June 2024

FONDAZIONE ENI ENRICO MATTEI

Working Paper

16.2024

Income taxation and labour response. Empirical evidence from a DID analysis of an income tax treatment in Italy

Bruno Bosco, Carlo Federico Bosco, Paolo Maranzano

Income taxation and labour response. Empirical evidence from a DID analysis of an income tax treatment in Italy

Bruno Bosco (Department of Economics, Management and Statistics, University of Milano-Bicocca), **Carlo Federico Bosco** (University of Pavia), **Paolo Maranzano** (Department of Economics, Management and Statistics, University of Milano-Bicocca and Fondazione Eni Enrico Mattei)

Summary

This paper uses the Italian income tax treatment of 2006/7 as a quasi-natural tax experiment to offer some fresh empirical evidence on how labour supply responds to exogenous income tax hikes. We adopt the identification strategy based on TWFE panel data Difference-in-Differences (DID) model to define the correct statistical framework of the study, and to benefit from the specific features of the above tax experiment, namely homogeneity and contemporaneity of the treatment. Results show that the extensive negative adjustments of various response variables measuring the supply of labour services offered by treated taxpayers are statistically significant, rapid, and strong but not long-time lasting. Not surprisingly, we also find that that treated families reduce in a similar manner their consumption with respect to families in the control groups. Analogous adjustment responses to tax hikes characterise the growth of per-capita regional GDP. The estimated aggregate effects of tax hikes are further compared with the spatial-temporal patterns observed for every response variable in treated and untreated regions.

Keywords: Income Taxation, extensive labour supply change, TWFE Panel Data DID, convergence tests, taxation and regional growth

JEL classification: C10, C18, C21 H2, E2, E32, E62, C23, C26

Corresponding Author:
Paolo Maranzano
Assistant Professor
Department of Economics, Management and Statistics, University of Milano-Bicocca
Piazza dell'Ateneo Nuovo 1, Milan (Italy)
e-mail: paolo.maranzano@unimib.it

Income taxation and labour response. Empirical evidence from a DID analysis of an income tax treatment in Italy

Bruno Bosco, DEMS, Department of Economics, Management and Statistics (DEMS), Piazza dell'Ateneo Nuovo, 1, Milano, 20126
bruno.bosco@unimib.it

Carlo Federico Bosco, University of Pavia, Italy
fede.bosco28@gmail.com

Paolo Maranzano*, University of Milano-Bicocca, Department of Economics, Management and Statistics (DEMS), Piazza dell'Ateneo Nuovo, 1, Milano, 20126 & Fondazione Eni Enrico Mattei (FEEM), Corso Magenta, 63, Milano, 20123
https://orcid.org/0000-0002-9228-2759
paolo.maranzano@unimib.it (*corresponding author)

Abstract

This paper uses the Italian income tax treatment of 2006/7 as a quasi-natural tax experiment to offer some fresh empirical evidence on how labour supply responds to exogenous income tax hikes. We adopt the identification strategy based on TWFE panel data Difference-in-Differences (DID) model to define the correct statistical framework of the study, and to benefit from the specific features of the above tax experiment, namely homogeneity and contemporaneity of the treatment. Results show that the extensive negative adjustments of various response variables measuring the supply of labour services offered by treated taxpayers are statistically significant, rapid, and strong but not long-time lasting. Not surprisingly, we also find that that treated families reduce in a similar manner their consumption with respect to families in the control groups. Analogous adjustment responses to tax hikes characterise the growth of per-capita regional GDP. The estimated aggregate effects of tax hikes are further compared with the spatial-temporal patterns observed for every response variable in treated and untreated regions.

Keywords: Income Taxation; Extensive labour supply change; TWFE Panel Data DID; Convergence tests; Taxation and regional growth;
JEL codes: C10, C18, C21 H2, E2, E32, E62, C23, C26

1. Introduction, motivation, and content of the paper

Labour supply response to income taxes is a fundamental and debated issue at both theoretical and empirical level. Using Italy as a case study, we contribute to the existing empirical literature and present the results of a two-way fixed effect (TWFE) panel data Difference-in-Differences (DID) regression model showing how labour supply reacts to exogenous regional income tax hikes. Despite Italy is not a federal country, she may

provide a good case for studying the labour and other growth-related variables reaction to income tax hikes when this only happens in some regions. Since 1970 (the year of the creation of the Italian Ordinary Regions as an implementation *in slow motion* of the 1948 democratic Constitution of the Republic) and particularly in the last years, Italian regions have accumulated large tax and expenditure powers, which must have affected somehow the behaviour of their taxpayers in their territories and the growth trajectories. Consequently, regional authorities may use different local income tax rates and adopt different tax expenditure measures. Hence, income tax heterogeneity exists, and exogenous regional income tax shocks may have affected the supply of market labour services and regional growth paths over the years.

We conduct the empirical study using a large and disaggregated Italian panel data containing regional and national tax and economic data and, to define the above DID empirical model, we exploit a quasi-natural tax experiment that occurred in Italy at the end of 2006. The characteristics of the tax treatment (timing and treatment assignment) and the time pattern that response variables follow before the treatment, as well as the absence of confounding factors, have important implications for our empirical approach. They make the statistical framework of our analysis consistent with the identification requirements needed for the efficient estimation of the Average Treatment Effect upon Treated (ATET) discussed by Roth, Sant'Anna, Bilinski, and Poe (2023, p. 2220) for the application of the above OLS-based method under treatment homogeneity. The treatment affected a subset of Italian regions (treatment group) from 2007 onwards and consisted in an exogenously imposed increase of their regional income tax rates with respect to the rates of the other regions (control group) which remained unchanged. Hence, the 2006/7 central government policy defines the framework of a quasi-natural tax experiment, fully described in Appendix 1, involving the taxpayers of five (not-randomly-chosen) treated regions vs the taxpayers of the remaining 15 untreated control regions. The goal of the estimation is to evaluate if exogenous differences in income taxation resulting from the centrally imposed income tax policy generated statistically significant responses of the labour supply of treated subjects. Our TWFE DID estimates of the ATET permit to study the timing of the transactions involved in income tax responses (reducing or delaying over time labour supply) and to identify other real responses such as per capita regional income growth and consumption. Results show that the extensive negative adjustments of various response variables measuring the supply of labour services of treated taxpayers are statistically significant, rapid, and strong but not long-time lasting. Important for tax policy are the results concerning the response to tax hikes of Self-Employed people –possibly the taxpayers mostly affected by the 2006/7 tax treatment. Our DID estimations show that the share of Self-Employed workers is reduced by the increase of income taxation. Not surprisingly, we also find that treated families reduce in a similar manner their consumption with respect to families in the control groups. Analogous adjustment responses to tax hikes characterise the growth of regional per-capita GDP. Altogether, our results show that exogenously imposed differences in income tax rates affect the behaviour of regional economic variables related to growth, starting from labour supply. The clearest effect is the negative and statistically significant impact of the income tax increase on Self-Employment, per-capita regional GDP growth and family consumption. The same result emerges using other labour response variables (e.g. the annual New VAT Certificates, necessary for

conducting self-employment activities). In general, the exogenous tax effect is time lasting but not long lasting. After treatment, adjustments of the labour supply to a long-run trend are observable, on average, 3 years after the introduction of the treatment but anticipatory effects are generally absent.

The paper is structured as follows. Relevant literature is discussed in section 2. The tax factors considered in this study are described in section 3, together with the short narrative of the Italian 2006/7 tax experiment. Details on the latter are in Appendix 1. Section 4 contains the discussion of the DID identification hypothesis in terms of a TWFE model as well as a statistical survey of the data with pre and post treatment plots of the response variables. It also contains the empirical results, which are commented in comparison with the results of the previous international literature. We also conduct a robustness analysis by using provincial, rather than regional, response variables measuring other indicators of labour activities and a larger data set. The robustness analysis makes use of an augmented version of the previous basic TWFE DID model and includes cofactors clustered at both provincial and regional levels. Results support the findings of the basic model that tax hikes negatively affect labour supply. We take advantage of the robustness analysis for a deeper discussion of the parallel trend and anticipation effects tests. Policy implications are also discussed. Section 5 includes some additional test procedures permitting to implement what we call a partial *falsification test* for the presence of pre-treatment parallel trends in our panel data response variables based on convergence analysis. Section 6 briefly concludes. The spatial-temporal patterns observed for every response variable in treated and untreated regions are illustrated in Appendix 2. Altogether, Appendix 1 and 2 form the supplementary material annexed to the paper.

2. Literature review

The empirical literature studying the relationships between income taxation and variables conducive to growth generally finds indications of heterogeneous aggregate effects of income tax changes. Zidar (2019, p. 1440) and others, for instance, provide evidence that lower income households tend to have higher marginal propensities to consume (Dynan, Skinner, & Zeldes, 2004; Jappelli & Pistaferri, 2010; Johnson, Parker, & Souleles, 2006; McCarthy, 1995; Parker, 1999; Parker, Souleles, Johnson, & McClelland, 2013) and so they respond to tax policy by increasing consumption more than high-income groups. In analogous way, labour supply and other growth-related variables of different income groups may respond to income tax policy in an opposite way. Focusing on labour supply, one may notice that on the extensive margin for lower-income groups, Eissa and Liebman (1996) and Meyer and Rosenbaum (2001) show that in the USA the Earned Income Tax Credit (a refundable tax credit enacted in 1975 for low-to-moderate-income working individuals and couples, particularly those with children) has increased labour force participation. Other authors (e.g., Romer & Romer, 2010; Saez, Matsaganis, & Tsakloglou, 2012) show that for high-income earners there is some evidence that the efficiency costs of raising taxes on top-income taxpayers expressed in terms of labour supply and other margins may be limited or can be offset by shifting in the timing or form/source of income acquisition (Auerbach & Siegel, 2000; Goolsbee, 2000). Ziliak and Kinierser (2005)

find that consumption and worked hours are direct complements in utility and increase (both) with a compensated increase in net wage. Other studies of labour response to tax changes analyse the potential adverse base effects of tax hikes such as tax avoidance, outright tax evasion, and a general reduction in economic activity (Piketty, Saez, & Stantcheva, 2014). They conclude that income tax can be responsible not only of a reduction of labour services supplied but also of a general contraction of the economic activity.

When territorial aspects are accounted for, the analysis considers the tax-induced mobility of inventors and discusses whether income tax rates can be employed by local authorities to attract highly qualified foreigners. The literature shows that local income taxes may affect local decisions (such as regional/local labour supply or residential location) and evaluates the possible consequences of these decisions on the level and geographical distribution of productive activities and on local technological spill over (Widmann, 2023). For instance, Akcigit, Baslandze, and Stantcheva (2016) find that "prolific" inventors migrate between countries in response to changes in personal top income tax rates and obtain an estimate of the elasticity of the number of foreign inventors (in a country) with respect to the top net-of-tax rate close to one. Moretti and Wilson (2017) examine the mobility responses of inventors to changes in personal top income tax rates across US states and find a corresponding elasticity with respect to the top net-of-tax rate of 1.8. In addition, since the market value of material goods, and its geographical distribution, may be affected by income taxation, other studies show that local income tax rates can generate tax externalities of various nature across local jurisdictions (Esteller-Moré & Solé-Ollé, 2002).

Implementing the DID approach, Jakobsen, Jakobsen, Kleven, and Zucman (2019) provide causal evidence of the impact of net-of-tax rate on wealth accumulation in Denmark whereas Kleven, Landais, and Saez (2013) study the causal relationship between tax rates and a particular subset of the population of migrants (football-players called super stars). Baskaran (2021) presents results from a set of municipalities who, in North Rhine-Westphalia, increased their local property and business tax rates. He studies the revenue and base effects of local property and business tax hikes in a generalized DID design. His results suggest that the property tax hikes had a revenue elasticity of unity even in the long run. He concludes that tax changes had no adverse effects on property tax bases. For the business tax, he finds no significant effects on revenues and tax bases.

2. The Italian income tax treatment of 2006/7

The measures adopted by the Italian government at the end of 2006 are a good example of a possible exogenous income tax treatment. In December 2006, the Italian central government decided to impose an increase - automatic and compulsory - of two regional income tax rates, namely IRAP tax rates and regional Income Tax Surcharge to regions who had increased, with respect to 2005, the deficit of their regional health care systems above a certain percentage. The other regions were not affected by the tax measure.[1] The

[1] If we call D ≡ Health Deficit = (Health expenditure – Health revenue)/Health expenditure, the criterion for eligibility was that $D_{it+1} \leq 1.07 \times D_{it}$. Those regions who did not respect that limit were assigned to the treated group. As discussed

measure was adopted to force those regions to collect additional local resources and use them *exclusively* to reduce the deficits of their health system. The government finally decided that the percentage increase to be used as a threshold for the inclusion in the treatment was the 7% of the previous year deficit. Five out of twenty Italian regions (Abruzzo, Campania, Lazio, Molise, and Sicily) fell into the treatment group. The content of the 2006 policy (described in details by Caruso and Dirindin (2019) is summarized below (a more complete description of the taxes involved in the treatment is in Appendix 1):

- IRAP tax rates. The ongoing standard IRAP tax rate existing in 2006 (4.5%) and employed by the five regions mentioned above was increased from the tax year 2007 onwards by almost a 1%. This implied that the actual tax rate for those regional taxpayers became equal to 5.25%. IRAP tax rates for taxpayers resident in other regions were not modified.
- Regional Income tax surcharge. From the tax year 2007 onwards, the five "treated" regions had to increase by 0.5% the ongoing rate of 1.23% they applied as a surtax on the personal income tax base (determined by the national Law) of their residents. Yet, since the five regions were under "controlled administration" by the central government, the surcharge rate was further increased by an additional 0.3%. The net total increase was equal to 0.8% leading to an income Tax Surcharge rate of 2.03%.
- A third element of the treatment is the following. Since the institution of the regional income surcharge, the Italian regions were permitted to allow deductions from the income surcharge payments instead of providing subsidies, vouchers, service vouchers and other social support measures determined by their regional legislation in favour of some taxpayers. These powers were taken away from the five treated regions. As a result, they could not *compensate* the surcharge increase with tax deductions.

For both taxations, the increase of the tax rates was compulsory, quantitatively relevant and finalized to collect an additional tax yield hopefully sufficient to cover the prolonged budget unbalances of their regional health care systems. No other uses were permitted. At the end of 2006, the law introducing the new tax regimes for the above five regions indicated that the measures had to be adopted for a 3-year period but the measures remained in force for the entire sample period used in this study (see below). Then, the condition of Irreversibility of Treatment applies: once a region became treated in 2007, that region will remain treated in the next periods.

The exogenous and unexpected increases of IRAP Tax and Personal Income surcharge mainly fell on Individual Business and Self-Employed Professionals (the main group of IRAP taxpayers within a region). Yet also the generality of treated regional taxpayers paying the national Income Tax were affected by the measure although less severely (they had to face only the increase of the regional personal income surcharge) unless in addition to paying the tax hike they had to forgo the above-mentioned tax deductions. Hence, based

in the text, the number and identity of regions who would integrate the treatment group was ex-ante uncertain. The idea was to punish regions who produced an "insufficient" fiscal effort in the year before the treatment and who had not reduced "appropriately" and autonomously the growth of their health deficit.

on the realization of a state of the word (the deficit of their regional health care systems) *determined when the measure was not in effect*, taxpayers, in the three-year period starting in 2007, were exposed to an exogenously determined tax increase of the main regional direct taxes.

Relevant for the applicability of our method is the circumstance that almost any Italian region could had been assigned to the treatment group. The health deficit was high and time increasing in the overwhelming majority of the regions and it was completely unknown ex-ante at what level the central government was going to put the above cut-off threshold determining a non-penalising annual increase. Some regions (e.g. Liguria) were expecting to be included but remained out whereas some other regions were expecting to remain out but were included (e.g. Lazio). The central government decided the above-mentioned 7% in the last days of 2006 when the budget law for 2007 was finally approved by the Parliament and the 7% value was quite unexpected, according to national newspapers. As a result, at the end of 2006 each region had a positive probability of being eligible as a member of the treatment group based on her previous and irreversible behaviour. Therefore, in the Italian tax experiment the randomness in the data did not come from a pure stochastic assignment of regions to treatment or from drawing regions from an infinite super-population of regions. Yet, the Italian tax experiment is a variant of a "design-based" experiment with (almost) random participation and permits conducting a "design-based inference" for DID. In this connection, Roth et al. (2023, p. 2219) recall that methods that are valid from the canonical sampling-based view are typically also valid from the design-based view as well, with the recommendation of clustering standard errors at the level at which the treatment is independently assigned (the regional level).

3 The DID identification

We follow Roth et al. (2023, p. 2220) and adopt the identification hypothesis of a TWFE panel data DID method to conduct the causal-effect study of the impact of a central government exogenous income tax policy. We employ response variables related to labour supply and other growth-related variables (such as family consumption) to evaluate the short and long run impact of income tax hikes taking both time and individual effects into account[2]. To illustrate the estimated model, we introduce the following notations:

$$y_{it} = \text{Observed response variable in region } i \text{ at time } t$$

$$D1 = \begin{cases} 0 \text{ when the region is untreated} & (\text{irrespective of time}) \\ 1 \text{ when the region is treated} & (\text{irrespective of time}) \end{cases}$$

[2] In the present case migration motivated by tax minimizing purposes had to be realised across regions (from treated to untreated regions) not municipalities, which may imply high moving costs. As a result, the estimation strategy may not consider the effects of the changes in the response variables induced by the inter-municipal migration motivated by municipal income tax differences as is it cleverly done by Rubolino and Giommoni (2023).

$$D2 = \begin{cases} 0 \text{ for years in which there was no treatment (irrespective of region)} \\ 1 \text{ for years in which there was treatment (irrespective of region)} \end{cases}$$

The total sample period is T = (1995… 2021) with t = (1995 … 2006) = No-Treatment period (dummy variable D2 = 1) and t = (2007 … 2021) = Treatment period (dummy variable D2 = 1)

$\boldsymbol{X}_{it}$ = vector of cofactors/controls (to be specified in each equation)

As noted above, we define the target or response variable y in terms of potential outcomes and estimate the average effect of treatment on the treated units. This compares the potential outcomes with treatment to the potential outcomes with no treatment, in the treated group. Written mathematically, the estimated effect of the treatment is the β coefficient of panel data OLS equation:

$$y_{it} = Constant + \alpha \times D1 + \delta \times D2 + \beta \times (D1 \times D2) + \boldsymbol{X}_{it}\gamma + \varepsilon_{it} \qquad (1)$$

where ε_{it} is a White Noise error term.

The crucial parameter to estimate is β. If we call R a binary income tax treatment indicator, we have $\hat{\beta} \equiv CATET = \mathbb{E}\left(y_{i,t}^{1} - y_{i,t}^{0} \middle| X_{it}; R_{i,t}\right)$ or $\hat{\beta} \equiv ATET = \mathbb{E}\left(y_{i,t}^{1} - y_{i,t}^{0} \middle| R_{i,t}\right)$ if cofactors are included or excluded. CATET (Conditional Average Effect upon Treated) or ATET will be interpreted as the mean effect (conditional or unconditional) of the "tax treatment" for those taxpayers who were compelled to participate in the central government program of income tax changes (i.e. the residents in the treated areas).

Unless it is not otherwise specified, in all estimates the average treatment effect is estimated by adjusting for both cross sectional and time effects. Notice that, given the design of the tax treatment, the response variables used in the estimations will all be independent on treatment conditional upon the $\boldsymbol{X}$ that will be used. Then in addition to the assumption of the parallel trend and the absence of anticipatory effects, our DID specification relies also on the *unconfoundedness* hypothesis (treatment participation is uncorrelated to the realizations of response variables). For (1) to identify the parameters of interest, tax shocks need to be exogenous conditional on fixed effects and controls. Intuitively, this identifying assumption is that the national tax shock of 2007 is not favouring regions that are doing poorly relative to how fast they normally perform in terms of response variables. Then, the validity of comparing outcomes of regions having different distributions of the response variables relies on three key assumptions: (1) the national tax shock is exogenous, (2) targeted tax shocks are unrelated to any possible targeted level of the response variables, and (3) outcomes from untreated regions provide a reasonable counterfactual since the 2007 tax shock was absent. Since we control for region and year fixed effects in equation (1), the first assumption maintains that Italian policy makers of the time were not systematically setting income tax policy to respond to idiosyncratic regional shocks other than the budget unbalance of their health care systems mentioned in section 2. Notice that the central government income tax policy introduces a contemporaneous treatment to a fixed group of

regions (treatment effect homogeneity). Then, the TWFE DID model (1) does not make "forbidden comparisons" between already-treated units (Roth, Sant'Anna, Bilinski and Poe, 2023, p. 2228).

Yet, the fact that assumptions (1), (2) and (3) above permit to use TWFE regression specification for the estimation of the treatment parameter does not exclude that other conditions for efficient parameter estimations of an OLS-based model should be met to obtain asymptotically valid inference, particularly when the sample period is large. Stationarity is not an absolute requirement for panel data analysis, but it is preferred in many cases. With a short time series of, for instance, up to 10 years, stationarity may not be critical. In that case, non-stationary of the data still allows using fixed effects (or first differences) models to control for unobserved heterogeneity and trends over time. Yet, with a sample like ours (27 years), it is necessary to conduct tests for checking the presence of panel unit root because the use of OLS-like procedures can produce invalid estimates. Granger and Newbold called such estimates 'spurious regression' results: high R^2 values and high t-ratios yielding results with no real meaning. As it shown in Table 1, we reject the panel unit root hypothesis at usual significance levels and interpret the result as evidence that a statistically significant proportion of the units (both treated and untreated) are stationary.

3.1 Data description

We collected a set of regional ($N = 20$) potential response variables on a yearly basis from 1995 to 2021 (i.e., $T = 27$ years). Descriptive statistics of the response and control variables are in Table 1 where stationarity and cross-sectional dependence tests are also reported. The relevance of these tests is discussed later. Data are divided between response variables

1. Self-employment (measured as the share of self-employed people out of the total annual employment in the region)
2. Family consumption
3. Number of new VAT certificates[3]
4. Regional GDP and GDP growth

and cofactors candidates

1. Internal Fixed Gross Investments[4] for non-financial sectors
2. Regional Value Added per worker (measure of labour productivity)

Descriptive statistics and tests are reported in Table 1 below.

[3] The new VAT certificates measures the annual new flow of operators who in their activities require a VAT identification number. The tax authorities grant this certification to entities wanting to start some business activity. It approximates the net flow of annual business entities entering the respective professional markets and it is computed as the Change of the annual stock of existing certificates = Existing Certificates at t-1 + New Solicited Certificates at t – (Expired Certificates at t – Cancelled Certificates at t).

[4] The Internal Fixed Gross Investments (non-financial sector), which consist of the acquisitions (net of sales) of fixed capital carried out by resident producers to which the increases in value of non-produced tangible goods is added.

Table 1: Descriptive statistics of response variables and cofactors (monetary variables are in 2015-chained values)

Variables	Obs.	Mean	SD	Min	Max	Panel unit roots tests H_0: Panels contain unit roots		Panel Stationarity	Pesaran CD test for cross-sectional dependence			
20 Units /Regions T = 1995, …, 2021						**Levin-Lin-Chu (LLC)**	**Harris-Tzavalis (HT)**		**Test statistics**	**P-value**	**Average joint T**	**Mean abs(ρ)**
Regional GDP (Euros)	540	8.35e+09	8.24e+09	4.25e+08	3.86e+10	t* = -2.429 PV = 0.0076	ρ = 0.6488 PV = 0.0057	Yes, both tests	43.348	0.000	27.00	0.64
Per-capita regional GDP (Euros)	540	27699.96	7453.601	15313	43103	t* = -1.928 PV = 0.0269	ρ = 0.6669 PV = 0.0210	Yes, both tests	53.49	0.000	27.00	0.75
Value Added per worker (Euros)	540	49832.05	13346.88	4253	76483	t* = -2.314 PV = 0.0103	ρ = 0.6615 PV = 0.0146	Yes, both tests	53.331	0.000	27.00	0.83
Gross Fixed Capital Formation GFCF (Euros)	540	1.62e+10	1.56e+10	8.45e+08	8.07e+10	t* = -2.594 PV = 0.0047	ρ = 0.7903 PV = 0.9093	Yes LLC No HT	48.668	0.000	27.00	0.68
Total regional Employment (Units)	540	1210098	1063136	58600	4884300	t* = -2.239 PV = 0.0126	ρ = 0.8682 PV = 0.9997	Yes LLC No HT	45.583	0.000	27.00	0.65
Self-employment (% of total employment)	355	13.55634	2.32783	6.9	20.4	t* = -5.298 PV = 0.000	t* = 0.995 PV = 1	Yes LLC No HT	19.578	0.000	17.51	0.37
New VAT Certificates (Units) (2001 – 2021)	421	25853.42	22523.57	879	90191	t* = -4.879 PV = 0.0000	ρ = 0.6287 PV = 0.1595	Yes LLC No HT	24.256	0.000	20.20	0.53
Family Consumption (Euros)	540	50667.72	43999.08	2593	204864	t* = -3.0287 PV = 0.0012	ρ = 0.6649 PV = 0.0185	Yes, both tests	43.009	0.000	27.00	0.60

Notes. Sources: ISTAT and Italian Treasury (Department of Finance). Mean, standard deviation (SD), min and max are computed using the overall sample. All monetary values are expressed as values linked to the 2015 reference year: the linking provides a measure of the economic aggregate of interest in terms of volume, i.e. net of the underlying price dynamics. Note that the sum of the chained values of the components of an aggregate is not equal to the chained value of the aggregate itself.
Under the null hypothesis of cross-section independence, the Pesaran CD test statistics is normally distributed, i.e., CD ~ $N(0,1)$, and p-values close to zero indicate that data are correlated across panel groups. In the Levin-Lin-Chu (LLC) unit root test, a common autoregressive parameter for all panels is assumed, so it does not allow for the possibility that some regional variables contain unit roots while others do not. To mitigate the effects of possible cross-sectional correlation among time series, we removed the cross-sectional means from the series (demeaning). Also, we included for all the tests a linear trend term. The LLC test requires that the ratio of the number of panels to time periods tend to zero asymptotically. Hence, it is not well suited with relatively few periods. Hence, we also present the Harris–Tzavalis (HT) unit root test, which assumes that the number of panels tends to infinity while the number of periods is fixed. For both tests, the null hypothesis is H_0: Panels contain unit roots.

The presence of parallel trend in pre-treatment periods for response variables is of great importance for DID identification since violation of parallel trend assumption will lead to biased estimation of the causal effect. For some of our response variables pre and post treatment period (end of 2006) are shown in the following Fig. 1. It is frequently assumed that the smaller the time period tested (for example a 2-year period), the more likely the assumption is to hold, but with a period of 27 years the possibility that treated and control groups have different outcome trends (which may generate time convergence) cannot be excluded.

Figure 1: *Time behaviour of response variables before and after the 2006/7 tax treatment* (Red line = Treated regions; Blue line = Control regions)

Notes. The plots permit a graphical diagnostic for parallel trend and show the time evolution of the response variables used in this study before and after the income tax treatment. 2006 is the last year before the treatment (vertical line). The first plot consists of two lines showing the mean of the outcome over time for the treatment and the control groups. The second plot augments the DID model to include interactions of time with an indicator of treatment and plots the predicted values of this augmented model for the treatment and control groups. See also the Notes below Fig. 5. The plots of Appendix 2 supply more graphical analysis of the time-space evolution of the variables.

Visual inspection of Fig. 1 shows that parallel variations prior to 2007 is a reasonable hypothesis for the response variables and it is suggestive that the identification proposition of the DID are valid as far as parallel trend is concerned. Plots reproduced in Appendix 2 (Fig. A6) show the time series of index numbers with baseline 2006 = 100 for all available response variables. Abruzzo, Campania, Lazio, Molise and Sicilia are the five treated regions, while the non-treated regions are averaged across year (solid black lines). Both groups of plots (Fig. 1 above and those reproduced in Appendix 2) indicate that the average change in the response variables for treated and untreated regions would have been the same in the absence of income tax hikes.
Yet, the parallel trend hypothesis will be subjected to statistical test based on an augmented version of eq. (1) and is discussed at the beginning of the next section. The presence of anticipation effects will be also tested. All the tests are reported in section 4, Table 2.

Note that the reported descriptive statistics in Table 1 are computed on the overall sample, and they do provide information about the overall scale and the variability of the macroeconomic indicators under consideration. However, they do not provide any insight about the territorial heterogeneity typically affecting the Italian economic measurements (see, for example, De Philippis, Locatelli, Papini, and Torrini (2022); Federici, Ferrante, and Parisi (2023); Lo Cascio, Mazzola, and Epifanio (2019)). In the Appendix 2 (Figures A2, A3 and A4) we show the evolution of the considered regional indicators for selected years along with the 1995-2020 average. For almost all indicators included in the sample, the charts show a pronounced regional heterogeneity, highlighting in many cases either the North-South gap (e.g. GDP per capita, value added and employment), which persists over the decades, or clear patterns linked to urbanisation degree and population density (e.g. regional household consumption). Although territorial heterogeneity is not explicitly modelled within our framework by including distances between regions (J Paul Elhorst, 2014) or geostatistical factors (Wikle, Zammit-Mangion, & Cressie, 2019), the cross-sectional (regional) fixed effects included in Equation (1) can still mitigate the effect of spatial dependence by capturing region-specific latent characteristics (J. Paul Elhorst, 2010, 2024) and preserving reliable estimates. In this sense, learning from examples like Merfeld (2019) and Muralidharan, Niehaus, and Sukhtankar (2023), a potential avenue for future investigations of Italian tax treatment could take into account the spatial structure between areas looking for spatial spill-overs and bordering effects of the policy.
As anticipated above, due to the panel macroeconomic nature of the dataset we also test for panel non-stationarity. Adding the cross-section dimension to the time series dimension offers an advantage in testing for nonstationary and cointegration since cross-section increases the data set used in those tests, thus improving their power. However, the cross-sectional dimension also brings some new problems to our TWFE DID model, namely the existence of cross-sectional dependency that can bias usual panel data unit root test results in small

samples. The Pesaran (2007) cross-sectional dependence test checks for this issue. Our findings show, as expected, the presence of high correlation among regional variables across panels/regions. The importance of a priori check of the existence of unit roots in the panel data comes from the already known effect that the presence of unit roots in time series may cause a misinterpretation of estimated results. The unit root hypothesis is tested through the Levin, Lin, and Chu (2002) and Harris and Tzavalis (1999) unit root tests (i.e., for both, the null hypothesis is that the panels contain unit roots) accounting for linear trend and cross-correlation adjustments. With the only exception of the variable "Regional Value Added per regional worker", we reject the null that the panel contains unit roots and conclude that the series are trend stationary. Since in the OLS estimation of (1), panels are considered to be homogeneous, the opposite would be a drawback. The absence of unit roots is important for our DID analysis. In the most intuitive sense, stationarity means that the statistical properties of a process generating a time series do not change over time because the statistical properties of the data generation process do not change over time. It does not mean that the series does not change over time, just that the way it changes does not itself change over time. If an exogenous event occurs (such as an exogenous tax change) changes in the response variables can be, loosely speaking, attributed to the influence of the event and not to some long run property of the series.

4. Results

TWFE panel DID results are presented in Table 2. The ultimate causal parameters of interest (ATET) are reported in the first row. Tests of parallel trend and anticipation effects are reported in the last rows and commented in the Notes below the table. We refer the reader to Table 3A in the Appendix for a detailed summary of the estimated coefficients, including those associated with temporal and cross-sectional fixed effects.

Table 2: TWFE DID estimations of the Italian income tax treatment

Response Variables	**Self-Employment**	**New VAT *Certificates***	**Regional GDP Growth per capita (HP filter)**	**Regional GDP per capita**	**Log Family Consumption**
ATET or CATET					
Treated vs untreated regions	-1.165**	-4,592**	-.032***	-629***	-.049**
Controls					
log(GFCF)	-0.8	-453.9	.09***	2241***	
Log(Value Added per worker)	-11.716371***	-13460.289	.23681953**	9867.1645**	
Const.	160.11159***	178491.2	5.672762***	-128581.29***	10.29218***
Time fixed Effects	√	√	√	√	√
Region fixed Effects	√	√	√	√	√
N	355	420	540	540	540
TESTS of Parallel Trend and Anticipation responses					
Ptrend Test: H_0 **Linear parallel trend**	F(1, 19) = 2.30 Prob > F = 0.1455	F(1, 19) = 2.02 Prob > F = 0.1715	F(1, 19) = 1.18 Prob > F = 0.2903	F(1, 19) = 0.34 Prob > F = 0.5674	F(1, 19) = 0.62 Prob > F = 0.4398

	H_0: cannot be rejected	H_0: cannot be rejected	H_0 cannot be rejected	H_0 cannot be rejected	H0 cannot be rejected
Granger Type Test: H_0 **No Anticipation effects**	F(3, 19) = 1.50 Prob > F = 0.2466 **H_0 cannot be rejected**	F(5, 19) = 1.34 Prob > F = 0.2893 **H_0 cannot be rejected**	F(11, 19) = 11.29 Prob > F = 0.0000 **Reject H_0**	F(11, 19) = 7.29 Prob > F = 0.0001 **Reject H_0**	F(11, 19) = 7.12 Prob > F = 0.0001 **Reject H_0**

Notes. Results from DID estimation of eq. (1) with robust SE. H_0 in the *Ptrend* test is that pre-treatment linear trends are parallel. *Granger-type* test is for anticipatory effects. H_0 is that there are no effects in anticipation of treatment. In the fourth and fifth columns, the dependent variable is the logged per-capita income of the regional residents. We used the Hodrick–Prescott high-pass filter to separate the time series into trend and cyclical components. The trend component may contain a deterministic or a stochastic trend. The smoothing parameter determines the periods of the stochastic cycles that drive the stationary cyclical component. With the exception of the first, columns report CATET estimates obtained using Controls in addition to Time. ATET estimates (not reported) are very similar. More comments and test discussion are in the Notes of Table 3. To save space, the analytics of the test procedures are discussed in the Notes below Table 3 and in *fn* 7. Data source: see previous Table.

Before commenting the specific results for each response variable, we discuss the test for the presence of the parallel trend and anticipation effects. The last two rows of Table 2 report the estimated F-tests for parallel trend hypothesis and the test for absence of anticipation effects. The former is based on the results of an augmented linear DID model based on our eq. (1) with additional terms that capture the differences in slopes between treated and controls. As for anticipation effects, Table 2 reports the results of a Granger type test obtained from a linear DID model based once again on our eq. (1) augment in this case by interacting the dummy variable that marks treated observations with dummy variables for time periods prior to the treatment to capture any potential anticipatory treatment effects. Yet, one should recall that the Granger-type test is less robust than the above parallel trend F-test since it consumes more degrees of freedom because it estimates a higher number of parameters. Hence, the parallel trends test has higher statistical power. T results allow us to conclude (at any significance level) that the data generation process is consistent with the hypothesis of parallel trend but not always with the absence of anticipation effects hypothesis. The latter result may be informative for the policy design of the tax treatment. In the following sections, the estimation results are commented for each response variable with greater emphasis on those related to labour supply.

4.1 Self-Employment and New VAT Certificates

If the determinants of employment decisions are based on the hypothesis that the individual responds to the risk adjusted relative earnings opportunities, then income tax hikes should affect negatively their supply behaviour at both intensive and extensive level. As for the duration of this effect, (Saez, 2002, p. 1043) rightly emphasises that the shorter the time period upon which the income tax is assessed the more relevant is the extensive effect of the tax changes. The results reported in the first two columns of Table 2 refer to different ways of measuring labour supply at the extensive margin. The first is the annual levels of *Self-Employment* (measured as the share of self-employed people out of the total annual employment in the region). The second (*New VAT Certificates*) is a response variable measuring the participation decision of a vast set of categories of people offering labour services.

4.1.1 Self-Employment

Self-Employment is affected by income tax hikes. Self-Employed people are the taxpayers mostly affected by the 2007 tax treatment since they will pay both taxes whose rates have being increased in the five treated regions (see Appendix 1). Then, our DID estimations show that the share of Self-Employed workers is reduced by the increase of income taxation. These results accord with previous literature (e.g., Heim, 2010), including those models which consider explicitly how stochastic shocks in demand or cost functions affect the decision to become self-employed (Appelbaum & Katz, 1986; Kihlstrom & Laffont, 1979). In those models, an asymmetric tax reduction, favouring the self-employed relative to other workers, increases self-employment. The converse result represented by the negative effect of income tax hikes is not a surprise[5] and apparently it is not off-set by the many "tax opportunities" open to self-employed people. Robson and Wren (1999) state that higher tax rates lead individuals to take up self-employment to take advantage of the greater opportunity for tax avoidance and evasion that self-employment offer relative to paid employment (i.e., pure wage earners). According to this latter view, one should expect that income tax hikes increase rather than decrease *Self-Employment*. Our results indicate that this effect should be excluded. Yet, the estimated effect may be compatible with reactions consisting in tax evasion behaviour. On the other hand, the findings presented by Agrawal, Foremny, and Martínez-Toledano (2024, p. 2) indicate that self-employed wealthy individuals are not significantly more likely than non-self-employed wealthy individuals to move to less taxed areas because of tax reasons (a wealth tax in their research). A possible interpretation of our findings may be that the reduction of self-employment depends upon a reduction of low-middle income self-employed individuals who decided not stay (at least openly) in business after the increase of the marginal income tax rate.

Anticipatory effects on *Self-Employment* are excluded by Granger-like test reported in Table 2. We interpret the test as the absence of anticipated avoidance-evasion behaviour on the part of self-employed taxpayers in the treated regions (the tax hikes arrived as a surprise) or as a clue that changing occupation in advance of the tax hikes was seen as not profitable.

As for the duration of the effect, Figure 3 (middle plot) shows that the impact of the tax treatment on *Self-Employment* is statistically significant in the first year of the treatment and during the 3 years following the treatment, but the coefficients are no longer statistically significant afterwards. Then, the tax effect loses strength and significance as we move away from the initial year of the treatment. Still, results indicate that income tax hikes affect the level of self-employment in the treated areas although the effect last for a short time.

The significant negative effect of tax hikes on *Self Employment* and the non-significant effect on *Total Employment* are also observable from the temporal trends reported in Appendix 2 that we use as an additional

[5] Hoover, in many models of the specification of the relationship between employment and income tax, no distinction is typically made between the marginal and average tax rate and this make the policy evaluation more difficult. For example, Blau (1987) uses measures of the marginal tax rate in his study of the determinants of U.S. self-employment, and Evans and Leighton (1989) use the average tax rate. The impression is that for these authors distinction is largely irrelevant.

source for the interpretation of the TWFE results. Figure A5 shows the absolute variation between 2006 and 2009 of the above response variables, while Figure A6 shows the time series of index numbers with base year 2006 (the last year prior to treatment). As for *Self-Employment*, the five treated regions show substantial reductions (between 20% and 30% in the four years following the treatment according to the reported index numbers). On the contrary, in the untreated regions we observe on average an increase of recorded values. The same Figure A6 also shows that *Total Employment* (reported in the Appendix for completeness) has changed heterogeneously within the treated group with some region experiencing an increase (Lazio +3%) and some others experiencing a reduction (Campania -6% and Sicily -3%) while some other regions (Abruzzo) showing no variation.

4.1.2 New VAT Certificates

The second labour response variable is the *New VAT Certificates*. Operators who want to start legally a (new) business activity must obtain a VAT certificate (similar to the tax code identification) from tax authorities. Hence, this flow variable measures the net increase of economic operators (plumbers, mechanics, shop owners, engineers, architects, etc.) requiring and obtaining a formal VAT Certificate from the tax authorities, net of the number of certificates returned back to tax authorities or extinguished during each year. Clearly, employees are excluded from this set of taxpayers. In a DID model a CATET/ATET parameter will correspond to the estimate of the effect of the income tax hikes on the (official) net flow of new labour supply provided in each region by professional business operators who need the above certification. Then, observing the post treatment behaviour of this response variable and the difference between treated and untreated taxpayers one may obtain indication on the extensive effect of income tax increase on that specific subset of taxpayers. As it was stressed in section 2, those who apply for a VAT certificate/permission are subjected to the payment of *both* IRAP and IRPEF. Therefore, they represent the labourer category most affected by the tax treatment of 2007. How this variable reacts to the tax treatment indicates an extensive marginal reaction to taxation that takes the form of a participation decisions: either apply or not apply for the certificate as a result of the tax change. As stressed by Saez (2002, p. 1048) this view is obviously a crude simplification of reality but captures the extensive margin labour supply decision. If one assumes away income effects and tax evasion, the decision to participate depends only on the difference between the (after tax) income with participation and unemployment transfers without participation.

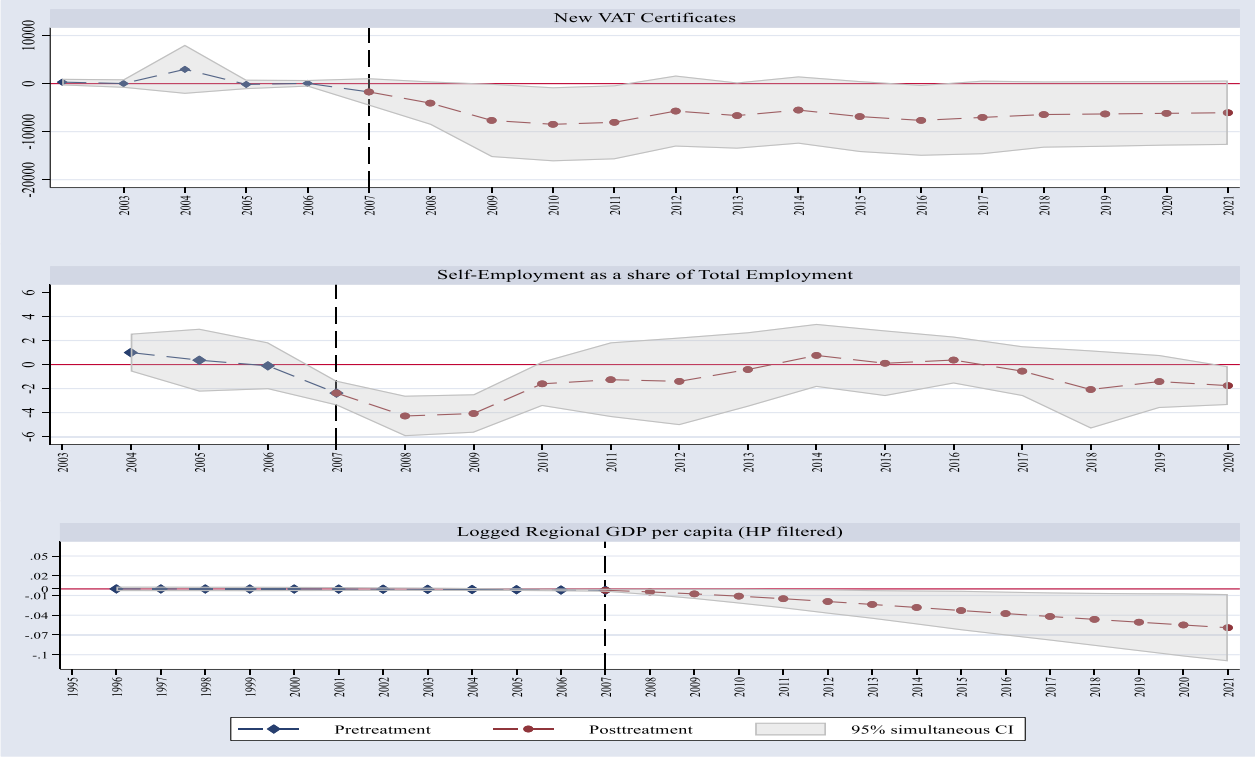
The estimation of the tax treatment indicates that an *average* decrease of the new applications for the VAT Certificates roughly corresponds to 4600 units (about the 6% of the average 2006 value), which is the estimate of the average annual reduction of new (official) operators in the treated regions with respect to the untreated ones. This means that the tax treatment has discouraged new professionals to enter the relevant markets in treated regions or has led some of the already active operators to leave the (official) markets.

Once again, the test reported in the last rows of Table 2 suggest that we cannot reject the null that trends are parallel and that anticipatory effects are absent even for this variable. Recall that operators (old and new)

requiring a VAT Certificate to carry a business are the main IRAP taxpayers, and the result seems to indicate that the tax base of IRAP is rigid to anticipatory adjustments.

As for the effect of the tax treatment over time, Figure 3 (upper plot) shows that the effect of the tax treatment on the supply of professional services is strong, statistically significant (slightly at the 95% all over the sample), and not restricted to the very short run.

Figure 3: *Estimated ATET/CATET over Time in treated regions for New VAT Certificates, Employment, and Self-Employment and per-capita regional GDP (HP filtered)*



Notes. In the plots, the vertical line corresponds to the first year of treatment and, therefore 2006 is the last year without treatment. As a result the 2006 value of the parameter is zero. Details on how the above parameters are estimated are given in the Notes below Table 3 where we report the results of the robustness analysis.

Altogether, our results show that income tax hikes might reduce *Self-Employment*. As in Wen and Gordon (2014) where in a *probit* model the tax convexity variable and the net-of-tax income difference between self- and paid employment have the predicted (negative) signs and there is a high level of statistical significance for the probability of self-employment, in our estimation Self-Employment negatively reacts to income tax increase. Then tax hikes are perceived as disincentive to entry the formal self-employment professional markets and produce negative effects at the extensive margin by reducing the creation of new companies by entrepreneurs who plan to exploit some business opportunities (Aghion and Howitt (1990). Our DID results accord with those obtained by some recent studies in which income/profit taxation is a determinant of the

incorporation decision (see Barrios, Huizinga, Laeven, & Nicodème, 2009; Djankov, Ganser, McLiesh, Ramalho, & Shleifer, 2010) and represents an instrument to support business conditions (Da Rin, Di Giacomo, & Sembenelli, 2011).

Graphical tendencies shown in Appendix 2 accord with our TWFE DID estimates. The index numbers, in particular, clearly show that the pattern of the treated group and of the control group are widely divergent and that the regions affected by the income tax hikes have experienced large reductions (around 20% in 2009) in the number of VAT certificates compared to the rest of Italy.

4.2 Per-capita income

Two income response variables are considered: *per-capita Regional GDP Growth* (Hodrick-Prescott filtered), and *per-capita Regional GDP* in levels, which proxies the income tax base for the generality of taxpayers. For the former we use the Hodrick–Prescott high-pass filter to separate the time series into trend and cyclical components. Cofactors are Gross Fixed Capital and Value Added per Worker (average labour productivity). Estimates show that personal income growth is negatively affected by income tax increase and this result accords with the previous extensive labour supply contraction shown in the first three columns of Table 2. As for the values of *per-capita Regional GDP* in levels, results show that it is negatively affected by the treatment by a value nearly equal 2.3% of the national mean value reported in Table 1 for the entire data set and to a 2.7% of the per capita income measured in treated regions only (Average Effect upon Treated). In this case, anticipated effects cannot be excluded (see the last row of the penultimate column) which implies that some taxpayers in treated regions may have anticipated the tax treatment and have taken advantage of any existing opportunities to plan tax-reducing strategies.

We may contrast the above results with Zidar (2019, p. 1423). A clear similarity emerges between our response results and his estimates, although, in our case, estimates indicate a more pronounced decline of the response variables in treated regions. The exogenous increase of both taxes (the regional income Surcharge and the increase of the business quasi-income tax IRAP), taken in isolation, apparently reduces the post treatment regional GDP levels and growth over time. Clearly this drastic conclusion should be moderated by recalling that the effects on income and growth, via the local multiplier effects, should be evaluated by weighting the effects of both taxes, on the one hand, and public expenditures on the other (Ramey, 2011a, 2011b).

As for the effects of the treatment over time, the lower plot of Fig. 2 shows the time value of the treatment coefficient. The effect of the tax hikes is not strong at the beginning but it has leads effects during the years following the treatment. These effects become stronger and statistically significant as time passes.

The graphical trends (Figures A5 and A6 in appendix 2) of the three indicators show a marked generalised reduction following the introduction of the treatment. While the average values in the untreated regions shows a pro-cyclical pattern with quick alternations of growth and stagnation, Sicily, Campania and Molise (three of the treated regions) follow a smooth process of GDP loss that is not reabsorbed in the following years. The general conclusion is that the GDP response variables negatively react to the income tax treatment.

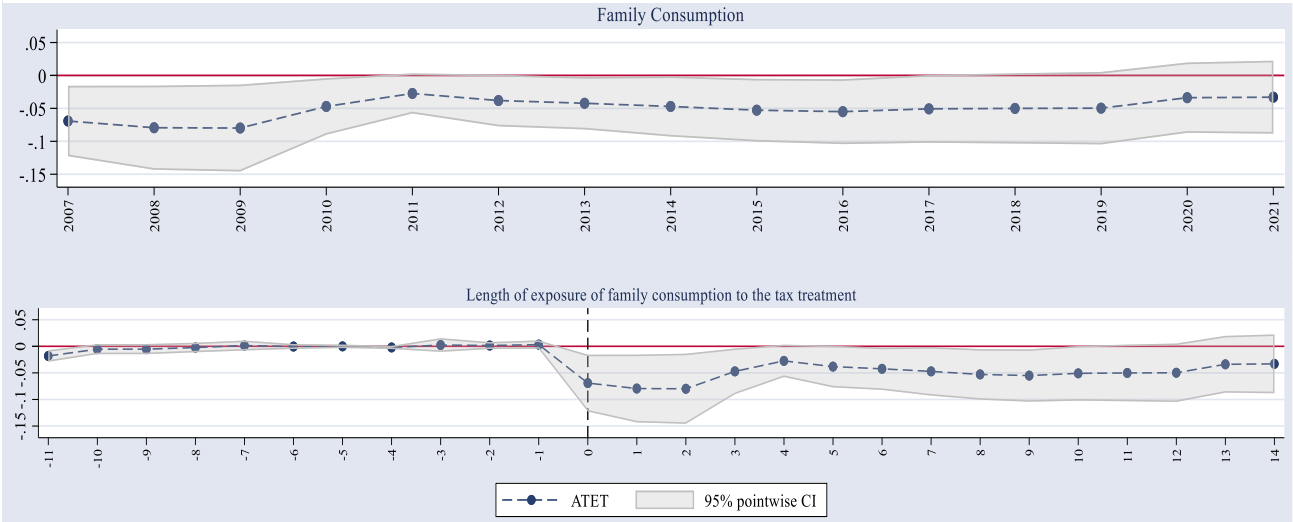
4.3 Family consumption

Results obtained for labour supply and GDP point in the direction that tax hikes should negatively affect also other labour and growth-related exogenous variable. Table 2 reports results for the response variable represented by family consumption. We observe that estimated parameters (last column of Table 2) show a negative value of the overall consumption elasticity in the treated region to be approximately 0.5%. The 2007 tax treatment clearly made the (national and regional) income taxation more progressive in treated regions. Thus, by changing the disposable income distribution in treated regions, a more redistributive taxation may have changed the relative aggregate demand for goods and services with possible decrease of consumption of taxpayers of the higher taxable brackets. This in turn may have led to changes in the demand for workers in various sectors, particularly those offering more income-elastic goods. With non-homothetic consumption preferences, when we allow consumption to adjust (i.e., when income effects are at work), any change in progressivity differentially affects the labour supply of skilled and unskilled workers because the compositions of agents' consumption baskets change with their disposable income. If this is the case, the marginal utility of income may vary across income groups, distinctively affecting the incentives to work of each group (Oni, 2023, p. 4). Consumption results seem consistent with the above commented labour supply results.

If we look at the two plots of Fig. 2 and at the *Granger-type test* results reported at the end of Table 2 (the null of absence of anticipatory effects of tax treatment on consumption must be rejected), we may state as an additional comment that in our data set the permanent income hypothesis is not consistent with the data[6].

The length of exposure of consumption to the tax treatment is shown in Fig. 4 below.

Figure 4: *Estimated ATET/CATET over Time in treated regions for the Family consumption.*



Notes. See the explicative Notes below Fig. 3.

[6] This is so because consumption should not respond to *predictable* fluctuations in income (see for example, Souleles, 1999, p. 947, who finds significant evidence of excess sensitivity in the response of households' consumption to their income tax refunds). Nguyen, Onnis, and Rossi (2021, p. 456 and Fig. 4) also find a significant sensitivity of consumption to general level of taxation.

The tax shock has a strong effect during the first years following the treatment and then loses intensity and significance (at the 95%).

Similar deductions can be made from the observed trends in household consumption exposed in Appendix 2 (Figures A5 and A6), which show that consumption in the treated regions has a behaviour comparable to that of GDP. In particular, the graphs suggest that the reduction of consumption in the treated regions was stronger than in the control group between 2008 and 2011, being partially reabsorbed by some regions (yet, Lazio, Sicily and Campania did not catch up). This may suggest that durable consumption might be affected by the treatment at the beginning of the post treatment period and then, since current consumption may be measured with error, its point estimates may be less reliable than other response variables such as, for instance, self-employment. However, the results are somewhat not clear-cut for no distinction is made for durable goods consumption vs current consumption and, above all, for different income groups.

4.4 A robustness analysis

We replicate the TWFE panel DID estimation using *Rate of Employment* (male and female, age from 15 to 64 years, Eurostat NUTS-3 data) recorded at the provincial level in treated and untreated regions from 2004 to 2021. Altogether, we use data of 103 provinces distributed among 20 regions: 5 treated and 15 untreated. The model is identified by the same basic TWFE DID assumptions of equation (1), plus the requirement that each province belonged to the same region at the beginning and at the end of the sample period 2004-2020 (no province left the original region). We also include cofactors relative to both provincial and regional levels. Equation (1) rewrites

$$y_{ijt} = \beta_0 \underbrace{+[\alpha(D1) + \delta(D2) + \beta(D1 \times D2)]}_{\text{Tax Treatment Component}} + \sum_{i=1}^{P} \beta_i X_{ijt} + \sum_{j=1}^{R} \beta_j W_{jt} + u_{ijt} \qquad (2)$$

where $t = (2004, \dots, 2020)$ and X_{ijt} is the year value of each of the possible P cofactors existing at the provincial level (i) (when i belongs to region j) and W_{jt} is the year value of each possible R cofactor existing at the regional level j and associated to the provinces nested in that region (for each $i \in j$). Finally, u_{ijt} is an error term normally distributed. Clearly, the Tax Treatment Component is the same as in (1).

All estimated versions reported below include both unit fixed effects and time fixed effects in ordinary least squares estimation. The following Table 3 repost results of various version of the model.

MOD1 is a no-cofactors (pure ATET) model; MOD2 includes provincial cofactors only; MOD3 has cofactors of both provincial and regional level; and MOD4 includes a variable computed as the annual average value of the employment rate and is included to incorporate the national mean trend of the response variable.

Tab 3. *DID ATET TWFE Estimations of various versions of Equation (2)*

Response RATE OF EMPLOYMENT	**MOD1**	**MOD2**	**MOD3**	**MOD4**
INCOME TAX TREATMENT (CATET Treated vs Untreated)	**-1.74*** **	**-1.69*** **	**-1.35*** **	**-1.35*** **
Constant	58.015534***	56.73562***	50.446486***	-77.124896
COFACTORS				
Provincial GDP PPS		.000097**	.00008323**	.00008323**
Regional GDP PPS			5.167e-10***	5.167e-10***
Regional Growth Rate			0.53433769	0.53433769
National Mean Employment Rate				2.2019134**
Time Fixed Effects	√	√	√	√
Individual Fixed Effects	√	√	√	√
Observations	1,751	1,751	1,648	1,648
Parallel Trends Test				
F(1, 102)	0.42	0.47	0.44	0.44
	Prob > F = 0.516	Prob > F = 0.4942	Prob > F = 0.5077	Prob > F = 0.5077
Anticipation Effects				
Granger causality test				
F(2, 102)	0.32	0.34	0.43	0.43
	Prob > F = 0.725	Prob > F = 0.7115	Prob > F = 0.514	Prob > F = 0.514

Notes. *Ho* of both Parallel Trends[7] and No Anticipatory Effects[8] cannot be rejected at any level of significance. Interpretation of ATET/CATET is easy: the rate of employment is negatively affected by income tax increases in any specification of the model. MOD4 includes as cofactor the population grand mean of the response variable, National Mean Employment Rate. It is computed as (103 ×

[7] To test for parallel trends we define the model

$$y_{ij,t} = \gamma_{00} + [\alpha(D1) + \delta(D2) + \beta(D1 \times D2)] + \gamma_{10} X_{ij,t} + \psi_1[S_0 \times w_i \times YEAR] + \psi_2[S_1 \times w_i \times YEAR] + r_{ij,t}$$

where S_0 is a time dummy indicating pre-treatment years, S_1 indicates treatment years and *w* indicates either treated provinces (belonging to a treated region) ($i = 1$) or untreated provinces ($i = 0$). If the estimated $\psi_1 \neq 0$, then a difference in the slope of the time trend between treated and untreated units in pre-treatment years is statistically consistent with the data. Table 1 reports the Wald test of the H_0: $\psi_1 \neq 0$ vs $\psi_1 = 0$. The null is that of linear parallel trends ($\psi_1 = 0$). This means that the test statistic is obtained by dividing the maximum likelihood estimate (MLE) of the slope parameter ψ_1 by the estimate of its standard error. Under the above null, this ratio follows a standard normal distribution. This model is also used to construct the right plot (linear trend) of Fig. 2; however, in latter case the variable YEAR enters as a difference with respect to the first YEAR included in the sample (2004).

[8] The last rows of Table 3 report the results of a Granger-type causality test to assess whether treatment effects are observed prior to the treatment. More on the test are in the Notes to Fig. 2.

$17)^{-1} \sum_{i=1}^{103} \sum_{t=2004}^{2020} y_{it}$. Including as a cofactor this "double demeaned" version of the response variable is motivated by the need to incorporate a long run trend of employment into a model such as (6) where time fixed effects are included, and therefore the long run changes in the general economic environment that have the same effect on all units are removed (Wooldridge, 2021). Then, since local employment is logically affected by the National Mean Employment, with the latter indicative of the national trend of employment, a factor incorporating the above general environment changes is needed to better define what affects local employment. Finally, when interpreting the estimated coefficient of the cofactors, the scale of measurement is to be considered. For example, since Provincial GDP PPS is measured in millions of euros, the reported coefficient .00008323 means that if that Provincial GDP increased by 1.000.000 euros, the rate of employment would increase by an 8% of its value. Legend: * p < 0.1; ** p < 0.05; *** p < 0.01.

Results reported in Tab. 3 indicate that the exogenous increase of the income taxation affects negatively the rate of employment in any specification of the estimated model. Interestingly, the presence of cofactors affects the magnitude of the estimated ATET (or CATET) coefficients. In MOD1 for example a 1% increase of the income tax rate leads to a 1.7% decrease of the ongoing mean employment rate in treated provinces whereas in MOD3 and MOD4 the effect reduces to a 1.3%. Province level cofactor (GDPs) is strongly statistically significant (for the interpretation of the magnitude, see the Notes below Table 2) and has the expected sign. As for regional level cofactors, GDP has effects analogous to the provincial GDP but its rate of growth has no statistically significant effect. Controlling for the National Mean Employment Rate has no consequences.
Graphical illustration of the existence of parallel trends is reproduced below and support the results of the tests reported in the last rows of Table X and commented about in the Notes below the Table and in the footnote. Yet, both procedures lead us to conclude that the pre-treatment parallel trend assumption is consistent with the data.

Fig. 5 *Provincial Employment Rate in Italy. Parallel Trends and time duration of exposure to income tax treatment*

Notes. The upper right plot shows that the before treatment period shows parallel trend and the post treatment period shows an increase of the mean values of untreated unis and a decrease of the mean values of the treated units. Centring around the minimum time value provides a common reference

point at the first observed time point such that deviations from parallelism are easily detectable. The graph then shows the predicted values from this model, evaluated at all observed time points for each of the treatment groups and at the means of the covariates. The graph at the bottom is obtained from the fits of MOD1 of Table 1 and is the plot of the estimated coefficients (including their 95% confidence intervals). The other MODs generate almost identical results and are not reproduced. The model is similar to the Granger model, but uses a different parameterization, and includes lags in addition to leads. To illustrate, let 2006 be the last year before treatment, m be the number of periods prior to 2006, $q \geq 0$ be the number of periods after 2006, and b is the baseline period. The fitted model is (using MOD1 of Table 2 as reference; no cofactors included)

$$y_{ij,t} = \gamma_i + \gamma_t + [\alpha(D1) + \delta(D2) + \beta(D1 \times D2)] + \sum_{\substack{k=m \\ k\neq 2006}}^{q} (B_{it}^{k} w_i)\tau_k + r_{ij,t}$$

where w_i has the same meaning as above (see fn. 4) and $B_{it}^{k} = (t_{it} = I_i + k)$ because we use all available leads and lags. The plot reproduces the estimated τ_k with the 95% CIs.

5. Additional statistical tests

As it was discussed in previous sections, the presence of different trends may jeopardise the parallel trend assumption of a DID. Given the macroeconomic nature of our regional data set, the presence of converging trends of some response variables cannot be excluded *a priori*. When the time period tested is short (for example a 2-year period) the assumption is the more likely to hold. Yet with a period of 27 years the possibility that treated and control groups have different outcome trends (which may generate time convergence or time divergence among them) cannot be excluded. Then, in addition to the tests reported in the last rows of Table 2 and the information provided by the visual inspection of the plots of Figure 1, in this section we propose *a falsification test* that *at least* excludes that the response variables follow some *converging trend*. Time convergence, i.e. converging trends of the response variables for treated and control groups (employment, labour supply, consumption, etc.), would then be at odds with the hypothesis of parallel trends. This is so because the change (or "path") in response variables over time that regions in the treated group would have experienced if they had not been included in the treatment may not be the same as the path of response variables that regions in the control group actually experienced. In the case of a quasi-natural experiment where assignment is possibly not entirely random, the treatment assignment must be mean-independent of factors that affect the trend in the response variables. Roth et al. (2023, p. 2222) clarify that if the bias for selecting into the treatment is the same each year, then parallel trends allow for selection bias. Excluding converging trends is not synonymous of validating parallel trends, however. Trends can be non-parallel because they might *diverge*. Also discovering that the trends diverge would be a sign that the treatment assignment may not be mean-independent of factors that affect the trend in the response variables and that there is absence of parallel trends. Hence, discovering that response variables do no converge is not necessarily a sign of parallel trends.

Bearing in mind the above limitation we propose an application of the *log*(t) convergence *club test* of Phillips and Sul (2007, 2009) analysis. Phillips and Sul (2007, 2009) proposed the novel "*log*(*t*)" regression test approach to test for the convergence hypothesis based on a nonlinear time-varying factor model. The proposed approach has the following merits: first, it accommodates heterogeneous agent behaviour and evolution in that behaviour. Second, the proposed test does not impose any particular assumptions concerning trend stationarity

or stochastic non-stationarity and consequently it is robust to the stationarity property of the series. In other words, it could be used whatever the unit root results reported in Table 1. Yet, to make the test more adherent to the time span of the Italian tax treatment we conduct it: a) using the entire sample period (1995 - 2021) and for the pre and post treatment periods; b) using two subsets (treated and control) of data.

Results of *log*(*t*) tests are reported below bin Tables 3a, 3b, and 3c.

Table 3a: Phillips-Sul log(t) tests of convergence. t-stat in parentheses (H0 is that there is convergence)

Response Variables (treated and untreated regions)	1995 - 2021	1995 - 2006	2007 – 2021
Total Employment	-0.7846 (-185.436) Rejected	-0-7846 (-185.436) Rejected	-0.9543 (-44.2028) Rejected
Self-Employment	-1.0471 (-3.9066) Rejected	-1-0471 (-3.9066) Rejected	-1.3775 (-5.4437) Rejected
New VAT Certificates	-0.8194 (-37.4199) Rejected		-0.9361 (-55.5262) Rejected
Regional GDP Growth p-c	2.7964 (1.85) **H_0 Not Rejected**	-4.6937 (-2.0528) H_0 Rejected	3.3804 (2.1556) **H_0 Not rejected**
log Regional GDP Growth p-c HP filter)	-0.8112 (-253.81) Rejected	-0.9415 (-40.8417) Rejected	-0.9571 (-53.9166) Rejected
Regional GDP p-c	-0.8233 (-23.3134) Rejected	-0.9610 (-26.2895) Rejected	-0.8512 (-32.0493) Rejected
Family Consumption	-0.8788 (-82.5425) Rejected	-1.0069 (-53.7084) Rejected	-0.9222 (-37.5706) Rejected

Notes. The Table shows the results of a *log*(t) test for overall convergence of the response variables used in the DID analysis (see Table 1) proposed by Phillips and Sul (2007, 2009) in order to investigate the presence of convergence by adopting the Andrews estimator of long-run variance (fixed or adaptive bandwidth of the kernel) based on Quadratic Spectral kernel. If the t-test of log(t) test (reported in parentheses) is smaller than -1.65, then the null hypothesis of convergence of the variable is rejected at 5% CI. Data availability of New VAT Certificates permits to run the test for the post treatment period only. Results show that during the pre-treatment period the hypothesis of convergence is not consistent with the data and this finding does not contradict the DID parallel trend hypothesis. In one case (regional GDP growth per capita), post treatment period and the entire period show convergence.

Table 3b: Phillips-Sul log(t) tests of convergence. t-stat in parentheses (Untreated regions only)

Response Variables (untreated regions)	1995 - 2021	1995 2006	2007 – 2021
Total Employment	-0.7667 (-160.2177) Rejected	-1.0048 (-33.1669) Rejected	-0.9503 (-44.9480) Rejected
Self-Employment	-1.2245 (-6.6226) Rejected	-2.6070 (-18.1581) Rejected	-1.0843 (-4.5208) Rejected
New VAT Certificates	-0.7446 (-30.1970) Rejected		-0.9188 (-48.3598) Rejected
Regional GDP Growth p-c	2.2234 (1.4464) **H_0 not rejected**	-4.6663 (-2.2425) Rejected	2.8482 (1.2841) **H_0 not rejected**
log Regional GDP Growth p-c HP filter)	-0.7526 (-153.3184) Rejected	-0.9037 (-37.4473) Rejected	-0.9260 (-50.2768) Rejected
Regional GDP p-c	-0.7894 (-18.6610)	-0.9299 (-23.6503)	-0.8154 (-34.8615)

	Rejected	Rejected	Rejected
Family Consumption	-0.8783 (-78.1308) Rejected	-1.0022 (-54.3171) Rejected	-0.9188 (-40.0264) Rejected

Notes. The Table shows the result of log(t) test of convergence for the untreated regions only. Data availability of New VAT Certificates permits to run the test for the post treatment period only. See notes below Table 2 for interpretation.

Table 3c: Phillips-Sul log(t) tests of convergence. t-stat in parentheses (Treated regions only)

Response Variables (untreated regions)	1995 - 2021	1995 2006	2007 – 2021
Total Employment	-0.8333 (-137.0381) Rejected	-1.0310 (-34.5509) Rejected	-0.9679 (-39.6925) Rejected
Self-Employment	-2.1374 (-2.3277) Rejected	-2.0493 (-1.5835) **H_0 not rejected**	-1.0843 (-4.5208) Rejected
New VAT Certificates	-0.8244 (-62.8409) Rejected		-0.9373 (-35.9240) Rejected
Regional GDP Growth p-c	3.2923 (2.1344) **H_0 not rejected**	-3.8758 (-1.6369) **H_0 not rejected**	1.4841 (0.8986) **H_0 not rejected**
log Regional GDP Growth p-c HP filter)	-0.7645 (-44.4023) Rejected	-1.0604 (-44.7164) Rejected	-0.8736 (-35.9767) Rejected
Regional GDP p-c	-0.6442 (-13.5016) Rejected	-1.2055 (-24.0175) Rejected	-0.7757 (-19.9970) Rejected
Family Consumption	-0.8118 (-82.0494) Rejected	-1.0345 (-51.2356) Rejected	-0.9468 (-33.0851) Rejected

Notes. The table shows the result of log(t) test of convergence for the untreated regions only. Data availability of New VAT Certificates permits to run the test for the post treatment period only. See notes below Table 2 for interpretation. The Table shows that convergence seems consistent with the data for Self-Employment among Treated regions during the pre-treatment period. If the t-test of *log*(t) test is smaller than -1.65, then the null of convergence of the variable is rejected at 5% CI. Data availability of New VAT Certificates permits to run the test for the post treatment period only.

Results show that during the pre-treatment period the hypothesis of convergence is not consistent with the data and this finding accords with the DID parallel trend hypothesis. In one case (*per-capita GDP growth*), the null hypothesis of convergence cannot be rejected at usual significance levels for the entire time period and for the post treatment period (this contradicts the P-trend test reported in Table 2 for that variable) but no convergence is found for the pre-treatment period.

Obtaining the rejection of the null of convergence of the response variables when pooling treated and control observations just indicate that the series do not converge. However, it does not exclude that trends could be not parallel because they *diverge*. Yet, a different result (not rejecting the null hypothesis of convergence) would be a clear contradiction of the parallel trend hypothesis and that is why we propose this test (admittedly having a possible indeterminate outcome) *in addition* to those reported in Table 2.

6. Conclusions

We have analysed the effects of an Italian binary income tax treatment that was adopted at the end of 2006 and remained on afterwards. It affected the same treated group of regions with no variation in treatment timing and assignment. TWFE panel DID results provide robust evidence that labour supply and other regional outcome growth-related variables are responsive to income tax hikes. Robustness analysis conducted by using a multi-level TWFE DID panel, which includes provincial and regional level cofactors, support the results.

Large labour supply extensive responses to income tax treatment are the main finding of our analysis. Results show that, in particular, self-employment is responsive to tax hikes. Our findings are consistent with the evidence presented on labour force participation by previous literature that the supply of professional self-employed person is adversely affected by the tax increase. In our estimates, readjustments of that segment of the labour supply to a long-run trend are observable 3 years after the treatment. Anticipatory effects are not consistent with the data, and this may be relevant in terms of future policy measures. Moreover, the absence of anticipatory effects somehow implies that the tax bases of the two income taxes analysed in this paper may be more rigid to anticipatory adjustments than the tax base given by the general personal income tax (called IRPEF by the Italian tax legislation). The specific result about New VAT Certificates indicates that exogenous modification of regional taxation has robust extensive repercussion on the supply of professional services, possibly delating or cancelling the decisions of potential operators to enter their specific official markets or leading some of them to enter the underground economy. Results on per-capita income and family consumption also show a negative effect of income tax hikes for they show a negative reaction to income tax hikes and are consistent with the above labour supply findings.

Declarations

- Conflict of interest/Competing interests: The authors have no competing interests to declare that are relevant to the content of this article.
- Funding: The authors did not receive support from any organization for the submitted work.
- Data availability and code: Data used in the paper are public (source: Eurostat and Italian National Statistics Office - ISTAT). All results presented in this paper can be reproduced using the R and Stata 18 software. The codes were developed entirely by the authors. Data and codes can be provided upon request to the authors.
- Figures: All images included in the paper were created by the authors and do not require any publication permission.

References

Aghion, P., & Howitt, P. (1990). A model of growth through creative destruction. In: National Bureau of Economic Research Cambridge, Mass., USA.

Agrawal D., D. Foremny, & C. Martínez-Toledano (2024). Wealth Tax Mobility and Tax Coordination, Cesifo w.p. 11048, April

Akcigit, U., Baslandze, S., & Stantcheva, S. (2016). Taxation and the International Mobility of Inventors. American economic review, 106(10), 2930-2981. doi:10.1257/aer.20150237

Appelbaum, E., & Katz, E. (1986). Measures of Risk Aversion and Comparative Statics of Industry Equilibrium. American economic review, 76(3), 524-529.

Auerbach, A. J., & Siegel, J. M. (2000). Capital-Gains Realizations of the Rich and Sophisticated. American economic review, 90(2), 276-282. doi:10.1257/aer.90.2.276

Barrios, S., Huizinga, H., Laeven, L., & Nicodème, G. (2009). International taxation and multinational firm location decisions. Retrieved from https://EconPapers.repec.org/RePEc:euf:ecopap:0356

Baskaran, T. (2021). The revenue and base effects of local tax hikes: evidence from a quasi-experiment. International Tax and Public Finance, 28(6), 1472-1518. doi:10.1007/s10797-021-09657-2

Blau, D. M. (1987). A Time-Series Analysis of Self-Employment in the United States. Journal of Political Economy, 95(3), 445-467. Retrieved from http://www.jstor.org/stable/1831972

Bosco B., C.F Bosco, and P. Maranzano (2024). Income taxation and regional growth: DID evidence from an income tax treatment in Italy. ResearchSquare https://doi.org/10.21203/rs.3.rs-3962177/v1

Caruso, E., & Dirindin, N. (2019). Salute ed Economia-Questioni di economia e politica sanitaria: Il mulino.

Da Rin, M., Di Giacomo, M., & Sembenelli, A. (2011). Entrepreneurship, firm entry, and the taxation of corporate income: Evidence from Europe. Journal of Public Economics, 95(9), 1048-1066. doi:https://doi.org/10.1016/j.jpubeco.2010.06.010

De Philippis, M., Locatelli, A., Papini, G., & Torrini, R. (2022). La crescita dell'economia italiana e il divario Nord-Sud: trend storici e prospettive alla luce dei recenti scenari demografici [Italian economic growth and the North-South gap: Historical trends and future projections in light of the recent demographic scenarios]. Bank of Italy Occasional Paper(683).

Djankov, S., Ganser, T., McLiesh, C., Ramalho, R., & Shleifer, A. (2010). The Effect of Corporate Taxes on Investment and Entrepreneurship. American Economic Journal: Macroeconomics, 2(3), 31-64. doi:10.1257/mac.2.3.31

Dynan, K. E., Skinner, J., & Zeldes, S. P. (2004). Do the Rich Save More? Journal of Political Economy, 112(2), 397-444. doi:10.1086/381475

Eissa, N., & Liebman, J. B. (1996). Labor Supply Response to the Earned Income Tax Credit*. The quarterly journal of economics, 111(2), 605-637. doi:10.2307/2946689

Elhorst, J. P. (2010). Applied Spatial Econometrics: Raising the Bar. Spatial Economic Analysis, 5(1), 9-28. doi:10.1080/17421770903541772

Elhorst, J. P. (2014). Spatial econometrics: from cross-sectional data to spatial panels (Vol. 479): Springer.

Elhorst, J. P. (2024). Raising the bar in spatial economic analysis: two laws of spatial economic modelling. Spatial Economic Analysis, 1-18. doi:10.1080/17421772.2024.2334845

Esteller-Moré, A., & Solé-Ollé, A. (2002). Tax Setting in a Federal System: The Case of Personal Income Taxation in Canada. International Tax and Public Finance, 9(3), 235-257. doi:10.1023/A:1016212110137

Evans, D. S., & Leighton, L. S. (1989). The determinants of changes in U.S. self-employment, 1968–1987. Small Business Economics, 1(2), 111-119. doi:10.1007/BF00398629

Federici, D., Ferrante, F., & Parisi, V. (2023). The educational roots of persistent territorial disparities: the case of Italy. GeoJournal, 88(2), 2063-2082. doi:10.1007/s10708-022-10740-2

Goolsbee, A. (2000). What Happens When You Tax the Rich? Evidence from Executive Compensation. Journal of Political Economy, 108(2), 352-378. doi:10.1086/262122

Harris, R. D. F., & Tzavalis, E. (1999). Inference for unit roots in dynamic panels where the time dimension is fixed. Journal of Econometrics, 91(2), 201-226. doi:https://doi.org/10.1016/S0304-4076(98)00076-1

Heim, B. T. (2010). The responsiveness of self-employment income to tax rate changes. Labour Economics, 17(6), 940-950. doi:https://doi.org/10.1016/j.labeco.2010.02.010
Jakobsen, K., Jakobsen, K., Kleven, H., & Zucman, G. (2019). Wealth Taxation and Wealth Accumulation: Theory and Evidence From Denmark*. The quarterly journal of economics, 135(1), 329-388. doi:10.1093/qje/qjz032
Jappelli, T., & Pistaferri, L. (2010). The Consumption Response to Income Changes. Annual Review of Economics, 2(1), 479-506. doi:10.1146/annurev.economics.050708.142933
Johnson, D. S., Parker, J. A., & Souleles, N. S. (2006). Household Expenditure and the Income Tax Rebates of 2001. American economic review, 96(5), 1589-1610. doi:10.1257/aer.96.5.1589
Kihlstrom, R. E., & Laffont, J.-J. (1979). A General Equilibrium Entrepreneurial Theory of Firm Formation Based on Risk Aversion. Journal of Political Economy, 87(4), 719-748. doi:10.1086/260790
Kleven, H. J., Landais, C., & Saez, E. (2013). Taxation and International Migration of Superstars: Evidence from the European Football Market. American economic review, 103(5), 1892-1924. doi:10.1257/aer.103.5.1892
Levin, A., Lin, C.-F., & Chu, C.-S. J. (2002). Unit root tests in panel data: asymptotic and finite-sample properties. Journal of Econometrics, 108(1), 1-24.
Lo Cascio, I., Mazzola, F., & Epifanio, R. (2019). Territorial determinants and NUTS 3 regional performance: A spatial analysis for Italy across the crisis. Papers in Regional Science, 98(2), 641-678. doi:https://doi.org/10.1111/pirs.12372
McCarthy, J. (1995). Imperfect insurance and differing propensities to consume across households. Journal of Monetary economics, 36(2), 301-327. doi:https://doi.org/10.1016/0304-3932(95)01214-1
Merfeld, J. D. (2019). Spatially heterogeneous effects of a public works program. Journal of Development Economics, 136, 151-167. doi:https://doi.org/10.1016/j.jdeveco.2018.10.007
Meyer, B. D., & Rosenbaum, D. T. (2001). Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers*. The quarterly journal of economics, 116(3), 1063-1114. doi:10.1162/00335530152466313
Moretti, E., & Wilson, D. J. (2017). The Effect of State Taxes on the Geographical Location of Top Earners: Evidence from Star Scientists. American economic review, 107(7), 1858-1903. doi:10.1257/aer.20150508
Muralidharan, K., Niehaus, P., & Sukhtankar, S. (2023). General Equilibrium Effects of (Improving) Public Employment Programs: Experimental Evidence From India. Econometrica, 91(4), 1261-1295. doi:https://doi.org/10.3982/ECTA18181
Nguyen, A. D. M., Onnis, L., & Rossi, R. (2021). The Macroeconomic Effects of Income and Consumption Tax Changes. American Economic Journal: Economic Policy, 13(2), 439-466. doi:10.1257/pol.20170241
Oni, M. H. (2023). Progressive income taxation and consumption baskets of rich and poor. Journal of Economic Dynamics and Control, 157, 104758. doi:https://doi.org/10.1016/j.jedc.2023.104758
Parker, J. A. (1999). The Reaction of Household Consumption to Predictable Changes in Social Security Taxes. American economic review, 89(4), 959-973. doi:10.1257/aer.89.4.959
Parker, J. A., Souleles, N. S., Johnson, D. S., & McClelland, R. (2013). Consumer Spending and the Economic Stimulus Payments of 2008. American economic review, 103(6), 2530-2553. doi:10.1257/aer.103.6.2530
Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. Journal of applied econometrics, 22(2), 265-312.
Phillips, P. C. B., & Sul, D. (2007). Transition Modeling and Econometric Convergence Tests. Econometrica, 75(6), 1771-1855. doi:https://doi.org/10.1111/j.1468-0262.2007.00811.x
Phillips, P. C. B., & Sul, D. (2009). Economic transition and growth. Journal of applied econometrics, 24(7), 1153-1185. doi:https://doi.org/10.1002/jae.1080
Piketty, T., Saez, E., & Stantcheva, S. (2014). Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities. American Economic Journal: Economic Policy, 6(1), 230-271. doi:10.1257/pol.6.1.230
Ramey, V. A. (2011a). Can Government Purchases Stimulate the Economy? Journal of Economic Literature, 49(3), 673-685. doi:10.1257/jel.49.3.673
Ramey, V. A. (2011b). Identifying Government Spending Shocks: It's all in the Timing*. The quarterly journal of economics, 126(1), 1-50. doi:10.1093/qje/qjq008

Robson, M. T., & Wren, C. (1999). Marginal and Average Tax Rates and the Incentive for Self-Employment. Southern Economic Journal, 65(4), 757-773. doi:https://doi.org/10.1002/j.2325-8012.1999.tb00198.x

Romer, C. D., & Romer, D. H. (2010). The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks. American economic review, 100(3), 763-801. doi:10.1257/aer.100.3.763

Roth, J., Sant'Anna, P. H. C., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature. Journal of Econometrics, 235(2), 2218-2244. doi:https://doi.org/10.1016/j.jeconom.2023.03.008

Saez, E. (2002). Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses*. The quarterly journal of economics, 117(3), 1039-1073. doi:10.1162/003355302760193959

Saez, E., Matsaganis, M., & Tsakloglou, P. (2012). Earnings Determination and Taxes: Evidence From a Cohort-Based Payroll Tax Reform in Greece *. The quarterly journal of economics, 127(1), 493-533. doi:10.1093/qje/qjr052

Souleles, N. S. (1999). The Response of Household Consumption to Income Tax Refunds. American economic review, 89(4), 947-958. doi:10.1257/aer.89.4.947

Wen, J.-F., & Gordon, D. V. (2014). An Empirical Model of Tax Convexity and Self-Employment. The Review of Economics and Statistics, 96(3), 471-482. doi:10.1162/REST_a_00388

Widmann, R. (2023). Immigrant inventors and local income taxes: Evidence from Swiss municipalities. Journal of Public Economics, 219, 104822. doi:https://doi.org/10.1016/j.jpubeco.2023.104822

Wikle, C. K., Zammit-Mangion, A., & Cressie, N. (2019). Spatio-temporal Statistics with R: Chapman and Hall/CRC.

Wooldridge J. M. (2021), Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Difference-in-Differences Estimators. doi: http://dx.doi.org/10.2139/ssrn.3906345

Zidar, O. M. (2019). Tax Cuts for Whom? Heterogeneous Effects of Income Tax Changes on Growth and Employment. Journal of Political Economy, 127(3), 1437-1472. doi:10.1086/701424

Ziliak J. P. & T. J. Kniesner (2005). The Effect of Income Taxation on Consumption and Labor Supply. Journal of Labor Economics, 23, 4, 769-796

Supplementary material

Appendix 1: The 2006/07 regional tax policy in Italy[9]

The appendix provides a brief description of the two Italian taxes involved in the treatment.

IRAP

From 1997 Italian corporate entities and individual natural persons generating income in the form of business income (*Reddito d'impresa*) as well as non-resident companies, but only on Italian source income, are subject to a regional production tax, IRAP. IRAP taxpayers include those carrying out business activities and self-employment, operating both individually and in partnership, private non-commercial entities as well as public administrations and other public bodies.

IRAP taxes the production activity, the exchange of goods and the supply of services. In the paper IRAP is viewed as an Income Tax *in disguise* because it is charged on the value of net production resulting from the business pursued within the region. For commercial business and professionals, state administrations, regions, provinces, municipalities and generally for public administrations whose main object is not commercial business, the basis of assessment is the value of remunerations. The 90% of its regional yield is destined to the financing of the health system in each regional (in addition to central transfers and other funds). As a rule, IRAP tax rates are fixed by the regional governments within an interval defined by the national law that leaves some degree of autonomy to the regional authorities. Hence, although the tax structure is completely defied by the parliamentary law of 1997, IRAP tax rates may vary across regions and time.

The regional Income Tax Surcharge

The Regional income tax surcharge was introduced in the Italian tax system in December 1997 (D.Lgs. number 446/1997) and it is applied from the 1998 tax year. The regional surcharge is a regional tax, whose rate is added on top of those rates taxing the income of natural persons at the national level (IRPEF). Then, it falls on the taxable income as it is defined for IRPEF purposes net of any applicable tax deductions and credits for incomes produced abroad. Tax rates vary from region to region and may increase in relation to the taxpayer's income bracket. The ordinary regional income tax rate ranges from 1.23% to 3.33%. In the event that the region decides not to adopt a single rate but a plurality of differentiated rates, the latter must be structured exclusively in relation to the same income brackets established for the IRPEF tax and then tax rates must be increasing according to each bracket. The regions can arrange tax deductions in favour of the family by increasing those already existing for national IRPEF purposes and can also adopt other economic support measures in favour of IRPEF subjects. During the period of the present study, the regional surcharge rates varied greatly between regions. In 2000, the surtax rate was determined for the entire Italian territory by the national law at the level of 0.9 but each regions had the power to increase that rate up to 1.4. In that year, only Lombardy, Marche and Veneto applied the additional taxation on a true progressive scale. At the end of our period (2021), some regions (Abruzzo, Basilicata, Calabria (t = 1.73), Campania (t = 2.03), Sardinia (t = 1.23), Sicily (t = 1.23), Valle d'Aosta (t = 1.23), Veneto (t = 1.23) maintained (or switched to) the proportional (constant tax rate) surcharge tax: https://www1.finanze.gov.it/ finanze2/ dipartimentopolitichefiscali/ fiscalita locale/ addregirpef/ sceltaregione.htm?cm=).

The following figure gives a synthetic view of the Italian Income Tax Treatment introduced at the end of 2006. From 2007 taxpayers in treated regions faced a uniform increase of the IRAP and Regional Surcharge rates. Uniformity means that the statutory rates were increased by the same percentage for all treated taxpayers without differences related the value of the tax base for bottom- and top-income groups.

[9] This Appendix relies on Bosco, Bosco, and Maranzano (2024).

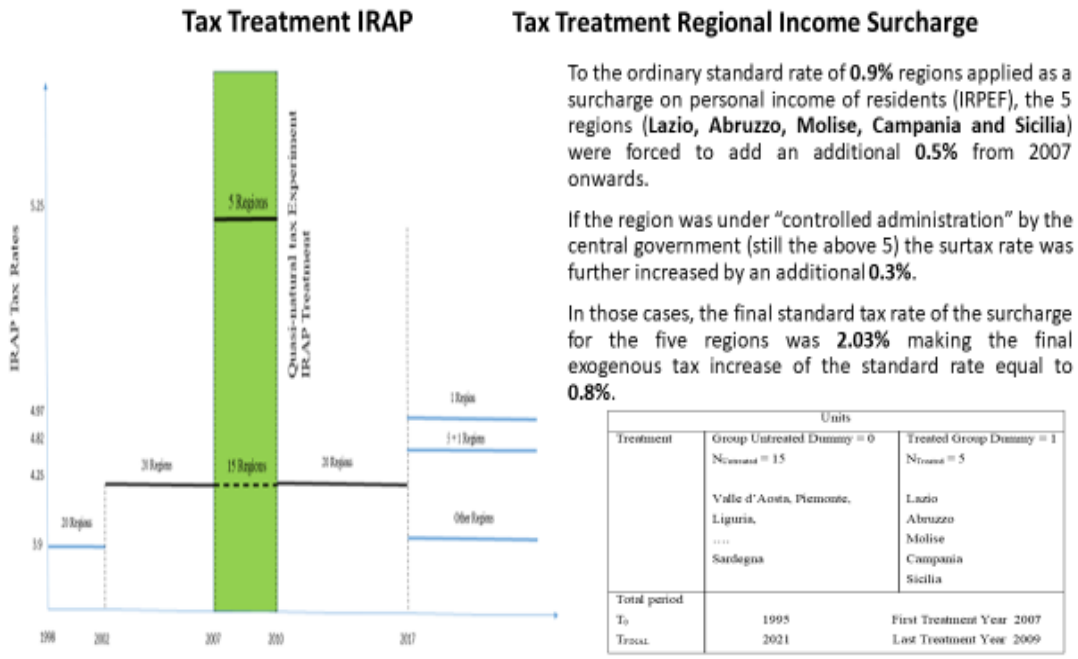


Figure A 1*: The Italian Tax Treatment of the end of 2006. The chart describes how the tax rates of IRAP (left) and Regional Income Surcharge (right) have been modified by the Italian Budget Law of 2006 (Law 27/12/2006, n. 296, Disposizioni per la formazione del bilancio annuale e pluriennale dello Stato; legge finanziaria 2007).). The figure includes a box with the indication of the five treated regions and the time span of our DID model. The last year before treatment is 2006.*

Appendix 2 Additional figures and results

This section of the appendix provides additional figures on the available data by depicting the regional distribution of the considered macroeconomic indicators from 1995 to 2021.

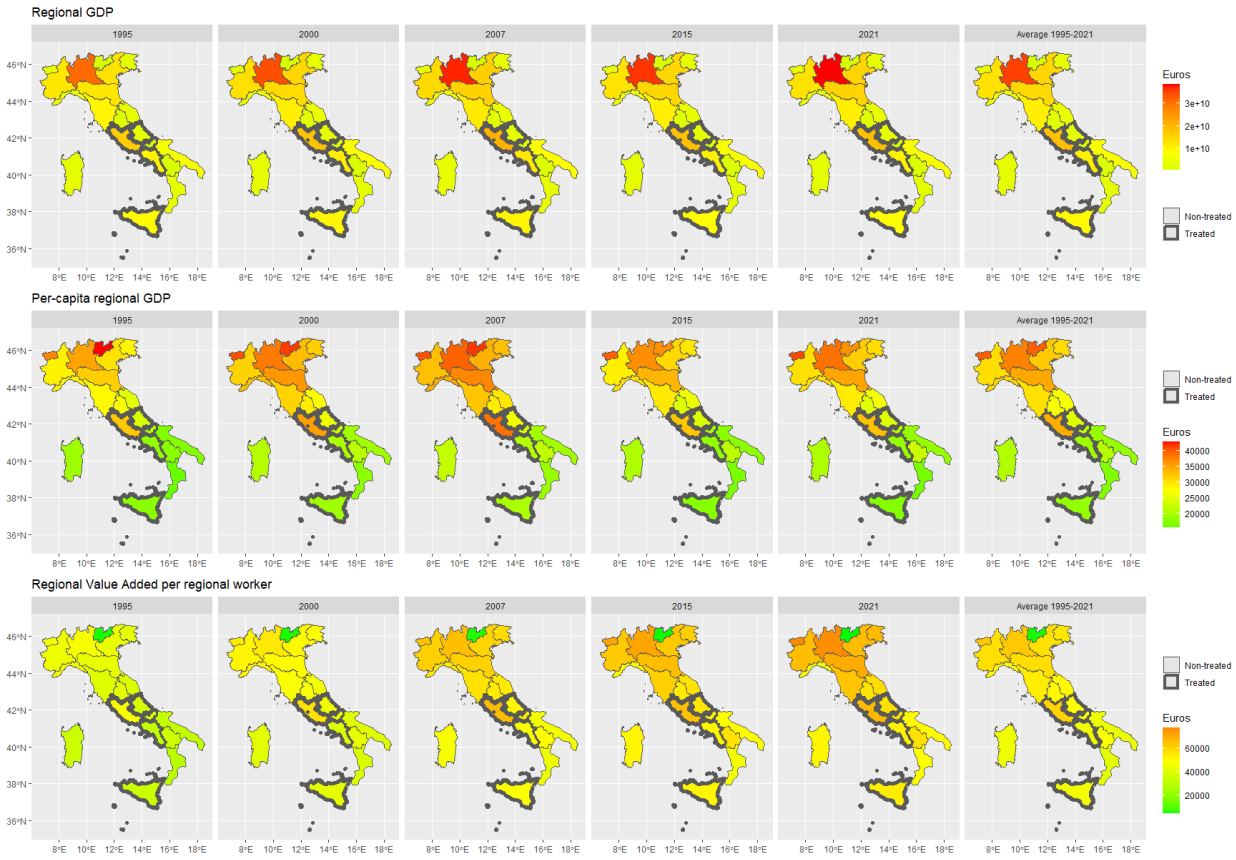


Figure A2: Observed values of overall GDP, per capita GDP, and value added per worker across the Italian regions from 1995 to 2021. The last column depicts the regional average value 1995-2021. Regions with a marked border are those treated (Lazio, Campania, Abruzzo, Molise, and Sicily).

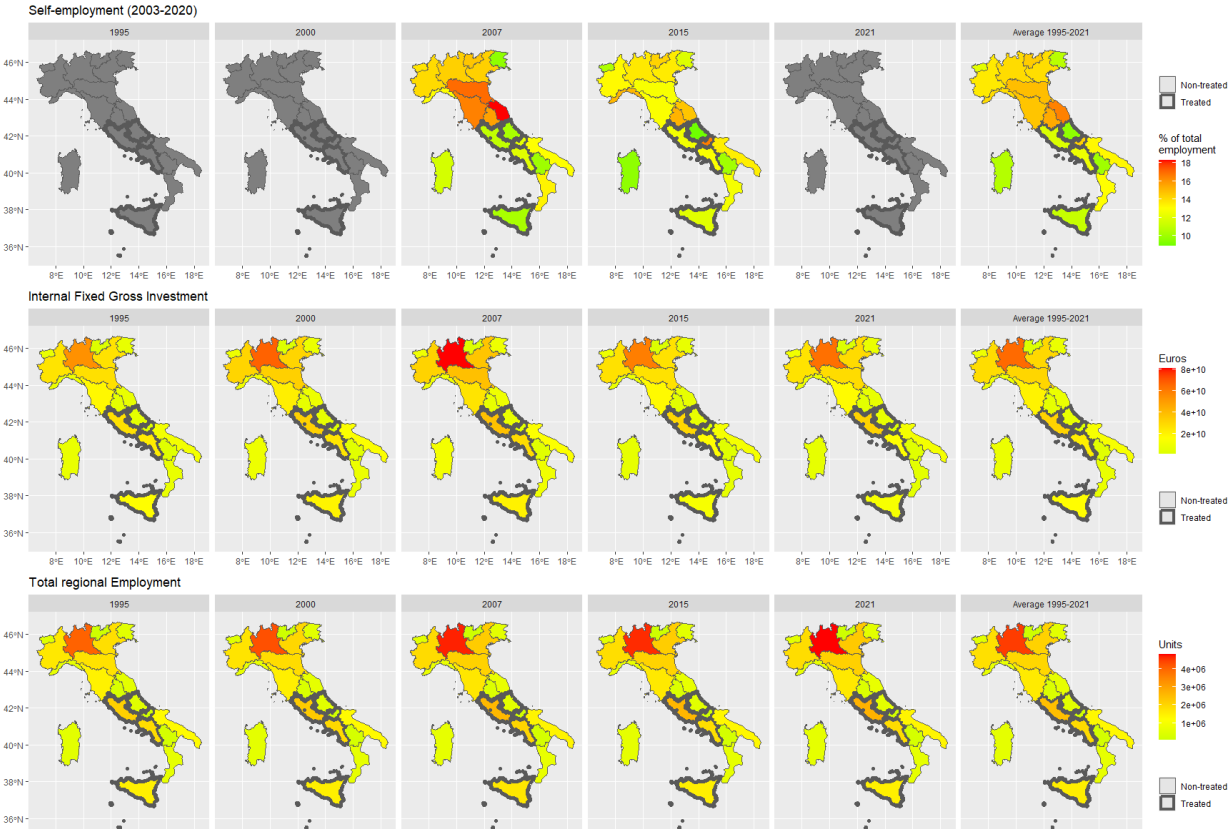


Figure A3: Observed values of self-employment, internal fixed gross investment, and total employment across the Italian re from 1995 to 2021. The last column depicts the regional average value 1995-2021. Regions with a marked border are treated (Lazio, Campania, Abruzzo, Molise, and Sicily)

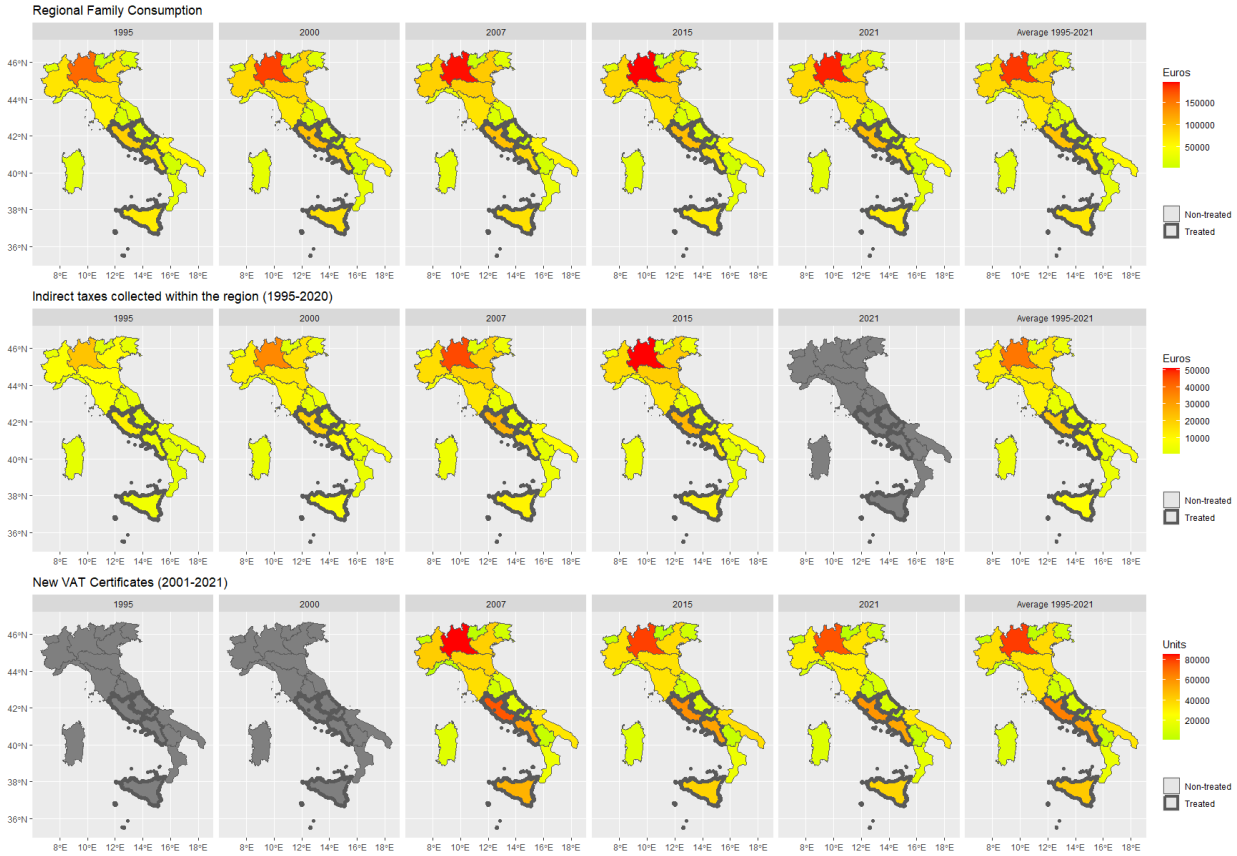


Figure A4: Observed values of family consumption, indirect taxes collected, and new VAT certificates across the Italian regions from 1995 to 2021. The last column depicts the regional average value 1995-2021. Regions with a marked border are those treated (Lazio, Campania, Abruzzo, Molise, and Sicily).

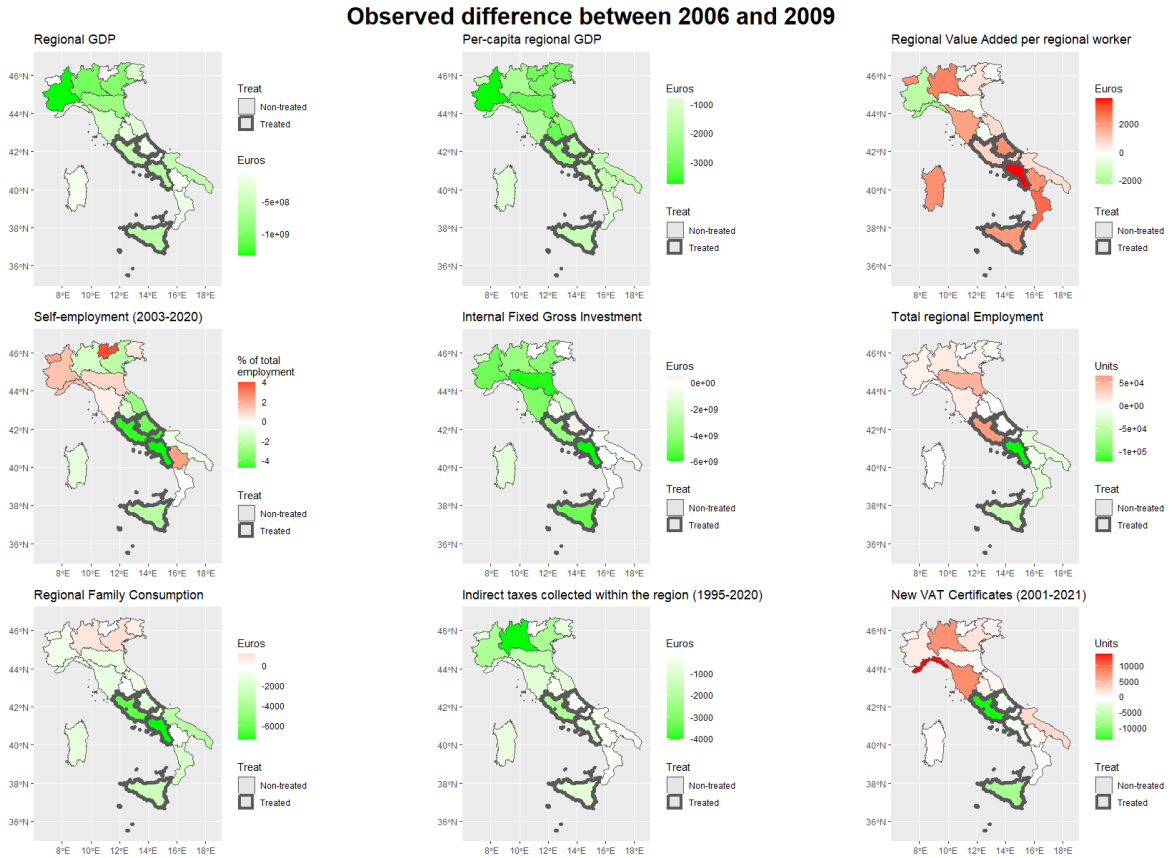


Figure A5: Absolute variation from 2006 to 2009 for the available macroeconomic indicators at the regional level. The five treated regions are highlighted using bold regional borders.

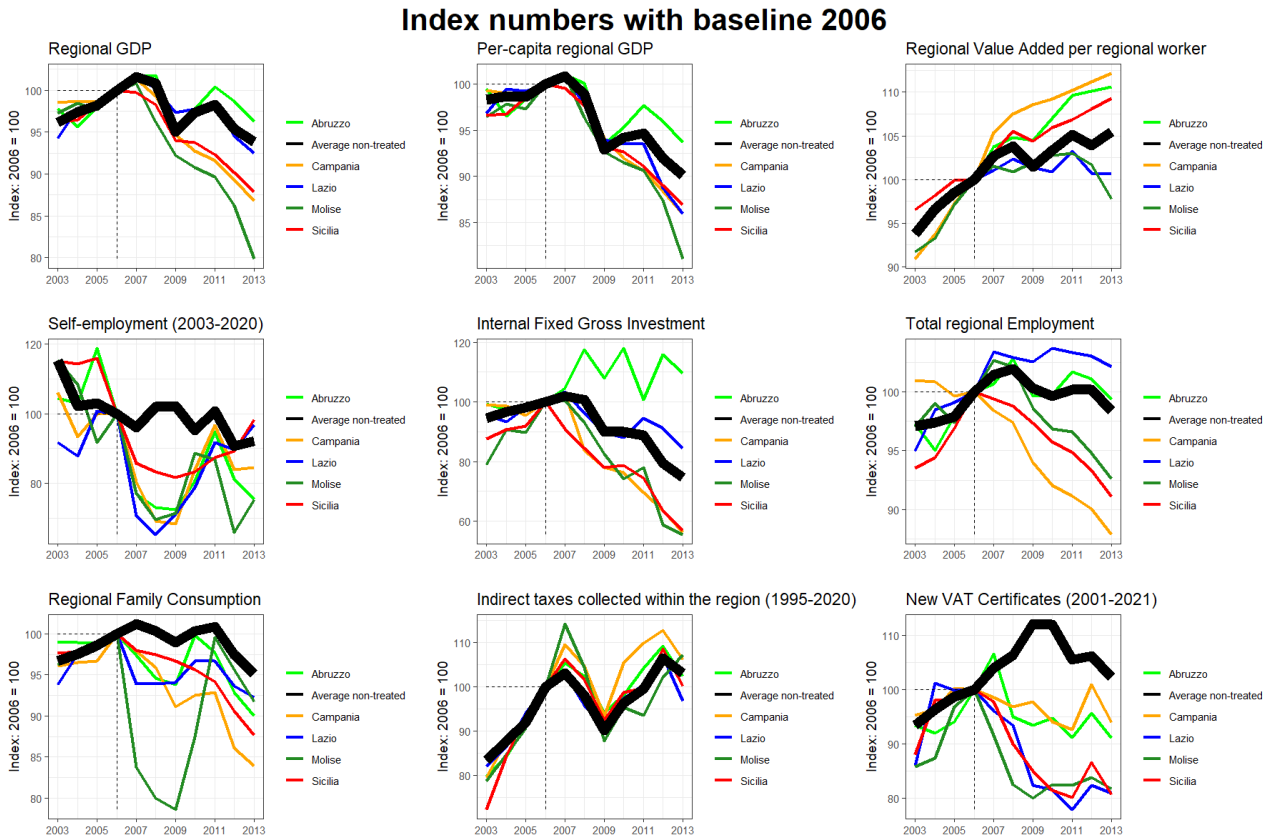


Figure A6: Time series of index numbers with baseline 2006=100 for the available macroeconomic indicators. Estimated values represent the percentage variation for a given year from the baseline year (2006). Abruzzo, Campania, Lazio, Molise and Sicilia are the five treated regions, while the non-treated regions are averaged across year (solid black lines).

Table 3A: DID estimations of the Italian income tax treatment (time effects included)

Response Variables	**Total Employment**	**Self-Employment**	**New VAT Certificates**	**Regional GDP Growth per capita**	**Regional GDP Growth per capita (HP filter)**	**Regional GDP per capita**	**Log Family Consumption**
ATET or CATET							
Treated vs untreated regions	-29401.111	-1.1650436**	-4592.5248**	-.00492499*	-.03186035***	-629.18586***	-.049**
1996	6690**				-0.0045002	-259.30395	.0064782*
1997	10230**			0.00385154	-0.00703847	-332.96859	.0391751***
1998	21495***			0.00627141	-0.00637149	-115.83798	.0688513***
1999	33675***			0.00675496	-0.00703858	93.668916	.0924181***
2000	55815***			.01882805***	-0.01289044	601.68919	.118144***
2001	78415***		(base)	0.00002573	-0.01804747	656.04639	.1233866***
2002	98220***		362.62472*	-.0215791***	-0.01728926	382.62437	.1187407***
2003	115750***		1009.5801**	-.0232905**	-0.01858051	73.910492	.1240574***
2004	123135***	-1.3032172***	2759.9581***	-.01468758**	-0.02405906	-107.40569	.1329757***
2005	129990***	-0.79481413	3532.96***	-.01878086**	-0.02915511	-312.69061	.1399013***
2006	154155***	-1.1050093**	4066.7757***	-0.00507515	-0.03699559	-116.92608	.1539774***
2007	176985.28***	-1.7357198***	6032.483***	-0.01214461	-0.03874279	-2.1277481	.156496***
2008	179790.28***	-1.2009668**	6118.9498***	-.04189799***	-0.04368337	-687.60307	.1438508***
2009	158895.28***	-1.4413136***	6257.9583**	-.07688143***	-0.03531405	-1955.6365	.1276458***
2010	150955.28***	-1.6152737***	6245.3666**	-0.01254494	-0.04403754	-1832.6498	.143082***
2011	154300.28***	-0.57854018	5098.0129*	-.01639645*	-0.05142562	-1823.484	.1476519**
2012	150960.28***	-2.1962287***	5637.865*	-.04757515***	-0.04597545	-2314.2495	.1080306***
2013	128780.28***	-1.896113**	4588.6282	-.04166368***	-0.04600308	-2823.6999*	.0820907***
2014	129685.28***	-2.5526536***	6938.0525**	-.02099425**	-0.04919986	-2855.4027*	.0806337***
2015	137625.28***	-2.0213505**	4246.2576	-0.00320001	-0.0591872	-2712.2537*	.101242***7
2016	154245.28***	-1.8697245**	3578.8324	-0.00947221	-.06638269*	-2543.8082	.1122534***
2017	168730.28***	-1.2597971	4181.602	-0.00137235	-.07464948*	-2204.5669	.1276559***
2018	180345.28***	-2.1817539**	3998.3382	-0.00729039	-.08428351**	-2043.9405	.1382362***
2019	186980.28***	-1.5534404*	3778.6483	-0.01148007	-.09223028**	-1955.0961	.1414657**
2020	159470.28***	-3.1028488***	2645.0976	-.10811968***	-.07597091**	-3679.2373**	.0235429
2021	166375.28***		3013.8352	.04887077***	-.09266339**	-2413.492	.0738984**
Controls							
log(GFCF)		-0.79632567	-453.96235	0.00989756	.08774462***	2241.2117***	
Log(Value Added per worker)		-11.716371***	-13460.289	0.03106403	.23681953**	9867.1645**	
Const.	1095230***	160.11159***	178491.2	-.54361653*	5.672762***	-128581.29***	10.29218***
N	540	355	420	520	540	540	540

Ptrend Test: H_0 **Linear parallel trend**	F(1,19) = 0.04 Prob > F = 0.8459 **H_0 cannot be rejected**	F(1, 19) = 2.30 Prob > F = 0.1455 **H_0: cannot be rejected**	F(1, 19) = 2.02 Prob > F = 0.1715 **H_0: cannot be rejected**	F(1, 19) = 0.02 Prob > F = 0.9030 **H_0: cannot be rejected**	F(1, 19) = 1.18 Prob > F = 0.2903 **H_0 cannot be rejected**	F(1, 19) = 0.34 Prob > F = 0.5674 **H_0 cannot be rejected**	F(1, 19) = 0.62 Prob > F = 0.4398 **H0 cannot be rejected**
Granger Type Test: H_0 **No Anticipation effects**	F(11, 19) = 4.00 Prob > F = 0.0040 **Reject H_0**	F(3, 19) = 1.50 Prob > F = 0.2466 **H_0 cannot be rejected**	F(5, 19) = 1.34 Prob > F = 0.2893 **H_0 cannot be rejected**	F(10, 19) = 9.99 Prob > F = 0.0000 **Reject H_0**	F(11, 19) = 11.29 Prob > F = 0.0000 **Reject H_0**	F(11, 19) = 7.29 Prob > F = 0.0001 **Reject H_0**	F(11, 19) = 7.12 Prob > F = 0.0001 **Reject H_0**

FONDAZIONE ENI ENRICO MATTEI WORKING PAPER SERIES

Our Working Papers are available on the Internet at the following address:
https://www.feem.it/pubblicazioni/feem-working-papers/

"NOTE DI LAVORO" PUBLISHED IN 2024

1. A. Sileo, M. Bonacina, The automotive industry: when regulated supply fails to meet demand. The Case of Italy
2. A. Bastianin, E. Mirto, Y. Qin, L. Rossini, What drives the European carbon market? Macroeconomic factors and forecasts
3. M. Rizzati, E. Ciola, E. Turco, D. Bazzana, S. Vergalli, Beyond Green Preferences: Alternative Pathways to Net-Zero Emissions in the MATRIX model
4. L. Di Corato, M. Moretto, Supply contracting under dynamic asymmetric cost information
5. C. Drago, L. Errichiello, Remote work admist the Covid-19 outbreak: Insights from an Ensemble Community-Based Keyword Network Analysis
6. F. Cappelli, Unequal contributions to CO2 emissions along the income distribution within and between countries
7. I. Bos, G. Maccarrone, M. A. Marini, Anti-Consumerism: Stick or Carrot?
8. M. Gilli, A. Sorrentino, The Set of Equilibria in Max-Min Two Groups Contests with Binary Actions and a Private Good Prize
9. E. Bachiocchi, A. Bastianin, G. Moramarco, Macroeconomic Spillovers of Weather Shocks across U.S. States
10. T. Schmitz, I. Colantone, G. Ottaviano, Regional and Aggregate Economic Consequences of Environmental Policy
11. D. Bosco, M. Gilli, Effort Provision and Incentivisation in Tullock Group-Contests with Many Groups: An Explicit Characterisation
12. A. Drigo, Environmental justice gap in Italy: the role of industrial agglomerations and regional pollution dispersion capacity
13. P. I. Rivadeneyra García, F. Cornacchia, A. G. Martínez Hernández, M. Bidoia, C. Giupponi, Multi-platform assessment of coastal protection and carbon sequestration in the Venice Lagoon under future scenarios
14. T. Angel, A. Berthe, V. Costantini, M. D'Angeli, How the nature of inequality reduction matters for CO2 emissions
15. E. Bacchiocchi, A. Bastianin, T. Kitagawa, E. Mirto, Partially identified heteroskedastic SVARs

FONDAZIONE ENI ENRICO MATTEI

Fondazione Eni Enrico Mattei
Corso Magenta 63, Milano – Italia

Tel. +39 02 403 36934

E-mail: letter@feem.it
www.feem.it

Fondazione Eni Enrico Mattei