UNIVERSITÀ DEGLI STUDI DI TORINO

ALMA UNIVERSITAS TAURINENSIS

Founded in 1404

DEPARTMENT OF
ECONOMICS AND STATISTICS
WORKING PAPER SERIES

Quaderni del Dipartimento di Scienze
Economico-Sociali e Matematico-Statistiche

ISSN 2279-7114

DOES SIZE REALLY AFFECT TURNOUT? EVIDENCE FROM ITALIAN MUNICIPAL AMALGAMATIONS

SILVIA BOLGHERINI

VINCENZO MOLLISI

Working paper No. 891- February 2024

DOES SIZE REALLY AFFECT TURNOUT? EVIDENCE FROM ITALIAN MUNICIPAL AMALGAMATIONS

Silvia Bolgherini, Vincenzo Mollisi [*]

April 29, 2022

Abstract

Evidence on the electoral participation at the municipal level usually points to a detrimental effect of an enlarged size (due to amalgamation) at the following municipal elections. Contrary to previous studies, our results show an overall positive effect of amalgamation on municipal turnout: larger units do not necessarily vote less than smaller ones. In a quasi-experimental Difference-in-difference design following Callaway and Sant'Anna (2020), we find that the final municipal size per se does not explain turnout after amalgamation. Hence the traditional claim that a larger size should depress municipal turnout does not always hold. Cross- and within municipal heterogeneity emerges instead as a crucial lens for explaining such evidence. In particular, municipalities with higher dissent towards amalgamation show higher turnouts at the following municipal election. This article is the first study relative to Southern Europe and considers all municipalities merged between 2013 and 2019 in Italy.

Keywords: Municipal amalgamations, Turnout, Local institutions' size, Referendum, Consent/dissent, Political efficacy, Difference-in-Difference
JEL: H7, H70, H77, D7, D72

[*]Silvia Bolgherini, University of Perugia, email:silvia.bolgherini@unipg.it; Vincenzo Mollisi, University of Mannheim/University of Turin, email: vincenzo.mollisi@gmail.com.

I Introduction

Over the past half-century, local government amalgamation reforms have occurred worldwide, aiming at improving the efficiency and effectiveness of local public service delivery through economies of scale in public expenditure. These reforms involve the classic trade-off between population size and local democracy (Dahl and Tufte (1973); Denters et al. (2014); Oliver (2000)), where the mainstream assumption is that "small municipalities enhance the democratic quality, whereas larger ones are conducive for economic efficiency" (Koch and Rochat (2017), p. 216). The effect of municipal amalgamations on electoral participation (turnout) is usually considered negative. This study's evidence shows two main points. First, contrary to the expectations and previous studies, we find an overall positive effect of an enlarged size (due to amalgamation[1]) on municipal turnout. Turnout rises in amalgamated municipalities in average around 6 per cent. Second, at a closer inspection on the reasons of such an against-the-current evidence, municipal heterogeneity emerges as a crucial lens for observing different effects of amalgamation on turnout. Using data on (non-binding) referenda concerning municipal mergers, we found that municipalities with higher dissent towards amalgamation show higher turnouts at the following municipal election. To our knowledge this is the first paper to use pre-merger referendum data for measuring the impact of amalgamation on electoral participation. Previous studies have explored pre-merger referenda but in order to assess their role in the amalgamation process at large (Strebel (2019); Miyazaki (2014)) and not their impact on turnout. This article is also one of the first studies on the impact of amalgamation on municipal turnout in Southern Europe. We explore this impact in the Italian municipalities merged between 2013 and 2019[2] and apply a quasi-experimental design following Callaway and Sant'Anna (2020)'s approach of groups treatments, which refines the Difference-in-difference (DiD) method. We then perform an analysis of heterogeneity, based on several dimensions, in order to investigate a possible explanation for the overall result.

This article contributes to the literature on size and local democracy and to the empirical studies on the effects of an enlarged size on electoral participation by rivaling the usual explanation based on the size-and-democracy/political efficacy assumption with an alternative explanation based on the level of dissent/consent towards the amalgamation.

[1] We employ here amalgamation and merger as synonyms, albeit technically merger is the unification of two or more pre-existing municipalities, which may occur either by amalgamation of more units that fuse to form a new body, or by incorporation by a municipality of one or more other units.

[2] Namely, all those for which pre- and post-amalgamation municipal electoral data were available.

II Literature review

The academic debate on territorial rescaling of local authorities is decades-long (Keating (1995). Amalgamation of territorial units, including municipalities, is a form of rescaling through consolidation (Swianiewicz (2010). Studies on municipal amalgamations have focused on different aspects, being clearly pre-eminent those concerning possible economies of scale and public expenditure savings (among many, Tran et al. (2019); Blom-Hansen et al. (2021); Dollery (2010); Roesel (2017); Blesse and Baskaran (2016)).

The studies concerning the effects on turnout at local elections are also present and they mostly argue that a larger size due to amalgamation corresponds to a lower electoral turnout at municipal elections (among others, Heinisch and Mühlböck (2016); Horiuchi et al. (2015); Roesel (2017); Allers et al. (2021)). Several explanations for this negative effect have been attempted. Some studies have shown that in the case of amalgamation, information costs for electoral participation are higher (Lapointe et al. (2018)), implying lower turnout rates in the first election after a merger (Koch and Rochat (2017); Heinisch et al. (2018)). Others have argued that mergers are detrimental to internal political efficacy -namely the belief about one's own competence to participate effectively to political life (Niemi et al. (1991)- as well as to satisfaction with different aspects of local democracy (Oliver (2000); Kjær et al. (2010); Hansen (2015); Zeedan (2017); Van Houwelingen (2017); Voda and Svačinová (2020)) and, as a consequence, mergers exert a depressive effect on turnout.

Nonetheless, rival evidence has also emerged: some empirical studies have found null (Cameron and Milne (2013), Mabuchi (2001), Lapointe et al. (2018)) or limited positive effect of merger on turnout (Kraaykamp et al. (2001), Hicks (2006), Bhatti and Hansen (2019)). According to the pioneer study by Dahl and Tufte (1973), a larger size of a polity should increase its "system capacity", that is the capacity of a political system to respond to, and account for, its citizens' demands, which is expected to have a positive effect on political and electoral participation. Larger municipalities are also expected to have more administrative capabilities to deal with a broader array of tasks (Allers and Geertsema (2016)) and this transfer should increase participation as citizens feel that there is "more at stake" in local politics.

Evidence and literature are therefore quite controverse, and the debate still open. Moreover, the geographical setting of those studies pointing towards a negative effect of mergers on turnout is mainly Continental and Northern European -although with some notable overseas exceptions (e.g., Australia, Japan, Israel, etc.). This may raise questions on the applicability and validity of their results in the Southern European countries, which, due to their different institutional traditions or their mainly centralistic model of local government (Kuhlmann and Wollmann (2019)), could feature different patterns towards amalgamation in respect to the so-called continental Europe federal model, or from the Nordic model.

Countries like Greece have experienced two main amalgamation waves in the 90s and in the 2010s (Hlepas (2010)), Portugal has merged its submunicipal units in 2013 (Rodrigues and Tavares (2020)), Italy has also started an amalgamation policy since the 1990s. And yet, empirical research on the impact of amalgamations on municipal turnout in this geographical area is rare.

Beyond the studies exploring the size/democracy trade-off and the amalgamation/turnout relation, another part of the literature investigates the factors explaining units' consolidation and its keys and obstacles. These studies emphasize the role of endogenous borders and thus heterogeneity, which may refer to cultural, social, economic features (religion, ethnicity, income, etc...), as well as to demographic and spatial features (Alesina and Spolaore (2005); Alesina and Spolaore (1997)). Other studies consider the aggregate effects of either consolidation of, or cooperation among, units (Di Porto and Paty (2018)), with mixed results on efficiency gains and cost reduction (Solé-Ollé and Bosch (2005); Reingewertz (2012); Blom-Hansen et al. (2016)), or on government's efficiency (Allers and Geertsema (2016); Bel and Warner (2016), among others).

Interestingly to our perspective, heterogeneity plays a role in explaining the drivers for merging (Gordon and Knight (2009); Saarimaa and Tukiainen (2014); Bhatti and Hansen (2011)), as well as the influence of 'supply-side' factors, namely how the institutional, political and social contexts of differently-sized municipalities vary, and to what extent these variations affect citizens' participatory attitudes (Oliver (2000). Hence, focusing on some heterogeneity features could reveal to be particularly useful to better investigate the effects of mergers on turnout.

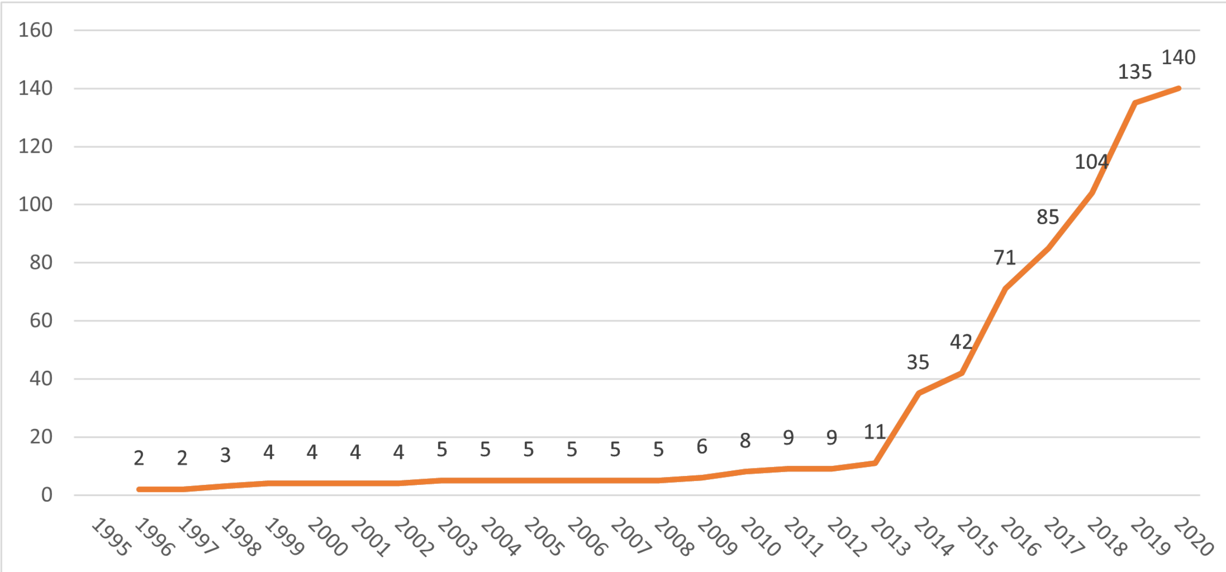
III Institutional Setting: The Italian Amalgamations

Italy is characterized by a high number of small and medium-sized municipalities: the largest share of Italian municipalities (slightly under 70%) still counts a population under 5,000 inhabitants. These figures reflect the historical legacy of Italian Middle Age and its traditionally rooted municipal structure and relevant local territorial attachment. According to Ladner et al. (2019), the number of municipalities in EU and EFTA countries decreased by about 9% in the 1990-2014 span. Italy experienced a reduction of 2.42% in the period 1990-2021 (8,100 municipalities in 1991, 7,904 in 2021), but over 95% of this reduction was achieved from 2011 to 2020. Italian amalgamations have concerned mostly small municipalities thus far, and the size of the merged municipalities remains quite small (Appendix tables A.1-A.2). Even if the number of amalgamations is still very limited, new attention toward defragmentation has undoubtedly emerged.

Italian local government landscape entails municipalities (*Comuni*), provinces (*Province*), metropolitan cities (*Città metropolitane*), and regions (*Regioni*). Five of the 20 Italian regions (Trentino-South Tyrol, Friuli-Venezia Giulia, Aosta Valley, Sardinia, and Sicily), due to the presence of linguistic minorities and/or their geographical peripherical position, were granted a special status (Special Status Regions - SSRs) by the Constitution, thus implying special forms and conditions of autonomy and legislative powers. The remaining 15 Ordinary Status Regions (OSRs) were (homogeneously) given fewer powers than the SSRs. Italy also features a historical divide between the Northern and the Southern area, due to different socio-economic and political culture's conditions. The formal establishment of an amalgamation depends on the local decisions (municipal and regional), and it is voluntary and bottom-up. Citizens express their consent or dissent through compulsory popular referenda, despite the final choice firmly remains in the hands of the regional policymaker that could opt for merging even in case of a negative outcome. Local voluntariness is not the most common feature in the European panorama (Steiner et al. (2016)), although several countries (Finland and Switzerland among others) experienced voluntary or mixed amalgamation forms. Amalgamations in Italy are dispersed in time, following different defragmentation policy periods at the national level. Three main phases (and relevant strategies) deployed, started in the early 1990s and intensified during the years of the economic crisis (Bolgherini et al. (2018), as showed in Fig. 1.

After two decades, an analysis of the effects of this policy tool appears necessary. Thus far, municipal amalgamations in Italy have never been studied in the international context. This article fills this gap by positing a series of theoretical expectations employing the current debate to build our theoretical framework.

Figure 1: Number of amalgamations in Italy per year. 1995-2021.



Notes: No amalgamation occurred in 2021. Source: official data from www.tuttitalia.it.

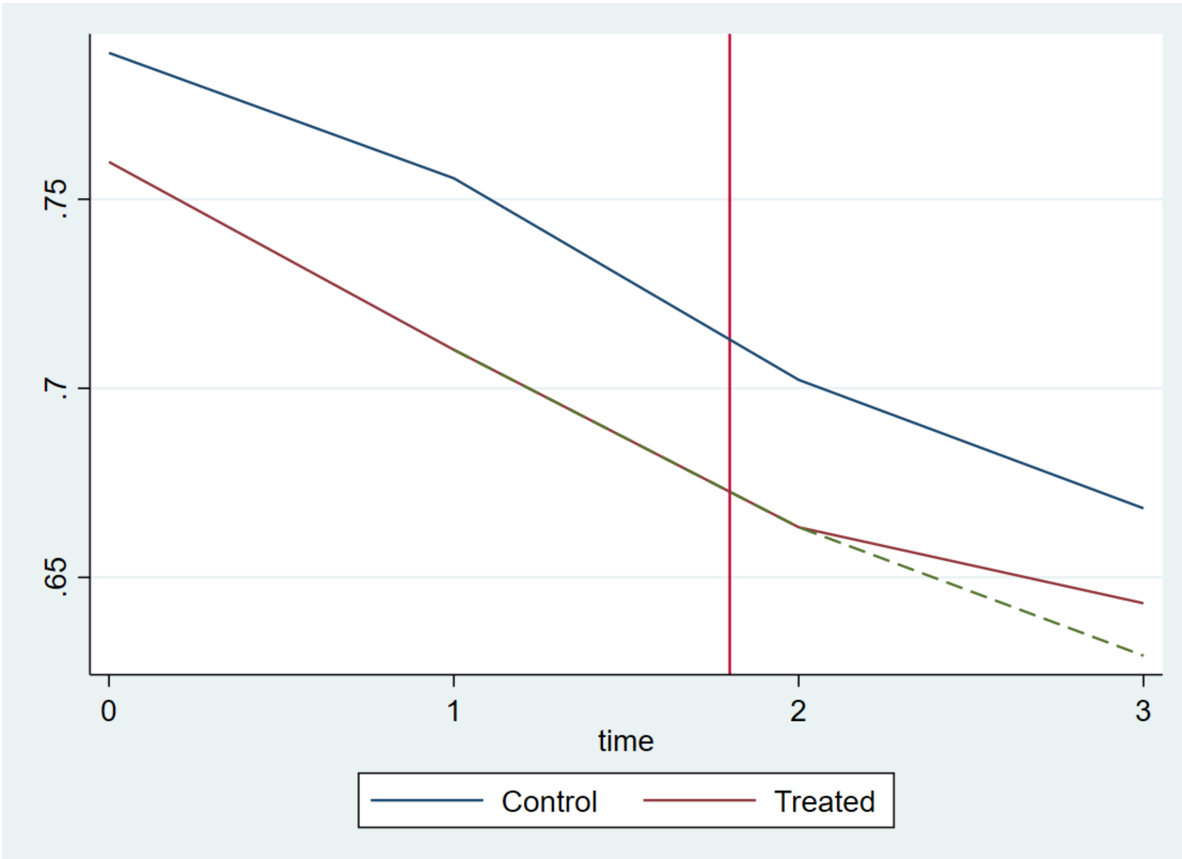
IV Municipal Mergers and Turnout: Theoretical Framework and Relevant Questions

In this section, we explicit the theoretical framework employed for our empirical investigation. Our goal is to test what is the effect of amalgamation on municipal turnout in Italy by focusing on the explicative role of some heterogeneity variables, and by employing the Difference-in-difference strategy, in line with a flourishing strand of local government literature.

We first performed a preliminary analysis, looking at the Italian case by taking electoral data on turnout at four points in time (the 2 pre- and the 2 post-merging elections). Figure 2 shows the average turnout pre- and post-amalgamated for both amalgamated and non-amalgamated municipalities. The post-amalgamation trend points to a direction, which contradicts what most of the literature (although not unanimously, as previously discussed) usually claims: namely, in Italy, merged (treated) municipalities (bottom line) increase their turnouts after merging, instead of decreasing it. Hence, it prompts for further inquiry.

Explanations usually given for the empirically pre-eminent negative amalgamation/turnout relationship are based on the already-mentioned size-and-democracy adagio (Dahl and Tufte (1973)), which on its turn has connections with the rational voting model (Downs et al. (1957)), and the political efficacy theory.

Figure 2: Average turnout pre- and post-merger for treated and non-treated (control) groups. Visual Inspection



Note: Staggered average turnout, two periods before (0; 1) and after (2; 3) the amalgamation. Vertical line indicates the treatment (amalgamation). We can identify pre- and post- period for the control group based on the fact that the amalgamation waves are fixed and non-merged municipalities went to the polls in the same years of the amalgamated ones. We excluded those municipalities that in our timespan held more than four elections (since these are mostly due to exceptional circumstances), and considered the third and fourth election as the two post-merger periods, while the first and the second elections as the pre-merger periods. This is quite a rough definition, but pretty useful for descriptive purposes. Although, we precisely identify the group trends in the following of the paper.

The rational voting theory assumes that voters look for utility and vote effectiveness (Downs et al. (1957); Riker and Ordeshook (1968)). Amalgamations may affect all the components of this voting model: information costs, the degree of the instrumental benefits (Harjunen et al. (2021)), the impact of a single voter (Lapointe et al. (2018)), the degree of sense of community (Koch and Rochat (2017)), and that of the social pressure to vote (Gerber et al. (2008)). These elements are likely to be more penalized, or penalizing, in small jurisdictions than in larger ones. The political efficacy theory relates to the beliefs about the responsiveness of government to citizen demand (external efficacy) and to one's capacity to be effective in politics (internal efficacy) (Niemi et al. (1991)). Merger, and its enlarged size, may affect citizens' evaluation of local politics (Koch and

Rochat (2017)): vote decisiveness and political efficacy are likely to be considerably higher in small communities than in larger ones (Lassen and Serritzlew (2011); McDonnell (2020); Strandberg and Lindell (2020)). Both these mechanism may negatively reflect on turnout in municipal elections.
However, as hinted before, the size-and-democracy paradigm, as well as the instrumental voting and the political efficacy theories may open the way to double-sided reading. As already argued, larger units could enhance systemic capacity (Dahl and Tufte (1973); Dahl (1994)), could rely on larger competence allocation, broader resources and thus provide better and more carefully targeted public policies (Allers et al. (2021); Dollery (2010)) and this should prompt a higher sense of effectiveness and influence on policy outcomes, and thus also a higher sense of political efficacy, which should positively rebound on (electoral) participation.
This competing theoretical arguments stress how multiple variables may contribute to an overall effect on turnout, hardly disentanglable at a first glance. Our first expectation thus plainly claims a non-null effect of merger on municipal turnout:

E1) after municipal merger a non-null effect on municipal turnout is expected

We then now move to the municipal heterogeneity variables, both cross- and within-amalgamation, to disentangle their possible impact on turnout and make sense of the non-mainstream evidence we found in our preliminary analysis (Fig. 2). By cross-amalgamation heterogeneity we mean the diversity among the post-merger final municipalities; by within-amalgamation, the diversity among the pre-merger units. As for the first, we rely, in line with Alesina and Spolaore (2005)'s distinction between demographic and spatial heterogeneity, to the final demographic size of the (merged) municipality. For the second, we consider political and demographic features of each of the pre-merger municipalities: namely, the degree of political dissent towards the amalgamation, the level of participation to the referendum on amalgamation, and the complexity of the amalgamation (different demographic ratios among the merging municipalities). All expectations rely on our theoretical framework's main references, namely the instrumental voting theory's and on the internal political efficacy's assumptions.
First, given the features of Italian amalgamations previously discussed, we are interested in testing to what extent mergers' impact on turnout may be differentiated according to the final (post-amalgamation) size. As previous studies have shown, this impact seems to be different and de-

pendent on the municipality's population final size (Kraaykamp et al. (2001)): hence it could be differently strong/weak across the whole spectrum of the post-amalgamation acquired municipal sizes. Koch and Rochat (2017) show that in Switzerland smaller municipalities experienced more significant turnout declines than larger municipalities did. Heinisch et al. (2018) find the same in the Austrian region of Styria. Independently from the amalgamations, Gendźwiłł and Kjaer (2021) comparatively confirmed that the larger the municipality, the lower the turnout at local elections–although this relation could vanish in very large units, where the numerical differences in size cease to be significant. Moreover, large and small municipalities (are expected to) have different stances towards amalgamation, either disregarding or fearing merger, or instead cherishing it for opposite reasons–according to each municipality's expectation about its own relative position within the merger (Miyazaki (2014); Strebel (2019); Steiner et al. (2016)), e.g. if they expect to be swallowed by larger/stronger/richer partners or to lose their identity or powers, or viceversa if they esteem to be reinforced and empowered. Relying on the above-mentioned studies and considering the size differences of the Italian municipalities -Appendix tables A.1-A.2-, we therefore expect that:

E2) turnout should differently vary in amalgamated smaller-sized and amalgamated larger-sized municipalities

Let us now devote to the within-amalgamation heterogeneity. We first test an expectation based on the complexity of an amalgamation. Being it the number of the merging municipalities (Allers et al. (2021); Bhatti and Hansen (2019)), their mutual population ratios, or the type of configuration (polycentric, satellites, etc.), complexity as a form of heterogeneity has been rarely considered. Several indicators may account for such heterogeneity: e.g., Harjunen et al. (2021) employed geographic political representation to account for it. We employ for this purpose the demographic unbalance among the components. We employ the population unbalance ratio among the merging municipalities as a measure of complexity by assuming that the higher the population unbalance among them, the higher the amalgamation's complexity[3]. To formulate our expectation we rely on the instrumental voting theory, which argues that voters cast the ballot when the perceived influence of the individual vote on government policy outcomes is strong. As said, the instrumental

[3]For the sake of clarity the within-merger heterogeneity is the different population of the merging units and is used to calculate each merged unit's complexity. Hence, complexity refers to the final merged municipality and as such could also be read as cross-amalgamation heterogeneity.

voting theory may have a double-side interpretation. On the one hand, the perceived influence in highly heterogenous and larger communities is expected to decrease: in this case turnout is expected to diminish when complexity increases. A high unbalance could in fact trigger identity conflicts, absorption syndromes, power allocation struggles, etc., among municipalities, thus hampering individual votes' (perceived) influence and political efficacy (Lassen and Serritzlew (2011); Jakobsen and Kjaer (2016)). On the other hand, larger (merged) units, having more influence over policies, could enhance effectiveness and thus turnout (Allers et al. (2021)). In Italy final merged municipalities remain small. That means that the benefits deriving from the heterogeneity of preferences, (Oliver (2000)), or from the possibility to better address citizens' needs, both typical in enlarged units, McDonnell (2020), seem unlikely to be realized. Be as it may, to account for both views, in presence of different degrees of complexity, we expect that:

E3) the higher the complexity of an amalgamation, the higher the variation in turnout after amalgamation

Italian national legislation requires compulsory referenda to be held in all concerned municipalities before the merging in order to poll the citizens' stances towards this territorial reform. The referendum may raise attention on the amalgamation process itself, thus making the merger a 'hot issue' in the eyes of the citizens. Both the degree of political consent/dissent towards the amalgamation, and the level of participation to the referendum in each pre-merging municipality, can be considered as indicators of societal connectedness (Bhatti and Hansen (2011)) and thus of political homo/heterogeneity among the merging units. Forms of identity protest driven by several factors such as fear of status' loss (Steiner et al. (2016)) may arise. Strebel (2019) argues in his study on a group of Swiss cantons that both functional pressures (e.g. the financial situation and the argument for attaining economies of scale) and self-determination concerns (e.g. the relative size within the new unit, or the future of the community's cohesion) may affect the results of a referendum on mergers - with the latter being subordinate to the former: self-determination matters only when functional pressures are contained. Weinbacher (2018) and Jenny and Ennser-Jedenastik (2014) analyze the different levels of dissent against mergers in Austria and find that smaller municipalities are more likely to show a higher level of dissent. In Canada Hicks (2006) finds a positive effect of amalgamation when strong public protests accompanies the reform. Others, e.g. Zeedan (2017), state instead that increased turnout reflects citizens' dissatisfaction, while Castater and Han (2018) argue that disagreement with a policy strategy - conceived, at large, as a social pact - can produce stronger effects on elections held close to the merger reform. Again, to account for both possible (positive or negative) directions, we expect that:

E4) the higher the dissent towards a merger, the higher the variation in turnout at the municipal election

Furthermore, exploiting the data collected on the pre-merger referenda, we additionally test the relation between the turnout at these referenda and the turnout at the post-merger municipal elections, although without setting any specific hypothesis on this since the mechanisms linking the two turnouts are not straightforward. The result will be shown in the models along with the other variables to support the analysis on the pre-merger phase.

V Data and Methods

Electoral participation is the activity to go to the polls and vote. Local electoral participation refers here to as the turnout at the municipal elections. More precisely, we focus on the turnout in municipal elections before and after the amalgamation and on its variation in the merged municipalities compared to the non-merged ones, which is our dependent variable. We considered all municipal amalgamations occurred in Italy from 2013 to 2019, namely all those whose electoral data were fully available; for previous amalgamations pre-merger electoral data could not be retrieved. We nonetheless covered over 90% of all Italian municipal amalgamations (Fig. 1).

The local electoral data in Italy does not allow to track the pre-merging polling districts of a municipality after it amalgamated into a new unit. Hence, after amalgamation, we assign the turnout of the merged municipality (final unit) to each municipality forming an amalgamation (constituent unit). Said otherwise, turnouts of each constituent unit are compared to the post-amalgamation (unique) turnout of the final unit, by assigning this post-amalgamation turnout to all constituent units for pre/post comparison. Alternative strategies, such as in Lapointe et al. (2018) for Finland, pay some arbitrariness downsides as well[4].

Turning to the independent variables, municipal final size is measured as the number of inhabitants in the final (merged) unit. This variable, accounting for the cross-amalgamation heterogeneity, is a dummy equal to 1 if the final population of the merged municipality is above the 50th percentile of the inhabitants distribution of the municipalities in the sample - where the median is 1,796 - equal to 0 if below. For the cross-amalgamation heterogeneity, we also controlled for the geographical position - accounted by a binary variable, which is a dummy equal to 1 if the final unit is classified as belonging to a northern-central region, and 0 for southern-central regions - and for the institutional status of the region that a merging municipality belongs to - accounted by a binary variable equal to 1 if the final unit is located in a Special Status Region (SSR), 0 if in an Ordinary Status Region (OSR).

We measure complexity as the (un)balance of population among the merging municipalities calculated through a Gini index. We split the complexity distribution into two sub-samples, namely under and above the median (21%), respectively. The complexity variable is a categorical variable equal to zero for the control group, 1 for treated municipalities with an (un)balance of population below the median (simple amalgamation, thus low complexity), 2 for treated municipalities with such (un)balance above the median (more complex amalgamation).

The degree of consent/dissent towards a merger accounts for the within-amalgamation heterogeneity and is computed starting from the number of Yes and No votes expressed by the citizens at the non-binding compulsory referendum to be held before a possible amalgamation. Data on the Italian municipal referenda (both on their results and participation) before amalgamation are reasonably a strong proxy for consent/dissent toward this territorial choice. Moreover, the peculiar Italian institutional setting allows strong validation since referenda data cover all merged municipalities' populations and not just a sample. The Yes votes in favor of the amalgamations ranged between 20 and almost 100 percent. A higher percentage of yes means a higher consent and satisfaction towards the merger and viceversa. Low consenters (or, reversely, stronger dissenters) are thus considered all those municipalities below the majority of yes, they supported the amalgamation only lukewarmly; stronger consenters (reversely, weaker dissenters) all those above the majority, namely their percentage of yes-votes is over 50%, and their satisfaction with the amalgamation project is higher. We employ a categorical variable equal to 0 for the control group, 1

[4]The reconstruction of the merged municipalities' turnout from voting data in pre-treatment units provides similar baseline results (Appendix table A.5).

for treated municipalities with referenda positive outcome below 50% (weaker consenters, lower satisfaction and higher dissent), 2 for treated municipalities with referenda positive outcome above 50% (stronger consenters, higher satisfaction thus weaker dissent).

In addition, we run a further test with the level of participation to the same non-binding referendum on amalgamation. We measured it with a categorical variable equal to 0 for the control group, 1 for treated municipalities with referenda turnout below the 50% (low Ref-participation), 2 for treated municipalities with referenda turnout above 50% (high Ref-participation). Table 1 shows some descriptive data on the independent variables.

Table 1: Heterogeneity variables: descriptives

		Freq.	Percent	Valid	Cum.
DDD: Final Size	0	1947	49.92	49.92	49.92
	1	1953	50.08	50.08	100.00
Total		3900	100.00	100.00	
DDD: North	0	937	24.03	24.03	24.03
	1	2963	75.97	75.97	100.00
Total		3900	100.00	100.00	
DDD: SSRs	0	3162	81.08	81.08	81.08
	1	738	18.92	18.92	100.00
Total		3900	100.00	100.00	
DDD: Complexity	0	3034	77.79	77.79	77.79
	1	436	11.18	11.18	88.97
	2	430	11.03	11.03	100.00
Total		3900	100.00	100.00	
DDD: Consent/Dissent	0	3034	77.79	78.34	78.34
	1	36	0.92	0.93	79.27
	2	803	20.59	20.73	100.00
Total		3873	99.31	100.00	
DDD: Ref-Participation	0	3034	77.79	78.10	78.10
	1	436	11.18	11.22	89.32
	2	415	10.64	10.68	100.00
Total		3885	99.62	100.00	

To test our expectations, we employ a Difference-in-difference (DiD) methodology. DiD is frequently employed to assess public policies and, in particular, the possible impact of amalgamations on different municipal aspects (among others Lassen and Serritzlew (2011); Reingewertz (2012); Blom-Hansen et al. (2014); Blom-Hansen et al. (2016); Blesse and Baskaran (2016); Allers and Geertsema (2016); Roesel (2017); Koch and Rochat (2017) ; Tricaud (2021)).

A very recently developing literature on this method (De Chaisemartin and d'Haultfoeuille (2020); Callaway and Sant'Anna (2020); Goodman-Bacon (2021); Sun and Abraham (2020) points to challenges in interpreting the estimated DiD effect in the traditional way, namely through the two-way fixed effects models when treatment effects are heterogeneous (across either groups or time periods). Contrarily to most studies that employ a staggered (two-way fixed effect) DiD, we use the waves of mergers as groups to identify the different effect of amalgamation on the turnout by year-group[5]. When there is a 'selective treatment timing', the estimated effect ends up being weighted averages of treatment effects across different lengths of exposures. In our case it means that municipalities in different amalgamation waves experience systematically different effects of the treatment from municipalities in other waves (e.g., municipalities choose to merge in earlier periods if they expect larger benefits from participating earlier in the treatment). This is possibly present in Italy, where several amalgamation waves and simultaneous municipal elections involving both treated and non-treated municipalities occur. The theoretical and methodological evolution of the last years came exactly from acknowledging the staggered nature of phenomena such as municipal amalgamations, and thus the risks of inappropriateness in employing the traditional DiD method.

We thus proceeded to a more accurate DiD estimation following Callaway and Sant'Anna (2020)'s application in case of multiple time periods. We assume the Neyman-Rubin potential outcome framework, where $Y_{it}(0)$ is the untreated unit i's potential outcome and $Y_{it}(g)$ is the unit i's potential outcome in time period t if treated in period g.[6] The parameter of interest -the Average Treatment effect on the Treated (ATT)- is a group-time average treatment effect, namely the average effect of participating in the treatment for units in group g at time period t:

$$ATT(g,t) = E[Y_t(g) - Y_t(0)|G = g] \tag{1}$$

Estimation of the average effect is provided under the following parallel trend assumption with the never-treated observations. [7]

[5]Namely this approach estimates the effects of the treatment for each year-group and then aggregates them in a single average effect, weighting the different effects in respect to their relevance in the sample.

[6]Groups are often defined by the time period when a unit becomes treated, so G_i is also the notation for groups.

[7]A different parallel trend assumption uses the not-yet-treated observations:

For all $t \geqslant g$:

$$E[Y_t(0) - Y_{t-1}(0)|G = g] = E[Y_t(0) - Y_{t-1}(0)|C = 1]. \quad (2)$$

where C is an indicator variable for whether unit i is in a never-treated group. This is a natural extension of the assumption of parallel trends in the case of two periods and two groups' approach. It states that, without treatment, the average untreated potential outcomes for the group first treated at time g and for the never-treated group would have been parallel at all post-treatment periods $t \geqslant g$. Applying the parallel trends assumption based on never-treated units, one has that for all $t \geqslant g$:

$$ATT(g, t) = E[Y_t - Y_{g-1}|G = g] - E[Y_t - Y g - 1|C = 1]. \quad (3)$$

With the DiD method it is assumed that the treatment explains the main difference between the groups. If the two groups had experienced similar trends in the pre-merger period, it is plausible that turnout would have continued to evolve in a parallel way in the post-treatment period in the absence of treatment (amalgamation) (dotted line Fig.2). As Kahn-Lang and Lang (2020) suggest, DiD will generally be more plausible if the treatment and control groups are similar in levels, not just in trends.

We thus employ a Propensity Score Matching (PSM) based on a nearest-neighbour matching algorithm to select municipalities (Section 2 Appendix) and create a suitable control group. We resort to a multivariate matching procedure (Diamond and Sekhon (2013)) based on pre-treatment outcomes and covariates (Steventon et al. (2013); Kreif et al. (2016)). Ryan et al. (2019) illustrate, via simulations in a context of health policy interventions, that a matched-DiD performs suitably even in presence of non-parallel trends, thus confirming the importance of the matching procedure in improving the DiD performance.

In order to proceed with our analysis, we provide pre-testing evidence of the impact of mergers

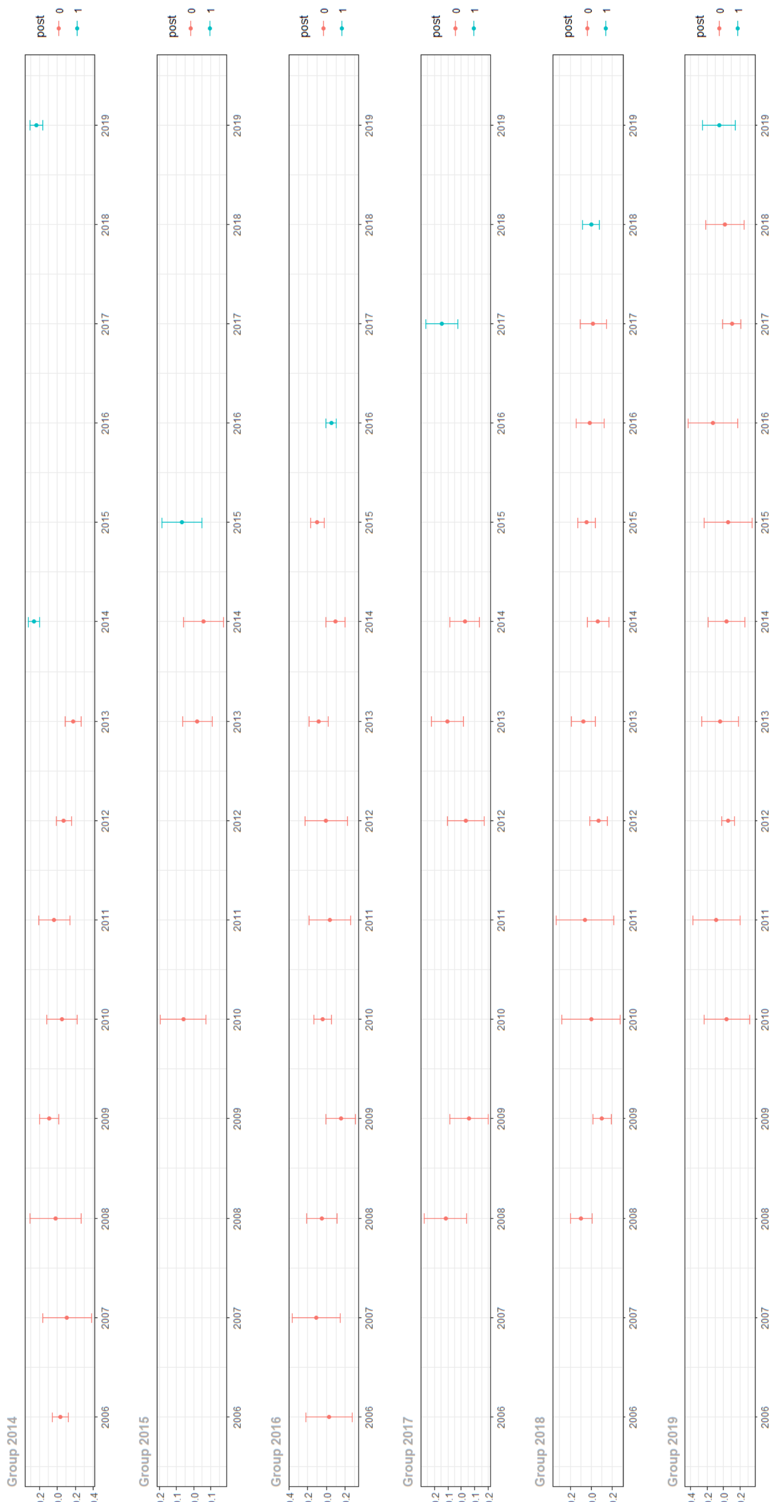
$$E[Y_t(0) - Y_{t-1}(0)|G = g] = E[Y_t(0) - Y_{t-1}(0)|D_s = 0, G \neq g].$$

where D_{it} is an indicator of whether unit i has been treated by time t. This assumption states that one can use the not-yet-treated by time s ($s \geqslant t$) units as valid comparison groups when computing the average treatment effect for the group first treated in time g. In general, this assumption uses more observations when building comparison groups than the smaller group of the never-treated. Details and estimates using this different parallel trend assumption are reported in Appendix Section 3. Results are very similar.

on turnout, through the implementation of an event study. This is just a pre-test, so that it is possible that the identifying assumptions are true in prior periods, but not in the current one -or viceversa. For these reasons, we consider the pre-test as a piece of evidence for the credibility of the DiD design. In our context, pre-tests based on group-time average treatment effects (or based on group-time average treatment effects that are aggregated into an event study plot) are required. Indeed, if the test holds for the single year-groups, then it also holds even in the presence of selective treatment timing.

Figure 3 shows separate event study plots for each group. The figure contains estimates of group-time average treatment effects for each group in each time period, along with a simultaneous confidence interval. The red dots in the plots are pre-treatment pseudo group-time average treatment effects and are useful for pre-testing the parallel trends assumption. The blue dots are post-treatment group-time average treatment effects and should be interpreted as the average effect of participating in the treatment for units in a particular group at a particular point in time. The uniform confidence bands of the pre-treatment period, most of the time, contain the zero: one thus fails to reject parallel trends in pre-treatment periods. The acceptance of the parallel trends is also confirmed by a Wald test that strongly rejects the null hypothesis of failure in the parallel trends. The graphs show different dynamic patterns of the turnout by groups, but a strong positive effect persists in most of the groups.

Figure 3: Event Study Plots by Year



Notes: The rows shows the groups of municipalities that merged from 2014 to 2019, respectively. Red lines give point estimates and uniform 95% confidence bands for pre-treatment periods. Under the null hypothesis of the parallel trends assumption holding in all periods, these should be equal to 0. Blue lines give point estimates and uniform 95% confidence bands for the post-treatment effect.

To investigate the different effect of amalgamation on turnout of the variables of cross- and within-amalgamation heterogeneity, we employ a triple difference-in-difference model (DDD) following Goodman-Bacon (2021)'s model as detailed in its Appendix D. Assume that units belong to either an affected group ($H_i = 1$) or an unaffected group ($H_i = 0$). When some units should not be affected by a given treatment (here amalgamation), they can be used as a falsification test and used to compute the DDD estimator. However, if the affected group would result more represented in some waves of the treatment (amalgamation waves), the classical DDD estimator may not capture relative trends by heterogenous dimension within waves. DDD specifications as in 4, which include a more saturated set of fixed effects, can overcome this problem. If treatment is rolled out by waves (w), for example, a DDD model can include waves-by-time fixed effects ($a_w a_t$) along with unit (a_i), time (a_t) and waves fixed effects (a_w) as follows:

$$y_{it} = a_i + H_i a_w + H_i a_t + a_w a_t + \beta^{DDD} D_{it} G_i + \epsilon_{it}, \tag{4}$$

where β^{DDD} equals the difference between the treatment effects of DDs estimated on the affected and unaffected sub-samples. Estimates from 4 net out changes across affected and unaffected group within waves/year cells[8].

VI Findings

In this section we present, first, our results concerning the overall impact of amalgamation on electoral municipal turnout; then we discuss the results concerning the expectations on the heterogeneity factors that may help disentangle the several determinants of this overall impact.

VI.1 Baseline results: the overall impact

In the bottom panel of table 2 we report the average treatment effect for each year-group. We observe a general positive effect of the amalgamations on municipal turnout. Still, there is some evidence of negative effects: one of the group-time average treatment effects (2016) is negative and statistically different from 0. In order to appreciate the overall effect and assess which is

[8]This changes the weights, though, because the introduction of variation across the third difference within a cell leads to the typical OLS result that cells with more variation get more weight. In this case, more variation in sample membership within a cell means approximately equal numbers of units with $G_i = 1$ and $G_i = 0$.

the prevalent direction, these group-time average treatment effects are aggregated. In the top panel (overall ATT), we present a more general-purpose overall treatment effect parameter, that is given by computing the average treatment effect for each year-group, and then averaging across groups. This sort of procedure provides an average treatment effect parameter with a very similar interpretation to the Average Treatment Effect on the Treated (ATT) of the traditional two-periods-and-two-groups DID case. Our results show that an amalgamation increases turnout by 5.7% and its effect is statistically significant, with confidence bands that do not contain the zero.

Table 2: Baseline. The effect of amalgamation on turnout.

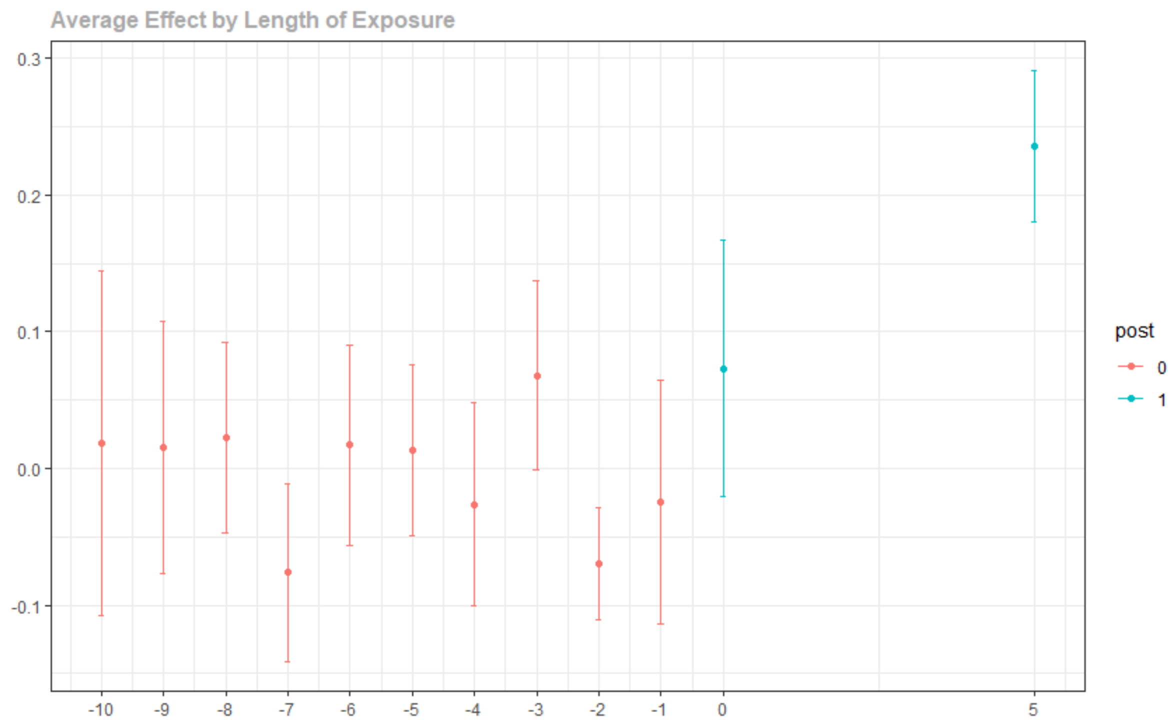
Overall ATT:	Std. Error		95% Conf. Int.
0.057	0.0069		[0.0439 , 0.0709]*
Group Effects:	**ATT**	**Std. Error**	**95% Simult. Conf. Int.**
2014	0.2485	0.0188	[0.2021 , 0.2949]*
2015	0.0680	0.0329	[-0.0133 , 0.1492]
2016	-0.0494	0.0181	[-0.0941 , -0.0046]*
2017	0.1430	0.0332	[0.0610 , 0.2251]*
2018	0.0013	0.0242	[-0.0584 , 0.0610]
2019	0.0010	0.0033	[-0.0072 , 0.0092]
—			

Notes: Top panel: the overall ATT displays the average treatment effects across all lengths of exposure to the treatment (across groups) following Callaway and Sant'Anna (2020). Bottom panel: the group-time effects show the average treatment effects specific to each year group. '*' = confidence band does not cover 0.

In fig. 4, we report the aggregated event study plot. The graph shows an even more clearer pattern of pre-treatment confidence bands including zero. The post-treatment dynamic is strongly positive. By using municipalities from group 2014, which were actually able to vote in a second post-treatment period, it is possible to evaluate the persistence of the effect. After five years (blue line, right-side of the chart) the effect is around 0.20 and 0.25, a huge effect indeed. However, it is not possible to distinguish the average effect across groups at this lead from the group time effect. This effect is in fact produced by the 2014-group only (those municipalities in our sample that voted twice) and it is not possible to generalize it to the whole sample.

In any case, a positive effect is clearly detected. This confirms our E1). It also confirms our preliminary analysis presented in Section IV, and thus the Italian case belonging to the minority of case studies where turnout seems to be positively affected by amalgamation and not the opposite, as instead claimed by most of the literature.

Figure 4: Effect of amalgamation on turnout (overall ATT). Aggregated Event Study Plot.



Notes: The overall ATT computes the average treatment effect across all lengths of exposure to the treatment (across groups) following Callaway and Sant'Anna (2020). Red lines give point estimates and uniform 95% confidence bands for pre-treatment periods. Under the null hypothesis of the parallel trends assumption holding in all periods, these should be equal to 0. Blue lines provide point estimates and uniform 95% confidence bands for the treatment effect. Lags and leads in the x axis indicate, respectively, the number of years before and after year 0, corresponding to the first post-amalgamation election.

VI.2 Testing the heterogeneity variables: evidence and possible explanations

Let us now devote to try to account for such an evidence, and to test our expectations about heterogeneity. To do that we run a DDD model (triple Difference-in-difference) based on Goodman-Bacon (2021), whose results are shown in table 3.

E2 concerned the final size of the merged municipalities in order to detect possible dimensional effects on turnout. We find a null effect of amalgamation on municipalities above the 50th percentile of our population distribution in the final merged municipality in respect to those below: the turnout is more or less equal, no matter the final size of the municipalities. This finding is particularly interesting since it contradicts the main assumption on the size effect, namely that the bigger the size the lower the turnout at the local polls. Our data show that in the Italian case the final size -how big the municipality becomes after merging- does not impact the turnout at the following municipal elections. The political efficacy theory and size-and-democracy assumptions do not hold here, an evidence we try to explain in the conclusive section.

In table 3, we also test E3. Findings disconfirm the expectation: amalgamation has not a statistically significant effect on more complex units (those above the median) in respect to less complex ones (=1 as baseline). It means that those municipalities members of more complex final units (i.e., more unbalanced in respect to their constituent units' population), tend to vote more similarly than expected to those in amalgamations formed by municipalities with a more balanced number of inhabitants. Also controlling for the geographical position of an amalgamated municipality (assigned through a binary classification of Italian regions in Northern and Southern ones) on the turnout after amalgamation, does not bring to any significant effect. The same applies when controlling for the institutional (Special or Ordinary) status of the region where the amalgamation is located.

Table 3: Heterogeneity Analysis

	(1) turnout	(2) turnout	(3) turnout	(4) turnout	(5) turnout	(6) turnout
DDD: Final Size	0.00673					
	(0.73)					
DDD: North		0.0165				
		(1.38)				
DDD: SSRs			0.0116			
			(0.60)			
DDD: Complexity				0.000294		
				(0.02)		
DDD: Consent/Dissent					-0.0504***	
					(-3.96)	
DDD: Ref-Participation						-0.0443***
						(-3.18)
Observations	3877	3877	3877	3877	3847	3847

Notes: All categorial variables have baseline=1, binary variables have baseline=0.
T statistics are in parenthesis. * $p < .1$, ** $p < .05$, *** $p < .01$.

Results displayed in columns (5) and (6) confirm instead our E4. In those municipalities where a higher consent towards the merger was present, the DDD turnout effect is negative of 5 percentage points and statistically significant. Said otherwise, a higher dissent boosts turnout variation of about 5 points in respect to those municipalities where the amalgamation has been more largely accepted. Statistically, moving from low consent (our baseline) to high consent depresses turnout of about five percentage points. Assuming - as we did - that consent towards a merger implies satisfaction for the amalgamation policy, the evidence means that the DDD marginal effect of the satisfaction with the merger (high consenters) on turnout, being negative of about 5%, almost counterbalances the overall positive effect (of about 6 points) found in the baseline model. Consent with the amalgamation policy lowers the overall effect to nearly zero[9], which would be a result in line with the literature on the scarce or even detrimental effect of amalgamation on turnout. On the other hand, this evidence points out how the consent/dissent variable is possibly a crucial one in explaining the overall positive - and against-the-current - effect that we found in the Italian case. This is confirmed when we look at column (6), where we additionally report the impact of participation to the local non-binding referenda concerning amalgamations: referenda with low participation are those with referendum turnout under 50% (scoring 1 as baseline; 2 if referendum turnout is high, i.e. over 50%). In those municipalities with low participation, a similar negative and statistically significant DDD marginal effect of 4.4 percentage points on turnout emerges. Analogously, moving to low participation (baseline) to higher participation depresses turnout of 4.4 percentage points. The positive (cor)relation between a high consent to the amalgamation policy and a high participation to the referendum concerning the amalgamation has of course to be tested both theoretically and empirically. Here it would suffice to note that also the Ref-participation variable provides some hints of the fact that the consent/dissent issue has several intriguing aspects -and probably explanatory ones, although possible double-sided as well- and thus needs to be further disentangled in future researches.

VII Conclusions

This article contributed to the debate on the relationship between municipal amalgamation and turnout by providing the first study on the Italian case, and one of the few on Southern Europe. The article examined the effects of the 2013-2019 municipal amalgamations (nearly 95% of all Italian

[9]This reasoning is possible since here we are looking at the effects of a triple DiD, namely at the differences not only between treated and non-treated pre- and post-amalgamation, but also between two groups of treated - here consenters and dissenters, in other variables, simple and complex, etc... - always pre- and post.

mergers) on the turnout in their subsequent municipal elections by employing the Difference-in-Differences method following Callaway and Sant'Anna (2020)'s group treatment approach.

The Italian case not only strongly disconfirms the claim that amalgamation has negative or null effect on turnout, but it also shows that this overall effect is positive, and of about six percentage points.

The empirical investigation to account for such an against-the-current evidence was conducted according to a theoretical framework elaborated starting from the size-and-democracy's, instrumental voting's and political efficacy's main arguments.

The possible explanation for turnout variation in municipal election after amalgamation relies on the consent/dissent issue. The degree of consent/dissent –measured through the pro/against-merger referendum vote– accounts for around five percentage points of turnout variation at the first municipal election after amalgamation. More precisely, where amalgamation was more accepted (namely where the consent to the amalgamation referendum was higher) the turnout at the elections was lower; conversely, where dissent towards the amalgamation policy was more robust (those municipalities with lower consent, namely lower yes-rate at the referendum) turnout increased at the following elections, probably to voice this dissent. This explanation is possibly supported (with the mentioned caveats) by the introduction of the variable on the participation at the referendum, which additionally points to this explanatory direction and further encourages this research path. Other heterogeneity factors we controlled for -such as the complexity of an amalgamation or its final size- could not satisfactorily account for such a post-merger increase in turnout.

This evidence points to the fact that the different degree of consent -and thus acceptance and satisfaction- of the citizen towards the merger policy is the variable that really matters in explaining turnout variation after amalgamation. Consent with the amalgamation policy lowers the overall effect to nearly zero–which would be a result in line with the literature on the scarce or even detrimental effect of amalgamation on turnout–while dissent increases turnout at the following municipal elections, and explains the positive variation we found in our analysis.

Moreover, as second takeaway point of this research, the final municipal size per se does not explain turnout after amalgamation, since this variable does not hold to a statistically significance test–as well as the complexity of an amalgamation, namely the population (un)balance among its constituent units. Hence the traditional claim that a larger size (caused by the merger) should depress municipal turnout does not hold–at least in the Italian case. This is in line with Kraaykamp et al. (2001) already hinted in the early 2000s, namely that (final) size should matter little on turnout.

Finally, our results allow to reconsider the classical explanations of the effect of municipal amalgamation on turnout based on both the size-and-democracy assumptions and on the political efficacy and the instrumental voting theories, namely that an enlarged size following a merger should weaken local democracy, lower the perception of own's capacity to be influential in politics and thus should depress the turnout. Consent/dissent may trigger different political efficacy or instrumental strategies, which may entail both voice and exit voting strategies with different and possibly rival voting behaviors, and thus effects on municipal turnout. Moreover, as our triple DDD analysis showed, these differences may also emerge within the group of the amalgamated, and not only between the merged and the non-merged municipalities. This suggests that the traditional explanations, if they hold, are not so straightforward. True, the role of both the perceived influence and of the instrumental voting choice are still in the background, but the direction of their effect is not necessarily towards a lower turnout.

Further research thus urges both on the consent/dissent issue and on further heterogeneity variables, which might account for different results among the treated group (the merged municipalities). Our contribution presents in fact some limitations. First, no inference about individual voting behavior can be drawn: our claims relate to the aggregate and not to the individual level. Secondly, specific factors of the single merger cases (e.g. historical rivalries among certain municipalities possibly affecting referenda's results and/or their turnout) cannot be addressed, or assessed, here. Moreover, we are aware of the limitations our data present: our municipal electoral data are available only for the final units and not also for each merging municipality, so that turnout variations relate to a unique post-merger turnout; then, the number of cases are quite low in absolute terms despite we covered 95% of all Italian amalgamations, meaning that some findings may result not as strong as desired (although robustness checks have been performed). Comparative studies should be carried out in order to further verify and assess the relevance of the degree of consent/dissent towards the municipal amalgamations, and its impact on the electoral turnout.

References

Alesina, A. and Spolaore, E. (1997), 'On the number and size of nations', *The Quarterly Journal of Economics* **112**(4), 1027–1056.

Alesina, A. and Spolaore, E. (2005), *The size of nations*, MIT Press, Cambridge, Mass.

Allers, M. A. and Geertsema, J. B. (2016), 'The effects of local government amalgamation on public spending, taxation, and service levels: Evidence from 15 years of municipal consolidation', *Journal of Regional Science* **56**(4), 659–682.

Allers, M., de Natris, J., Rienks, H. and de Greef, T. (2021), 'Is small beautiful? transitional and structural effects of municipal amalgamation on voter turnout in local and national elections', *Electoral Studies* **70**, 102284.
URL: *https://www.sciencedirect.com/science/article/pii/S0261379421000068*

Angelucci, D. and Paparo, A. (2019), 'Elezioni in italia–le elezioni comunali del 2019', *Italian Journal of Electoral Studies IJES-QOE* **82**(2), 191–217.

Bel, G. and Warner, M. E. (2016), 'Factors explaining inter-municipal cooperation in service delivery: a meta-regression analysis', *Journal of Economic Policy Reform* **19**(2), 91–115.

Bhatti, Y. and Hansen, K. M. (2011), 'Who 'marries' whom? the influence of societal connectedness, economic and political homogeneity, and population size on jurisdictional consolidations', *European Journal of Political Research* **50**(2), 212–238.

Bhatti, Y. and Hansen, K. M. (2019), 'Voter turnout and municipal amalgamations—evidence from denmark', *Local Government Studies* **45**(5), 697–723.

Blesse, S. and Baskaran, T. (2016), 'Do municipal mergers reduce costs? evidence from a german federal state', *Regional Science and Urban Economics* **59**, 54–74.

Blom-Hansen, J., Houlberg, K. and Serritzlew, S. (2014), 'Size, democracy, and the economic costs of running the political system', *American Journal of Political Science* **58**(4), 790–803.

Blom-Hansen, J., Houlberg, K. and Serritzlew, S. (2021), 'Jurisdiction size and local government effectiveness: Assessing the effects of municipal amalgamations on performance', *European Journal of Political Research* **60**(1), 153–174.

Blom-Hansen, J., Houlberg, K., Serritzlew, S. and Treisman, D. (2016), 'Jurisdiction size and local government policy expenditure: Assessing the effect of municipal amalgamation', *American Political Science Review* **110**(4), 812–831.

Bolgherini, S., Casula, M. and Marotta, M. (2018), 'Pursuing defragmentation at the municipal level: signs of a changing pattern?', *Modern Italy* **23**(1), 85–102.

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

Econometrics .

URL: *https://www.sciencedirect.com/science/article/pii/S0304407620303948*

Cameron, R. and Milne, C. (2013), 'Size, efficiency and local democracy in south africa: A preliminary assessment', *African Journal of Public Affairs* **6**(4).

Castater, E. G. and Han, K. J. (2018), 'Trying not to lose: The electoral consequences of unilateral reform efforts and the social pact formation process', *European Journal of Political Research* **57**(1), 171–193.

Dahl, R. A. (1994), 'A democratic dilemma: System effectiveness versus citizen participation', *Political Science Quarterly* **109**(1), 23–34.

URL: *http://www.jstor.org/stable/2151659*

Dahl, R. A. and Tufte, E. R. (1973), *Size and democracy*, Vol. 2, Stanford University Press.

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

Denters, B., Goldsmith, M., Ladner, A., Mouritzen, P. E. and Rose, L. E. (2014), *Size and local democracy*, Edward Elgar Publishing.

Di Porto, E. and Paty, S. (2018), 'Cooperation among local governments to deliver public services', *Politics & Policy* **46**(5), 790–820.

Diamond, A. and Sekhon, J. S. (2013), 'Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies', *The Review of Economics and Statistics* **95**(3), 932–945.

URL: *https://doi.org/10.1162/REST_a_00318*

Dollery, B. (2010), 'Are there tensions between democracy and efficiency in local government? a conceptual note on the structural reform debate', *Urban Policy and Research* **28**(1), 117–123.

URL: *http://dx.doi.org/10.1080/08111140903552704*

Downs, A. et al. (1957), 'An economic theory of democracy'.

Gendźwiłł, A. and Kjaer, U. (2021), 'Mind the gap, please! pinpointing the influence of municipal size on local electoral participation', *Local Government Studies* **47**(1), 11–30.

Gerber, A. S., Green, D. P. and Larimer, C. W. (2008), 'Social pressure and voter turnout: Evidence from a large-scale field experiment', *American political Science review* **102**(1), 33–48.

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

URL: *https://www.sciencedirect.com/science/article/pii/S0304407621001445*

Gordon, N. and Knight, B. (2009), 'A spatial merger estimator with an application to school district consolidation', *Journal of Public Economics* **93**(5-6), 752–765.

Hansen, S. W. (2015), 'The democratic costs of size: How increasing size affects citizen satisfaction with local government', *Political studies* **63**(2), 373–389.

Harjunen, O., Saarimaa, T. and Tukiainen, J. (2021), 'Political representation and effects of municipal mergers', *Political Science Research and Methods* **9**(1), 72–88.

Heinisch, R., Lehner, T., Mühlböck, A. and Schimpf, C. H. (2018), 'How do municipal amalgamations affect turnout in local elections? insights from the 2015 municipal reform in the austrian state of styria', *Local Government Studies* **44**(4), 465–491.

Heinisch, R. and Mühlböck, A. (2016), 'Auf die Größe kommt es an! neue empirische evidenz zur wahlbeteiligung in gemeinden', *Zeitschrift für vergleichende Politikwissenschaft* **10**(2), 165–190.

Hicks, B. M. (2006), *Are Marginalized Communities Disenfranchised?: Voter Turnout and Representation in Post-merger Toronto*, Citeseer.

Hlepas, N. (2010), 'Uncompleted greek territorial consolidation: two waves of reforms', *Territorial consolidation reforms in Europe* pp. 45–74.

Horiuchi, Y., Saito, J. and Yamada, K. (2015), 'Removing boundaries, losing connections: Electoral consequences of local government reform in japan', *Journal of East Asian Studies* **15**(1), 99–125.

Jakobsen, M. and Kjaer, U. (2016), 'Political representation and geographical bias in amalgamated local governments', *Local Government Studies* **42**(2), 208–227.

Jenny, M. and Ennser-Jedenastik, L. (2014), The power of an angry base: Examining the electoral consequences of forced municipal mergers, Tag der Politikwissenschaft, Vienna.

Kahn-Lang, A. and Lang, K. (2020), 'The promise and pitfalls of differences-in-differences: Reflections on 16 and pregnant and other applications', *Journal of Business & Economic Statistics* **38**(3), 613–620.
URL: *https://doi.org/10.1080/07350015.2018.1546591*

Keating, M. (1995), 'Size, efficiency and democracy: Consolidation, fragmentation', *Theories of urban politics* p. 117.

Kjær, U., Hjelmar, U. and Leth Olsen, A. (2010), 'Municipal amalgamations and the democratic functioning of local councils: the case of the danish 2007 structural reform', *Local government studies* **36**(4), 569–585.

Koch, P. and Rochat, P. E. (2017), 'The effects of local government consolidation on turnout: Evidence from a quasi-experiment in switzerland', *Swiss Political Science Review* **23**(3), 215–230.

Kraaykamp, G., van Dam, M. and Toonen, T. (2001), 'Institutional change and political participation: The effects of municipal amalgamation on local electoral turnout in the netherlands', *Acta Politica* **36**(4), 402–418.

Kreif, N., Grieve, R., Hangartner, D., Turner, A. J., Nikolova, S. and Sutton, M. (2016), 'Examination of the synthetic control method for evaluating health policies with multiple treated units', *Health economics* **25**(12), 1514–1528.

Kuhlmann, S. and Wollmann, H. (2019), *Introduction to comparative public administration: Administrative systems and reforms in Europe*, Edward Elgar Publishing.

Ladner, A., Keuffer, N., Baldersheim, H., Hlepas, N., Swianiewicz, P., Steyvers, K. and Navarro, C. (2019), *Patterns of local autonomy in Europe*, Springer.

Lapointe, S., Saarimaa, T. and Tukiainen, J. (2018), 'Effects of municipal mergers on voter turnout', *Local Government Studies* **44**(4), 512–530.

Lassen, D. D. and Serritzlew, S. (2011), 'Jurisdiction size and local democracy: Evidence on internal political efficacy from large-scale municipal reform', *American Political Science Review* **105**(2), 238–258.

Leuven, E. and Sianesi, B. (2003), 'PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing'.
URL: *https://ideas.repec.org/c/boc/bocode/s432001.html*

Mabuchi, M. (2001), 'Municipal amalgamation in japan', *World Bank* .

Massetti, E. (2018), 'Regional elections in italy (2012–15): Low turnout, tri-polar competition and democratic party's (multi-level) dominance', *Regional & Federal Studies* **28**(3), 325–351.

McDonnell, J. (2020), 'Municipality size, political efficacy and political participation: a systematic review', *Local Government Studies* **46**(3), 331–350.

Miyazaki, T. (2014), 'Municipal consolidation and local government behavior: evidence from japanese voting data on merger referenda', *Economics of Governance* **15**(4), 387–410.

Niemi, R. G., Craig, S. C. and Mattei, F. (1991), 'Measuring internal political efficacy in the 1988 national election study', *American Political Science Review* **85**(4), 1407–1413.

Oliver, J. E. (2000), 'City size and civic involvement in metropolitan america', *American Political Science Review* **94**(2), 361–373.

Reingewertz, Y. (2012), 'Do municipal amalgamations work? evidence from municipalities in israel', *Journal of Urban Economics* **72**(2-3), 240–251.

Riker, W. H. and Ordeshook, P. C. (1968), 'A theory of the calculus of voting', *The American Political Science Review* **62**(1), 25–42.
URL: *http://www.jstor.org/stable/1953324*

Rodrigues, M. and Tavares, A. F. (2020), 'The effects of amalgamations on voter turnout: Evidence from sub-municipal governments in portugal', *Cities* **101**, 102685.

Roesel, F. (2017), 'Do mergers of large local governments reduce expenditures?–evidence from germany using the synthetic control method', *European Journal of Political Economy* **50**, 22–36.

Ryan, A. M., Kontopantelis, E., Linden, A. and James F Burgess, J. (2019), 'Now trending: Coping with non-parallel trends in difference-in-differences analysis', *Statistical Methods in Medical Research* **28**(12), 3697–3711. PMID: 30474484.
URL: *https://doi.org/10.1177/0962280218814570*

Saarimaa, T. and Tukiainen, J. (2014), 'I don't care to belong to any club that will have me as a member: empirical analysis of municipal mergers', *Political Science Research and Methods* **2**(1), 97–117.

Solé-Ollé, A. and Bosch, N. (2005), 'On the relationship between authority size and the costs of providing local services: lessons for the design of intergovernmental transfers in spain', *Public Finance Review* **33**(3), 343–384.

Steiner, R., Kaiser, C. and Eythórsson, G. T. (2016), A comparative analysis of amalgamation reforms in selected european countries, *in* 'Local public sector reforms in times of crisis', Springer, pp. 23–42.

Steventon, A., Tunkel, S., Blunt, I. and Bardsley, M. (2013), 'Effect of telephone health coaching (birmingham ownhealth) on hospital use and associated costs: cohort study with matched controls', *Bmj* **347**.

Strandberg, K. and Lindell, M. (2020), 'Citizens' attitudes towards municipal mergers. individual-level explanations', *Scandinavian Political Studies* **43**(4), 296–316.

Strebel, M. A. (2019), 'Why voluntary municipal merger projects fail: Evidence from popular votes in switzerland', *Local Government Studies* **45**(5), 654–675.

Sun, L. and Abraham, S. (2020), 'Estimating dynamic treatment effects in event studies with heterogeneous treatment effects', *Journal of Econometrics* .

Swianiewicz, P. (2010), 'If territorial fragmentation is a problem, is amalgamation a solution? an east european perspective', *Local Government Studies* **36**(2), 183–203.
URL: *https://doi.org/10.1080/03003930903560547*

Tran, C., Kortt, M. and Dollery, B. (2019), 'Population size or population density? an empirical examination of scale economies in south australian local government, 2015/16', *Local Government Studies* **45**(5), 632–653.

Tricaud, C. (2021), 'Better alone? evidence on the costs of intermunicipal cooperation'.

Van Houwelingen, P. (2017), 'Political participation and municipal population size: A meta-study', *Local Government Studies* **43**(3), 408–428.

Voda, P. and Svačinová, P. (2020), 'To be central or peripheral? what matters for political representation in amalgamated municipalities?', *Urban Affairs Review* **56**(4), 1206–1236.

Weinbacher, G. (2018), Unterschiedliche Ebenen, unterschiedliches Wahlverhalten-politische Folgen von Kommunalstrukturreformen am Beispiel der Steiermark, PhD thesis, uniwien.

Zeedan, R. (2017), 'Bigger but not always better: Size and democracy in israeli amalgamated local governments', *Journal of Urban Affairs* **39**(5), 711–728.

A Appendix

A.1 More on Municipalities

As Table 1 shows, Italian mergers have so far mainly involved municipalities with fewer than 5,000 inhabitants: almost 40% of Italian municipalities involved in a merger process had fewer than 1,000 inhabitants, and almost 90% fewer than 5,000; the average population of a merging municipality is 2,362. The policy of municipal defragmentation has led to the emergence of new municipalities (see Table 2), of which just under 30% are still under 2,000 inhabitants, and about 60% under 5,000; the average population of a new merged municipality is 5,226, more than the double of the original municipalities. Moreover, the number of municipalities involved in a merger is quite limited: most of them (about 70%) consist of only two municipalities and only about 10% of more than three.

Table A.1: Size of Italian Municipalities involved in a merger process (1995-2021)

Population	N Municipalities	%	Cumulative %
Under 1,000	141	40.99	40.99
1,001-2,000	79	22.97	63.95
2,001-3,000	44	12.79	76.74
3,001-4,000	30	8.72	85.47
4,001-5,000	16	4.65	90.12
5,001-6,000	7	2.03	92.15
6,001-7,000	9	2.62	94.77
7,001-8,000	4	1.16	95.93
8,001-9,000	4	1.16	97.09
9,001-10,000	3	0.87	97.97
More than 10,000	7	2.03	100.00
Total	344	100.00	

Notes: Authors' elaboration of official data by Italian National Institute of Statistics (ISTAT).

Table A.2: Size of Italian Municipal Amalgamations (1995-2020)

Population	N Amalgamations	%	Cumulative %
Under 1,000	14	10.00	10.00
1,001-2,000	24	17.14	27.14
2,001-3,000	16	11.43	38.57
3,001-4,000	15	10.71	49.28
4,001-5,000	14	10.00	59.28
5,001-6,000	15	10.71	69.99
6,001-7,000	11	7.86	77.85
7,001-8,000	7	5.00	82.85
8,001-9,000	1	0.71	83.56
9,001-10,000	5	3.57	87.14
More than 10,000	18	12.86	100.00
Total	140	100.00	

Notes: Authors' elaboration official data by Italian National Institute of Statistics (ISTAT)

In table A.3 we present data on turnout for amalgamated (treated) and non-amalgamated (non-treated) observations, before and after the merger (treatment). We can identify pre- and post- period for the control group based on the fact that the amalgamation waves are fixed and non-merged municipalities went to the polls in the same years of the amalgamated ones. We excluded those municipalities that in our timespan held more than four elections (since these are mostly due to exceptional circumstances), and considered the third and fourth election as the two post-merger periods, while the first and the second elections as the pre-merger periods. This is quite a rough definition, but pretty useful for descriptive purposes. Electoral data show an overall turnout decrease in Italian elections (at all levels) detected over time (Massetti (2018); Angelucci and Paparo (2019)). In line with the overall decline in local participation, the average turnout in all municipalities - be they amalgamated or not, - in a subsequent period (indicated as post-merger) is lower. Nonetheless, amalgamated municipalities decrease less their turnout compared to the non-merged units. These data are hence in line with Fig. 2, namely that merged municipalities tend to increase (or better to decrease less in the whole period) their turnout after merging.

Table A.3: Average turnout in municipal elections (percentage points)

	All municipalities	Non-Amalgamated	Amalgamated
Overall (2001-2019)			
Average Turnout	71.2	71.5	70.2
	(0.110)	(0.113)	(0.0991)
N	3900	3034	866
Pre-merger period			
Average Turnout	75.0	75.8	73.0
	(0.102)	(0.103)	(0.0971)
N	1725	1202	523
Post-merger period			
Average Turnout	68.2	68.7	66.0
	(0.107)	(0.110)	(0.0865)
N	2175	1832	343

Notes: All Italian municipalities, amalgamated and non-amalgamated. Overall period (2001-2019), pre- and post-merger periods. Standard deviations in parentheses.

A.2 Propensity Score Matching - PSM

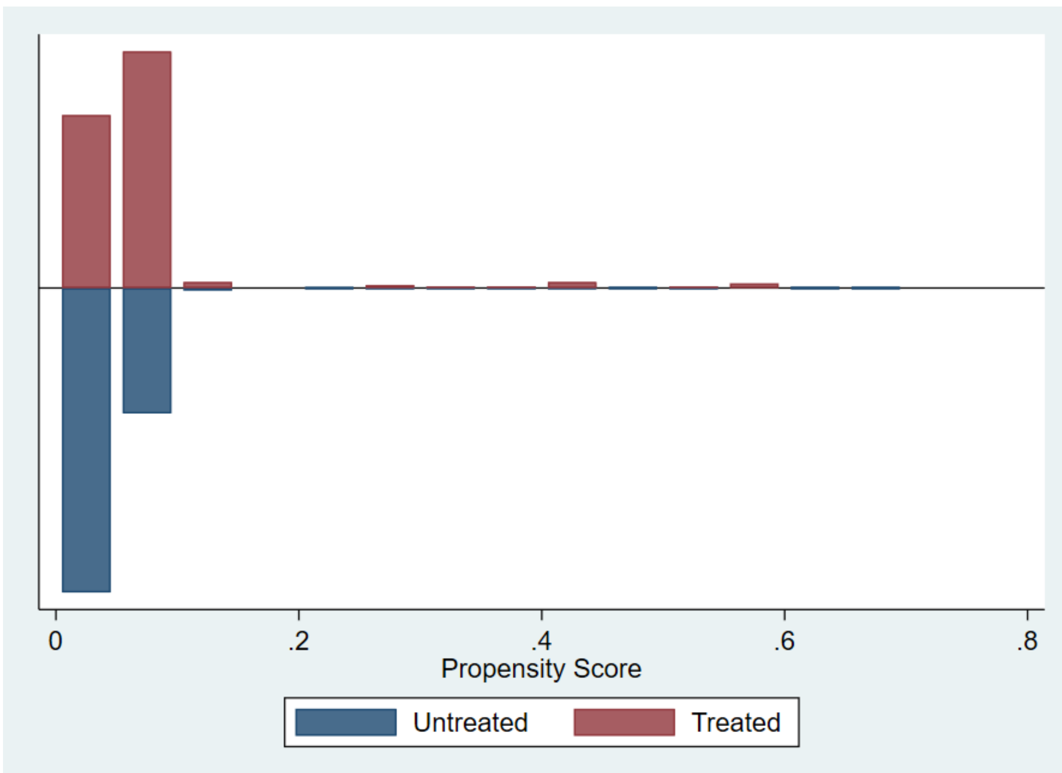
We employ a Propensity Score Matching (PSM) based on a nearest-neighbour matching algorithm to select pre-merging municipalities and create a suitable control group. This approach will generally make more plausible our DiD, making the treatment and control groups more similar in levels. First, to make more similar the voting behavior evolution of municipalities before the amalgamation period starts, data are matched on turnouts in the three big round of election: 2005, 2009, and 2013. Second, data are matched on sensible dimensions of heterogeneity among municipalities that can explain differences in voting behavior: financial autonomy and rigidity, and pre-merging population. Financial autonomy is the proportion of tax revenues over the total municipal income, that measures the capacity of a municipality to finance itself independently of transfers from the higher levels of governance. Financial rigidity is the proportion of fixed expenditures (salaries, loan refunds, and current incomes) over the total current municipal income, that explains the capacity of each municipality to direct resources when it needed for new expanses. Data for covariates have been retrieved by different sources: population from the Italian National Statistics Institute (ISTAT) database; financial data and turnout from the Department of Internal and Territorials Affairs of the Ministry of Interior. The common support condition requires that there exist treated and non-treated at any value of the propensity score. In figure A.1 there exist treated and non-treated with close propensity score values for propensity scores between .0 and .7. None of the treated are disregarded because they cannot be matched. In table A.4, we provide a balancing test before and after the matching. In column 1, we report the means of each covariate used to retrieve the propensity score. Post-matching, the means of treated and controls are more equalized. Column 2 reports the standardized percentage bias together with the achieved percentage reduction in bias. On average the bias reduction is around 83%. In Column 3, the t-tests for equality of means in the two samples are reported. Excluding the financial autonomy measure, all the other variables were strongly unbalanced before-matching, and the matching strongly reduces the bias.

Table A.4: Matching quality indicators before and after matching

		(1) Mean		(2) % reduct		(3) t-test	
Variable	Unmatched Matched	Treated	Control	%bias	|bias|	t	p>|t|
2005 Turnout	U	.13663	.08828	17.6		3.19	0.001
	M	.13663	.15538	-6.8	61.2	-0.73	0.468
2009 Turnout	U	.46092	.42032	10.6		1.74	0.082
	M	.46092	.45998	0.2	97.7	0.03	0.976
2013 Turnout	U	.03203	.00303	27.3		9.23	0.000
	M	.03203	.02741	4.4	84.0	0.40	0.687
Financial	U	73.882	72.59	5.9		0.98	0.325
Autonomy	M	73.882	73.388	2.3	61.8	0.26	0.795
Financial	U	76.065	111.25	-45.1		-6.11	0.000
rigidity	M	76.065	74.035	-2.6	94.2	0.50	0.620
Population	U	2394.5	7608.7	-16.6		-1.99	0.046
	M	2394.5	2466.6	-0.2	98.6	-0.23	0.817

Notes: Turnout in 2005, 2009, and 2013 are imputed with zero when the municipality did not have an election in one of these years. Financial autonomy is the proportion of tax revenues over the total municipal income. Financial rigidity is the proportion of fixed expenditures (salaries, loan refunds, and current incomes) over the total current municipal income. Population is the number of resident inhabitants. Data, elaborated using pstest (Leuven and Sianesi (2003))

Figure A.1: Common Support.



Notes: Propensity score distribution and common support for propensity score estimation. Treated indicates the observations in the adoption group that have a suitable comparison. Data plotted using psgraph (Leuven and Sianesi (2003))

A.3 Other approach data

A.3.1 Merged ex-post

In the bottom panel of table A.5, we report the average treatment effect for each year-group by using fictitiously merged municipalities as treated group. The results are very similar to the baseline estimates in table 2. In this specification the 95% confidence bands of the year-group-2016 contain a zero. In the top panel, the Average Treatment Effect on the Treated (ATT) is very close to the baseline estimate.

Table A.5: Baseline - Fused

Overall ATT:	Std. Error		95% Conf. Int.
0.0582	0.0068		[0.0449 , 0.0714]*
Group Effects:	**ATT**	**Std. Error**	**95% Simult. Conf. Int.**
2014	0.2535	0.0175	[0.2105 , 0.2966]*
2015	0.0662	0.0328	[-0.0145 , 0.1468]
2016	-0.0525	0.0176	[-0.0957 , -0.0094]
2017	0.1470	0.0327	[0.0666 , 0.2275*
2018	0.0024	0.0241	[-0.0568 , 0.0616]
2019	0.0012	0.0033	[-0.0070 , 0.0093]
—			

Notes: Top panel: the overall ATT displays the average treatment effects across all lengths of exposure to the treatment (across groups) following Callaway and Sant'Anna (2020). Bottom panel: the group-time effects show the average treatment effects specific to each year group. '*' = confidence band does not cover 0.

A.3.2 Not-Yet-Treated Control Group

A different parallel trend assumption uses the not-yet-treated observations. D_{it} is an indicator of whether unit i has been treated by time t:

$$E[Y_t(0) - Y_{t-1}(0)|G = g] = E[Y_t(0) - Y_{t-1}(0)|D_s = 0, G \neq g].$$

This assumption states that one can use the not-yet-treated by time s ($s \geqslant t$) units as valid comparison groups when computing the average treatment effect for the group first treated in time g. In general, this assumption uses more observations when building comparison groups than the smaller group of the never-treated. When one imposes the parallel trends assumption based on not-yet-treated units, one has that for all $t \geqslant g$:

$$ATT(g,t) = E[Y_t - Y_{g-1}|G = g] - E[Y_t - Yg - 1|D_t = 0, G \neq g].$$

In the bottom panel of table A.6, we report the average treatment effect for each year-group. The results are very similar to the baseline estimates that use the never-treated control group in table 2. In this specification the year-group-2016 95% confidence bands contain a zero. In the top panel the Average Treatment Effect on the Treated (ATT) is very close to the baseline estimate.

Table A.6: Baseline - Not-Yet-Treated Control Group

Overall ATT:	Std. Error		95% Conf. Int.
0.0587	0.0083		[0.0424 , 0.075]*
Group Effects:	**ATT**	**Std. Error**	**95% Simult. Conf. Int.**
2014	0.2444	0.0204	[0.1893 , 0.2996]*
2015	0.0610	0.0424	[-0.0536 , 0.1757]
2016	-0.0439	0.0242	[-0.1092 , 0.0215]
2017	0.1293	0.0344	[0.0363 , 0.2444]*
2018	0.0012	0.0272	[-0.0723 , 0.0747]
2019	0.0007	0.0039	[-0.0099 , 0.0113]
—			

Notes: Top panel: the overall ATT displays the average treatment effects across all lengths of exposure to the treatment (across groups) following Callaway and Sant'Anna (2020). Bottom panel: the group-time effects show the average treatment effects specific to each year group. '*' = confidence band does not cover 0.