Wage and Employment Effects of Wage Subsidies*

Maxime Gravoueille

November 2022

[LINK TO THE LATEST VERSION]

Abstract

This paper estimates the wage and employment effects of wage subsidies using a large 2015 national-level reform in France that provides additional financial support to poor working households. While the aim of this policy is to promote work, it can incidentally reduce wages in response to an increase in the labor supply. Using administrative data and a shift-share IV design leveraging variation in the exposure to the reform based on the socio-economic composition of the local working-age population, I show that labor markets exposed to an increase in wage subsidies experience an increase in the growth rate of the number of hours worked and a decrease in the growth rate of the average hourly wage. I find no significant effect on pre-tax labor earnings growth at the local labor market level, as the wage and employment effects are of similar magnitude. These effects suggest a pass-through of wage subsidies to wages equal to 37% on average.

JEL classification codes: H22, H23, H24, H31, J22, J31.

Keywords: Wage subsidies, wage effect, employment effect, tax incidence, shift-share IV.

*Maxime Gravoueille: Paris School of Economics (maximegravoueille@gmail.com). I am indebted to Philippe Aghion, Victor Augias, Antoine Bozio, Sir Richard Blundell, Léo Czajka, Simon Gleyze, Maxence Gouleau, Xavier Jaravel, Attila Lindner, David Margolis, Antoine Mayerowitz, Daniel Prinz, Sarah Schneider-Strawczynski and Stefanie Stantcheva for their advice and comments. I thank numerous seminar participants at the Institute for Fiscal Studies, Paris School of Economics, Collège de France and INSEAD.

1 Introduction

In developed countries, government transfers in the form of wage subsidies are widely implemented anti-poverty programs. They directly provide financial support to poor working families and give them additional incentives to work. A prime example of their popularity is the Earned Income Tax Credit in the United States, which has experienced multiple federal and state expansions since its implementation in 1975. In 2021, it represented approximately $60 billion in expenditure for 25 million workers. Despite a large body of literature evaluating the employment effects of these transfers for workers, there is little evidence on the labor market equilibrium effects of these programs. Indeed, they can produce opposing and unintended effects: by making work more profitable and increasing the number of hours worked in the economy, employers can reduce the hourly wage growth. Ultimately, wage subsidies may not fully benefit workers, as employers are able to capture part of it.

This paper challenges the idea that the incidence of wage subsidies falls entirely on workers by departing from the conventional no wage effect assumption. I present novel causal estimates of the effect of wage subsidies on wage and employment using a shift-share IV design that exploits variation in the exposure to a national reform in wage subsidies in France, based on the socio-economic compositions of the local working-age population.

Most of the recent microeconomic literature focuses on labor supply, namely workers, casting aside equilibrium effects on the labor market. It is equivalent to the implicit assumption that the hourly wage rate is fixed or that labor demand is completely elastic. In this context, causal evidence on the extensive margin (the probability of participating in the labor market) and intensive margin (the number of hours worked conditional on employment) are estimated by comparing individuals from the same labor market, some experiencing an increase in wage subsidies and others not. While this estimation strategy is reasonable to examine the employment effect of wage subsidies, it is unable to identify the wage effect. Because they participate in the same labor market, an increase in employment from the treated group also reduces the hourly wage rate in the control group. By extension, it is also unhelpful to estimate the overall effect on labor earnings. Taking into account this channel has substantial implications in accessing who really benefits from these policies between employers and workers.

Disentangling the wage and employment channels of wage subsidies in the presence of labor market equilibrium effects is challenging. One needs a research design that allows for both labor supply and labor demand responses to the policy. This paper sheds a new light on this question by using a novel identification strategy and a unique reform of wage subsidies in France in 2015. In France, wage subsidies are set at the national level and are paid directly to workers. Although they depend on individual and household characteristics (such as labor earnings, marital status or the number of children), they do not depend on specific local labor market characteristics. Conditional on having similar characteristics, an individual living in the north of France receives the same amount of wage subsidy as someone living in the south of France. The reform merged two wage subsidy programs, creating a shock that differs along individual and household characteristics. As a result, some labor markets were more exposed to the reform on average. Intuitively, the identification strategy compares two labor markets facing the same reform, but for which the change in wage subsidies will be different because of initial differences in these socio-economic characteristics. I take advantage of a high-quality dataset on a representative sample of French individuals, combining administrative data matching employer-employee information with income tax returns and social agencies claims. In particular, sampled individuals are followed over time, allowing me to precisely track their labor market outcomes. This unique combination of nationwide reform and panel data on individuals and households is ideal for studying the wage and employment effects of wage subsidies.

In the first part of this paper, I outline a simple conceptual framework that provides the essence of the rest of the empirical analysis. I begin by building a competitive labor market model, based on Rothstein (2010), allowing for equilibrium effects in the presence of wage subsidies. The model features labor supply responses at the intensive margin and extensive margin. Each labor market contains many agents that differ in socio-economic characteristics used for computing wage subsidies in a non-linear schedule (e.g., household income, marital status, or having children). This simple model highlights how a labor market level analysis provides sufficient statistics to assess the wage and employment effects of wage subsidies. It especially shows that an increase in labor supply at the labor market level is partly offset by a decrease in the hourly wage rate if labor demand is less than perfectly elastic.

Then, I develop a quasi-experimental research design to identify the wage and employment effects of wage subsidies. A key contribution of this study is to show that causal estimates for both channels can be recovered via labor market level regression using a shift-share instrumental variable design. This identification stems from two factors. First, eligibility requirements and wage subsidies are set at the national level. It depends mainly on individual characteristics (such as labor earnings) and household characteristics (such as income, marital status, and number of dependents). Importantly, it does not depend on which labor market people are in. For example, wage subsidies are not higher in depressed areas than in prosperous ones. Second, individual and household characteristics are heterogeneously distributed across labor markets. The combination of these two features makes some labor markets more exposed than others to a change in wage subsidies. I construct two exposure measures that are relevant for my analysis: the hour-weighted change in the marginal tax rate and the hour-weighted change in the average tax rate due to the reform, at the labor market level.

The validity of this research design relies on the quasi-random assignment of shocks (Borusyak, Hull, and Jaravel (2022)). Intuitively, this means that a change in wage subsidies should not have been chosen strategically, based on labor market trends, or in a way that correlates with such trends. This assumption naturally holds in my design, as the wage subsidy schedule is set at the national level and is therefore not directly indexed to labor market characteristics. A threat to the analysis is the reverse causality between wage subsidies and labor earnings. I build on the simulated instruments literature and compute the tax rates under the assumption of no behavioral responses to the reform. To validate my research design and show that changes in wage subsidies are not correlated with other unobservable labor market features also affecting wages and employment, I construct falsification tests based on a regression of past outcomes on current shocks.

In the final part of the paper, I use my quasi-experimental research design to evaluate the effect of the 2015 French reform on wage subsidies. I find sizeable wage and employment responses with respect to the average tax rate. On the contrary, I do not find any significant responses with respect to the marginal tax rate. I find that a decrease in 10% of the average tax rate increases by 3.3% the number of hours worked and decreases by 3.7% the hourly wage rate, relative to the situation without any change in wage subsidies. Overall, I find no significant effect on pre-tax labor earnings growth at the local labor market level, as the wage effect and the employment effect are of similar magnitude. These effects suggest a pass-through of wage subsidies to wages equal to 37% on average.

Related literature This paper builds on and contributes to several strands of literature. First, there is a vast body of literature on the effects of wage subsidies on the labor market. Most of the microeconomic literature has focused on the effect of these programs on the labor supply of individuals (surveyed by V. Joseph Hotz (2003), Eissa and H. W. Hoynes (2006), Meyer (2010), Nichols and Rothstein (2015), Brewer and H. Hoynes (2019), H. Hoynes (2019)). See, in particular, Eissa and Liebman (1996), Meyer and Rosenbaum (2001), Grogger (2003), V Joseph Hotz and Scholz (2006), Gelber and Mitchell (2012), Gelber and Mitchell (2012), Kleven (2019), Bastian (2020), Whitmore Schanzenbach and Strain (2021) for the extensive margin responses, and Bollinger, Gonzalez, and Ziliak (2009), Rothstein (2010), Chetty, Friedman, and Saez (2013) for the intensive margin responses. This literature considers the equilibrium effects to be negligible, implicitly assuming fixed wage rates. This paper departs from this assumption by allowing for equilibrium effects and empirically investigating both wage and employment channels. Although some studies have moved away from this canonical setting (Rothstein (2010), Leigh (2010), Azmat (2019)), they remain inconclusive because of the difficulty in identifying such effects using credible research designs and the unavailability of administrative and/or panel data. The closest empirical study to mine is that of Leigh (2010). He finds that an increase of 10% in the EITC decreases the hourly wage rate by 5 % for high school dropouts. However, this result is not compatible with reasonable incidence parameters for wage subsidies. Relative to this literature, the key contribution of this paper is the development of a quasi-experimental research design that can credibly estimate wage and employment effects separately at the labor market level. To do so, I combine recent advances in Bartik/shift-share IV design (Goldsmith-Pinkham, Sorkin, and Swift (2020), Borusyak, Hull, and Jaravel (2022)), panel data from high-quality French administrative sources, and a large reform in wage subsidies that occurred in 2015 in France.

Second, this paper contributes to the scarce literature on the incidence of wage subsidies. By estimating wage and employment effects, I can provide bounds on the share of subsidy going to employers and workers. Rothstein (2010) investigates incidence of the EITC using a competitive partial equilibrium model for the labor market and calibrations along a range of plausible values for the labor supply and demand elasticities. Focusing on the labor market for women, he finds a fall in total earnings primarily due to a decrease in the wage rate, leading to only 70% of a dollar increase in the EITC going to low-skilled mothers. Azmat (2019) shows, in a specific context of the tax credit being paid through employers, that firms cut by 7% the wage of claimants relative to nonclaiments, which is suggestive of a negative spillover between the two groups. My results are in line with this literature. I show that employers are able to capture a sizeable part of wage subsidies, up to 37% on average, through reduced wage growth.

Finally, this paper contributes to the growing literature estimating the macroeconomic effects of taxes and transfers, including wages subsidies (Froemel and Gottlieb (2021), Ortigueira and Siassi (2022), Ferriere et al. (2021)). I show how my estimation strategy is compatible with a model aggregating individual and household responses at the relevant market level. I also provide a set of reduced form elasticities at the market level, separately for the wage and employment channels, that can be used as targeted moments.

The remainder of this paper is organized as follows. Section 2 presents the French institutional background, with a particular focus on the 2015 reform of wage subsidies. Section 3 presents the conceptual framework and the quasi-experimental research design based on shift-share IV. Section 4 describes the data, variables construction and provides summary statistics. Section 5 reports causal estimates at the labor market level. Finally, section 6 concludes.

2 Institutional Background

2.1 Wage subsidies

Wages subsidies are widely used anti-poverty and social inclusions programs. They target low-earnings individuals or households through a government transfer conditional on working. The most famous program is maybe the US *Earned Income Tax Credit* (EITC henceforth) that was implemented in 1975 and has seen several expansions since (in 2021, it represented approximately \$60B for 25M workers).

France has had a similar program in France since 2001. It has an increasing phase-in and a decreasing phase-out as a function of earnings. The main goal is to promote and encourage work, not only by increasing financial incentives in-work, but also by minimizing the loss in welfare benefits (transfers not conditioned on working) induced by returning to work. The system has remained fairly stable up to a salient reform in 2015, explained in more detail below. The following formula summarizes the benefit level individuals are entitled to given their earnings and their socio-economic characteristics:

$$B_{i,t} = b_t(e_{i,t}, \boldsymbol{E}_{i,t}, \boldsymbol{\Omega}_{i,t}, \boldsymbol{\phi}_t) \tag{1}$$

where $B_{i,t}$ stands for the amount of benefit individual i is eligible for in year t, that depends on her individual labor earnings e, her other household revenues $\boldsymbol{E}$, her household characteristics $\boldsymbol{\Omega}$ (such as the number of dependents or marital status for example) and some institutional parameters $\boldsymbol{\phi}$ (such as eligibility thresholds and the benefit schedule).

Before the reform Prior the 2015 reform, two wage subsidies programs were available. First, a tax credit[1] that was refundable through the income tax system on a yearly basis. Individuals received the tax credit with a one year delay relative to the income year and all parameters were computed by the tax administration using tax returns. Note that contrary to the EITC, the take-up rate was nearly complete as individuals were automatically given the tax credit if eligible by the tax administration. Secondly, an in-work benefit[2] was available for individuals through social programs on a monthly basis and targeted a lower part of the earnings distribution compared to the tax credit program. The take-up rate of this complement was lower (close to 32%, see Bourguignon (2011)). Finally, in-work benefits that had been received were deducted from the income tax credit individuals were eligible to, creating no overlap between the two systems.

After the reform Following the 2015 reform, the income tax credit and the in-work benefit have been merged into a unique in-work benefit. The reform's goal was to simplify the system by having a unique and accessible in-work benefit, available through social programs on a monthly basis[3]. Contrary to the main tax credit available before 2015, it is not automatically distributed to eligible individuals. The take-up significantly decreases but still remained at a high level (73% in 2016).

[1] *Prime Pour L'emploi*.

[2] *Revenu de Solidarité Activité d'Activité*.

[3] While it is paid on a monthly basis, parameters used for the computation of the benefit are updated on a quarterly basis. The data is only a yearly basis, such that we cannot fully capture adjustments between different quarters.

2.2 French Labor Market

The French labor market is characterized by high participation rates of its population since the early 1990s, both for men and women. Figure 18 shows the labor force participation rates of prime-age workers, by sex and over time. Men (panel (a)) have participation rates between 90%-95%, with similar magnitudes between France, the United Kingdom and the United States. Women (panel (b)) have seen a large increase in labor force participation in France and in the United Kingdom, to rise to 80%-85% in 2019. However, the United States have experienced a stagnation in women participation at around 75%.

The difference in labor market outcome between sex is even more pronounced for part-time employment: Figure 19 shows the share of employed workers in part-time, by sex and over time. Men (panel (a)) are mostly working full-time, while women (panel (b)) are more likely to work part-time. The difference between countries in the share of women in part-time is also striking, with only 10% for the United States compared to 20% for France and 30% for the United Kingdom in 2019.

A potential explanation for these differences is childcare costs. Figure 20 shows the childcare costs for two children in percentage of the net household income, separately for single parents (panel (a)) and couples (panel (b)). It is low for France (around 10% in both cases) but very high for the United States (around 40% for singles and 20% for married). The United Kingdom seems to be an intermediate case.

3 Conceptual Framework and Quasi-Experimental Research Design

This section introduces the empirical methodology used in the remainder of this paper. I start with a stylized model, in the spirit of Rothstein (2010), to provide insights into wage effects emerging from equilibrium in the labor market. Although the model abstracts from many potentially important labor market features, it provides useful reduced-form formulas to understand how both channels can be estimated from labor market level regressions. Finally, I demonstrate how a quasi-experimental shift-share instrumental variable design based on Borusyak, Hull, and Jaravel (2022) can be used to claim causal estimates.

3.1 A Simple Model of Wage Subsidies with Equilibrium Effects

3.1.1 Setup

Consider a static economy with individuals belonging to the labor market $m = 1, \ldots, M$. For example, it is a combination of the locations and bins of hourly wage rates. Production uses labor as the only input. The output and labor markets are perfectly competitive.

Production A representative firm produces with constant elasticity of substitution between the m labor market groups:

$$L = \left(\sum_m \beta_m L_m^{\frac{1+\varepsilon^d}{\varepsilon^d}} \right)^{\frac{\varepsilon^d}{1+\varepsilon^d}}$$

and total cost is $\sum_m w_m L_m = C$. Labor demand for the m group is:

$$L_m = \underbrace{\frac{C}{\sum_m w_m^{1-\varepsilon^d} \beta_m^{-\varepsilon^d}}}_{\varphi(\boldsymbol{w})} \cdot \beta_m^{-\varepsilon^d} \cdot w_m^{\varepsilon^d}$$

The change in wage at the group level is then:

$$\frac{\mathrm{d}w_m}{w_m} = \frac{1}{\varepsilon^d} \cdot \frac{\mathrm{d}\varphi(\boldsymbol{w})}{\varphi(\boldsymbol{w})} + \frac{1}{\varepsilon^d} \cdot \frac{\mathrm{d}L_m}{L_m} \tag{2}$$

Household preferences Within labor market m, individuals belong to a socio-economic group $g = 1, \ldots, G$ that defines their level of taxes and benefits. It is a function, for example, of their own labor earnings, total household income, marital status, or the number of children. These groups can be considered measures of the treatment intensity of subsidy reform. Individuals participate in the labor market by paying an entry cost $q_{m,g}$. Then, they choose the number of hours worked $h_{m,g}$, conditional on participation. The representative individual maximizes utility $U(c_{m,g}, h_{m,g}) = v(c_{m,g}, h_{m,g}) - q_{m,g} \cdot 1\left[h_{m,g} > 0\right]$ under budget constraint $c_{m,g} = e_{m,g} - T_{m,g}(z_{m,g})$. This depends on disposable income $c_{m,g}$, labor earnings $e_{m,g} = w_m h_{m,g}$ and the tax and benefit paid $T_{m,g}(z_{m,g})$. Because individuals belong to the same labor market, they get the same hourly wage w_m

The optimal number of hours is given by $h_{m,g} = h(w_m(1 - \text{MTR}_{m,g}))$ and the optimal participation rate is $P_{m,g} = P(w_m h_{m,g}(1 - \text{ATR}_{m,g}))$. The first condition captures the intensive margin, which depends positively on the wage rate and negatively on the marginal tax rate. The second condition captures the extensive margin, which depends positively on the wage rate and the number of hours worked but negatively on the average tax rate (capturing the difference in taxes and subsidies between participation and non-participation situations).

Labour supply A treatment group in a given labor market is populated by $N_{m,g}$ potential workers. Individuals can adjust their labor supply along two margins. Thus, the labor supply for group $m \times g$ is given by $L_{m,g} = N_{m,g} P_{m,g} h_{m,g}$ and the labor supply in the labor market m by $L_m = \sum_g L_{m,g}$. The growth rate in labor supply at the market level is given by:

$$\frac{\mathrm{d}L_m}{L_m} = \sum_g \frac{L_{m,g}}{L_m} \frac{\mathrm{d}P_{m,g}}{P_{m,g}} + \sum_g \frac{L_{m,g}}{L_m} \frac{\mathrm{d}h_{m,g}}{h_{m,g}} \tag{3}$$

with $\frac{\mathrm{d}h_{m,g}}{h_{m,g}} = \varepsilon^c \left[\frac{\mathrm{d}(1 - \text{MTR}_{m,g})}{1 - \text{MTR}_{m,g}} + \frac{\mathrm{d}w_m}{w_m} \right], \frac{\mathrm{d}P_{m,g}}{P_{m,g}} = \varepsilon^p \left[\frac{\mathrm{d}w_m}{w_m} + \frac{\mathrm{d}h_{m,g}}{h_{m,g}} + \frac{\mathrm{d}(1 - \text{ATR}_{m,g})}{1 - \text{ATR}_{m,g}} \right]$

For the sake of simplicity, I assume homogeneous compensated elasticity of labor supply ε^c and participation elasticity ε^p across groups. Note that, without the wage effect ($\mathrm{d}w_m / w_m = 0$), we have the classic employment effect of a change in taxes asssuming no equilibrium effects.

3.1.2 Wage and employment effects

Starting with the wage effect and combining the labor supply and the wage rate at the market level:

$$\frac{\mathrm{d}w_m}{w_m} = \frac{1}{\chi} \cdot \frac{\mathrm{d}\varphi(\boldsymbol{w})}{\varphi(\boldsymbol{w})} + \frac{\varepsilon^c(1+\varepsilon^p)}{\chi} \sum_g \frac{L_{m,g}}{L_m} \frac{\mathrm{d}(1-\mathrm{MTR}_{m,g})}{1-\mathrm{MTR}_{m,g}} + \frac{\varepsilon^p}{\chi} \sum_g \frac{L_{m,g}}{L_m} \frac{\mathrm{d}(1-\mathrm{ATR}_{m,g})}{1-\mathrm{ATR}_{m,g}}$$

with $\chi = \varepsilon^d - \varepsilon^c - \varepsilon^p - \varepsilon^c\varepsilon^p$. Finally we have the labor supply side:

$$\begin{aligned}\frac{\mathrm{d}L_m}{L_m} = \frac{\varepsilon^d - \chi}{\chi} \cdot \frac{\mathrm{d}\varphi(\boldsymbol{w})}{\varphi(\boldsymbol{w})} + \left[\frac{\varepsilon^c(1+\varepsilon^p)\varepsilon^d}{\chi}\right] \sum_g \frac{L_{m,g}}{L_m} \frac{\mathrm{d}(1-\mathrm{MTR}_{m,g})}{1-\mathrm{MTR}_{m,g}} \\ + \left[\frac{\varepsilon^p\varepsilon^d}{\chi}\right] \sum_g \frac{L_{m,g}}{L_m} \frac{\mathrm{d}(1-\mathrm{ATR}_{m,g})}{1-\mathrm{ATR}_{m,g}}\end{aligned}$$

These two equations are interesting for two reasons. First, it shows how the effect of a change in wage subsidies on employment and wage rate growth can be directly estimated using labor market level regressions. The total response for each of them depends on the sum of the growth rate of the net-of-tax rate (one minus the marginal tax rate) for each treatment group weighted by the initial number of hours in these groups, and on the sum of the growth rate of the participation tax rate (one minus the average tax rate) for each treatment group weighted by the initial number of hours in these groups. In subsection 3.2, I show how this analysis is similar to a shift-share research design when the wage subsidy schedule is set at the national level. The equivalence follows when $(1-\mathrm{MTR}_{m,g}) = (1-\mathrm{MTR}_g)$ and $(1-\mathrm{ATR}_{m,g}) = (1-\mathrm{ATR}_g), \forall m$.

Second, the magnitude of the response to these wage subsidies shocks depends on the magnitude of the demand elasticity ε^d, the participation elasticity ε^p and the compensated elasticity of labor supply ε^c. Note that when $\varepsilon^d \to -\infty$, the labor supply is completely elastic such that there is no wage effect. Labor market level regressions only estimate a mix of these elasticities. A change in the net-of-tax rate or the participation tax rate only has an effect on the total number of hours worked in labor market m.

3.2 Quasi-Experimental Research Design

Building on the insights from the simple model in subsection 3.1, I show that the differences in treatment intensity between labor markets arising from a nationwide reform in wage subsidies leverage two components. First, differences in initial exposure to the reform are determined because of the heterogeneous socio-economic composition across labor markets. Second, exogenous shocks to wage subsidies are set at the national level.

3.2.1 Setting

We consider panel data on individuals indexed by $i = 1, \ldots, N$ observed in two subsequent periods t and $t+h$. Individuals are split into distinct labor markets $m = 1, \ldots, M$. They are also in different treatment groups $g = 1, \ldots, G$ that determine the level of tax and benefit they are entitled to in year t. For example, it is a function of their own labor earnings, total labor earnings at the household level, marital status, and number of children. Importantly, treatment group g does not depend on the labor market components because the wage subsidy is a nationwide program.[4]

3.2.2 Treatment definition

I start by measuring the variation in taxes for different treatment groups within a labor market. The shock $\theta_{m,g,t}$ is a hours-weighted aggregation of the shocks at the individual level:

$$\theta_{m,g,t} = \sum_i \frac{h_{i,m,g,t}}{L_{m,g,t}} \times \theta_{i,m,g,t}$$

[4] I also report results using an alternative definition of the labor market level variables, with cross-sectional weights. The procedure is described in detail in subsection C.3.

where $h_{i,m,g,t}/L_{m,g,t}$ is the share of hours worked by individual i in labor market $\times$ treatment group total number of hours. For example, the net-of-tax rate shock is $\Delta \ln(1 - \text{MTR}_{m,g,t}) = \sum_i (h_{i,m,g,t}/L_{m,g,t}) \times \Delta \ln(1 - \text{MTR}_{i,m,g,t})$.

Assumption 1. *Treatment varies across the g groups but is not specific to labor market m, such that each $\theta_{m,g,t}$ is a noisy version of the latent shock $\theta_{g,t}$.*

$$\theta_{m,g,t} = \theta_{g,t} + \nu_{m,g,t}$$

where $\nu_{m,g,t}$ is an estimation error and $\theta_{g,t}$ is the shock at the national level. The wage subsidy schedule is set at the national level such that the shock for each initial treatment group g is not correlated with specific labor market characteristics by design.

Consider now the share of individuals in treatment bin g in labor market m, $S_{m,g,t} = L_{m,g,t}/L_{m,t}$. This share varies across labor markets because their socio-economic composition is different. Consequently, they have heterogeneous exposure to the same nationwide set of shocks. Formally, distinct labor market indexed by m have treatment for each year t:

$$X_{m,t} = \sum_g S_{m,g,t} \times \theta_{g,t} \tag{4}$$

This treatment variable measures the labor market exposure to the nationwide shock in taxes and benefits. For example, $\Delta \ln(1 - \text{MTR}_{m,t}) = \sum_g S_{m,g,t} \times \Delta \ln(1 - \text{MTR}_{g,t})$ is the hours-weighted growth rate in 1 minus the marginal tax rate for labor market m at initial year t. Similarly, $\Delta \ln(1 - \text{ATR}_{m,t}) = \sum_g S_{m,g,t} \times \Delta \ln(1 - \text{ATR}_{g,t})$ is the hours-weighted growth rate in 1 minus the average tax rate for labor market m at initial year t.

Instrument A large literature has emphasized issues arising when using Equation 4 directly as a treatment variable. Because individuals adjust their labor earnings (either through the number of hours worked or the wage rate) in response to an increase or decrease in tax liabilities, taxes and benefits are a direct function of labor earnings. In this case, reverse causality is a particular threat to identification.

To circumvent this problem, I build on the simulated instruments literature by computing the tax rates under the assumption of no behavioral responses to the reform. For the net-of-tax rate, the associated simulated rate is:

$$\Delta \ln\big(1 - \mathrm{MTR}_{i,t}^{sim}\big) = \ln\left[1 - \mathrm{MTR}_{i,t+h}(e_{i,t})\right] - \ln\left[1 - \mathrm{MTR}_{i,t}(e_{i,t})\right]$$

I use individuals' income and demographics in the initial year t to compute the the counterfactual marginal tax rate $\mathrm{MTR}_{i,t+h}(e_{i,t})$. Intuitively, this method construct the tax rates by switching off labor market responses from the reform such that $\Delta \ln\big(1 - \mathrm{MTR}_{i,t}^{sim}\big)$ is the mechanical change in net-of-tax rate. More precisely, it is exogenous to potential changes in the wage rate and number of hours worked.

Finally, I follow the same procedure as in the last paragraph to construct instruments at the labor market level:

$$Z_{m,t} = \sum_{g} S_{m,g,t} \times \theta_{g,t}^{sim}$$

3.2.3 Specification

I estimate a linear regression of outcome $Y_{m,t}$ on treatment $X_{m,t}$. In this paper, I consider two sets of outcomes: the change in the total number of hours worked $\Delta \ln(L_{m,t})$ and the change in average hourly wage rate $\Delta \ln(w_{m,t})$ between t and $t+h$. Specifically, I fit the following models via 2SLS:

$$Y_{m,t} = \beta \cdot \Delta \ln(1 - \mathrm{MTR}_{m,t}) + \gamma \cdot \Delta \ln(1 - \mathrm{ATR}_{m,t}) + \Lambda_{m,t} + \epsilon_{m,t}$$

The coefficient of interests are β and γ. They capture, respectively, the causal effect of a change in the labor market level marginal tax rate and average tax rate, conditional on a set of time-varying controls $\Lambda_{m,t}$. I instrument $\Delta \ln(1 - \mathrm{MTR}_{m,t})$ with $\Delta \ln\big(1 - \mathrm{MTR}_{m,t}^{sim}\big)$ and $\Delta \ln(1 - \mathrm{ATR}_{m,t})$ with $\Delta \ln\big(1 - \mathrm{ATR}_{m,t}^{sim}\big)$.

Following Borusyak, Hull, and Jaravel (2022), I define the equivalent shock-level

IV regression, weighted by $S_{g,t} = \sum_m S_{m,g,t}$. This is the baseline regression for the remaining of the paper.

$$\tilde{Y}_{g,t} = \beta \cdot \tilde{\Delta} \ln(1 - \text{MTR}_{g,t}) + \gamma \cdot \tilde{\Delta} \ln(1 - \text{ATR}_{g,t}) + \tilde{\Lambda}_{g,t} + \tilde{\epsilon}_{g,t} \tag{5}$$

where $\tilde{v}_{g,t} = (\sum_m S_{m,g,t} v_{m,t})/(\sum_m S_{m,g,t})$. This shock-level formulation will prove particularly useful to discuss the source of identification. Regressions include labor market and year fixed-effects, a set of year-specific fixed-effects related to the treatment group, as well as controls for socio-economic characteristics in the initial period.

3.2.4 Identification

My instruments are a combination of two features. First, exposure shares determined by the specific socio-economic composition within labor markets. Second, shocks in wage subsidies arising from a change in the national tax schedule. The literature on shift-share (or Bartik) instruments has emphasized two sources of identification are possible in this research design. Either the shares are considered exogenous (Goldsmith-Pinkham, Sorkin, and Swift (2020)) or the shocks are assumed to be quasi-randomly assigned (Borusyak, Hull, and Jaravel (2022)). I argue that my quasi-experimental research design falls in the second category.

Assumption 2. *(Conditional quasi-random shock assignment).*

$$\mathbb{E}[\tilde{Z}_{g,t} | \tilde{\epsilon}_{g,t}, \tilde{\Lambda}_{g,t}, S_{g,t}] = \tilde{\Lambda}'_{g,t} \cdot \mu$$

This condition implies that each shock has the same expected value, conditional on the shock-level unobservables $\tilde{\epsilon}_{g,t}$, average exposure $S_{g,t}$ and shock-level observables $\tilde{\Lambda}_{g,t}$. Intuitively, it means that change in wage subsidies should not have been chosen strategically, based on labor market trends, or in a way that is correlated with such trends. This assumption holds naturally in my design as the wage subsidy schedule is set at the national level and is therefore not directly indexed on labor market characteristics. For similar household characteristics (such as household income, marital status and number of children), an individual receives the same amount of wages subsidies in every labor market. Thus, the size of the shock (conditioned on a set of controls including shock-level fixed-effects) is unlikely to be correlated with unobservable labor market features affecting outcomes.

Assumption 3. *(Many uncorrelated shock residuals).*

$$\mathbb{E}[\sum_g \sum_t S_{g,t}^2] \to 0 \text{ and } \mathbb{C}\text{ov}(\tilde{Z}_{(g,t)}, \tilde{Z}_{(g,t)'} | \tilde{\epsilon}_{g,t}, \tilde{\Lambda}_{g,t}, S_{g,t}) = 0, \forall ((g,t),(g,t)') \text{ with } (g,t) \neq (g,t)'$$

The first part gives an intuitive measure of the effective sample size. It states that the shocks should not be concentrated in relatively few treatment groups. This is equivalent to say that the largest importance weight in the regression, $S_{g,t}$, becomes vanishingly small with the number of observations. The second part states that the shocks are mutually uncorrelated given the unobservables and $S_{g,t}$.

Finally, instruments should satisfy the relevance condition. It implies that instruments have sufficient statistical power.

Assumption 4. *(Relevance condition).*

$$\mathbb{E}[\tilde{X}_{g,t} \cdot \tilde{Z}_{g,t} | \tilde{\Lambda}_{g,t}, S_{g,t}] \neq 0$$

This condition is easily verifiable by checking the first-stage F statistic. More generally, $\tilde{Z}_{g,t}$ is a strong instruments as it relies only on initial period individual characteristics to predict what would have been the change shocks if individuals had not change their behavior.

Robustness tests So far, I have discussed why assumptions 2 and 3 are likely to hold in my setting. The key feature of my shift-share IV research design relies on being able to isolate labor supply shocks stemming from a plausibly exogenous reform in wage subsidies. Still, one might worry about year-specific unobservable shocks correlated with the labor demand, affecting in return labor market outcomes. It would introduce bias in my estimates. To confirm the validity of the empirical strategy, I implement several robustness checks.

First, I construct falsification tests by regressing lagged outcomes variables on current shocks. This placebo regression checks if past outcomes are correlated with current change in wage subsidies or not. Coefficients are expected to be non statistically significant if assumption 2 holds. Second, I directly test assumption 3 by computing the inverse Herfindal index $1/\sum_{g,t} S_{g,t}^2$. Borusyak, Hull, and Jaravel (2022) show that the shock-level estimation performs well even with an effective sample size of 20.

4 Data, Variables Construction and Summary Statistics

In this section, I briefly describe the data, how labor-market variables are constructed and provide some summary statistics. Additional details about the data and the construction of the variables are available in Appendix C.

4.1 Data

Data The main dataset is the *Echantillon Démographique Permanent* (EDP henceforth), a large individual-level panel following approximatively 4% of the French population at random.[5] It is a rich dataset linking several administrative datasets such as the census, matched employer-employee data[6], income tax returns and information from social agencies. The census provides detailed information on individuals' demographics: birth and death date, place of birth and death, sex and education among others. The employer-employee dataset contains detailed information on labor earnings, number of hours worked, type of contract, occupation and sector.[7] Finally, I

[5]Up to 2008, it sampled 1% of the French population: every individual born the first 4 days of October. Since 2008, the population has been extended to individuals born in the first 4 days of January, April, July and October.

[6]*DADS* (*Déclaration annuelle des données sociales*) database.

[7]Main activity and main firm refer to the situation where an individual has multiple employment spells in different firms in a given year. In this case, the type of contract, occupation, sector and number of workers are the ones from the longest spell (or the most paid one if two spells are of equivalent time).

have additional information on individual and household revenues from income tax returns (labor earnings, capital income, unemployment benefits, taxes and tax credits) and information on welfare benefits claimed by individuals through social agencies.

Sample I focus on the population of prime-age individuals aged 25-55 and living in the French metropolitan area between 2011 and 2017. This restriction ensures that I focus on a population that is plausibly in the labor force, and not in education or retired. Finally, I restrain my sample to individuals for which I can identify the tax household precisely. This embodies individuals who are single and living alone and couples who are living together and are married or in a civil union. Because the French redistribution system depends on household characteristics such as the marital status or the number of dependents, this restriction enables me to compute taxes and benefits with more precision.

4.2 Variable Construction

Labor market definition To identify wage and employment effects of wage subsidies, the quasi-experimental research design needs distinct labor markets. In practice, as noted by Hamermesh (1996) and Rothstein (2010), it requires labor markets within which workers are substitute. To do so, I adopt a simple decomposition based on the intersection of the geographical residence of individuals.

I start by selecting low-wage earners based on their hourly wage rate in the initial period. Specifically, I consider individuals with a pre-tax hourly wage below 14 euros per hour.[8] For non-working individuals, I construct their counterfactual group by predicting their hourly wage rate based on their socio-economic characteristics.

Then, I split individuals based on their geographical residence and segment my population by *departements*. It is the administrative division of France in 101 geographical area. I focus on 94 metropolitan areas due to data constraints.[9] This decomposition

[8]Based on https://www.insee.fr/fr/statistiques/1280904, the average hourly wage is 14.15 euros for someone working full-time.

[9]I exclude overseas departements and the two departements in Corse because my sample do not have enough observation for them.

maximizes the number of observations by location, while preserving variation in labor market characteristics to carry out the identification strategy.

Treatment groups Equation 1 describes the set of variables used to compute the level of wage subsidies in France. It mainly depends on an individual labor earnings, other household revenues, marital status and the number of dependents. I build treatment groups based on the interaction between bins of household labor earnings, a dummy for being in a couple and a dummy for having children.

I compute an equal split measure of household income. This is the sum of labor earnings at the household level, divided by 2 if people are in a couple. For non-working individuals, I add their predicted labor earnings based on their socio-economic characteristics to the household total labor earnings. Finally, I split this measure in bins of 1000 euros from 0 up to 30000 euros (the last bin contains all individuals above).

Tax rates Taxes and benefits for the current year are directly reported in the main dataset, but marginal and average tax rates are not. I use a publicly available tax simulator[10] to generate them. To consider the full effect of an increase of in-work subsidies, I compute the marginal and average tax rates for individual i in year t using the maximum amount she is entitled to. I simulate the marginal tax rates $\mathrm{MTR}_{i,t}$ and the average tax rate $\mathrm{ATR}_{i,t}$ using the tax and benefit system (see C.3). I compute the same set of tax variables considering the full tax system. I conduct the same analysis with it, since changes in labor market participation and earnings might also affect the eligibility and amount received from other programs.

Then, I compute the individual level log-growth rate for the marginal and average tax rates between year t and year $t+h$, $h>0$:

$$\Delta \ln(1-\mathrm{MTR}_{i,t}) = \ln\left[\frac{1-\mathrm{MTR}_{i,t+h}(e_{i,t+h}, \boldsymbol{E}_{i,t+h}, \boldsymbol{\Omega}_{i,t+h}, \boldsymbol{\phi}_{t+h})}{1-\mathrm{MTR}_{i,t}(e_{i,t}, \boldsymbol{E}_{i,t}, \boldsymbol{\Omega}_{i,t}, \boldsymbol{\phi}_{t})}\right]$$

[10] Openfisca is an open source taxes and benefits simulator with an extension specific for France (Openfisca France).

$$\Delta \ln(1 - \text{ATR}_{i,t}) = \ln \left[\frac{1 - \text{ATR}_{i,t+h}(e_{i,t+h}, \boldsymbol{E}_{i,t+h}, \boldsymbol{\Omega}_{i,t+h}, \boldsymbol{\phi}_{t+h})}{1 - \text{ATR}_{i,t}(e_{i,t}, \boldsymbol{E}_{i,t}, \boldsymbol{\Omega}_{i,t}, \boldsymbol{\phi}_t)} \right]$$

where, following notation from section 2, e is the individual labor earnings, $\boldsymbol{E}$ other household revenues, $\boldsymbol{\Omega}$ relevant household characteristics (such as the number of dependents or marital status) and some institutional parameters $\boldsymbol{\phi}$ (such as eligibility thresholds and the benefit schedule). I also construct the simulated instruments for the marginal tax rate $\Delta \ln\big(1 - \text{MTR}^{sim}_{i,t}\big)$ (respectively $\Delta \ln\big(1 - \text{ATR}^{sim}_{i,t}\big)$) at the individual level by replacing $\text{MTR}_{i,t+h}(e_{i,t+h}, \boldsymbol{E}_{i,t+h}, \boldsymbol{\Omega}_{i,t+h}, \boldsymbol{\phi}_{t+h})$ (resp. $\text{ATR}_{i,t+h}(e_{i,t+h}, \boldsymbol{E}_{i,t+h}, \boldsymbol{\Omega}_{i,t+h}, \boldsymbol{\phi}_{t+h})$) for $\text{MTR}_{i,t+h}(e_{i,t}, \boldsymbol{E}_{i,t}, \boldsymbol{\Omega}_{i,t}, \boldsymbol{\phi}_{t+h})$ (resp. $\text{ATR}_{i,t+h}(e_{i,t}, \boldsymbol{E}_{i,t}, \boldsymbol{\Omega}_{i,t}, \boldsymbol{\phi}_{t+h})$). For the simulated instruments, I inflate labor earnings and other household revenues by CPI evolution between t and $t + h$.

Finally, I construct the market level treatment variables and corresponding instruments using the methodology presented in subsection 3.2. Note that the individuals are split into labor markets and treatment groups based on their characteristics in the initial year t.

Labor market outcomes The last set of variables required for the analysis is the change in wage and employment at the labor market level. To compute them, I keep only individuals that are observed in both year t and $t + h$. I also classify individuals' labor market based on their characteristics in the initial period. By doing so, I avoid composition effects that might bias my measure.

I start with the definition of the change in employment by taking the log-growth rate of the total number of hours worked in a given labor market, $\Delta \ln(L_{m,t})$. I also construct a measure of wage rate by computing the hours-weighted wage rate. Formally, this measure for a given labor market in a given year is given by $w_{m,t} = (\sum_i w_{i,m,t} h_{i,m,t})/(\sum_i h_{i,m,t})$ with $w_{i,m,t}$ the hourly wage rate and $h_{i,m,t}$ the number of hours workerd by individual i. Then, I take the log-growth rate of the total number of hours worked in a given labor market, $\Delta \ln(w_{m,t})$.

4.3 Summary Statistics

Shocks due to the reform I start with a brief description of the effect of the reform on the change in participation and net-of-tax rates across treatment groups. Figure 1 plots the 2-year log-growth of the simulated participation tax rate for all the different treatment groups, for base year 2011 (pre-reform) and 2015 (reform). For the pre-reform period, the change in the participation tax rate is close to 0 and fairly constant for all groups. On the contrary, the variation in the participation tax rate for the reform period is substantial. These changes are particularly important for households in the bottom part of the labor earnings distribution, more exposed to wage subsidies. Figure 2 reproduces the same analysis for the 2-year log-growth of the simulated net-of-tax rate for all the different treatment groups. Again, changes are particularly important for households in the bottom part of the labor earnings distribution.[11]

A potential concern about this measure is that the change in the wage subsidy schedule do not necessarily translates in the total change in tax and benefits, implying a weak correlation between the two measures. To address it, I report the same plots as before in Figure 9 and Figure 10, but only for the wage subsidy (weighted by the initial share of the wage subsidy in total taxes and benefits). The shape of the distribution of shocks are similar to the baseline plots: differences in the change in participation and net-of-tax rates across treatment groups considering the full tax and benefit system is mainly driven by the change in the wage subsidy schedule. This is confirmed by Figure 3, that reports the correlation between the two measures. Results from an OLS regression (with intercept) reports a correlation coefficient of 0.80 for the net-of-tax rate and of 0.77 for the participation tax rate.

Shocks distribution My quasi-experimental research design and the relevance of my exposure measure depends on shocks being exogenous. More precisely, it depends on variation of log-growth of net-of-tax rates $\Delta \ln(1 - \mathrm{MTR}_{g,t})$ and participation rates $\Delta \ln(1 - \mathrm{ATR}_{g,t})$, as well as the distribution of the average exposure $s_{g,t}$ across treat-

[11] I report the same plots, but using an alternative definition of the labor market level variables, using cross-sectional weights. They can be found in Figure 11 and Figure 12. The results are similar.

ment groups. Following Borusyak, Hull, and Jaravel (2022), I access the validity of the shift-share IV design by summarizing shocks at the treatment group levels. This highlights the source of variation used for identification, while preserving its interpretation of labor market level causal effects.

Table 1 reports the summary statistics for the net-of-tax rate and participation rate. It also provides additional information about the effective sample size. All statistics are weighted by the average exposure $s_{g,t}$. I report two sets of statistics for the shocks. Panel (a) reports statistics using raw values and panel (b) reports statistics after residualization on year fixed-effects.[12]

Starting with the raw shocks (panel (a)), the distribution of the log-growth of net-of-tax rates (resp. participation rates) is left-skewed. It exhibits a mean of -0.023 (resp. -0.012), median of -0.007 (resp. -0.010), with a standard deviation of 0.055 (resp. 0.017). The 5$^{\text{th}}$ percentile is equal to -0.112 (resp. -0.044) and the 95$^{\text{th}}$ percentile is equal to 0.055 (0.014). This suggest a sizeable degree of variation: there is a 17 percentage point difference between the log-growth rates of the 5$^{\text{th}}$ and 95$^{\text{th}}$ percentile for the net-of-tax rate, and a 6 percentage point difference for the participation rate.

Then, I access the distribution of shock residuals from a regression on year fixed-effects (panel (b)). The distribution of shocks is more regular with a mean and median close to 0. It also confirms there is still sizeable variation, with standard variations close to those in panel (a).

Finally, I report summary statistics for the sample in panel (c). The effective sample size, as measured by the inverse Herfindal index, is 275. This confirms that the effective sample size is large. Consistent with this result, the largest share accounts forless than 1%. The number of unique treatment groups (combination of household income bins, a couple dummy and having children dummy) is 144 per year, with a total of 720 pooled observations over 5 years.

[12] I also provide summary statistics using an alternative definition of the labor market level variables, using cross-sectional weights. It can be found in Table 5. This is not my preferred measures because the weights are made to match the aggregate distribution of incomes at the national level (not only labor earnings) and are not designed to be longitudinal weights. Still, the results are similar.

5 Results

This section presents the main estimates using the shift-share IV research design presented in section 3. Specifically, I report estimates from equation (5), weighted by the average exposure of treatment group $s_{g,t}$. My outcomes of interest are the 2-year log-growth rates in the total number of hours worked $\Delta \ln$(hours)$= \Delta \ln(L_{m,t})$, in the average hourly wage $\Delta \ln$(wage)$= \Delta \ln(w_{m,t})$ and the sum of labor earnings $\Delta \ln$(labor earnings)$= \Delta \ln(L_{m,t}) + \Delta \ln(w_{m,t})$ in a given labor. In my main analysis, I report standard errors clustered at the household income-level.

5.1 Underlying Variation Behind the Shift-share IV

I start by providing graphical evidence for the variation underlying the shift-share IV.

First-stage correlation I start by investigating the relevance condition for both the net-of-tax rate instrument and the participation tax rate instrument, using a graphical depiction of the first-stage. Figure 4 plots the correlation between the instruments and observed tax shocks. Panel (a) and panel (b) report correlations for the net-of-tax rate and the participation rate, respectively. For each tax measure, I adjust for labor market level fixed-effects, year fixed-effects, base-year socio-economic controls and the other tax measure. Both observed tax shocks exhibit a salient positive relationship with respect to their respective instrument, indicative of strong first-stages.

Correlation between outcomes and instruments Starting with the employment side, Figure 5 depicts the relationship between the instrument (simulated tax shocks) and the employment response. Panel (a) (resp. panel (b)) first reports the response with respect to from a change in the net-of-tax rate (resp. participation rate). For each tax instrument, I adjust for labor market level fixed-effects, year fixed-effects, base-year socio-economic controls and the other tax instrument. Hours are not correlated with the net-of-tax rate, but they are strongly positively correlated with the participation rate.

Then, Figure 6 depicts the same analysis but for the wage response. Panel (a) (resp. panel (b)) reports the response with respect to from a change in the net-of-tax rate (resp. participation rate). Similarly, the hourly wage is not correlated with the net-of-tax rate, but it is strongly negatively correlated with the participation rate.

Finally, Figure 7 (resp. Figure 8) reports falsification tests for the employment response (resp. the wage response). It correlates past outcomes on current shocks, providing a intuitive placebo test. For both the net-of-tax rate and the participation rate, there is no correlation between instruments and outcomes.

5.2 Shift-share IV Results

Baseline results Table 2 reports the estimates for the wage and employment effects from the shift-share IV research design for my preferred specification. It uses simulated tax shocks as instruments for change in net-of-tax rates and participation rates. First-stage F-statistics are high, 44.1 for the net-of-tax rate and 224.9 for the participation rate, indicating that the shift-share instruments are strong.

Column (1) reports employment effects. The point estimate is 0.034 for the elasticity with respect to the net-of-tax rate and 0.325 for the elasticity with respect to the participation rate. Only the second coefficient is statistically significant. Column (2) reports similar elasticities for the hourly wage. The point estimate is 0.043 for the elasticity with respect to the net-of-tax rate and -0.367 for the elasticity with respect to the participation rate. Again, only the second coefficient is statistically significant. An increase in 10% of the participation rate increases by 3.25% the number of hours worked and decreases by 3.67% the average hourly wage, compared to the situation with a change in the wage subsidy schedule. Finally, column (3) reports the results for labor earnings, which is the sum of the wage and employment effects. Points estimates are 0.077 and and -0.042, and are both non-significant. The wage effect is approximately equal to the employment effect, such that labor earnings are not positively affected by an increase in wage subsidies. These effects suggest a pass-through of wage subsidies to wages equal to 37% on average.[13]

A note of caution about the interpretation of the results. The null effects on labor earnings should not be interpreted as a stagnation in absolute value. The set of fixed-effects (including year, labor market level and treatment groups) controls for specific labor supply and labor demand trends and year-specific shocks. In France, labor earnings are growing positively on average. An increase in the net-of-tax rate or participation rate leads to a decrease in the prospective wage. Put in another way, absent wage subsidies, labor market outcomes would have experienced higher growth rates.

My results are qualitatively consistent with the scarce literature on wage and employment effects of wage subsidies. Leigh (2010) finds that an increase of 10% in the EITC decreases the hourly wage rate by 5 % for high school dropouts using variation in US states EITC. His analysis uses wage and employment levels, while mine uses log-growth rates. As a result, my estimates are compatible with reasonable incidence parameters for wage subsidies. Moreover, the magnitude of my wage and employment effects are consistent with Rothstein (2010) calibration using a competitive labor market model. His analysis focuses on the labor market for women, while my paper uses all individuals who are low-wage earners. Still, my results are similar for a reasonable set of micro elasticities.

Robustness tests To access the plausibility and the robustness of my results, I implement two checks. First, I run falsification tests consisting in regressing past outcomes on currents shocks. Second, I run alternative specifications than my baseline estimation. In particular, I introduce less stringent set of fixed-effects and I use alternative regression weights.

I start with the falsification tests. Table 3 reports results from regressions of past outcomes on current tax shocks, using the same shift-share IV research design as in the baseline results. Intuitively, pre-reform outcomes should not be correlated with shocks from the reform period. I implement this test for the two-years following the reform,

[13] Results are quantitatively similar when including only the log-growth in participation tax rate. They are available upon request.

reducing the sample size to 288 observations. In all specification, I cannot reject that there is no relationship between current tax shocks and past outcomes. In addition standard errors are high, suggesting this regression only captures noise. Overall, it validates the shift-share IV design as credible identification strategy.

Then, I investigate if less stringent sets of fixed-effects affect my results. More precisely, I split the treatment group fixed-effects into three separate sets of fixed-effects: household income bin fixed-effects, a dummy for being in a couple or not and a dummy for having children or not. Table 4 reports the results for these alternative specifications. For wage, employment and labor earnings responses, coefficients are almost similar to the baseline specification.

Finally, I report additional results using alternative regression weights with the same baseline specification. I use the initial share of a local labor market in the national labor supply (number of hours worked), instead of the initial share of a local labor market in the national population. Table 6 reports the results for the main specification and Table 7 reports the results for the falsification tests. Again, for wage, employment and labor earnings responses, coefficient are almost similar to the baseline specification.[14]

5.3 Discussion and Limitations

Minimum wage Wage subsidies target low-wage earners, who are also more likely to be close to the minimum wage. In particular, the minimum wage in France is more binding than in other countries such as the United States. As a result, it could reduce both the wage effect and the employment effect. My estimation strategy minimizes this concern by including a dummy for being close to the hourly minimum wage in the initial period.

I provide additional elements that goes in favor of the results not being driven by

[14] I also provide results using an alternative definition of the labor market level variables, using cross-sectional weights. They can be found in Table 8, Table 9, Table 10 and Table 11. This is not my preferred measures because the weights are made to match the aggregate distribution of incomes at the national level (not only labor earnings) and are not designed to be longitudinal weights. Still, the results are qualitatively similar.

changes in the hourly minimum wage. Figure 15 shows the cross-sectional distribution of the share of workers, by distance to the minimum wage and over time. All series are stable over time. Individuals below 1.1 times the minimum wage represent close to 10% of the working population. Figure 16 plots the transition between bins of distance to the minimum wage between two consecutive periods. Individuals starting below 1.1 the minimum wage have 50 % chance remaining in the same bin, suggesting that an important share of them grow more than the minimum wage between two years. The likelihood to remain in the same bin is larger for higher hourly wage. Figure 17 confirms this result by plotting the number of consecutive years spent close to the minimum wage, conditional on starting a period at the minimum wage. 70% of individuals only spent a year or less at the minimum wage.

Micro elasticities My finding that the wage effect offsets the employment effect is interesting for discussing micro elasticities. It suggests that labor demand is far for being completely elastic and that labor supply micro elasticities (at the intensive and/or extensive margin) are sizeable. Using the simple model from subsection 3.1, the total number of hours worked in a labor market is a decreasing function of the labor demand elasticity. Intuitively, as the labor demand becomes more inelastic, employers do not adjust their number of hours worked. In return, the wage rate is very sensitive to an increase in labor supply. This result is consistent with sizeable wage and employment effects at the labor market level. Using my preferred specification (from Table 2) and assuming a compensated elasticity of labor supply $\varepsilon^c = 0$ (consistent with no response to change in the net-of-tax rate), it suggests an elasticity of labor demand of -0.9 and a labor supply participation elasticity of 0.5. The magnitude of the participation elasticity is consistent with results from Eissa and Liebman (1996) and Whitmore Schanzenbach and Strain (2021).[15]

[15] Another potential explanation for the magnitude of the response with respect to the participation tax rate is that individuals are "ironing": they linearizes their tax schedule using their average tax rate. Rees-Jones and Taubinsky (2020) find that 43% of the US population is ironing. In this context, wage and employment responses with respect to the participation tax rate is a combination of the labor supply elasticity at the intensive and extensive margin, and the elasticity of labor demand.

Limitations Results from my research design have two limitations. First, I only have two post-reform years due to data limitations. While this time window is reasonable to infer short-run wage and employment responses to change in wage subsidies, results may differ in the long-run. Indeed, labor demand can be relatively inelastic in the short-run. On the contrary, employers can possibly adjust their production in the long-run to a larger extent. In this case, my results provide an upper-bound for the wage and employment effects in the long-run: wage effects can be even more negative and employment effects get closer to zero.

Second, individuals might adjust gradually to the reform. Labor supply responses can be delayed if individuals do not respond immediately to a change in wage subsidies. This is likely in a context with salience effects, information frictions about wage subsidy programs or infrequent wage renegotiation. My research design minimizes these concerns by taking two-year differences in log wage, hours and labor earnings. Still, the employment effect might be higher in the long-run. Again, my results should be interpreted as upper-bounds for the wage and employment effects in the long-run.

6 Conclusion

This paper evaluates the wage and employment effects of wage subsidies at the labor market level. I depart from the conventional assumption of the absence of equilibrium effects in the labor market by allowing for labor demand and labor supply responses. I leverage a unique combination of rich administrative data on individuals, a reform in wage subsidies in France in 2015, and a novel quasi-experimental research design to quantify wage and employment effects separately.

This paper shows that a more generous wage subsidy increases the number of hours worked in the labor market, but it is offset by a decrease in wage rate growth. The labor market level elasticities for wage (resp. employment) is close to 0 with respect to the marginal tax rate and equal to -0.367 (resp. 0.325) with respect to the average tax rate. Overall, the wage effect is equal to the employment effect such that labor earnings growth is non-significantly different from zero. These effects suggest a pass-through of wage subsidies to wages equal to 37% on average.

These findings highlight that employers can capture part of the increase in wage subsidies through reduced wage rate growth. This has important implications for the design of programs incentivizing individuals to increase their labor supply. There is a hidden and incidental cost, as the targeted population does not receive full benefit from them. This is particularly true because these programs often concern working-poor households. A negative income tax, as discussed by Rothstein (2010), can be a more effective tool to redistribute to the bottom part of the income distribution.

References

Azmat, Ghazala (2019). "Incidence, salience, and spillovers: The direct and indirect effects of tax credits on wages". In: *Quantitative Economics* 10.1, pp. 239–273.

Bastian, Jacob (2020). "The rise of working mothers and the 1975 earned income tax credit". In: *American Economic Journal: Economic Policy* 12.3, pp. 44–75.

Bollinger, Christopher, Luis Gonzalez, and James P Ziliak (2009). "Welfare reform and the level and composition of income". In: *Welfare reform and its long-term consequences for America's poor*, pp. 59–103.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel (2022). "Quasi-experimental shift-share research designs". In: *The Review of Economic Studies* 89.1, pp. 181–213.

Bourguignon, François (2011). "Rapport final du comité national d'évaluation du RSA". In: *La documentation française*.

Brewer, Mike and Hilary Hoynes (2019). "In-Work Credits in the UK and the US". In: *Fiscal Studies* 40.4, pp. 519–560.

Chetty, Raj, John N Friedman, and Emmanuel Saez (2013). "Using Differences in Knowledge across Neighborhoods to Uncover the Impacts of the EITC on Earnings". In: *American Economic Review* 103.7, pp. 2683–2721.

Eissa, Nada and Hilary W Hoynes (2006). "Behavioral responses to taxes: Lessons from the EITC and labor supply". In: *Tax policy and the economy* 20, pp. 73–110.

Eissa, Nada and Jeffrey B Liebman (1996). "Labor supply response to the earned income tax credit". In: *The quarterly journal of economics* 111.2, pp. 605–637.

Ferriere, Axelle et al. (2021). "Larger transfers financed with more progressive taxes? On the optimal design of taxes and transfers". In:

Froemel, Maren and Charles Gottlieb (2021). "The Earned Income Tax Credit: Targeting the poor but crowding out wealth". In: *Canadian Journal of Economics/Revue canadienne d'économique* 54.1, pp. 193–227.

Gelber, Alexander M and Joshua W Mitchell (2012). "Taxes and time allocation: Evidence from single women and men". In: *The Review of Economic Studies* 79.3, pp. 863–897.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift (2020). "Bartik instruments: What, when, why, and how". In: *American Economic Review* 110.8, pp. 2586–2624.

Grogger, Jeffrey (2003). "The effects of time limits, the EITC, and other policy changes on welfare use, work, and income among female-headed families". In: *Review of Economics and statistics* 85.2, pp. 394–408.

Hamermesh, Daniel S (1996). *Labor demand*. princeton University press.

Hotz, V Joseph and John Karl Scholz (2006). *Examining the effect of the earned income tax credit on the labor market participation of families on welfare*.

Hotz, V. Joseph (2003). "The Earned Income Tax Credit". In: *Means-Tested Transfer Programs in the United States*. National Bureau of Economic Research, Inc, pp. 141–198. URL: https://EconPapers.repec.org/RePEc:nbr:nberch:10256.

Hoynes, Hilary (2019). "The earned income tax credit". In: *The Annals of the American Academy of Political and Social Science* 686.1, pp. 180–203.

Kleven, Henrik (2019). *The EITC and the extensive margin: A reappraisal*. Tech. rep. National Bureau of Economic Research.

Leigh, Andrew (2010). "Who benefits from the earned income tax credit? Incidence among recipients, coworkers and firms". In: *The BE Journal of Economic Analysis & Policy* 10.1.

Meyer, Bruce D (2010). "The effects of the Earned Income Tax Credit and recent reforms". In: *Tax policy and the economy* 24.1, pp. 153–180.

Meyer, Bruce D and Dan T Rosenbaum (2001). "Welfare, the earned income tax credit, and the labor supply of single mothers". In: *The quarterly journal of economics* 116.3, pp. 1063–1114.

Nichols, Austin and Jesse Rothstein (2015). "The earned income tax credit". In: *Economics of Means-Tested Transfer Programs in the United States, Volume 1*. University of Chicago Press, pp. 137–218.

Ortigueira, Salvador and Nawid Siassi (2022). "The us tax-transfer system and low-income households: Savings, labor supply, and household formation". In: *Review of Economic Dynamics* 44, pp. 184–210.

Rees-Jones, Alex and Dmitry Taubinsky (2020). "Measuring "schmeduling"". In: *The Review of Economic Studies* 87.5, pp. 2399–2438.

Rothstein, Jesse (2010). "Is the EITC as good as an NIT? Conditional cash transfers and tax incidence". In: *American economic Journal: economic policy* 2.1, pp. 177–208.

Whitmore Schanzenbach, Diane and Michael R Strain (2021). "Employment effects of the Earned Income Tax Credit: Taking the long view". In: *Tax Policy and the Economy* 35.1, pp. 87–129.

Tables

Table 1: Shock and labor market outcomes summary statistics

	Mean	Standard deviation	p5	Median	p95
(a) Shocks					
$\Delta \ln(1\text{-MTR}^{sim})$	-0.023	0.055	-0.112	-0.007	0.055
$\Delta \ln(1\text{-ATR}^{sim})$	-0.012	0.017	-0.044	-0.010	0.014
(b) Shocks with year F.E					
$\Delta \ln(1\text{-MTR}^{sim})$	0	0.053	-0.077	0.003	0.068
$\Delta \ln(1\text{-ATR}^{sim})$	0	0.015	-0.024	0	0.025
(c) Sample size					
1/HHI	275	275	275	275	275
Largest share	0.008	0.008	0.008	0.008	0.008
Treatment groups	144	144	144	144	144
Treatment groups x year	720	720	720	720	720

Notes: This table summarizes the distribution of instruments (net-of-tax rate and participation rate) across treatment groups. Shocks are two-year difference in log. All statistics are weighted by the average treatment group exposure share $s_{g,t}$ for 2012-2015. In the second panel, shocks are first residualized on year fixed-effects. I also report information about the sample: the inverse of the Herfindal index, the largest average exposure share, the number of units and the total number of observations.

Table 2: Shift-share IV estimates

	$\Delta\ln(\text{hours})$	$\Delta\ln(\text{wage})$	$\Delta\ln(\text{labor earnings})$
	(1)	(2)	(3)
$\Delta\ln(\text{1-MTR})$	0.034	0.043	0.077
	(0.083)	(0.080)	(0.144)
$\Delta\ln(\text{1-ATR})$	0.325***	-0.367***	-0.042
	(0.057)	(0.045)	(0.071)
Observations	720	720	720
F-test (1st stage), $\Delta\ln(\text{1-MTR})$	44.1	44.1	44.1
F-test (1st stage), $\Delta\ln(\text{1-ATR})$	224.9	224.9	224.9
Treatment group F.E	✓	✓	✓
Period F.E	✓	✓	✓
Labor market F.E	✓	✓	✓
Base-year controls	✓	✓	✓

Notes: This table reports coefficients from shift-share IV regressions of the two-year change in the log of labor market outcomes on tax shocks, weighted by the share of the labor market in the national population in the initial period. The two tax shocks are the two-year change in the log of the net-of-tax rate and participation rate, instrumented by the corresponding simulated instruments. Labor market outcomes are the number of hours worked (panel (1)), the hourly wage rate (panel (2)) and the sum of labor earnings (panel (3)). Base-year controls include a set of socio-economic characteristics in the initial year. Results come from equivalent treatment group level regressions in order to obtain exposure-robust standard errors and F-statistics, and to include treatment group fixed-effects. Standard errors are clustered at the household income level.

Table 3: Falsification tests for the shift-share IV

	$\Delta\ln$(hours)	$\Delta\ln$(wage)	$\Delta\ln$(labor earnings)
	(1)	(2)	(3)
$\Delta\ln$(1-MTR)	-0.465	0.198	-0.267
	(0.509)	(0.445)	(0.761)
$\Delta\ln$(1-ATR)	-0.581	0.490	-0.092
	(0.724)	(0.417)	(0.970)
Observations	288	288	288
F-test (1st stage), $\Delta\ln$(1-MTR)	4.79	4.79	4.79
F-test (1st stage), $\Delta\ln$(1-ATR)	127.7	127.7	127.7
Treatment group F.E	✓	✓	✓
Period F.E	✓	✓	✓
Labor market F.E	✓	✓	✓
Base-year controls	✓	✓	✓

Notes: This table reports coefficients from shift-share IV falsification tests. I regress the two-year change in the log of past labor market outcomes on current tax shocks, weighted by the share of the labor market in the national population in the initial period. The two tax shocks are the two-year change in the log of the net-of-tax rate and participation rate, instrumented by the corresponding simulated instruments. Labor market outcomes are the number of hours worked (panel (1)), the hourly wage rate (panel (2)) and the sum of labor earnings (panel (3)). Base-year controls include a set of socio-economic characteristics in the initial year. Results come from equivalent treatment group level regressions in order to obtain exposure robust standard errors and F-statistics, and to include treatment group fixed-effects. Standard errors are clustered at the household income level.

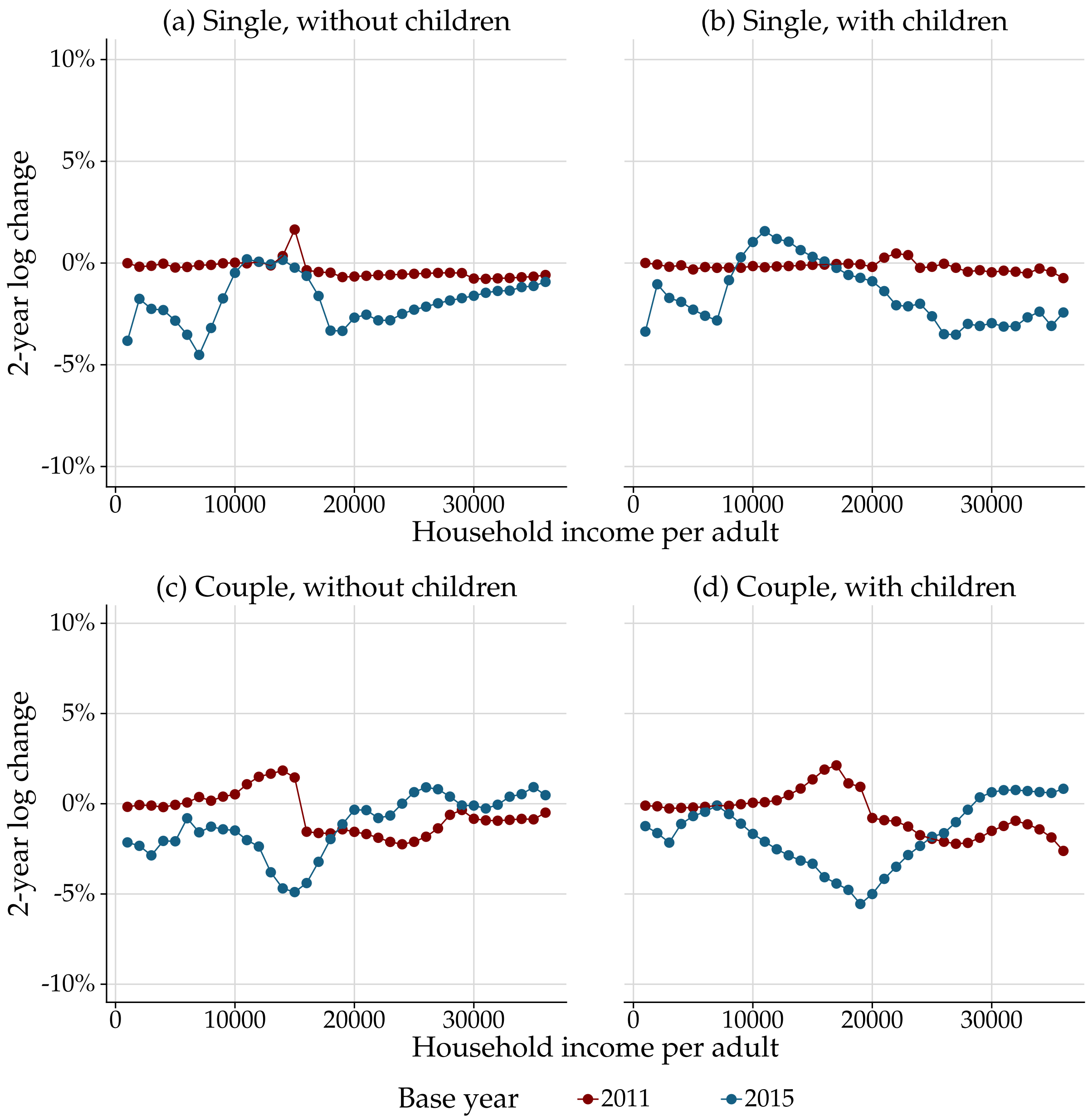
Table 4: Shift-share IV estimates, alternative specifications

	$\Delta \ln$(hours)			$\Delta \ln$(wage)			$\Delta \ln$(labor earnings)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\Delta \ln$(1-MTR)	0.046	0.047	0.045	0.051	0.048	0.049	0.097	0.094	0.093
	(0.104)	(0.102)	(0.100)	(0.077)	(0.078)	(0.077)	(0.151)	(0.151)	(0.150)
$\Delta \ln$(1-ATR)	0.329***	0.328***	0.328***	-0.376***	-0.370***	-0.370***	-0.047	-0.042	-0.042
	(0.059)	(0.061)	(0.061)	(0.044)	(0.044)	(0.044)	(0.072)	(0.075)	(0.075)
Observations	720	720	720	720	720	720	720	720	720
F-test (1st stage), $\Delta \ln$(1-MTR)	35.3	36.2	36.2	35.3	36.2	36.2	35.3	36.2	36.2
F-test (1st stage), $\Delta \ln$(1-ATR)	179.1	195.2	195.1	179.1	195.2	195.1	179.1	195.2	195.1
Children FE			✓			✓			✓
Couple F.E		✓	✓		✓	✓		✓	✓
HH income F.E	✓	✓	✓	✓	✓	✓	✓	✓	✓
Period F.E	✓	✓	✓	✓	✓	✓	✓	✓	✓
Labor market F.E	✓	✓	✓	✓	✓	✓	✓	✓	✓
Base-year controls	✓	✓	✓	✓	✓	✓	✓	✓	✓

Notes: This table reports coefficients from shift-share IV regressions of the two-year change in the log of labor market outcomes on tax shocks, weighted by the share of the labor market in the national population in the initial period. The two tax shocks are the two-year change in the log of the net-of-tax rate and participation rate, instrumented by the corresponding simulated instruments. Labor market outcomes are the number of hours worked (panel (1), (2) and (3)), the hourly wage rate (panel (4), (5) and (6)) and the sum of labor earnings (panel (7), (8), (9)). Base-year controls include a set of socio-economic characteristics in the initial year. Results come from equivalent treatment group level regressions in order to obtain exposure-robust standard errors and F-statistics, and to include treatment group fixed-effects. Standard errors are clustered at the household income level.

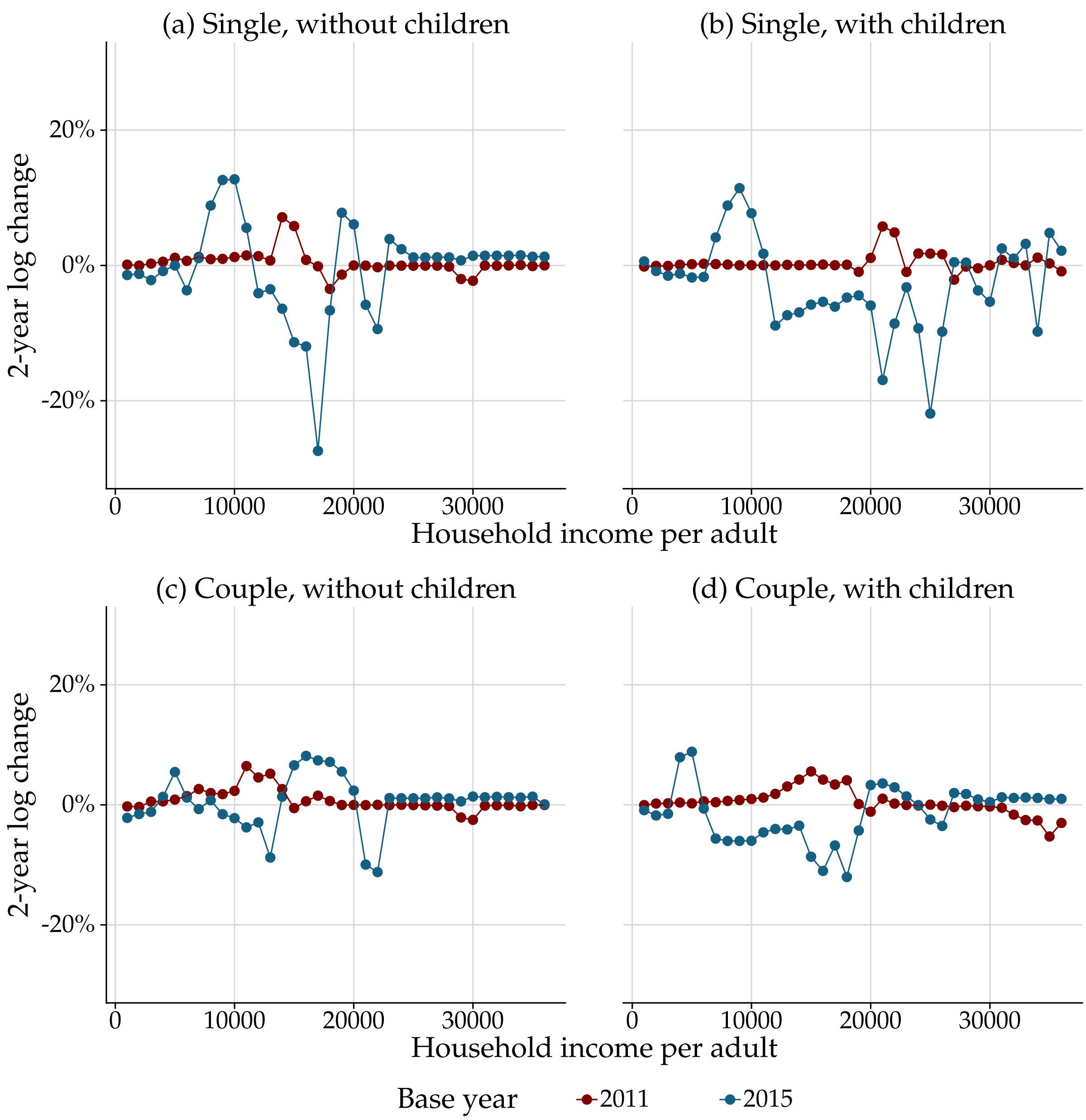
Figures

Figure 1: $\Delta\ln$(1-ATR) by treatment group, for base year 2011 and 2015



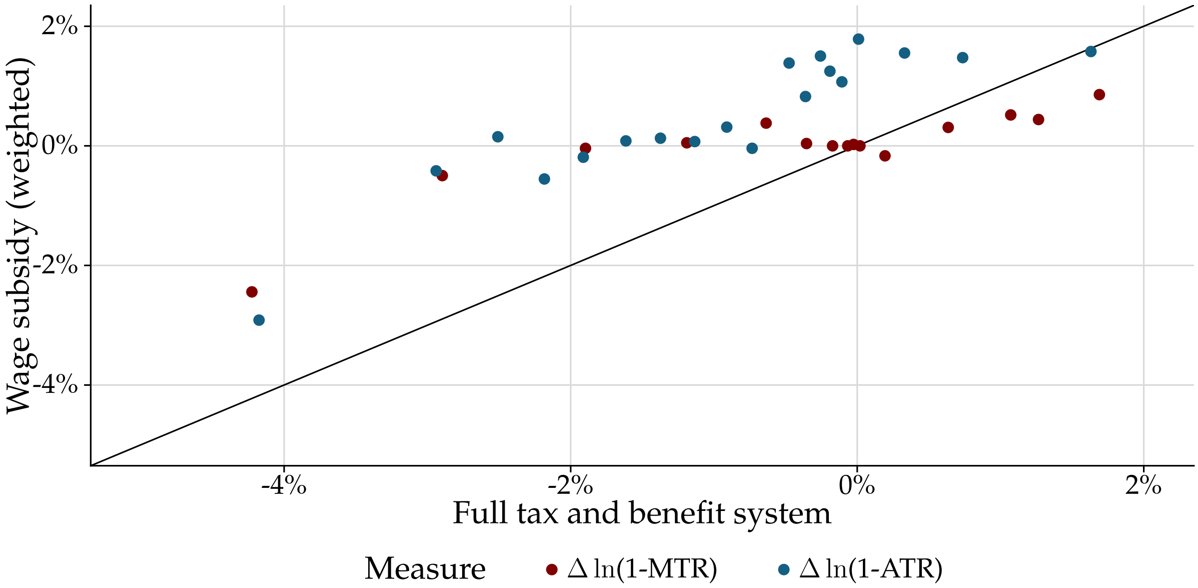
Notes: The figure plots the two-year change in the log of the participation tax rate at the treatment group-level, for base year 2011 and 2015 and considering the full tax and redistribution system. The tax groups are defined by the combination of being single (panel (a) and (b)) or in a couple (panel (c) and (d)), having children (panel (a) and (c)) or not (panel (c) and (d)) and the household income per adult in euros (x-axis).

Figure 2: Δ ln(1-MTR) by treatment group, for base year 2011 and 2015



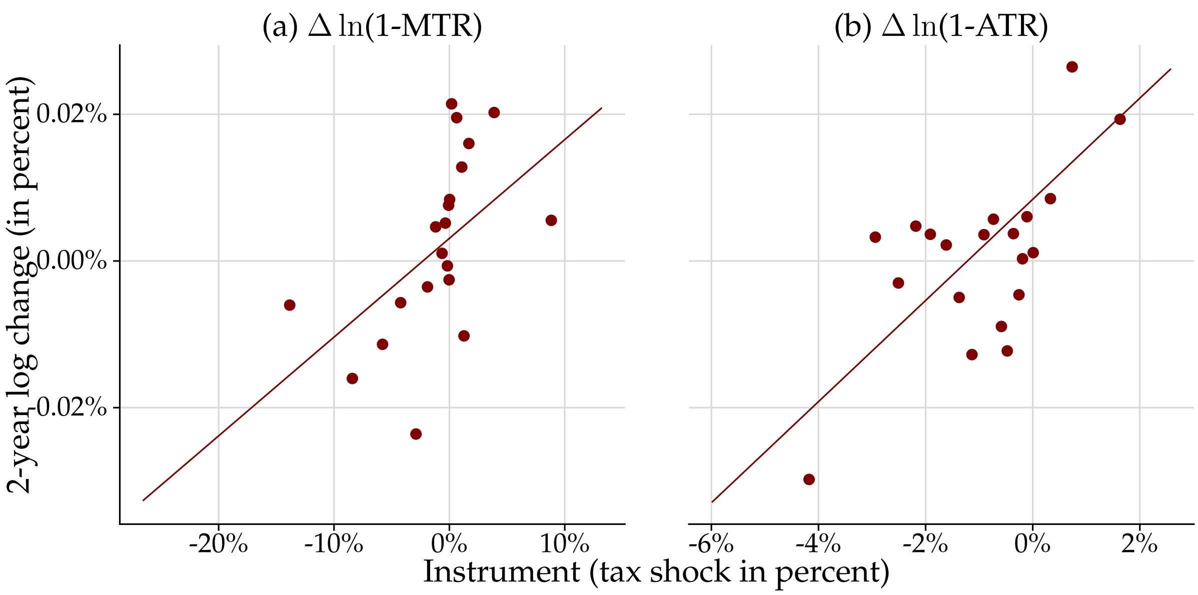
Notes: The figure plots the two-year change in the log of the net-of-tax rate at the treatment group-level, for base year 2011 and 2015 and considering the full tax and redistribution system. The tax groups are defined by the combination of being single (panel (a) and (b)) or in a couple (panel (c) and (d)), having children (panel (a) and (c)) or not (panel (c) and (d)) and the household income per adult in euros (x-axis).

Figure 3: Correlation between exposure measures



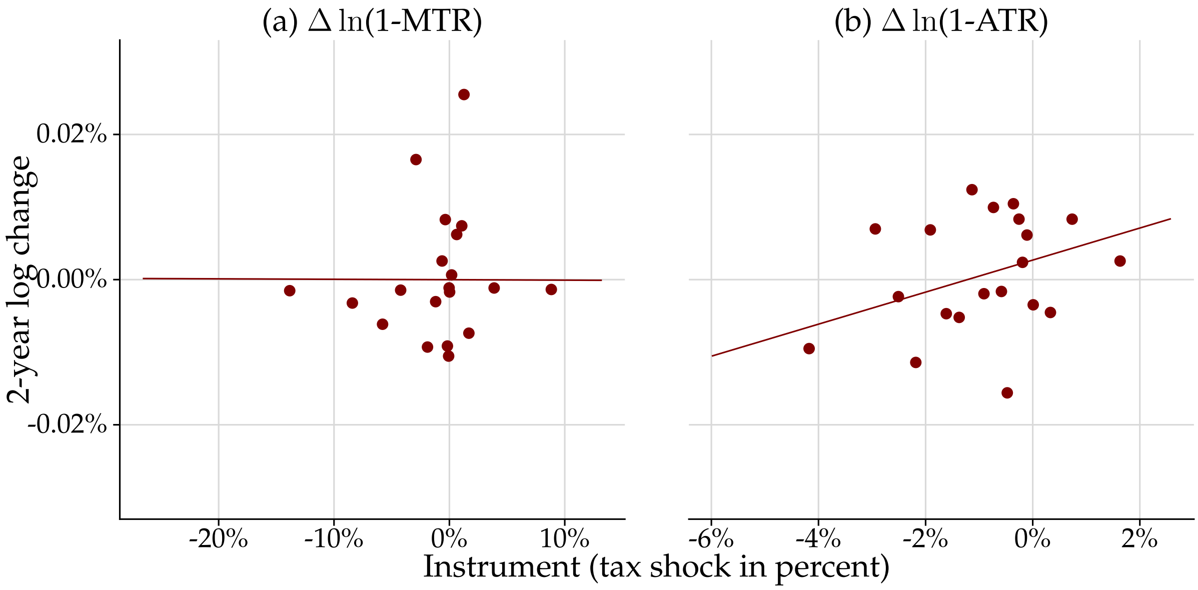
Notes: The figure plots the correlation between the two-year change in tax shocks only considering the wage subsidies (y-axis) and the full tax and benefit system (x-axis) at the treatment group-level, pooling all years together. Results are reported separately for the participation tax rate and the net-of-tax rate. Log-growth in wage subsidies in weighted by their share in the total taxes and benefits in the initial year. Each dot represents 5% of the data. Results from an OLS regression (with intercept) reports a correlation coefficient of 0.80 for the net-of-tax rate and of 0.77 for the participation tax rate.

Figure 4: First-stage estimations, SSIV research design



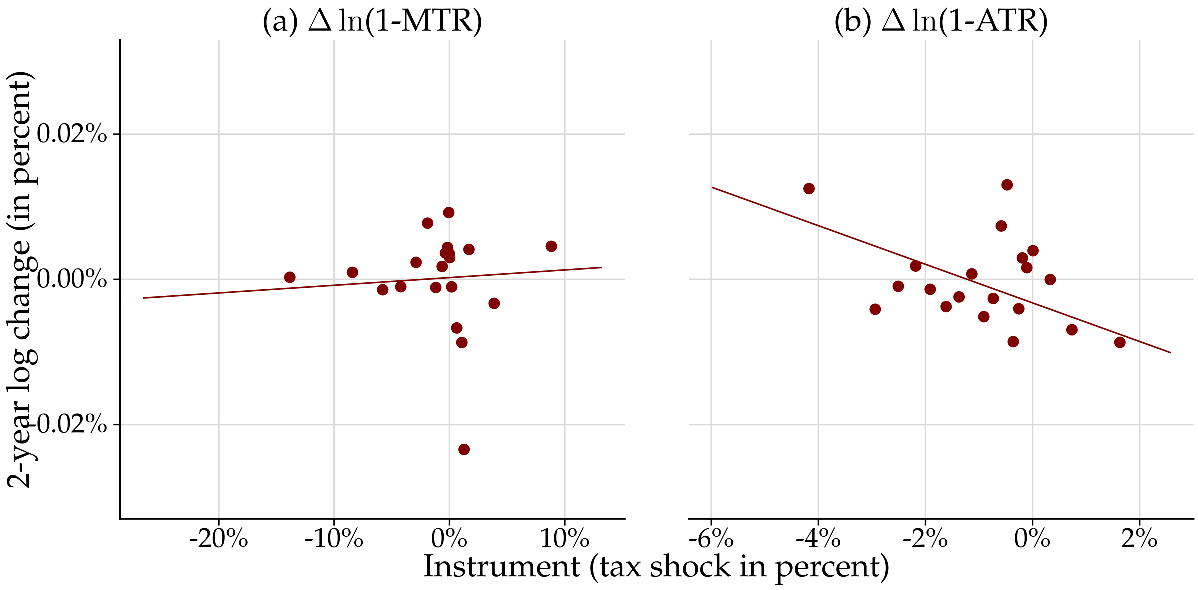
Notes: The figure plots the reduced-form relationship underlying the shift-share IV research design. It plots the correlation between the observed tax shocks and the corresponding instruments. Panel (a) shows the correlation for the participation tax rate, controlling for the net-of-tax rate. Panel (b) shows the correlation for net-of-tax rate, controlling for the participation rate. Observation are weighted by the average treatment group exposure share $s_{g,t}$. The x-axis shows the simulated instruments and the y-axis the average observed tax shocks. Each dot represents 5% of the data.

Figure 5: Reduced-form employment relationships



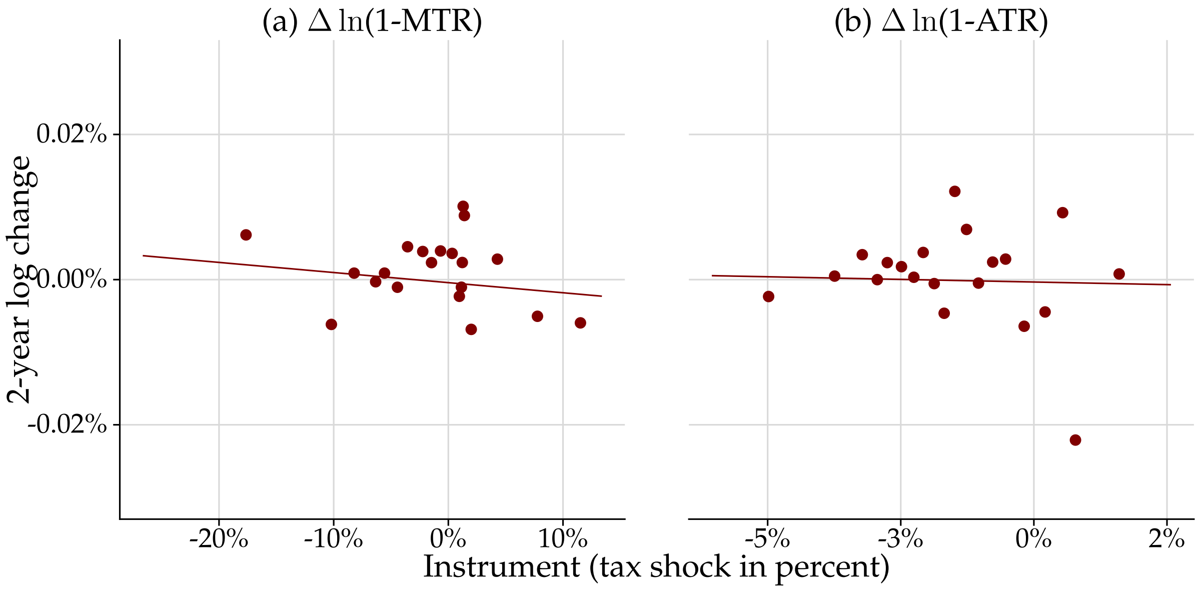
Notes: The figure plots the reduced-form relationship between two-year change in the log of hours at the labor market level and the two instruments. The two tax shocks are the two-year change in the log of the participation rate and net-of-tax rate. Panel (a) shows the correlations with respect to the participation tax rate and panel (b) shows the correlations with respect to the net-of-tax rate. Outcomes are first residualized on a set of labor market fixed-effects, year fixed-effects, a set of socio-economic characteristics in the initial year and the other tax shock. The corresponding regression is weighted by the average treatment group exposure share $s_{g,t}$. The x-axis shows the shock size (simulated instruments) and the y-axis the average change in outcome. Each dot represents 5% of the data.

Figure 6: Reduced-form wage relationships



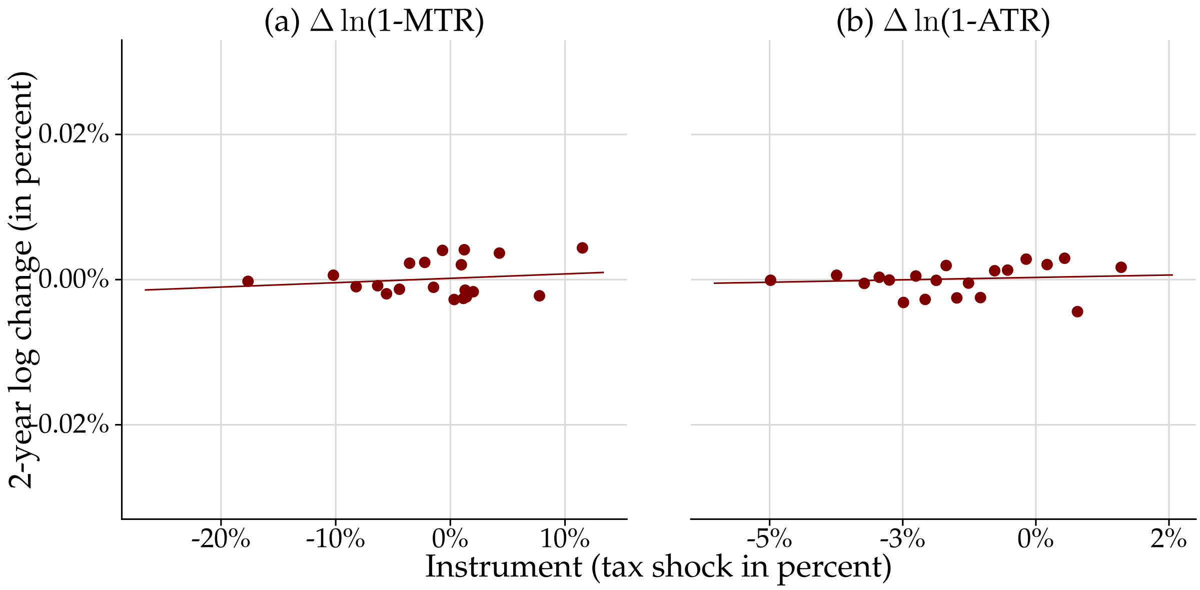
Notes: The figure plots the reduced-form relationship between two-year change in the log hourly wage at the labor market level and the two instruments. The two tax shocks are the two-year change in the log of the participation rate and net-of-tax rate. Panel (a) shows the correlations with respect to the participation tax rate and panel (b) shows the correlations with respect to the net-of-tax rate. Outcomes are first residualized on a set of labor market fixed-effects, year fixed-effects, a set of socio-economic characteristics in the initial year and the other tax shock. The corresponding regression is weighted by the average treatment group exposure share $s_{g,t}$. The x-axis shows the shock size (simulated instruments) and the y-axis the average change in outcome. Each dot represents 5% of the data.

Figure 7: Reduced-form employment falsification tests



Notes: The figure plots the falsification tests for the two-year change in the log of hours at the labor market level. The two tax shocks are the two-year change in the log of the participation rate and net-of-tax rate. Outcomes are the change in number of hours 3 years before the shocks. Panel (a) shows the correlations with respect to the participation tax rate and panel (b) shows the correlations with respect to the net-of-tax rate. Outcomes are first residualized on a set of labor market fixed-effects, year fixed-effects, a set of socio-economic characteristics in the initial year and the other tax shock. The corresponding regression is weighted by the average treatment group exposure share $s_{g,t}$. The x-axis shows the shock size (simulated instruments) and the y-axis the average change in outcome. Each dot represents 5% of the data.

Figure 8: Reduced-form wage falsification tests



Notes: The figure plots the falsification tests for the two-year change in the log hourly wage at the labor market level. The two tax shocks are the two-year change in the log of the participation rate and net-of-tax rate. Outcomes are the change in number of hours 3 years before the shocks. Panel (a) shows the correlations with respect to the participation tax rate and panel (b) shows the correlations with respect to the net-of-tax rate. Outcomes are first residualized on a set of labor market fixed-effects, year fixed-effects, a set of socio-economic characteristics in the initial year and the other tax shock. The corresponding regression is weighted by the average treatment group exposure share $s_{g,t}$. The x-axis shows the shock size (simulated instruments) and the y-axis the average change in outcome. Each dot represents 5% of the data.

A Tables

Table 5: Shock and labor market outcomes summary statistics, with cross-sectional weights

	Mean	Standard deviation	p5	Median	p95
(a) Shocks					
$\Delta \ln(1\text{-MTR}^{sim})$	-0.023	0.055	-0.113	-0.007	0.056
$\Delta \ln(1\text{-ATR}^{sim})$	-0.012	0.017	-0.045	-0.010	0.014
(b) Shocks with year F.E					
$\Delta \ln(1\text{-MTR}^{sim})$	0	0.053	-0.077	0.003	0.068
$\Delta \ln(1\text{-ATR}^{sim})$	0	0.015	-0.024	0	0.025
(c) Sample size					
1/HHI	276	276	276	276	276
Largest share	0.008	0.008	0.008	0.008	0.008
Treatment groups	144	144	144	144	144
Treatment groups x year	720	720	720	720	720
Cross-sectional weights	Yes	Yes	Yes	Yes	Yes

Notes: This table summarizes the distribution of instruments (net-of-tax rate and participation rate) across treatment groups. Shocks are two-year difference in log. All statistics are weighted by the average treatment group exposure share $s_{g,t}$ for 2012-2015. Variables are constructed using cross-sectional administrative weights. In the second panel, shocks are first residualized on year fixed-effects. I also report information about the sample: the inverse of the Herfindal index, the largest average exposure share, the number of units and the total number of observations.

Table 6: Shift-share estimates, with hours-weighted labor markets

	$\Delta\ln(\text{hours})$	$\Delta\ln(\text{wage})$	$\Delta\ln(\text{labor earnings})$
	(1)	(2)	(3)
$\Delta\ln$(1-MTR)	0.024	0.037	0.061
	(0.090)	(0.084)	(0.151)
$\Delta\ln$(1-ATR)	0.350***	-0.376***	-0.026
	(0.056)	(0.045)	(0.074)
Observations	720	720	720
F-test (1st stage), $\Delta\ln$(1-MTR)	42.9	42.9	42.9
F-test (1st stage), $\Delta\ln$(1-ATR)	223.0	223.0	223.0
Treatment group F.E	✓	✓	✓
Period F.E	✓	✓	✓
Labor market F.E	✓	✓	✓
Base-year controls	✓	✓	✓

Notes: This table reports coefficients from shift-share IV regressions of the two-year change in the log of labor market outcomes on tax shocks, weighted by the share of the labor market in the total number of hours worked at the national level in the initial period. The two tax shocks are the two-year change in the log of the net-of-tax rate and participation rate, instrumented by the corresponding simulated instruments. Labor market outcomes are the number of hours worked (panel (1)), the hourly wage rate (panel (2)) and the sum of labor earnings (panel (3)). Base-year controls include a set of socio-economic characteristics in the initial year. Results come from equivalent treatment group level regressions in order to obtain exposure-robust standard errors and F-statistics, and to include treatment group fixed-effects. Standard errors are clustered at the household income level.

Table 7: Falsification tests for the shift-share IV, with hours-weighted labor markets

	$\Delta\ln(\text{hours})$	$\Delta\ln(\text{wage})$	$\Delta\ln(\text{labor earnings})$
	(1)	(2)	(3)
$\Delta\ln(\text{1-MTR})$	-0.363	0.151	-0.212
	(0.478)	(0.436)	(0.730)
$\Delta\ln(\text{1-ATR})$	-0.747	0.503	-0.245
	(0.711)	(0.404)	(0.943)
Observations	288	288	288
F-test (1st stage), $\Delta\ln(\text{1-MTR})$	4.25	4.25	4.25
F-test (1st stage), $\Delta\ln(\text{1-ATR})$	119.4	119.4	119.4
Treatment group F.E	✓	✓	✓
Period F.E	✓	✓	✓
Labor market F.E	✓	✓	✓
Base-year controls	✓	✓	✓

Notes: This table reports coefficients from shift-share IV falsification tests. I regress the two-year change in the log of past labor market outcomes on current tax shocks, weighted by the share of the labor market in the total number of hours worked at the national level in the initial period. The two tax shocks are the two-year change in the log of the net-of-tax rate and participation rate, instrumented by the corresponding simulated instruments. Labor market outcomes are the number of hours worked (panel (1)), the hourly wage rate (panel (2)) and the sum of labor earnings (panel (3)). Base-year controls include a set of socio-economic characteristics in the initial year. Results come from equivalent treatment group level regressions in order to obtain exposure-robust standard errors and F-statistics, and to include treatment group fixed-effects. Standard errors are clustered at the household income level.

Table 8: Shift-share estimates, with cross-sectional weights

	$\Delta\ln$(hours)	$\Delta\ln$(wage)	$\Delta\ln$(labor earnings)
	(1)	(2)	(3)
$\Delta\ln$(1-MTR)	-0.173	0.025	-0.148
	(0.147)	(0.077)	(0.209)
$\Delta\ln$(1-ATR)	0.521^{***}	-0.311^{***}	0.209^{*}
	(0.083)	(0.045)	(0.111)
Observations	720	720	720
F-test (1st stage), $\Delta\ln$(1-MTR)	45.1	45.1	45.1
F-test (1st stage), $\Delta\ln$(1-ATR)	221.6	221.6	221.6
Treatment group F.E	✓	✓	✓
Period F.E	✓	✓	✓
Labor market F.E	✓	✓	✓
Base-year controls	✓	✓	✓

Notes: This table reports coefficients from shift-share IV regressions of the two-year change in the log of labor market outcomes on tax shocks, weighted by the share of the labor market in the national population in the initial period. The two tax shocks are the two-year change in the log of the net-of-tax rate and participation rate, instrumented by the corresponding simulated instruments. Labor market outcomes are the number of hours worked (panel (1)), the hourly wage rate (panel (2)) and the sum of labor earnings (panel (3)). Variables are constructed using cross-sectional administrative weights. Base-year controls include a set of socio-economic characteristics in the initial year. Results come from equivalent treatment group level regressions in order to obtain exposure-robust standard errors and F-statistics, and to include treatment group fixed-effects. Standard errors are clustered at the household income level.

Table 9: Falsification tests for the shift-share IV, with cross-sectional weights

	$\Delta\ln(\text{hours})$	$\Delta\ln(\text{wage})$	$\Delta\ln(\text{labor earnings})$
	(1)	(2)	(3)
$\Delta\ln(\text{1-MTR})$	-1.72	0.055	-1.66
	(1.11)	(0.412)	(1.38)
$\Delta\ln(\text{1-ATR})$	0.294	0.664	0.958
	(1.38)	(0.405)	(1.69)
Observations	288	288	288
F-test (1st stage), $\Delta\ln(\text{1-MTR})$	4.99	4.99	4.99
F-test (1st stage), $\Delta\ln(\text{1-ATR})$	126.9	126.9	126.9
Treatment group F.E	✓	✓	✓
Period F.E	✓	✓	✓
Labor market F.E	✓	✓	✓
Base-year controls	✓	✓	✓

Notes: This table reports coefficients from shift-share IV falsification tests. I regress the two-year change in the log of past labor market outcomes on current tax shocks, weighted by the share of the labor market in the national population in the initial period. The two tax shocks are the two-year change in the log of the net-of-tax rate and participation rate, instrumented by the corresponding simulated instruments. Labor market outcomes are the number of hours worked (panel (1)), the hourly wage rate (panel (2)) and the sum of labor earnings (panel (3)). Variables are constructed using cross-sectional administrative weights. Base-year controls include a set of socio-economic characteristics in the initial year. Results come from equivalent treatment group level regressions in order to obtain exposure-robust standard errors and F-statistics, and to include treatment group fixed-effects. Standard errors are clustered at the household income level.

Table 10: Shift-share estimates, with hours-weighted labor markets and cross-sectional weights

	$\Delta\ln(\text{hours})$	$\Delta\ln(\text{wage})$	$\Delta\ln(\text{labor earnings})$
	(1)	(2)	(3)
$\Delta\ln$(1-MTR)	-0.207	0.018	-0.189
	(0.151)	(0.080)	(0.215)
$\Delta\ln$(1-ATR)	0.508***	-0.324***	0.185
	(0.086)	(0.046)	(0.117)
Observations	720	720	720
F-test (1st stage), $\Delta\ln$(1-MTR)	44.0	44.0	44.0
F-test (1st stage), $\Delta\ln$(1-ATR)	218.9	218.9	218.9
Treatment group F.E	✓	✓	✓
Period F.E	✓	✓	✓
Labor market F.E	✓	✓	✓
Base-year controls	✓	✓	✓

Notes: This table reports coefficients from shift-share IV regressions of the two-year change in the log of labor market outcomes on tax shocks, weighted by the share of the labor market in the total number of hours worked at the national level in the initial period. The two tax shocks are the two-year change in the log of the net-of-tax rate and participation rate, instrumented by the corresponding simulated instruments. Labor market outcomes are the number of hours worked (panel (1)), the hourly wage rate (panel (2)) and the sum of labor earnings (panel (3)). Variables are constructed using cross-sectional administrative weights. Base-year controls include a set of socio-economic characteristics in the initial year. Results come from equivalent treatment group level regressions in order to obtain exposure-robust standard errors and F-statistics, and to include treatment group fixed-effects. Standard errors are clustered at the household income level.

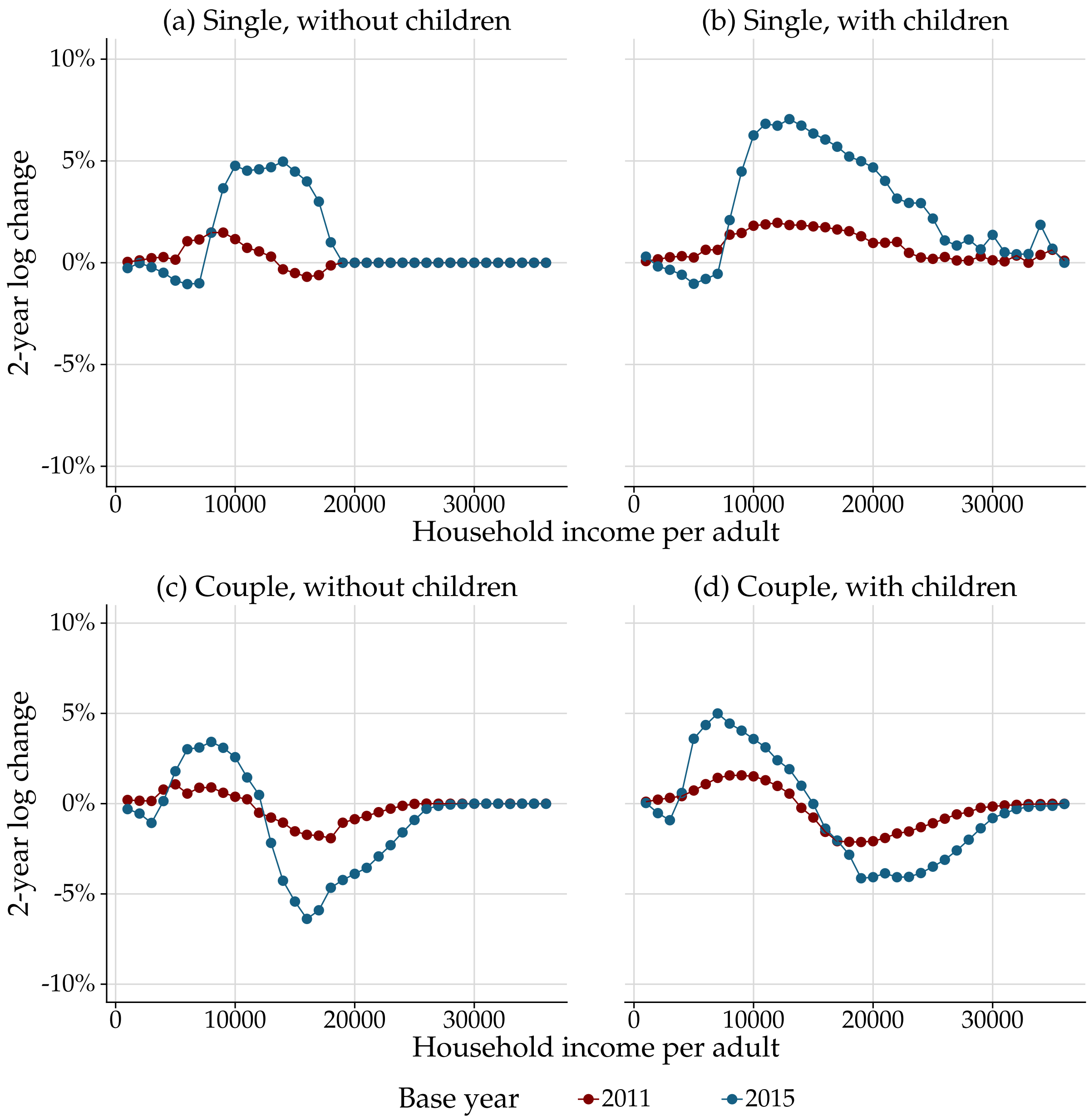
Table 11: Falsification tests for the shift-share IV, with hours-weighted labor markets and cross-sectional weights

	Δ ln(hours) (1)	Δ ln(wage) (2)	Δ ln(labor earnings) (3)
Δ ln(1-MTR)	-1.59	-0.006	-1.59
	(1.07)	(0.406)	(1.34)
Δ ln(1-ATR)	0.027	0.678	0.705
	(1.32)	(0.391)	(1.62)
Observations	288	288	288
F-test (1st stage), Δ ln(1-MTR)	4.37	4.37	4.37
F-test (1st stage), Δ ln(1-ATR)	120.0	120.0	120.0
Treatment group F.E	✓	✓	✓
Period F.E	✓	✓	✓
Labor market F.E	✓	✓	✓
Base-year controls	✓	✓	✓

Notes: This table reports coefficients from shift-share IV falsification tests. I regress the two-year change in the log of past labor market outcomes on current tax shocks, weighted by the share of the labor market in the total number of hours worked at the national level in the initial period. The two tax shocks are the two-year change in the log of the net-of-tax rate and participation rate, instrumented by the corresponding simulated instruments. Labor market outcomes are the number of hours worked (panel (1)), the hourly wage rate (panel (2)) and the sum of labor earnings (panel (3)). Variables are constructed using cross-sectional administrative weights. Base-year controls include a set of socio-economic characteristics in the initial year. Results come from equivalent treatment group level regressions in order to obtain exposure-robust standard errors and F-statistics, and to include treatment group fixed-effects. Standard errors are clustered at the household incomelevel.

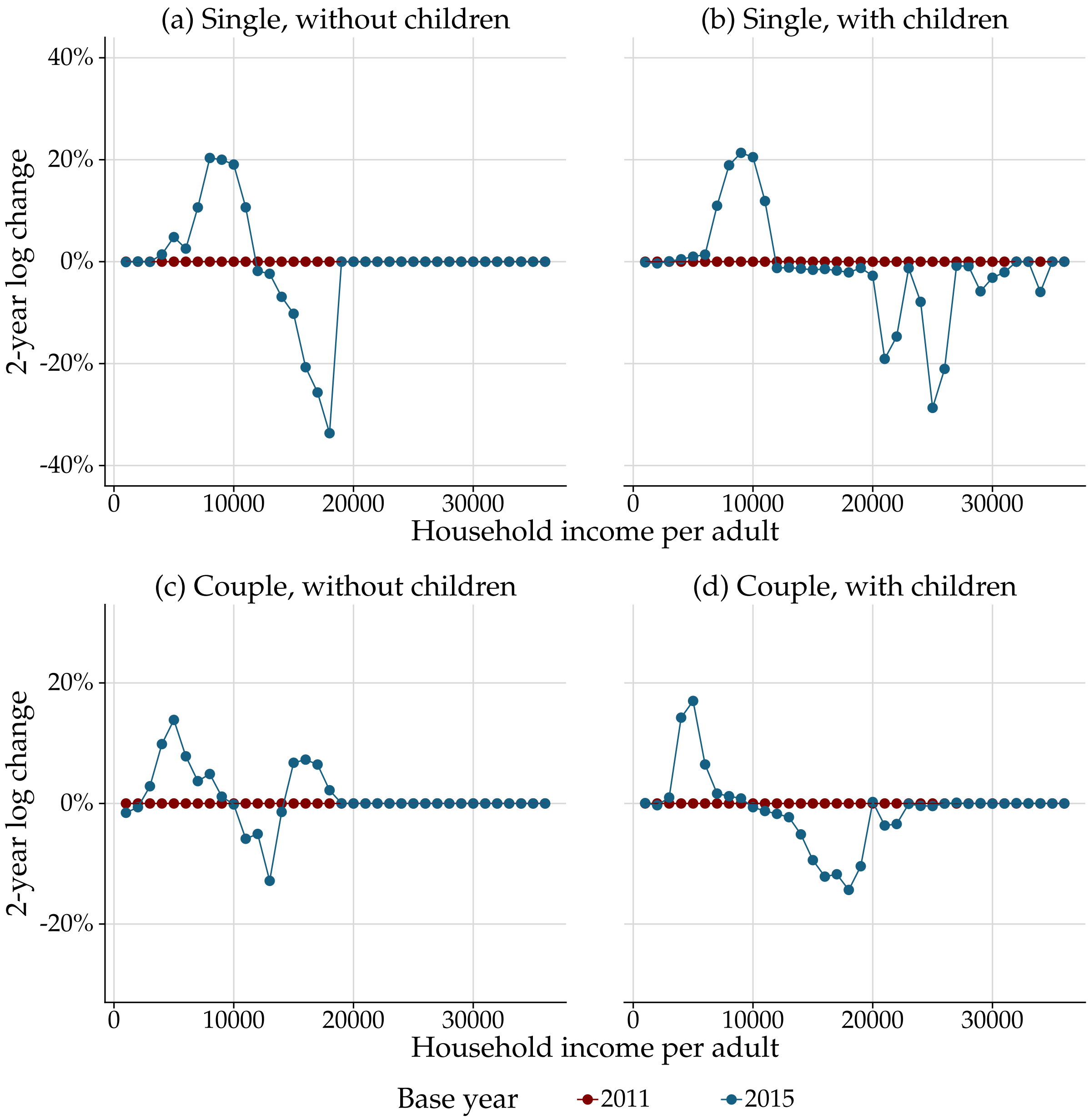
B Figures

Figure 9: $\Delta \ln$(1-ATR) by treatment group, for base year 2011 and 2015 and with only wage subsidies



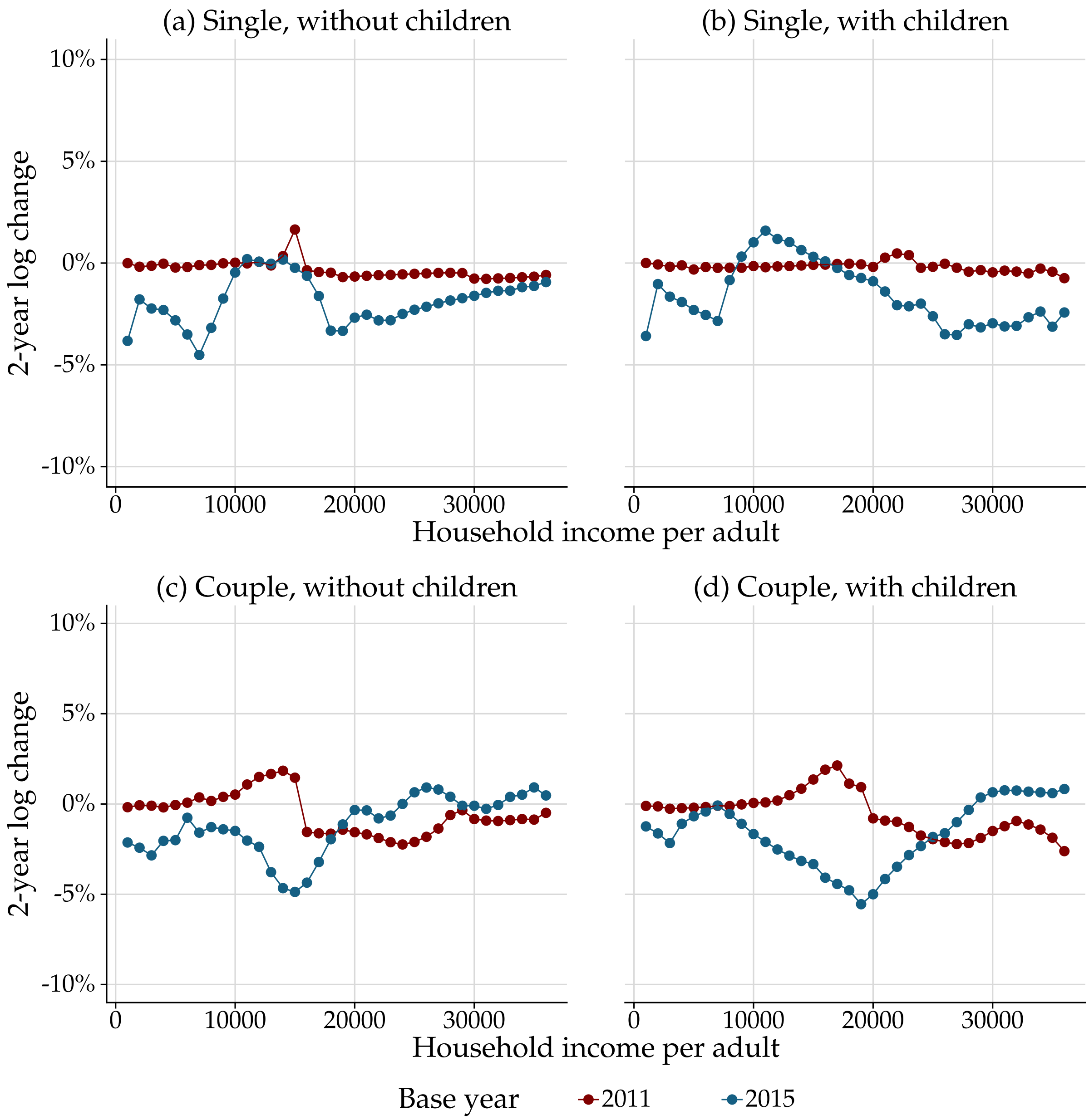
Notes: The figure plots the two-year change in the log of the participation tax rate at the treatment group-level, for base year 2011 and 2015 and considering only wage subsidies. The tax groups are defined by the combination of being single (panel (a) and (b)) or in a couple (panel (c) and (d)), having children (panel (a) and (c)) or not (panel (c) and (d)) and the household income per adult in euros (x-axis).

Figure 10: $\Delta\,\ln$(1-MTR) by treatment group, for base year 2011 and 2015 and with only wage subsidies



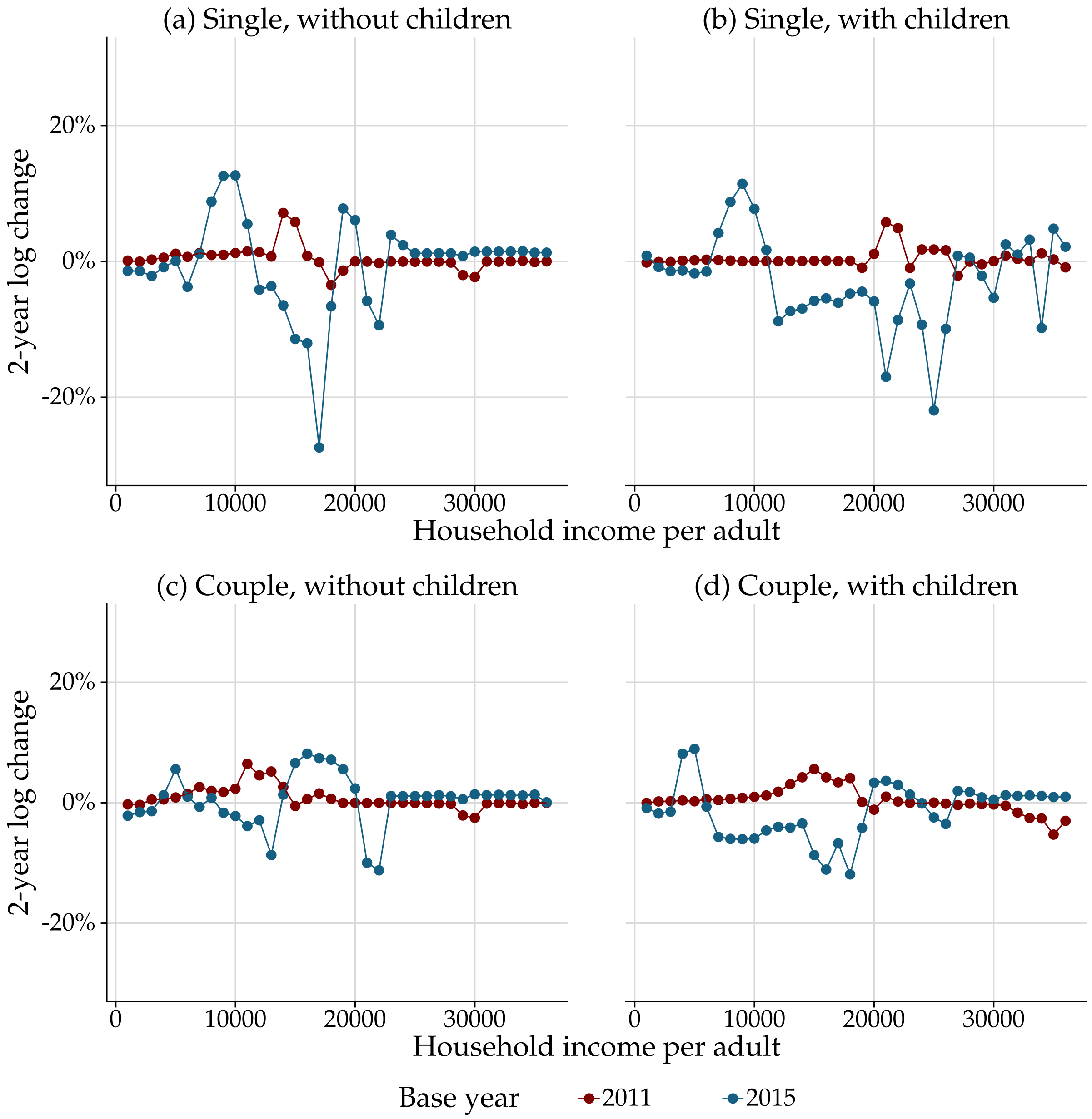
Notes: The figure plots the two-year change in the log of the net-of-tax rate at the treatment group-level, for base year 2011 and 2015 and considering only wage subsidies. The tax groups are defined by the combination of being single (panel (a) and (b)) or in a couple (panel (c) and (d)), having children (panel (a) and (c)) or not (panel (c) and (d)) and the household income per adult in euros (x-axis).

Figure 11: $\Delta \ln$(1-ATR) by treatment group, for base year 2011 and 2015 and using the alternative definition



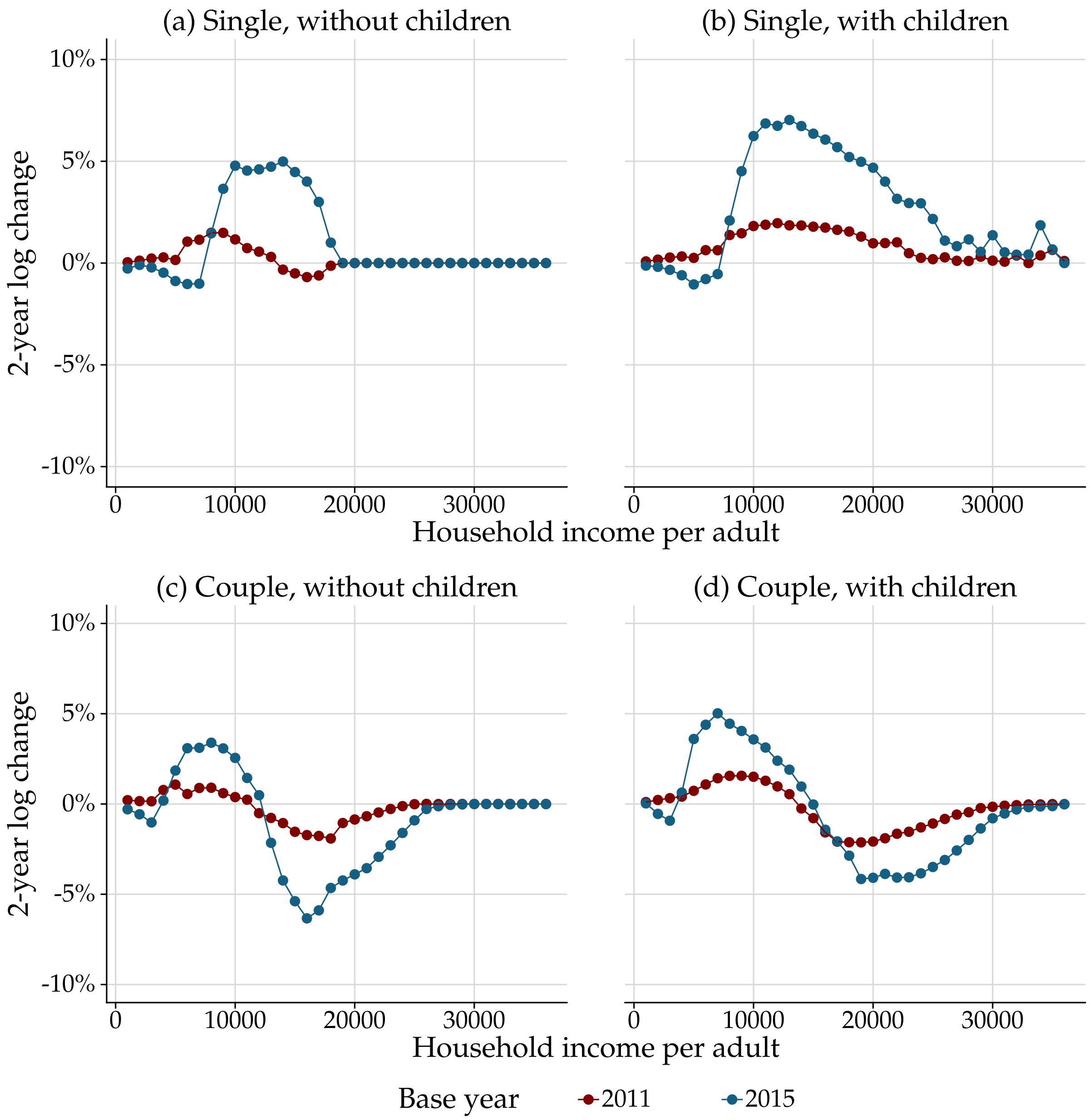
Notes: The figure plots the two-year change in the log of the participation tax rate at the treatment group-level, for base year 2011 and 2015 and considering the full tax and redistribution system. The tax groups are defined by the combination of being single (panel (a) and (b)) or in a couple (panel (c) and (d)), having children (panel (a) and (c)) or not (panel (c) and (d)) and the household income per adult in euros (x-axis). Variables are constructed using cross-sectional administrative weights.

Figure 12: $\Delta \ln$(1-MTR) by treatment group, for base year 2011 and 2015 and using the alternative definition



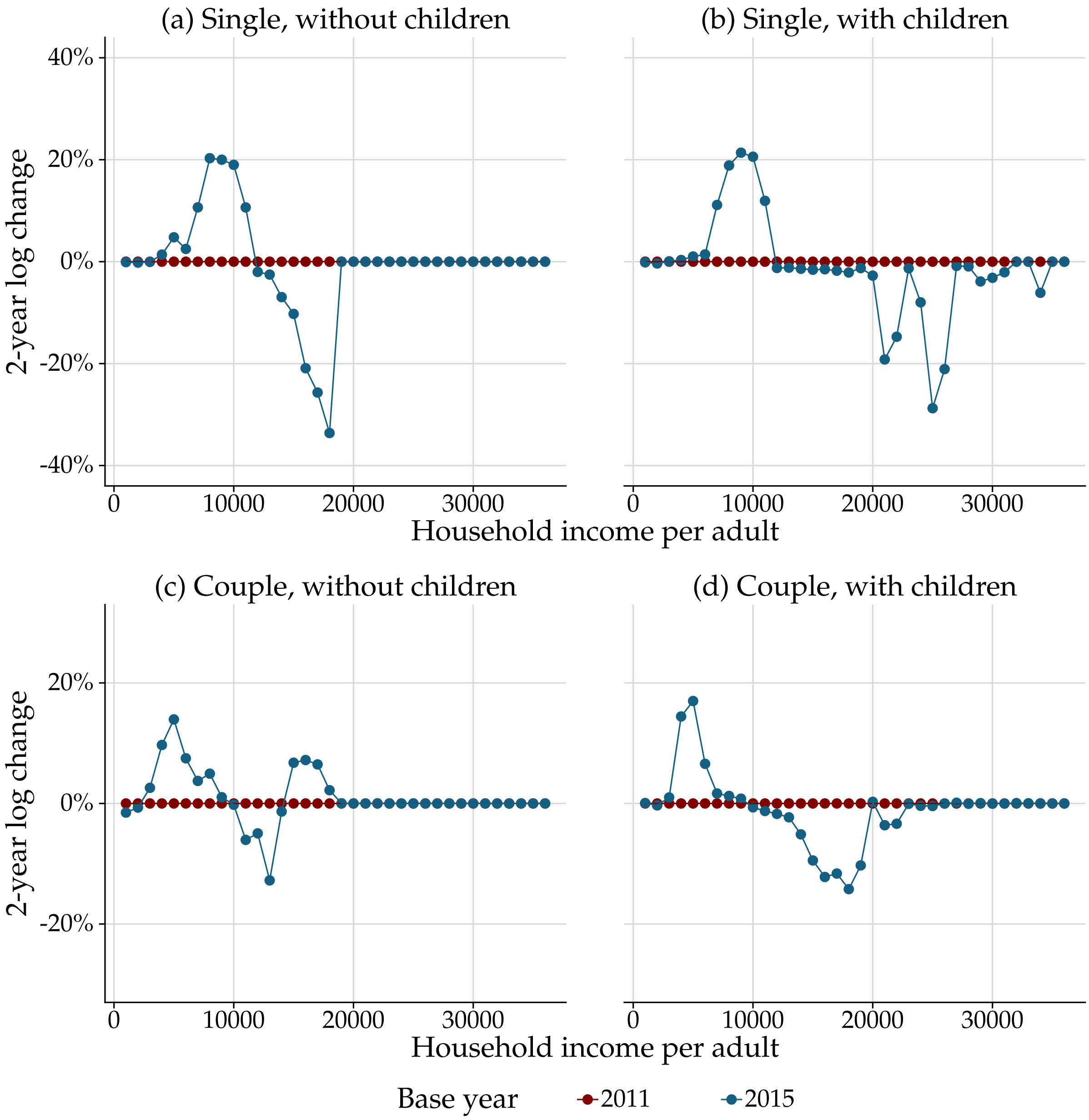
Notes: The figure plots the two-year change in the log of the net-of-tax rate at the treatment group-level, for base year 2011 and 2015 and considering the full tax and redistribution system. The tax groups are defined by the combination of being single (panel (a) and (b)) or in a couple (panel (c) and (d)), having children (panel (a) and (c)) or not (panel (c) and (d)) and the household income per adult in euros (x-axis). Variables are constructed using cross-sectional administrative weights.

Figure 13: $\Delta \ln$(1-ATR) by treatment group, for base year 2011 and 2015, with only wage subsidies and using the alternative definition



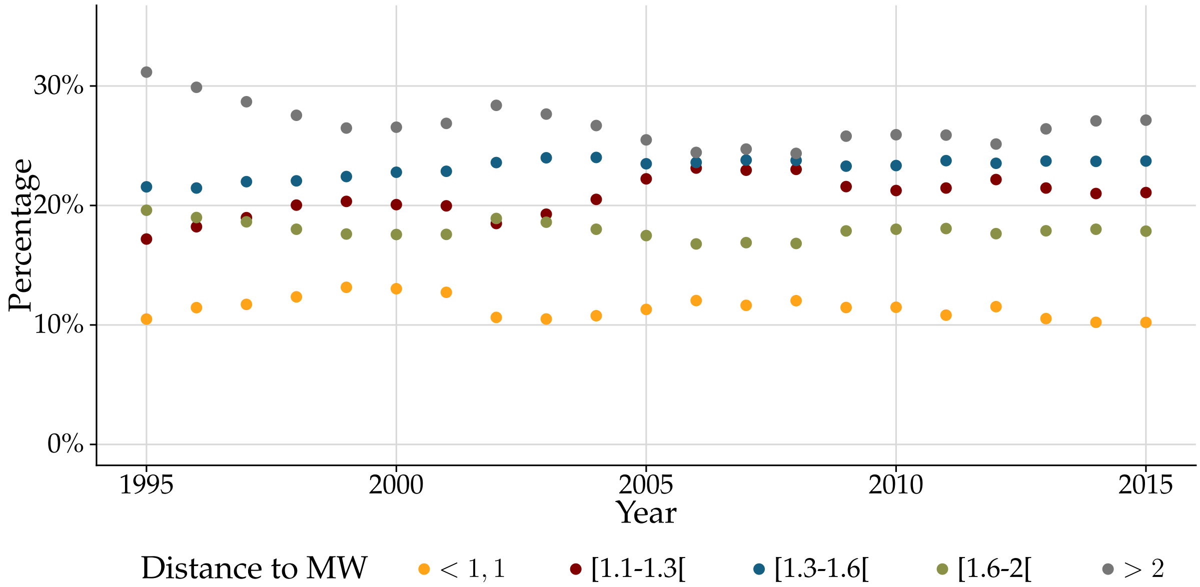
Notes: The figure plots the two-year change in the log of the participation tax rate at the treatment group-level, for base year 2011 and 2015 and considering only wage subsidies. The tax groups are defined by the combination of being single (panel (a) and (b)) or in a couple (panel (c) and (d)), having children (panel (a) and (c)) or not (panel (c) and (d)) and the household income per adult in euros (x-axis). Variables are constructed using cross-sectional administrative weights.

Figure 14: $\Delta \ln$(1-MTR) by treatment group, for base year 2011 and 2015, with only wage subsidies and using the alternative definition



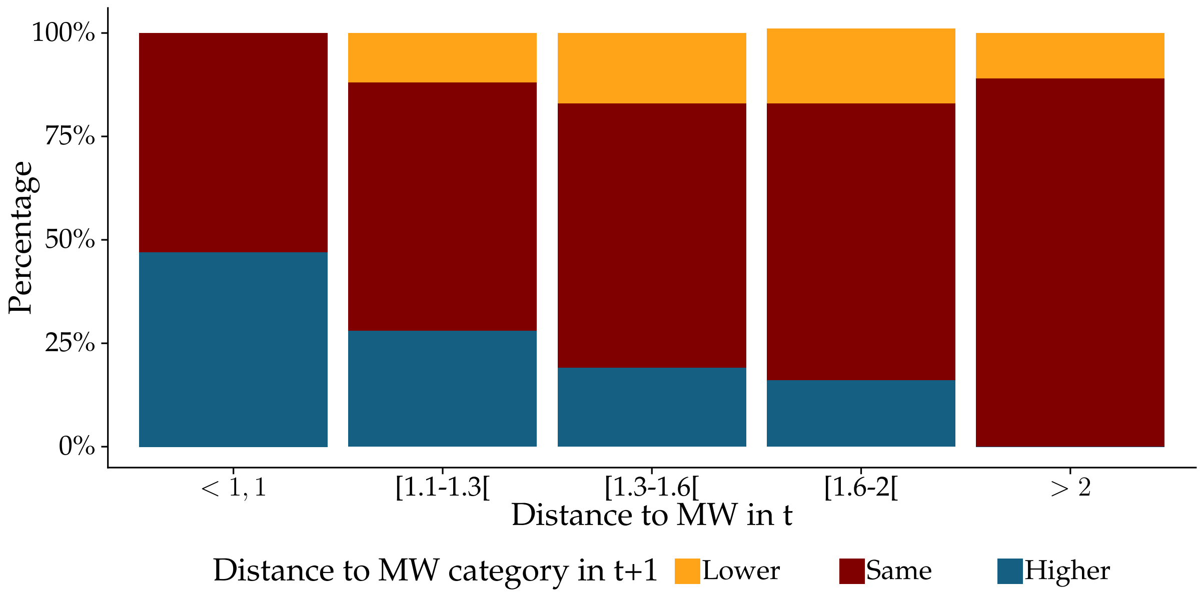
Notes: The figure plots the two-year change in the log of the net-of-tax rate at the treatment group-level, for base year 2011 and 2015 and considering only wage subsidies. The tax groups are defined by the combination of being single (panel (a) and (b)) or in a couple (panel (c) and (d)), having children (panel (a) and (c)) or not (panel (c) and (d)) and the household income per adult in euros (x-axis). Variables are constructed using cross-sectional administrative weights.

Figure 15: Distribution of the distance to the minimum wage over time



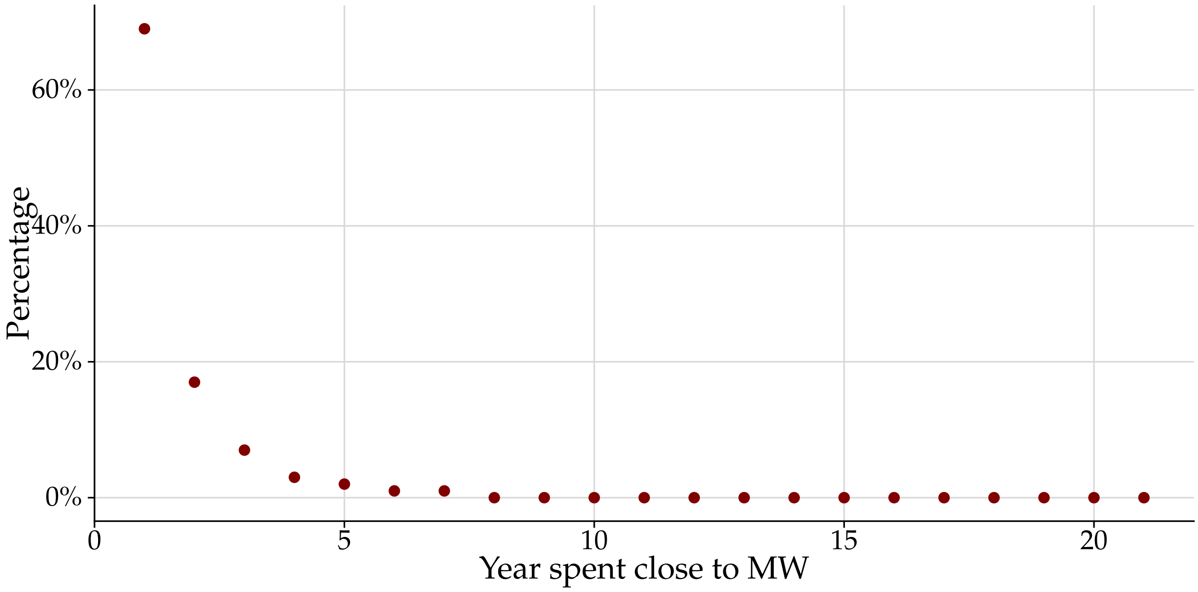
Notes: The figure plots the percentage of the salaried population in bins of distance to the minimum wage over time. Data and methodology is available here.

Figure 16: Transition between distances to the minimum wage



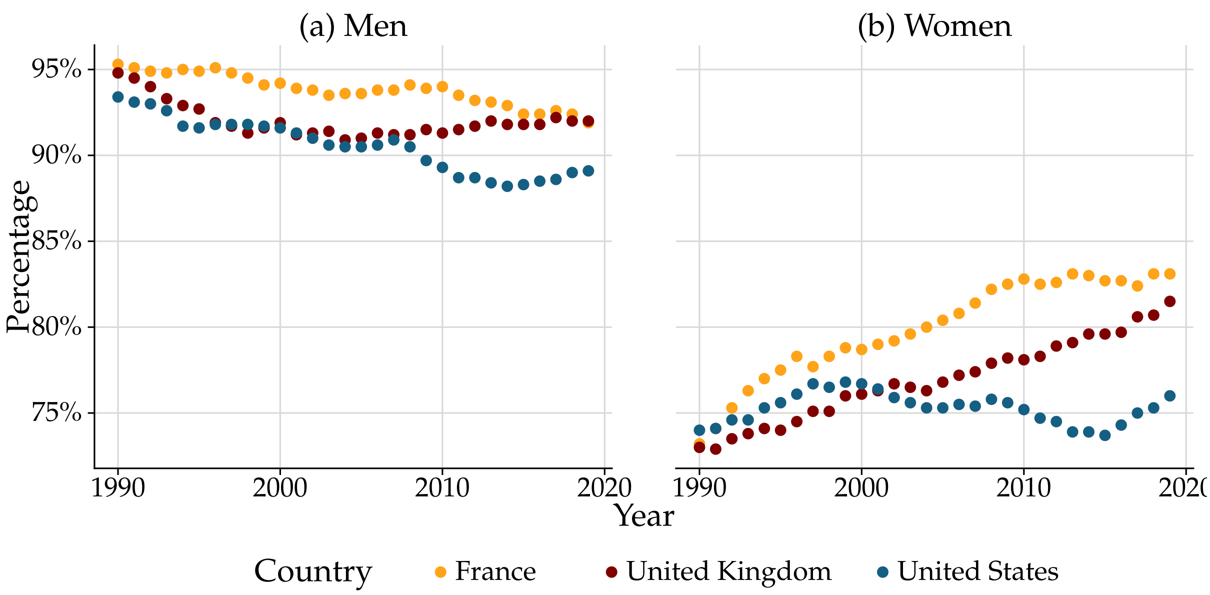
Notes: The figure plots the transition rates between bins of distance to the minimum wage between two consecutive years. Data and methodology is available here.

Figure 17: Years spent close to the minimum wage



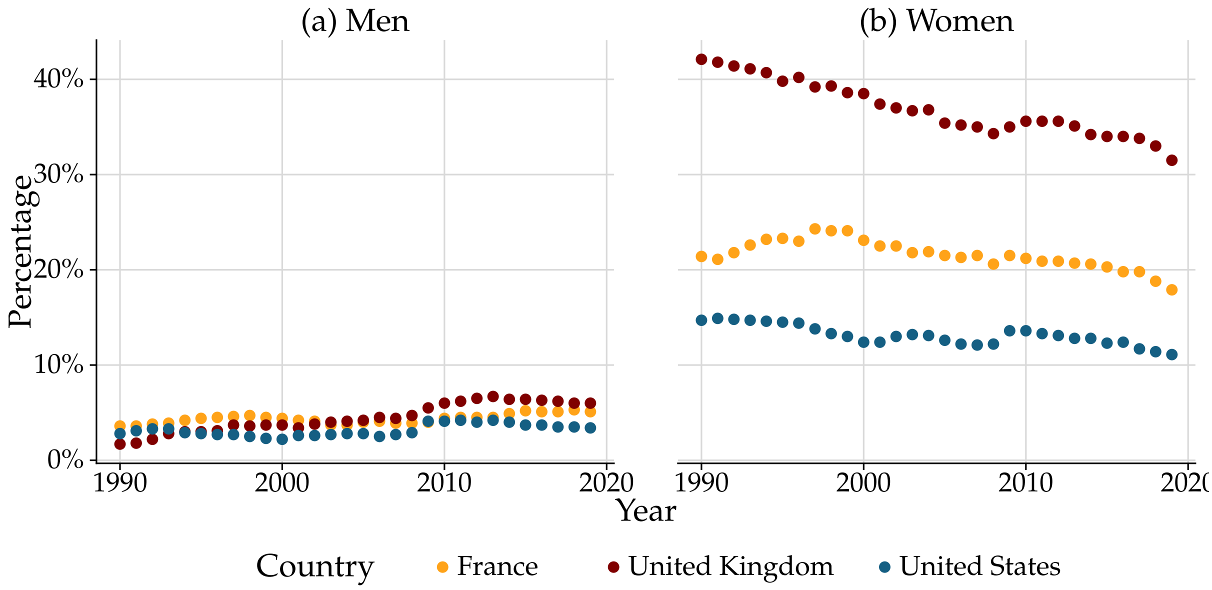
Notes: The figure plots the distribution of years spent close to the minimum wage, as defined by being below 1.1 times the minimum wage, conditional on starting a period at the minimum wage. Data and methodology is available here.

Figure 18: Labor force participation, by sex



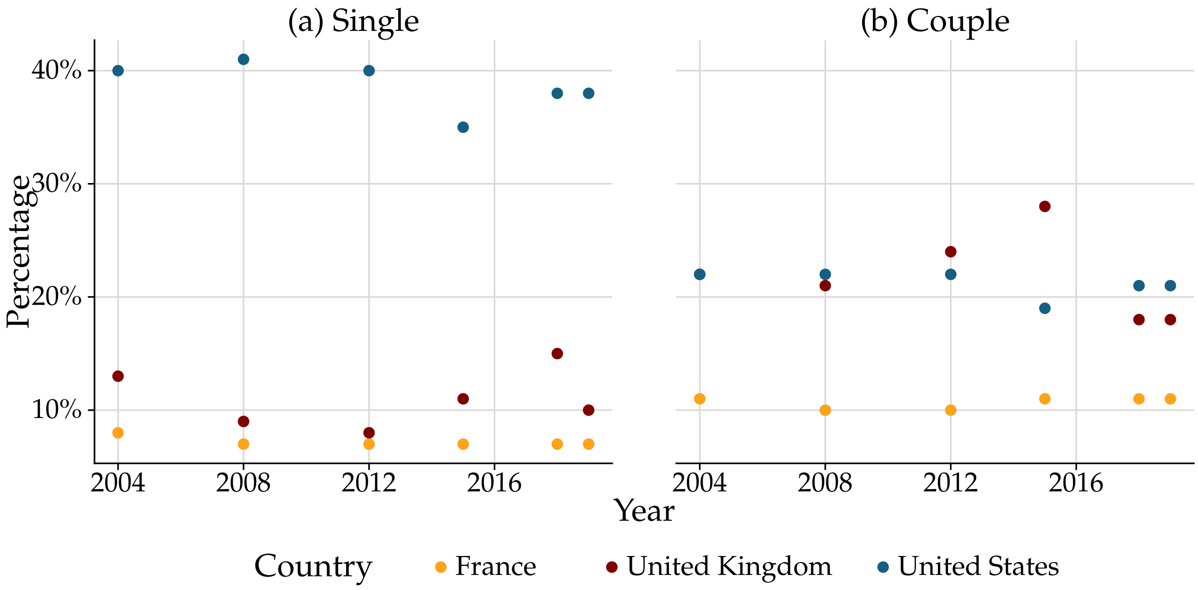
Notes: The figure plots times series of labor force participation rates for France, the United Kingdom and the United States at the yearly level. The labour force participation rates is calculated as the labour force divided by the total working-age population, separately for the men (panel (a)) and for women (panel (b)). The reference population is people aged 25 to 54. Data and methodology are available on OECD.Stat.

Figure 19: Share of employed in part-time employment, by sex



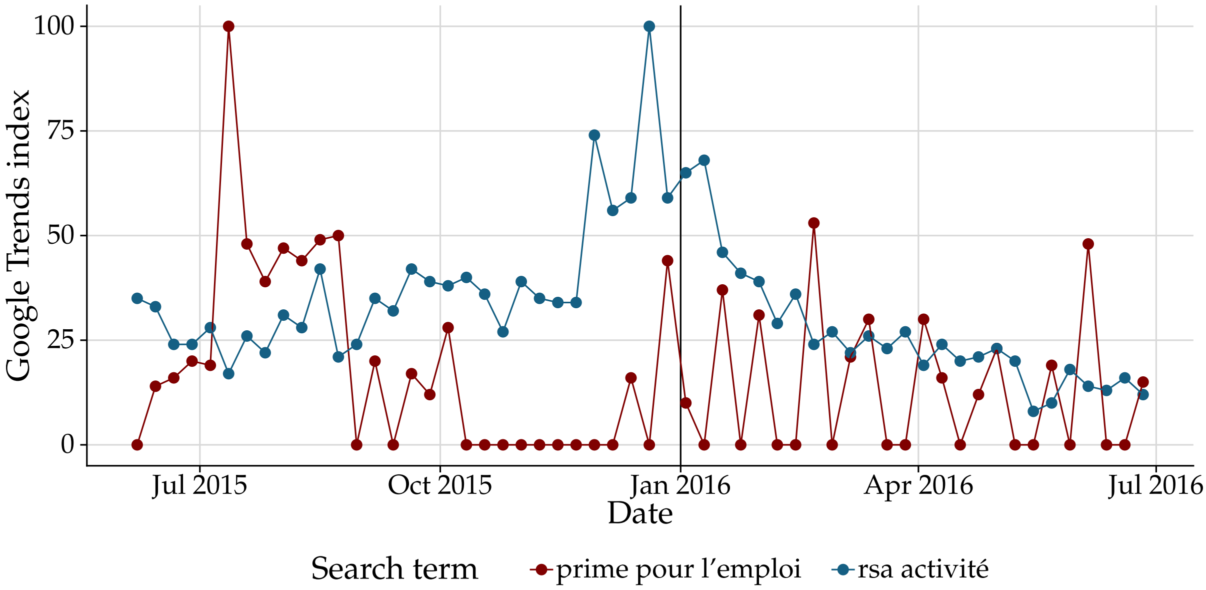
Notes: The figure plots times series of labor force participation rates for France, the United Kingdom and the United States at the yearly level. Part-time employment is defined as people in employment (whether employees or self-employed) who usually work less than 30 hours per week in their main job. Employed people are those who report that they have worked in gainful employment for at least one hour in the previous week or who had a job but were absent from work during the reference week while having a formal job attachment. The shares are calculated as the labour force divided by the total working-age population, separately for the men (panel (a)) and for women (panel (b)). The reference population is people aged 25 to 54. Data and methodology are available on OECD.Stat.

Figure 20: Childcare costs in net household income for parents with two children



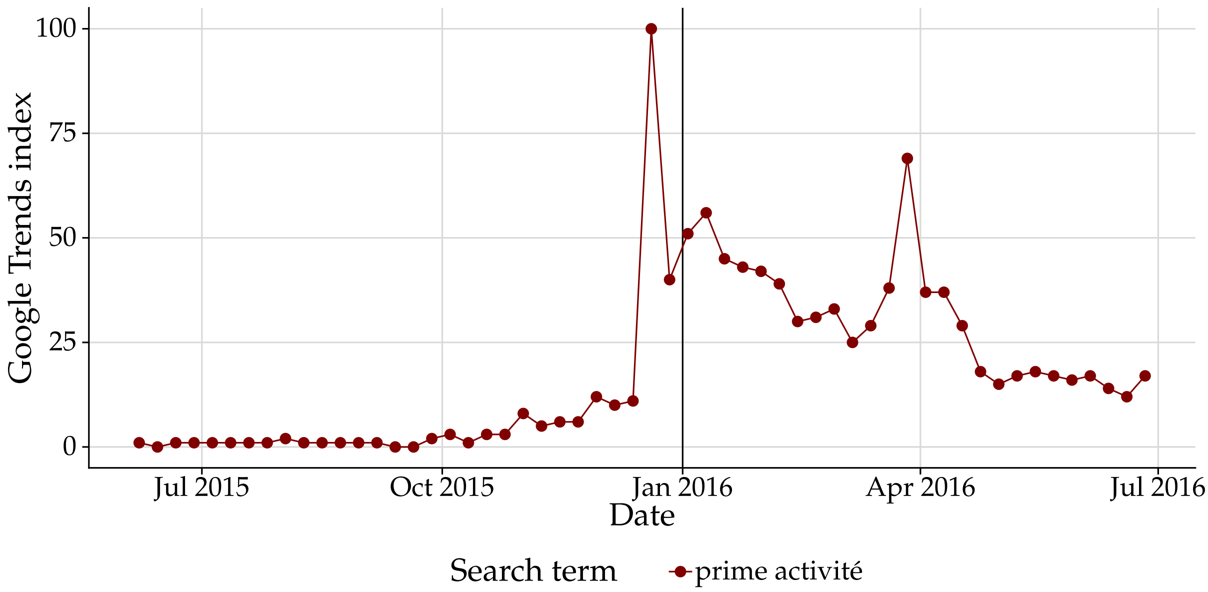
Notes: The figure plots times series of net childcare costs in percentage of net household income for France, the United Kingdom and the United States at the yearly level. Parents earn the average wage at full-time work. The net childcare cost is the difference between the gross childcare fee and childcare benefits (any types). Panel (a) plots the percentage for single parents and panel (b) for couples. Data and methodology are available on OECD.Stat.

Figure 21: Evolution of the Google Trends index for the two before-reform wage subsidy programs over time



Notes: The figure plots times series for the Google Trends index for the two before-reform wage subsidy programs. More precisely, the two search terms are "prime pour l'emploi" and "rsa activité". Each index is the result of a normalization between 0 and 100 of the number of search for these terms. 100 indicates the day where the number of search are the highest. The vertical black line is the date of implementation of the reform. Data and methodology are available on Google Trends.

Figure 22: Evolution of the Google Trends index for the after-reform wage subsidy program over time



Notes: The figure plots times series for the Google Trends index for the two before-reform wage subsidy programs. More precisely, the search term is "prime activité". Each index is the result of a normalization between 0 and 100 of the number of search for these terms. 100 indicates the day where the number of search are the highest. The vertical black line is the date of implementation of the reform. Data and methodology are available on Google Trends.

C Data and Variable Construction

C.1 Data

The main dataset is the *Echantillon Démographique Permanent* (EDP henceforth), a large individual-level panel following approximatively 4% of the French population at random. Individuals are sampled based on their date of birth (first 4 days of every quarter in a calendar year). It collects data on individuals from multiple data sources. It includes the census, matched employer-employee data[16] and other administrative datasets such as income tax returns and information from social agencies. Note that the EDP does not only collect individual-level variables, but also household-level variables. Only sampled individuals have a unique identifier over time. Other members of the household do not.

The census provides detailed information on individuals' demographics: birth and death date, place of birth and death, and sex among others. I use it to define age and gender of individuals.

The employer-employee dataset contains detailed information on labor earnings, number of hours worked, type of contract, occupation and sector. The main activity and the main firm refer to the situation where an individual has multiple employment spells in different firms in a given year. In this case, the type of contract, occupation, sector and number of workers are the ones from the longest spell (or with the highest labor earnings if two spells are of equivalent time). I use it to define the number of hours worked in a given year, the number of days worked, the type of contract (full-time/part-time) and the hourly wage rate.

Finally, I have additional information on individual and household revenues from income tax returns (labor earnings, capital income, unemployment benefits, taxes and tax credits) and information on welfare benefits claimed by individuals through social agencies. I use individual and household incomes, as well as other socio-economic characteristics, to compute the wage subsidy and disposable income for individuals. I

[16] *DADS* (*Déclaration annuelle des données sociales*) database.

also use it to define the place of residence.

C.2 Construction of the Sample

To construct the dataset used for the empirical analysis in this paper, I implement the following steps:

- **Step 1**: selection of the population
- **Step 2**: simulation of the tax and benefit system
- **Step 3**: construction of the estimation samples
- **Step 4**: data aggregation and binning.

I describe each step in details in the following sections.

C.2.1 Step 1: selection of the population

I start by selecting the population on a yearly basis, considering only the cross-sectional dimension of the data. For years between 2011 and 2017, I consider individuals filling at least an income tax return as primary or secondary filer, not dying within the year, with a place of residence in the French metropolitan area and between 25 and 55 years of age. For years before 2011 (included), I drop individuals getting married within the year. Indeed, individuals were required to file multiple tax returns, one for each specific marital status period, making it complicated to simulate their tax rates. Specifically, I define couples are individuals who are in civil unions or married. For years between 2011 and 2017, I also drop individuals getting divorced or experiencing a death of the spouse for similar reasons.

Next, I restrict the population to individuals for which I can identify a single and unique household in a given year. The tax household definition used in the income tax returns is different from conventional household definition. More precisely, household can have multiple tax households. For example, if two occupants who are single live in the same dwelling, they count for one household but two tax households. Since my data aggregate all tax households at the household level, I restrict my population to individuals for which their tax household is also their unique household. To do so, I start by restricting my sample to households with a unique combination of primary and secondary filers. Second, I keep households where being in a couple is equivalent to being married or in a civil union. Hence, the number of tax filers is two if in a couple and one if single. Third, I delete individuals present in at least two different households.

Finally, I restrict my sample to households for which I have information on the number of dependants. Since my analysis is about salaried workers, I also discard households with self-employed revenues, retirement incomes and foreign incomes.

C.2.2 Step 2: simulation of the tax and benefit system

In order to construct marginal and average tax rates, as well as the corresponding simulated instruments, I use a tax simulator of the French tax and benefit system. Openfisca is freely available online and is implemented in Python.

Before going to the simulation, I estimate the counterfactual labor earnings for the population not working at all in a given year (i.e reporting zero labor earnings). I explain the procedure in details in subsection C.3.

Openfisca requires several information on incomes and socio-economic characteristics. At the indvidual level, I use the marital status (single/divorced/widowed/civil union/married), labor earnings (with the counterfactual equivalent for the non-working population), unemployment insurance, retirement income, alimonies and effective number of hours worked. At the household level, I use number of dependents, real estate incomes, capital incomes, the received housing benefits and the corresponding area.

Next, I simulate the disposable income and the wage subsidy for each year separately, under the assumption of the full take-up of welfare benefits. To do so, I make two assumptions. First, I split yearly labor earnings equally between each month within a year. Second, I compute wage subsidies and welfare benefits only depend on monthly earnings. This is a simplifying assumption because two of the wage subsidy programs (the RSA activité and the Prime d'Activité) are based on the average earnings in the past three months. Since Openfisca only allows static simulations, I do not take into account this rule. It gives a reasonable first-order approximation of the wage subsidy under full take-up at the yearly level.

Finally, I compute two sets of marginal and average tax rates. I first consider the marginal and average tax rate with the full tax and benefit system. Second, I compute the tax rates only for the wage subsidies. Exact formulas are presented in subsection C.3.

C.2.3 Step 3: construction of the estimation samples

At this stage, I focus on sampled individuals which have a unique identifier over time. I keep observations based on the consistency between observed net labor earnings and taxable labor earnings, as the taxable labor earnings should be higher than its net counterpart because it includes non-deductible social contributions. I keep individuals both individuals with non-zero labor earnings for which this condition is verified and individuals with zero labor earnings. Monetary variables are expressed in real values, base 2011.

Then, I construct two version of the data. First, I consider only the cross-sectional version of the data. It is used for the construction of the shares, in order to maximize sample size and thus have a representative distribution of individuals across labor markets and treatment groups. I keep individuals with an hourly wage strictly below 14 euros. Second, I consider a panel version version of the data. I match variables from year t to year $t-h$ for each individual observed in the two periods. In my baseline analysis, $h=2$. It enables me to construct two sets of individual-level shocks: changes in tax rates with respect to the wage subsidy (weighted by the initial share of wage subsidy in total taxes and benefits) and changes in tax rates with respect to the full tax and benefit system. Exact formulas are presented in subsection C.3. I keep individuals with an hourly wage strictly below 14 euros in $t-h$.

C.2.4 Step 4: data aggregation and binning

In the final step I aggregate shocks at the labor market level, indexed by m. To do so, I construct three sets variables. First, the log-growth between year t and $t-h$ for several outcomes (hourly wage rage, number of hours worked, total labor earnings), denoted by $\Delta \ln(Y_{m,t})$. Second, the share of treatment group g in the total labor supply at the market level in year t, $S_{m,g,t}$. Finally, the log-growth in net-of-tax rate and participation rate at the treatment group level between year t and $t-h$. Details about the construction of these variables is available in subsection C.3.

C.3 Variable Construction

C.3.1 Counterfactual labor earnings

In order to properly classify households into treatment groups, I need to know their labor earnings. For individuals reporting zero earnings within a given year, this is not possible to have a good approximation for it directly in the data. To circumvent this problem, I build on the wage subsidy literature to predict labor earnings for the *sampled* population that is not working. Specifically, I only estimate the counterfactual labor earnings for individuals that have a unique identifier over time, and not for their spouse for example.

I follow a similar procedure as Kleven (2019). I start by estimating the relationship between the log of labor earnings $\ln(Y_i)$ and a set of fixed-effects, conditional on having positive earnings. The estimation is done on the cross-sectional version of the data to maximize sample size, and separately for each year:

$$\ln(Y_i) = \alpha_{sex} + \alpha_{ms} + \alpha_{age} + \alpha_{res} + \lambda_{sex,ms} + \lambda_{sex,age} + \lambda_{ms,age} + \lambda_{res,sex} + \lambda_{res,ms} + \lambda_{res,age} + \epsilon_i$$

where sex is a categorical variable for men/women, ms is a categorical variable for the marital status (single/divorced/widowed/civil union/married), res is a categorical variable for the place of residence based on the French *departements* and age is a categorical variable containing the age of individuals. The estimated coefficients are used to predict labor earnings for the non-working population: $\exp\left(\widehat{\ln(Y_i)}\right)$.

Finally, I follow the same procedure for the number of hours worked and the hourly wage rate.

C.3.2 Tax rates

Marginal tax rate Marginal tax rates are computed by simulating the tax system twice. First, with the observed individual and household values. Second, by adding 100 euros to labor earnings for sampled individuals. I define the simulated marginal tax rate using the following formula:

$$\text{MTR}_{i,t}(y_{i,t}) = \frac{T_{i,t}(y_{i,t}+100,\cdot) - T_{i,t}(y_{i,t},\cdot)}{100}$$

where $y_{i,t}$ is the labor earnings for individuals i in year t. I apply directly this formula for the wage subsidy, but use an alternative formulation for the marginal tax rate with respect to the full tax and benefit system:

$$\text{MTR}_{i,t}(y_{i,t}) = 1 - \frac{Z_{i,t}(y_{i,t}+100,\cdot) - Z_{i,t}(y_{i,t},\cdot)}{100}$$

where $Z_{i,t}$ is the disposable income. The equivalence with the previous formula comes from $T_{i,t}(y_{i,t}) = R_{i,t}(y_{i,t}) - Z_{i,t}(y_{i,t})$ and $T_{i,t}(y_{i,t}+100) = R_{i,t}(y_{i,t}+100) - Z_{i,t}(y_{i,t}+100) = R_{i,t}(y_{i,t}) + 100 - Z_{i,t}(y_{i,t}+100)$, with $R_{i,t}$ the pre-redistribution income.

The counterfactual marginal tax rate used for the simulated instrument is computed using the same method:

$$\text{MTR}_{i,t}(y_{i,t-h}) = \frac{T_{i,t}(k_{t,t-h}y_{i,t-h}+100,\cdot) - T_{i,t}(k_{t,t-h}y_{i,t-h},\cdot)}{100}$$

where $y_{i,t-h}$ is the labor earnings for individuals i in year $t-h, h>0$. All incomes are multiplied by the inflation coefficient $k_{t,t-h}$ between period t and $t+h$ based on CPI series computed by the INSEE. I apply the same procedure as before to compute the marginal tax rate with respect to the full tax and benefit system.

Average tax rate Average tax rates are computed by simulating the tax system twice. First, with the observed individual and household values. Second, by considering labor earnings equal to zero for sampled individuals. I define the simulated tax rate using the following formula:

$$\text{ATR}_{i,t}(y_{i,t}) = \frac{T_{i,t}(y_{i,t}, \cdot) - T_{i,t}(0, \cdot)}{y_{i,t}}$$

where $y_{i,t}$ is the labor earnings for individuals i in year t. I apply directly this formula for the wage subsidy, but use an alternative formulation for the marginal tax rate with respect to the full tax and benefit system:

$$\text{ATR}_{i,t}(y_{i,t}) = 1 - \frac{Z_{i,t}(y_{i,t}, \cdot) - Z_{i,t}(0, \cdot)}{y_{i,t}}$$

where $Z_{i,t}$ is the disposable income. The equivalence with the previous formula comes from $T_{i,t}(y_{i,t}) = R_{i,t}(y_{i,t}) - Z_{i,t}(y_{i,t}) = R_{i,t}(0) + y_{i,t} - Z_{i,t}(y_{i,t})$ and $T_{i,t}(0) = R_{i,t}(0) - Z_{i,t}(0)$, with $R_{i,t}$ the pre-redistribution income.

The counterfactual average tax rate used for the simulated instrument is computed using the same method:

$$\text{ATR}_{i,t}(y_{i,t-h}) = \frac{T_{i,t}(k_{t,t-h} y_{i,t-h}, \cdot) - T_{i,t}(0, \cdot)}{k_{t,t-h} y_{i,t-h}}$$

where $y_{i,t-h}$ is the labor earnings for individuals i in year $t-h, h > 0$. All incomes are multiplied by the inflation coefficient $k_{t,t-h}$ between period t and $t+h$ based on CPI series computed by the INSEE. I apply the same procedure as before to compute the marginal tax rate with respect to the full tax and benefit system.

C.3.3 Individual-level shocks

I compute the log-growth in net-of-tax and participation tax rates at the individual-level (indexed by i) for two sets of variables: with respect to the full tax and benefit system and with respect to the wage subsidy. I explain below the construction for the observed shocks, but the procedure remains the same for the simulated instruments.

Full tax and benefit system The log-growth of the net-of tax rate is defined as follows:

$$\Delta \ln\Big(1 - \text{MTR}_{i,t}^{full}\Big) = \ln\left(\frac{1 - \text{MTR}_{i,t}^{full}}{1 - \text{MTR}_{i,t-h}^{full}}\right)$$

where $\text{MTR}_{i,t}^{full}$ is the marginal tax rate in year t. The log-growth rate of the participation rate is defined by:

$$\Delta \ln\Big(1 - \text{ATR}_{i,t}^{full}\Big) = \ln\left(\frac{1 - \text{ATR}_{i,t}^{full}}{1 - \text{ATR}_{i,t-h}^{full}}\right)$$

where $\text{ATR}_{i,t}^{full}$ is the average tax rate in year t.

Wage subsidy The log-growth of the net-of tax rate is defined as follows:

$$\Delta \ln\big(1 - \text{MTR}_{i,t}^{ws}\big) = \ln\left(\frac{1 - \text{MTR}_{i,t}^{ws}}{1 - \text{MTR}_{i,t-h}^{ws}}\right) \times \frac{1 - \text{MTR}_{i,t-h}^{ws}}{1 - \text{MTR}_{i,t-h}^{full}}$$

where $\text{MTR}_{i,t}^{ws}$ is the marginal tax rate in year t. It is weighted by the share of the marginal tax rate with respected to the wage subsidy in the full tax and benefit marginal tax rate, such that $\sum_k \Delta \ln\big(1 - \text{MTR}_{i,t}^{k}\big) = \Delta \ln\Big(1 - \text{MTR}_{i,t}^{full}\Big)$. The same logic applies for the log-growth of the participation tax rate:

$$\Delta \ln\big(1 - \text{ATR}_{i,t}^{ws}\big) = \ln\left(\frac{1 - \text{ATR}_{i,t}^{ws}}{1 - \text{ATR}_{i,t-h}^{ws}}\right) \times \frac{1 - \text{ATR}_{i,t-h}^{ws}}{1 - \text{ATR}_{i,t-h}^{full}}$$

C.3.4 Shares

To obtain $s_{m,g,t}$, the share of treatment group g in total labor supply of labor market m in year t, I use the cross-section of the data. For each year, I use the following formula:

$$S_{m,g,t} = \frac{L_{m,g,t}}{L_{m,t}} = \frac{\sum_i a_{i,m,g,t} h_{i,m,g,t}}{\sum_g \sum_i a_{i,m,g,t} h_{i,m,g,t}}$$

where $h_{i,m,g,t}$ is the number of hours worked by individual i and $a_{i,m,g,t}$ is an administrative (frequency) weight for the sample to be nationally representative.

C.3.5 Outcomes

I have three outcomes in this paper at the labor market level m: the log-growth in the hourly wage rate, the log-growth in the total number of hours worked and the log-growth in total earnings. I use the panel data version of the sample to compute them.

I start with the definition of the change in employment by taking the log-growth rate of the total number of hours worked in a given labor market:

$$\Delta \ln(L_{m,t}) = \ln\left(\sum_i h_{i,m,t}\right) - \ln\left(\sum_i h_{i,m,t-h}\right)$$

with $h_{i,m,t}$ the number of hours worked for individuals i in year t. Note that labor market m is defined in $t-h$ and kept fixed for t. I winsorize the top 1% of the distribution for both $h_{i,m,t}$ and $h_{i,m,t-h}$ by initial year $t-h$ to avoid extreme values.

Then, I construct a measure of wage rate by computing the hours-weighted wage rate. Formally, this measure for a given labor market in a given year is given by:

$$w_{m,t} = (\sum_i a_{i,m,t} w_{i,m,t} h_{i,m,t}) / (\sum_i a_{i,m,t} h_{i,m,t})$$

with $w_{i,m,t}$ the hourly wage rate and $h_{i,m,t}$ the number of hours workerd by individual i. Then, I take the log-growth rate of the total number of hours worked in a given labor market:

$$\Delta \ln(w_{m,t}) = \ln(w_{m,t}) - \ln(w_{m,t-h})$$

where I winsorize the bottom 1% of the distribution for both $w_{i,m,t}$ and $w_{i,m,t-h}$ by initial year $t-h$ to avoid extreme values.

Finally, I compute the log-growth in total earnings at the labor market level as follows:

$$\Delta \ln(w_{m,t}L_{m,t}) = \Delta \ln(L_{m,t}) + \Delta \ln(w_{m,t})$$

taking into account the winsorization on the hourly wage and number of hours described above.

C.3.6 Treatment group-level shocks

The goal is to compute the average log-growth for the net-of-tax and participation tax rates at the treatment group-level g. Consider the shock $\theta_{g,t}$. It is a hours-weighted aggregation of the shocks at the individual level:

$$\theta_{g,t} = \sum_i \frac{a_{i,g,t-h}h_{i,g,t-h}}{\sum_i a_{i,g,t-h}h_{i,g,t-h}} \times \theta_{i,g,t}$$

where $h_{i,g,t-h}/\sum_i h_{i,g,t-h}$ is the share of hours worked by individual i in the treatment group total number of hours in the initial year $t-h$. Note that consistent with the construction of outcomes, I winsorize the top 1% of the distribution for both $h_{i,g,t-h}$ by initial year $t-h$ to avoid extreme values. For similar reasons, I winsorize the top 5% and bottom 5% of the distribution of shocks by initial year $t-h$.

I follow the same procedure to construct the instruments at the treatment group-level, replacing the observed shock $\theta_{i,g,t}$ by the simulated instruments $\theta^{sim}_{i,g,t}$.