Schooling Provision and Residential Choices: Evidence from Italy

Agar Brugiavini, Marco Di Cataldo, Giulia Romani

August, 2020

Abstract

The availability of local public goods can orient residential choices. Schools are especially relevant in this respect, since households with children have a daily need for school services. An aspect so far overlooked is the role played by economies of scale in public schooling provision. However, their presence lies at the core of rationalisation policies, which aim to increase efficiency by cutting on undersized nodes of the school network. If households relocate in response to service cuts, rationalisation policies can have the effect of promoting population decline. This paper investigates the demographic impact of primary school closures in the Italian context, where the Gelmini reform of 2008 produced a significant contraction of the school network. We want to assess whether service cuts have an effect on top and beyond preexisting negative population trends which motivate school closures. We address the fundamental empirical challenge of services-population endogeneity through two complementary strategies, both accounting for population pre-trends. Our preliminary findings suggest that school closures further promote population decline. Moreover, the effect appears stronger for municipalities endowed with few schools in absolute value and it intensifies in the first few years after the closure. This evidence provides highly relevant policy implications, namely that rationalisation policies foster depopulation and therefore increase regional divergence.

Key words: school closures, residential choices, regional divergence
JEL: H40; H52; R23

1 Introduction

Theory and common sense suggest that access to services plays a key role in orienting residential choices. People decide where to live taking into account job opportunities, idiosyncratic preferences, but also the availability of near and good quality services. The idea of people moving to access services has a long academic tradition, stemming

from Tiebout's contribution (1956)[30]. Despite its long history, Tiebout's proposition has just recently undergone empirical tests, mainly in the US context[1]. In that academic tradition, school service attracted particular attention, because of its relevance for residential choices. However, the focus has mainly been on school quality differentials and related dynamics of households sorting by socio-economic status. In this article, instead, we take a different perspective and consider the role played by scale economies in shaping the distribution of school services. Despite largely neglected by the above literature, scale economies are significantly present in the provision of school service. This fact is highlighted by influential theoretical references, such as Urquiola (2005)[31] and Alesina, Baqir and Hoxby (2004)[1]. Moreover, the implementation of 'rationalisation policies' represents in itself a proof of existing scale economies. Indeed, rationalisation policies consist of eliminating undersized service centres in order to increase economic efficiency. If this in turn induced people to move, the dynamics generated would very much resemble the core-periphery pattern described by the New Economic Geography (NEG). That literature, stemming from Krugman (1991)[24], depicts a circular dynamics in which industries locate where product demand is higher; people move to follow job opportunities and this causes a further increase in product demand. The difference, here, is that we look at public services distribution, instead of industry location[2].

This article draws insights from both the above academic traditions. In few words, we investigate whether people vote with their feet in favour of school access, in a context where rationalisation policies cut on undersized nodes of the school network

[1] Authoritative references in this respect are Banzhaf and Walsh (2008)[2] and Gamper-Rabindran and Timmins (2011)[13] in the field of environmental amenities; Kahn (2007)[23] for transports; Baum-Snow and Lutz (2011)[3], Brunner, Cho and Reback (2012)[8], Epple and Romano(2003)[10] and Hoxby(2000)[21] precisely regarding schools. As for school service, indirect tests looking at housing prices are also noteworthy. A seminal contribution is provided by Black (1999)[5]. More recent ones are Fack and Grenet (2010)[11], Gibbons and Machin (2006)[15] and Gibbons, Machin and Silva (2013)[17].

[2] In that literature, the public sector enters only through the provision of infrastructure to firms (for a review, see Ottaviano, 2008[29]). Residential choices, instead, are confined to wage differentials or neglected, assuming immobile workers. However, core-periphery dynamics involve both households and firms location choices and the public sector is likely to affect either side of the dynamics separately.

to exploit economies of scale.

For our purposes, the Italian context represents an interesting and unique analytical setting. On the one side, despite the traditionally low mobility of Italian people, there is evidence of rising internal migrations, mainly directed towards big urban centres. That trend especially concerns people in working and fertile age: indeed, young adults with children constitute the highest fraction of internal migrants (ISTAT, 2019)[22]. On the other side, in the last decades, austerity measures have produced a deep rationalisation of key services. This process regarded also public education: in this respect, the Gelmini reform of 2008 represents the most decisive and effective push towards the rationalisation of the school network. The objective of that reform was to cut on public spending by eliminating undersized centers of service provision. Indeed, the high per-pupil spending was mainly attributed to the excessively dispersed configuration of the Italian schooling system.
The link between population dynamics and access to services lies at the core of the National Strategy for Inner Areas (SNAI)[28], which constitutes the largest scale policy effort to address the Italian regional divergence. That cohesion program starts from the verification that Italy has progressively evolved from a polycentric configuration towards a core-periphery structure, where people, opportunities and services are always more concentrated in few attractive centres, to the detriment of the rest of the country[3].

The high policy relevance of the issue is therefore explicitly recognised at national level. However, to the best of our knowledge, there has been no systematic assessment of the population effects of rationalisation policies in the field of public services. In this article, we try to fill this gap, investigating whether service cuts to undersized service centres have the effect of further promoting population decline. As for the research question, the closest contribution to our work is provided for the British context by Gibbons, Heblich and Pinchbeck (2018)[14]. In that article, the authors

[3] For an historical perspective on Italy's territorial issues, see Borghi (2017)[6].

assess the population effect of rail cuts directed towards non-profitable lines and motivated by the dire financial situation of the British rail transport industry over the period 1951-1981. In the present work, we focus on school service and, in particular, on primary school closures. The main reason is that we believe school service is especially relevant for residential choices. Indeed, households with children have a daily need of school services. Moreover, with specific regard to primary schools, these are mandatory, lasts five years and children of that age still depend on their parents for daily commuting. Conversely, preschool and lower secondary school lack some of these features and therefore seem less suitable for our analysis. Finally, the Gelmini reform offers an interesting analytical setting and it has mostly intervened on preschools and primary schools, which displayed a more dispersed configuration.

Our analysis entails a fundamental empirical challenge in that school closures precisely concern municipalities experiencing negative population trends. This clearly create an endogeneity issue, which we address through two complementary identification strategies. Firstly, drawing from Gibbons, Heblich and Pinchbeck (2018)[14], we investigate the overall population effect of experiencing closures, flexibly controlling for historical population pre-trends. Secondly, we use the information on population pre-trends to preprocess data using the entropy balancing algorithm; then, we estimate the dynamic effect of closures on current and forward population outcomes.

Our preliminary findings suggest that school closures negatively affect population dynamics on top and beyond preexisting trends. As expected, the effect is highly significant for the sub-population of residents below fifty years old, which represent the potential recipients of school services. Conversely, placebo tests on elder population display non-significant estimates. Moreover, we find stronger impacts in municipalities which, since the beginning, were poorly endowed with schools; namely, closures contribute more to population decline where there are few schools in absolute values. Finally, looking at the dynamics of the effect, estimated coefficients increase in magnitude over the first years after school closure and remain significant up until eight years from the cut.

The paper is structured as follows: in the next section (2), we describe the institutional context of the Italian schooling system and the key elements of the recent reform process; in section 3, we present the dataset and some descriptive statistics; then (section 4), we outline the two empirical strategies adopted and (section 5) the main results obtained; finally, section 6 reports few preliminary conclusions.

2 Institutional context

Despite recent trends towards decentralisation, the Italian schooling system still displays a considerably centralised and unitary configuration[4]. The national government has authority over the general norms in the field of education, including the definition of school programs, quality standards and their evaluation (MEF, 2012)[25]. Moreover, it regulates and directly manages the recruitment and payment of the schooling personnel, which constitutes the largest component of the expenditure for education[5].

Our focus is on the first educational cycle, which comprehends preschool (*scuola dell'infanzia*), primary school (*scuola elementare*) and lower secondary school (*scuola secondaria di primo grado*). Primary school and lower secondary school are mandatory, whereas preschool is not. This latter, however, came to be perceived as an important educational service, so that by 2010 the participation into preschool reached the 100% of the admissible population (MIUR, 2010)[27]. The vast majority of pupils of the relative schooling ages attend public schools[6]. These are mainly managed by the central government, with the exception of some residual municipal preschools

[4]For an historical perspective on Italian school design and achievements, see Checchi et al. (2006)[9]. In more recent years, in line with the trend towards "regionalisation" of the overall public system, some jurisdictional powers have been transferred from the central government to local authorities. To such process, contributed since the '90s the institution of school autonomy and the 2001 reform of the Italian Constitution.

[5]In all OECD countries, school expenditure is for the 90% current expenditure. The four fifths of that amount consists of personnel wages. Compared to other OECD countries, in Italy the unbalanced expenditure distribution in favour of school personnel is even more marked (MEF-MPI, 2007)[26].

[6]More than 70% of pupils enrolled in preschools attend public schools. The percentage rises to over 90% for primary and lower secondary education (ISTAT data, available at: http://dati.istat.it/).

and schools of any order in the autonomous Regions of Trentino-Alto Adige and Valle d'Aosta[7].

Over the first educational cycle, there is basically no tracking in schooling offer, so that in principle school quality should be equalised. In the Italian context, indeed, the strongest evidence of sorting across schools is found at higher secondary school (*scuola secondaria di secondo grado*)[8]; whereas it seems not particularly relevant for the first educational cycle. Moreover, for these orders of school, choice is made by parents, who have to combine work and family needs. As a consequence, house-school commuting time becomes particularly relevant in orienting parents' choice. The Italian system, indeed, allows for school choice. Parents can enrol children in their preferred school, even in municipalities different from the one they reside in. However, if the chosen school happens to be oversubscribed, the priority is given to pupils residing in the school's catchment area[9]. In conclusion, at least for the first educational cycle, residence and school choice are not completely independent. In particular, it seems plausible that households take into account distance to school when evaluating residence decisions.

The distribution of schooling services relates to two normative institutions: the criteria for class formation and the guidelines for the organisation of school network.

Class formation is regulated at national level by the D.P.R. n.81/2009, part of the Gelmini reform, which substituted the previous D.P.R. n.331/1998. For what concerns, instead, the organisation of school network, guidelines are provided by each Italian Region for its territory and contain directives on activation, suppression and

[7]From our elaborations on MIUR data, excluding Trentino and Valle d'Aosta, public preschools managed by the national government range from 75% to almost 100% across Italian regions. In Trentino and Valle d'Aosta schools are under the jurisdiction of the autonomous Provinces.

[8]For the second educational cycle, the differentiation between vocational (*istituti tecnici-professionali*) and high schools (*licei*) leads to self-selection of students according to their family background. On school tracking and social segregation, see Bertola and Checchi (2004)[4] and also Brunello and Checchi (2007)[7].

[9]Each school institution has to declare its admission criteria in case of over-subscription. On admission rules, see MIUR enrolment circular n. 22994 for school year 2020-21 at the link https://www.miur.gov.it/web/guest/-/iscrizioni-alle-scuole-dell-infanzia-e-alle-scuole-di-ogni-ordine-e-grado-anno-scolastico-2020-2021. At least since 2008, these rules have been the same.

merger of school complexes[10]. According to such guidelines, Provinces elaborates sizing plans, negotiating variations with school institutions and local authorities. If consistent with regional guidelines, provincial plans are approved and all together constitute the annual school sizing regional plan (*Piano di dimensionamento scolastico regionale*).

It is important to note that regional authorities are constrained in their decision making to the workforce assigned to each Region by the central government. Only within those limits, Regions can organise the school network in their territory. In fact, the truly binding constraint to class and school activation is represented by the scarcity of teachers and janitors, which are the more valuable and costly resource of the schooling system[11].

In this framework, the single school has little control over activations and suppressions. Indeed, school workforce is assigned on the basis of student enrolments (*organico di diritto*) and then adjusted to cover particular and transitory needs, so to get the effective personnel for the school year (*organico di fatto*). Therefore, despite the formal decentralisation of authority, central reforms affect the organisation of school network. For this reason, we focus on the reform process realised under the Education Minister Gelmini, which constitutes the most recent and decisive effort of school network's rationalisation.

Rationalisation policies: the Gelmini Reform

The Italian school system was historically characterised by a high degree of territorial dispersion, following the polycentric distribution of the Italian population. However, since the second postwar, Italian demography has considerably changed, enlarging the population of biggest cities to the detriment of more peripheral areas, nowadays at risk of depopulation. This phenomenon was mainly due to changes in the

[10] The normative references in this regard are D.L. n.112/1998 and the reformed title V of the Constitution.

[11] Those resources are financed by the national government, whereas local authorities - for the first educational cycle, municipalities - are in charge of school buildings and finance their maintenance.

economic structure of the country, which directed residential mobility towards cities where there existed better job opportunities. In addition, since the '90s policies of rationalisation started to be implemented in the field of public services. This kind of policies interested also public education, which in recent decades experienced continuous expenditure cuts. In this regard, the last noticeable turn occurred after the 2008 crisis with the Gelmini reform, which produced a relevant contraction of the school network, both in terms of school complexes and classes activated (MIUR , 2010)[27]. Indeed, by 2008 rationalisation policies had mostly intervened on reducing autonomous school institutions (therefore, cutting on school directors), but they had not strongly affected the distribution of school complexes. The territorial fragmentation of school complexes and the limited class size were identified as the main reasons for the high per-pupil expenditure compared to OECD countries[12].

The reform process started with law n. 133 of August 2008 art. 64, which established the increase of one percentage point of the pupils-teacher ratio and the elaboration of a strategy plan (*"piano programmatico"*) to achieve a "more rational use of human and material resources" in the schooling system, from which should derive public savings for 8 billions of euros by 2012. In the resulting strategy plan of the Ministry of Education it is written that "by almost one decade the school network has been substantially unchanged in its school complexes (*plessi*) and institutions (*istituzioni scolastiche*); this, despite the demographic dynamics which have emptied or excessively filled school classrooms, and made difficult or superfluous the management and coordination of schools". In the same document, the Ministry declares the need to eliminate undersized school complexes; despite clarifying that "the institution, suppression or merge of schools is under the jurisdiction of regions and local authorities according to D.L. n.112/1998 and to the reformed title V of the Constitution, on the basis of sizing criteria defined by MIUR with the specific Regulation planned by art.

[12] These arguments are maintained by Fontana (2008)[12] and supported by the analysis of MEF-MPI (2007)[26]. This report states that the 2007 school network was historically inherited, since the consolidation process - by that time - had not reduced the number of school complexes.

64"[13]. The cited Regulation is the already mentioned D.P.R. n. 81/2009, which revised the numerical limits to form initial classes, ordered the increase in pupils/class ratio and allowed for exceptions only in case of effective increase in schooling population[14]. That Regulation still constitutes the normative reference for class formation in all regional guidelines for the elaboration of sizing plans. In addition, some Regions establishes numerical criteria for the activation or suppression of school complexes, whereas others simply formulate general norms for the organisation of the school network. Among these, we always find directives towards a more rational distribution of school complexes, to be achieved through the suppression of the undersized ones. In those decisions, it is generally recommended to consider the distance to further schools and the overall characteristics of the territory (in particular, in case of mountain areas or small islands)[15].

3 Data and Descriptives

Data on schools have been provided to us by the Italian Ministry of Education (MIUR - *Ufficio Gestione Patrimonio Informativo e Statistica*) and refer to preschools, primary and lower secondary schools, over the period 2009-2019. Those administrative data cover the entire population of Italian schools and represent the most official source of information about the Italian schooling system. At this preliminary stage, we exploit information only regarding public schools, which essentially consist of government schools. We decided to exclude from our analysis the Regions Trentino-Alto Adige and Valle d'Aosta, due to some data issues and, more importantly, because in those Regions school policy and funding are under the direct jurisdiction of the autonomous Provinces and therefore follow considerably different dynamics compared to the rest of the country. At the present stage, we only look at primary school

[13] *Schema di Piano Programmatico del Ministero dell'Istruzione, dell'Università e della Ricerca di concerto col Ministero dell'Economia e delle Finanze.*

[14] *Norme per la riorganizzazione della rete scolastica e il razionale ed efficace utilizzo delle risorse umane nella scuola, ai sensi art. 64, comma 4, del decreto legge 25 giugno 2008 n.112, convertito, con modificazioni, dalla legge 6 agosto 2008, n.133.*

[15] Guidelines for regional sizing plans can be found on Regions' websites.

closures. This choice is motivated by the fact that primary school is mandatory, lasts five years, has a considerably dispersed territorial distribution and children of that age still depend on their parents for daily commuting. Conversely, preschool and lower secondary school lack some of these features and therefore seem less suitable for our analysis.

To identify school closures, we exploit the universal coverage of our data and the information about the exact location of each school (street address). Indeed, each school complex is identified by an administrative code (*codice meccanografico*), whose disappearance from the dataset signals the closure of the school. Since there have been changes of codes for administrative reasons, we employ the ArcGIS software to precisely geolocate schools, so to avoid the risk of overstating school closures. By doing so, primary school closures over the period 2009-2019 distributes across Italian Regions as follows:

Table 1: Primary school closures by Region (2009-2019)

Region	Closures	Percent
Abruzzo	66	5.99
Basilicata	28	2.54
Calabria	126	11.44
Campania	169	15.35
Emilia-Romagna	30	2.72
Friuli-Venezia Giulia	16	1.45
Lazio	68	6.18
Liguria	27	2.45
Lombardia	94	8.54
Marche	14	1.27
Molise	30	2.72
Piemonte	67	6.09
Puglia	53	4.81
Sardegna	58	5.27
Sicilia	133	12.08
Toscana	29	2.63
Umbria	18	1.63
Veneto	75	6.81
Total	1,101	100.00

Note that data regarding school year 2009-2010 are associated to outcome variables referring to 2010, and similarly for subsequent years. Moreover, if a given school is absent for school year 2010-2011, we consider the related observational unit 'treated' in 2010. The reason behind this choice is that school sizing plans are approved by the end of December of the year before. Therefore, if the school complex will not be activated for school year 2010-2011, the announcement is given by the beginning of 2010, and we take that moment as the timing of treatment.

For what concerns, instead, the outcome variable, we collected data on residential population at municipal level from ISTAT - Residential population by sex, age and marital status[16]. These are administrative data, which provide yearly information on residents in each municipality. We refer to that data source for the period 2009-2019, for which we have school data. Regarding instead historical population, employed in the identification strategy, we use data from the censuses of 1971, 1981, 1991, 2001. In most of our study, we use as outcome variable residential population below fifty years old, so to restrict the analysis to the sub-population which is more likely to benefit from school services. That information is obtained by summing residential population by age classes. Moreover, from ISTAT - Main geographical statistics on municipalities[17], we got further control variables, as land area in square kms, municipal centre's elevation and an indicator of whether a municipality is an island one. Always from ISTAT we draw data on local labour systems, so to control for labour market conditions[18]. At this stage, we confine our analysis to municipalities that have not undergone processes of administrative reorganisation, so that we can easily track the municipal unit by its ISTAT code over the entire period considered. After that sample reduction, we are left with 6,227 over 7,547 Italian municipalities, excluding Trentino-Alto Adige and Valle d'Aosta.

More than the 40% of municipalities in our sample has just one primary school and

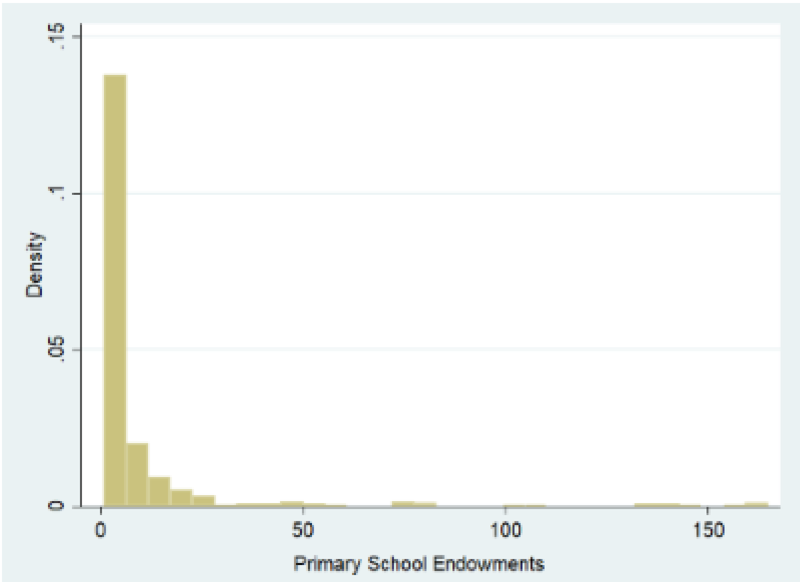
[16]These information are available at ISTAT warehouse http://dati.istat.it/ from 2012; whereas, for more historical data, at http://demo.istat.it/archivio.html.

[17]Data are available at https://www.istat.it/it/archivio/156224.

[18]Data can be found at https://www.istat.it/it/informazioni-territoriali-e-cartografiche/sistemi-locali-del-lavoro.

closures occur for the 60% in municipalities with at most three primary schools (more than 30% has just one). Moreover, where closures take place, above the 60% of municipalities are endowed with at most three preschools and one lower secondary school. To provide a visual intuition, we report a graphical illustration of the distribution of primary school endowments in municipalities experiencing closures[19].

Figure 1: Number of primary schools in municipalities where closures took place



That evidence is consistent with the small size of most Italian municipalities and with the dispersed configuration of the schooling system. At the same time, the purpose of rationalisation policies is precisely to overcome such dispersed configuration. Therefore, it is no surprise that service cuts largely concern municipalities with few schools in absolute values. Since our analysis is at municipal level, we can exploit this setting to investigate the demographic impact of school closures where we can effectively associate service centres and geographical units observed. In other words, for municipalities with few schools, catchment areas basically coincide with municipal boundaries. Therefore, residents (below a certain age) of those municipalities are indeed the potential recipients of a given school's services and they represent the population truly affected by school closure. Following these considerations, in most

[19]Analogous graphs for preschools and lower secondary schools are reported in the Appendix.

part of our analysis we focus on municipalities endowed with one primary school at the beginning of the period considered, investigating the effect of having that single school closed. This choice clearly reduces the amount of 'treated' observations; however, at this stage we prefer to adopt a conservative approach and to look for an effect where it can be more clearly identified. Moreover, we can anyway rely on a large number of potential controls for each treated observation. Indeed, if we confine the analysis to municipalities with a single primary school in 2010, we are left with 344 municipalities with primary school closures over a total of 3,916. That high controls-to-treated ratio will result particularly useful for the counterfactual analysis; moreover, it has to be noted that a similar sample restriction results in closely comparable observational units.

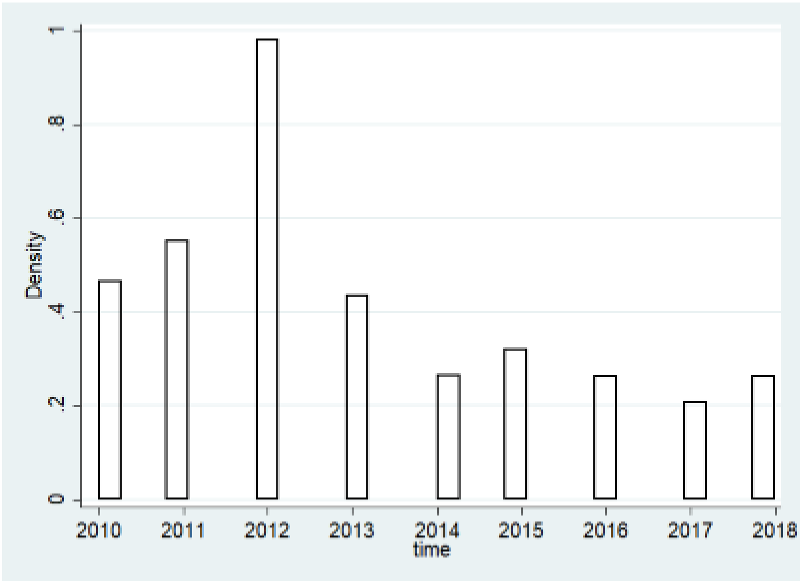
Table 2: Single-Primary-School Municipalities by Region

Region	Not Closure	Closure	Total
Abruzzo	156	32	188
Basilicata	86	10	96
Calabria	143	29	172
Campania	260	35	295
Emilia-Romagna	136	2	138
Friuli-Venezia Giulia	126	6	132
Lazio	201	27	228
Liguria	124	11	135
Lombardia	776	53	829
Marche	112	4	116
Molise	78	20	98
Piemonte	565	38	603
Puglia	104	5	109
Sardegna	135	46	181
Sicilia	164	11	175
Toscana	97	2	99
Umbria	48	2	50
Veneto	261	11	272
Total	3,572	344	3,916

When considering the panel dimension of our data, the timing of closures becomes also relevant. Not surprisingly, we observe a sharp increase in primary school closures

in the first three years observed, which correspond to school years 09-10, 10-11 and 11-12. That period represents the horizon indicated by the Gelmini reform to rise 8 billions euros of public savings. As explained, Gelmini reform intervened mostly on criteria for class formation, while school network is under Regions' jurisdiction. However, those reformed criteria entered the guidelines for school sizing of all Regions and constituted a more stringent constraint to the activation of school complexes.

Figure 2: Primary School Closures by year



If we focus on the reduced sample of municipalities with a single primary school in 2010, the time distribution of closures appears less skewed. However, we can still note a concentration of cases in the two first years observed (see the graph in the Appendix).

4 Empirical Strategy

Dealing with the demographic impact of rationalisation policies entails a fundamental empirical challenge: by the same definition of rationalisation policies, service cuts are directed towards locations where there are few potential recipients, which makes the

treatment highly endogenous. To address that endogeneity issue, we exploit population pre-trends, so to compare municipalities which followed similar demographic paths in the pre-treatment period. In the following, we present two complementary strategies, both accounting for population pre-trends.

First of all, drawing from Gibbons, Heblich and Pinchbeck (2018)[14], we regress the difference in log-population over the whole period on an indicator for school closure. We include logs and square logs of historical population as controls, to flexibly account for population evolution. We then add controls for municipal endowments of preschool and lower secondary school at the beginning of the period and their variations during the span considered. Indeed, different school orders can be though as complementary services, which can jointly enter households' residence decisions. Moreover, we further include control variables for municipal characteristics, such as centre's elevation, land area and an indicator for island municipalities. Finally, regional and local labour system (SLL) dummies are added to the regression. The inclusion of those fixed effects allows to compare municipalities exposed to the same regional school policies and facing homogeneous labour market conditions[20].

Essentially, we estimate the following regression model:

$$lnP_{i2019} - lnP_{i2010} = \beta treat_i + \mathbf{x}_i'\gamma + \epsilon_i \qquad (1)$$

where the $\mathbf{x}$ vector progressively includes the controls just described. The baseline estimation restricts to municipalities with one primary school in 2010, so that the treatment variable takes values $\{0, 1\}$ depending on whether the single primary school has been maintained or closed over the period considered. Despite the simple specification, we argue that including a fairly long series of lagged outcomes can account for any pre-treatment observable and unobservable confounder. Here, in fact, we are

[20]We opt for *regional* fixed effects since Regions have authority on the guidelines for school sizing plans, which constitutes a relevant source of heterogeneity in school policies. Regarding SLL, these represent the most accurate definition of homogeneous labour market areas. Their inclusion allows to clean the estimated population effect from mobility due to labour opportunities (for a deeper discussion, see Hoxby (2000)[21] or Brunner, Cho and Reback (2012)[8]). Given the period considered, we use the SLL definition of 2011.

comparing municipalities endowed with one primary school in 2010, which followed similar population pre-trends, and differ with respect to the fortunes of that single primary school. Therefore, the β coefficient is really able to capture whether school closure has an effect in terms of depopulation on top and beyond historical trends.
Adopting the same identification strategy, we repeat the estimation for a larger sample, which includes municipalities with at most four primary schools in 2010. In that extension, we add a further control variable for initial 'school density', computed as the ratio between resident children in primary school age and number of primary schools in 2010. We run the regression for the overall sample and, again, restricting to different levels of primary school endowments in 2010. That larger sample estimation contributes to disentangle absolute and relative effects of school closures. Alternatively, we want to investigate whether population effects are driven by the absence of a primary school or by its relative scarcity when compared to the number of potential recipients.
In the baseline specification, we use as dependent variable the population below fifty years old, which is more likely affected by school closures. However, we also conduce a placebo test using as dependent variable the resident population above fifty years old.

The other strategy we present is based on entropy balancing as matching technique prior to parametric estimation. Here, instead of controlling for pre-trends, we use those information to balance treated and control distributions before proceeding with parametric regression. This empirical strategy present numerous advantages. Indeed, as stated in Ho et al. (2007)[20], preprocessing data through matching can help to solve issues of model dependence, related to the difficulty of correctly specify the parametric model. At the same time, extending matching to parametric estimation - instead of a simple mean comparison of matched data - makes the identification more precise in case covariates balancing were not perfect. Therefore, the combination of non-parametric and parametric methods reduces the concerns arising from

the separate application of each specific technique. Essentially, we exploit information deriving from pre-treatment covariates and, mainly, lagged outcomes to balance treated and control distributions, and then run an OLS regression including balancing weights to compute ATT. As matching technique, we choose to employ entropy balancing (EB)[21]. That method is based on the selection of weights (w_i) for each observation in the control group that minimise an entropy distance metric $H(w)$:

$$\min_{w_i} H(w) = \sum_{[i]T_i=0]} w_i \log\left(w_i/q_i\right) \tag{2}$$

where T_i is an indicator taking value 1 for treated and 0 for controls and $q_i = 1/n_0$ is a base weight. Minimisation is subject to a set of balance constraints imposed on the covariates moments, depending on whether the researcher wants to balance only means or also variance and skewedness of the distributions[22]. In that procedure, the availability of a large sample of potential controls enable us to consider many confounding variables, which makes our matching exercise more robust.

We balance pre-treatment covariates including the second and third order polynomial of all variables except dummies. By doing this, EB on first moment also balances variance and skewedness of all first order variables. As balancing covariates, we select population pre-trends and some other variables for relevant municipality characteristics, such as elevation, land area and an indicator for island municipality. Moreover, we balance on time and on a set of regional dummies, so to compare municipalities close in time and space. Thereafter, in the OLS regression we include as further controls the endowment of preschools and lower secondary schools over the period, on top of balancing covariates and weights, and we cluster by municipal identifier.

In this second estimation strategy, we are really able to exploit the longitudinal nature

[21]For more details of EB method, see Hainmueller (2012)[18] and Hainmueller and Xu (2013)[19].

[22]Compared to alternative matching techniques, EB overcomes the tedious procedure of "'manually' search[ing] for a suitable weighting that balances the covariate distributions" which generally results in poor overlap, in practice (Hainmueller and Xu, 2013)[19]. Conversely, EB is based on a reweighting scheme that can satisfy a large set of constraints which involve exact matching on distribution moments in the treatment and reweighted control group. Therefore, covariate balance directly derives from the weight function employed to adjust the control group.

of our data, so to investigate the dynamics of treatment effect. In this specification, indeed, we define treatment as an entry status; that is, treatment dummy takes value 1 only for the municipality-year combination where closure occurs and it is missing in pre-post years. As controls, we use only municipalities that have never been treated in the period considered. Then, we run the OLS regression for current and forward values of the outcome variable. In this way, we also observe the time profile of treatment effect in the years following school closure.

This empirical strategy is applied only to the preferred sub-sample of municipalities with a single primary school in 2010. In the baseline specification, we use as dependent variable the population below fifty years old. Moreover, we repeat the EB exercise and the parametric estimation using as dependent variable the resident population above fifty years old, as placebo test.

5 Results

In table 3 and 4 we present estimates obtained adopting the first empirical strategy. Since we collapse our panel dataset to a cross-section, treatment is a dummy variable taking value 1 if the municipality has ever experienced a primary school closure during the period observed. The outcome variable of the estimates presented in table 3 is the logarithm of resident population below fifty years old in 2019. Note that these estimates are exactly the same as β in equation (1), since we simply include among regressors initial population instead of expressing the outcome in differential terms. As explained in the above section, we control for pre-trends, including second order terms, as for 1971. Table 3 reports estimates for the progressive inclusion of controls, focusing on the restricted sub-sample of municipalities endowed with a single primary school in 2010. Column 1 only includes population pre-trends; column 2 controls also for initial endowments of preschool and lower secondary school, together with their variations in the period observed; column 3 adds municipal characteristics and col-

umn 4 and 5, respectively, regional and SLL fixed effects. The estimated coefficient of school closure seems robust to the progressive inclusion of controls. The magnitude of the effect stays around a 2.5% decrease in young population where the single primary school has been closed.

In table 4 we enlarge the sample, including municipalities with up to four primary schools in 2010. The reported coefficient refer to different samples: column 1 results relate to municipalities with a single primary school in 2010; column 2 to the largest sample of municipalities with up to four primary schools; column 3 to the sample of municipalities with one or two schools; column 4 to that of more endowed municipalities, with three or four primary schools. Not surprisingly, the effect is stronger for municipalities which since the beginning were poorly endowed with school services and looses its significance if we focus on observational units fairly provided with primary schools[23]. Moreover, in the specifications of column 2 to 4, we also included among regressors the ratio between children in primary school age and number of primary schools, as a measure of 'school density'. That variable displays a zero coefficient in all regressions. These findings seems to suggest that the absolute dimension of the impact is more relevant than its relative one. In other words, population is negatively affected by the absence of school services more than by their relative scarcity.

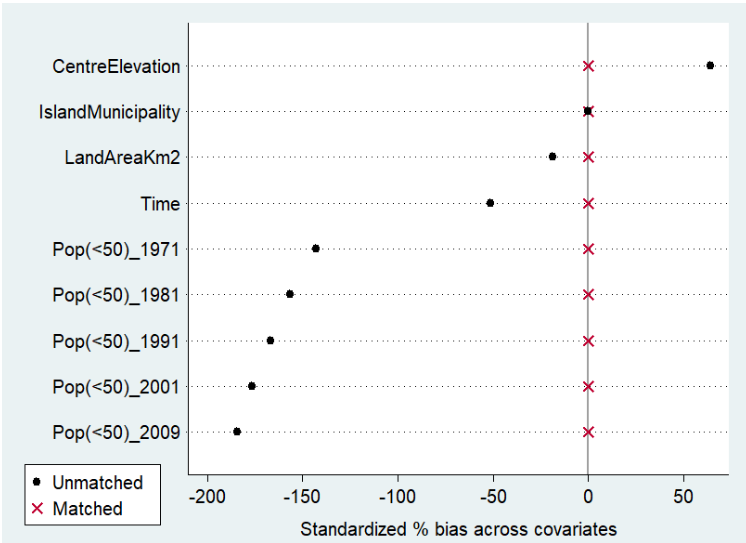
Finally, table 5 reports the coefficients estimated using resident population above fifty years old as dependent variable. That placebo test shows no effect of school closure on elder population for all samples considered. That evidence provides robustness to our results, since it is consistent with *a priori* knowledge about the sub-population of interest.

We then present results as for the second empirical strategy. Here, we truly exploit the longitudinal nature of our data, which enable us to investigate the temporal dynamics of the effect. In the following, estimates refer to the restricted sample of municipalities with a single primary school in 2010.

[23] As can be noticed, in this latter estimation, the number of observations reduces considerably, while standard errors increase. Therefore, this last result should be considered with caution.

First of all, we start by showing the need for covariates balancing. As can be seen from the pstest graph, population pre-trends are significantly unbalanced between treated and controls prior to EB. More specifically, municipalities experiencing closures display more negative population dynamics compared to those who have maintained their single primary school. This evidence is a further confirmation of the fact that rationalisation policies intervened where service demand was shrinking, making inefficient its provision. Moreover, compared to other key covariates, such as land area or elevation, population pre-trends display higher standardised percentage bias among distributions. This finding supports our choice to build the identification strategies on population pre-trends, which in fact represent the main source of selection bias.

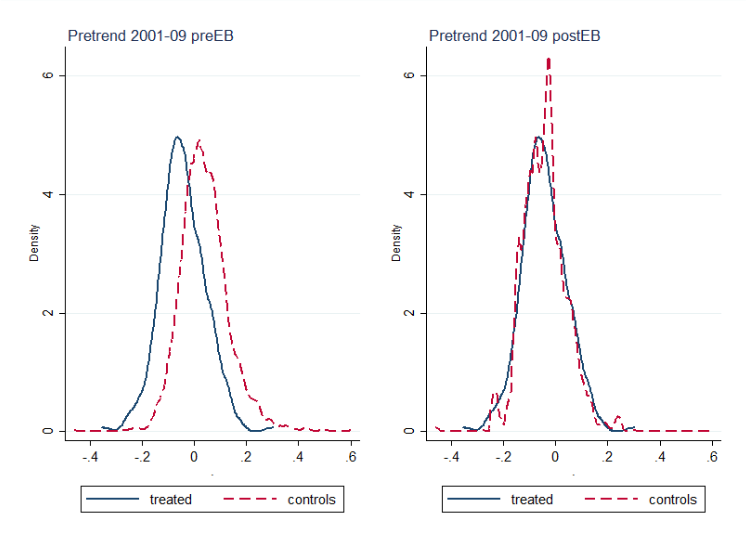
Figure 3: Covariates (un)balancing (pre)post EB



By construction, EB achieves a very good overlap between treated and control distributions, since the algorithm extracts balancing weights directly from the data. The great reduction in means misalignment between distributions is already clear from the above pstest graph. However, in the matching procedure, we included the second and third order terms of relevant covariates, so to balance also variance and skewedness of the distributions. In the following, we present one pre-post EB graphical evidence of the overlap. Kernel distributions refer to the difference in log-population between

2001 and 2009, which displayed the highest bias according to the pstest.

Figure 4: Population trend 2001-2009 pre-post EB



Once data have been preprocessed, we can run an OLS regression including entropy balancing weights and investigate the temporal dynamic of the effect. In table 6 are reported the estimated coefficients for the treatment variable on current log-population below fifty years old (Column 1) and on its forward values (Column 2 to 10). Closures mainly occur in the first years considered, which allows us to assess the *ex-post* consequences of school cuts for a considerably large sample. As it can be noticed, the treatment effect intensifies in the subsequent years after (single) school closure, reaching a peak of 4.3% decrease in young population. The insignificant coefficient in the ninth year from school cut is little informative, since observations reduce considerably and the standard error is high. However, this evidence suggests that school closures have an immediate and lasting effect on population dynamics. Considering that these are municipalities with a single primary school, that immediate response is not particularly surprising: potential recipients must find a substitute school already for the following school year and, if residence-school distance is significant, they could plausibly consider to move. The increasing effect in the first years after closure could itself account for moving costs, which are likely to delay the full arising of the impact.

Moreover, the persistence of the effect after eight years from the closure calls for a non-transitory shock, able to feed itself and to drive the municipality towards a new peripheral equilibrium; in line with the cumulative core-periphery process described by the New Economic Geography.

Finally, in table 7 we present the OLS coefficients using as dependent variable the resident population above fifty years old. For this placebo test, we repeat the entropy balancing procedure using the same covariates as before, with population pre-trends referring to elder residents. OLS estimates are largely insignificant, with few exceptions (at 0.10 significance) which are plausibly due to statistical noise. Again, these findings support our claim that we are really capturing the effect of school closure on top of preexisting population dynamics. Indeed, treated municipalities display more negative pre-trends even regarding elder population (see the pstest graph in the Appendix); however, the population dynamics of that subgroup result unaffected by the treatment, as suggested by prior knowledge of the phenomenon.

6 Preliminary Conclusions

The analysis presented is still very preliminary: we confine the investigation to rather specific contexts and, more importantly, we have not produced yet sufficient robustness tests. Some extensions are already in our agenda: for example, we plan to control for substitute services, such as private schools, and to include some local measures of other services endowment. Despite being at its initial stage, the analysis offers some interesting insights on the demographic impact of rationalisation policies.

First of all, school closures mainly concerned municipalities already experiencing population decline and poorly endowed with schools in absolute terms. Secondly, school cuts negatively affect population dynamics on top and beyond preexisting trends. The magnitude of the effect oscillates around a 2.5% reduction in population below fifty years old, increasing up to 4,3% in the first years after the closure of a single primary school. Indeed, population effect is stronger in municipalities with just one

school at the beginning of the period. For that sub-sample, the population response is immediate and persistent up until the eighth year after the school cut. Finally, placebo tests display a non significant effect on elder population. This supports our claim that we are truly capturing a school closure effect, which adds up to preexisting population trends. These findings provide highly relevant policy implications, namely that closing undersized schools further promotes population decline. That detrimental effect of rationalisation policies arises quite rapidly, intensifies in subsequent years and regards population in fertile and working age, which represents the most valuable resource for areas at risk of depopulation.

References

[1] Alesina, A., Baqir, R. and Hoxby, C., 2004. Political jurisdictions in Heterogeneous Communities. Journal of Political Economy 112(2), 348-396.

[2] Banzhaf, H.S. and Walsh, R.P., 2008. Do People Vote with Their Feet? An Empirical Test of Tiebout. American Economic Review 98(3), 843-863.

[3] Baum-Snow, N. and Lutz, B.F., 2011. School Desegregation, School Choice, and Changes in Residential Location Patterns by Race. American Economic Review 101(7), 3019-46.

[4] Bertola G. and Checchi D., 2004. Sorting and Private Education in Italy. In: Checchi D., Lucifora C. (eds) Education, Training and Labour Market Outcomes in Europe. Palgrave Macmillan, London

[5] Black, S.E., 1999. Do better schools matter? Parental Evaluation of Elementary Schools. Quarterly Journal of Economics 114(2), 577-599.

[6] Borghi, E., 2017. Piccole Italie. Le aree interne e la questione territoriale. Roma: Donzelli Editore.

[7] Brunello, G. and Checchi, D., 2007. Does school tracking affect equality of opportunity? New international evidence, Economic Policy 22(52), 782–861.

[8] Brunner, E.J., Cho, S.W. and Reback, R., 2012. Mobility, housing markets, and schools: Estimating the effects of inter-district choice programs. Journal of Public Economics 96(7-8), 604-614.

[9] Checchi, D., Fiorio C.V. and Leonardi, M., 2006. Sessanta anni di istruzione scolastica in Italia. Rivista di politica economica.

[10] Epple, D., and Romano, R., 2003. Neighborhood Schools, Choice, and the Distribution of Educational Benefits, ch. 7 The Economics of School Choice.

[11] Fack, G. and Grenet, J., 2010. When do better schools raise housing prices? Evidence from Paris public and private schools 94(1-2), 59-77.

[12] Fontana, A., 2008. La rete scolastica: un modello per il numero di studenti per classe nella scuola primaria e scondaria di I grado. XX Conferenza Società Italiana di Economia Pubblica, Università di Pavia.

[13] Gamper-Rabindran, S. and Timmins, C., 2011. Hazardous Waste Cleanup, Neighborhood Gentrification, and Environmental Justice: Evidence from Restricted Access Census Block Data. American Economic Review 101(3), 620-24.

[14] Gibbons, S., Heblich, S. and Pinchbeck, T., 2018. The spatial impacts of a massive rail disinvestment program: the Beeching Axe. CEP Discussion Papers.

[15] Gibbons, S. and Machin, S., 2003. Valuing English primary schools. Journal of Urban Economics 53(2), 197-219.

[16] Gibbons, S. and Machin, S., 2006. Paying for Primary Schools: Admission Constraints, School Popularity or Congestion?. The Economic Journal 116(510), C77-C92.

[17] Gibbons, S., Machin, S. and Silva, O., 2013. Valuing school quality using boundary discontinuities. Journal of Urban Economics 75, 15-28.

[18] Hainmueller, J., 2012. Entropy balancing for causal effects: a multivariate reweighting method to produce balanced samples in observational studies. Political Anal. 20 (1), 25–46.

[19] Hainmueller, J., Xu, Y., 2013. ebalance: A Stata package for entropy balancing. J. Stat. Softw. 54 (Issue 7), 1–18.

[20] Ho, D., Imai, K., King, G., Stuart, E.A., 2007. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. Political Anal. 15, 99–236.

[21] Hoxby, C. M. (2000). Does competition among public schools benefit students and taxpayers?. American Economic Review, 90(5), 1209-1238.

[22] ISTAT, 2019. Iscrizioni e cancellazioni anagrafiche della popolazione residente anno 2018. Statistiche report.

[23] Kahn, ME., 2007. Gentrification Trends in New Transit-Oriented Communities: Evidence from 14 Cities That Expanded and Built Rail Transit Systems. Real Estate Economics 35(2), 155-182.

[24] Krugman, P., 1991. Increasing Returns and Economic Geography. Journal of Political Economy 99(3), 483-499.

[25] MEF, 2012. Analisi dell'efficienza delle scuole italiane rispetto agli apprendimenti degli studenti, differenze territoriali e possibili determinanti. Studi e Pubblicazioni del Dipartimento della Ragioneria Generale dello Stato.

[26] MEF-MPI, 2007. Quaderno bianco sulla scuola.

[27] MIUR, 2010. La scuola statale: sintesi dei dati a.s. 2009/2010.

[28] MUVAL, 2014. Strategia Nazionale per le Aree Interne: definizione, obietizzi, strumenti e governance. materiali UVAL Numero 31, 2014.

[29] Ottaviano, G.I., 2008. Infrastructure and economic geography: An overview of theory and evidence. EIB Papers 13(2), 8-35.

[30] Tiebout, C.M., 1956. A pure theory of local expenditures. Journal of Political Economy 64, 416-424.

[31] Urquiola, M., 2005. Does School Choice Lead to Sorting? Evidence from Tiebout Variation. American Economic Review 95(4), 1310-1326.

Table 3: Overall population effect of school closure in municipalities endowed with a single primary school

Population below 50	(1)	(2)	(3)	(4)	(5)
Single school closure	-0.027***	-0.024***	-0.025***	-0.022***	-0.024***
	(0.0059)	(0.0059)	(0.0059)	(0.0059)	(0.0059)
Population pre-trends	yes	yes	yes	yes	yes
Other school endowments	no	yes	yes	yes	yes
Municipal characteristics	no	no	yes	yes	yes
Regional fe	no	no	no	yes	yes
SLL fe	no	no	no	no	yes
R-squared	0.992	0.992	0.992	0.993	0.993
N	3837	3837	3837	3837	3765

Note: Robust standard errors in parentheses * p<0.10, ** p<0.05, *** p<0.01.
Dependent variable is log-population below fifty years old in 2019. All columns includes as controls initial log-population and historical trends since 1971. Second order terms are included to allow for non-linearities in population evolution. We then add controls for other school order endowments in 2010, their variation in the period considered, and a measure of school service density relative to recipients' population. From column 3 onward we control also for municipal characteristics, such as centre's elevation, land area and island municipality indicator. In the last two columns are included, respectively, regional and local labour market (SLL) dummies.

Table 4: Overall population effect of school closure: municipalities with up to four primary schools

Population below 50	(1)	(2)	(3)	(4)
	One school	Up to 4	Up to 2	More than 2
School closure	-0.024***	-0.015***	-0.019***	-0.002
	(0.0059)	(0.0034)	(0.0041)	(0.0080)
Population pre-trends	yes	yes	yes	yes
Other school endowments	yes	yes	yes	yes
Municipal characteristics	yes	yes	yes	yes
Regional fe	yes	yes	yes	yes
SLL fe	yes	yes	yes	yes
R-squared	0.993	0.996	0.995	0.992
N	3765	5367	4658	500

Note: Robust standard errors in parentheses * p<0.10, ** p<0.05, *** p<0.01.
Dependent variable is log-population below fifty years old in 2019. Regressors includes initial log-population and historical trends since 1971. Second order terms are included to allow for non-linearities in population evolution. We also control for other school order endowments in 2010, their variation in the period considered, and a measure of school service density relative to recipients' population. Moreover, we add regressors for municipal characteristics, such as centre's elevation, land area and island municipality indicator. Finally, we include regional and local labour market (SLL) dummies. Column 1 estimates refer to municipalities with one single primary school at the beginning of the period; column 2 to those with up to four schools; column 3 and 4 to municipalities endowed with at most or more than two, respectively.

Table 5: Overall population effect of school closure on residents above 50 years old (placebo test): municipalities with up to four primary schools

Population above 50	(1)	(2)	(3)	(4)
	One school	Up to 4	Up to 2	More than 2
School closure	-0.002	-0.001	0.001	-0.006
	(0.0037)	(0.0021)	(0.0025)	(0.0043)
Population pre-trends	yes	yes	yes	yes
Other school endowments	yes	yes	yes	yes
Municipal characteristics	yes	yes	yes	yes
Regional fe	yes	yes	yes	yes
SLL fe	yes	yes	yes	yes
R-squared	0.997	0.998	0.997	0.997
N	3765	5367	4658	500

Note: Robust standard errors in parentheses * p<0.10, ** p<0.05, *** p<0.01. Dependent variable is log-population above fifty years old in 2019. Regressors includes initial log-population and historical trends since 1971. Second order terms are included to allow for non-linearities in population evolution. We also control for other school order endowments in 2010, their variation in the period considered, and a measure of school service density relative to recipients' population. Moreover, we add regressors for municipal characteristics, such as centre's elevation, land area and island municipality indicator. Finally, we include regional and local labour market (SLL) dummies. Column 1 estimates refer to municipalities with one single primary school at the beginning of the period; column 2 to those with up to four schools; column 3 and 4 to municipalities endowed with at most or more than two, respectively.

Table 6: Dynamic effect of single school closure on resident population below 50 years old: OLS estimates after balancing procedure

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Year0	Year1	Year2	Year3	Year4	Year5	Year6	Year7	Year8	Year9
Single school closure	-0.029***	-0.035***	-0.040***	-0.043***	-0.041***	-0.043***	-0.035***	-0.032**	-0.039***	-0.033
	(0.0079)	(0.0084)	(0.0091)	(0.0096)	(0.0105)	(0.0111)	(0.0119)	(0.0134)	(0.0146)	(0.0203)
Balancing covariates	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
EB weights	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Other school endowments	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
R-squared	0.974	0.972	0.971	0.969	0.966	0.963	0.960	0.957	0.956	0.952
N	36543	32905	29252	25600	21937	18266	14608	10950	7284	3630

Note: Robust standard errors in parentheses * $p<0.10$, ** $p<0.05$, *** $p<0.01$.
Dependent variable is log-population below fifty years old, respectively referring to the time of school closure (Column 1) and subsequent years (Column from 2 to 10). Regressors includes balancing weights and controls for other school orders endowments. Balancing covariates are population pre-trends, municipal land area, elevation and an indication for island municipalities. Moreover, we balance on years and on a set of regional dummies. For all variables, except dummies, we include second and third order polynomials.

Table 7: Dynamic effect of single school closure (placebo test): OLS estimates after balancing procedure using as dependent variable population above 50 years old

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Year0	Year1	Year2	Year3	Year4	Year5	Year6	Year7	Year8	Year9
Single school closure	0.011	0.010	0.011	0.014*	0.012	0.014*	0.011	0.009	0.009	-0.004
	(0.0074)	(0.0072)	(0.0071)	(0.0072)	(0.0072)	(0.0074)	(0.0075)	(0.0084)	(0.0091)	(0.0110)
Balancing covariates	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
EB weights	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Other school endowments	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
R-squared	0.972	0.974	0.975	0.976	0.979	0.979	0.979	0.976	0.977	0.980
N	36543	32905	29252	25600	21937	18266	14608	10950	7284	3630

Note: Robust standard errors in parentheses * $p<0.10$, ** $p<0.05$, *** $p<0.01$.
Dependent variable is log-population above fifty years old, respectively referring to the time of school closure (Column 1) and subsequent years (Column from 2 to 10). Regressors includes balancing weights and controls for other school orders endowments. Balancing covariates are population pre-trends, municipal land area, elevation and an indication for island municipalities. Moreover, we balance on years and on a set of regional dummies. For all variables, except dummies, we include second and third order polynomials.

Appendix

Figure 5: Single-Primary-School Municipalities: closures by year

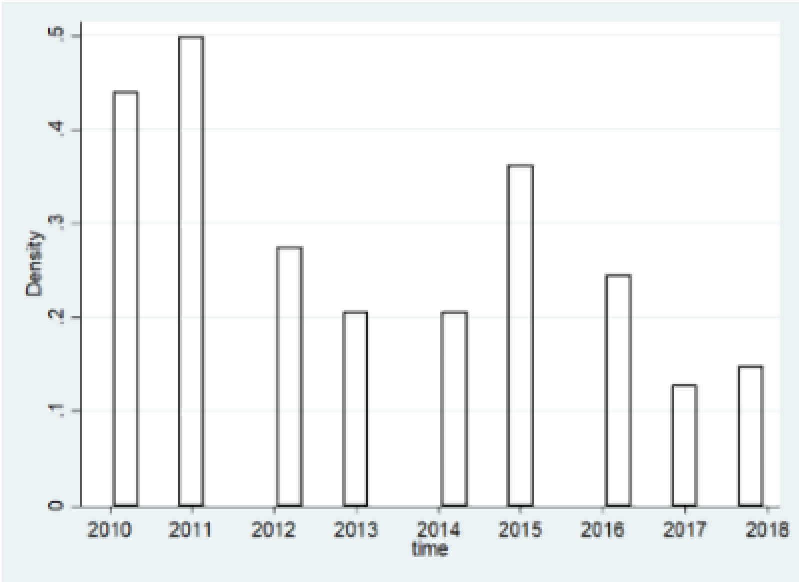


Figure 6: Number of preschools and lower secondary schools in municipalities where closures took place

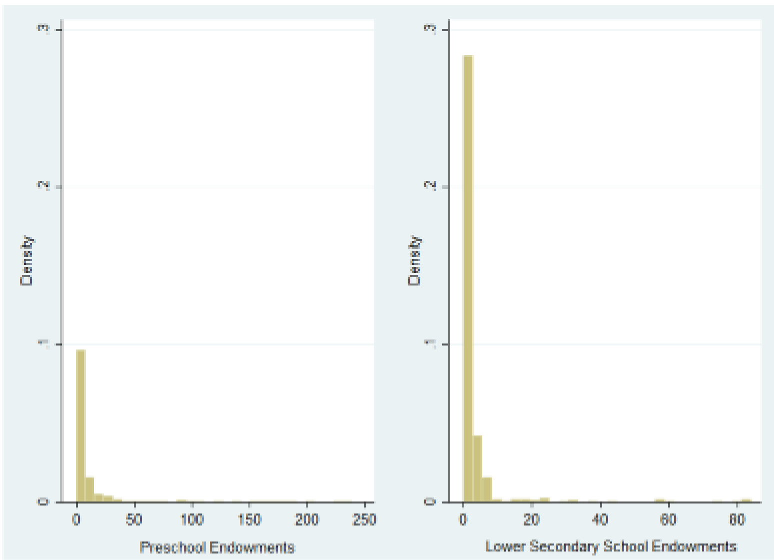


Figure 7: Covariates (un)balancing (pre)post EB for population above 50 years old