The impact of Covid-19 restrictions on economic activity: evidence from the Italian regional system

Brian Cepparulo*

Robert Calvert Jump†

August 24, 2022

Abstract

Non-pharmaceutical interventions adopted by governments to halt the spread of Sars-Cov2 are thought to have non-trivial consequences for the economy. The purpose of this paper is to estimate the economic impact of non-pharmaceutical interventions in Italy, by taking advantage of timing differences in their implementation across regions and employing mobility data to proxy activity. To achieve this, we estimate one-way and two-way fixed effects models on a large sample of Italian provinces. We also isolate a set of well-defined quasi-natural experiments in which one region goes from a lower to a higher tier of restrictions, while a neighbouring region remains in the lower tier, for which we can estimate difference-in-differences and continuous treatment models. Moreover, in order to observe whether the impact of restrictions has changed over time, we split the sample around December 2020 and replicate the analysis in each subsample. Our case studies indicate that an Italian province moving from tier 2 to tier 3 in the system of restrictions can expect a fall in mobility of between 12 and 18 percentage points. Thus, we provide evidence of the negative effects of non-pharmaceutical interventions on economic activity. Finally, we provide some evidence that the effectiveness of NPIs in reducing mobility is likely to reduce over time, which has important policy implications. Our estimations are robust to a number of checks.

Keywords: Covid-19; Lockdown; Non-pharmaceutical intervention; Mobility.

JEL Codes: C21; C23; H12; I18; R12.

Acknowledgements: We would like to thank Özlem Onaran and Alex Guschanski for extensive advice and comments. We are also extremely grateful to Andrea Fracasso, Colombe Ladreit, Karsten Kohler and Andreas Maschke for precious comments and revisions, as well as to Clément de Chaisemartin and Xavier D'Haultfœuille for econometric guidance. Moreover we would like to thank the participants in the 5th Nordic Post Keynesian Conference, the 13th PKES PhD conference in Greenwich, the 8th Annual Conference of the International Association for Applied Econometrics, the 2022 French Stata conference in applied econometrics. All remaining errors are the sole responsibility of the authors.

*Institute for Political Economy, Governance, Finance and Accountability, Old Royal Naval College, Park Row, Greenwich, London SE10 9LS. Email: b.cepparulo@greenwich.ac.uk

†University of Greenwich, Institute for Political Economy, Governance, Finance and Accountability, Old Royal Naval College, Park Row, Greenwich, London SE10 9LS. Email: r.g.calvertjump@greenwich.ac.uk

1 Introduction

Do lockdowns come with economic costs? As outbreaks of Covid-19 infections spread out around the world, governments set up extraordinary measures to tackle their proliferation. Many countries implemented non-pharmaceutical interventions (NPIs), including personal hygiene advice, social distancing, test and contact tracing, restrictions on school and office attendance, the implementation of stay-at-home policies, and travel bans. These restrictive measures are commonly referred to in public as 'lockdowns', although the term is somewhat vague and has been used to represent different policies in different countries. Independent of their epidemiological efficacy, NPIs are thought to have non-trivial consequences for the economy. In their *World Economic Outlook*, published in October 2020, the International Monetary Fund found that countries that adopted severe restrictions experienced larger drops in GDP, even after controlling for the epidemiological situation (WEO, 2020).

As NPIs are thought to have economic costs, a heated debate around the trade-off between 'saving lives and livelihoods' has emerged (Islamaj et al., 2021). The empirical research on this topic faces serious challenges, particularly regarding how to disentangle the effects of NPIs from those of the epidemic itself. Despite the scale of the challenge, several studies have attempted to identify the impact of NPIs on economic activity. The results are mixed, with some pointing to a greater role of individual behaviour, such as changes in consumption patterns triggered by fear of infection or voluntarily compliance with public guidelines (Sheridan et al., 2020; Maloney and Taskin, 2020; Goolsbee and Syverson, 2021; Mendolia et al., 2021; Kong and Prinz, 2020; Chetty et al., 2020; Guglielminetti and Rondinelli, 2021), and others stressing the role of government policies (Caselli et al., 2020; Coibion et al., 2020; Benzeval et al., 2020; Boone and Ladreit, 2021; Kok, 2020). The first position implies that any trade-off between the incidence of Covid-19 and the economy is limited, while the second tends to emphasise the importance of such a trade-off. A second question considered in the literature is whether or not the effects of NPIs have changed over time, perhaps due to 'lockdown fatigue' (Goldstein et al., 2021). There is, in fact, some evidence that the impact of restrictions on GDP declined over the course of 2020 (OECD, 2021). An extended review of the literature can be found in appendix B.

Most of the existing studies of the economic effects of NPIs are based on cross-country analyses, and often downplay country-specific heterogeneity in policy implementation. However, after the first wave of infections, many countries started to apply regional or locally-based system of interventions, recognizing the need of balancing the public health benefits of these policies with their economic costs. For instance, Boone and Ladreit (2021) consider the possibility that regional heterogeneity in Spain and the UK might have affected their regression estimates based on aggregate national data. Interestingly, the authors do not mention the case of Italy, which was a prominent example of a country adopting rather different systems of restrictions between its so-called first and second waves of Covid-19. In fact, while the first wave was characterized by a national lockdown, the second wave was approached with a regional-based system of interventions that allowed territories characterized by different epidemiological risks to respond with different levels of restrictions. This system, known colloquially as the color-based system, provides an ideal set of case studies to investigate the impact of NPIs on economic activity.

In this paper, we take advantage of timing differences in the implementation of NPIs in Italian regions to estimate the impact of such policies on economic activity. Broadly speaking, we address the following questions:

1. What is the impact of NPIs on economic activity?
2. Does the effect of NPIs change over time?

In answering these questions, we contribute to the literature in several ways. First of all, we contribute to the strand of literature that attempts to estimate the short-term impact of NPIs on economic activity, and shed light on the debate around the trade-off between public health and the economy. Second, we contribute to the literature investigating whether or not the effects of restrictions have changed over time. Third, we contribute to the growing literature that employs high frequency mobility data to evaluate the economic consequences of the pandemic (Deb et al., 2020; Famiglietti and Leibovici, 2022). Lastly, this paper adds to the empirical literature that estimates locally differentiated restrictions on individual mobility by means of sub-national data (Caselli et al., 2020; Gathergood and Guttman-Kenney, 2020; Walker and Hurley, 2021). Our paper complements Caselli et al. (2020), in particular, who esimate the effects of NPIs on mobility in Italy during the first wave of Covid-19.

To assess the causal impact of NPIs on economic activity, we employ a diversified empirical approach. We first analyze a large sample of Italian provinces on which we run one-way fixed effects and two-way fixed effects models. Secondly, to account for possible bias of the two-way fixed effect estimator, we isolate a set of case studies characterized by well defined quasi-natural experiments in which one region goes from a lower to a higher level of restrictions, while a neighbouring region remains in the lower tier. We estimate difference-in-differences and continuous treatment models on each of these episodes. In addition, we split the sample around December 2020, and replicate the analysis in each subsample. This method allows us to observe whether the impact of restrictions has changed over time. In all specifications we control for regional and provincial public health outcomes. Finally, we employ Google mobility data for trips to retail and recreational venues as our dependent variable.

In line with a large part of the existing literature, we interpret mobility data as an indicator of economic activity (WEO, 2020; Deb et al., 2020; Boone and Ladreit, 2021; Buono and Conteduca, 2020). Although mobility does not necessarily correspond to expenditure, it represents a useful indicator of activity in light of its strong correlation with both GDP and consumption (OECD, 2020). Mobility data also performs well in out-of-sample forecasts of growth compared with other benchmarks (Gamtkitsulashvili and Plekhanov, 2021), as well as being a good predictor of industrial production (Sampi Bravo and Jooste, 2020). Moreover, given its high frequency and geographical granularity, these data are particularly useful in the context of Covid-19 restrictions.

Weill et al. (2021) points out that parts of the existing literature investigating the impact of NPIs on mobility are not robust to minor changes in research design. One issue is that simple transformations of the outcome variable can deliver different results. More relevant to this paper, and in line with a large emerging literature, two-way fixed effects models might fail to identify an average treatment effect for treated units in the presence of heterogeneity (see among others, De Chaisemartin and d'Haultfoeuille, 2020). We discuss this in more detail below, and compute the empirical weights that arise in the estimation of two-way fixed effects models recently proposed by de Chaisemartin and D'Haultfoeuille (2020a). In doing so, we demonstrate that our full-sample results are likely to be robust to the types of two-way fixed effects biases recently highlighted in the literature.

Our estimates point to a negative and significant impact of NPIs on economic activity. Specifically, our case study estimates indicate that an Italian province moving from the lower level (tier 2) to the intermediate level (tier 3) in the system of restrictions can expect a significant fall in mobility. Moreover, we observe a systematic decline in the magnitude of the impact of NPIs in the second subsample, suggesting that the impact of NPIs does indeed lessen over time. These findings provide evidence of short-term economic consequences of Covid-19 restrictions on economic activity, indicating the need for policy makers to implement policies to mitigate the effects on employment and income. Our findings also suggest that the effectiveness of restrictions in reducing mobility is likely to fade away over time, suggesting that rolling systems of restrictions are less likely to be effective than shorter, sharper restrictions. We test the robustness of our results performing a series of checks which we display in appendix D.

The remainder of the paper is structured as follows. Section 2 describes the epidemiological context and the Italian regional system of NPIs. Section 3 discusses our methodology and section 4 the data. Section 5 presents the results, and section 6 concludes the paper. The paper is complemented with an appendix that includes tables, an extensive literature review, additional analysis, and robustness checks.

2 Epidemiological context and the Italian regional restriction system

The first two cases of Sars-Cov2 in Italy were two Chinese tourists, who were hospitalized and tested positive on the 30th January 2020 (Vicentini and Galanti, 2021). On the same day, the Italian government interrupted flight connections with China, and on the following day, it declared a state of emergency. Between the 21st and the 22nd of February, significant outbreaks were registered in parts of Lombardy and Veneto, and on the 23rd of February, eleven municipalities were put into severe restrictions. By the end of the month, new cases topped 1000 per day, and new measures were put in place to contain the virus on the 4th of March, including closing schools. Ultimately, the government decided to impose a national lockdown on the 9th of March, and was the first to do so outside of China.

The epidemiological situation continued to worsen, and by the 21st of March the daily number of individuals in need of intensive care hospitalization was over 2800. The government imposed further restrictions as a result, and only very essential or strategic activities were allowed to stay open. On the 28th of March daily deaths attributed to Covid-19 peaked at 928 (ISTAT, 2020), and on the 3rd of April hospitalizations peaked with over 4000 people in need of intensive care (see figure 1). From the 14th of April, containment measures were gradually lifted, and with the Prime Ministerial Decree of April 26 the government enacted 'phase-two', which officially ended the lockdown through a plan of gradual reopening of activities (Barbieri and Bonini, 2021).

After a relatively quiet Summer, a new surge in cases occurred in the Autumn of 2020. Initially, between the 14th of October and 5th of November, the extension of restrictions by the Italian government took place at the national level, and included mandatory face masks in outdoor spaces, reduced capacity for recreational venues, and targeted reductions of business opening hours (Manica et al., 2021). From November 6th, however, the government established a three-tiered (and subsequently four tiered) set of restrictions at the regional level. The tier system was conventionally known as the color-based system, as tiers

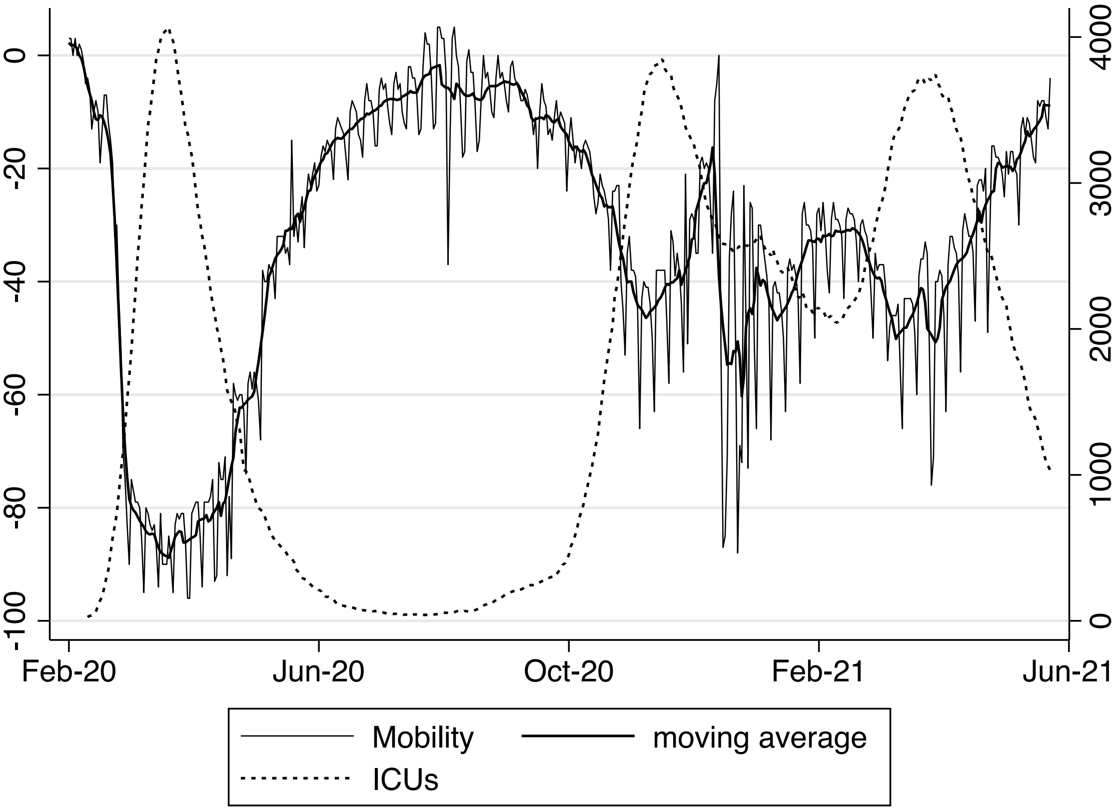

Figure 1: Google mobility index (retail and recreation) and its seven day moving average (left axis), and number of individuals in intensive care units (right axis), between February 2020 and June 2021.

corresponded to different colors: yellow for low risk, orange for medium risk and red for high risk. The tiered system involved, among other rules, limitations to retail and service activities, individual movement restrictions and reinforced distance learning in schools, with stringency increasing in higher tiers (see e.g., Manica et al., 2021). Figure 2 illustrates the progression of restrictions in each region from November 2020 up to June 2021.

Although this system was based on local restrictions, the decisions regarding its implementation were taken by the central government. These decisions were based on a number of criteria regarding epidemiological risk, which included recent trends in cases, the number of symptomatic individuals, the number of available hospital beds, and the R number (the reproduction rate). With the decree of the Prime Minister (Decreto del Presidente del Consiglio, DPCM) of the 14th of January, the government introduced an additional parameter consisting of a threshold of 50 cases for every 100,000 habitants.

Despite the fact that decision-making was centralized, there were also potential sources of heterogeneity in stringency levels between regions and across time. One source stems from the fact that local authorities enjoyed some degree of discretion, limited to increasing the stringency of government measures. Therefore, while central government could force a region into tier 3 from tier 2, for example, the region could itself decide to impose stricter measures on top of this. This was generally done at province or sub-province level, and there have been a number of episodes in which regions assigned to yellow or orange zones have applied

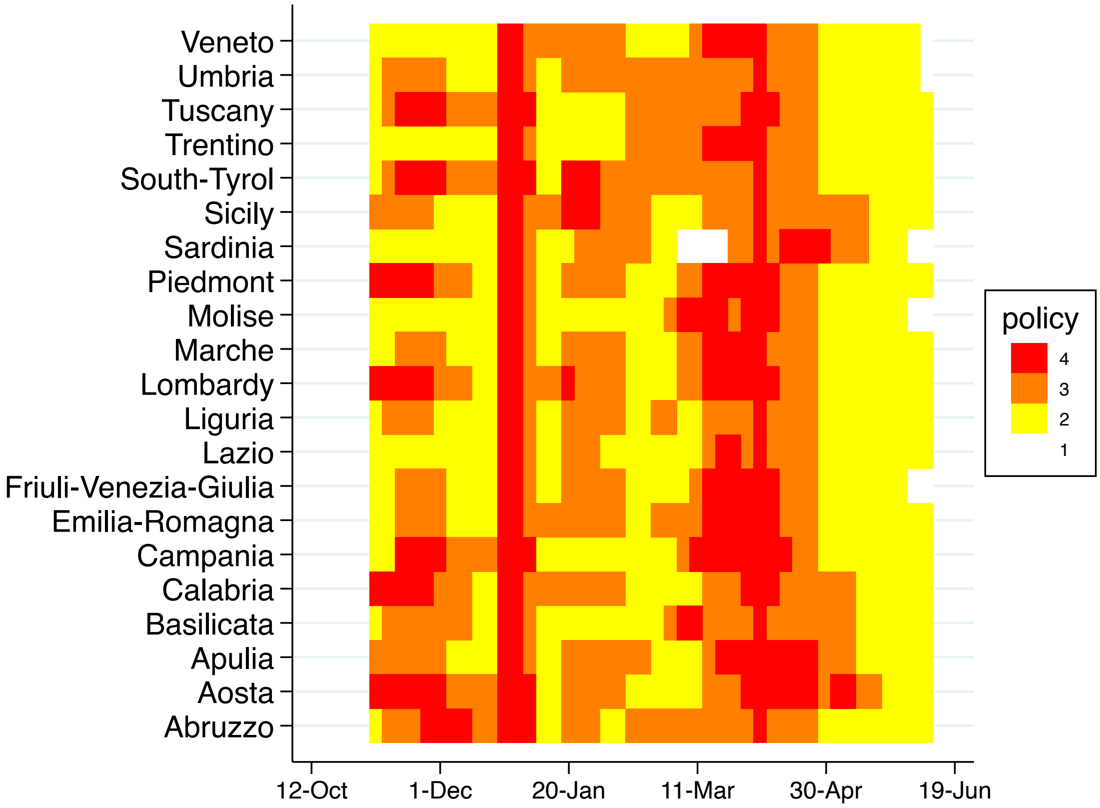

Figure 2: Progression of restrictions in each of 19 Italian regions and the autonomous provinces of Trento and Bolzano between October 2020 and June 2021. White denotes level 1, yellow denotes level 2, orange denotes level 3, and red denotes level 4.

local red zones in some areas. Regions could also increase the stringency of specific rules, for instance, remote schooling. In fact, local authorities had the possibility to impose schools closing also in zones where the general rules allowed for in-presence attendance.

Finally, an additional element of interest regards the time-varying nature of the rules. Overall, the tier system was fairly stable across time. Nevertheless, slight changes that might have had a potential impact on mobility have occurred. For instance, the DPCM of the 14 of January 2021 restricted inter-regional mobility among yellow regions, as well as within-regional and within-province mobility. These potential sources of heterogeneity between regions in apparently equal tiers should be borne in mind in what follows. The key takeaway, however, is that NPIs varied systematically between Italian regions after November 2020, and it is the resulting heterogeneity that we leverage to estimate the effects of Covid-19 restrictions on economic activity.

3 Empirical approach

3.1 Full sample estimation

As discussed in the previous section, we exploit geographic variation in the Italian tier system to identify the effects of NPIs on economic activity proxied by mobility.

The first model we employ is a one-way fixed effects (FE) model, controlling for time-invariant province-specific unobservables and a rich set of epidemiological observables:

$$Mobility_{irt} = \alpha_i + \rho D_{rt}^{orange} + \tau D_{rt}^{red} + \beta' X_{irt} + \epsilon_{irt}, \tag{1}$$

where $Mobility_{irt}$ denotes a mobility index (to be defined in more detail below) for province i in region r at time t and α_i represents cross-section fixed effects. D_{rt}^{orange} is a policy dummy equal to one when region r is subject to tier three of restrictions, and zero otherwise; D_{rt}^{red} is a policy dummy equal to one when region r is subject to tier four of restrictions, and zero otherwise. X_{irt} is a set of control variables measuring the severity of the Covid-19 pandemic at province, regional and national level. It also includes Saturday and Sunday dummies to catch seasonality in mobility data.

The second model that we employ is a two-way fixed effects (TWFE) model with variable treatment timing and intensity, to control for cross-sectionally invariant unobserved time patterns:

$$Mobility_{irt} = \alpha_i + \alpha_t + \rho D_{rt}^{orange} + \tau D_{rt}^{red} + \beta' X_{irt} + \epsilon_{irt}, \tag{2}$$

where α_t represent time fixed effects, and X_{irt} only includes Covid-19 controls at regional and province level, but not at national level. For both models 1 and 2 we split the sample around December 2020 to examine time variation in treatment effects.

3.2 Case studies

If restriction tiers are conditionally exogenous, given the time-invariant province-level characteristics and observables control variables, then the FE model should capture the causal effects of NPIs on mobility. If cross-sectionally invariant time-varying unobservable are important, then the conventional belief was, until recently, that a TWFE model such as model (2) would deliver the ATT under a parallel trend assumption.

However, in a setup with more than two groups and two periods, the TWFE model is now known to deviate from the canonical 2×2 difference-in-differences (DiD) model. In fact, as shown in Goodman-Bacon (2021), the TWFE model yields a variance-weighted average of all possible 2×2 DiD in the data. In case of heterogeneous treatment effects across groups and across time, which can be assumed to apply to most cases, the TWFE model yields a biased estimator of the ATT (De Chaisemartin and d'Haultfoeuille (2020), Goodman-Bacon (2021), Callaway and Sant'Anna (2020)). In addition, model (1) and model (2) include multiple treatments. In such cases, de Chaisemartin and D'Haultfoeuille (2020a) shows that the TWFE estimator is also affected by a contamination effect arising from the other treatments.

In combination, these elements raise concerns on the reliability of running a TWFE model on our data. To minimize these concerns, as suggested by Callaway and Sant'Anna (2020), one solution would be to break the sample into all possible episodes for which we can retrieve canonical 2×2 DiD, with a control group remaining in a single tier while a treatment group changes from one tier to the next. We broadly follow this approach with a small number of extra restrictions. Specifically, we select canonical DiD episodes based on the following criteria:

1. A treatment group has to move from a low lockdown tier to a high lockdown tier;
2. A control group has to remain in the same tier;
3. The control group has to be geographically contiguous to the treatment group;
4. The episode has to be repeated in both the November–December 2020 and January–March 2021 subsamples.

The first requirement and second requirements simply define a 2×2 DiD episode that identifies, in principle, the treatment effect of entering a lockdown tier. The third requirement is essentially a 'choice of control donor pool' requirement – geographically contiguous regions are likely to be more similar to each other than more distant regions, and therefore more likely to be suitable controls than more distant regions (see e.g., Abadie, 2021). Finally, the fourth requirement allows us to measure the extent to which the effect of NPIs on economic activity changes over time. To do this effectively, we need to control for unobserved heterogeneity in treatment effects across units, and thus we need to compare the same treatment and control groups at two different points in time.

There are three pairs of regions that satisfy these requirements, thus, we conduct the analysis on these three case studies. The first case involves the regions Emilia-Romagna and Veneto, where the former goes from yellow to orange on the 15th of November 2020, and again on the 21st of February 2021, while the latter stays yellow. The second case involves Toscana and Lazio, with the former going from yellow to orange on the 11th of November 2020, and again on the 14th of February 2021, while the latter stays yellow. The third case involves Lazio as control group and Abruzzo as treated, as the latter passes from yellow to orange on November 11th 2020, and again on February 14th 2021. As is evident from these case studies, a limitation of this approach is that there are no episodes satisfying the conditions above in which a treatment group moves from orange to red. However, by relaxing condition 4, we are able to select two episodes of passage from yellow to red. Given the particularities of these cases we discuss them directly in the results section.

For each of these case studies we estimate the following two specifications:

$$Mobility_{irt} = \alpha_i + \alpha_t + \rho D_{rt} + \beta' X_{irt} + \epsilon_{irt}, \tag{3}$$

$$Mobility_{irt} = \alpha_i + \alpha_t + \rho Stringency_{irt} + \beta' X_{irt} + \epsilon_{irt}. \tag{4}$$

Model (3) essentially replicates model (2), a TWFE, in each case study. Model (4) employs as regressor $Stringency_{irt}$, a province level stringency index developed by Banca d'Italia economist Conteduca (2021), which applies a similar methodology to the Oxford Stringency Index of Hale et al. (2020). This index takes into account potential heterogeneity between provinces and across time that might arise as local authorities impose further restrictions at province or municipality level, as well as heterogeneity that emerges as characteristics of the tier system change over time.

Finally, we also estimate the following generalized (or event-study) difference-in-differences to measure the post-treatment dynamic evolution of the treatment:

$$Mobility_{irt} = \alpha_i + \alpha_t + \sum_{j=1}^{T} \rho_j d_{rt} + \beta' X_{it} + \epsilon_{irt}, \tag{5}$$

in which there is now a sequence of dummy variables d_{rt} that take the value one for a single day in the sample $t = 1, ..., T$, and zero otherwise. Note that the model is unidentified without at least one restriction on these coefficients; we return to this point below. Also, by construction we cannot include region-specific variables in model (5).

To validate our empirical strategy, we must be able to motivate the assumption of parallel trends necessary for the standard difference-in-differences model to identify the average treatment effect on treated units. This assumption states that the outcome in the treatment and control groups would have had the same evolution, on average, in the absence of treatment. Hence, any deviation from the trend occurring at the time in which the policy is implemented can be attributed to the policy (the treatment). In a simple 2×2 DiD with pre- and post-treatment periods, evidence in support of the parallel trend assumption can be obtained by visually inspecting the trends prior to treatment.

Employing DiD methods requires a careful selection of standard error estimators. If heteroskedasticity is rarely an issue for inference (Angrist and Pischke, 2008), our data is likely to feature within-cluster correlation as well as broader forms of cross-sectional dependence, as commonly occurs in spatial panel data. Given that the policy is mostly implemented at regional level (and thus regressors are correlated within regions), if the model fails in fitting within-region correlations then regular estimators will underestimate SEs, even if cluster fixed effects are included (Cameron and Miller, 2015). The ideal solution would be to cluster at the level where the policy is implemented, but this is not feasible in our case studies because we would only have two clusters: the control region and the treated region (Garmann, 2017). Thus, the only option would be to cluster at province-level, which implicitly assumes that errors are uncorrelated between provinces. However, even clustering at province level only implies a number of clusters per case study between nine and sixteen given our sample, which is still thought to be less than optimal (Cameron and Miller, 2015).

In addition, it is also possible that broader forms of cross-sectional dependence might exist in our sample. Specifically, we might expect the existence of spatial correlation between provinces that belong to different regions, particularly as we have imposed geographical proximity as a group selection criteria. Given the foregoing, we proceed by employing the most conservative SEs in each model. In our FE and TWFE models these are Driscoll–Kraay standard errors, which are robust to spatial correlation, autocorrelation and heteroskedasticity (Hoechle, 2007). While in model (5) the largest SEs are provided by clustering. Table D1 in the appendix displays the standard errors based on model (2) obtained under different assumptions. The significance of the estimator is not impacted by the choice of the standard errors.

Lastly, as noticed by Buono and Conteduca (2020), Boone and Ladreit (2021) and Weill et al. (2021) among others, linear models attempting to estimate the impact of the epidemic or containment policies on mobility might suffer from endogeneity. In fact, policies are based on observed epidemiological outcomes, which can be expected to affect mobility. However, the latter might also affect epidemiological outcomes, and thus policy adoption, assuming that mobility is correlated with social distancing. Moreover, mobility might also be affected by observed epidemiological outcomes, as these can impact individuals' understanding of, and attitude to, infection risk. Our identification strategy addresses some of these problems. First, our case studies are analyzed over a short time frame, reducing the likelihood that mobility affects epidemiological outcomes (which tends to take place with a lag, see e.g., ISS, 2020). In addition, Maruotti et al. (2021) reports that the reproduction rate, one of the

Figure 3: Geographical distribution of treated and control groups in the case studies.

crucial parameter for policy adoption, "may be biased and in significant delay with respect to the current evolution of the epidemic process". Secondly, we include lags of epidemiological outcomes to reduce the likelihood of simultaneity bias between epidemiological development and policy adoption. Third, the presence of common trends in the pre-treatment period suggests that previous outcomes are conditionally independent from policy adoption.

4 Data and descriptive statistics

To estimate the impact of NPIs on mobility, we collect data on the 107 Italian provinces[1] from the 28th of October to the 23rd of December 2020, and from the 7th of January until the 15th of March 2021. We have excluded the period spanning December 24th to January 6th, as during this period the regional system was suspended, and national measures were applied for the Christmas holidays. When running our FE and TWFE models we exclude provinces belonging to Sardegna, as it is the only region passing into tier 1 in March 2021. We also exclude the provinces of Aosta, Bolzano and Trento as these represent autonomous

[1]Italy is composed by 20 regions, divided into 107 provinces. One average each region is composed by 5 provinces, ranging between the one-province region of Valle D'Aosta and the twelve provinces region of Lombardia. In the European administrative system, Italian regions are NUTS2 units while provinces are NUTS3. In comparison with the US, a proper term of comparison the Italian provinces are the American counties.

administrative entities, and therefore the group level where the policy is implemented coincides with the observational unit level. In addition, the provinces belonging to Piemonte are also excluded because of persistent reporting errors in Covid-19 province cases. However, including Piemonte does not affect our results, as reported in Appendix D.

As explained in the previous section, for our DiD specification we consider 6 specific case studies in which a treatment group goes from level 2 to level 3 (i.e., from yellow to orange) and a control group remains in level 2. In the remainder of this section, we focus on descriptive statistics for these case studies, as these constitute the weakest identification condition in our battery of empirical approaches.

In principle the sample selection of each individual case study should include a sufficient number of pre-treatment observations for which the common trend assumption holds, and should stop at the time when either the control or treatment group changes policy regime. However, to balance pre-treatment and post-treatment periods, we have restricted the latter to at most three weeks after the policy implementation.

Table 1 provides details on the six specific samples including starting and end dates, the date in which the policy is implemented and when the policy is announced, as well as the date in which the next policy regime change occurs.

Table 1: Sample start and end-points in the six case studies

	start of sample	announcement	implementation	end of sample	policy regime change
Veneto - E.R.	Nov 1	Nov 13	Nov 15	Dec 5	Dec 6
	Feb 1	Feb 19	Feb 21	Mar 7	Mar 8
Lazio - Toscana	Oct 28	Nov 10	Nov 11	Nov 14	Nov 15
	Feb 1	Feb 12	Feb 14	Mar 1	Mar 15
Lazio - Abruzzo	Oct 28	Nov 10	Nov 11	Nov 17	Nov 18
	Feb 1	Feb 12	Feb 14	Mar 1	Mar 15

Notes: E.R. stands for Emilia Romagna.

4.1 NPI data

Information on restrictions is taken from the Italian ministry of health. In regards to model (4), we employ a province-level stringency index developed within Banca d'Italia by Conteduca (2021). This index follows the criteria of the Oxford Stringency Index developed by Hale et al. (2020). The latter is a cross-country and cross-time measure of government pandemic-related interventions, based on a number of standardized metrics such as school closing, stay home requirements and travel restrictions from which the composite index is generated (Hale et al., 2020). Conteduca (2021) also takes into account potential asymmetries between regions and across time, as it is built at municipality-level and then aggregated up to province-level.

Figure 4 shows the evolution of the average stringency index by province around the date of the policy implementation in each case study, and reports the minimum and maximum values. In all cases but one, the index displays a near-identical pattern for treated and control groups before implementation. There is also little apparent variation in restrictions among provinces, with two noticeable exceptions in panel B and F. In fact, in the latter two cases there is significant variation in stringency between provinces after the implementation, which is due to region-specific restrictions being imposed at province-level. For example, when the Italian government imposed the orange tier restrictions for the region of Abruzzo in February, the regional authorities imposed local red zones in the province of Chieti and Pescara from February 14th (Regione Abruzzo, 2021).

4.2 Mobility

We obtain mobility data at province-level from the Google Mobility Report, which aggregates anonymized information regarding trends in visits to categorized places from Google users that have opted in using their 'location history' setting[2]. The data are expressed as percent deviations from baseline values that represent norms for each day of the week. Specifically, the baselines are the median values over the five-week period spanning 3rd January to 6th February 2020, i.e., just prior to the onset of the pandemic in Europe. These data are classified into five categories: retail and recreational places, groceries and pharmacies, parks, transit stations, and places of residence. We use visits to retail and recreational places as our benchmark measure, since it is likely to best capture the effect of restrictions targeting non-essential activities, which was an extremely common feature of most countries' NPIs and of the Italian regional system.

To understand these data, suppose that the retail and recreational places mobility index is equal to minus 10 on a given day in the province of Rome. This means that visits, meaning the number of trips made by Android devices with an activated history location setting, to recreational places (including restaurants, bars and retailers) in Rome are 10 percent below their pre-pandemic median value for that specific day of the week.

Mobility data have two key advantages for our purposes, namely, that they are available at very high frequency and at a granular level, which makes them particularly suitable for an investigation of the impact of NPIs. In fact, mobility data have been widely used in the economic literature on Covid-19 (WEO, 2020; Deb et al., 2020; Boone and Ladreit, 2021; Buono and Conteduca, 2020). Importantly, the OECD has found strong correlations between mobility, GDP and private consumption in a large sample of countries, with the index able to explain around 75% of cross-country variation in consumption (OECD, 2020). Mobility data has also proved to perform well in out-of-sample forecasts of growth compared to other benchmarks (Gamtkitsulashvili and Plekhanov, 2021), as well as a good predictor of industrial production (Sampi Bravo and Jooste, 2020). The implication is that mobility is correlated with economic activity, e.g., visits to retailers are correlated with consumption expenditure. In this regard, Buono and Conteduca (2020) employs Google visits to retail and recreational places as a proxy for consumption. On the other hand, it is important to note that reductions in mobility will overstate the economic impact if consumers maintain the same level of spending in a smaller number of trips (Goolsbee and Syverson, 2021), and mobility data cannot reveal potential substitution effects occurring as consumers switch from in-person to on-line shopping.

[2]For a detailed explanation of the data aggregation process, see Aktay et al. (2020)

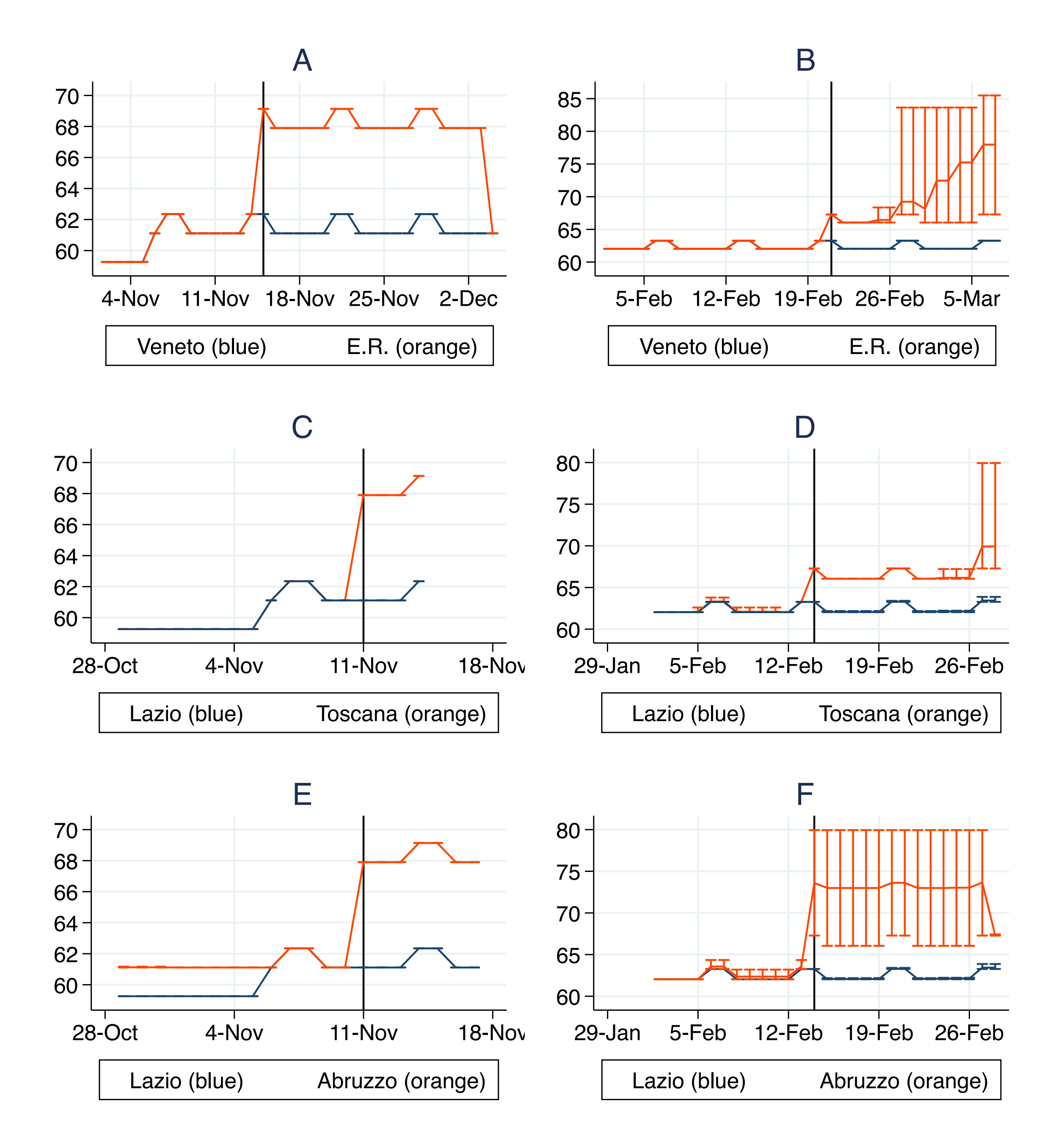

Figure 4: Government Stringency Index. The blue and orange lines represent respectively the mean value of provinces belonging to the control group and treatment group. Bands are given by the minimum and maximum value. The black vertical line indicates the treatment date.

Figure 5 shows the evolution of the benchmark mobility index in the six case studies around the date of the implementation. This figure helps us to visually investigate the existence of common trends in the pre-treatment period. In all cases, the evolution of mobility in the control and treatment group coincides very closely in the pre-treatment periods, and then decouples when the policies are implemented. In some case, for instance in panel F, the decoupling of the two lines appears to begin at the time when the policy is announced. As announcements might play a role in the overall impact of NPIs on activity, it is investigated

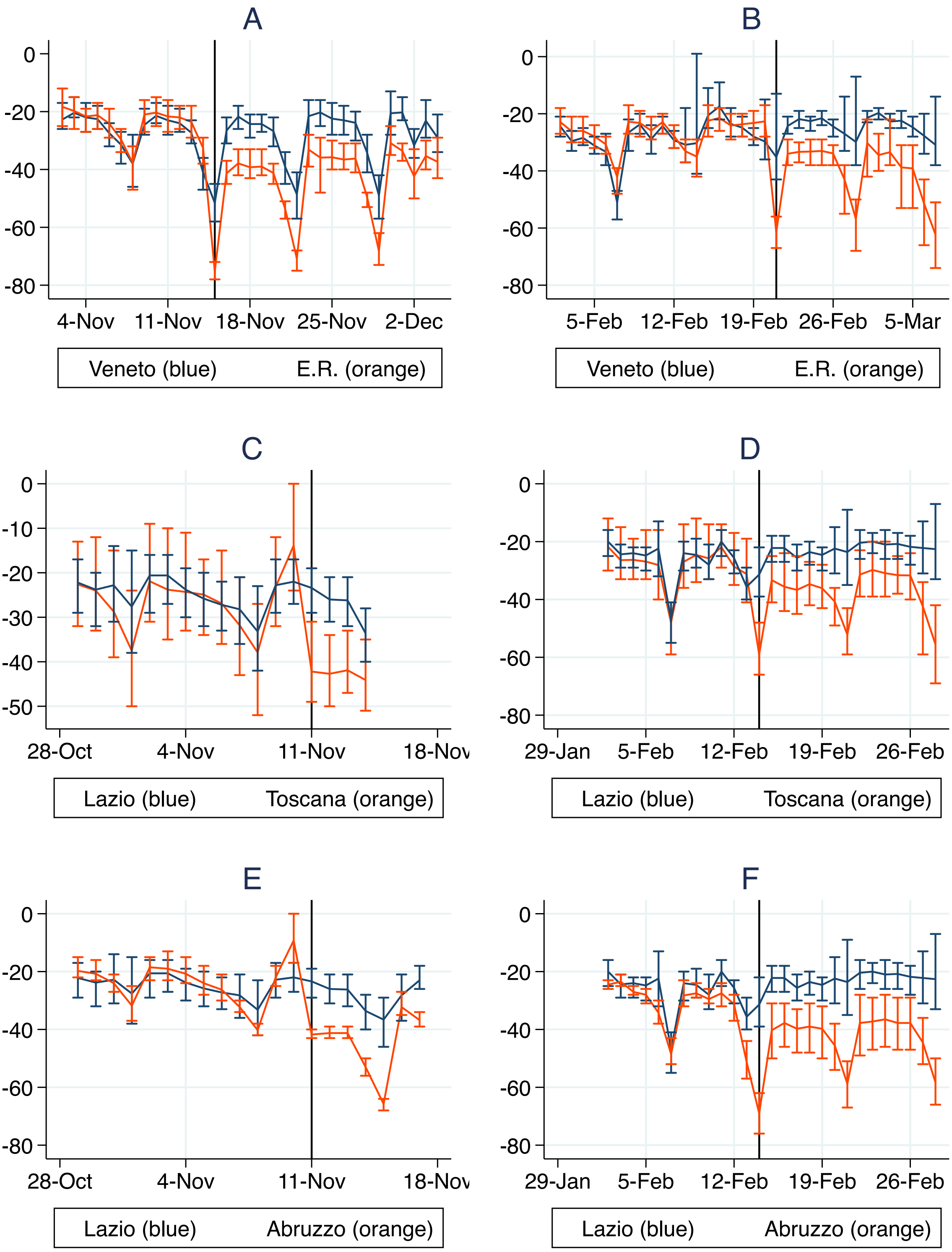

Figure 5: Google Mobility Index. The blue and orange lines represent respectively the mean value of provinces belonging to the control group and treatment group. Bands are given by the minimum and maximum value. The black vertical line indicates the treatment date.

in our analysis. Allowing for this, however, figure 5 is strongly suggestive that common trends existed in the pre-treatment periods in all of our case studies.

4.3 Epidemiological data

Finally, as NPIs are implemented based on regional epidemiological data, we control for the extent of the Covid-19 pandemic, including a set of variables at the regional level that includes Covid-19 related deaths, cases and ICUs. To account for the impact of voluntary social distancing or fear of infections, we also employ the number of daily new Covid-19 cases per 100,000 people at province level. The latter is obtained as the first difference of cumulative cases, which is the only series available at province level (Morettini et al., 2020). Unfortunately, at the time of writing, there are no time series available regarding hospitalizations and fatalities for most of the Italian provinces. In addition, province data for certain units and certain dates display reporting errors, as the first difference in cumulative cases delivers negative numbers. In some cases such errors are corrected by extracting the relevant information from local authority sources. In other cases where is not possible to retrieve the information or where data revision occurs, the negative values are substituted with zeros. Finally, systematic reporting errors appear in the series regarding the provinces of Piemonte. Specifically, as these data differ from local authority sources for long periods, we have excluded the region from our full sample specification, with no consequences on our results (see figure D1 in appendix D). Aside from these corrected observations, all the Covid-19 data are taken from Morettini et al. (2020). When we run the regression on the full sample using model (1) we also include a set of Covid-19 controls at national level, employing cases, fatalities and ICUs. Including national level epidemiological data helps to catch potential correlation between national epidemiological development and local mobility.

Our case studies baseline specifications include a seven day moving average of daily province cases per 100,000 people, and a one day lag of the logarithm of regional ICUs. The latter is selected among a richer set of transformed series of regional data included in model (1), as it appears to be the most significant region-level predictor. However, we explore alternative specifications in appendix D and show that our results are not affected by the choice of Covid-19 regressor. Figure 6 shows the progression of mean new cases per 100,000 people at province level within our estimation samples in each case study. Figure 6 indicates no obvious connection between surge in infections and policy adoption.

5 Results

5.1 Full sample estimation

We first report the results based on model (1), a FE model, and (2), a TWFE model, in the entire sample of regions. As explained above, we split the estimation for the periods November-December 2020, and January-March 2021, to test whether the effect has changed over time. Figure 7 displays the point estimates and confidence intervals of the policy dummies. The entire regression output can be found in table A1 in appendix A. The point estimates represent the effect of NPIs on our Google Mobility benchmark measure, that is, deviations from the pre-pandemic median in visits to retail and recreational places. We can immediately notice that the point estimates are all negative and significant. Moreover, the FE and TWFE models deliver very similar coefficients, particularly when estimating the impact of tier 3 on mobility. The estimates indicate that passing into tier 3 triggers a decline

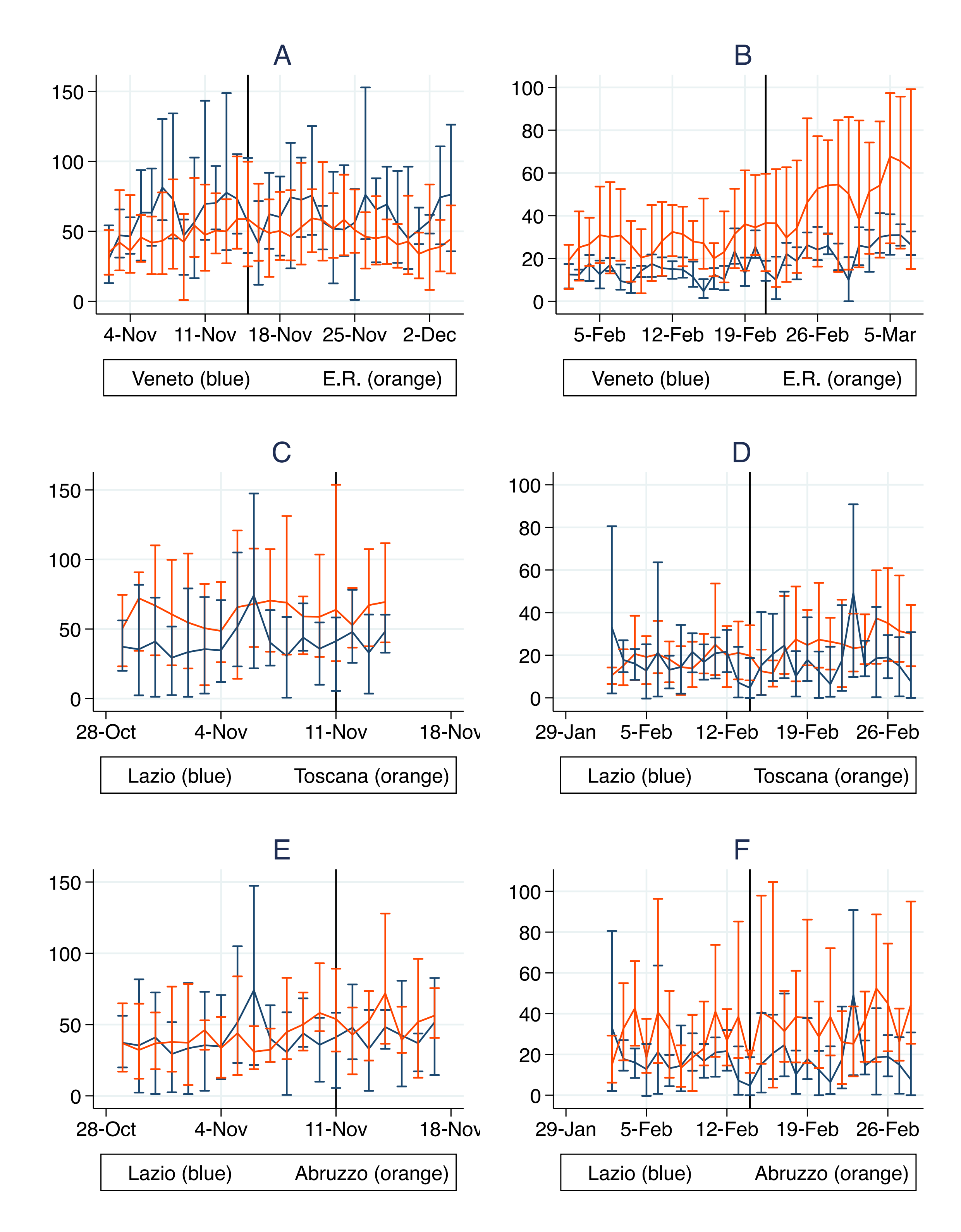

Figure 6: Province Covid-19 cases per 100,000 people. The blue and orange lines represent respectively the mean value of provinces belonging to the control group and treatment group. Bands are given by the minimum and maximum value. The black vertical line indicates the treatment date.

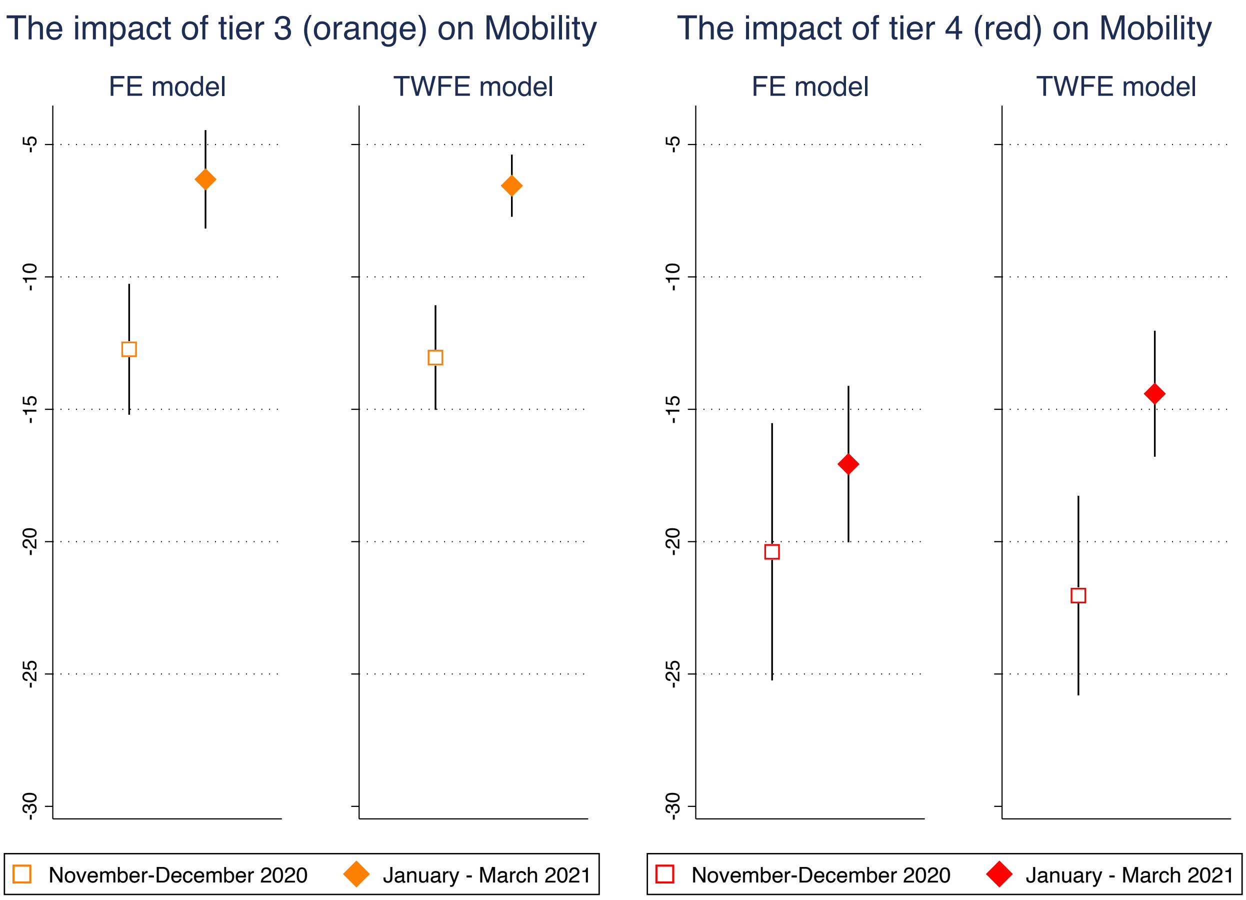

Figure 7: The impact of NPIs on province Google mobility Index estimated with FE and TWFE models. The figure is based on the estimation reported on table A1 in the appendix.

in the expected mobility index of about 13 percentage points in the first subsample, and 7 percentage points in the January-March period. Instead, the impact of tier 4 on expected mobility is between -20 p.p. and -22 p.p. in the first period, and between -14 p.p. and -17 p.p. in the second subsample. Therefore, these figures suggest that NPIs negatively impact mobility, even after controlling for the local epidemiological situation. Also, based on these results the impact of restrictions appear to have declined in the second period of analysis, and the difference between the coefficients in the two subsamples is statistically significant.

Table A1 reports the coefficient of the moving average of new cases per 100,000 people, our proxy for individual voluntary social distancing or fear of infection. The estimates are all negative and precisely estimated, suggesting that for every new case per 100,000 people, the expected mobility index decreases between -0.08 to -0.2 percentage points. In addition, the magnitude increases in the second period. This suggests that epidemic-related individual motives are also an important predictor of mobility, particularly in the second subsample, although we have not attempted to identify the causal effect associated with increases in the severity of the pandemic.

We check whether our estimates could be biased because of heterogeneous treatment effects by assessing to what extent negative weights impact the TWFE estimator (Jakiela, 2021). As shown in De Chaisemartin and d'Haultfoeuille (2020), the TWFE estimator is a

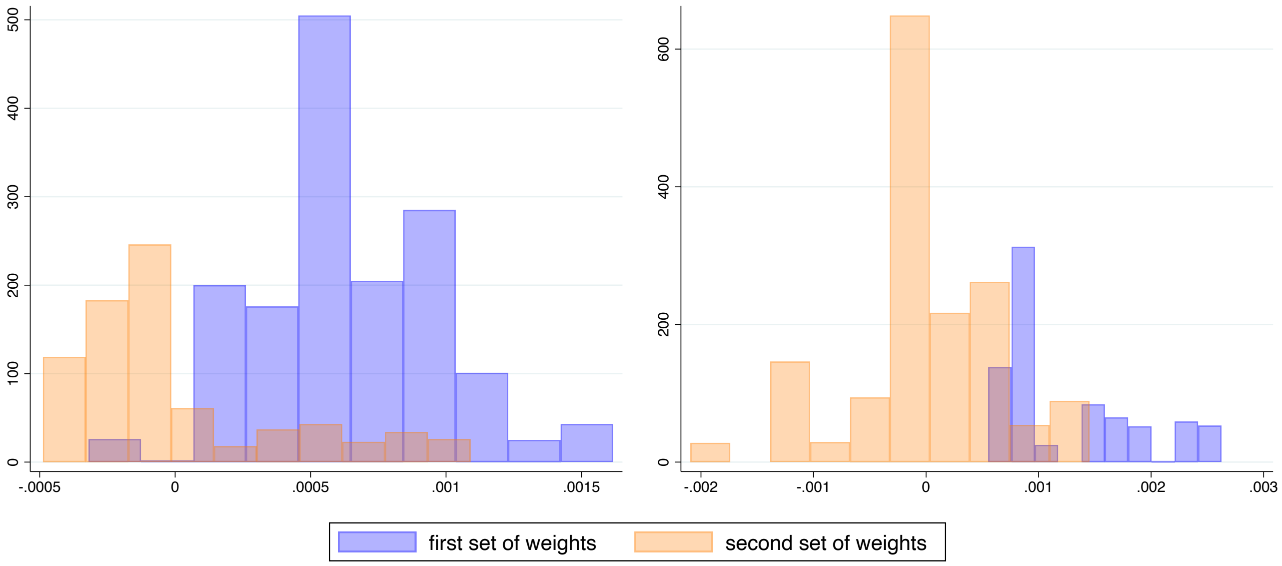

Figure 8: The left hand panel plots the histogram of the two set of weights associated with the estimation of D_{rt}^{orange} in model (2), while the right hand panel those associated with D_{rt}^{red}, in the first subsample. The first and second set of weights corresponds respectively to those in the first and second term in the expression of the estimator in theorem 1 of de Chaisemartin and D'Haultfoeuille (2020a).

weighted sum of several average treatment effects, with some negative weights. In case of heterogeneous treatment effects, the presence of negative weights raises the possibility that the TWFE estimator may not satisfy the "no-sign reversal property" (De Chaisemartin and D'Haultfoeuille, 2022). In a specification with two treatments, the estimator might also suffer from a contamination due to the second treatment (de Chaisemartin and D'Haultfoeuille, 2020a). We thus compute both the weights associated with the first and the second term of the estimator, for both dummies of model (2), as in theorem 1 of de Chaisemartin and D'Haultfoeuille (2020a). The weights are proportional to the residuals of the regression of the treatment on time and group fixed effects, the other treatment and controls, as in equation 1 in de Chaisemartin and D'Haultfoeuille (2020a).

Figure 8 shows the histograms of the two sets of weights for the two coefficients. If we look at the first set of weights, we can observe that there are very few treatment effects associated with negative weights in the estimation of the parameter of D_{rt}^{orange}, and in the that of D_{rt}^{red}. On the other hand, there is a significant number of negative weights associated with the second term, suggesting the presence of contamination effects of the other treatment. However, these weights are small in magnitude compared to the likely effect sizes. Moreover, as the treatment effect is expected to be negative, the contamination weights imply that the TWFE estimates are biased towards zero. These estimates are therefore conservative, and thus we expect our results to be robust to the types of TWFE biases recently highlighted in the literature.

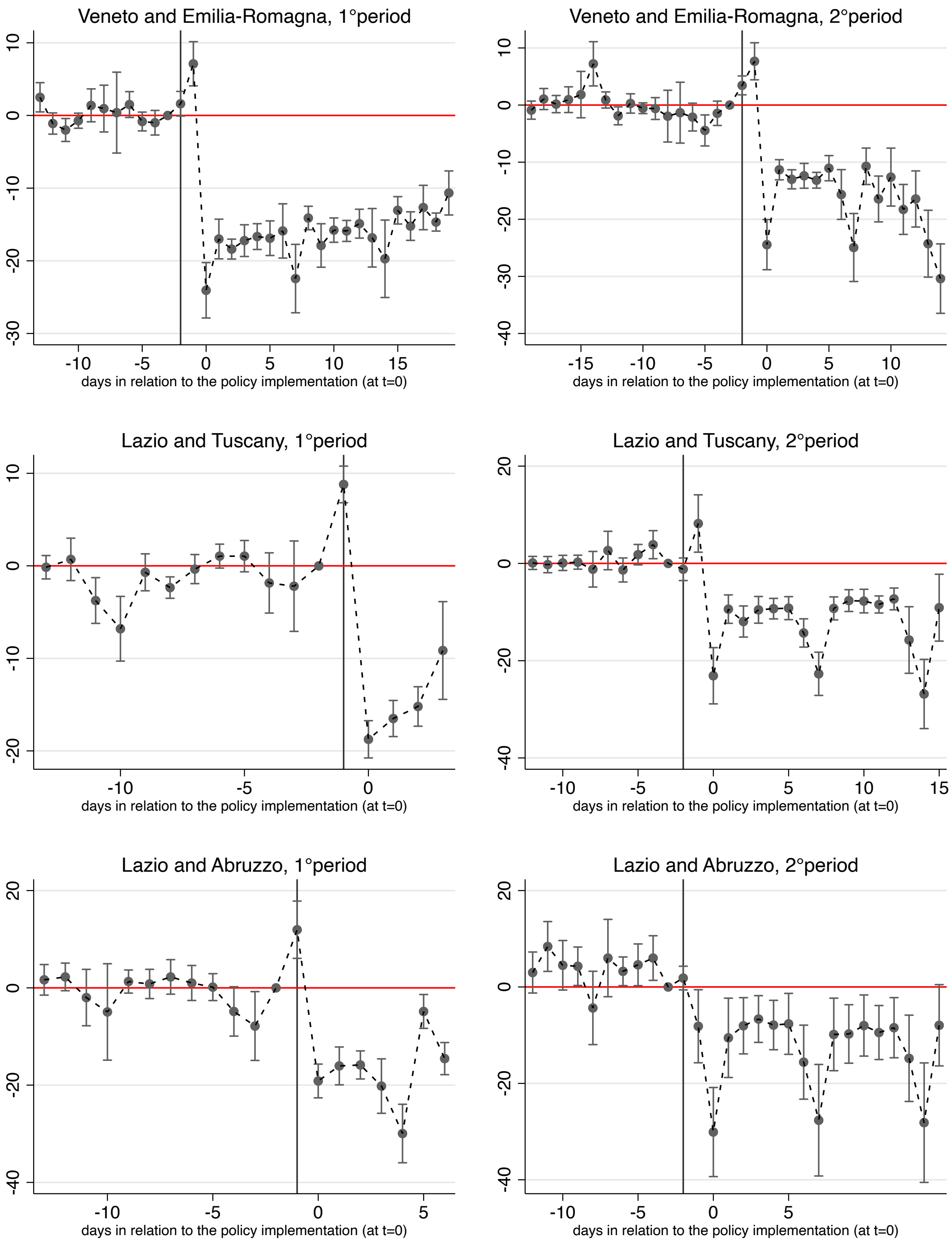

Figure 9: Lead and lag effects of NPIs on mobility. The horizontal axis represent the days in relation to the policy implementation ($t = 0$). The vertical black line represents the day in which the policy is announced.

5.2 Case studies

5.2.1 Specification 1: binary treatment

We turn now to our case studies analysis, and report the dynamic DiD results obtained by estimating model (5) on each case study. All point estimates are expressed as percentage

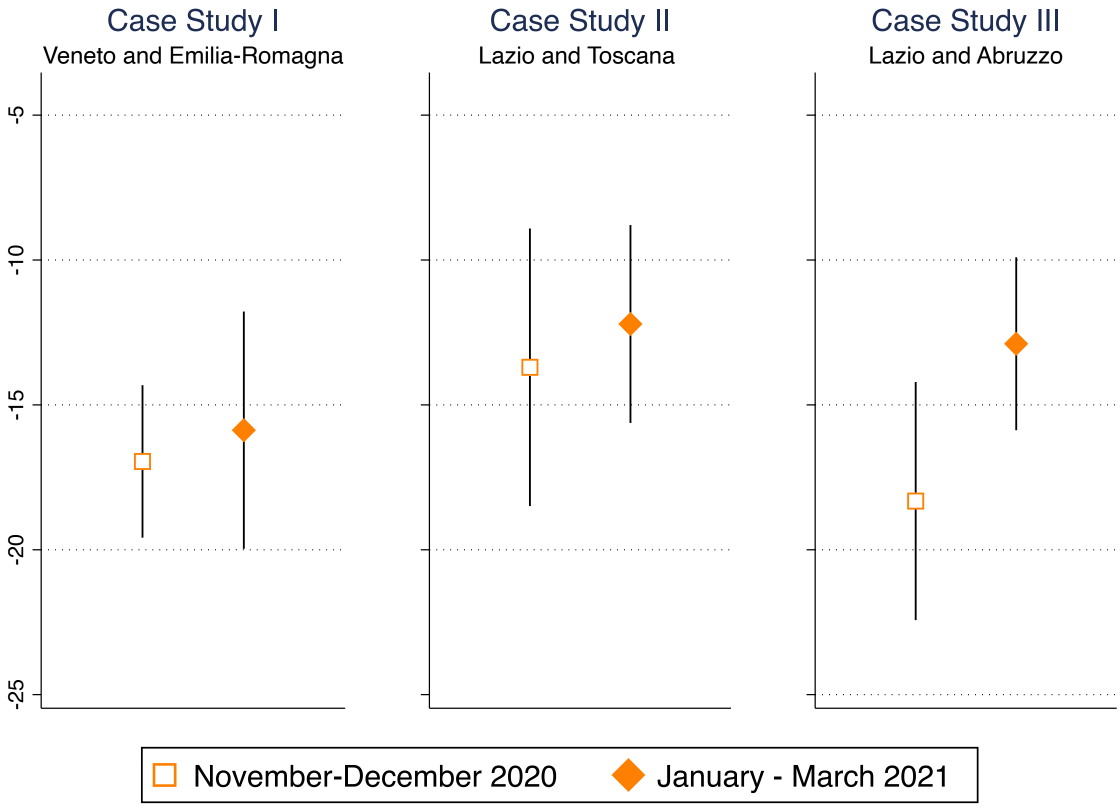

Figure 10: Estimated impact of NPIs on mobility. The figure is based on the estimation reported on table A2 in appendix A

point differences of the mobility index with respect to one day before the announcement of new restrictions. The dynamic plots allow us to verify the presence of common trends in the pre-treatment period.

Figure 9 confirms what was suggested in figure 5, as all plots show close-to-zero coefficients prior to the implementation date, followed by significantly negative coefficients after the implementation date.

The policy effects appear to be heterogeneous across regions, and are not constant within regions over the treatment period. Nevertheless, the estimated effects of NPIs on mobility are clearly large and negative, and appear to be larger in the first period than the second. Among other things, these results indicate that treatment effects are heterogeneous between groups.

Another key takeaway from figure 9 is the presence of positive announcement effects, characterized by unusually higher mobility in the day before the implementation. This result does not come as a particular surprise, as individuals anticipating upcoming restrictions adjust their mobility behavior, for instance, by concentrating consumption before business closures. The only exception appears to be the case of Lazio and Abruzzo in the second period (see Goodman-Bacon and Marcus, 2020, for an analysis of the impact of announcement effects from a methodological perspective).

Figure 10 summarises the dynamic DiD results in figure 9 using the static DiD specification

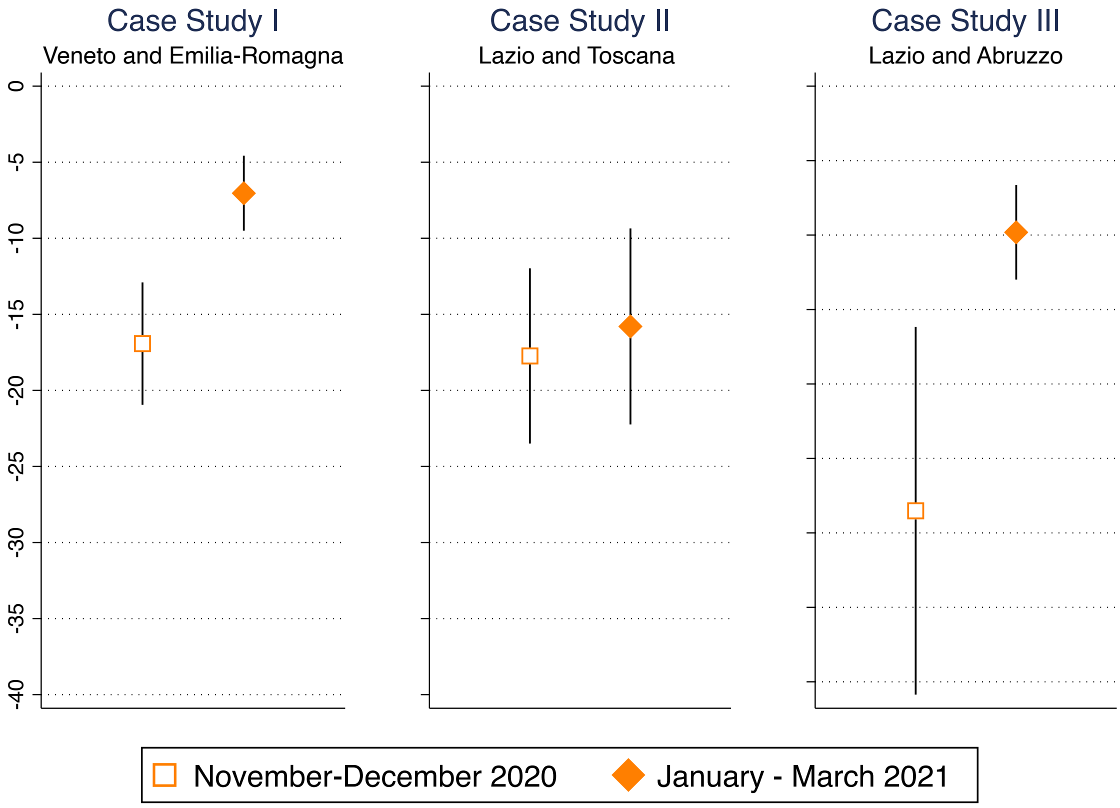

Figure 11: Estimated impact of NPIs on mobility. The figure is based on the estimation reported in table A3 in the appendix A.

from model (3). The full regression results are presented in table A2 in appendix A. As before, the point estimates represent the effect of NPIs on our Google Mobility benchmark measure, that is, deviations from the pre-pandemic median in visits to retail and recreational places. As we can notice from figure 10, the estimated coefficients are negative and significant, with their magnitude somewhat heterogeneous between cases and ranging from -18 p.p. to -12 p.p.. Nevertheless, it is possible to observe a decreasing effect in the second sub-sample in each of the three case studies, which suggests that the impact of NPIs on province mobility has decreased over time, confirming our findings obtained on the full sample of regions. However, the difference between the two periods is not statistically significant. Finally, we can compare figure 10 with figure 7. It emerges that our case study estimates are broadly in line with the coefficients obtained in the FE and TWFE models in the first period. However, the case study estimates are significantly higher than the full sample coefficients in the second subsample. These higher estimates could be explained by the trimmed post-treatment samples in two case studies, or conservatism in the first set of estimates.

Table A2 in appendix A reports the full regression output of running model (3) in each case study. In panel A the model is regressed assuming the treatment starts on the implementation date. The estimated impact of NPIs on mobility in the first and second subsample is about -17 p.p. and -16 p.p. in the first case study, about -14 p.p. and -12 p.p. in the second case, while it is -18 p.p. and about -12 p.p. for the third case study.

In panel B the model assumes that treatment starts at the announcement. As a matter of fact, the positive announcement effect observed in figure 9 are expected to partially offset the negative impact occurring after the policy is implemented. For instance, if announcements induce consumers to stockpile purchases before restrictions are implemented, it impacts overall economic performance. By estimating the parameters relative to the announcement our findings change in two ways. First, we obtain lower estimates for five out of six coefficients. Hence, accounting for the announcement slightly mitigates the policy impact on activity. Secondly, we do not observe a decreasing magnitude in the second subsample in two of the case studies. Thus, by including the announcement effects in the regression we find evidence of a decreasing impact of NPIs over time only in the case of Veneto and Emilia-Romagna.

Table A2 reports the estimates of our province level Covid-19 control, that is, the moving avarage of new cases per 100,000 people. Contrary to our full sample results, the proxy for individual fear or voluntary social distancing is significant only in three case studies, but the estimates are in line with our previous findings.

5.2.2 Specification 2: Continuous treatment

We now turn to our continuous variable specification, i.e., model (4). To compare the estimates of the latter with those obtained with the binary variable specification, we have computed in each case study the marginal difference between the sample mean stringency index for treated and untreated units. The results are summarized in figure 11 and are broadly in line with those estimated with the binary specification. In fact, the estimates point to a significant decline in mobility in each of the six cases, as well as suggesting a decreasing impact over time. In terms of magnitude the coefficients are mostly comparable with those of figure 10. Major differences in the magnitude of the estimates are the second period in the first case study and the first period in the second case study. As outlined above, by including the stringency index we take into account potential heterogeneity in regards to NPIs severity between regions and over time that are necessarily omitted in the binary specification.

Table A3 in appendix A contains the full output of the regressions employing the stringency index. The estimated coefficients of NPIs in the first case study indicate that a one percent increase in stringency is associated with a drop in mobility of respectively -2.3 and -1 percentage points in the first and second subsamples. In the second case study the point estimates are -1.9 and -1.6, while in the third case are -2.5 and -0.9 percentage points.

5.2.3 Yellow to red case studies

Based on the four conditions that we have imposed in section 3, we are able to select episodes in which a region goes from yellow to orange while a control region remains yellow. If we relax condition 4 we can identify two additional case studies in which a treatment region goes from tier 2 into tier 4, thus from yellow to red. The first episode involves the case of Veneto and Lombardia, with the latter passing into red as soon as the color-based system was imposed on the 6th of November 2020. The rules applied in the yellow tier were fairly consistent with those existing prior to the regional system. Thus, as Veneto was assigned to the yellow tier, we do not expect a significant impact on regional mobility. Similarly, the second episode involves the regions of Campania and Calabria, as the latter was assigned into red tier on the 6th of November 2020 while the former into yellow tier. In this case we

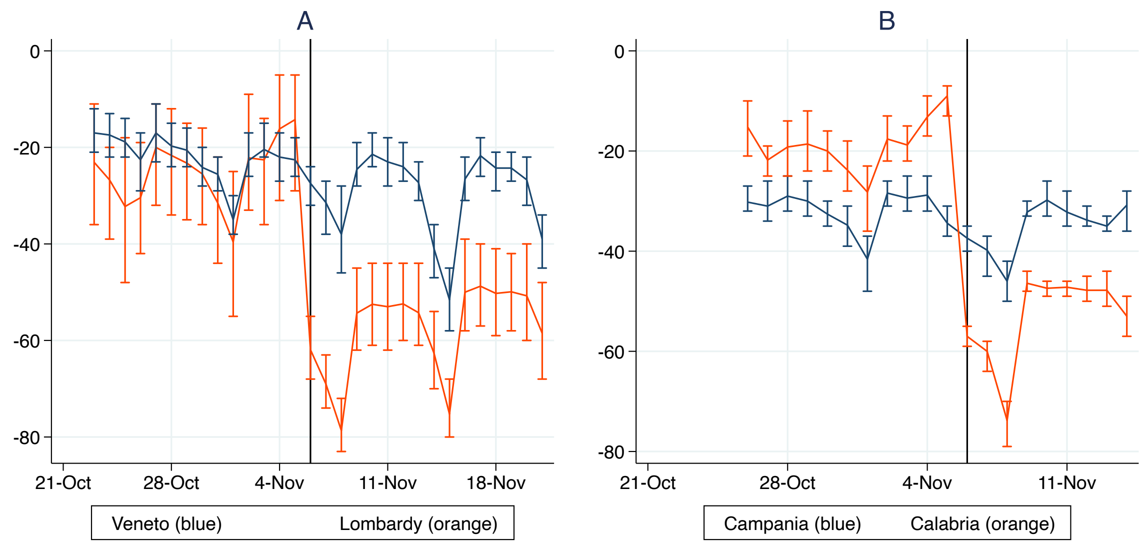

Figure 12: Mobility data in yellow to red case studies

also relax condition 3, as the two regions do not directly share a border. However, as they both belong to the macro area of southern Italy, they ought to be broadly comparable.

As we can appreciate from figure 12, parallel trends seem to hold in the pre-treatment period, while a significant decline is observed after the implementation date in the treated group. Moreover, no significant effect is observed in the control group, suggesting that in fact the rules in the yellow tier were fairly consistent with those before the regional system was set. In figure 13 we plot the estimated coefficients based on model (3). The impact of NPIs is negative and significant, with a magnitude of about 25 and 29 percentage points, which is somewhat higher than the estimates obtained in the full sample estimation. But, as discussed in section 5.1, we expect the latter to be conservative, so this is not altogether surprising.

5.3 Other robustness checks

To assess the robustness of our full sample estimation we run a number of checks, which are shown in appendices C and D. In appendix C we show that estimates obtained with alternative Google mobility indices, as well as with the Apple indices, are consistent with our baseline results. In appendix D, we show that our results are robust to changes in the composition of the sample, as well as to changes in the set of Covid-19 controls. We run the procedure proposed by Hadi (1992) to find outliers in multivariate models and we find zero observations deemed as outliers. Lastly, we run model (2) under different assumptions for the standard errors and show that our coefficients are always precisely estimated.

6 Conclusion

In this paper, we document that NPIs implemented by the Italian government are associated with a significant decline in province mobility, after controlling for region-specific epidemiological data. Among other things, we document announcement effects prior the implementation of the policy which partially offset the negative decline in economic activ-

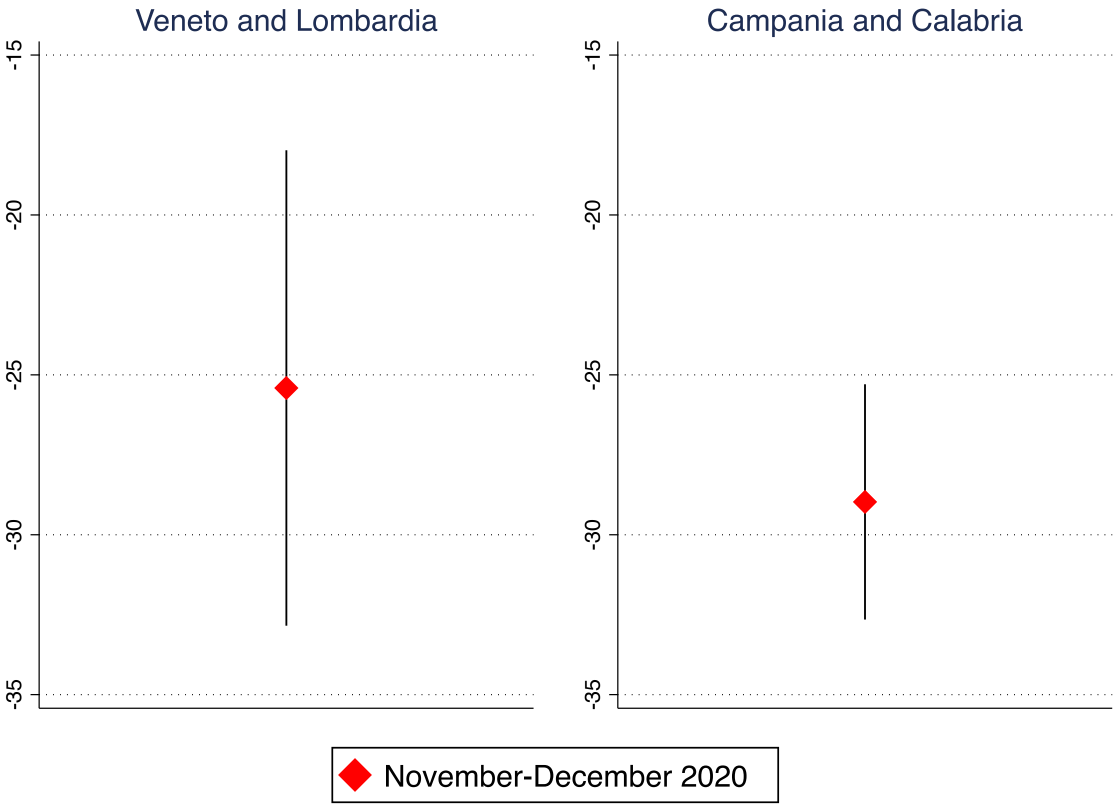

Figure 13: Estimated impact of NPIs (tier 4) on mobility, based on model (3).

ity, but also have obvious health implications. In fact, the spikes in mobility around the announcements are likely to be correlated with declining social distancing, in direct contrast to the policy objective.

Our estimates suggest a stronger impact on mobility than those obtained by Caselli et al. (2020), who, using a fixed effects model, estimate a drop in mobility of 7% in Italian municipalities that experienced the very first wave of restrictions in March 2020. By employing a stringency index designed based on the Italian system of restrictions, our estimates suggest that a 10 percentage point increase in stringency is associated with a decrease of mobility that ranges between 9 percentage points and 23 percentage points. We can also compare these estimates with those obtained by Boone and Ladreit (2021). In their cross-country regression on the second wave of Covid-19, they estimate that in Italy a 10 percentage point increase in their stringency index is correlated with around a 4 percentage point decline in mobility.

We also investigate whether the impact of the Italian restrictions changed over time. The existing literature has underlined that the efficacy of restrictions is likely to be correlated with the duration of the measures, as fatigue motives increase over time (Goldstein et al., 2021). By estimating our models in two subsequent subsamples, we find consistent evidence of a declining impact of NPIs on mobility. Our confidence in this finding is slightly reduced by models that take into account announcement effects, in which the declining impact only appears in one case study. Nevertheless, some of these results are affected by the need to balanc pre-treatment and post-treatment sample spans, which reduce the latter in some cases and yield higher estimates in the second subsample.

To summarise, our findings provide evidence of a short-term trade-off between lives and livelihoods. The existence of this trade-off suggests that policymakers need to balance economic and public health costs when implementing NPIs. Moreover, our findings indicate that the efficacy of NPIs is likely to decline over time, suggesting that rolling systems of restrictions are less likely to be effective than shorter, sharper lockdowns. Given this, our data and empirical approach do not provide insights into the existence of a long-term trade-off, and more research is needed in this direction to expand our knowledge of the economic effects of NPIs.

References

A. Abadie. Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature*, 59(2):391–425, 2021.

A. Aktay, S. Bavadekar, G. Cossoul, J. Davis, D. Desfontaines, A. Fabrikant, E. Gabrilovich, K. Gadepalli, B. Gipson, M. Guevara, et al. Google covid-19 community mobility reports: anonymization process description (version 1.1). *arXiv preprint arXiv:2004.04145*, 2020.

J. D. Angrist and J.-S. Pischke. *Mostly harmless econometrics*. Princeton university press, 2008.

Apple. Covid-19 mobility trends reports. https://covid19.apple.com/mobility, 2022.

P. N. Barbieri and B. Bonini. Political orientation and adherence to social distancing during the covid-19 pandemic in italy. *Economia Politica*, pages 1–22, 2021.

A. Bauer and E. Weber. Covid-19: how much unemployment was caused by the shutdown in germany? *Applied Economics Letters*, pages 1–6, 2020.

M. Benzeval, J. Burton, T. F. Crossley, P. Fisher, A. Jäckle, H. Low, and B. Read. The idiosyncratic impact of an aggregate shock: the distributional consequences of covid-19. *Available at SSRN 3615691*, 2020.

L. Boone and C. Ladreit. Fear of covid and non-pharmaceutical interventions: An analysis of their economic impact among 29 advanced oecd countries. *Centre for Economic Policy Research, March*, 2021.

I. Buono and P. Conteduca. Mobility before government restrictions in the wake of covid-19. 2020.

B. Callaway and P. H. Sant'Anna. Difference-in-differences with multiple time periods. *Journal of Econometrics*, 2020.

A. C. Cameron and D. L. Miller. A practitioner's guide to cluster-robust inference. *Journal of human resources*, 50(2):317–372, 2015.

M. Caselli, A. Fracasso, and S. Scicchitano. From the lockdown to the new normal: An analysis of the limitations to individual mobility in italy following the covid-19 crisis. Technical report, GLO Discussion Paper, 2020.

R. Chetty, J. Friedman, N. Hendren, M. Stepner, et al. How did covid-19 and stabilization policies affect spending and employment? a new real-time economic tracker based on private sector data. *NBER working paper*, (w27431), 2020.

O. Coibion, Y. Gorodnichenko, and M. Weber. The Cost of the COVID-19 Crisis: Lockdowns, Macroeconomic Expectations, and Consumer Spending. Technical report, 2020.

T. G. Conley. Gmm estimation with cross sectional dependence. *Journal of econometrics*, 92(1):1–45, 1999.

F. P. Conteduca. Measuring covid-19 restrictions in italy during the second wave. Technical report, BANCA D'ITALIA, 2021.

C. Cot, G. Cacciapaglia, and F. Sannino. Mining google and apple mobility data: temporal anatomy for covid-19 social distancing. *Scientific reports*, 11(1):1–8, 2021.

C. De Chaisemartin and X. d'Haultfoeuille. Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–96, 2020.

C. de Chaisemartin and X. D'Haultfoeuille. Two-way fixed effects regressions with several treatments. *arXiv preprint arXiv:2012.10077*, 2020a.

C. De Chaisemartin and X. D'Haultfoeuille. Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey. Technical report, National Bureau of Economic Research, 2022.

P. Deb, D. Furceri, J. D. Ostry, and N. Tawk. The Economic Effects of COVID-19 Containment Measures. CEPR Discussion Papers 15087, C.E.P.R. Discussion Papers, July 2020. URL https://ideas.repec.org/p/cpr/ceprdp/15087.html.

M. Famiglietti and F. Leibovici. The impact of health and economic policies on the spread of covid-19 and economic activity. *European economic review*, 144:104087, 2022.

Q. Feng, G. L. Wu, M. Yuan, and S. Zhou. Save lives or save livelihoods? a cross-country analysis of covid-19 pandemic and economic growth. *Journal of economic behavior & organization*, 197:221–256, 2022.

T. Gamtkitsulashvili and A. Plekhanov. Mobility and economic activity around the world during the covid-19 crisis. *Applied Economics Letters*, pages 1–7, 2021.

S. Garmann. The effect of a reduction in the opening hours of polling stations on turnout. *Public Choice*, 171(1-2):99–117, 2017.

J. Gathergood and B. Guttman-Kenney. The english patient: Evaluating local lockdowns using real-time covid-19 & consumption data. *arXiv preprint arXiv:2010.04129*, 2020.

J. Gibson and X. Sun. Understanding the economic impact of covid-19 stay-at-home orders: A synthetic control analysis. *Available at SSRN 3601108*, 2020.

P. Goldstein, E. L. Yeyati, and L. Sartorio. Lockdown fatigue: The diminishing effects of quarantines on the spread of covid-19. 2021.

A. Goodman-Bacon. Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 2021.

A. Goodman-Bacon and J. Marcus. Using difference-in-differences to identify causal effects of covid-19 policies. 2020.

A. Goolsbee and C. Syverson. Fear, lockdown, and diversion: Comparing drivers of pandemic economic decline 2020. *Journal of Public Economics*, 193(C), 2021. doi: 10.1016/j.jpubeco.2020.10. URL https://ideas.repec.org/a/eee/pubeco/v193y2021ics0047272720301754.html.

P.-O. Gourinchas. Flattening the pandemic and recession curves. *Mitigating the COVID economic crisis: Act fast and do whatever*, 31(2):57–62, 2020.

E. Guglielminetti and C. Rondinelli. Consumption and saving patterns in italy during covid-19. *Bank of Italy Occasional Paper*, (620), 2021.

A. S. Hadi. Identifying multiple outliers in multivariate data. *Journal of the Royal Statistical Society: Series B (Methodological)*, 54(3):761–771, 1992.

T. Hale, A. Petherick, T. Phillips, and S. Webster. Variation in government responses to covid-19. *Blavatnik school of government working paper*, 31:2020–11, 2020.

D. Hoechle. Robust standard errors for panel regressions with cross-sectional dependence. *The stata journal*, 7(3):281–312, 2007.

E. Islamaj, D. T. Le, and A. Mattoo. Lives versus livelihoods during the covid-19 pandemic. 2021.

ISS. Characteristics of sars-cov-2 patients dying in italy report based on available data on july 22nd, 2020. Technical report, ISTITUTO SUPERIORE DI SANITA', 2020.

ISTAT. Impatto dell'epidemia covid-19 sulla mortalitÀ totale della popolazione residente anno 2020. Technical report, Il quinto Rapporto prodotto congiuntamente dall'Istituto nazionale di statistica (Istat) e dall'Istituto Superiore di Sanità (Iss), 2020.

P. Jakiela. Simple diagnostics for two-way fixed effects. *arXiv preprint arXiv:2103.13229*, 2021.

J. L. C. Kok. Short-term trade-off between stringency and economic growth. *CEPR Covid Economics*, (60), 2020.

E. Kong and D. Prinz. Disentangling policy effects using proxy data: Which shutdown policies affected unemployment during the covid-19 pandemic? *Journal of Public Economics*, 189:104257, 2020.

J. Kurita, Y. Sugishita, T. Sugawara, and Y. Ohkusa. Evaluating apple inc mobility trend data related to the covid-19 outbreak in japan: Statistical analysis. *JMIR public health and surveillance*, 7(2):e20335, 2021.

W. Maloney and T. Taskin. Determinants of social distancing and economic activity during covid-19. 2020.

M. Manica, G. Guzzetta, F. Riccardo, A. Valenti, P. Poletti, V. Marziano, F. Trentini, X. Andrianou, A. M. Urdiales, M. Del Manso, et al. Impact of tiered restrictions on human activities and the epidemiology of the second wave of covid-19 in italy. *medRxiv*, 2021.

A. Maruotti, M. Ciccozzi, and F. Divino. On the misuse of the reproduction number in the covid-19 surveillance system in italy. *Journal of Medical Virology*, 2021.

S. Mendolia, O. Stavrunova, and O. Yerokhin. Determinants of the community mobility during the covid-19 epidemic: the role of government regulations and information. *Journal of Economic Behavior & Organization*, 184:199–231, 2021.

L. F. Morales, L. Bonilla-Mejía, J. Pulido, L. A. Flórez, D. Hermida, K. L. Pulido-Mahecha, and F. Lasso-Valderrama. Effects of the covid-19 pandemic on the colombian labour market: Disentangling the effect of sector-specific mobility restrictions. *Canadian Journal of Economics/Revue canadienne d'économique*, 55:308–357, 2022.

M. Morettini, A. Sbrollini, I. Marcantoni, and L. Burattini. Covid-19 in italy: Dataset of the italian civil protection department. *Data in brief*, 30:105526, 2020.

OECD. *OECD Economic Outlook, Volume 2020 Issue 2*. 2020. doi: https://doi.org/https://doi.org/10.1787/39a88ab1-en. URL https://www.oecd-ilibrary.org/content/publication/39a88ab1-en.

OECD. *OECD Economic Outlook, Interim Report March 2021*. 2021. doi: https://doi.org/https://doi.org/10.1787/34bfd999-en. URL https://www.oecd-ilibrary.org/content/publication/34bfd999-en.

C. Regione Abruzzo. Covid-19: zona rossa a chieti e pescara. comunicazione del presidente marsilio, 2021. URL https://www.regione.abruzzo.it/content/covid-19-zona-rossa-chieti-e-pescara-comunicazione-del-presidente-marsilio.

J. R. E. Sampi Bravo and C. Jooste. Nowcasting economic activity in times of covid-19: An approximation from the google community mobility report. *World Bank Policy Research Working Paper*, (9247), 2020.

A. Sheridan, A. L. Andersen, E. T. Hansen, and N. Johannesen. Social distancing laws cause only small losses of economic activity during the COVID-19 pandemic in Scandinavia. *Proceedings of the National Academy of Sciences*, 117(34):20468–20473, August 2020.

statista. Leading mobile operating systems ranked by market share in italy from 2010 to 2021. https://www.statista.com/statistics/623153/leading-mobile-operating-systems-ranked-by-market-share-in-italy/, 2022.

G. Vicentini and M. T. Galanti. Italy, the sick man of europe: Policy response, experts and public opinion in the first phase of covid-19. *South European Society and Politics*, pages 1–27, 2021.

D. Walker and J. Hurley. Did the covid-19 local lockdowns reduce business activity? evidence from uk smes. 2021.

J. A. Weill, M. Stigler, O. Deschenes, and M. R. Springborn. Researchers' degrees-of-flexibility and the credibility of difference-in-differences estimates: Evidence from the pandemic policy evaluations. Technical report, National Bureau of Economic Research, 2021.

WEO. *Chapter 2 The at Lockdown: Dissecting the Economic Effects*. INTERNATIONAL MONETARY FUND, USA, 2020. ISBN 9781513556055.

Appendix

A Output tables

Table A1: One way and two way fixed effect models

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				DEPENDENT VARIABLE: Mobility				
NPIs (dummy = 1 if Orange)	-17.60***	-12.59***	-12.73***	-13.05***	-9.348***	-9.163***	-6.314***	-6.554***
	(1.320)	(0.910)	(1.232)	(0.984)	(1.265)	(0.843)	(0.931)	(0.588)
NPIs (dummy = 1 if Red)	-30.93***	-22.36***	-20.38***	-22.04***	-22.93***	-18.43***	-17.07***	-14.41***
	(2.848)	(1.734)	(2.421)	(1.878)	(2.025)	(1.394)	(1.480)	(1.191)
m.a. cases per 100.000 people			-0.0847***	-0.0905***			-0.213***	-0.230***
			(0.0114)	(0.0109)			(0.0302)	(0.0329)
1 day lag of log of ICUs (Region)			-0.961	-0.339			-4.199***	-5.289***
			(2.516)	(1.595)			(1.268)	(0.797)
7 days lag of log of ICUs (Region)			-1.644				-4.140*	
			(2.682)				(2.455)	
14 days lag of log of ICUs (Region)			0.0605				1.939	
			(2.306)				(1.865)	
1 day lag of log fatalities (Region)			-0.295				-0.204	
			(0.503)				(0.333)	
7 days lag of log fatalities (Region)			0.199				0.0744	
			(0.377)				(0.312)	
14 days lag of log fatalities (Region)			0.116				-0.0742	
			(0.339)				(0.352)	
m.a. fatalities (National)			-0.0391***				-0.0362***	
			(0.0133)				(0.00987)	
m.a. delta ICUs (National)			-0.0380*				0.00476	
			(0.0214)				(0.0412)	
7days lag of growth of cases (National)			-1.417**				-0.539	
			(0.626)				(0.732)	
dummy = 1 if Saturday	-9.488***		-8.215***		-9.309***		-8.612***	
	(1.240)		(1.069)		(1.911)		(1.682)	
dummy = 1 if Sunday	-21.73***		-20.37***		-22.51***		-21.80***	
	(2.083)		(1.868)		(2.342)		(2.013)	
Constant	-18.44***	-34.33***	21.28	-29.38***	-25.48***	-23.47***	22.64***	7.724**
	(2.023)	(0)	(14.75)	(7.110)	(0.749)	(0.176)	(6.419)	(3.667)
R^2 (within) 0.703	0.897	0.780	0.903	0.646	0.796	0.749	0.835	
Observations	4,732	4,732	4,443	4,732	6,006	6,006	5,780	6,006
Number of groups	91	91	91	91	91	91	91	91
Province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	No	Yes	No	Yes	No	Yes	No	Yes
Period	Nov 1 - Dec 23	Nov 1 - Dec 23	Nov 1 - Dec 23	Nov 1 - Dec 23	Jan 1 - Mar 15	Jan 1 - Mar 15	Jan 1 - Mar 15	Jan 1 - Mar 15

Driscoll-Kraay Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A2: case studies binary treatment

	(1)	(2)	(3)	(4)	(5)	(6)
	DEPENDENT VARIABLE: Mobility					
	Veneto and Emilia-Romagna		Lazio and Toscana		Lazio and Abruzzo	
	PANEL A: Estimation at the implementation date					
NPIs (dummy =1 if Orange)	-16.95***	-15.87***	-13.74***	-12.21***	-18.37***	-12.89***
	(1.295)	(2.015)	(2.274)	(1.668)	(1.923)	(1.457)
m.a. cases per 100.000 people	-0.0446**	-0.197**	-0.0403	-0.178	-0.00182	-0.300***
- province level	(0.0197)	(0.0903)	(0.0310)	(0.135)	(0.0329)	(0.0745)
7 days lag of log fatalities	-0.735	-1.471	-1.648	-4.333**	1.174	-3.422***
- region level	(0.574)	(1.548)	(1.325)	(1.673)	(1.302)	(0.939)
Constant	-31.18***	-2.802	-15.82***	-19.35***	-20.19***	-13.47***
	(0.833)	(2.948)	(2.562)	(2.828)	(1.918)	(1.692)
R^2 (within)	0.953	0.817	0.880	0.813	0.896	0.789
Treatment date	Nov 15	Feb 21	Nov 11	Feb 14	Nov 11	Feb 14
	PANEL B: Estimation at the announcement date					
NPIs (dummy = 1 if Orange)	-14.64***	-12.24***	-8.521**	-9.773***	-13.07***	-12.84***
	(1.560)	(2.474)	(3.776)	(1.491)	(4.066)	(1.316)
m.a. cases per 100.000 people	-0.0838**	-0.291***	-0.0540	-0.257*	-0.0678	-0.322***
- province level	(0.0333)	(0.0975)	(0.0348)	(0.132)	(0.0590)	(0.0773)
7 days lag of log fatalities	0.0895	-1.938	-2.613	-4.865**	2.128	-3.568***
- region level	(0.950)	(1.743)	(2.189)	(1.967)	(2.321)	(1.076)
Constant	-29.39***	-4.003	-13.71***	-17.65***	-19.47***	-12.89***
	(1.428)	(2.997)	(4.105)	(3.142)	(2.265)	(1.686)
R^2 (within)	0.919	0.769	0.816	0.782	0.799	0.786
Announcement date	Nov 13	Feb 19	Nov 10	Feb 12	Nov 10	Feb 12
Observations	560	560	270	435	173	261
Number of groups	16	16	15	15	9	9
Period	Nov 1 - Dec 5	Feb 1 - Mar 8	Nov 1 - Nov 14	Feb 1 - Mar 1	Nov 1 - Nov 17	Feb 1 - Mar 1

Panel A displays the full regression output obtained by running model (3) in each case studies assuming the treatment begins on the date the policy is implemented. Panel B replicates Panel A assuming treatment starts at the announcement. Columns (1) and (2) report the finding for Veneto and Emilia-Romagna, columns (3) and (4) those of Lazio and Toscana, instead columns (5) and (6) report the findings of Lazio and Abruzzo. Driscoll-Kraay Standard errors in parentheses.
**** p<0.01, ** p<0.05, * p<0.1

Table A3: case studies continuous treatment

	(1)	(2)	(3)	(4)	(5)	(6)
	DEPENDENT VARIABLE: Mobility					
	Veneto and Emilia-Romagna		Lazio and Toscana		Lazio and Abruzzo	
NPIs (Stringency Index)	-2.364***	-0.983***	-2.023***	-1.797***	-2.863***	-0.980***
	(0.287)	(0.176)	(0.335)	(0.374)	(0.619)	(0.162)
m.a. cases per 100.000 people	0.0136	-0.122	-0.0403	-0.166*	-0.00372	-0.207*
- province level	(0.0411)	(0.0895)	(0.0310)	(0.0869)	(0.0455)	(0.106)
7 days lag of log fatalities	-0.802	-1.640	-1.648	-4.416**	0.925	-4.262***
- region level	(0.539)	(1.853)	(1.325)	(1.845)	(1.282)	(1.218)
Constant	106.9***	40.70***	104.1***	92.04***	152.1***	46.29***
	(17.85)	(9.586)	(20.07)	(21.73)	(35.79)	(9.061)
Observations	560	560	270	435	173	261
Number of groups	16	16	15	15	9	9
R^2 (within)	0.943	0.762	0.880	0.825	0.866	0.783
Period	Nov 1 - Dec 5	Feb 1 - Mar 8	Nov 1 - Nov 14	Feb 1 - Mar 1	Nov 1 - Nov 17	Feb 1 - Mar 1
Treatment date	Nov 15	Feb 21	Nov 11	Feb 14	Nov 11	Feb 14

Columns (1) and (2) report the finding for Veneto and Emilia-Romagna, columns (3) and (4) those of Lazio and Toscana, instead columns (5) and (6) report the the findings of Lazio and Abruzzo. Driscoll-Kraay Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

B Literature Review

Since the outbreak of Sars-Cov2, a growing literature has investigated the economic consequences of the Covid-19 epidemic, with many papers trying to estimate the impact of restrictions on economic activity. This literature contributes to the debate on the potential trade-off that policymakers face between saving lives and livelihood when adopting NPIs (Gourinchas, 2020; Islamaj et al., 2021). Evaluating the economic impact of containment measures raises several challenges for an empirical researcher. The main issue regards disentangling the effect of the disease itself from that of the policy. It is reasonable to assume that negative effects on economic activity manifest via two channels: individual behaviours (fear triggered by the epidemics or by information related to the epidemics, or voluntary social distancing) and the policy intervention (Sheridan et al., 2020). Many studies employ some form of regression technique in cross-country analysis where a mobility indicator is regressed on a proxy of NPIs, generally the Oxford stringency Index developed by Hale et al. (2020), (Goolsbee and Syverson, 2021; Deb et al., 2020; WEO, 2020; Mendolia et al., 2021; Boone and Ladreit, 2021; Buono and Conteduca, 2020; Goldstein et al., 2021). Most studies proxy the impact of the virus on individual behavior by employing data on Covid-19 cases and related fatalities (Goolsbee and Syverson, 2021; Deb et al., 2020; WEO, 2020; Mendolia et al., 2021; Boone and Ladreit, 2021; Buono and Conteduca, 2020; Islamaj et al., 2021; Kong and Prinz, 2020).

Part of the literature documents that most of the impact must be attributed to the epidemic itself. In summary, these papers conclude that, to a great extent, the economic fallout would have occurred independently of government interventions. For instance, Sheridan et al. (2020) study the first wave of the epidemic in Denmark and Sweden, as the former Scandinavian country adopted stricter measures than the latter. Their findings show that only a small fraction of the decline in activity can be attributed to government policies, but the fear of contagion and voluntarily social distancing contributed the most. In the same line, Maloney and Taskin (2020) find that movie spending in Sweden and restaurant reservations in the US fall independently of NPIs. Goolsbee and Syverson (2021) analyze whether shelter in place policies or fear of contagion impacted consumer confidence. They make use of mobile phone data to identify consumer mobility for 2.25 million US businesses at local level instead of state level, allowing them to identify the policy variable. The latter is regressed on consumer mobility as well as controlling for the number of deaths and fixed effects. They find that only 7% of the observed economic decline is attributed to shelter-in-place policies, while most of the economic downturn is determined by waning consumer confidence because of individual fear.

Similarly to the foregoing papers, Mendolia et al. (2021) separate information-led social-distancing effects from those induced by government policies. They find that Covid-19 cases (their proxy for information-led social-distancing) have had a long-lasting moderate effect on mobility that is independent of NPIs implemented by the government. This channel can explain about 15 percentage points of the overall reduction in mobility across the affected countries. Morales et al. (2022) exploit sectoral heterogeneity in policy adoption to assess the causal impact of NPIs on employment in Colombia. Based on their estimates they conclude that only one fourth of the job loss can be attributed to the policy, while the remaining is explained by epidemic-related factors. Kong and Prinz (2020) confirm these findings by focusing on labour market outcomes and arguing that most of the short run surge in UI claims in the US cannot be explained by NPIs. Chetty et al. (2020) also argue that health concern is the primary driver of reduced spending in the first stages of the pandemic in the US. Guglielminetti and Rondinelli (2021) explore the drivers of consumption and savings pattern in Italy in 2020. They found that a large part of the drop is due to unexplained macro factors. The authors further disentangle these factors by using survey data and find that fear and uncertainty contribute the most rather than government restrictions.

Other studies, which are closer to our paper, focus on locally differentiated system of restrictions. For example, Walker and Hurley (2021) use firm level data on SMEs, and compare the turnover between geographically contiguous businesses located in areas with and without restrictions in the UK. Their findings, obtained via a regression discontinuity design attributes only two fifth of decline in turnover to lockdown policies. Gathergood and Guttman-Kenney (2020) investigate the economic impact of restrictions in the context of local lockdowns in the UK. They employ transaction level data obtained from *Fable Data* disaggregated at local level and estimate difference-in-differences models between geographically contiguous units that feature heterogeneous restrictions. They find no evidence that local lockdown determined large decline in consumption. However, the authors do not interpret their estimates as causal, following the argument of Goodman-Bacon and Marcus (2020) that the counterfactual might not represent a proper control group. Finally, Feng et al. (2022) evaluate the trade-off between health and economic costs by constructing a pandemic containment effectiveness measure inspired by the stochastic cost frontier literature. They then regress such measure on quarterly GDP and find that higher GDP is correlated with a more effective epidemic containment. Based on that the authors conclude that policymakers do not face a trade-off between saving lives and livelihood.

Other papers reach different conclusions, attributing a more important role to governments measures. Coibion et al. (2020) use surveys to assess the macroeconomic expectations of households in the US. They attribute the drop in consumption, employment, inflationary expectations, mortgage payments and increased uncertainty primarily to lockdowns rather than the Covid-19 infections themselves. Benzeval et al. (2020) reach similar conclusions by means of UK survey data, and show that the decline in working hours was driven by the restrictions. Bauer and Weber (2020) finds that 60% of the surge in unemployment in Germany was caused by shutdown policies. Boone and Ladreit (2021) try to estimate the elasticity of fear and lockdown on economic activity, proxied by mobility, in a sample of 29 OECD countries with particular focus on six European countries, and look at differences between the first and second wave of infections. Their findings suggest that, overall, the greater effect on mobility was due to restrictions and not fear. When they look at subsamples, they find a stronger elasticity of restrictions during the first wave and a stronger elasticity of fear in the second wave. However, the authors acknowledge in the case of the UK and Spain that regional differences in the restrictions might have affect the validity of the national stringency index. Moreover, they also underline the problem of simultaneity, and thus of the difficulties of interpreting their results as causal. Kok (2020) investigate the short-term economic impact of NPIs by directly employing GDP growth in place of the common high frequency proxies. Their estimate points out that restrictions cost about 8 p.p of growth in advanced economies and 5 p.p in emerging economies. Caselli et al. (2020), who we complement, estimate the impact of locally differentiated restrictions on mobility during the first lockdown in Italy. They employ a fixed effects model, using mobility data from *Enel X* at municipality level. They identify the causal impact comparing municipalities belonging to the same local labour market area, but to different provinces under different restrictions. Their findings suggest that municipalities in lockdown experience a 7% decline in mobility, controlling for epidemiological data.

It is probably true to say, as WEO (2020) acknowledges, that lockdowns and voluntary social distancing are both important in explaining the Covid-19 recession. Gibson and Sun (2020), for example, employ a synthetic control method to identify the effect of stay-home policies on weekly jobless claims in the US. The control group is composed by the 7 states that never implemented stay-home policies up to the 25th April 2020. Their findings show higher spikes in unemployment claims in the treatment group, however the synthetic series also were affected. Their conclusion is that, although stay-home-policies have amplified the economic damage, this would have occurred anyway. Likewise, Buono and Conteduca (2020) acknowledge the importance of both voluntary measures and restrictions in explaining the drop in activity, and the challenge of estimating the effect of each channel individually, because they can occur simultaneously. , caselli2021protecting find that both social distancing and lockdown had significant and negative impacts on economic activity using Google mobility and Indeed job posting as outcome variables. They also conclude that short-lived, stricter lockdowns are preferable than prolonged, mild restrictions. Lastly, Famiglietti and Leibovici (2022) estimate the joint dynamics of the Covid-19 epidemic, associated containment policies, economic support policies and economic activity, by implementing a structural vector autoregression model and monthly export data to isolate the exogenous supply-side component of the shock. Among their results, they find that containment policies impact the economy with a transitory negative effect, but positively impact the spread of the epidemic by reducing fatalities.

C Additional Analysis

C.1 Other Google mobility data

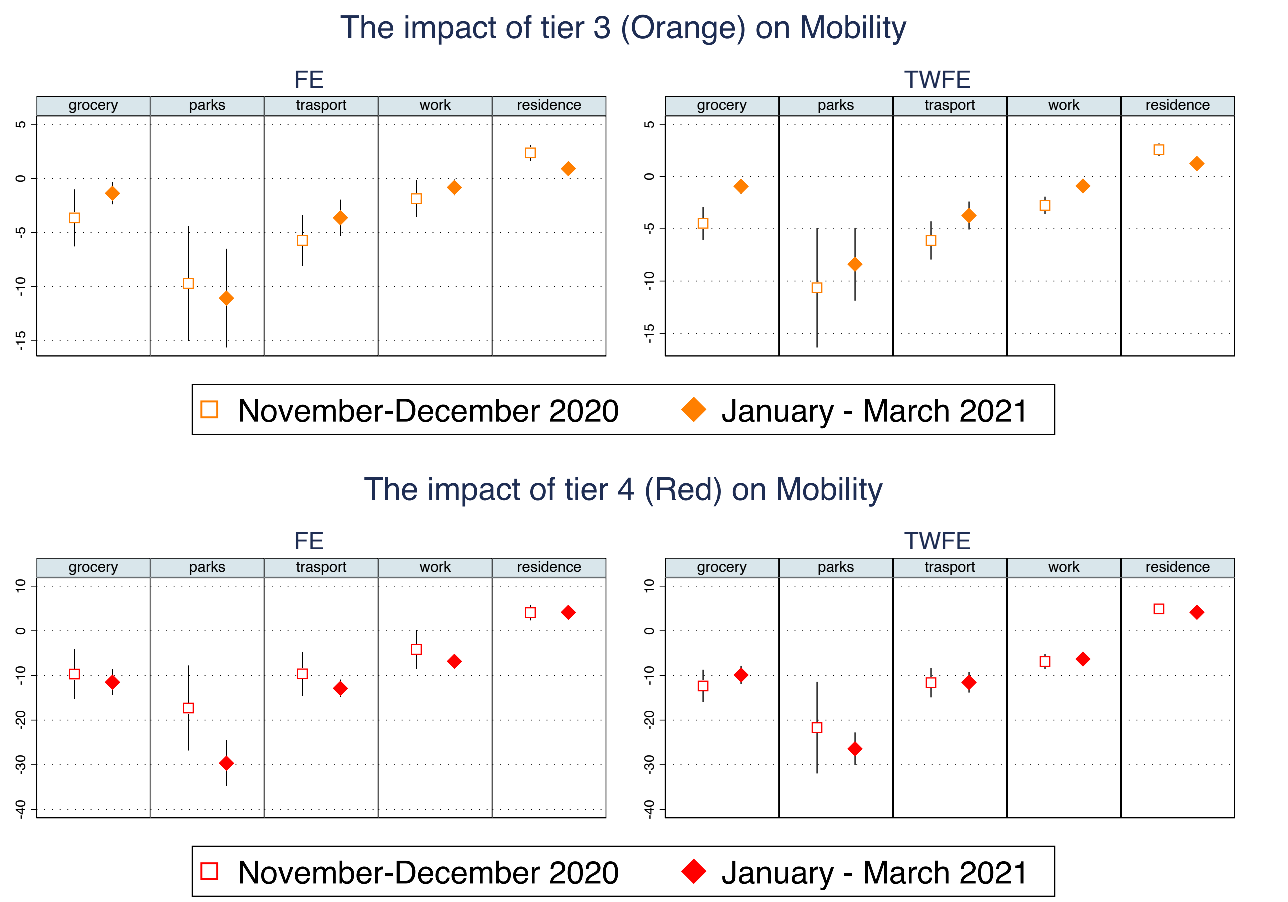

Figure C1: The impact of NPIs on province Google mobility Index estimated with FE and TWFE models.

In this section we corroborate our main findings employing additional outcome series, and specifically other google mobility Indices regarding *grocery and pharmacy, parks, transit station, work place* and *residence*. These series are all highly correlated among each other, with Pearson correlation coefficients above 0.5. With the exclusion of *residence* these indices all show negative averages, in line with increasing social distancing. Thus we run model (1) and model (2) on the full sample and replicate figure 7 with the remaining mobility indices. The results are exposed in figure C1, and are overall broadly in line with figure 7. The impact of NPIs on mobility is negative and significant in all mobility cluster but for *residence*, where the coefficient is, as expected, positive. We can also immediately spot that the two models deliver almost identical coefficients in most cases. We can also see that the magnitude of the coefficient decreases in the January - March subsample, confirming our main findings that the effect of NPIs appear to have diminished over time. The only exception is observable when we employ *parks* as outcome, as the coefficient of NPIs display growing magnitude in absolute terms over time in three out of four cases.

C.2 Apple mobility data

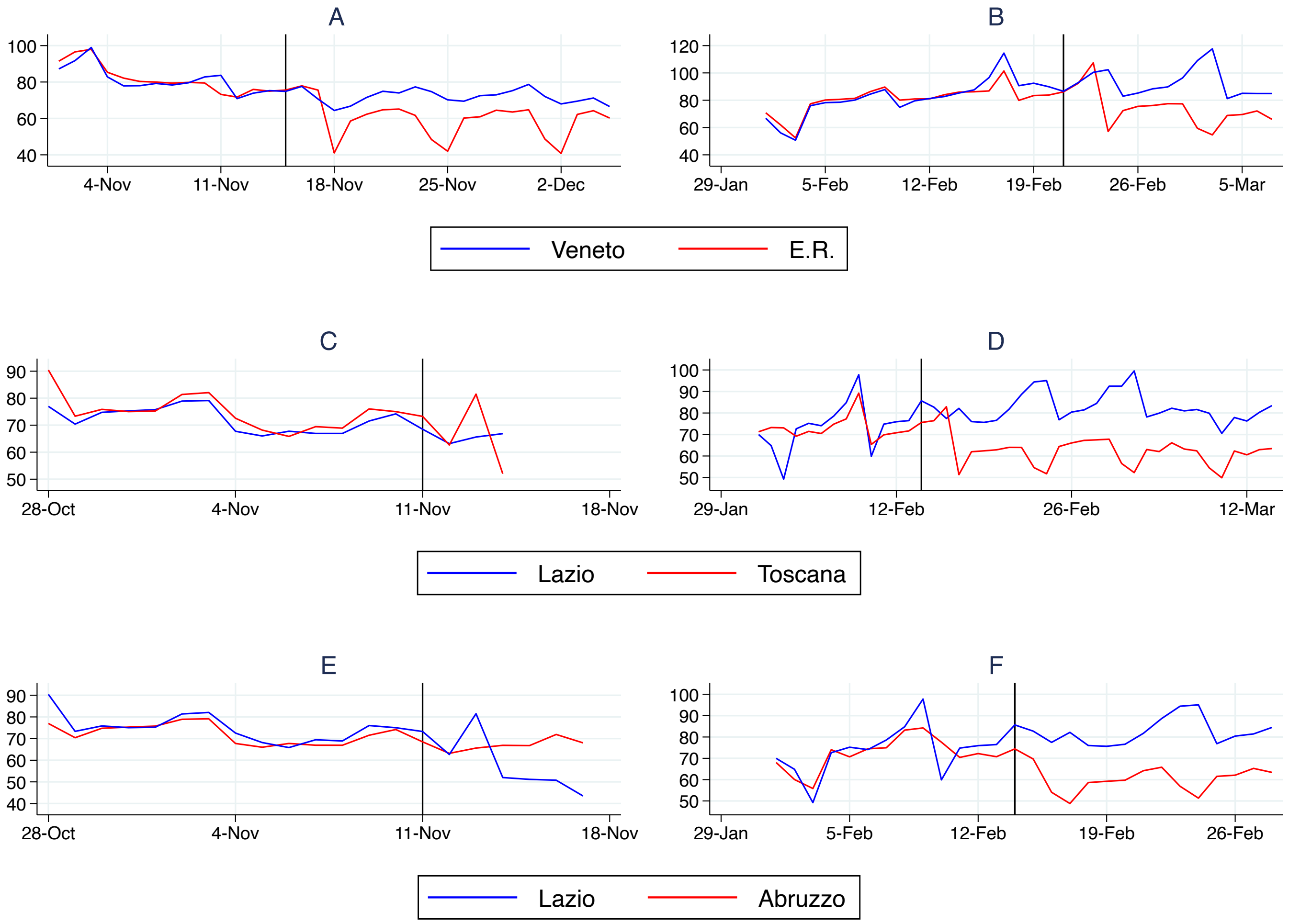

Figure C2: Apple mobility Index - driving

Apple, the major Android's competitor in mobile software devices, has also publicly provided mobility data used by researchers to study Covid-19 (Cot et al. (2021), Kurita et al. (2021)). Apple has produced two type of indices: *driving* and *walking*. There are significant differences between Google and Apple data. First of all, Apple does not provide detailed information about the data collection process and the aggregation (Apple, 2022). Secondly, iOS the operating system running on Apple devices represents only around 24% of market share in Italy between 2020 and 2021 (statista, 2022), contrary to Android that counts for more that 74%. But most importantly Apple mobility data are not available at province level for Italy. Given this limitations we do not include Apple data as dependent variable inside a regression, but instead we visually inspect the data at region level in our case studies. Figure C2 displays the evolution of Apple driving index for treated and control group in the individual case studies. By inspecting figure C2 we can observe that, prior to the policy implementation, the mobility index follows an almost identical path in the treated and control groups. Instead, with a few days of lag after the policy is implemented, the treated units experience a level shift and record lower mobility compared to the counterfactual. Figure

C2 closely resembles the dynamic observed in figure 5.

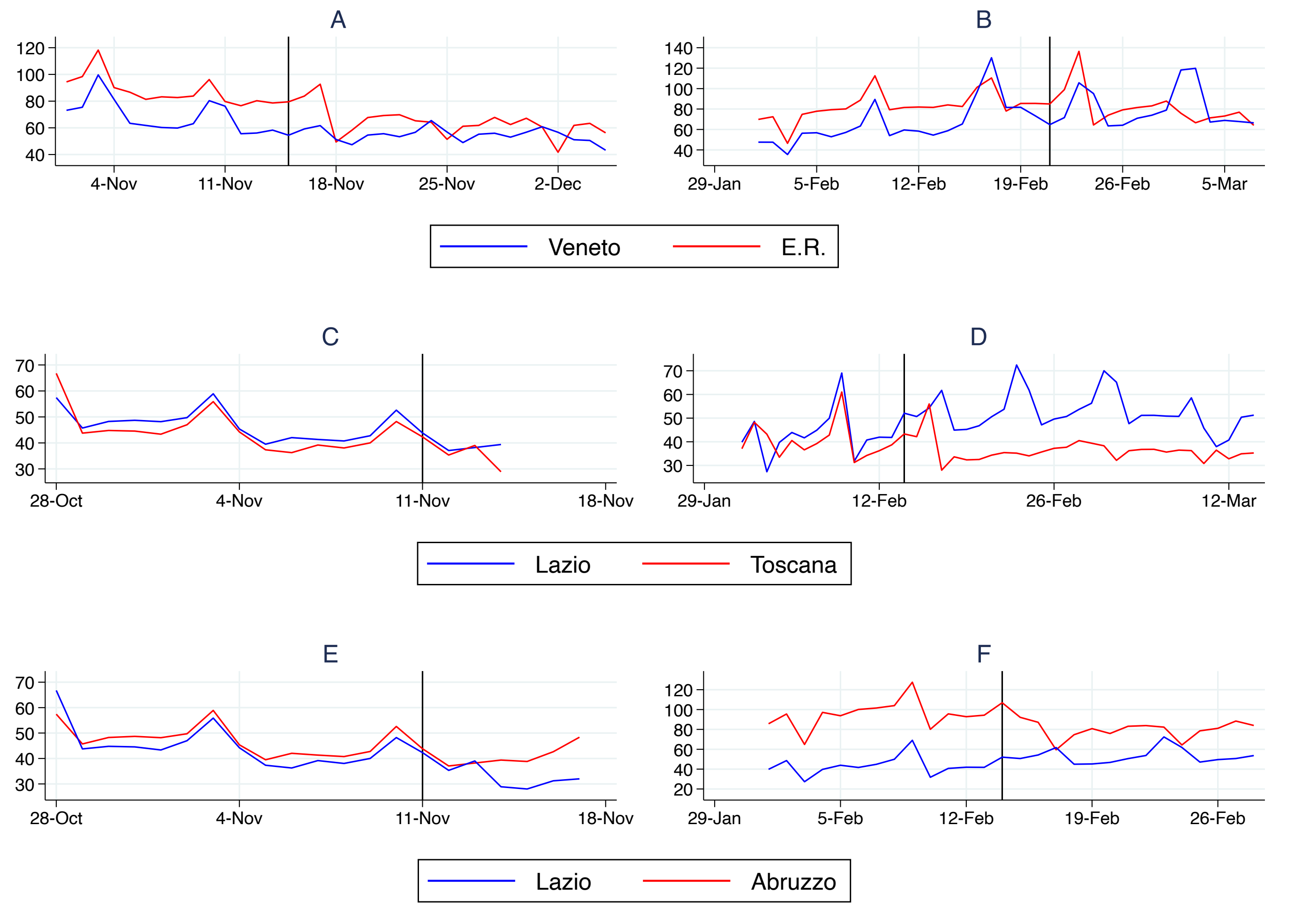

Figure C3: Apple mobility Index - walking

On the other hand, when we inspect the evolution of the Apple walking Index, as displayed in figure C3, evidence of treatment effect appear only in two of six cases. In light of the poor information regarding these data, we cannot speculate about the reason behind this result.

D Robustness Checks

To assess the robustness of our estimation we run a number of checks. First of all, we verify whether our results are driven by a specific group. Hence, we exclude in turn each region from the sample and compare it with the baseline. We also check whether including Piemonte, which we have excluded for inconsistencies in the regional epidemiological data, affects our estimations. Figure D1 plots the estimates of the policy dummies based on model (2) in the two subsamples but excluding individual regions from the samples. The second row in each panel shows instead the estimates obtained in the full sample but including

Piemonte. The main takeaway of fig D1 is that altering the composition of the sample does not dramatically affect our baseline results.

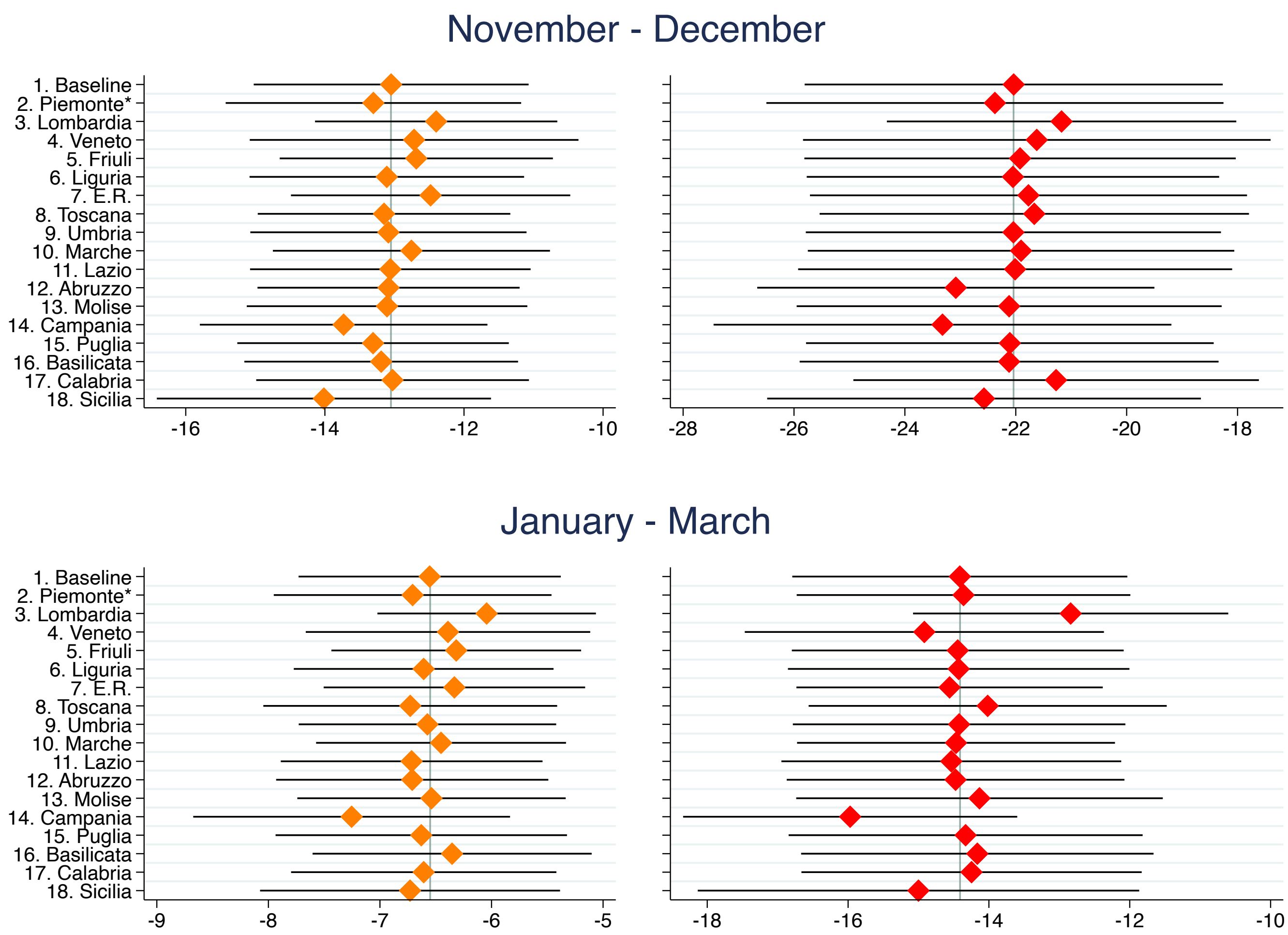

Figure D1: Model (2) estimates in different samples. * The second row of each panel presents the estimates obtained by including Piemonte in the baseline sample. E.R. stands for Emilia-Romagna.

A second check that we perform consists in testing the robustness of the estimates with different regional Covid-19 series. The results are shown in figure D2. The top panel includes the estimates based on model (2) in the first subsample, while the bottom panel includes those obtained in the second subsample. According to figure D2, the choice of regional epidemiological data appears to be irrelevant in the November-December period. Instead, more heterogeneity can be observed in the January-March period. Particularly, the one day lag of the logarithm of ICUs, which we included in our baseline model, appears to be the regressor producing the most conservative estimates. Despite this result, the choice of Covid-19 regional regressor does not alter our main findings.

Furthermore, as in Garmann (2017) we check for potential outliers that might influence our results. We run the procedure proposed by Hadi (1992) to find outliers in multivariate models, and we find zero observations deemed as outliers.

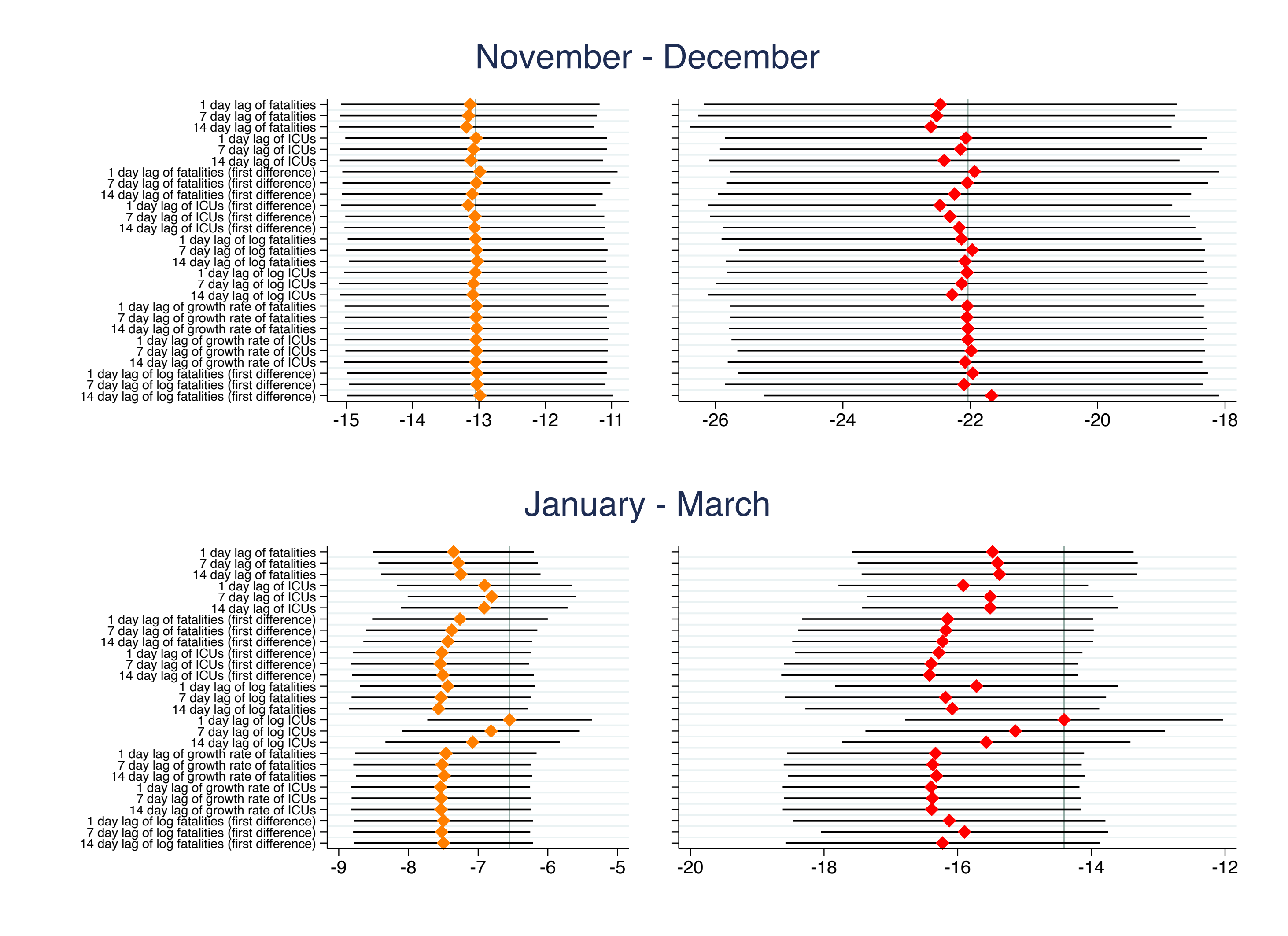

Figure D2: Model (2) estimates with with different regional Covid-19 series.

Lastly, we run model (2) under different assumptions for the standard errors. Table D1 displays the results. We can observe that the choice of estimating the standard errors does not impact our findings as the coefficient of our policy dummies are always significant at 99% confidence level. Column (4) shows the estimates obtained in our baseline, which emerge as the largest standard errors. Column (5) employs the procedure suggested by Conley (1999) that incorporates geographical information to estimate a measure of economic distance.

Table D1: Standard Errors

	(1)	(2)	(3)	(4)	(5)
VARIABLES	ols	robust	cluster	Driscoll-Kraay	Conley
		November 1 - December 23			
dummy = 1 if Orange	-13.05***	-13.05***	-13.05***	-13.05***	-13.05***
	(0.217)	(0.230)	(0.550)	(0.986)	(0.450)
dummy = 1 if Red	-22.05***	-22.05***	-22.05***	-22.05***	-22.05***
	(0.324)	(0.375)	(0.820)	(1.875)	(0.795)
		January 8 - March 15			
dummy = 1 if Orange	-6.548***	-6.548***	-6.548***	-6.548***	-6.548***
	(0.178)	(0.189)	(0.404)	(0.591)	(0.372)
dummy = 1 if Red	-14.41***	-14.41***	-14.41***	-14.41***	-14.41***
	(0.383)	(0.347)	(0.790)	(1.189)	(0.689)

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

All regressions include, but do not report, province and time fixed effects. Columns (1) employes regular ols standard errors, while column (2) and column (3) use respectively robust standard errors and clustered at province. Column (4) uses our baseline Driscoll-Kray standard errors, and column (5) displays the ones obtained with the procedure suggested by Conley (1999).

*** p<0.01, ** p<0.05, * p<0.1