Stronger together? Assessing the causal effect of inter-municipal cooperation on the efficiency of small Italian municipalities

(Accepted for publication by the Journal of Regional Science, August 2020)

Davide Luca*, Felix Modrego†

Abstract

When local governments are small and fragmented, promoting inter-municipal cooperation (IMC) among them is seen as a tool to improve the management of public services by reaping economies of scale and scope. Yet, the empirical evidence on the impacts of IMC on local governments' efficiency is scarce and inconclusive. In this paper we investigate the experience of Italy's municipal unions (*Unioni di comuni*). We develop an index of technical efficiency by means of Robust Data Envelopment Analysis. We then exploit Nearest-Neighbor Matching and Fuzzy Regression Discontinuity Design estimators to explore whether municipal unions have any impacts on the administrative efficiency of member municipalities. We fail to find any strong, significant effect.

Keywords: efficiency; inter-municipal cooperation; Data Envelopment Analysis; Regression Discontinuity Design; matching; Italy.

JEL Classification: D73, D74, H70, H77, R51.

* Corresponding author. Department of Land Economy, University of Cambridge, 19 Silver Street, Cambridge, UK; Middle East Centre, London School of Economics and Political Science, Houghton Street, London, UK, dl622@cam.ac.uk.

†Institute of Social Sciences, Universidad de O'Higgins, Libertador Bernardo O'Higgins 611, Rancagua, Chile; felix.modrego@uoh.cl.

1. INTRODUCTION

There is a long debate on what the optimal size of jurisdictions should be (Epple & Romer, 1989; E. Ostrom, 2010; V. Ostrom, Tiebout, & Warren, 1961). On the one hand, public choice scholars and the traditional theory on fiscal federalism suggest that the provision of public goods and services should be located at the lowest level of government which has the capacity to achieve specific objectives – an exception being a limited set of basic responsibilities prerogative of the central government (Oates, 1972, 1999; Ostrom, Tiebout, & Warren, 1961). On the other hand, however, excessive fragmentation may prevent economies of scale and scope, increase transaction costs, and give rise to problems linked to sub-optimal jurisdiction size (Ferraresi, Migali, & Rizzo, 2018; Oates, 1999), also in light of often outdated municipal boundaries (Howell-Moroney, 2008; Luca & Salone, 2013).[1]

Identifying optimal governance arrangements is particularly relevant in periods of limited resources. Across the EU, for example, the budget constraints caused by austerity following the 2008 recession have imposed strain on local administrations, for whom providing key goods and services to citizens under scarcer resources is increasingly difficult (Bel & Warner, 2015). Numerous governments around the world have identified excessive fragmentation as a problem to tackle, and promoting inter-municipal cooperation (IMC) among local administrations has been increasingly seen as a tool to improve the management of public services by reaping economies of scale and scope.

The literature mapping the diffusion of IMC, and exploring its effects on local management practices has grown significantly over the last decade (Allers & Geertsema, 2016; Allers & van Ommeren, 2016; Bel, Dijkgraaf, Fageda, & Gradus, 2010; Bel & Warner, 2015; Breuillé, Duran-Vigneron, & Samson, 2018; Di Porto, Parenti, Paty, & Abidi, 2017; Hulst & van Montfort, 2007, 2012; Ivaldi, Marinuzzi, Quintiliani, & Tortorella, 2016; Kwon & Feiock, 2010). Nevertheless, rigorous empirical analyses on the effects of IMC on local efficiency are still scarce and reach contradictory conclusions (cf. for example: Ferraresi et al., 2018; Frère, Leprince, & Paty, 2014).

The current project aims to contribute to filling this gap. It focuses on Italy's municipal unions (MU, *Unioni di comuni*)[2] and its goal is to measure their impact on the efficiency of member municipalities. We define technical efficiency as a production possibility for which it is not possible to increase any output without, *ceteris paribus*, simultaneously reducing any other output and/or increasing inputs (Koopmans, 1951).

The paper exploits Robust Data Envelopment Analysis (R-DEA) – a micro-economically-grounded, non-parametric, distribution-free, multi-input/output method based on mathematical programming – to build a measure of technical administrative efficiency for Italian municipalities. Combining municipal expenditure data with novel information on local public service provision, we are able to develop a rigorous measure of efficiency which links inputs (municipality expenditure) to a detailed set of outputs for public goods and services across 12 key policy areas. We then deploy a Nearest-Neighbor Matching (NNM) estimator to measure the link between participating in a MU and efficiency. We alternatively adopt a Fuzzy Regression Discontinuity Design (F-RDD, or FDD), exploiting the arbitrary threshold of 5,000 inhabitants (included in different legislative acts) under which local administrations have been incentivized to join an inter-municipal body. We focus on small municipalities, which not only are the target of the Italian national policy, but which may also benefit the most from inter-municipal cooperation (Ermini & Fiorillo, 2009).

The analysis adds to the growing efforts interested in identifying the effects of IMC (Allers & de Greef, 2018; Bel & Warner, 2015; Breuillé et al., 2018; Di Porto et al., 2017; Ferraresi et al., 2018; Ivaldi et al., 2016; Manestra, Messina, & Peta, 2018). Our results, which are robust to a host of sensitivity checks and alternative estimators, fail to find any strong or statistically significant effect of inter-municipal cooperation on local efficiency.

The paper is structured as follows. Section 2 briefly discusses the related literature. Section 3 presents the empirical case, providing background information on inter-municipal cooperation in Italy. Section 4 describes the data, and presents the methodology adopted to construct the municipal efficiency score, our main dependent variable in the rest of the analysis. Section 5 details the identification strategy, while section 6 presents the results. Section 7 concludes.

2. INTER-MUNICIPAL COOPERATION AND ADMINISTRATIVE EFFICIENCY

The choice of governance arrangements has significant consequences for the well-being of citizens, economic performance, and environmental outcomes (OECD, 2002, 2015). Public choice scholars and the traditional theory on fiscal federalism suggest that the provision of public goods and services should be located at the lowest level of government which has the capacity to achieve specific objectives (Oates, 1972; Ostrom et al., 1961). As Ostrom (2010) points out, polycentric governance arrangements and smaller jurisdictions may optimize welfare by: (1) matching local public goods to local preferences. Decentralizing the provision of goods and services would be particularly relevant when preference heterogeneity is high, and when public goods and services have highly localized effects; (2) allowing politicians and civil servants to monitor more easily the performance of their service provision; (3) letting citizens have a stronger say in the decision process, reducing potential “political distortions” (Besley, 2006) and “bureaucratic slack” (Luca, 2017; Niskanen, 2001).

At the same time, however, excessive territorial fragmentation may affect administrative outcomes by missing economies of scale in the delivery of public services, reducing efficient land use planning, and transport system management (Ahrend, Farchy, Kaplanis, & Lembcke, 2014); increasing bureaucratic costs for businesses (Djankov, McLiesh, & Ramalho, 2006); and leading to failures in the implementation of growth promoting strategies (Cheshire & Gordon, 1996). To this aim, governments around the world have increasingly explored tools such as the amalgamation of municipalities and inter-municipal cooperation. Theoretical contributions suggest that small municipalities – in particular those unable to provide goods and services on their own – should be the main beneficiaries of any potential increase in efficiency determined by cooperation (Bartolini & Fiorillo, 2011; Ermini & Fiorillo, 2009). The rationale behind promoting IMC is that it may be a tool to reap the benefits of returns to scale (in sectors where average costs decrease as production increase), exploit economies of density/scope (in policy areas where network effects are important, particularly when these networks go beyond the municipal jurisdictional borders), and reduce negative externalities resulting from uncoordinated actions of individual jurisdictions (Bel & Warner, 2015; Hulst & van Montfort, 2007).

Compared to municipal mergers,[3] cooperation has been identified as a potentially more attractive tool, in that it is more flexible (Feiock & Scholz, 2009; Ferraresi et al., 2018) – it avoids the frequent opposition of local governments to measures which would reduce their powers and autonomy – and it circumvents free-riding and the common-pool problems (Weingast, Shepsle, & Johnsen, 1981) repeatedly observed with municipality mergers (Harjunen, Saarimaa, & Tukiainen, 2017; Hinnerich, 2009; Jordahl & Liang, 2010). Among the different types of territorial coordination arrangements,[4] inter-municipal cooperation is a governance structure where municipalities collaborate with the goal of providing shared public goods/services without completely renouncing their decision-making powers. More precisely, we define an inter-municipal cooperation venture as "*the fulfillment of a public municipal task by an individual municipality, by two or more municipalities jointly or by a third legal entity, whereby the task fulfillment simultaneously serves at least two municipalities and the participating municipalities participate directly ('performing') or indirectly ('organizing')*" (Steiner, 2003, p. 553).

Nevertheless, despite frequent claims that local government cooperation may be an effective and efficient tool of territorial governance, until recently there has been insufficient rigorous empirical evidence on the subject, and findings are contradictory. On the one hand, recent pieces of research show that cooperation is positively associated with lower cost levels. As an example, Bel and Mur (2009) focus on the Spanish Region of Aragon, and provide preliminary OLS evidence that cooperation among small municipalities is associated with lower cost for their waste collection services. Dijkgraaf and Gradus (2013) find similar results for a panel of Dutch municipalities over the period 1998-2010. The results of these studies may however suffer from significant endogeneity concerns. Ferraresi et al. (2018) adopt quasi-experimental techniques and analyze the effects of cooperation on spending across the municipalities of the Italian Region of Emilia-Romagna. They find that being in a municipality union reduces the total per-capita current expenditure by around 5%, while not affecting service outputs in the areas of waste collection, road safety, and kindergarten supply.

On the other hand, however, other scholars find that cooperation is either associated with higher municipal expenditure, or has no significant impacts. Allers & de Greef (2018) for example analyze panel data for 2005-2013 on inter-municipal cooperation in the Netherlands to uncover that cooperation is associated with increased spending in the case of small and large local governments, while having no relationship in medium-sized municipalities. Frère, Leprince, & Paty (2014) apply spatial econometric techniques to study spending patterns of

French municipalities over the period 1994-2003, and do not find any effect of cooperation on the levels of local public spending. Analogous results are provided for the case of Italy by Ivaldi et al. (2016) and Ferraresi and Secomandi (2015).

The current paper aims to make sense of these contradictory findings by providing more novel, rigorous evidence on Italy. Furthermore, most papers published on the topic primarily focus on municipal expenditure. However, cooperation may not necessarily result in lower spending. This may be the case if large cooperation structures lead to higher agency and information costs, or simply if public administrations are reluctant to reduce spending, either for self-interest, or because they prioritize upgrading service levels to meet citizens' needs rather than reducing costs – especially when average service costs decline as a result of scale economies (Buettner & Holm-Hadulla, 2013). Hence, analyzing spending alone is not sufficient to measure whether municipal cooperation has any impacts (Allers & Geertsema, 2016). The potential effects of cooperation should be measured on indicators of administrative efficiency based on the joint analysis of policy inputs and outputs (e.g. Ivaldi et al., 2016; Rouse & Putterill, 2005). Our paper contributes to this endeavor by developing a rigorous measure of administrative efficiency based on detailed output indicators across 12 key expenditure categories. This is an advancement compared to most studies in the literature, where the lack of detailed data on policy outputs has been a key problem.

In the specific case of Italy, efforts have been recently made to assess the effects of municipal unions. Ferraresi and Secomandi (2015), and Ivaldi et al. (2016) offer preliminary insights on the potential impacts of such bodies. They respectively suggest that, contra to expectations, local administrations part of inter-municipal authorities show higher per-capita expenditures and lower quality in the provision of public services than independent municipalities. Manestra et al. (2018) argue that municipal unions are associated to reduced spending in only a limited set of spending categories, and that such effects require years to mature. Giacomini, Sancino, & Simonetto (2018) explore potential mechanisms behind the higher/lower successfulness of MU, but rely on the perception of reform adopters to measure the effects of cooperation. While these papers provide novel insights, they do not measure efficiency taking into account both policy inputs and outputs, while their identification strategies are unable to control for systematic differences across the two groups of municipalities (i.e. for potential endogenous sorting of municipalities into the "treated" and "untreated" groups). Ferraresi et al. (2018) overcome these limitations by analyzing both financial inputs and a small set of service outcomes, and by adopting a diff-in-diff Propensity-score Matching (PSM) estimator. Our paper improves their analysis in three ways. First, we measure the effect of municipal cooperation on efficiency and not only on expenditure/service outputs separately. Relatedly, we are able to measure outputs across a much broader set of policy lines. Second, their study exclusively focuses on the municipalities that are part of the Emilia-Romagna Region (they also include Tuscany in the robustness checks). While internally valid, their findings may yet have limited external validity with respect to other parts of Italy, especially considering that Emilia-Romagna (and, in a similar way, Tuscany) is considered a "virtuous" case characterized by "good governance" and effective public administrations (inter alia: Putnam, 1993). We hence cover the entirety of Italy's 15 Italian Ordinary Statute Regions.[5] Last but not least, we overcome the potential pitfalls associated with Propensity-score Matching, as underlined by recent international literature (King & Nielsen, 2016; Nielsen, Findley, Davis, Candland, & Nielson, 2011), by also adopting Mahalanobis-distance NNM and F-RDD estimators.

3. INSTITUTIONAL BACKGROUND

Italy is composed of four administrative layers: Regions, Provinces, Metropolitan Municipalities and, at the lowest level, Municipalities. The country has inherited from its pre-unitary past a highly fragmented administrative structure. It currently counts 7,954 municipalities, many of which are very small and are marked by outdated boundaries dating back to pre-1861 borders (Gambi & Merloni, 1995).[6] The median and mean population size of municipalities were, in 2017, 2,522 and 7,694 inhabitants respectively. In light of this fragmentation, and inspired by theories of functional federalism, the Italian legislator has increasingly fostered measures of municipal cooperation as a way to improve the efficiency of the country's local governments (Ermini & Fiorillo, 2009).[7]

Municipal unions were introduced in Italy three decades ago through Law 142/1990. By joining a municipal union, municipalities transfer part of their decision-making powers and financial resources in specific pre-agreed policy areas to the newly established administrative entity which, in exchange, provides the corresponding services. Italian municipal unions can be compared, among others, to Belgium's (Flanders) *Opdrachthoudende & dienstverlenende verenigingen*, France's *Sivu, Sivom,* and *Syndacats mixtes*, Germany's *Zweckverbände*, and Spain's *Mancomunidades* (Hulst & van Montfort, 2012). The new tool was developed for small municipalities with a population below 5,000 inhabitants – one local government with up to 10,000 inhabitants also being allowed to join – as a temporary, intermediate measure to prepare each member to fully merge into a single municipality within 10 years. The temporary nature of the coordination tool, combined with limited economic incentives, did not attract much interest among local administrators. As a matter of fact, by 1999 there were only 16 municipal unions in the whole country (Ivaldi et al., 2016). Both the temporary nature of the bodies and, temporarily, the population size limits were abolished by Law 265/1999. Different measures were subsequently taken to promote the new tool, e.g. providing financial incentives for municipalities (Law 42/2009). This led to a surge in the popularity of municipal unions, whose total number increased significantly throughout the 2000s (cf. Figure 1).

[Figure 1 about here]

Lastly – and importantly for part of our identification strategy – Law 122/2010 reintroduced the obligation for local governments below 5,000 inhabitants (3,000 if in mountainous areas) to start delivering public services in 6 (10 following Law 135/2012) key policy areas through municipal unions.[8] This policy shift was taken by the legislator following the post-2008 recession (Crescenzi, Luca, & Milio, 2016) with the explicit twin "extraordinary" objectives of: "urgently" developing measures to "*reduce public spending*" and "*improving efficiency*".[9] It is hence important to causally assess whether the policy – which provides a plausible source of exogenous variation in municipal participation to IMC – has achieved its intended goals.

Over the years – and especially following the Constitutional reform of 2001, which has strengthened the role of Regions as pivots of local governance –, regional governments have been endowed with significant regulatory powers in implementing the national general legislation regarding IMC, as well as with a strong role in monitoring and evaluating processes linked to municipal cooperation. Between 2011 and 2016, in particular, different regions have implemented regulations which are not always fully homogeneous (Manestra et al., 2018). There is therefore regional policy heterogeneity over regulation and the specific types of incentives adopted to foster IMC (see Manestra et al., 2018, for a detailed discussion of regional differences). This implies that in the empirical analysis we will have to take into account this regional heterogeneity).[10] Overall, however, almost all Ordinary-statute Regions have legislated to comply with the national government requirements (ibid.). Similarly, while the nature and amounts vary across regions, in most cases regional governments have set up incentives towards unionization, mostly in the form of direct financial contributions (ibid.).

MUs' revenues mainly come from transfers from member municipalities (whose budgets include the transfers to the union, if part of one), while other sources are financial incentives from central and regional governments.[11] In 2013, the total expenditure of municipal unions accounted to about 0.3% of the total expenditure of Italian local governments, that is, 970 million out of 334 billion of euro (Ferraresi et al., 2018). While this percentage underestimates the real expenditure because municipalities do not frequently write off the amounts "allocated" to the MUs, their budget is overall small. Nevertheless, IMC could contribute to increasing spending efficiency not only through direct budget effects, but also indirectly through collective learning and knowledge transfers. Most commonly, the functions transferred to municipal unions include management and administration (management of personnel, recruiting, training, etc.), municipal police (road safety, security, application of municipal regulation, civil protection, etc.), education (especially kindergartens and child care), transport services and planning (urban planning, management of public works plans), and social welfare (social inclusion measures, support for vulnerable groups, health structures).

While the net number of MUs has steadily increased over the years, there has also been considerate membership churn (cf. Manestra et al., 2018). In fact, municipalities can both join and leave a union, for example because their union cease to exist, or to join another one. While it is impossible know the exact number of existing unions since there is no official updated registry, it is estimated that, in 2015, across the whole country, 36.9% of municipalities were part of one of the existing 524 municipal unions (Ivaldi et al., 2016). Yet, participation varied significantly across regions (cf. Figure 2). Excluding Special Statute Regions (which will not be part of the empirical analysis), participation ranged from almost 87% in Emilia-Romagna, close to 55% in Tuscany and Umbria, to less than 13% in Calabria and Basilicata (ibid.). Among small municipalities (i.e. those below 5,000 inhabitants) participation fluctuated from almost 90% and 70% in Emilia-Romagna and Tuscany respectively, to less than 15% in Umbria, Basilicata, and Calabria (ibid.). Overall, in 2015 18.5% of Italian citizens lived in municipalities part of a MU.

[Figure 2 about here]

4. DATA SOURCES AND THE EFFICIENCY SCORE

In this section, we present the data and explain the methods we employ to develop our efficiency score.

4.1. Data sources

We combine secondary statistical data on Italian municipalities from two sources. Data on municipality inputs (expenditure) and outputs (services provided) for the year 2013 comes from Opencivitas,[12] an online portal developed by SOSE (a publicly listed firm owned by Italy Italy's Ministry of Economy and Finance and Bank of Italy). The dataset covers most municipalities part of Italy's 15 Ordinary Statute Regions (i.e. it excludes the five Special Statute Regions). Input data covers current expenditure, that is, spending on personnel, purchase of goods and services (including the outsourcing of services such as waste collection to private companies), passive interests incurred in the purchase/maintenance of fixed capital investments, current transfers and depreciations. Most of the services provided by unions are financed by the transfers made by municipalities. MUs also receive a small share of additional resources from regional governments and the central state. These are pro-rated and added to each municipality, so that the expenditure of each local administration reflects all the running costs incurred in the provision of services to citizens.[13]

SOSE also provides detailed indicators measuring the level of goods and services delivered by each municipality across 12 key policy areas (as of Law 42/2009). These categories are tax collection, civil registry, technical office, other general management services, local police, public schools, road maintenance, local public transports, territorial planning and environmental protection, waste management, welfare support, and public kindergartens. These outputs are collected by SOSE through questionnaires administered to municipalities, as foreseen by Legislative Decree 216/2010.[14] Finally, all other socio-economic controls come from Italy's National Institute of Statistics (ISTAT). Appendix A.1 summarizes the variables, while Appendix A.2 provides key summary statistics.

4.2. The municipal efficiency score

Theoretical considerations about how to measure technical efficiency have been a subject of economic research for many decades (c.f. Ruggiero, 2000). We calculate the spending efficiency of each municipality using Data Envelopment Analysis (DEA) (Banker, Charnes, & Cooper, 1984; Charnes, Cooper, & Rhodes, 1978; Farrell, 1957). DEA is a micro-economically-grounded, non-parametrical, distribution-free and multi-input/multi-output method based on mathematical programming, developed for the analysis of relative technical efficiency of decision-making units (DMU). The intuition behind DEA is to empirically build a pricewise-linear production frontier by finding optimal linear combinations of the different DMUs (in this case municipalities) and comparing these composite or "virtual" DMUs to each real DMU in the sample. An observed DMU producing less outputs with the same amount of inputs (output-oriented DEA) or the same outputs using more inputs (input-oriented DEA) is said to be "dominated" and lies at the interior of the empirical efficiency frontier. Making these comparisons for all the DMUs in the sample, the efficiency frontier enveloping the data is built and radial technical efficiency scores ranging from 0 to 1 are calculated. The score represents the relative distance of each DMU to the efficiency frontier, where the DMUs at the frontier have an efficiency score of 1. More intuitively, the score can also be interpreted as the proportional increase in outputs that the DMU could reach with its observed level of inputs (output-oriented) or the proportional reduction in inputs that the DMU could achieve to get the same outputs (input-oriented) had it achieved the maximal efficiency level in the sample.[15]

Among the many variants of the DEA model, we use the input-oriented, variable-returns-to-scale version (I-O VRS).[16] Formally, it solves independently for each DMU i = 1,..., I, the following linear programming problem:

$$min_{\theta,\lambda}\,\theta, \tag{1}$$

Subject to:

$$\theta x_i - X\lambda_i - s^I = 0, \tag{2}$$

$$y_i - Y\lambda_i - s^0 = 0, \tag{3}$$

$$\textstyle\sum_i \lambda_i = 1, \tag{4}$$

$$\lambda_i \geq 0;\ s^I \geq 0;\ s^0 \geq 0, \tag{5}$$

where θ is the efficiency score, x and y are, respectively, a vector of K inputs and M outputs for DMU i, s^I and s^O are vectors of input and output slacks, λ is a vector of weights and X and Y are $I * K$ and $I * M$ matrix of inputs and outputs of all DMUs respectively.

DEA has a series of advantages that makes it useful to analyze municipal spending efficiency. First, it accommodates well to the multi-output structure of our data (see below), reducing the dimensionality of the problem into a single efficiency score. Second, not being based on regression analysis, it is flexible in the sense that it does not impose any *ad-hoc* functional form to the production function, nor does it make any assumption about an error structure.

At the same time, “standard” DEA is quite sensitive to outliers and ignores stochastic phenomena like measurement errors, impeding statistical inference and yielding a downward-biased estimation of the “true” production frontier (and therefore an upward-biased metric of the true technical efficiency of the DMU’s) (Simar and Wilson, 1998, 2000). Therefore, we use the *Robust* Data Envelopment Analysis (R-DEA) method by Simar and Wilson (1998; 2002; 2011), which was developed as a solution to the drawbacks of the standard DEA. In a nutshell, R-DEA bootstraps the DEA mathematical programming problem under the assumption that the empirical distribution of the difference between the standard and the robust efficiency indices mirrors that of the difference between the “true” and the estimated scores (Simar & Wilson, 1998).

The input variable is the per-capita municipal spending in general social services in each municipality for 2013, taken from SOSE’s Opencivitas database. With respect to the output variables, unlike previous analyses in Italy, we use measured levels of provision for municipal services, one for each of the 12 spending categories that make up the municipal expenditure in general social services, as described in Section 4.1.[17] Indicators for each axis were selected based on their relevance, but also based on their completeness. We calculate the robust (bias-corrected) spending efficiency scores using 100 bootstrap replications, and obtain the scores´

confidence intervals[18] (Simar & Wilson, 2000, 2011). R-DEA was implemented in the statistical software *R* using the rDEA package (Simm & Besstremyannaya, 2016).[19]

Figure 3 summarizes the distribution of the robust efficiency score for 5,644 Italian municipalities with complete information in the Opencivitas database. The average efficiency of Italian municipalities is 85.6% and the median 86.8%. The maximum efficiency in the sample is 99.0% and the minimum 24.7%. The interquartile range is 9.3 percentage points. The efficiency score of the first municipality in the fourth quartile (90.9%) lays outside the 95% confidence interval of the first in the third one ([85.4 – 88.3%]). The score of the first municipality in the third quartile (86.8%) is larger than the upper limit of the confidence interval of the first municipality in the second quartile (83.4%), and the first in the second quartile (81.6%) is also larger than the upper limit of the smallest score (27,6%). Still, around half of the municipalities have an efficiency score between 80 and 90% and 76% between 80 and 95%.

[Figure 3 about here]

Although not fully comparable to Ivaldi et al.'s indicator (2016) due to methodological differences,[20] our efficiency score reveals less mean distance to the efficiency frontier (in their study the average efficiency is 66%) and less dispersion. DEA may suffer from a limited discriminatory power when applied to too many inputs and/or outputs (a problem of “overspecialization” of DMUs). Yet, this should not be the case here, as we resort to a very large number of observations (5,644) (see Podinovski & Thanassoulis, 2007) which is way beyond, for instance, the rule of thumb proposed by Dyson et al. (2001). Instead, we believe that our criterion of focusing on the 12 key municipal spending categories reflects a right balance between the potential risk of losses in discriminatory power and a comprehensive depiction of the municipal public provision technology.

[Figure 4 about here]

Figure 4 depicts the spatial distribution of the estimated robust efficiency scores. Consistent with abundant characterizations of regional socioeconomic and political differences in the country, the figure portraits a quite sharp North-South divide, with more efficient municipalities clustered in richer northern regions (particularly in Lombardy, Veneto, and Emilia-Romagna). Among the 100 most efficient municipalities in the sample according to the robust efficiency score, 39 are in Lombardy and 32 in Veneto. On the contrary, among the 100 less efficient, 72 are in the southern regions of Calabria (32), Basilicata (20) and Molise (20). Yet, clusters of highly efficient neighboring municipalities can also be found in some areas of the south, particularly in Puglia and Campania. Our results are relatively consistent with Ivaldi et al. (2016) who report Veneto as the Region with the greatest mean municipal efficiency followed by Lombardy, whereas Molise and Basilicata are both among those with the lowest average. (In our case, however, Puglia fares among the least efficient regions, not among the best performers.) Regarding a hypothetical relationship between population and municipal efficiency, although positive and statistically significant (at the 5% confidence level), the simple correlation between the two variables is only 6%.[21]

Finally, Appendix A.3 shows the pairwise correlation between our efficiency measure and an alternative statistic provided by SOSE and calculated as the ratio of two indicators. These are the % difference between municipal total expenditure and theoretical needs (a measure calculated through regression analysis considering local socioeconomic characteristics), and the % difference between the level of services offered and the level of services theoretically needed. We believe that our measure is more valid on two main grounds. First, the two indicators by SOSE are calculated via regression analysis, hence imposing a functional form while also making strong assumptions about the error structure. Second, the two indicators in Opencivitas are potentially endogenous if one aims at understanding the determinants of efficiency. As an example, the measure of municipal expenditure's "theoretical needs" is calculated considering local socioeconomic characteristics as well as municipal managerial practices which, in our view, should be considered as a determinant of efficiency rather than an outcome.

5. EMPIRICAL APPROACH

5.1. Baseline model

Our goal is to estimate the effect of MUs on the administrative efficiency of Italian municipalities. To this aim we define the following model:

$$ES_i = \beta_1 M_i + \Sigma_k^K \beta_2 X_i + \alpha_r + \varepsilon_i, \tag{6}$$

where:

ES_i is the efficiency score for municipality i. M_i, a dummy variable equal to 1 if the municipality i is part of a municipality union and 0 otherwise. $\Sigma_k^K X_i$ is a vector of covariates which may affect the efficiency of local administrations. In particular, we control for three main groups of variables: (1) geographical attributes, namely altitude and population density; (2) economic structure, namely unemployment, and employment in agriculture and secondary sectors; (3) social structure, namely share of people with university degrees, shares of youngsters and elders, share of non-natives residents, and tax income base; (4) institutional conditions, namely a dummy equal to one for municipalities which went through local elections between 2010 and 2013, and an index of local mafia presence developed by Crime&tech.[22] α_r are regional fixed-effects, included to control for regional idiosyncrasies, particularly considering how Regions can significantly influence municipal unionisation through regulations and incentives (cf. Casula, 2014). Finally, ε_i is the error term.

5.2. Matching estimator

The concern with regressing our indicator of efficiency on MU membership is that the latter may be correlated with spurious factors that also affect efficiency, while there may be reverse causality between efficiency and unionization, since more efficient municipalities may have better complied with the national legal requirements after 2010. We attempt to address these issues by first adopting a Matching estimator. Matching estimators reweight observational data as a way to achieve experimental-like balanced samples. They are nonparametric, since they do not impose any functional form to either the treatment or the outcome. King & Nielsen (2016) stress the pitfalls of Propensity-Score Matching pointing out that, compared to other matching methods, PSM increases imbalance, inefficiency, model dependence, and bias. We hence adopt a Nearest-Neighbor Matching estimator. In a first stage, we calculate the "Mahalanobis distance" between pairs of observations drawing on the set of covariates discussed in section 5.1, and match each treated municipality to comparable control observations that are closest to it. Similarity between units is based on a weighted function of the covariates for each observation. Weights are based on the inverse of the variance-covariance matrix of the covariates. We apply a bias-correction to continuous variables (Abadie & Imbens, 2011), and a bias-adjustment term to all covariates to correct for large-sample bias (Drukker, 2016). We then compare treated municipalities part of a MU with untreated ones, where the treatment effect is computed as the average of the difference between the observed and imputed potential outcomes for each unit.

5.3. Fuzzy-RDD estimator

Matching estimators are yet unable to account for unobservable factors that may be correlated with both participating in a MU and the error term. To account for unobserved differences, we implement a Fuzzy Regression Discontinuity Design exploiting the population cut-off of 5,000 inhabitants discussed in section 3. The key identifying assumption of the RDD strategy is that local governments which are close to the arbitrary cut-off are similar on all unobserved characteristics which may be correlated with the dependent variable. A consequence is that, if individuals are unable to exactly manipulate the assignment (the running variable), the variation in the treatment close to the threshold is randomly distributed similarly to a randomized experiment (Lee & Lemieux, 2010). We will provide below a battery of tests of this assumption.

We adopt a fuzzy specification since not all municipalities which, according to Law 122/2010 should have joined a MU have done so. Besides, numerous municipalities complied with Law 122/2020 by joining a milder form of IMC called Inter-municipal Agreement (*Convenzione*) (cf. Bartolini & Fiorillo, 2011, for a theoretical discussion).[23] It is worth remembering that F-RDD does not require the predictor x_i to perfectly or monotonically predict treatment assignment D_i, but rests on the milder assumption that there is a (possibly small) shift in the probability of assignment to the treatment at the threshold c:

$$Pr\ (D_i = 1|x_i) = \begin{cases} g_1\ (x_i)\ if\ x_i\ \leq c \\ g_0\ (x_i)\ \ if\ x_i\ > c \end{cases}, \tag{7}$$

Where $g_1\ (x_i 1)\ \neq\ g_0\ (x_i)$.

Specifically, we will analyze municipalities which joined a union before 2013, that is, the year for which the data on policy outputs used to construct the efficiency score is available. We discard municipalities which had joined MUs but left them before 2013. We estimate the following specification:

$$ES_i = \beta_0 + \beta_{LATE}\, M_i + \beta_2 f(pop_i - c) + M_i \beta_3 f(pop_i - c) + \Sigma_k^K \beta_4 X_i + \alpha_r + \varepsilon_i, \qquad (8)$$

where M_i, a dummy variable equal to 1 if the municipality i is part of a municipality union in 2013 and 0 otherwise, is instrumented using D_i, a dummy variable indicating whether each municipality should by law have joined a MU. The variable is defined as follows:

$$D_i = \begin{cases} 1\ if\ pop_i \ \leq 5{,}000 \\ 0\ if\ pop_i \ > 5{,}000 \end{cases}. \qquad (9)$$

In the first stage of the 2SLS implementation of the F-RDD, we estimate:

$$M_i = \gamma_0 + \beta_{ITT}\, D_{i,f} + \Sigma_k^K \beta_4 X_i + \alpha_r + \varepsilon_i. \qquad (10)$$

In equation (8), β_{LATE} identifies the effect of being part of a municipality union on the municipal efficiency of small municipalities. The equation allows the regression function to differ on both sides of the cut-off point c by including interaction terms between $(pop_i - c)$ and M_i. We also examine the robustness of different functional forms of f. In particular, when $f = 1$, equation (7) corresponds to a local linear regression, in which observations close to the discontinuity selected according to a specific bandwidth are used. Alternatively, we estimate equation (7) using higher-order polynomials, which can help absorb the variation of the observations which are far away from the discontinuity. Following Gelman and Imbens' (2017) recommendations to avoid high-order polynomials, we focus the discussion on the results of the second order polynomial functions. $\Sigma_k^K X_i$ is our vector of covariates. While municipal controls and regional fixed-effects are not necessary for identification, they improve the efficiency of the estimation. Standard errors are clustered again at the NUTS2 regional level.

In an RDD setting the choice of the bandwidth plays a key role and involves striking an optimal balance between precision and bias. Lower bandwidths reduce the bias in estimates but may increase standard errors because of the lower number of observations considered. By contrast, larger bandwidths reduce variance (hence standard errors), but may also increase bias in point estimates. We follow the data-driven methods proposed by Calonico, Cattaneo, and Titiunik (2014b) based on asymptotic Mean Squared Error (MSE) minimization.

6. RESULTS

6.1. Matching results

Table 1 presents the NNM outputs for the Average Treatment Effects on the Treated (ATTs).[24] As it can be seen, there is limited evidence of any effects of IMC on municipal efficiency. Model one shows a positive and statistically significant effect. The magnitude, however, is negligible. Being part of a MU is linked to an increase in efficiency of less than half a percentage point, that is, less than one tenth of a standard deviation in the efficiency score. In model two we then replicate the NNM analysis replacing the efficiency measure with total per-capita municipal expenditure. On average, unionized municipalities show higher levels of expenditure compared to their matched controls. At least in part, the coefficient may capture the negative short-term effects of IMC on expenditure also uncovered by Manestra et al. (2018). Results not presented but available on request indeed show that when running the analysis only on the municipalities which have unionized before 2010 (which may have hence had more time to reap potential positive effects, the coefficient lowers to 0.038 (significant at the 1% confidence level). Nevertheless, the overall evidence supporting any strong effects of IMC on municipal efficiency remains weak.[25]

[Table 1 about here]

Treatment-effects models reweight observational data to achieve experimental-like balanced samples. A key assumption to check if the reweighting is successful is hence to test whether the weighted distribution of each covariate is the same across “treatment” and “control” groups. Appendix A.4 reports, for specifications one of Table 1, the model-adjusted differences in means (standardized differences) in the treatment groups as well as the ratio of variances. A perfectly balanced covariate has a standardized difference of zero and variance ratio of one. Outputs provide evidence of a better balance in the matched samples compared to the unmatched ones.

6.2. Fuzzy-RDD results

This section presents the Fuzzy-RDD results. Appendix A.5 shows the main RD plot, using binned local averages and our benchmark second-order polynomial specification. As it can be seen, smaller municipalities below 5,000 inhabitants have on average lower efficiency scores than larger ones. This is in line with theoretical contributions, which suggest that small municipalities should be the main beneficiaries of any potential increase in efficiency determined by cooperation (Bartolini & Fiorillo, 2011; Ermini & Fiorillo, 2009). Nevertheless, there is limited evidence of any discontinuity at the threshold of 5,000 inhabitants.

[Table 2 about here]

Table 2 presents the Fuzzy-RDD regression outputs, showing conventional, bias-corrected, and robust estimates – which are our preferred ones being the most conservative (Calonico et al., 2014b). The middle part of the table presents the first-stage results. The bottom of the table also reports the order of the polynomial, the bandwidth size as well as the effective number of observations included by the optimal-bandwidth selector on each side of the cut-off, and the first-stage F test. We report the Montiel and Pflueger F statistic (rather than non-robust or Kleibergen-Paap Fs), as suggested by Andrews and Stock (2018). Compared to column one, column two controls for covariates. As expected, being above the population cut-off significantly reduces the likelihood of being part of a MU: the first-stage coefficients are negative and statistically significant at the 5% confidence level. The coefficients from the second stage are positive and point to a plausible positive effect of unionization on efficiency comprised between 3 and 6.7% points. The results, however, are statistically insignificant. Column three restricts the analysis to the municipalities of North and Central Italy which, on average, better complied with the legal requirement to unionize compared to those in the Southern part of the country. Indeed, the first-stage coefficients become stronger and more precise. Results point to a positive effect which, however, is again statistically insignificant. Findings confirm the exploratory results of Ivaldi et al. (2016), and are in line with other international assessments. As an example, Frère et al. (2014) do not find any statistically significant effect of municipal cooperation on spending for French municipalities. Relatedly, Breuillé et al. (2018) find that joining an IMC body leads to higher local tax rates, contradicting the frequent claim of tax cuts induced by expected economies of scale. The results are also in line with the analysis carried out on Italy by Manestra et al. (2018) who find a short-term negative effect of IMC.

A caveat must be stressed here. It must in fact be remembered that our RDD analysis captures the *short-term* impacts of IMC. The potential positive effects of MUs may yet materialize over a number of years. For example, there may be significant costs linked to setting up new organizational structures – e.g. administrative systems may have to be integrated, new offices may be needed, and regulations and administrative practices may require harmonization (Allers & Geertsema, 2016). Further rigorous research on the medium-long term impacts of unions on administrative efficiency is part of our future research agenda. Besides, our F-RDD estimator allows the identification of a local average treatment effect (LATE) for all Italian municipalities around the cut-off of 5,000 inhabitants. Yet, like any LATE it has strong internal validity but may not be representative of all municipalities. Under such light, our results may not be inconsistent with those of Ferraresi et al. (2018), who may have instead uncovered a positive *local* average treatment effect of IMC on a subset of municipalities (in Emilia-Romagna) intrinsically more efficient than many other Italian counterparts (cf. Figure 4).

6.3. Fuzzy-RDD: sensitivity checks

It is common in RDD settings to test for the exogeneity of the running variable by plotting its density $d(.)$ around the cut-off point c to ensure no manipulation or non-random sorting (McCrary, 2008). We follow Cattaneo, Jansson, & Ma (2018), who provide a set of tests based on a local-polynomial density estimator. Eggers, Freier, Grembi, & Nannicini (2018) argue that strategic sorting in municipal population figures is common across countries, and show that it was strong in Italy during the period 1960-2001. Our result, presented in Appendix A.6, reassuringly suggests that there is no statistically significant evidence of manipulative sorting around such threshold using more recent population data from the 2011 census.[26]

Another of the key assumptions of the RDD is that municipalities close to either side of the threshold predicting participation have similar characteristics. We test for continuity of all our baseline covariates around the threshold of 5,000 inhabitants. A discontinuity would indicate a violation of the underlying assumption that predicts local random assignment. Appendix A.7 presents the results. As it can be observed, across Italy there are no statistically significant differences between municipalities below/above the treatment threshold. (In the reduced sample for North and Centre Italy, an exception is population density, significant at the 5% level and, partly, industry, for which the random assignment hypothesis is rejected but only at the 10% confidence level.)

We then explore the sensitivity of the baseline estimates to alternative bandwidths and a range of orders to the polynomial. Appendix A.8 presents the results adopting three different data-driven optimal-bandwidth selection methods (cf. Calonico et al., 2014), both for the full Italian sample and for the municipalities of Northern and Central Italy. Columns one and four report the main results from Table 2, based on one MSE-optimal bandwidth selector common to each side of the cut-off. Columns two and five adopt a less restricting selector, and allow for different bandwidths on each side of the cut-off. Columns three and six allow for different bandwidths on either side, and are based on Coverage Error Rate (CER), rather than MSE, optimal selectors. Overall, the results are very similar to those of Table 2.

Appendix A.9 measures the sensitivity of the baseline results to alternative polynomial specifications. We test orders one, two, and three. In the case of North and Central Italy, the specification with third-order polynomials provides positive and significant second-stage coefficients. Following Gelman and Imbens (2017), who recommend avoiding high-order polynomials since they can lead to confidence intervals that fail to include zero with higher probability, we yet interpret this result with care.

Finally, in Appendix A.10 we run tests at placebo thresholds, to see whether the 5,000-inhabitant cut-off is truly driving the first-stage likelihood of being part of a MU. As expected, results find no effects at any other thresholds.

6.4. Addressing multiple treatments

The 5.000-population threshold adopted in the F-RDD analysis may be inappropriate to identify the specific effect of MUs: the same cut-off is used to determine other policies, making it difficult to interpret the RDD coefficients as a measure of the effect of one particular policy (Eggers et al., 2018). As a matter of fact, Grembi, Nannicini, and Troiano (2016) show that in Italy the threshold informs two main other policy changes.

First, the salary of the major and their executive officers increases sharply in municipalities above 5,000 inhabitants. Higher wages may attract more educated and capable individuals into politics, improving their performance once elected (Gagliarducci & Nannicini, 2013). To disentangle any potential efficiency effect driven by wages from that of IMC, we adopt a cross-sectional Difference-in-Discontinuity design (Eggers et al., 2018), and compare RDD effects measured at different thresholds in order to “difference out” the potential impact of the confounding policy. We exploit the alternative 1,000- and 3,000-inhabitants’ cut-offs – around which wages also change – and run sharp-RDD regressions. A significant effect of wages on efficiency at these alternative cut-offs would be evidence of a potential confounding impact also at the 5,000 threshold – and, hence, would require to “difference it out” from the estimation of the IMC coefficients (Eggers et al., 2018). The results, presented in the first two columns of appendix A.11, fail to find any significant impact.

Second, the 5,000-resident threshold also affects the ability of municipalities to accumulate debt, as local governments below the cut-off are exempted from state-imposed fiscal rules (Grembi et al., 2016). While there are no other thresholds to exploit for a Diff-in-Disc strategy, an indirect way to see if fiscal rules may confound our main results is to run a sharp RDD analysis using, as a dependent variable, per-capita expenditure, the variable through which borrowing/deficit restrictions would potentially impact on our measure of efficiency. The results, presented in column 3 of appendix A.11, find as expected a negative effect on expenditure for municipalities just above the 5,000 threshold. The effect is in any case hardly significant, suggesting that fiscal rules should not be a concern for our identification.

6.5. Heterogeneity of impacts

Our baseline analysis measures the average effect of municipal cooperation on an overall measure of efficiency calculated by pooling together output indicators across all 12 expenditure categories. Yet, different types of public goods and services may be differently subject to economies of scale and scope. We hence explore whether the effect of municipal cooperation on administrative efficiency depends on the type of public good/service provided. Parks et al. (1981) put forward the concept of “co-production”, i.e. the idea that, compared for example to the provision of physical infrastructure, the optimal production of some public services requires the active involvement of citizens. We split the 12 policy areas into two groups, distinguishing between those involving significant face-to-face delivery (e.g. policing, education, social welfare, etc.), and those where the active involvement of citizens is less relevant (e.g. transportation, infrastructures, etc.). We expect that the former may benefit less from economies of scale than the latter, as it has been shown that policy areas involving co-production are more effective in smaller communities while they are more likely to suffer from diseconomies of scale (Ostrom, 2010; Ostrom et al., 1961). We hence re-calculate our Robust-DEA measure of administrative efficiency for each of the sub-groups.[27] Table 3 presents the results. Again, we do not find any statistically significant effects of IMC.

[Table 3 about here]

Second, the size of municipal cooperation bodies may affect their efficiency (Ermini & Fiorillo, 2009). We divide the sample of municipalities into two groups, distinguishing between those which joined a municipal union with a number of members above / below the national average level. We then re-run the NNM estimator considering, as "treated", each group separately. Results, presented in columns 1 and 2 of Appendix A.12, show a moderate but significantly positive effect on efficiency among larger MUs.[28] While there are not enough observations below 5,000 inhabitants to runt the F-RDD analysis on large MUs, and hence care is needed in interpreting the matching results in a causal way, the outputs may indicate an area for future research where municipal unionization may have determined an efficiency dividend. Similarly, in column 3 of Appendix A.12 we estimate the NNM estimator restricting the sample to municipalities part of a MU where at least one member has a population above the Italian average. We do so since efficiency gains may be conditional on small municipalities cooperating with larger ones. Results preliminarily suggest that this may be the case.[29]

7. CONCLUSION

While it has been frequently posited that cooperation among local governments can be used as a tool to improve administrative efficiency, rigorous empirical evidence is still scarce. Drawing on the experience of Italy's *municipal unions*, this paper explores whether joining an inter-municipal cooperation arrangement has any impacts on the efficiency of small municipalities. We first adopt a Nearest-Neighbor Matching estimator, and then exploit an arbitrary threshold set by the national legislator to develop a Fuzzy Regression Discontinuity Design. Our results, which are robust to a host of sensitivity checks, do not however find any strong statistically significant effect of IMC on local efficiency.

The analysis contributes to the literature in two main ways. First, we join recent efforts to rigorously assess the potential impacts of IMC by adopting quasi-experimental tools and applying them, for the first time, to almost 5,000 Italian municipalities for which data is available. Second, we advance on previous contributions by providing a rigorous analysis of spending efficiency measured as the relationship between policy inputs (expenditure) and outputs (amount of goods and services provided). In particular, we combine data on local expenditure with a novel dataset on municipal outputs for public goods and services across 12 key policy areas to build an indicator of administrative efficiency by means of Robust Data Envelopment Analysis. As suggested by Ermini and Fiorillo (2009), joining a MU implies costs for member municipalities, which frequently decide to enter into a municipal cooperation body as a way to increase the number of services/goods provided rather than as a way to cut expenditure. Hence, any analysis on the efficiency impacts of IMC should jointly consider policy inputs (expenditure) and outputs, since any potential effects may derive from lower costs per output produced rather than lower overall expenditure.

The analysis has implications for scholars and policymakers interested in whether IMC can be an effective tool to maximize the efficiency of local administrations. Findings confirm the exploratory results of Ivaldi et al. (2016), and are in line with other international assessments. As an example, focusing on France, Frère et al. (2014) reach conclusions similar to ours. Relatedly, Breuillé et al. (2018) find that in France joining an IMC body leads to higher local tax rates, contradicting the frequent claim of tax cuts induced by expected economies of scale.

We can speculate on two potential explanations behind the failure to uncover any strong "IMC dividend". The first, most straightforward one is that the theories foreseeing an "efficiency dividend" of inter-municipal cooperation may be flawed or incomplete. For example, as suggested by the literature (Bel & Warner, 2015; Feiock & Scholz, 2009; Sorensen, 2007), the (potential) benefits associated with cooperation may be offset – or even outweighed – by potential diseconomies and inefficiencies linked to high transaction costs and institutional collective action problems arising during cooperation efforts. A second, complementary hypothesis is that the Italian case – as well as the others analyzed by the literature – may have suffered from implementation failures. These factors may reconcile our findings with the outputs of Ferraresi et al. (2018), who uncover a positive effect for the Italian municipalities part of Emilia-Romagna and Tuscany. These two regions are frequently considered as particularly "virtuous" cases characterized by high administrative capacity and strong local institutions, and may hence suggest how, to be effective, IMC measures require careful design and a strong local institutional capacity.

Two caveats must be underlined. First, our analysis is unable to distinguish between municipal unions only involving a service delivery agreement and those determining the creation of a service delivery organization – that is, a second-level organization independent of each municipality and aimed at delivering services on behalf of each member (Hulst & van Montfort, 2012). As suggested by Giacomini et al. (2018), these institutional differences may yet play a role in producing any (potential) advantage. Second, and most importantly, the potential "efficiency dividend" of inter-municipal cooperation may require time to come about (Ermini & Fiorillo, 2009). When data will become available, future research should be devoted to replicating our efficiency analysis to explore any potential dynamic effects of cooperation over the medium-long term, for example by adopting longitudinal difference-in-discontinuity research designs.

---

[1] There may hence be a potential trade-off between on the one hand allowing local differentiation and on the other internalizing spill-overs as well as letting economies of scale/scope develop (Oates, 1999).

[2] As of 2015, more than 11 million, that is 18.5%, of Italian inhabitants lived in municipalities part of a Union. 79.3% of them are small municipalities with a population below 5,000 (Ivaldi et al., 2016).

[3] The literature on the effects of municipal mergers is quite developed, and frequently finds negative or null impacts on local governments' costs and efficiency (inter alia: Allers & van Ommeren, 2016; Blesse & Baskaran, 2016; Hirota & Yunoue, 2017; Lima & Silveira Neto, 2018; Reingewertz, 2012).

[4] The OECD (2015) for example identifies four arrangements: (1) ad-hoc, soft coordination bodies; (2) inter-municipal authorities; (3) supra-municipal entities; (4) and metropolitan cities. Relatedly, Hulst & van Montfort (2012) identify the following types of co-operation: (1) quasi-regional governments; (2) planning fora; (3) service delivery agreements; (4) service delivery organizations. Italy's municipal unions are an example of the latter.

[5] Our dataset does not cover Italy's five Autonomous Regions (Valle d'Aosta, Trentino Alto Adige, Friuli Venezia Giulia, Sicily and Sardinia), which enjoy greater autonomy and special fiscal arrangements.

[6] Municipalities are highly heterogeneous in numerous dimensions, as the country was born in 1861 out of several States with significantly different territorial governance traditions.

[7] In parallel, the legislator has also stimulated processes of municipal amalgamation which, however, have not proven particularly popular.

[8] The initial deadline set to local governments for 31 December 2016 was eventually postponed one day before the deadline by Legislative Decree 244/2016. Until then, however, joining a MU was a compulsory requirement.

[9] http://www.gazzettaufficiale.it/eli/id/2010/07/30/010G0146/sg, accessed on 7/10/2018.

[10] It was particularly after 2014, with the approval of Law 56/2014 (the so-called "Delrio Law"), that regions were allowed to transfer functions from provinces to municipal unions, while the latter were granted significantly higher regional financial incentives.

[11] Central government incentives remain however very limited. For instance, Law 147/2013 (the so-called "Stability Law 2014" introduced a yearly fund in support of unionisation of only 30 million euro for the whole of Italy.

[12] http://www.opencivitas.it/open-data. Accessed in January 2018.

[13]For a detailed description of the methodology used to calculate municipal expenditures, readers can refer to the following document prepared by SOSE Spa: http://www.mef.gov.it/ministero/commissioni/ctfs/documenti/Nota_revisione_metodologia_FS2017_SOSE_13_settembre_2016.pdf, accessed on 27/04/2020.

[14] For each of the 12 service categories we analyse, we select an indicator based both on their relevance and their completeness, trying to favour "hard" measures across all areas. e.g. Tax collection: "average time needed to collect local taxes"; Civil registry: "number of certificates prepared directly at the counter every 1000 inhabitants"; Technical office: "number of executive projects for public works approved every 1000 persons"; etc. These outputs are collected through questionnaires sent by SOSE to municipalities. While we have no way to externally assess the quality of the data collected by SOSE, we have to somehow 'trust' the dataset, and rely on the fact that such data is officially used by the Ministry of Interior and the Ministry of Economy and Finance to allocate parts of the central government transfers to municipalities (the *Fondo di Solidarietá Comunale*, through the "*fabbisogni standard*" procedure, as introduced by Legislative Decree 216/2010). The lack of data availability in empirical research is alas a key issue. Most papers in the literature exclusively explore the effects of inter-municipal cooperation on expenditure rather than on the combined interplay expenditure/service outputs (or consider only a small subset of outputs) exactly because of the lack of output data availability. Aware that SOSE outputs are not perfect, we believe that they are the best data on municipal delivery of public services currently available.

[15] The indicators aim to capture the quality of services by quantitatively measuring the amount of outputs. Of course, these indicators are unable to account for the intrinsic quality of the services provided.

[16] The decision of using a VRS version instead of the constant (CRS) or non-increasing-returns to scale (NIRS) DEA was primarily because the former impose fewer restrictions to the underlying production

technology (Banker et al., 1984). Since CRS and NIRS are restricted version of the VRS-DEA, we run the tests proposed by Simar & Wilson (2002, 2011). Both tests rejected at a 5% confidence level the hypotheses of CRS and NIRS.

[17] Opencivitas (cf. https://www.opencivitas.it/progetto-fabbisogni-standard, accessed on 9/11/2018) follow a different approach to measure efficiency. They develop an indicator constructed as the difference between local expenditure and a measure of "local need", in turn calculated through a regression-based cost approach which, however, rests in our view on strong assumptions.

[18] They are available on request.

[19] We also applied the alternative Free Disposal Hull method (FDH) (Deprins & Tulkens, 1984). Unlike DEA, however, for more than seven output variables, FDH was unable to tell the differences in spending efficiency between municipalities. For instance, for eight output variables, FDH yielded more than 75% of the municipalities laying at the efficiency frontier.

[20] Ivaldi et al (2016) calculate their score with data for 2010, using the Free Disposal Hull Method, applying it to a case with two inputs (current expenditure and personnel expenditure) but only two synthetic output measures without bootstrapping. Besides, their output measures may be less reliable. The first one, municipal population, may be a *determinant* of efficiency rather than an outcome (cf. our Appendix A.5). The second, a synthetic index developed by Opencivitas measuring how municipalities are able to satisfy their citizens' needs, is potentially endogenous, as it is calculated taking into account municipal expenditure plus a host of indicators of 'municipal economic need' which includes, among others, municipal managerial practices (which could be again a determinant of bureaucratic efficiency rather than an outcome).

[21] Indeed, the five most efficient municipalities according to our robust efficiency score are all below 11,000 inhabitants: Examples: Bonata Sopra (Lombardy), Sirignano (Campania), Borgarello (Lombardy), Camisano Vicentino (Veneto) and Bosaro (Veneto).

[22] www.crimetec.it.

[23] Future work should explore the impacts of these alternative arrangements on efficiency. We thank one anonymous referee for flagging out such issue.

[24] It's important to remember that Average Treatment Effects (ATEs) would be inconsistent and biased, since a significant stratum of the municipal population never receive the treatment (Winship & Morgan, 2014): municipal unions were indeed designed for small and medium municipalities.

[25] In contrast to NN-Matching, tests not presented but available on request show how the Propensity-score Matching finds a small but positive and significant effect of unionisation on efficiency and a negative and significant effect on expenditure. Yet, we are afraid that these results may be biased, as strongly argued by King & Nielsen (2016).

[26] Results not presented but available on request interestingly find evidence of non-random sorting around the 5,000 inhabitants' threshold for the 1971 and 1981 censuses (in line with Eggers et al., 2018), but not for the consequent 1991 and 2001 ones. Eggers et al. (2018) underline how there are two biases in the 'standard' McCrary test which tend to increase the likelihood of falsely detecting sorting even when this is not present. We do not consider these biases since our results do not anyway find evidence of strategic manipulation.

[27] In particular, we consider as low economy of scale the following policy areas: tax collection, civil registry, technical office, other general management services, local police, public schools, welfare support, and public kindergartens. By contrast, we consider as high economy of scale the following sectors: road maintenance, local public transports, territorial planning and environmental protection, and waste management.

[28] Balance tests for these two extra specifications are available on request.

[29] We thank one anonymous referee for suggesting us this hypothesis.

ACKNOWLEDGEMENTS

The authors are grateful to J. Vernon Henderson, Fabio Fiorillo, Nicola Matteucci, Antonio Páez and Olmo Silva for helpful comments and suggestions, as well as to Massimo Tatarelli and his team from the Italian Ministry of Interior's Central Directorate for Local Finance for sharing data on municipal unions. The authors are also thankful to participants at the UNIVPM Economics seminar, AISRe 2018 Conference, and LARCSA 2018 Conference, and to three anonymous referees whose comments and suggestions significantly improved the manuscript.

DATA STATEMENT

The data used in the research is available on request.

REFERENCES

Abadie, A., & Imbens, G. W. (2011). Bias-corrected matching estimators for average treatment effects. *Journal of Business & Economic Statistics*, *29*(1), 1–11. https://doi.org/10.1198/jbes.2009.07333

Ahrend, R., Farchy, E., Kaplanis, I., & Lembcke, A. C. (2014). *What makes cities more productive? Evidence on the role of urban governance from five OECD Countries* (No. 2014/5). https://doi.org/10.1787/5jz432cf2d8p-en

Allers, M. A., & de Greef, J. A. (2018). Intermunicipal cooperation, public spending and service levels. *Local Government Studies*, *44*(1), 127–150. https://doi.org/10.1080/03003930.2017.1380630

Allers, M. A., & 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. https://doi.org/10.1111/jors.12268

Allers, M. A., & van Ommeren, B. (2016). Intermunicipal cooperation, municipal amalgamation and the price of credit. *Local Government Studies*, *42*(5), 717–738. https://doi.org/10.1080/03003930.2016.1171754

Andrews, I., & Stock, J. (2018). Weak instruments and what to do about it. *NBER Summer Institute 2018. Methods Lectures*. Retrieved from https://www.nber.org/econometrics_minicourse_2018/2018si_methods.pdf

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in Data Envelopment Analysis. *Management Science*, *30*(9), 1078–1092. https://doi.org/10.1287/mnsc.30.9.1078

Bartolini, D., & Fiorillo, S. (2011). Cooperation among local councils for the provision of public goods. *Rivista Italiana Degli Economisti*, *XVI*(1), 85–108.

Bel, G., Dijkgraaf, E., Fageda, X., & Gradus, R. (2010). Similar problems, different solutions: Comparing refuse collection in The Netherlands and Spain. *Public Administration*, *88*(2), 479–495. https://doi.org/10.1111/j.1467-9299.2009.01808.x

Bel, G., & Mur, M. (2009). Intermunicipal cooperation, privatization and waste management costs: Evidence from rural municipalities. *Waste Management*, *29*(10), 2772–2778. https://doi.org/10.1016/J.WASMAN.2009.06.002

Bel, G., & Warner, M. E. (2015). Inter-municipal cooperation and costs: expectations and evidence. *Public Administration*, *93*(1), 52–67. https://doi.org/10.1111/padm.12104

Besley, T. (2006). *Principled agents? The political economy of good government*. Oxford: Oxford University Press.

Blesse, S., & Baskaran, T. (2016). Do municipal mergers reduce costs? Evidence from a German federal state. *Regional Science and Urban Economics*, *59*, 54–74. https://doi.org/10.1016/j.regsciurbeco.2016.04.003

Breuillé, M.-L., Duran-Vigneron, P., & Samson, A.-L. (2018). Inter-municipal cooperation and local taxation. *Journal of Urban Economics*, *107*, 47–64. https://doi.org/10.1016/J.JUE.2018.08.001

Buettner, T., & Holm-Hadulla, F. (2013). City size and the demand for local public goods. *Regional Science and Urban Economics*, *43*(1), 16–21. https://doi.org/10.1016/J.REGSCIURBECO.2012.10.003

Calonico, S., Cattaneo, M. D., Farrell, M. H., & Titiunik, R. (2017). Rdrobust: Software for regression-discontinuity designs. *Stata Journal*, *17*(2), 372–404. https://doi.org/10.1177/1536867x1701700208

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014a). Robust data-driven inference in the regression-discontinuity design. *Stata Journal*, *14*(4), 909–946. https://doi.org/10.1257/jel.48.2.281

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014b). Robust nonparametric confidence intervals for Regression-Discontinuity Designs. *Econometrica*, *82*(6), 2295–2326. https://doi.org/10.3982/ECTA11757

Casula, M. (2014). Politiche regionali di incentivazione all'esercizio associato di funzioni e servizi a confronto: i casi di Veneto ed Emilia-Romagna. *Istituzioni Del Federalismo*, *3*, 667–697.

Cattaneo, M. D., Jansson, M., & Ma, X. (2018). Manipulation testing based on density discontinuity. In *The Stata Journal* (Vol. 18). Retrieved from https://sites.google.com/site/rdpackages/.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, *2*(6), 429–444. https://doi.org/10.1016/0377-2217(78)90138-8

Cheshire, P. C., & Gordon, I. R. (1996). Territorial competition and the predictability of collective (in)action. *International Journal of Urban and Regional Research*, *20*(3), 383–399. https://doi.org/10.1111/j.1468-2427.1996.tb00324.x

Crescenzi, R., Luca, D., & Milio, S. (2016). The geography of the economic crisis in Europe: National macroeconomic conditions, regional structural factors and short-term economic performance. *Cambridge Journal of Regions, Economy and Society*, *9*(1), 13–32. https://doi.org/10.1093/cjres/rsv031

Deprins, D., & Tulkens, H. (1984). Measuring labour efficiency in post offices. In M. Marchand & H. Tulkens (Eds.), *The performance of public enterprises: Concepts and measurements* (pp. 243–267). Amsterdam: North-Holland.

Di Porto, E., Parenti, A., Paty, S., & Abidi, Z. (2017). Local government cooperation at work: a control function approach. *Journal of Economic Geography*, *17*(2), 435–463. https://doi.org/10.1093/jeg/lbw008

Dijkgraaf, E., & Gradus, R. H. J. M. (2013). Cost advantage cooperations larger than private waste collectors. *Applied Economics Letters*, *20*(7), 702–705. https://doi.org/10.1080/13504851.2012.732682

Djankov, S., McLiesh, C., & Ramalho, R. M. (2006). Regulation and growth. *Economics Letters*, *92*(3), 395–401. https://doi.org/10.1016/j.econlet.2006.03.021

Drukker, D. M. (2016). A generalized regression-adjustment estimator for average treatment effects from panel data. *The Stata Journal*, *16*(4), 826–836.

Dyson, R. G., Allen, R., Camanho, A. S., Podinovski, V. V., Sarrico, C. S., & Shale, E. A. (2001). Pitfalls and protocols in DEA. *European Journal of Operational Research*, *132*(2), 245–259. https://doi.org/10.1016/S0377-2217(00)00149-1

Eggers, A. C., Freier, R., Grembi, V., & Nannicini, T. (2018). Regression Discontinuity Designs based on population thresholds: pitfalls and solutions. *American Journal of Political Science*, *62*(1), 210–229. https://doi.org/10.1111/ajps.12332

Epple, D., & Romer, T. (1989). On the flexibility of municipal boundaries. *Journal of Urban Economics*, *26*(3), 307–319. https://doi.org/10.1016/0094-1190(89)90004-1

Ermini, B., & Fiorillo, F. (2009). L'assetto federale e le forme di associazionismo intercomunale. *Prisma. Economia Società Lavoro*, *J*(2), 79–96.

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society Series A*, *120*, 153–290.

Feiock, R. C., & Scholz, J. T. (2009). *Self-organizing federalism: Collaborative mechanisms to mitigate institutional collective action dilemmas*. Cambridge: Cambridge University Press.

Ferraresi, M., Migali, G., & Rizzo, L. (2018). Does inter-municipal cooperation promote efficiency gains? Evidence from Italian Municipal Unions. *Journal of Regional Science*, *58*, 1017–1044.

Ferraresi, M., & Secomandi, R. (2015). Unioni di comuni: risparmi di spesa? Il caso della Lombardia. In *La Finanza Territoriale. Rapporto 2015* (pp. 203–226). Bologna: Franco Angeli.

Frère, Q., Leprince, M., & Paty, S. (2014). The Impact of intermunicipal cooperation on local public spending. *Urban Studies*, *51*(8), 1741–1760. https://doi.org/10.1177/0042098013499080

Gagliarducci, S., & Nannicini, T. (2013). Do better paid politicians perform better? Disentangling incentives from selection. *Journal of the European Economic Association*, *11*(2), 369–398. https://doi.org/10.1111/jeea.12002

Gambi, L., & Merloni, F. (Eds.). (1995). *Amministrazioni pubbliche e territorio in Italia*. Bologna: Il Mulino.

Gelman, A., & Imbens, G. (2017). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*, 0–0. https://doi.org/10.1080/07350015.2017.1366909

Giacomini, D., Sancino, A., & Simonetto, A. (2018). The introduction of mandatory inter-municipal cooperation in small municipalities. Preliminary lessons from Italy. *International Journal of Public Sector Management*, *31*(3), 331–346.

Grembi, V., Nannicini, T., & Troiano, U. (2016). Do fiscal rules matter? *American Economic Journal: Applied Economics*, *8*(3), 1–30. https://doi.org/10.1257/app.20150076

Harjunen, O., Saarimaa, T., & Tukiainen, J. (2017). *Political representation and effects of municipal mergers* (No. 98). Helsinki.

Hinnerich, B. (2009). Do merging local governments free ride on their counterparts when facing boundary reform? *Journal of Public Economics*, *93*(5–6), 721–728. https://doi.org/10.1016/J.JPUBECO.2009.01.003

Hirota, H., & Yunoue, H. (2017). Evaluation of the fiscal effect on municipal mergers: Quasi-experimental evidence from Japanese municipal data. *Regional Science and Urban Economics*, *66*, 132–149. https://doi.org/10.1016/J.REGSCIURBECO.2017.05.010

Howell-Moroney, M. (2008). The Tiebout hypothesis 50 years later: Lessons and lingering challenges for metropolitan governance in the 21st Century. *Public Administration Review*, *68*(1), 97–109. https://doi.org/10.1111/j.1540-6210.2007.00840.x

Hulst, J. R. (Rudie), & van Montfort, A. (Eds.). (2007). *Inter-municipal cooperation in Europe*. Amsterdam: Springer Netherlands.

Hulst, J. R. (Rudie), & van Montfort, A. J. G. M. (André). (2012). Institutional features of inter-municipal cooperation: Cooperative arrangements and their national contexts. *Public Policy and Administration*, *27*(2), 121–144. https://doi.org/10.1177/0952076711403026

Ivaldi, G., Marinuzzi, G., Quintiliani, F., & Tortorella, W. (2016). Le unioni di comuni in Italia: strumenti efficienti a quali condizioni? *XXXVII Conferenza Italiana Di Scienze Regionali*. Ancona.

Jordahl, H., & Liang, C.-Y. (2010). Merged municipalities, higher debt: on free-riding and the common pool problem in politics. *Public Choice*, *143*(1–2), 157–172. https://doi.org/10.1007/s11127-009-9495-y

King, G., & Nielsen, R. (2016). Why propensity scores should not be used for matching. In *Unpublished manuscript*. Retrieved from https://gking.harvard.edu/files/gking/files/psnot.pdf?m=1481894836

Koopmans, T. C. (1951). An analysis of production as an efficient combination of activities. In T. C. Koopmans & C. C. for R. in Economics (Eds.), *Activity analysis of prodution and allocation: Proceedings of a conference* (pp. 33–97). New York: Wiley.

Kwon, S.-W., & Feiock, R. C. (2010). Overcoming the barriers to cooperation: Intergovernmental service agreements. *Public Administration Review*, *70*(6), 876–884. https://doi.org/10.1111/j.1540-6210.2010.02219.x

Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, *48*(2), 281–355. https://doi.org/10.1257/jel.48.2.281

Lima, R. C. de A., & Silveira Neto, R. da M. (2018). Secession of municipalities and economies of scale: Evidence from Brazil. *Journal of Regional Science*, *58*(1), 159–180. https://doi.org/10.1111/jors.12348

Luca, D. (2017). Boon or bane for development? Turkey's central state bureaucracy and the management of public investment. *Environment and Planning C: Politics and Space*, *35*(6), 939–957. https://doi.org/10.1177/0263774X16670666

Luca, D., & Salone, C. (2013). Regional theories and institutional regions. For an ontology of the relationships between government and collective action spaces | Teorie regionali e regioni istituzionali. Per un'ontologia del rapporto tra spazi di governo e spazi di azione collettiva. *Rivista Geografica Italiana*, *120*(3), 209–224.

Manestra, S., Messina, G., & Peta, A. (2018). *L'unione (non) fa la forza? Alcune evidenze empiriche sull'associazionismo comunale in Italia*. Rome.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, *142*(2), 698–714. https://doi.org/10.1016/J.JECONOM.2007.05.005

Nielsen, R. A., Findley, M. G., Davis, Z. S., Candland, T., & Nielson, D. L. (2011). Foreign aid shocks as a cause of violent armed conflict. *American Journal of Political Science*, *55*(2), 219–232. https://doi.org/10.1111/j.1540-5907.2010.00492.x

Niskanen, W. A. (2001). Bureaucracy. In W. F. Shugart & L. Razzolini (Eds.), *The Elgar companion to Public Choice* (pp. 258–270). Cheltenham, UK: Edward Elgar.

Oates, W. E. (1972). *Fiscal federalism*. New York: Harcourt, Brace and Jovanovich.

Oates, W. E. (1999). An essay on fiscal federalism. *Journal of Economic Literature*, *37*(3), 1120–1149. https://doi.org/10.1257/jel.37.3.1120

OECD. (2002). *Redifining territories: The functional regions*. Paris: OECD Publishing.

OECD. (2015). *Governing the city*. https://doi.org/http://dx.doi.org/10.1787/9789264226500-en

Ostrom, E. (2010). A Long Polycentric Journey. *Annual Review of Political Science*, *13*(1), 1–23. https://doi.org/10.1146/annurev.polisci.090808.123259

Ostrom, V., Tiebout, C. M., & Warren, R. (1961). The organization of government in metropolitan areas: A theoretical inquiry. *The American Political Science Review*, *55*(4), 831. https://doi.org/10.2307/1952530

Parks, R. B., Baker, P. C., Kiser, L., Oakerson, R., Ostrom, E., Ostrom, V., … Wilson, R. (1981). Consumers as producers of public services: Some economic and institutional considerations. *Policy Studies Journal*, *9*(7), 1001–1011. https://doi.org/10.1111/j.1541-0072.1981.tb01208.x

Podinovski, V. V., & Thanassoulis, E. (2007). Improving discrimination in data envelopment analysis: some practical suggestions. *Journal of Productivity Analysis*, *28*(1–2), 117–126. https://doi.org/10.1007/s11123-007-0042-x

Putnam, R. D. (1993). *Making democracy work*. Princeton, NJ: Princeton University Press.

Reingewertz, Y. (2012). Do municipal amalgamations work? Evidence from municipalities in Israel. *Journal of Urban Economics*, *72*(2–3), 240–251. https://doi.org/10.1016/J.JUE.2012.06.001

Rouse, P., & Putterill, M. (2005). Local government amalgamation policy: A highway maintenance evaluation. *Management Accounting Research*, *16*(4), 438–463. https://doi.org/10.1016/J.MAR.2005.07.003

Ruggiero, J. (2000). Measuring technical efficiency. *European Journal of Operational Research*, *121*(1), 138–150. https://doi.org/10.1016/S0377-2217(99)00010-7

Simar, L., & Wilson, P. W. (1998). Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. *Management Science*, *44*(1), 49–61. https://doi.org/10.1287/mnsc.44.1.49

Simar, L., & Wilson, P. W. (2000). A general methodology for bootstrapping in non-parametric frontier models. *Journal of Applied Statistics*, *27*(6), 779–802. https://doi.org/10.1080/02664760050081951

Simar, L., & Wilson, P. W. (2002). Non-parametric tests of returns to scale. *European Journal of Operational Research*, *139*(1), 115–132. https://doi.org/10.1016/S0377-2217(01)00167-9

Simar, L., & Wilson, P. W. (2011). Inference by the m out of n bootstrap in nonparametric frontier models. *Journal of Productivity Analysis*, *36*(1), 33–53. https://doi.org/10.1007/s11123-010-0200-4

Simm, J., & Besstremyannaya, G. (2016). *Robust Data Envelopment Analysis (DEA) for R*. Retrieved from https://cran.r-project.org/web/packages/rDEA/rDEA.pdf

Sorensen, R. J. (2007). Does dispersed public ownership impair efficiency? The case of refuse collection in Norway. *Public Administration*, *85*(4), 1045–1058. https://doi.org/10.1111/j.1467-9299.2007.00681.x

Steiner, R. (2003). The causes, spread and effects of intermunicipal cooperation and municipal mergers in Switzerland. *Public Management Review*, *5*(4), 551–571. https://doi.org/10.1080/1471903032000178581

Weingast, B. R., Shepsle, K. A., & Johnsen, C. (1981). The political economy of benefits and costs: a Neoclassical approach to distributive politics. *Journal of Political Economy*, *89*(4), 642–664.

Winship, C., & Morgan, S. L. (Eds.). (2014). Matching estimators of causal effects. In *Counterfactuals and Causal Inference: Methods and Principles for Social Research* (2nd ed., pp. 140–187). https://doi.org/DOI: 10.1017/CBO9781107587991.006

TABLES

TABLE 1: Municipal cooperation and administrative efficiency: NN-Matching estimates

	(1) Efficiency score	(2) Expenditure
MU member	0.006**	0.036***
	(0.003)	(0.010)
Observations	5,644	5,644

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.

TABLE 2: Municipal cooperation and administrative efficiency: Fuzzy-RDD estimates

	All Italy		North/Centre
Second stage	(1)	(2)	(3)
Conventional	0.030	0.040	0.049
	(0.060)	(0.054)	(0.035)
Bias-corrected	0.063	0.067	0.050
	(0.060)	(0.054)	(0.035)
Robust	0.063	0.067	0.050
	(0.068)	(0.064)	(0.040)
First stage			
Conventional	-0.157**	-0.153**	-0.204**
	(0.078)	(0.073)	(0.087)
Bias-corrected	-0.186**	-0.178***	-0.225***
	(0.078)	(0.073)	(0.087)
Robust	-0.186**	-0.178**	-0.225**
	(0.088)	(0.082)	(0.097)
Observations	5,644	5,644	4,064
Order Loc Poly (p)	2	2	2
Bandwidth (h) L	1596	1611	1661
Bandwidth (h) R	1596	1611	1661
Eff. nr of obs L	625	630	501
Eff. nr of obs R	398	400	298
1-st. Montliel-Pflueger F	2.50	24.51	19.43
Region FE	yes	yes	yes
Controls	no	yes	yes

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Controls and regional FE are not reported. The models are estimated adopting a triangular kernel type and a one common MSE-optimal bandwidth selector. Counterintuitively, first-stage coefficients show a negative sign since *rdrobust*, the package we use for estimation (cf. Calonico, Cattaneo, Farrell, & Titiunik, 2017), calculates them for municipalities *above* – rather than below – the 5,000 threshold.

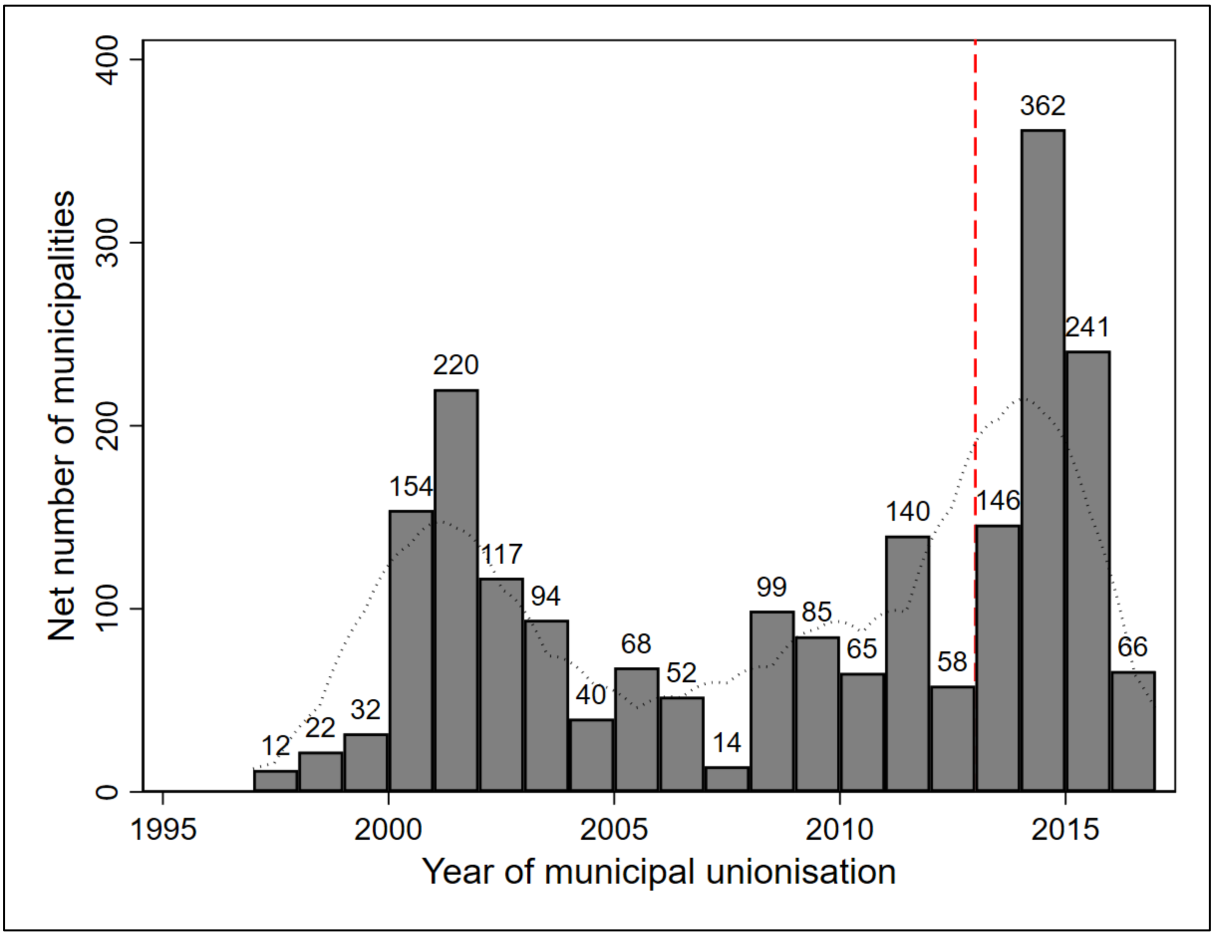
TABLE 3: Municipal cooperation and administrative efficiency: Fuzzy-RDD estimates distinguishing between policy areas characterized by low- and high-economies of scale (municipalities from North and Centre Italy)

	(1) LES	(2) HES	(3) LES	(4) HES
Second stage	All Italy		North/Centre Italy	
Conventional	-0.159	-0.057	0.042	0.057
	(0.217)	(0.100)	(0.170)	(0.071)
Bias-corrected	0.016	-0.049	0.091	0.057
	(0.217)	(0.100)	(0.170)	(0.071)
Robust	0.016	-0.049	0.091	0.057
	(0.266)	(0.111)	(0.187)	(0.078)
First stage				
Conventional	-0.121*	-0.137**	-0.215***	-0.151**
	(0.069)	(0.071)	(0.087)	(0.075)
Bias-corrected	-0.165**	-0.158**	-0.241***	-0.170**
	(0.069)	(0.071)	(0.087)	(0.075)
Robust	-0.165**	-0.158**	-0.241***	-0.170**
	(0.082)	(0.080)	(0.096)	(0.085)
Observations	5,476	5,476	3,921	3,921
Region FE	yes	yes	yes	yes
Controls	yes	yes	yes	yes
Order Loc Poly (p)	2	2	2	2
BW estim (h) L	1878	1740	1490	2065
BW estim (h) R	1878	1740	1490	2065
Eff. nr of obs L	753	675	412	622
Eff. nr of obs R	453	425	278	360

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. The models are estimated adopting a triangular kernel type and a one common MSE-optimal bandwidth selector.

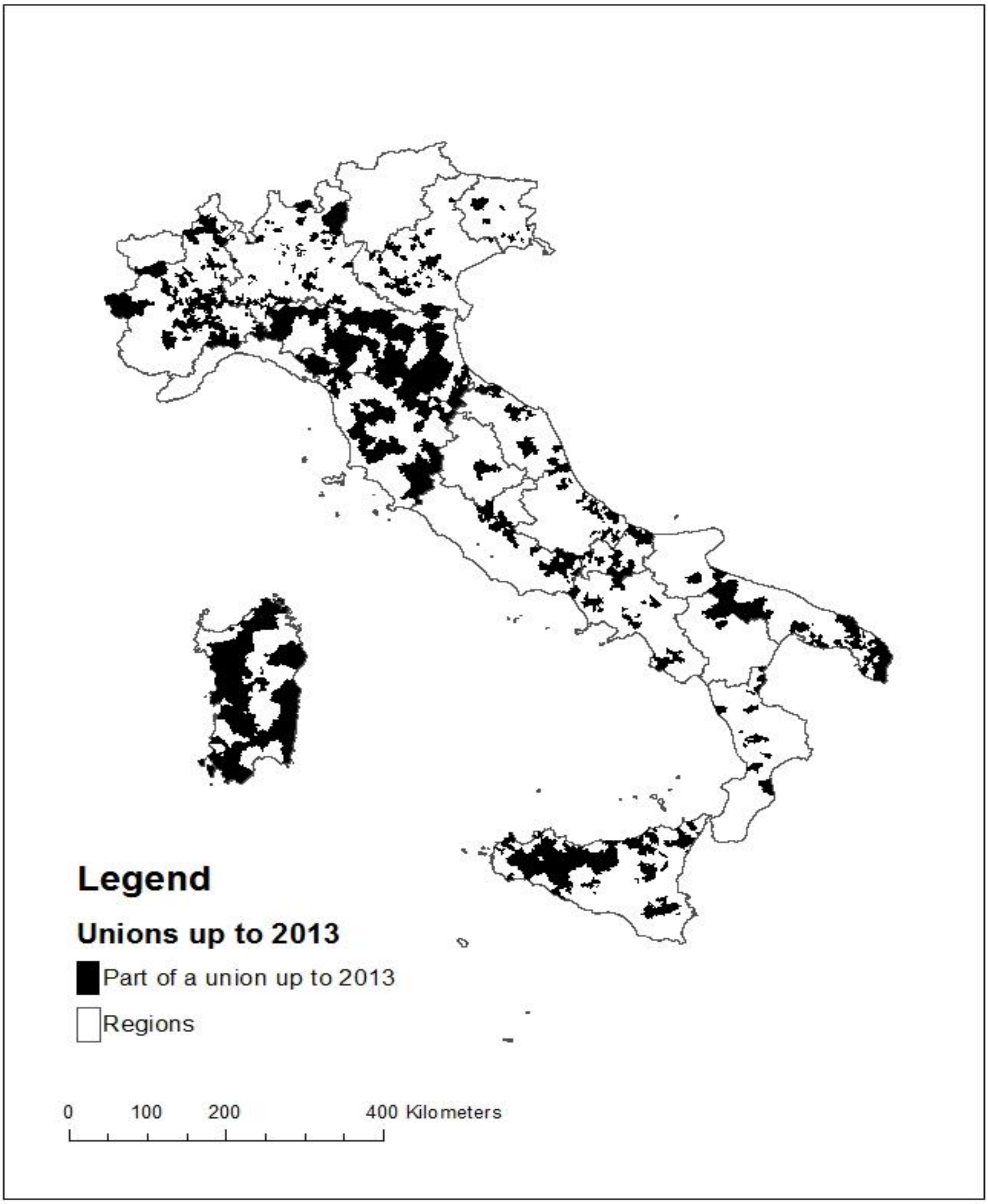
FIGURES

FIGURE 1: Year of municipal unionization in Italy's 15 Ordinary Statute Regions



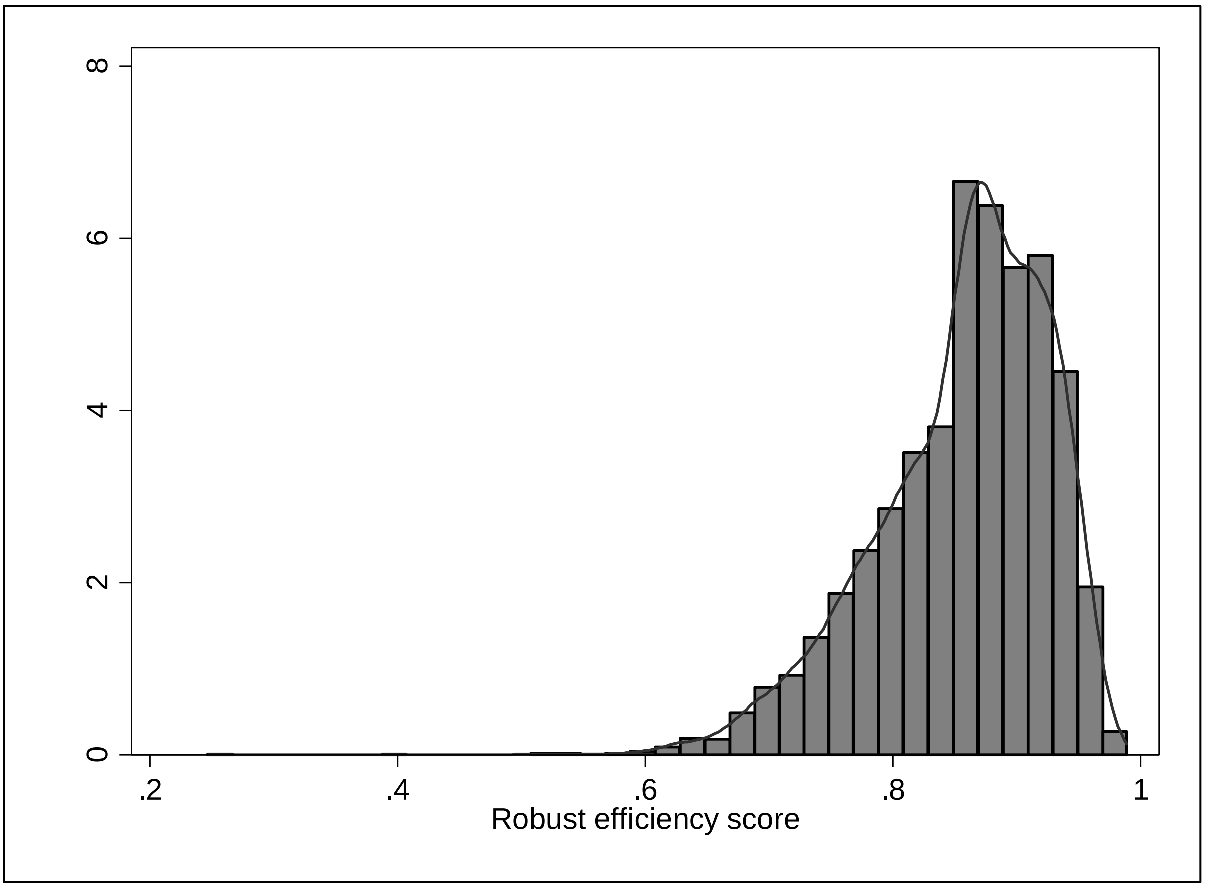
Source: own elaboration. The vertical dashed line indicates 2013, the year covered by our analysis. Note: the graph plots the year in which municipalities joined the unions existing in 2013, hence discarding entry into/exit from previous unions, as well as municipal unions which had ceased to exist before 2013. We do not unfortunately have data on union membership churn before then.

FIGURE 2: Municipalities which were part of a municipal union (by the end of 2013)



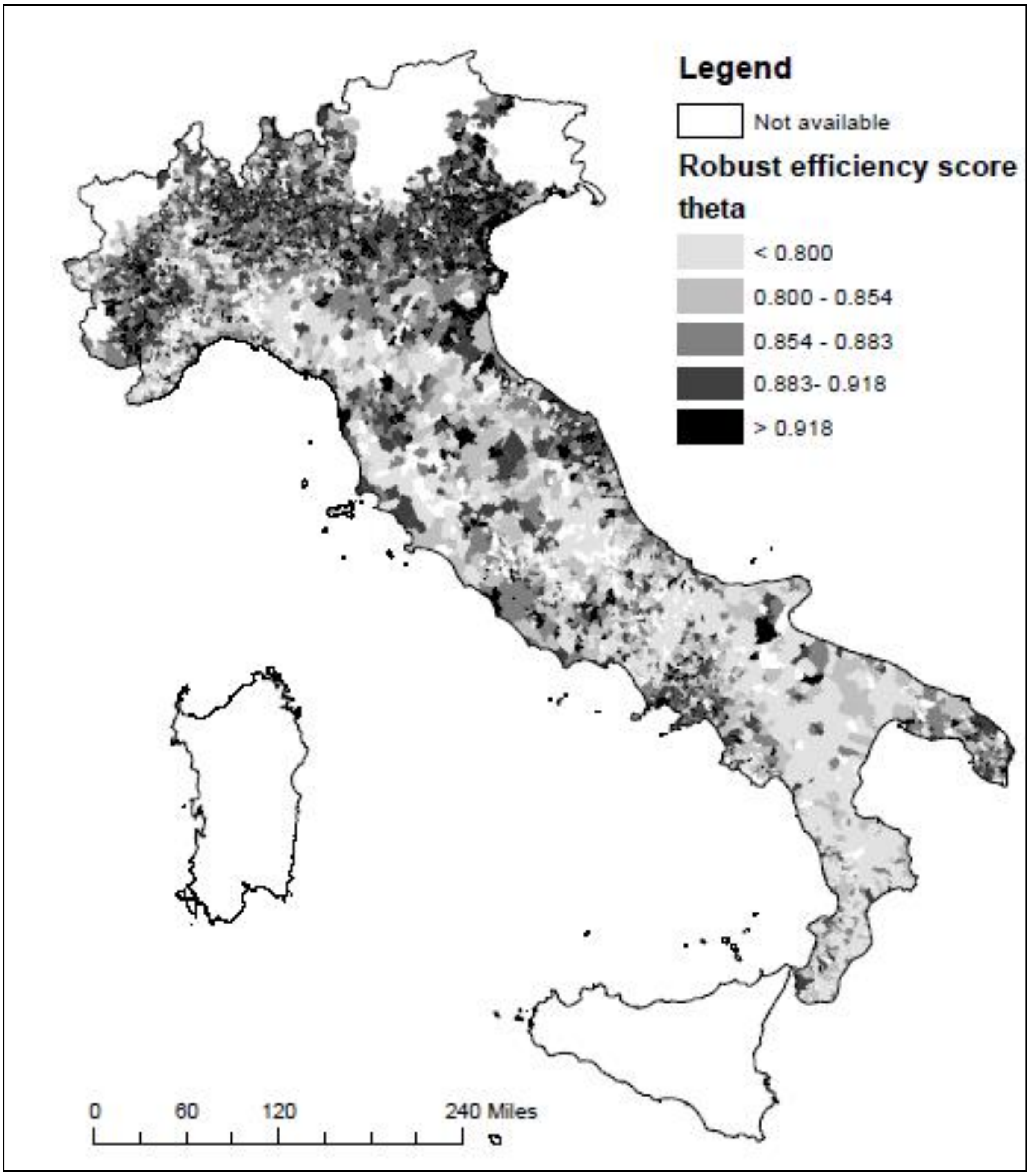
Source: own elaboration.

FIGURE 3: Distribution of the municipal efficiency score



Source: own elaboration.

FIGURE 4: Spending efficiency scores for Italian Municipalities



Source: own elaboration.

APPENDICES

Appendix A.1. Variables' description and source.

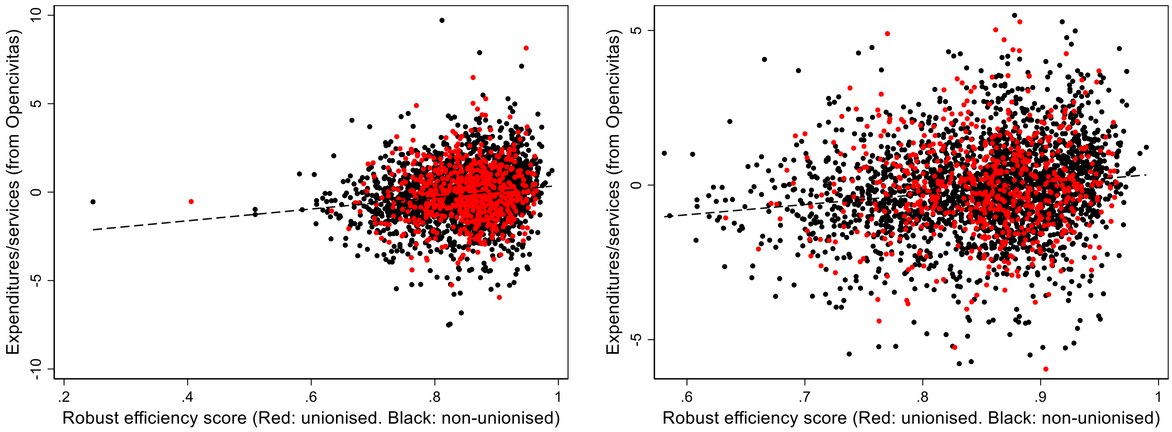
Variables	Description	Year	Source
Efficiency score	Robust efficiency score	2013	Own calculation
Population	Total municipal population	2011	Istat
Population density	Population per square Km	2011	Istat
Altitude zone	Categorical variable (1 to 5) on the altitude of the municipal centre	2011	Istat
Young people	Percent share of young people	2011	Istat
Old people	Percent share of old people	2011	Istat
Foreigners	Percent share of non-natives	2011	Istat
Unemployment	Unemployment share	2011	Istat
University graduates	Percent share of university graduates	2011	Istat
Primary sector	Percent share of workers in primary sector	2011	Istat
Secondary sector	Percent share of workers in secondary sector	2011	Istat
Fiscal base	Average municipal taxable income	2013	Openpolis
Election dummy	D = 1 if the municipality went through local elections between 2010-13	2010/13	Own calculation
Mafia intensity	Indicator measuring on a scale 0-100 the local infiltration of organized crime	2000-2015	Crime&tech

Appendix A.2. Descriptive statistics.

	Treated (1.418)				Untreated (4.226)			
Variables	Mean	SD	Min	Max	Mean	SD	Min	Max
Efficiency	0,858	0,063	0,406	0,979	0.856	0.073	0.247	0.990
Population	5.042,547	8.340,298	114,000	117.946,000	9,922.255	56,813.751	72.000	2873494.000
Pop. density	224,179	389,385	3,390	7.906,340	404.617	783.483	1.090	12,220.690
Altitude zone	2,854	1,680	1,000	5,000	2.880	1.598	1.000	5.000
Young people	0,238	0,027	0,119	0,324	0.242	0.026	0.130	0.324
Old people	0,176	0,051	0,046	0,467	0.170	0.052	0.047	0.437
Foreigners	7,352	4,335	0,000	31,299	6.985	4.348	0.000	36.274
Unemployment	9,209	5,152	0,000	35,458	9.994	5.900	0.000	42.182
Univ. grad.	7,420	2,325	1,049	27,323	7.679	2.878	0.840	26.897
Primary sector	8,331	6,955	0,334	55,589	8.403	8.700	0.000	78.476
Sec. sector	33,603	9,139	8,637	64,615	32.467	11.140	5.532	74.969
Fiscal base	17.325,15	3.114,98	8.776,00	34.070,00	16,925.94	3,671.50	6,796.00	53,973.00
Elect. dummy	0,007	0,084	0,000	1,000	0.008	0.086	0.000	1.000
Mafia intensity	29,971	15,758	0,000	81,980	32.950	18.947	0.000	100.000

The table presents statistics for two groups of municipalities: all those which by the end of 2013 had joined a MU and those which by the end of 2013 were not part of a MU.

Appendix A.3. Pairwise correlation between the robust efficiency score and an alternative efficiency measure based on indicators from Opencivitas: full sample (left), winsorizing the outliers (right).



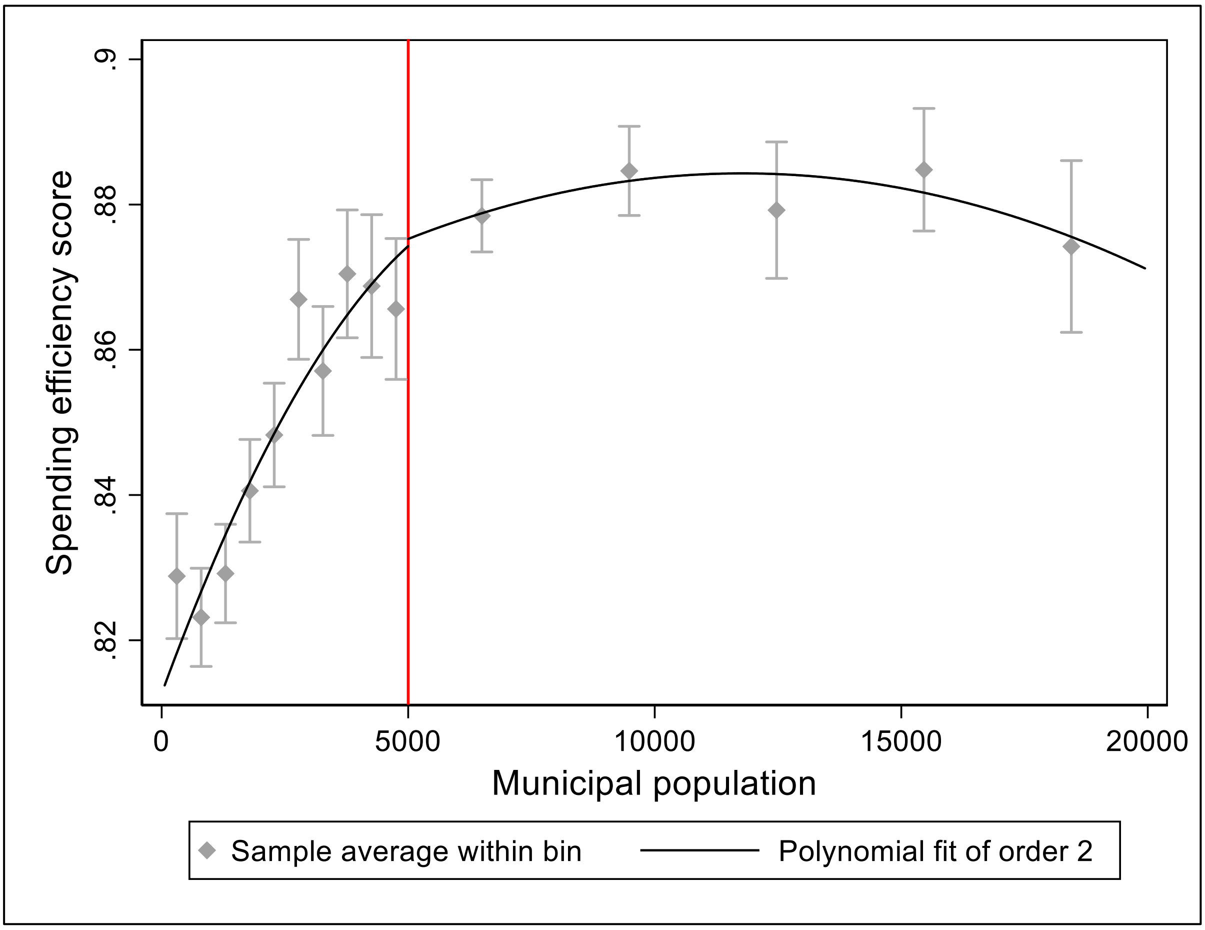
The alternative indicator is calculated as the logarithm of the ratio of two variables: the percent difference between total municipal actual expenditure and theoretical expenditure needs, and the percent difference between the level of services actually offered and the level of services theoretically needed. The plot on the right winsorizes 14 outliers whose robust efficiency score is below 0.55 or whose alternative measure is either below -6 or above 6.

Appendix A.4. Balance tests for model one of Table 1.

	Standardized differences		Variance ratio	
	Raw	Matched	Raw	Matched
Population	-0.193	-0.146	0.717	1.032
Population density	-0.292	-0.006	0.247	1.123
Altimetric class	-0.015	0.010	1.104	0.964
% of university graduates	-0.099	-0.021	0.653	1.190
Unemployment	-0.142	0.026	0.763	1.030
% of workers in primary sector	-0.009	0.041	0.639	1.175
% of workers in secondary sector	0.111	0.003	0.673	1.117
% of young people	-0.153	-0.024	1.079	1.224
% of old people	0.128	0.067	0.959	1.340
% of foreigners	0.084	0.085	0.994	1.305
Fiscal base	-0.004	-0.025	0.697	1.197
Election dummy	-0.049	0	0.599	1
Mafia intensity	-0.171	-0.021	0.692	0.963

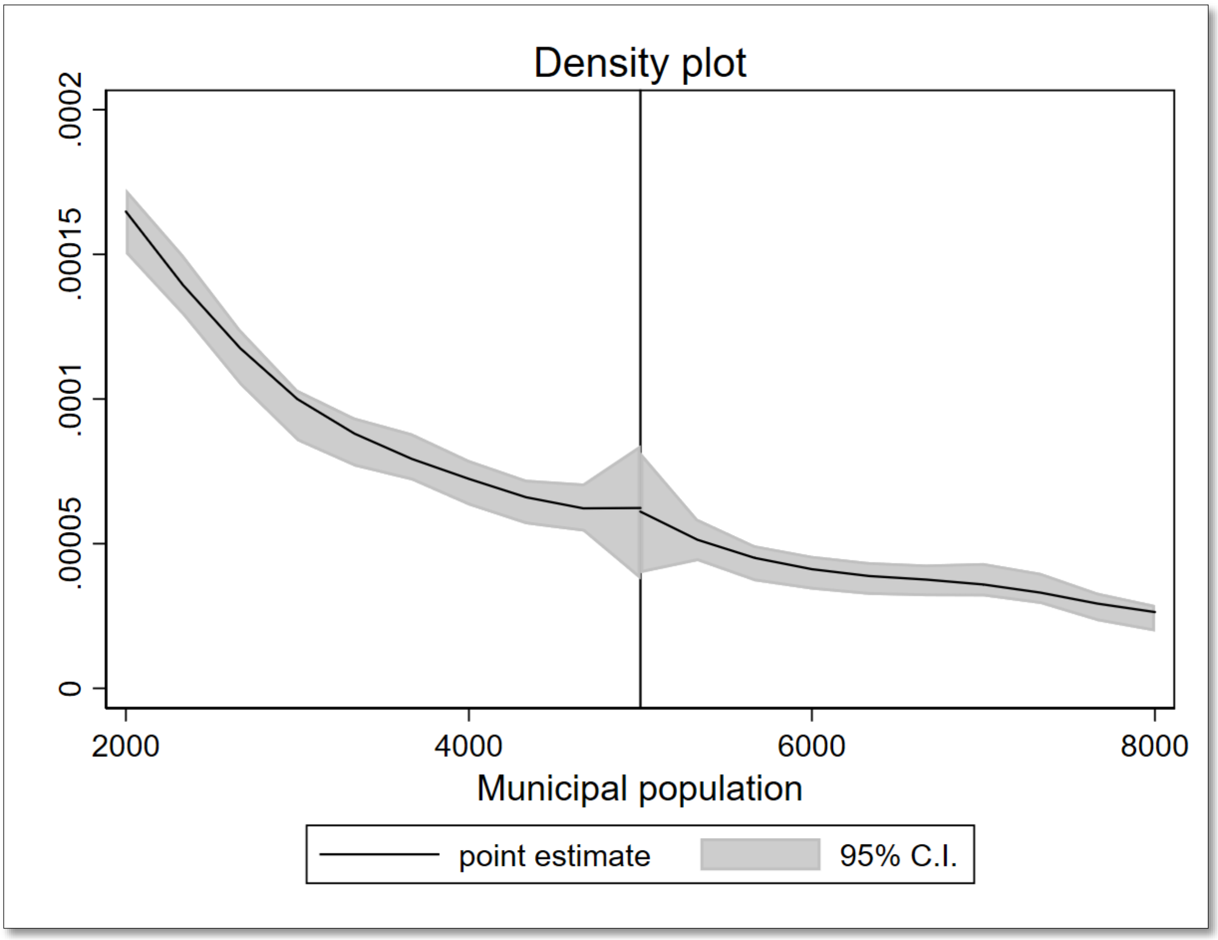
Tests not presented but available on request show that results are virtually unchanged when reducing the number of covariates used to match 'treatment' and 'control' groups.

Appendix A.5. Spending efficiency scores for Italian Municipalities: RDD graph.



Source: own elaboration. The RDD plot is constructed netting out regional fixed-effects and covariates, while the polynomial specification is based on uniform kernel function. The number of bins is selected using an IMSE-optimal evenly-spaced method (Calonico et al., 2014a). For easier readability the graph trims municipalities above 20,000 inhabitants (which represents 7.33% of the total number of municipalities). A graph on the full population is available on request.

Appendix A.6. Test for non-random sorting around the cutoff.



T-stat at the discontinuity -0.018 (p-val 0.986). There is no statistically significant evidence of non-random sorting. The test is based on a restricted local second-order polynomial triangular-kernel density estimator (with jackknife standard errors, and bandwidths selected as min on MSE of difference/sum of densities).

Appendix A.7. Balancing tests for the Fuzzy-RDD.

All Italy												
Dependent var:	Alt	P. dens	Univ.	Agric.	Ind.	Unempl.	Young	Elder	Foreign	Elect.	Mafia	Tax base
5.000 cut-off	0.087	-0.069	0.156	1.270	0.909	-0.049	0.004	-0.004	0.618	0.004	-2.308	185.584
	(0.316)	(47.727)	(0.353)	(1.821)	(1.214)	(0.462)	(0.003)	(0.005)	(0.780)	(0.006)	(3.667)	(362.411)
Observations	5,644	5,644	5,644	5,644	5,644	5,644	5,644	5,644	5,644	5,644	5,644	5,644
Region FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Order Loc Poly	2	2	2	2	2	2	2	2	2	2	2	2
BW estim (h) L	1925	2204	2013	2241	1754	1977	1978	2397	2189	1629	1719	2260
BW estim (h) R	1925	2204	2013	2241	1754	1977	1978	2397	2189	1629	1719	2260
Eff. nr of obs L	794	931	838	952	704	825	826	1069	921	643	684	966
Eff. nr of obs R	465	532	486	539	431	472	473	570	526	402	425	542

North/Central Italy												
Dependent var:	Alt	P. dens	Univ.	Agric.	Ind.	Unempl.	Young	Elder	Foreign	Elect.	Mafia	Tax base
5.000 cut-off	0.127	-61.96**	0.134	-0.080	2.280*	0.125	0.002	-0.008	0.478	0.006	-4.698	422.035
	(0.285)	(29.615)	(0.346)	(0.530)	(1.318)	(0.546)	(0.003)	(0.007)	(0.721)	(0.008)	(5.649)	(380.933)
Observations	4,064	4,064	4,064	4,064	4,064	4,064	4,064	4,064	4,064	4,064	4,064	4,064
Region FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Order Loc Poly	2	2	2	2	2	2	2	2	2	2	2	2
BW estim (h) L	1826	2071	2653	1569	1606	2426	2294	1433	2254	1028	1516	2200
BW estim (h) R	1826	2071	2653	1569	1606	2426	2294	1433	2254	1028	1516	2200
Eff. nr of obs L	560	648	914	457	475	814	741	418	716	282	435	696
Eff. nr of obs R	319	363	442	287	289	416	395	269	387	192	280	379

The table reports robust estimates adopting a triangular kernel, a second-order local polynomial, and one common-MSE-optimal bandwidth selector. Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Appendix A.8. Municipal cooperation and administrative efficiency: Fuzzy-RDD estimates with alternative bandwidths.

	(1)	(2)	(3)	(4)	(5)	(6)
Second stage		All Italy			North/Centre Italy	
Conventional	0.040	0.046	0.033	0.049	0.047	0.049
	(0.054)	(0.053)	(0.056)	(0.035)	(0.033)	(0.035)
Bias-corrected	0.067	0.062	0.067	0.050	0.048	0.050
	(0.054)	(0.053)	(0.056)	(0.035)	(0.033)	(0.035)
Robust	0.067	0.062	0.067	0.050	0.048	0.050
	(0.064)	(0.060)	(0.067)	(0.040)	(0.039)	(0.040)
First stage						
Conventional	-0.153**	-0.163**	-0.147**	-0.204**	-0.220**	-0.204**
	(0.0733)	(0.0774)	(0.0719)	(0.0865)	(0.0919)	(0.0865)
Bias-corrected	-0.178***	-0.179**	-0.177***	-0.225***	-0.232***	-0.225***
	(0.0733)	(0.0774)	(0.0719)	(0.0865)	(0.0919)	(0.0865)
Robust	-0.178**	-0.179**	-0.177**	-0.225**	-0.232**	-0.225**
	(0.0823)	(0.0832)	(0.0821)	(0.0967)	(0.0975)	(0.0967)
Observations	5,644	5,644	5,644	4,064	4,064	4,064
BW Type	mserd	cercomb2	msecomb2	mserd	cercomb2	msecomb2
Region FE	yes	yes	yes	yes	yes	yes
Controls	yes	yes	yes	yes	yes	yes
Order Loc Poly (p)	2	2	2	2	2	2
BW estim (h) L	1611	1390	1688	1661	1408	1661
BW estim (h) R	1611	1390	1688	1661	1408	1661
Eff. nr of obs L	630	533	673	501	408	501
Eff. nr of obs R	400	359	417	298	265	298

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. The models are estimated adopting a triangular kernel type. The Mserd bandwidth type refers to a one common MSE-optimal bandwidth selector. The Msecomb2 is the median between: the one-common MSE-optimal bandwidths selector; the asymmetric (i.d. different on each side of the cutoff) MSE-optimal bandwidths; one-common MSE-optimal selector for the sum of regression estimates. The Cercomb2 bandwidth is similar to the Msecomb2 but it’s estimated for each side of the cutoff separately (Calonico et al., 2014a).

Appendix A.9. Municipal cooperation and administrative efficiency: Fuzzy-RDD estimates with alternative polynomial specifications.

	(1)	(2)	(3)	(4)	(5)	(6)
		All Italy			North/Centre Italy	
Conventional	-0.193	0.040	0.073	-0.017	0.049	0.062*
	(0.332)	(0.054)	(0.062)	(0.059)	(0.035)	(0.032)
Bias-corrected	-0.156	0.067	0.100	-0.016	0.050	0.072**
	(0.332)	(0.054)	(0.062)	(0.059)	(0.035)	(0.032)
Robust	-0.156	0.067	0.100	-0.016	0.050	0.072**
	(0.381)	(0.064)	(0.070)	(0.065)	(0.040)	(0.034)
First stage						
Conventional	-0.0323	-0.153**	-0.176**	-0.0714*	-0.204**	-0.257**
	(0.0410)	(0.0733)	(0.0855)	(0.0440)	(0.0865)	(0.109)
Bias-corrected	-0.0412	-0.178***	-0.195**	-0.0882**	-0.225***	-0.272***
	(0.0410)	(0.0733)	(0.0855)	(0.0440)	(0.0865)	(0.109)
Robust	-0.0412	-0.178**	-0.195**	-0.0882*	-0.225**	-0.272**
	(0.0492)	(0.0823)	(0.0915)	(0.0531)	(0.0967)	(0.118)
Observations	5,644	5,644	5,644	4,064	4,064	4,064
Region FE	yes	yes	yes	yes	yes	yes
Controls	yes	yes	yes	yes	yes	yes
Order Loc Poly (p)	1	2	3	1	2	3
BW estim (h) L	1945	1611	2199	1776	1661	1914
BW estim (h) R	1945	1611	2199	1776	1661	1914
Eff. nr of obs L	806	630	927	540	501	595
Eff. nr of obs R	465	400	530	311	298	331

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. The models are estimated adopting a triangular kernel type and a one common MSE-optimal bandwidth selector.

Appendix A.10. Municipal cooperation and administrative efficiency: First-stage Fuzzy-RDD estimates comparing the 5,000-inhabitant cut-off with alternative placebo thresholds.

	(1)	(2)	(3)	(4)	(5)
Cut-offs	4,000	4,500	5,000	5,500	6,000
Conventional	0.0481	0.0893	-0.153**	0.0582	0.0363
	(0.0870)	(0.0753)	(0.0733)	(0.0491)	(0.0547)
Bias-corrected	0.0740	0.0988	-0.178***	0.0355	0.0515
	(0.0870)	(0.0753)	(0.0733)	(0.0491)	(0.0547)
Robust	0.0740	0.0988	-0.178**	0.0355	0.0515
	(0.109)	(0.0831)	(0.0823)	(0.0537)	(0.0610)
Observations	5,644	5,644	5,644	5,644	5,644
Region FE	yes	yes	yes	yes	yes
Controls	yes	yes	yes	yes	yes
Order Loc Poly (p)	2	2	2	2	2
BW estim (h) L	1772	1474	1611	5429	2359
BW estim (h) R	1772	1474	1611	5429	2359
Eff. nr of obs L	975	646	630	3867	782
Eff. nr of obs R	582	436	400	869	442

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. The models are estimated adopting a triangular kernel type and a one common MSE-optimal bandwidth selector.

Appendix A.11. Effect of crossing other population thresholds on efficiency and expenditure: sharp-RDD estimates.

	(1)	(2)	(3)
Cut-offs:	1.000	3.000	5.000
Dependent variable:	Efficiency measure		Pc expenditure
Conventional	-0.001	-0.003	-0.058
	(0.006)	(0.009)	(0.037)
Bias-corrected	0.004	-0.002	-0.066*
	(0.006)	(0.009)	(0.037)
Robust	0.004	-0.002	-0.066
	(0.008)	(0.010)	(0.040)
Observations	5,644	5,644	5,644
Region FE	yes	yes	yes
Controls	yes	yes	yes
Order Loc Poly (p)	2	2	2
BW estim (h) L	1127	1336	1438
BW estim (h) R	1127	1336	1438
Eff. nr of obs L	1011	1042	552
Eff. nr of obs R	1282	604	369

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. The models are estimated adopting a triangular kernel type and a one common MSE-optimal bandwidth selector.

Appendix A.12. Municipal cooperation and administrative efficiency: NN-Matching estimates diving the sample between municipalities part of a small or large municipal union (columns 1 and 2), or including only municipalities in unions with at least one large member with a population above the Italian average (8.696 inhabitants).

	MU size		MU with large municipality
	(1) < average	(2) > average	(3)
	-0.000	0.014***	0.008**
	(0.003)	(0.004)	(0.004)
Observations	4,809	4,383	4,763
Region FE	yes	yes	yes
Controls	yes	yes	yes

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.