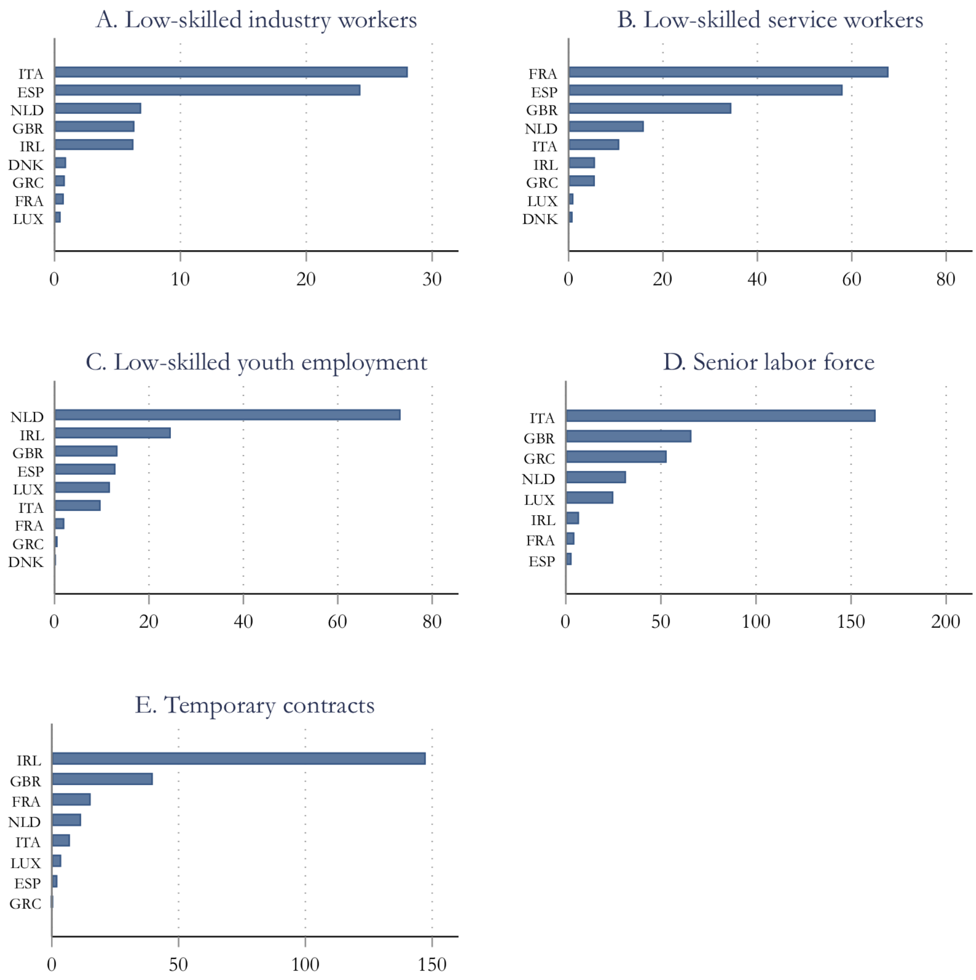
In order to study the policy's impact on the prevalence of temporary contracts, we use ILOSTAT's measurement of annual employees by temporary or permanent contract in thousands (ILOSTAT, 2022b). We look only at number of employees on temporary contracts, and divide the measurement by the total working age population. Data for this variable for the donor pool is available from 1987, meaning we have eleven pre-treatment periods. Luxembourg is missing data from 1995 on the number of employees with temporary contracts. We circumvent this by interpolating, and replace the 1995 value with the average of 1996 and 1994.

We collect the predictors GDP per capita, GDP growth, employment protection legislation for temporary contracts (OECD, 2021b), trade union density, the labor force participation rate, as well as the unemployment rate for this variable. In addition, we also include the share of the workforce with a temporary contract, which is identical to our outcome variable, except that it is divided by the labor force, as opposed to the working age population. This predictor could explain changes in the composition of contracts within the workforce, which may be related to structural changes in the working age population.

## B Extension for In-Space Placebo Tests

The in-space can be complemented with the calculation of p-vales. Figure 13 presents the ratios between the post-intervention and pre-intervention RMSPE for unit placebos, which are used to compute p-values. In Figure 13 Panel A, we see that France has the next to lowest ratio for low-skilled industry workers. This gives a p-value of $8/9 \approx 0.88$. In Panel B, the p-value corresponds to $1/9 \approx 0.11$. For panels C, D and E, the p-values are $7/9 \approx 0.77$, $7/8 \approx 0.88$ and $3/8 \approx 0.38$ respectively.

**Figure 13:** Ratio of pre and post-RMSPE




# C Coding of Skill by Education Level

Within the event study, we proxy skill level by educational attainment according to the classification shown in Table 9. Individuals with an education level lower than high school are coded as low-skilled. Medium-skilled individuals are those with an educational attainment equivalent to high school. Finally, the individuals with an educational level above high school are coded as high-skilled. The education levels are the ones reported in the FLFS.

**Table 9:** Skill level classification based on education

| Education level in FLFS | English Translation | Skill-level |
|---|---|---|
| Troisième cycle universitaire | Doctoral studies | High |
| Grande école, diplôme d'ingénieur | Certain business school diploma or engineer diploma | High |
| Maîtrise | Master's degree | High |
| Ecoles niveau licence et au-delà | Bachelor level school or beyond | High |
| Premier cycle universitaire, licence | Bachelor's degree | High |
| Autre diplôme (niveau bac+2) | Other diploma, post-high school degree of two years | High |
| BTS[1], DUT[2] | Technical university diploma | High |
| Paramédical ou social avec baccalauréat général | Paramedic with general high school diploma | High |
| Paramédical ou social sans baccalauréat général | Paramedic without general high school diploma | High |
| Baccalauréat général et diplôme technique secondaire | High school degree or secondary technical diploma | Medium |
| Baccalauréat général seul | High school diploma | Medium |
| Baccalauréat technologique, baccalauréat professionnel et brevet professionnel | Technological or professional high school diploma | Medium |
| BEI[3], BEC[4], BEA[5] | Certificate of industrial, commercial and agricultural education | Low |
| Brevet de technicien, brevet professionnel | Technological or professional diploma | Low |
| CAP[6], BEP[7] et BEPC [8] | Certificate of professional aptitude, certificate of professional studies and certificate of middle school studies | Low |
| CAP, BEP seul | Certificate of professional aptitude, certificate of professional studies | Low |
| BEPC seul, Brevet des collèges | Certificate of middle school studies | Low |
| Certificat d'études primaires | Certificate of primary school studies | Low |
| Aucun diplôme | No diploma | Low |

Abbreviations: [1] Brevet de technicien supérieur [2] Diplôme universitaire de technologie [3] Brevet d'enseignement industriel [4] Brevet d'enseignement commercial [5] Brevet d'enseignement agricole [6] Certificat d'aptitude professionnelle [7] Brevet d'études professionnelles [8] Brevet d'études du premier cycle

Sources: *Enquête Emploi* [French Labour Force Survey] 1993-2007. Producer: INSEE, Distributor: ADISP, Kieffer (2011), Ministère de l'Education (2021), Ministère de l'Education (2022), Ministère de l'Enseignement supérieur (2021), Campus France (n.d.)

# D Event Study Estimates

**Table 10:** Event study estimates

| | Low-skilled | Medium-skilled | High-skilled |
|---|---|---|---|
| 1993 | 0.0102 | -0.00589 | -0.00778* |
| | (0.00564) | (0.00572) | (0.00367) |
| 1994 | -0.00765 | -0.0179*** | -0.0154*** |
| | (0.00439) | (0.00527) | (0.00295) |
| 1995 | -0.00293 | -0.0117* | -0.00652* |
| | (0.00413) | (0.00587) | (0.00250) |
| 1996 | -0.000135 | -0.00685 | -0.00462 |
| | (0.00415) | (0.00526) | (0.00258) |
| 1997 | -0.00339 | -0.0111* | -0.00737** |
| | (0.00383) | (0.00466) | (0.00225) |
| 1999 | 0.00369 | 0.00213 | 0.00399 |
| | (0.00317) | (0.00421) | (0.00217) |
| 2000 | 0.0219*** | 0.0212*** | 0.0182*** |
| | (0.00381) | (0.00498) | (0.00247) |
| 2001 | 0.0295*** | 0.0354*** | 0.0280*** |
| | (0.00395) | (0.00590) | (0.00324) |
| 2002 | 0.0239*** | 0.0319*** | 0.0310*** |
| | (0.00445) | (0.00606) | (0.00315) |
| 2003 | 0.0116 | 0.0161* | 0.0350*** |
| | (0.00591) | (0.00737) | (0.00430) |
| 2004 | 0.00491 | 0.0113 | 0.0365*** |
| | (0.00618) | (0.00803) | (0.00382) |
| 2005 | 0.00293 | 0.00476 | 0.0349*** |
| | (0.00530) | (0.00772) | (0.00447) |
| 2006 | -0.000196 | 0.00471 | 0.0269*** |
| | (0.00561) | (0.00682) | (0.00406) |
| 2007 | 0.00800 | 0.0123 | 0.0274*** |
| | (0.00456) | (0.00671) | (0.00442) |
| Children | 0.0747*** | 0.0906*** | 0.0658*** |
| | (0.00656) | (0.00868) | (0.00406) |
| Married | 0.0791*** | 0.117*** | 0.121*** |

Continued on next page

Table 10 – continued from previous page

| | Low-Skilled | Medium-Skilled | High-Skilled |
|---|---|---|---|
| | (0.00468) | (0.00567) | (0.00504) |
| Married × Children | 0.0871*** | 0.126*** | 0.143*** |
| | (0.00428) | (0.00706) | (0.00736) |
| Female | -0.00952* | -0.0363*** | -0.0955*** |
| | (0.00385) | (0.00539) | (0.00459) |
| Female × Children | -0.0258*** | -0.125*** | -0.315*** |
| | (0.00692) | (0.0106) | (0.00758) |
| Female × Married | -0.0540*** | -0.0658*** | -0.106*** |
| | (0.00655) | (0.00556) | (0.00536) |
| Female × Married × Children | -0.123*** | -0.210*** | -0.360*** |
| | (0.00776) | (0.00734) | (0.00685) |
| Constant | 0.103*** | 0.107*** | 0.249*** |
| | (0.0237) | (0.00529) | (0.00846) |
| Age | Yes | Yes | Yes |
| Department | Yes | Yes | Yes |
| Observations | 237,882 | 170,253 | 798,428 |

Cluster-robust standard errors in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Dependent variable is employment probability.

Source: *Enquête Emploi* [French Labour Force Survey] 1993-2007. Producer: INSEE. Distributor: ADISP.