Do municipal investments in preschool education affect student achievement?

Giuseppe Rose[1] | Carmela Ciccarelli[1]

Abstract

The importance of public spending on preschool education and interventions for early childhood, such as kindergartens and other dedicated programs, is crucial for the cognitive and social development of children in the early stages of life. These investments are also key to long-term academic outcomes. In this regard, the present study aims to analyze the effect of public spending on preschool education and early childhood services, kindergartens, and minors in Italian municipalities on the INVALSI test results for second-grade students in Italian and Mathematics for the period 2013-2018. Endogeneity issues have been minimized using the instrumental variables method, leveraging the exogenous variation introduced by law no. 215/2012. Increased female representation in municipal councils significantly impacts the overall volume of local public expenditure and its distribution. Overall, the results of our analysis show that an increase in public spending on preschool education and childcare improves INVALSI scores in Mathematics.

JEL classification: I20 I21 I26 D72 J16

KEYWORDS : INVALSI, Gender Quotas, Preference Votes, Instrumental Variables

[1]Department of Economics, Statistics and Finance, University of Calabria, Rende, Italy

Email: giuseppe.rose@unical.it - carmela.ciccarelli@unical.it

1. Introduction

This study aims to examine the influence of local public spending on the development of human capital, specifically in relation to the educational outcomes of Italian elementary school pupils who have been evaluated by the INVALSI tests. The aim of this research is to analyze the impact of two specific categories of local public expenditure, managed by Italian municipalities, on the learning outcomes of primary school students in Italian and mathematics who took the INVALSI tests at GRADE 2 during the period 2013-2018. Local public expenditure is classified by Mission and Programmes[1]. The categories of local public expenditure used in the analysis refer to Pagamenti in conto competenza[2]. We focus on two types of expenditure, Mission 4, Programme 01, Preschool education and Mission 12, Programme 01, Interventions for children, minors and kindergartens[3]. We believe that investments in these spending programmes have a fundamental impact on educational outcomes. The early years of life are crucial for the development of children's cognitive, social and emotional skills. Children who attend quality pre-school programmes are more likely to complete higher education and get a good job. In addition, investments in preschools help families reconcile work and private life, facilitating the participation of parents, especially mothers, in the labour market. In the analysis, a distinction is made between ordinary and extraordinary local public expenditure[4]. Studies confirm that when a government invests in early childhood education, the investments will really make a difference in children's outcomes (Burchinal et al., 2008; Gilliam and Zigler, 2004; Wong, Luo, Zhang, & Rozelle, 2013). OECD data (OECD 2010) showed that public investment in preschool education in Italy was among the lowest in Europe even years ago. Education is widely considered a crucial and potent tool in mitigating economic inequality (Rocha et al., 2023). Since 2005, municipal investment spending has seen a significant decline, both in real terms and in relation to GDP (World Economic Forum 2017). This decline, which is widespread across regions, has highlighted high territorial variability (Viesti 2016); there are significant disparities in educational inputs and outputs between regions (Sibiano et al., 2013). From 2004 to 2010, the ability of municipalities to translate expenditure commitments into actual payments for planned works was greater in municipalities in the North than in those in the Centre and, especially, in the South (Chiades et al. ,2013). The Domestic Stability

[1] General government must adopt a representation of budget data that highlights the purposes of expenditure through a breakdown by missions and programmes defined by Annexes 9 and 10 of Legislative Decree 118/2011, coordinated and supplemented by Legislative Decree 126/2014. Missions are the main functions and strategic objectives that public administrations pursue using the resources at their disposal. Programmes are sets of homogeneous activities, organised to achieve the objectives set in the missions.

[2] Payments on accrual basis, i.e., the amount budgeted for expenditure that represents, with reference to the financial year.

[3] Mission 4, Programme 01, Preschool education includes the management of personnel, registrations, fees, relationships with users, cleaning and sanitizing of environments, and the recording of student attendance; Mission 12, Programme 01, Interventions for children, minors and kindergartens includes expenses for allowances for families with children, maternity, birth contributions, family leave, family allowances, and support for single-parent families or families with disabled children.

[4] Ordinary expenditure concerns current expenses necessary for the regular and continuous operation of the schools; Extraordinary expenses concern non-recurring investments and interventions that often require long-term planning.

Pact has slowed down investment planning and it has severely limited the possibility of reallocating existing expenditure and reducing inefficiencies[5]. According to Corte dei Conti (2012) and Grembi et al., (2016), about two-thirds of municipal expenditures are considered rigid. The extent of structural budget rigidity significantly undermines the financial autonomy of a municipality. Since salary expenditure is the main source of budget rigidity, difficulties in redundancies prevent municipalities from passing on austerity cuts to public sector employees (Pavese et al., 2024). This article emphasises the importance of investment in preschool education and childcare, particularly in Italian municipalities, where public spending is low and exacerbated by severe fiscal consolidation policies (Dipartimento Per Le Politiche Della Famiglia 2020). Reduced budgetary allocations for municipal spending can have a negative impact on children's educational outcomes, given the potential deterioration of public services such as schools, resulting in compromised learning conditions that hinder cognitive and developmental progress. The Italian municipalities are directly involved in this process by supervising the fundamental foundations that facilitate the development of children during early childhood. Specifically, they are accountable for funding facilities that provide childcare, primary and secondary education, with special emphasis on the construction, renovation, and maintenance of school premises. Indeed, it has been demonstrated in the literature that investment in school facilities improves student achievement (Jackson 2018; Jackson and Mackevicius 2021). Moreover, recent data from the INVALSI 2023 report[6] reveals a decline in the academic performance of Italian students, particularly in primary school subjects. Therefore, it is imperative to evaluate the effect of public spending on the human capital formation process to provide recommendations for the efficient allocation of resources, such as investments in education, to address inequalities in contemporary economic systems. It is important to note that an analysis of this kind may encounter methodological challenges, as a higher level of public spending on education could represent an endogenous characteristic of municipalities considered more virtuous. This virtuousness is often associated with a greater accumulation of human and social capital, which in turn leads to a higher demand for public goods and services in the education sector. To mitigate these potential endogeneity issues, we propose using an exogenous variation, determined by a specific law, in female participation in local decision-making bodies to evaluate its impact on the allocation of local public spending. Women's participation in local decision-making bodies refers to policies aimed at increasing women's representation in municipal councils, in our case, we refer to Law No. 215 of 2012, whose effects began in 2013, which introduces gender-conditional double preference voting in municipal elections. According to this law, voters can express two preferences, instead of one, if they vote for candidates of different gender (Baltrunaite et al., 2019). Considering this position, current literature has emphasised the economic repercussions of gender quotas in some countries (Cabaleiro and Buch 2018, 2020; Hern´andez-Nicol´as et al., 2018; Priyanka's 2022). The rationale is that, through the promotion of greater female participation in politics, they may exert an influence

[5] Since 1999, the DSP (Domestic Stability Pact) aimed to empower Italian municipalities through a set of local fiscal rules.
[6] On 12 July 2023, the Presentation of the INVALSI Report 2023 took place in the Chamber of Deputies.

on the political process and economic policy preferences regarding the allocation of resources, such as public spending (Andreoli et al., 2022). In Italy, various measures[7] have been implemented in recent decades, such as the introduction of gender quotas, to reduce the gender gap in politics. By examining the exogenous change in women's participation in local political bodies introduced by Law No. 215/2012, we expect a positive impact on local public spending in preschool education and in interventions for childhood, nurseries, and minors. To understand how potential variations in public spending, resulting from increased female political representation, influence the educational outcomes of second-grade students, particularly in mathematics and Italian, we employ the instrumental variables method. The analysis integrates data from Italian municipal budgets, focusing on Missions 4 and 12, along with information from the registry of local administrators and INVALSI data provided at the municipal level by the National Institute for the Evaluation of the Education and Training System. The first-stage regressions indicate that the selected instrument is highly effective. Additionally, we present evidence that the conditions of relevance and exogeneity are met. Each instrumented expenditure category is utilized in the second-stage regressions to evaluate the effect on the performance of Grade 2 students in Italian and mathematics. Our findings reveal that increased ordinary expenditure on preschool education and childcare positively influences INVALSI scores in Mathematics. These conclusions are supported by several robustness checks. The article is structured as follows: section two provides an overview of the literature, which is then divided into two subsections: austerity and achievement; gender quotas and public expenditure; section number three elaborates on the institutional framework of Law 215/2012 and the Italian education system, the data used and the related descriptive statistics; section number four discusses the econometric strategy; section five presents the results obtained; section number six provides details on the robustness checks; section number seven contains concluding remarks; finally, section number eight includes the appendix, while the last section is devoted to bibliographical references.

2. Literature Review

In literature, there is first a discussion on how austerity policies have negatively affected students' academic performance and widened the gap between economically disadvantaged students and others. Secondly, the role of gender-based preference differences in determining resource allocation for preschool education, childcare programs, and nurseries is analyzed.

[7] Law No. 81 of 1993, which was repealed in 1995 due to its unconstitutionality, mandated that neither sex could represent more than two-thirds of each municipal electoral list. Law No. 56 of 2014 stipulates that in municipalities with more than 3,000 inhabitants, neither sex can be represented to an extent less than 40 percent in the councils.

2.1 Austerity and Achievement

Student test scores are a function of school resources, teacher quality, and family background (Hanushek 2008). Policymakers can directly influence the allocation of resources for school inputs and infrastructure useful for learning and they should consider the potential social and economic costs of austerity policies when making decisions on public spending. However, the presence of budget rigidity can greatly limit the extent to which one can reallocate current expenses and eliminate inefficiencies. Austerity policies can have negative effects on student achievement, particularly for children from disadvantaged backgrounds. Austerity measures, which entail a reduction in public spending, have a negative impact on student performance in national standardised tests and spending cuts have a disproportionate impact on disadvantaged students (Pavese et al., 2023). Access to educational resources is a critical factor in shaping the outcomes of economically disadvantaged children, resulting in a significant reduction in the intergenerational transmission of poverty. The allocation of resources is equally important, and any increase in expenditure must be accompanied by efficient systems to ensure optimal utilization of funds. Increased spending per student significantly improves education levels, wages, household income and reduces adult poverty among low-income households (Jackson et al., 2015). However, the effects of increased spending on academic performance and economic outcomes in adulthood are less pronounced for children from non-poor families. Additional increases in spending have been associated with significant improvements in school quality, such as the reduction in student/teacher ratios, an increase in teacher salaries, and the extension of the school year (Jackson et al., 2015). Decision-makers should consider implementing school finance reforms that aim to equalize funding across districts to improve educational outcomes for disadvantaged students. Spending equalization leads to lower test scores among lower family background groups in particular, and it reduced the gap in average SAT scores (Scholastic Assessment Test) between children with highly educated parents and those with less educated parents (Card et al., 2002). The strength of local teachers' unions can play a crucial role in determining the effectiveness of school finance reforms and the allocation of resources to improve student outcomes. School finance reforms caused a significant increase in state aid to local governments in the US, where districts with strong unions increased spending on teacher pay, resulting in increases in student achievement (Brunner et al., 2018). The literature points to the effects of early education and access to public early childhood services on child outcomes. The provision of childcare services can be a powerful policy tool to enhance the language skills of immigrant children, particularly those whose mothers have limited education and those who speak a language very different from that of the host country. Maternal education has a multiplier effect on society: educated mothers tend to raise more educated children with greater prospects for success (Cui et al., 2019). Consequently, this can contribute to the advancement of the integration and assimilation of immigrant children in the host country. Italy continues to show a significant surplus in the demand for early childhood services, as well as a significant disparity in the provision of these services between different parts of the country (Del Boca, Pronato et al., 2016). The effect of childcare

attendance on the achievement gap between immigrant and native children is positive and significant, suggesting that childcare services can help to reduce the achievement gap between these two groups (Corazzini et al., 2021). The availability of public childcare plays a significant role in the maternal employment and cognitive outcomes of children in Italy. Politicians should consider expanding public childcare services to increase maternal employment and enhance children's cognitive development, especially in areas where childcare options are limited. Investments in childcare policies can help alleviate intergenerational persistence, especially for children from low-income families (Brilli et al., 2013).

2.2 Gender quotas and Public Spending

Women's representation in politics improves the quality of local politics. However, some studies are in conflict with the rest of the literature, providing no evidence that an increase in female representation has an impact on the composition of municipal expenditure (Campa and Bagues 2021, Cavallini et al., 2023, Gago and Carozzi 2021). The gender quota reform (law no. 81/1993) implemented in Italy in 1993 aimed to establish a more balanced gender composition across each list of candidates within municipal elections. Nevertheless, in 1995, the Constitutional Court deemed this law unconstitutional, asserting that the fundamental right to equal access to elected office cannot be subjected to selective treatment based on gender. Law No. 215/2012, introduced in Italy to increase women's representation in municipal councils, took effect during the 2013 elections. It implemented gender quotas and dual preference voting based on gender in municipalities with populations over 5,000. The law has successfully reduced gender gaps in political empowerment within Italian municipalities (Baltrunaite et al., 2019). Increasing the proportion of female councillors can lead to increased funding for local security, education, social services, water supply, health and the environment (Clots-Figueras, 2011, Funk and Gathmann, 2008, Rehavi, 2007, Andreoli et al., 2022). The augmentation in female representation predominantly stems from the preferential votes cast in favor of female candidates, thus highlighting the influential role of dual preference voting (Baltrunaite et al., 2019; Andreoli et al., 2022). Gender composition of municipal councils exerts an influence over the patterns of local expenditure (Svaleryd 2009, Cabaleiro and Buch 2018, 2020; Hern´andez-Nicol´as et al., 2018; Andreoli et al., 2022). The gender composition within institutions influences the structure of public spending, the budget policy cycle, and the size and structure of public spending in areas typically affected by PBC, with the exception of those related to women's needs (Ordine et al., 2022). The importance of fostering female leadership at different levels of government influences the accumulation of human capital in different ways, in particular the learning outcomes of primary school children in reading and mathematical competence (Priyanka's 2022). Women politicians increase the likelihood that public schools receive subsidies and useful inputs for learning. In addition, female politicians lead families to show greater confidence in the ability of public schools to provide a better education (Clots-Figueras 2011, 2012). Hessami and Baskaran (2019) observe that female electoral

success leads to an acceleration of about 40 per cent in the expansion of public childcare services and intensifies discussions on childcare policies in council meetings. Our study is in this line of research, focusing on gender-sensitive policy spending and extending the analysis not only to childcare, but also to pre-school education.

3. The institutional framework and data

3.1 Law no. 215/2012

Subnational levels of government in Italy include regions, provinces, and municipalities. There are 20 regions, 110 provinces, and approximately 7,904 municipalities, of which 70 percent have fewer than 5,000 inhabitants (Annuario statistico italiano, Istat 2021). Electoral rules are independently established for each level of government, and Law 215/2012 applies at the municipal level. Municipal administration manages the registry of births and deaths, the registry of deeds, and decides the level and allocation of local spending for various purposes, such as administration, education, and social services. Expenditures are financed by own taxes and fees as well as transfers from the central government. Municipalities are led by a mayor, assisted by a legislative body - the City Council - and an executive body - the City Board. Municipal elections are held every five years and municipal governments cannot alter the schedule (Baltrunaite et al., 2019). Law No. 215 of 23 November 2012, which came into force on 26 December 2012 but will be applied as of the 2013 elections, contains provisions to promote the rebalancing of gender representation in local councils and councils and regional councils The most significant novelty is the amendment of the law for the election of municipal councils with the introduction of measures to strengthen the presence of women, but also of considerable importance are the interventions aimed at consolidating gender equality in the juntas and, more generally, in all non-elective collegiate bodies of municipalities and provinces. For the election of municipal councils, in municipalities with a population of more than 5,000 inhabitants, the law provides for a double measure to ensure gender balance: the cd. list quota: in the lists of candidates neither of the two sexes can be represented by more than two thirds (with rounding to the upper unit for the least represented gender, even in case of decimal place less than 0,5); the introduction of the cd. double gender preference, which allows the voter to express two preferences (instead of one, as provided for in previous legislation) as long as they concern candidates of different sex, otherwise the cancellation of the second preference. In the event of infringement of the provisions on the quota list, provision is made for a differentiated sanction mechanism, depending on whether the population exceeds 15000 inhabitants or not.

3.2 Italian School System

Compulsory education in Italy lasts for 10 years, from ages 6 to 16, and includes the eight years of the first cycle of education and the first two years of the second cycle (Law 296 of 2006). In detail, the education system is organized as follows: *Preschool education*, which can be managed by the State, local authorities, directly or through agreements, by other public entities or private institutions, and accommodates children between the ages of three and six. *First cycle of education*, compulsory, lasting a total of 8 years, divided into: Primary school for students aged 6 to 11 (lasting five years) and Lower secondary school for students aged 11 to 14 (lasting three years). *Second cycle of education*, divided into two types of pathways: Upper secondary school for students who have successfully completed the first cycle of education (lasting five years). Schools offer courses in high schools (licei), technical institutes, and vocational institutes for students aged 14 to 19; Three-year and four-year vocational education and training (IeFP) courses, under regional authority, also aimed at students who have successfully completed the first cycle of education. Finally, higher education is offered by Universities, institutions of Higher Artistic, Musical and Coreutic Education (AFAM), and Higher Technical Institutes (ITS) with various types of courses.

3.3 Data

Combining data on Italian municipal budgets, data on the registry of local administrators provided by the AIDA PA database and the Ministry of the Interior, respectively, with INVALSI data provided at municipal level per student by the National Institute for the Evaluation of the Education and Training System[8], we obtain a repeated cross-section. The aim is to establish the effects of changes in local public spending, resulting from exogenous changes in the share of female participation in municipal councils, on the educational achievements of primary school students, using these as a proxy for the level of human capital. The Institute provides data on mathematics and Italian test scores for second grade students at the municipal level, but the municipality is identified when there are at least three schools. This does not allow us to conduct the analysis on all Italian municipalities, but only on a sample of 498 municipalities[9]. INVALSI data provides a means to objectively measure the performance of both individual students and schools. In addition, this database includes demographic and socio-economic data on students, which allows for a high degree of control over the heterogeneities typically present in individual student circumstances, facilitating an effective match with the municipal data available to us. The analysis focuses on Grade 2 education in mathematics and

[8] INVALSI tests consist of standardised performance tests administered to the entire population of Italian students.

[9] It is worth noting that 70 per cent of the total number of Italian municipalities - there are approximately 7904 of them - have fewer than 5,000 inhabitants. The average municipalities, which have between 5,000 and 250,000 inhabitants, total 2,357 and correspond to 30 per cent of the total number of Italian municipalities: 68.3 per cent of the country's population resides in them. (Annuario statistico italiano, Istat 2021).

Italian, considering the academic years 2013/2014, 2014/2015, 2015/2016, 2016/2017, 2017/2018, 2018/2019[10]. The items of ordinary and extraordinary municipal expenditure in Mission 4 (Programme 01, expenditure on preschool education) and in Mission 12 (Programme 01, expenditure on interventions for children, kindergartens and minors), a classification introduced by Article 1, paragraph 2, of Legislative Decree No. 118 of 23 June 2011, known as the Public Accounting Reform, were extracted from municipal budgets in the form of "pagamenti in conto competenza" from 2013 to 2018. We define "pagamenti in conto competenza" as payments actually made by the municipality from 2013 to 2018[11]. Finally, we gather data on elected politicians in elections conducted from 2013 to 2018 and utilize information on both the total number and the number of female councillors elected to calculate the proportion of women elected[12]. The change in the proportion of women elected from 2013 to 2018 will serve as the exogenous instrument for the endogenous variable "Expenditure on preschool education and Interventions for Children and Kindergartens."

3.4 Descriptive Statistics

Table 1 contains some descriptive statistics on the INVALSI score by geographic Area, while Figure 1 plots the difference between the average number of women councillors[13] in municipal councils before and after the reform at the 5,000-resident threshold. There has been a noticeable increase in the proportion of women elected to municipal councils and executive committees. Figure 2 presents the McCrary test, which shows no discontinuity at the population threshold. Figure 3 shows how near the 5,000-resident threshold, as the size of the municipality increases, there is a noticeable discontinuity caused by the reform, especially in the difference in regular spending on pre-school education before and after the implementation of Law No. 215/2012. A slight discontinuity is also observed in extraordinary spending on pre-school education. However, no discontinuity at the threshold is found, as influenced by the law, in regular and extraordinary spending on childcare, kindergartens and minors. Table A1 list the variables used in the model.

[10] We do not consider later years to avoid confounding factors in the analysis due to Covid-19.

[11] In the analysis, expenditure on preschool education and childcare interventions, expressed in euro units, are deflated to 2015 base year using the GDP deflator for Italy provided by the Federal Reserve Bank of St. Louis. We divide both types of expenditure by the population aged 0 to 5 years, as provided by ISTAT (Istituto Nazionale di Statistica), to obtain the amount of expenditure per child.

[12] The data are provided by the Ministry of Interior.

[13] The share of women was calculated as the average of the number of women in the municipal council and executive committee.

TABLE 1 *Invalsi Score Summary Statistics by Geographic Area*

	Area Geografica a 5 aree Istat					
	Nord ovest	Nord est	Centro	Sud	Isole	Total
N	344,559	192,891	366,414	400,626	26,967	1,331,457
	(25.9%)	(14.5%)	(27.5%)	(30.1%)	(2.0%)	(100.0%)
Italian	56.063	55.878	57.155	55.113	56.862	56.063
score[14]	(22.882)	(22.854)	(22.393)	(22.231)	(22.000)	(22.546)
Maths	56.234	56.155	57.747	56.603	56.245	56.749
score[15]	(20.358)	(20.449)	(19.965)	(20.321)	(19.873)	(20.254)

Notes. Mean (sd)
Special-status regions are not eliminated in the following descriptive statistics, period 2013-2018.

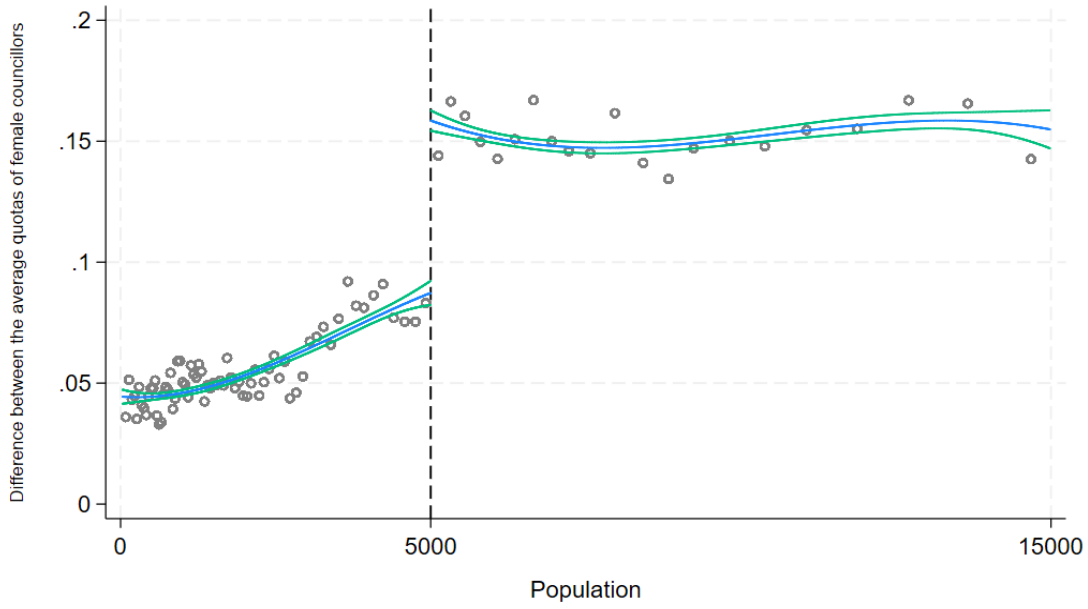

Fig.1. *Female representation in Italian municipalities*

Notes. The figure plots the difference between the averages of female councillors in municipal councils before and after the reform (considering the period from 2008 to 2018), along with the third-order polynomial spline on either side of the 5,000-resident threshold and the 95 percent confidence intervals. The sample includes Italian municipalities, excluding those belonging to special-status regions except for Sardinia, as Law No. 215/2012 does not apply to the latter.

[14] Standardised score in Italian adjusted for cheating.

[15] Standardised score in Mathematics adjusted for cheating.

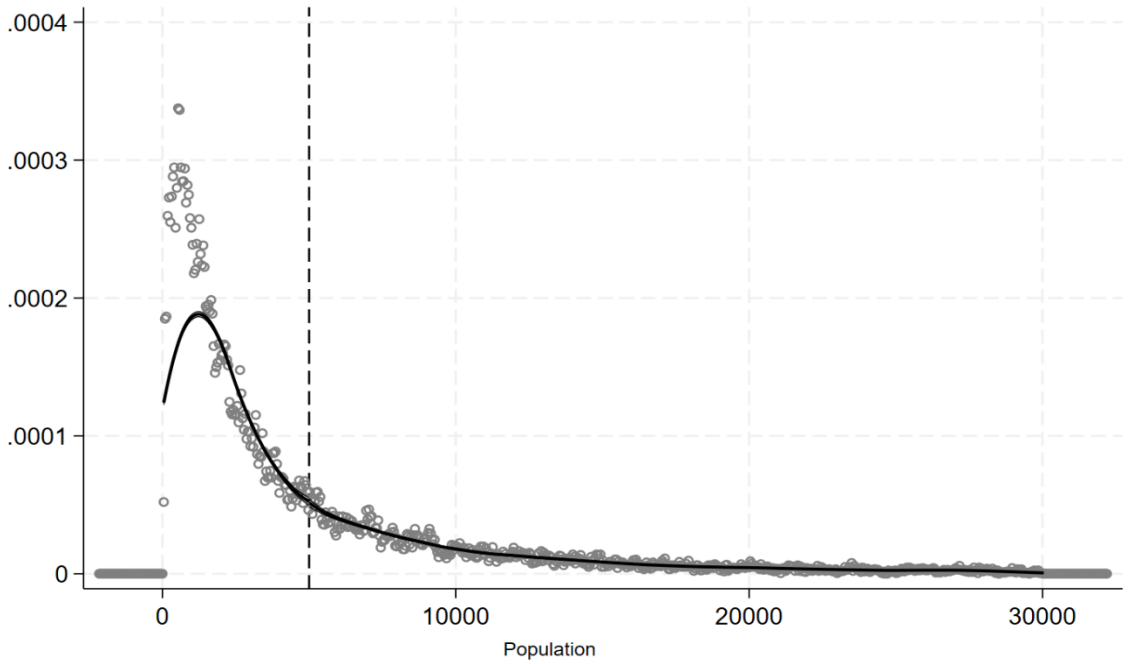

Fig.2 *McCrary test*

The figure plots municipal population density from 2012-2018. The population shows no discontinuities above and below the threshold.

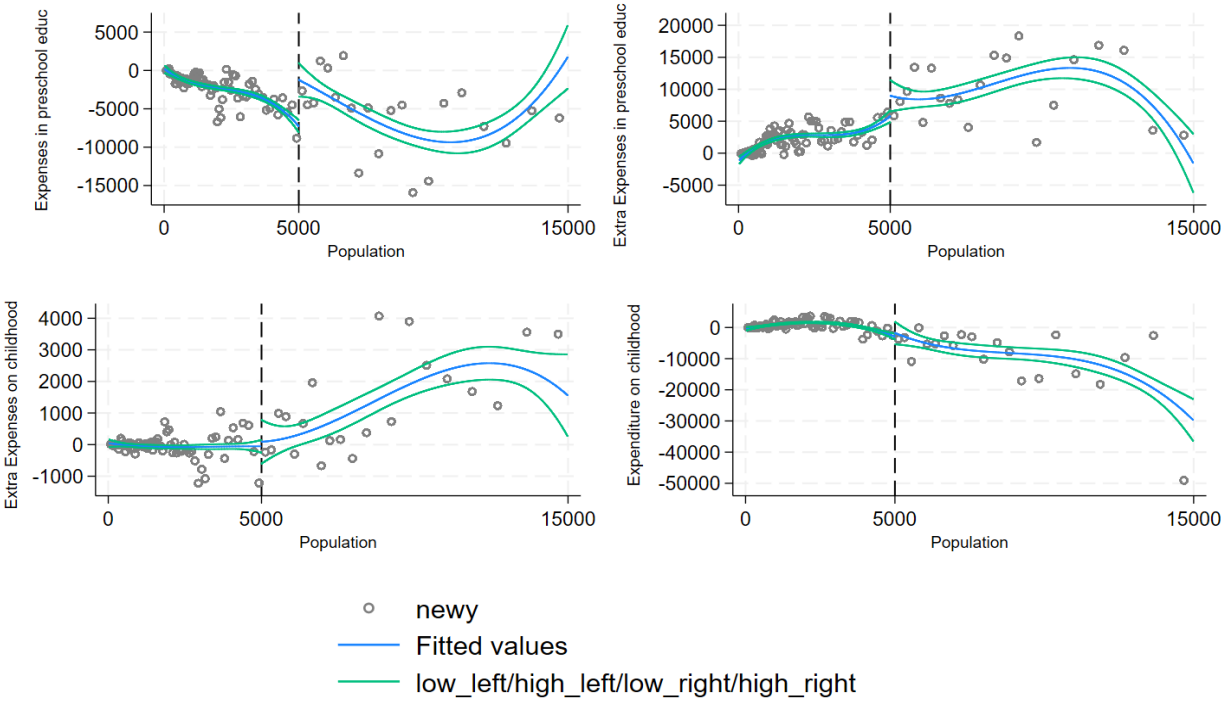

Fig.3 *Ordinary and Extraordinary Expenditure on Preschool Education (Mission4, Programme 01) and Interventions for Childhood and Kindergartens (Mission 12, Programme 01)*

Notes. Difference between the averages of Ordinary and Extraordinary Expenditure on Preschool Education and Interventions for Childhood and Kindergartens before and after the implementation of Law No. 215/2012. Special statute regions are eliminated in the following descriptive statistics except for Sardinia (period 2008-2018).

4. Econometric Strategy

This session will discuss the econometric problems of the analysis, the methodology used and the empirical model. The effects of local public spending on the human capital formation process in the early years of an individual's life could not only reduce the significant inequalities of economic systems but also improve the skills of new generations. Human capital is approximated by the INVALSI school results achieved by Italian elementary school students, as well as our main dependent variable. In our case, we focus on Grade 2 because an increase in spending on preschool education and childhood interventions, kindergartens, and minors, induced by greater female participation in municipal councils, could be evident from the first cycle of compulsory education in Italy, namely primary schools. The added value of this analysis, compared to the highlighted literature, lies in the fact that, unlike the work of Pavese et al. (2023), the allocation of spending not subject to the Domestic Stability Pact rules could be more efficient due to the increase in female representation in municipal councils, as these areas of spending might be better aligned with the needs of women.To estimate the effect of spending on the INVALSI score, we use the ORDINARY and EXTRAORDINARY spending types of Missions 4 and 12 (programme 01 both), leveraging the exogeneity provided by the variation in the presence of women in municipal councils. We interact spending with the child's exposure time to the reform. The child's exposure to the reform is calculated as the difference between when the child attended second grade (when we observe the test scores) and the year the reform was introduced. The exposure variable ranges from 0 (all those children who took the tests during the year of the reform) to 5 (children who had a maximum exposure time to the reform). Since the exposure time and spending represent interactions between an endogenous variable and an exogenous one, it is necessary to use appropriate instruments for these interactions as well. For this reason, we use the interaction between the variation in the presence of women in municipal councils and the child's exposure time to the reform years as an instrument for the endogenous interaction term of spending and exposure to the reform. As stated in the literature, the interactions between the instrument “female presence” and the exogenous variables, such as the child's exposure to the reform in question, serve as effective instruments for the endogenous EXPxEXPEND variables (Wooldridge, 2000). Finally, the individual expenditure items used in the analysis, ordinary and extraordinary, will be instrumented in all specifications with the change in the presence of women in municipal councils. We expect that as the average number of women elected to municipal councils increases, spending on preschool education and childhood interventions per pupil will increase, with a consequent positive impact on school results (Baskaran and Hessami, 2018).

4.1. Endogeneity problem

When trying to assess the impact of public spending on the INVALSI score, some methodological difficulties are encountered due to reverse causality problems and omitted variables. Municipalities with higher employment rates generate more demand for public day-care centres from policy makers. This is due to the fact that working parents are not able to take care of their children once the statutory holiday periods are over. Consequently, a reverse causality problem may arise due to the higher proportion of working people in relation to the population, resulting in local spending on education. At the same time, higher local spending on education has a lasting effect on human capital through the acquisition of more individual skills, which does not discourage employment, particularly female employment, due to the balance between work and the cost of enrolling children in a private kindergarten. Furthermore, municipalities with a more virtuous institutional environment are less likely to evade taxes. As a result, they have more revenue to finance public goods and services, including the creation of new kindergartens or primary schools, with the consequent hiring of new teachers. In addition, we use the change in the presence of women in municipal councils induced by the introduction of law 215/2012 and not simply the share of elected women within of the administration. The proportion of women elected to municipal councils may be due to socio-cultural reasons and gender stereotypes, which is why it is considered endogenous (Ordine et al., 2022). In particular in the southern regions, the proportion of women may be lower than in the northern regions. Women in southern Italy are more likely to be confined to traditional social roles and candidates are often perceived as violating these roles. As a result, southern voters are likely to show more prejudice against women and may choose not to vote or express their discontent when they are faced with a Increase in female candidates for elections (Alesina et al., 2008; De Paola et al., 2014).

4.2. Instrumental variables

From an econometric modelling perspective, a two-stage regression model is needed to study the relationship between INVALSI results in Italian and mathematics (second grade) and local public spending on preschool education and childhood interventions, kindergartens, and minors. For this purpose, Law 215/2012 is used as an exogenous variation in the analysis. Specifically, we use the change in the average share of women elected to the various municipal councils between 2013 and 2018 as an exogenous instrument for public spending on preschool education (Mission 4, Programme 01) and interventions for children, kindergartens and minors (Mission 12, Programme 01). To calculate this change, we take the average number of women on municipal councils during the 2013-2018 elections and compare it to the average number of women on municipal councils in 2012. Figure 4 shows a map of the municipalities where elections were held during this period. It can be seen that most municipalities held elections in 2014. In our model, we include two types of public

expenditure, ordinary and extraordinary, which correspond respectively to Missions 4 and 12 of Legislative Decree No 118/2011. These expenditure items, scaled by the total 0-5 year old population of each municipality, are interacted with the time the children have been exposed to the reform, which in turn is instrumented by the interaction term given by the change in the presence of women in municipal councils and the exposure of children to Law 215/2012. The exposure variable represents the period during which the child was exposed to the reform in the second grade (Table A2). We consider the exposure time to Law No. 215/2012 starting from 2013, as the reform came into effect at the end of December 2012. To confirm the reliability of the results, validity tests were conducted on the instruments (A3-A6; A9-A12).

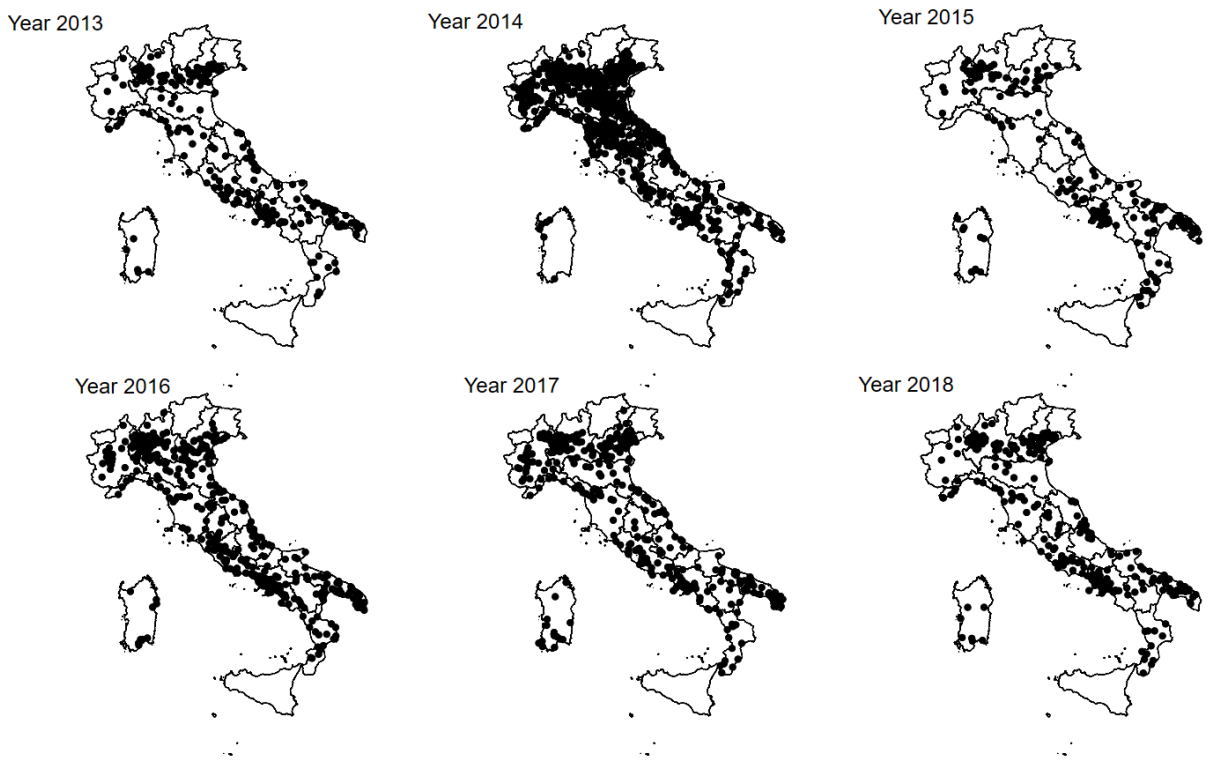

Fig.4 *Municipal elections in Italian municipalities*

Source: Eligendo a cura del DAIT - Ministero dell'Interno, Direzione Centrale per i Servizi Elettorali - S.I.E.C.C.

Notes: The map shows the municipalities that participated in the local elections between 2013 and 2018. In the creation of this map, municipalities with a population of less than 5,000 inhabitants and regions with special statutes were excluded. However, Sardinia was included because, unlike the other regions with special statutes, it was subject to Law No. 215/2012.

4.2.1 The Empirical Model

Equations 1-4 refer to the second stage regression. Y is the standardised test score in mathematics and Italian adjusted for cheating (grade II); EXPENDord[1] is ordinary spending per pupil[16] on preschool education (Mission 4, Programme 01);EXPENDextra[1] is extraordinary spending per pupil on preschool education (Mission 4, Programme 01); EXPENDord[2] is ordinary spending per pupil on childhood interventions, kindergartens, and minors (Mission 12, Programme 01); EXPENDextra[2] is extraordinary spending per pupil on childhood interventions, kindergartens, and minors (Mission 12, Programme 01); EXP is the time of the child's exposure to the reform; EXPxEXPENDord[1,2] and EXPxEXPENDextra[1,2] are the interaction between the individual items of expenditure in preschool education and childcare, day care and children per pupil, (distinguishing between ordinary and extraordinary expenditure), and the child's exposure to reform at the time of the INVALSI tests.;ϒ are province fixed effects to account for any unobserved heterogeneity over time; τ are time fixed effects; X are different controls(gender, pre-school attendance, parental background, nationality); μ is the error term.

$$Y_{i,t} = \alpha + \beta_1 \cdot EXPxEXPENDord_{i,t}^{1IV} + \beta_2 EXPENDord_{i,t}^{1IV} + \beta_3 EXP_{i,t} + \tau_t + \Upsilon_i + \pi \cdot X_{i,t} + \mu_{i,t} \quad (1)$$

$$Y_{i,t} = \alpha + \beta_1 \cdot EXPxEXPENDextra_{i,t}^{1IV} + \beta_2 EXPENDextra_{i,t}^{1IV} + \beta_3 EXP_{i,t} + \tau_t + \Upsilon_i + \pi \cdot X_{i,t} + \mu_{i,t} \quad (2)$$

$$Y_{i,t} = \alpha + \beta_1 \cdot EXPxEXPENDord_{i,t}^{2IV} + \beta_2 EXPENDord_{i,t}^{2IV} + \beta_3 EXP_{i,t} + \tau_t + \Upsilon_i + \pi \cdot X_{i,t} + \mu_{i,t} \quad (3)$$

$$Y_{i,t} = \alpha + \beta_1 \cdot EXPxEXPENDextra_{i,t}^{2IV} + \beta_2 EXPENDextra_{i,t}^{2IV} + \beta_3 EXP_{i,t} + \tau_t + \Upsilon_i + \pi \cdot X_{i,t} + \mu_{i,t} \quad (4)$$

Consider the following first-stage equations (5-8):

$$EXPENDord_{i,t}^{1,2} = \alpha + \delta_1 \, Female \; Presence_{\,i,t} + \varphi_1 Covariates_{i,t} + v_{i,t} \quad (5)$$

$$EXPENDextra_{i,t}^{1,2} = \alpha + \delta_2 \, Female \; Presence_{\,i,t} + \varphi_2 Covariates_{i,t} + v_{i,t} \quad (6)$$

$$EXPxEXPENDord_{i,t}^{1,2} = \alpha + \delta_5 EXPxFemale \; Presence_{\,i,t} + \varphi_5 Covariates_{i,t} + v_{i,t} \quad (7)$$

$$EXPxEXPENDextra_{i,t}^{1,2} = \alpha + \delta_6 EXPxFemale \; Presence_{\,i,t} + \varphi_6 Covariates_{i,t} + v_{i,t} \quad (8)$$

In equations 5-6, we regress the two expenditure items (distinguishing between ordinary and extraordinary expenditures) per pupil i at time t on our Female Presence instrument, and on all variables included in

[16] We scale spending by the total 0-5 population of each municipality using ISTAT data (Pavese and Rubolino 2023).

equations 1-4, summarized here as Covariates; v_{it} represents the first-stage residuals. Provincial fixed effects and time fixed effects are included in equations. Female presence is the change in the average female share of each municipality resulting from elections held between 2013 and 2018, influenced by Law 215/2012. In equations 7-8 we regress the expenditure items used in the analysis interacted with the time of exposure to the reform for each child i at time t, during the INVALSI of grade 2. By including the variables included in equations 1-4, temporal and provincial fixed effects are included in the covariates. EXPxFemale presence is the instrument for the endogenous interaction term EXPxEXPEND in each equation for the two items of ordinary and extraordinary expenditures.

5. Results

In Tables 2 and 3, we present the estimates for equations 1-4 using a simple Ordinary Least Squares (OLS) regression model, where the dependent variable is the standardized INVALSI score, first in mathematics and then in Italian. We aim to test the effect of public spending on student performance. The model accounts for temporal effects and fixed province effects. In Table 2 the results show a negative effect of spending on preschool education and child care interventions, including nurseries and services for minors, on INVALSI scores in mathematics, except in column 4, where the coefficient for extraordinary spending on child care interventions, nurseries, and services for minors is close to zero and not significant. The coefficients of the interaction terms between the two items of ordinary and extraordinary expenditure per pupil and the time of exposure to reform of the pupil are positive in columns 1, 2 and 3, except in column 4, where the coefficient of interaction between extraordinary expenditure per pupil, in child care, day care and minors, with the time of exposure to reform of the pupil is negative. Table 3 performs the same analysis using the Italian language score. In this specification as well, the coefficients related to spending variables are negative, and in column 2, extraordinary spending on preschool education seems not to be significant, as are the coefficients for the various interactions. These results do not show a positive and significant impact of spending on children's performance as found in the literature. Estimates derived from the OLS estimator could be biased and conditioned by some confounding factors, making it difficult to identify the true effect of public spending on INVALSI performance. To identify the true causal effect of spending, we adopted the IV procedure. In Tables 4 and 5, we present the results of the first stages related to the standardized mathematics scores. In each equation, the endogenous variables were regressed on the instruments and other exogenous covariates included in the second stage. In table 4, column 1, the ordinary expenditure on preschool education is regressed on the change in the proportion of women in municipal councils; in column 2, the extraordinary expenditure on preschool education is regressed on the change in the proportion of women in municipal councils. In columns 3 and 4, the ordinary and extraordinary expenditure on interventions for children, kindergartens, and minors are regressed on the change in the proportion of women in municipal councils. In this specification the effect of the instrument is positive on ordinary per pupil spending in preschool education

and ordinary spending in childcare interventions, which means that the dual preference law, which sees a greater presence of women in municipal councils, actually increased ordinary spending in preschool education and ordinary spending in childcare, preschool and childcare interventions. While no effect is found on extraordinary spending in preschool education, but we detect a negative effect on per-pupil extraordinary spending in interventions for infancy, kindergartens and children. In table 5 we regress the interactions between the two types of expenditure (always distinguishing between ordinary and extraordinary) and the child's exposure time to the reform on the instrument, which is the interaction between the change in the proportion of women in municipal councils and the child's exposure time to the reform. In the first-stage regressions with interaction terms, a negative effect of the presence of women in municipal councils is found on extraordinary spending per pupil in preschool education interacted with the time of pupil exposure to reform. In tables 4 and 5, we present the results of the Kleibergen-Paap rk Wald test, which provide evidence supporting the robustness of the instruments used. In Tables A3, A4, A5 and A6 we report the tests performed on the various specifications. In panel A of Table A3-A6, the results of the Durbin-Wu-Hausman test suggest the presence of endogeneity resulting from the expenditure variable and its interactions with the child's exposure time to the reform. In panel B of Tables A3-A6, the Anderson-Rubin Wald and Stock-Wright LM S tests are used to assess the significance of the endogenous regressors, both confirming that the endogenous regressors are not jointly equal to zero. In panel C of Tables A3-A6, the Kleibergen-Paap rk LM test is presented, indicating that the model is not under-identified. The Hansen J test statistic is equal to zero, as our model is exactly identified. The second-stage estimates for the mathematics score are shown in Table 6, which includes an empirical specification with time and province fixed effects. The results highlight a positive effect of ordinary spending on preschool education (Mission 4, Programme 01) and interventions for early childhood, kindergartens, and minors (Mission 12, Programme 01) on mathematics test scores. Specifically, an increase of one euro per pupil in these areas raises mathematics scores by 0.18% and 0.11%, respectively. Conversely, extraordinary spending on preschool education leads to lower test scores in Mathematics, while spending on interventions for early childhood, kindergartens, and minors does not appear significant. However, the interaction coefficients show a positive sign and remain significant in all specifications. We replicate the analysis by changing the dependent variable to the standardized score in Italian. The first stage results for the Italian score are presented in Tables A7 and A8, along with the tests on the validity of the instruments (Tables A9-A12), while Table 7 displays the second stage estimates for the score in Italian. In this case, the coefficients of the expenditure in preschool education and in childcare, nurseries and minors are not statistically significant. The extraordinary expenditure per pupil in preschool education seems to reduce the INVALSI score in Italian. In light of these results, it appears that ordinary spending on preschool education, as part of Mission 4, has a positive impact on INVALSI scores in Mathematics, without, however, affecting scores in Italian. On the contrary, extraordinary spending on preschool education tends to reduce INVALSI scores in both Mathematics and Italian. Similarly, the ordinary per pupil spending under Mission 12, related

to early childhood interventions, including daycare and minors, contributes to increasing INVALSI scores in Mathematics but does not show significant effects on scores in Italian; while the extraordinary spending under Mission 12 is not statistically significant in either case. Therefore, ordinary spending has a positive effect on Mathematics test scores compared to extraordinary spending, but not on Italian scores, for both preschool education and early childhood interventions. The effect is observed in the ordinary spending of both missions, likely because it exerts a cumulative and consistent influence, impacting the educational budget on a daily basis. In contrast, extraordinary spending represents an isolated or rare event with effects that occur in the long run not the short run.

TABLE 2 *Outcome: standardized II grade Math test score*

OLS ESTIMATES	(1) Standardized second-grade math test score (adjusted for cheating)	(2) Standardized second-grade math test score (adjusted for cheating)	(3) Standardized second-grade math test score (adjusted for cheating)	(4) Standardized second-grade math test score (adjusted for cheating)
$EXPEND^{1ord}$	-0.000883***			
	(0.000)			
$EXPxEXPEND^{1ord}$	0.0000690**			
	(0.000)			
$EXPEND^{1extra}$		-0.00271**		
		(0.001)		
$EXPx\ EXPEND^{1extra}$		0.00145***		
		(0.000)		
$EXPEND^{2ord}$			-0.000606***	
			(0.000)	
$EXPxEXPEND^{2ord}$			0.0000651**	
			(0.000)	
$EXPEND^{2extra}$				-0.00153
				(0.002)
$EXPx\ EXPEND^{2extra}$				-0.00191**
				(0.001)
EXP	0.210***	0.218***	0.205***	0.259***
	(0.019)	(0.015)	(0.022)	(0.015)
Nationality	5.592***	5.684***	5.637***	5.680***
	(0.086)	(0.086)	(0.086)	(0.086)
Regular Studies	2.885***	3.028***	2.952***	3.006***
	(0.445)	(0.445)	(0.445)	(0.445)
Gender	2.182***	2.181***	2.182***	2.182***
	(0.043)	(0.043)	(0.043)	(0.043)
Father low education	-1.092***	-1.067***	-1.080***	-1.062***
	(0.106)	(0.106)	(0.106)	(0.106)
Father mid education	2.185***	2.197***	2.193***	2.201***
	(0.102)	(0.102)	(0.102)	(0.102)
Father high education	4.532***	4.497***	4.520***	4.504***
	(0.111)	(0.111)	(0.111)	(0.111)
Mother low education	-1.507***	-1.482***	-1.497***	-1.477***
	(0.118)	(0.118)	(0.118)	(0.118)
Mother mid education	2.822***	2.847***	2.834***	2.849***
	(0.112)	(0.112)	(0.112)	(0.112)
Mother high education	5.538***	5.530***	5.535***	5.534***

	(0.118)	(0.118)	(0.118)	(0.118)
_cons	45.84***	45.38***	45.74***	45.33***
	(0.534)	(0.532)	(0.535)	(0.532)
Mean of dep. var	56.74	56.74	56.74	56.74
Controls	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes
Province Effects	Yes	Yes	Yes	Yes
Obs.	814715	814715	814715	814715
R^2	0.0768	0.0766	0.0767	0.0766

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$

Note: Dependent variable is the corrected total percentage score for cheating in Math. EXPEND^{1ord} and EXPEND^{1extra} is the ordinary and extraordinary expenditure per pupil in preschool education (Mission 4, Programme 01), expressed in 2015 euros. EXPEND^{2ord} EXPEND^{2extra} is the ordinary and extraordinary expenditure per pupil in childcare, kindergartens and minors (Mission 12, Programme 01), expressed in 2015 euros. Exposure is the time of exposure of the child to the reform. The sample expands over the period 2013–2018. Standard errors are robust to heteroscedasticity.

TABLE 3 *Outcome: standardized II grade Italian test score*

OLS ESTIMATES	(1) Standardized second-grade Italian test score (adjusted for cheating)	(2) Standardized second-grade Italian test score (adjusted for cheating)	(3) Standardized second-grade Italian test score (adjusted for cheating)	(4) Standardized second-grade Italian test score (adjusted for cheating)
EXPEND^{1ord}	-0.000427***			
	(0.000)			
EXPxEXPEND^{1ord}	0.0000263			
	(0.000)			
EXPEND^{1extra}		-0.000776		
		(0.001)		
EXPxEXPEND^{1extra}		0.000456		
		(0.000)		
EXPEND^{2ord}			-0.000219**	
			(0.000)	
EXPxEXPEND^{2ord}			0.00000364	
			(0.000)	
EXPEND^{2extra}				-0.00364*
				(0.002)
EXPx EXPEND^{2extra}				-0.000575
				(0.001)
EXP	-1.470***	-1.466***	-1.455***	-1.450***
	(0.020)	(0.016)	(0.024)	(0.017)
Nationality	8.448***	8.495***	8.472***	8.492***
	(0.093)	(0.093)	(0.093)	(0.093)
Regular Studies	5.182***	5.254***	5.218***	5.240***
	(0.464)	(0.464)	(0.464)	(0.464)
Gender	-1.026***	-1.026***	-1.026***	-1.026***
	(0.047)	(0.047)	(0.047)	(0.047)
Father low education	-1.138***	-1.125***	-1.132***	-1.123***
	(0.115)	(0.115)	(0.115)	(0.115)
Father mid education	2.409***	2.416***	2.413***	2.418***
	(0.111)	(0.111)	(0.111)	(0.111)
Father high education	5.063***	5.046***	5.057***	5.051***
	(0.121)	(0.121)	(0.121)	(0.121)
Mother low education	-1.227***	-1.214***	-1.220***	-1.211***

	(0.128)	(0.128)	(0.128)	(0.128)
Mother mid education	3.177***	3.190***	3.183***	3.191***
	(0.122)	(0.122)	(0.122)	(0.122)
Mother high education	6.581***	6.577***	6.580***	6.580***
	(0.129)	(0.129)	(0.129)	(0.129)
_cons	48.61***	48.38***	48.51***	48.37***
	(0.562)	(0.560)	(0.563)	(0.560)
Mean of dep. var	56.06	56.06	56.06	56.06
Controls	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes
Province Effects	Yes	Yes	Yes	Yes
Obs.	808136	808136	808136	808136
R^2	0.128	0.128	0.128	0.128

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$

Note: Dependent variable is the corrected total percentage score for cheating in Italian. $EXPEND^{1ord}$ and $EXPEND^{1extra}$ is the ordinary and extraordinary expenditure per pupil in preschool education (Mission 4, Programme 01), expressed in 2015 euros. $EXPEND^{2ord}$ and $EXPEND^{2extra}$ is the ordinary and extraordinary expenditure per pupil in childcare, kindergartens and minors (Mission 12, Programme 01), expressed in 2015 euros. Exposure is the time of exposure of the child to the reform. The sample expands over the period 2013–2018. Standard errors are robust to heteroscedasticity.

TABLE 4 *First stage related to the Standardized second-grade Math test score*

Outcome: expenditure per pupil in preschool education and expenditure in childcare, kindergartens, and minors

	(1) $EXPEND^{1ord}$	(2) $EXPEND^{1extra}$	(3) $EXPEND^{2ord}$	(4) $EXPEND^{2extra}$
$EXPxFemale\ presence^{IV}$	20.26***	-4.605***	-17.48***	2.014***
	(1.229)	(0.093)	(1.331)	(0.060)
$Female\ presence^{IV}$	**737.5*** **	**0.113**	**797.4*** **	**-2.496*** **
	(6.175)	(0.477)	(6.523)	(0.310)
EXP	0.0624	4.631***	34.92***	0.641***
	(0.378)	(0.048)	(0.363)	(0.017)
_cons	476.9***	-5.391***	518.8***	1.276***
	(6.080)	(0.880)	(6.333)	(0.362)
N	771409	771409	771409	771409
Kleibergen-Paap rk Wald statistic:	9855.63	269.68	6093.51	37.31
Controls	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes
Province Effects	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$

Note: $EXPEND^{1ord}$ and $EXPEND^{1extra}$ is the ordinary and extraordinary expenditure per pupil in preschool education (Mission 4, Programme 01), expressed in 2015 euros. $EXPEND^{2ord}$ and $EXPEND^{2extra}$ is the ordinary and extraordinary expenditure per pupil in childcare, kindergartens and minors (Mission 12, Programme 01), expressed in 2015 euros. Exposure is the time of exposure of the child to the reform. The main instrument for each First-stage regression is written in bold. Standard errors are adjusted for heteroscedasticity.

TABLE 5 *First stage related to the Standardized second-grade Math test score*
Outcome: interaction term of the expenditure variables per pupil with the child's exposure time to the law

	(1) EXPxEXPEND^{1ord}	(2) EXPxEXPEND^{1extra}	(3) EXPxEXPEND^{2ord}	(4) EXPxEXPEND^{2extra}
EXPxFemale presence^{IV}	**803.1*****	**-39.69*****	**451.1*****	**1.113*****
	(5.095)	(0.377)	(5.669)	(0.179)
Female presence^{IV}	192.3***	63.16***	990.6***	4.642***
	(17.822)	(1.509)	(20.341)	(0.717)
EXP	355.5***	30.97***	773.1***	8.509***
	(1.631)	(0.222)	(1.895)	(0.051)
_cons	303.7***	-44.83***	-311.4***	-6.407***
	(27.273)	(3.066)	(32.990)	(1.114)
N	771409	771409	771409	771409
Kleibergen-Paap rk Wald statistic:	9855.63	269.68	6093.51	37.31
Controls	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes
Province Effects	Yes	Yes	Yes	Yes

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$
Note: EXPEND^{1ord} and EXPEND^{1extra} is the ordinary and extraordinary expenditure per pupil in preschool education (Mission 4, Programme 01), expressed in 2015 euros. EXPEND^{2ord} and EXPEND^{2extra} is the ordinary and extraordinary expenditure per pupil in childcare, kindergartens and minors (Mission 12, Programme 01), expressed in 2015 euros. Exposure is the time of exposure of the child to the reform. The main instrument for each First-stage regression is written in bold. Standard errors are adjusted for heteroscedasticity.

TABLE 6 *Outcome: standardized II grade Math test score*

2SLS ESTIMATES	(1) Standardized second-grade math test score (adjusted for cheating)	(2) Standardized second-grade math test score (adjusted for cheating)	(3) Standardized second-grade math test score (adjusted for cheating)	(4) Standardized second-grade math test score (adjusted for cheating)
EXPEND$^{1ord\ IV}$	0.00179***			
	(0.000)			
EXPxEXPEND$^{1ord\ IV}$	0.000186*			
	(0.000)			
EXPEND$^{1extra\ IV}$		-0.229***		
		(0.035)		
EXPxEXPEND$^{1extra\ IV}$		0.0219***		
		(0.005)		
EXPEND$^{2ord\ IV}$			0.00114**	
			(0.000)	
EXPxEXPEND$^{2ord\ IV}$			0.000455**	
			(0.000)	
EXPEND$^{2extra\ IV}$				-0.0537
				(0.054)
EXPxEXPEND$^{2extra\ IV}$				0.264***
				(0.052)
EXP	0.148**	0.598***	-0.177	-1.997***
	(0.051)	(0.051)	(0.115)	(0.419)
Nationality	5.950***	5.466***	5.919***	6.161***
	(0.095)	(0.098)	(0.093)	(0.136)
Regular Studies	3.355***	2.090***	3.307***	4.634***
	(0.454)	(0.479)	(0.452)	(0.587)
Gender	2.225***	2.241***	2.223***	2.173***

	(0.044)	(0.047)	(0.044)	(0.055)
Father low education	-0.940***	-0.967***	-0.932***	-0.988***
	(0.109)	(0.115)	(0.109)	(0.130)
Father mid education	2.294***	2.413***	2.305***	2.212***
	(0.105)	(0.113)	(0.105)	(0.127)
Father high education	4.450***	4.860***	4.457***	4.068***
	(0.114)	(0.127)	(0.114)	(0.161)
Mother low education	-1.445***	-1.452***	-1.454***	-1.793***
	(0.121)	(0.128)	(0.121)	(0.152)
Mother mid education	2.922***	2.904***	2.917***	2.772***
	(0.114)	(0.121)	(0.114)	(0.137)
Mother high education	5.534***	5.736***	5.529***	5.283***
	(0.120)	(0.130)	(0.120)	(0.152)
_cons	44.36***	45.02***	44.83***	47.03***
	(0.582)	(0.557)	(0.632)	(0.685)
N	771409	771409	771409	771409
R^{2UN}	0.900	0.886	0.900	0.848
R^{2C}	0.0761	-0.0576	0.0760	-0.414
Controls	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes
Province Effects	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$

Note: Dependent variable is the corrected total percentage score for cheating in Math. $EXPEND^{1ord}$ and $EXPEND^{1extra}$ is the ordinary and extraordinary expenditure per pupil in preschool education (Mission 4, Programme 01), expressed in 2015 euros. $EXPEND^{2ord}$ $EXPEND^{2extra}$ is the ordinary and extraordinary expenditure per pupil in childcare, kindergartens and minors (Mission 12, Programme 01), expressed in 2015 euros. Exposure is the time of exposure of the child to the reform. The expenditure variables and the interactions between expenditure and the exposure time to the law are instrumented respectively by the change in the proportion of women induced by Law 215/2012 and by the interaction term between the change in the proportion of women and the exposure time to the reform. The sample expands over the period 2013–2018. Standard errors are robust to heteroscedasticity.

TABLE 7 *Outcome: standardized II grade Italian test score*

2SLS ESTIMATES	(1) Standardized second-grade Italian test score (adjusted for cheating)	(2) Standardized second-grade Italian test score (adjusted for cheating)	(3) Standardized second-grade Italian test score (adjusted for cheating)	(4) Standardized second-grade Italian test score (adjusted for cheating)
$EXPEND^{1ord\ IV}$	0.000483			
	(0.000)			
$EXPxEXPEND^{1ord\ IV}$	0.000269**			
	(0.000)			
$EXPEND^{1extra\ IV}$		-0.106**		
		(0.037)		
$EXPxEXPEND^{1extra\ IV}$		0.00679		
		(0.005)		
$EXPEND^{2ord\ IV}$			-0.0000999	
			(0.001)	
$EXPxEXPEND^{2ord\ IV}$			0.000493**	
			(0.000)	
$EXPEND^{2extra\ IV}$				0.0594
				(0.055)
$EXPxEXPEND^{2extra\ IV}$				0.106**
				(0.047)
EXP	-1.610***	-1.234***	-1.891***	-2.461***
	(0.056)	(0.053)	(0.125)	(0.374)
Nationality	8.593***	8.355***	8.580***	8.707***

	(0.102)	(0.100)	(0.101)	(0.124)
Regular Studies	5.472***	4.825***	5.449***	6.139***
	(0.474)	(0.488)	(0.472)	(0.552)
Gender	-1.015***	-1.009***	-1.017***	-1.040***
	(0.048)	(0.049)	(0.048)	(0.051)
Father low education	-1.069***	-1.071***	-1.063***	-1.094***
	(0.118)	(0.119)	(0.118)	(0.123)
Father mid education	2.467***	2.535***	2.477***	2.416***
	(0.114)	(0.117)	(0.114)	(0.119)
Father high education	5.057***	5.277***	5.064***	4.855***
	(0.124)	(0.132)	(0.124)	(0.149)
Mother low education	-1.237***	-1.234***	-1.247***	-1.404***
	(0.131)	(0.132)	(0.131)	(0.142)
Mother mid education	3.211***	3.201***	3.209***	3.146***
	(0.124)	(0.126)	(0.124)	(0.130)
Mother high education	6.583***	6.682***	6.579***	6.456***
	(0.131)	(0.135)	(0.131)	(0.143)
_cons	48.11***	48.17***	48.63***	48.99***
	(0.616)	(0.571)	(0.673)	(0.654)
N	764956	764956	764956	764956
R^{2UN}	0.884	0.880	0.884	0.870
R^{2C}	0.129	0.0998	0.128	0.0280
Controls	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes
Province Effects	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$

Note: Dependent variable is the corrected total percentage score for cheating in Italian. $EXPEND^{1ord}$ and $EXPEND^{1extra}$ is the ordinary and extraordinary expenditure per pupil in preschool education (Mission 4, Programme 01), expressed in 2015 euros. $EXPEND^{2ord}$ $EXPEND^{2extra}$ is the ordinary and extraordinary expenditure per pupil in childcare, kindergartens and minors (Mission 12, Programme 01), expressed in 2015 euros. Exposure is the time of exposure of the child to the reform. The expenditure variables and the interactions between expenditure and the exposure time to the law are instrumented respectively by the change in the proportion of women induced by Law 215/2012 and by the interaction term between the change in the proportion of women and the exposure time to the reform. The sample expands over the period 2013–2018. Standard errors are robust to heteroscedasticity.

6. Robustness Checks

In this section, we present the robustness checks of the analysis. We decided to re-estimate the model for each individual year of exposure to the law under study. In Tables A13-A16, we replicate the analysis from Table 6 for children exposed to Law No. 215/2012 for periods from 1 to 5 years, using expenditure on preschool education and interventions in favor of childhood, nurseries, and minors, instrumented by the variation in the share of women from 2013 to 2018, distinguishing between ordinary and extraordinary expenditure. The same procedure is applied for the outcome related to the Italian score (A17-A20). The robustness checks generally confirm the results of the main analyses. The effect of ordinary expenditure for both missions is very pronounced for children who take the INVALSI mathematics tests up to 2016; thereafter, the effect vanishes. In particular, for spending on interventions for children, kindergartens and minors, the (instrumented) expenditure coefficient increases with the length of exposure to the reform when the outcome is the mathematics score. Extraordinary expenditure for Mission 12 has positive effects on the mathematics score in 2015, while for Mission 4 in 2016, it positively impacts the INVALSI mathematics

outcome. In robustness checks we find a positive effect of ordinary spending for Missions 4 and 12 on the score in Italian, except for children who took the tests in the year 2015 (EXP=2) and in the year 2018 (EXP=5). The negative effect of extraordinary spending for Mission 12 on the score in Italian for children who took INVALSI tests in the years 2016 and 2017 is confirmed, while extraordinary spending for mission 4 seems to have a positive effect on the score in Italian for children who took INVALSI tests in 2016, but negative for those who took tests in the years 2014 and 2017. Once again, a more pronounced effect of ordinary spending compared to extraordinary spending is observed across both missions, likely because extraordinary spending does not occur regularly and involves larger-scale interventions with long-term effects. Additionally, the effects are primarily concentrated in the years 2014-2016, probably because most of the elections took place in 2014 and, to a lesser extent, in 2016.

7. Conclusion

The importance of public spending on preschool education and childcare is crucial for several reasons, particularly when considering the long-term effects on students' educational outcomes, such as those measured by INVALSI scores. Investing in the early stages of education lays a strong foundation for children's cognitive and socio-emotional development, better equipping them for the challenges they will face later in their schooling. Women in municipal councils are often more inclined to allocate resources towards preschool education programs, recognizing the importance of quality early education for the long-term social and economic development of the community. A greater representation of women in these councils can initiate a process that not only promotes greater equity in political representation but also significantly influences the composition of public spending, particularly in the area of preschool education and childcare. The introduction of Law 215/2012, which promotes a more balanced gender representation in decision-making processes, has led to greater gender diversity in municipal councils. This diversity often results in a heightened sensitivity to social issues, such as education and child welfare, which are frequently prioritized in women's political agendas. Consequently, an increased presence of women in municipal councils can positively influence decisions on public spending related to preschool education and childcare, as evidenced in the literature reviewed. Our study stands out for its combined analysis of the effects of public spending on the academic performance of second-grade children and the impact of the introduction of Law 215/2012, which established gender-based double preference in municipalities with over 5,000 inhabitants. The introduction of the gender quota reform caused an exogenous change in the gender composition of municipal councils in municipalities that held municipal elections in 2013 or later (Andreoli et al., 2022). The allocation of spending on preschool education and interventions in favor of childhood, nurseries, and minors, may be influenced by the presence of a greater number of women on municipal councils, as these issues are often considered to be of particular interest to women. In this context, we examine the effect of local public spending on

preschool education (expenditure category under Mission 4, Programme 01) and interventions in favor of childhood, nurseries, and minors (expenditure category under Mission 12, Programme 01) on the academic performance in Italian and mathematics of second-grade children. We use an instrumental variables model to address the endogeneity issues related to spending variables, leveraging the introduction of the gender-based double preference as a source of exogenous variation. The results of the first stage show that a greater female presence on municipal councils tends to increase local (ordinary) public spending on preschool education and childcare services. Therefore, the estimated effects on municipal expenditure are reinforced by the increase in the share of women. The results of the second stage indicate that an increase in ordinary public spending per student in preschool education and early childhood services, such as day care and minors, has a positive impact on INVALSI scores in mathematics. However, replicating the analysis using the INVALSI score in Italian as the dependent variable, we do not observe the same positive effect found in math scores. Robustness checks largely confirm the results of the main analysis, especially for the period after 2014 (the year in which most Italian municipalities held elections) up until 2016. This work can make a significant academic contribution to the existing literature, highlighting how targeted and adequate investment in the early stages of education can create a better-prepared and more competitive human capital, with benefits that extend not only to individuals but to society as a whole.

8. Appendix

TABLE A1 *Variable Descriptions*

Variable	Description
Dependent variable	
Standardized second-grade math test score[a] Standardized second-grade Italian test score[a]	*Corrected total percentage score for cheating in Italian and Mathematics*
Invalsi Variables	
Gender[a]	Dummy variable that takes the value 1 if the student is Male, 0 otherwise
Nationality[a]	Dummy variable that takes the value 1 if the student is native, 0 otherwise
Regular studies[a]	Dummy variable that takes the value 1 if the student is regular or anticipatory, 0 otherwise
Education Father[a]	Dummies for the respective qualifications: Bachelor's degree (high), diploma (mid) and lower secondary school diploma (low)
Education Mother[a]	Dummy variables for the respective qualifications: Bachelor's degree (high), diploma (mid) and lower secondary school diploma (low)
EXP	Years of exposure of the child to Law 215/2012 during grade 2.
Balance sheet variables	
EXPEND[1, ab]	Mission 4, Programme 01 (Preschool education) in accrual payments: per capita ordinary and extraordinary expenditure in relation to the population aged 0-5 years.
EXPxEXPEND[1]	Interaction terms between spending per pupil (distinguishing between ordinary and extraordinary expenditure) and the exposure to the reform the interaction terms between spending and exposure to the reform
EXPEND[2, ab]	Mission 12, Programme 01 (Childhood, kindergartens and minors) in accrual payments: per capita ordinary and extraordinary expenditure in relation to the population aged 0-5 years.
EXPxEXPEND[2]	Interaction terms between spending per pupil (distinguishing between ordinary and extraordinary expenditure) and the exposure to the reform the interaction terms between spending and exposure to the reform
Demographic variables	
Population[d]	Resident population in the municipality as at 31 December of year t
Pop 0-5[d]	Population by age group 0-5 years resident in the municipality on 31 December of year t
Electoral Variables	

Female share[e]	Proportion of women in the municipal government. The share of women was calculated as the average of the share of women defined for each year.
Female Presence	Change in the share of women in the years 2012-2018
EXPxFemalePresence	The instrument for endogenous interaction term EXPxEXPEND[1,2] (in this case as well, a distinction is made between ordinary and extraordinary expenses).

Notes. [b] Expenditure expressed in euro units, per capita and deflated using the GDP deflator from the St. Louis FRED (https://fred.stlouisfed.org/) base year 2015
Data source:
[a] *INVALSI SERVIZIO STATISTICO* (https://serviziostatistico.invalsi.it/catalogo-dati/)
[c] *AIDA PA (*https://login.bvdinfo.com/R1/AidaPANeo*)*
[d] *ISTAT (*http://demo.istat.it/*)*
[e] *Ministero dell'Interno – archivio dei risultati elettorali (*http://elezionistorico.interno.it*)*

TABLE A2 *Child's exposure time to law 215/2012 (EXP)*

INTRODUCTION OF THE LAW 215/2012	EXPOSURE TO REFORM	YEAR BIRTH	YEAR[17]
26 DECEMBER 2012	0	2006	2013
	1	2007	2014
	2	2008	2015
	3	2009	2016
	4	2010	2017
	5	2011	2018

Notes. Anyone born in 2006 cannot be affected by the reform, as it came into effect on December 26, 2012, and this cohort took the tests in 2013. Those born in 2007 have about a year of exposure to the reform, and so on.

[17] Year in wich they support the INVALSI test at Grade 2

TABLE A3 *Instruments validation (Math Grade II)*

COLUMN 1

Panel A: Durbin-Wu-Hausman test
H_0: OLS is consistent and fully efficient
H_1: IV must be preferred

Robust score chi2(2) = 105.695 (p = 0.0000)
Robust regression F(2,771305) = 52.8678 (p = 0.0000)

Panel B: Stock-Wright LM S test
H_0: Endogenous regressors are jointly equal to zero
H_1: Endogenous regressors are different than zero

Anderson-Rubin Wald test Chi-sq(2)= 65.30 P-val=0.0000
Stock-Wright LM S statistic Chi-sq(2)= 65.33 P-val=0.0000

Panel C: Kleibergen-Paap rk LM test
H_0: Matrix of reduced form coefficients has not full rank
H_1: Matrix of reduced form coefficients has full rank

Underidentification test (Kleibergen-Paap rk LM statistic): 14000
Chi-sq(1) P-val = 0.0000

Notes. The Durbin-Wu-Hausman test offers a test statistic for the null hypothesis that the OLS estimator is reliable and entirely efficient (and thus there are no biases from endogeneity). If the null hypothesis is rejected, the IV estimator should be favored as it is reliable. The OLS estimator is entirely efficient under the null hypothesis, but unreliable if the null hypothesis is false; the IV estimator is reliable both under the null hypothesis and the alternative hypothesis. For the Stock-Wright S LM test and the Anderson-Rubin Wald test, the null hypothesis in both cases is that the coefficients of the endogenous regressors in the structural equation are jointly equal to zero. Both tests are robust to the presence of weak instruments. The Kleibergen-Paap rk LM test allows us to assess whether the model is not underidentified and is adequately identified. The Hansen J statistic is not reported because the model is exactly identified.

TABLE A4 *Instruments validation (Math Grade II)*

COLUMN 2

Panel A: Durbin-Wu-Hausman test
H_0: OLS is consistent and fully efficient
H_1: IV must be preferred

Robust score chi2(2) = 69.7049 (p = 0.0000)
Robust regression F(2,771305) = 34.8409 (p = 0.0000)

Panel B: Stock-Wright LM S test
H_0: Endogenous regressors are jointly equal to zero
H_1: Endogenous regressors are different than zero

Anderson-Rubin Wald test Chi-sq(2)= 65.30 P-val=0.0000

Stock-Wright LM S statistic Chi-sq(2)= 65.33 P-val=0.0000

Panel C: Kleibergen-Paap rk LM test
H_0: Matrix of reduced form coefficients has not full rank
H_1: Matrix of reduced form coefficients has full rank

Underidentification test (Kleibergen-Paap rk LM statistic): 540.277

Chi-sq(1) P-val = 0.0000

Notes. The Durbin-Wu-Hausman test offers a test statistic for the null hypothesis that the OLS estimator is reliable and entirely efficient (and thus there are no biases from endogeneity). If the null hypothesis is rejected, the IV estimator should be favored as it is reliable. The OLS estimator is entirely efficient under the null hypothesis, but unreliable if the null hypothesis is false; the IV estimator is reliable both under the null hypothesis and the alternative hypothesis. For the Stock-Wright S LM test and the Anderson-Rubin Wald test, the null hypothesis in both cases is that the coefficients of the endogenous regressors in the structural equation are jointly equal to zero. Both tests are robust to the presence of weak instruments. The Kleibergen-Paap rk LM test allows us to assess whether the model is not underidentified and is adequately identified. The Hansen J statistic is not reported because the model is exactly identified.

TABLE A5 *Instruments validation (Math Grade II)*

COLUMN 3
Panel A: Durbin-Wu-Hausman test
H_0: OLS is consistent and fully efficient
H_1: IV must be preferred

Robust score chi2(2) = 87.2462 (p = 0.0000)
Robust regression F(2,771305) = 43.6271 (p = 0.0000)

Panel B: Stock-Wright LM S test
H_0: Endogenous regressors are jointly equal to zero
H_1: Endogenous regressors are different than zero

Anderson-Rubin Wald test Chi-sq(2)= 65.30 P-val=0.0000

Stock-Wright LM S statistic Chi-sq(2)= 65.33 P-val=0.0000

Panel C: Kleibergen-Paap rk LM test
H_0: Matrix of reduced form coefficients has not full rank
H_1: Matrix of reduced form coefficients has full rank

Underidentification test (Kleibergen-Paap rk LM statistic): 9123.065

Chi-sq(1) P-val = 0.0000

Notes. The Durbin-Wu-Hausman test offers a test statistic for the null hypothesis that the OLS estimator is reliable and entirely efficient (and thus there are no biases from endogeneity). If the null hypothesis is rejected, the IV estimator should be favored as it is reliable. The OLS estimator is entirely efficient under the null hypothesis, but unreliable if the null hypothesis is false; the IV estimator is reliable both under the null hypothesis and the alternative hypothesis. For the Stock-Wright S LM test and the Anderson-Rubin Wald test, the null hypothesis in both cases is that the coefficients of the endogenous regressors in the structural equation are jointly equal to zero. Both tests are robust to the presence of weak instruments. The Kleibergen-Paap rk LM test allows us to assess whether the model is not underidentified and is adequately identified. The Hansen J statistic is not reported because the model is exactly identified.

TABLE A6 *Instruments validation (Math Grade II)*

COLUMN 4
Panel A: Durbin-Wu-Hausman test H_0: OLS is consistent and fully efficient H_1: IV must be preferred Robust score chi2(2) = 66.3346 (p = 0.0000) Robust regression F(2,771305) = 33.1561 (p = 0.0000)
Panel B: Stock-Wright LM S test H_0: Endogenous regressors are jointly equal to zero H_1: Endogenous regressors are different than zero Anderson-Rubin Wald test Chi-sq(2)= 65.30 P-val=0.0000 Stock-Wright LM S statistic Chi-sq(2)= 65.33 P-val=0.0000
Panel C: Kleibergen-Paap rk LM test H_0: Matrix of reduced form coefficients has not full rank H_1: Matrix of reduced form coefficients has full rank Underidentification test (Kleibergen-Paap rk LM statistic): 74.355 Chi-sq(1) P-val = 0.0000

Notes. The Durbin-Wu-Hausman test offers a test statistic for the null hypothesis that the OLS estimator is reliable and entirely efficient (and thus there are no biases from endogeneity). If the null hypothesis is rejected, the IV estimator should be favored as it is reliable. The OLS estimator is entirely efficient under the null hypothesis, but unreliable if the null hypothesis is false; the IV estimator is reliable both under the null hypothesis and the alternative hypothesis. For the Stock-Wright S LM test and the Anderson-Rubin Wald test, the null hypothesis in both cases is that the coefficients of the endogenous regressors in the structural equation are jointly equal to zero. Both tests are robust to the presence of weak instruments. The Kleibergen-Paap rk LM test allows us to assess whether the model is not underidentified and is adequately identified. The Hansen J statistic is not reported because the model is exactly identified.

TABLE A7 *First stage related to the Standardized second-grade Italian test score*
Outcome: expenditure per pupil in preschool education and expenditure in childcare, kindergartens, and minors

	(1) EXPEND^{1ord}	(2) EXPEND^{1extra}	(3) EXPEND^{2ord}	(4) EXPEND^{2extra}
EXPxFemale presence^{IV}	21.89***	-4.719***	-14.92***	1.967***
	(1.236)	(0.093)	(1.336)	(0.061)
Female presence^{IV}	**733.5*** **	**0.196**	**793.2*** **	**-2.191*** **
	(6.198)	(0.478)	(6.527)	(0.314)
EXP	-0.297	4.662***	34.40***	0.658***
	(0.379)	(0.048)	(0.364)	(0.017)
_cons	479.5***	-5.252***	520.1***	1.331***
	(6.142)	(0.894)	(6.322)	(0.374)
N	764956	764956	764956	764956
Kleibergen-Paap rk Wald statistic:	9718.60	280.62	6058.64	39.04
Controls	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes
Province Effects	Yes	Yes	Yes	Yes

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$
Note: EXPEND^{1ord} and EXPEND^{1extra} is the ordinary and extraordinary expenditure per pupil in preschool education (Mission 4, Programme 01), expressed in 2015 euros. EXPEND^{2ord} and EXPEND^{2extra} is the ordinary and extraordinary expenditure per pupil in childcare, kindergartens and minors (Mission 12, Programme 01), expressed in 2015 euros. Exposure is the time of exposure of the child to the reform. The main instrument for each First-stage regression is written in bold. Standard errors are adjusted for heteroscedasticity.

TABLE A8 *First stage related to the Standardized second-grade Italian test score*
Outcome: interaction term of the expenditure variables per pupil with the child's exposure time to the law

	(1) EXPxEXPEND^{1ord}	(2) EXPxEXPEND^{1extra}	(3) EXPxEXPEND^{2ord}	(4) EXPxEXPEND^{2extra}
EXPxFemale presence^{IV}	**806.6*** **	**-39.84*** **	**458.6*** **	**1.041*** **
	(5.096)	(0.377)	(5.672)	(0.181)
Female presence^{IV}	201.4***	63.15***	990.0***	5.070***
	(17.943)	(1.516)	(20.467)	(0.724)
EXP	354.5***	31.03***	771.6***	8.540***
	(1.636)	(0.222)	(1.899)	(0.051)
_cons	312.4***	-44.56***	-309.0***	-6.097***
	(27.676)	(3.117)	(33.324)	(1.142)
N	764956	764956	764956	764956
Kleibergen-Paap rk Wald statistic:	9718.60	280.62	6058.64	39.04
Controls	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes
Province Effects	Yes	Yes	Yes	Yes

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$
Note: EXPEND^{1ord} and EXPEND^{1extra} is the ordinary and extraordinary expenditure per pupil in preschool education (Mission 4, Programme 01), expressed in 2015 euros. EXPEND^{2ord} and EXPEND^{2extra} is the ordinary and extraordinary expenditure per pupil in childcare, kindergartens and minors (Mission 12, Programme 01), expressed in 2015 euros. Exposure is the time of exposure of the child to the reform. The main instrument for each First-stage regression is written in bold. Standard errors are adjusted for heteroscedasticity.

TABLE A9 *Instruments validation (Italian Grade II)*

COLUMN 1

Panel A: Durbin-Wu-Hausman test
H_0: OLS is consistent and fully efficient
H_1: IV must be preferred

Robust score chi2(2) = 31.768 (p = 0.0000)
Robust regression F(2,764852) = 15.8889 (p = 0.0000)

Panel B: Stock-Wright LM S test
H_0: Endogenous regressors are jointly equal to zero
H_1: Endogenous regressors are different than zero

Anderson-Rubin Wald test Chi-sq(2)= 22.11 P-val=0.0000
Stock-Wright LM S statistic Chi-sq(2)= 22.12 P-val=0.0000

Panel C: Kleibergen-Paap rk LM test
H_0: Matrix of reduced form coefficients has not full rank
H_1: Matrix of reduced form coefficients has full rank

Underidentification test (Kleibergen-Paap rk LM statistic): 1.4e+04

Chi-sq(1) P-val = 0.0000

Notes. The Durbin-Wu-Hausman test offers a test statistic for the null hypothesis that the OLS estimator is reliable and entirely efficient (and thus there are no biases from endogeneity). If the null hypothesis is rejected, the IV estimator should be favored as it is reliable. The OLS estimator is entirely efficient under the null hypothesis, but unreliable if the null hypothesis is false; the IV estimator is reliable both under the null hypothesis and the alternative hypothesis. For the Stock-Wright S LM test and the Anderson-Rubin Wald test, the null hypothesis in both cases is that the coefficients of the endogenous regressors in the structural equation are jointly equal to zero. Both tests are robust to the presence of weak instruments. The Kleibergen-Paap rk LM test allows us to assess whether the model is not underidentified and is adequately identified. The Hansen J statistic is not reported because the model is exactly identified.

TABLE A10 *Instruments validation (Italian Grade II)*

COLUMN 2

Panel A: Durbin-Wu-Hausman test
H_0: OLS is consistent and fully efficient
H_1: IV must be preferred
Robust score chi2(2) = 24.1315 (p = 0.0000)
Robust regression F(2,764852) = 12.0673 (p = 0.0000)

Panel B: Stock-Wright LM S test
H_0: Endogenous regressors are jointly equal to zero
H_1: Endogenous regressors are different than zero

Anderson-Rubin Wald test Chi-sq(2)= 22.11 P-val=0.0000
Stock-Wright LM S statistic Chi-sq(2)= 22.12 P-val=0.0000

Panel C: Kleibergen-Paap rk LM test
H_0: Matrix of reduced form coefficients has not full rank
H_1: Matrix of reduced form coefficients has full rank

Underidentification test (Kleibergen-Paap rk LM statistic): 562.341
Chi-sq(1) P-val = 0.0000

Notes. The Durbin-Wu-Hausman test offers a test statistic for the null hypothesis that the OLS estimator is reliable and entirely efficient (and thus there are no biases from endogeneity). If the null hypothesis is rejected, the IV estimator should be favored as it is reliable. The OLS estimator is entirely efficient under the null hypothesis, but unreliable if the null hypothesis is false; the IV estimator is reliable both under the null hypothesis and the alternative hypothesis. For the Stock-Wright S LM test and the Anderson-Rubin Wald test, the null hypothesis in both cases is that the coefficients of the endogenous regressors in the structural equation are jointly equal to zero.

Both tests are robust to the presence of weak instruments. The Kleibergen-Paap rk LM test allows us to assess whether the model is not underidentified and is adequately identified. The Hansen J statistic is not reported because the model is exactly identified.

TABLE A11 *Instruments validation (Italian Grade II)*

COLUMN 3

Panel A: Durbin-Wu-Hausman test
H_0: OLS is consistent and fully efficient
H_1: IV must be preferred

Robust score chi2(2) = 26.959 (p = 0.0000)
Robust regression F(2,764852) = 13.4823 (p = 0.0000)

Panel B: Stock-Wright LM S test
H_0: Endogenous regressors are jointly equal to zero
H_1: Endogenous regressors are different than zero

Anderson-Rubin Wald test Chi-sq(2)= 22.11 P-val=0.0000
Stock-Wright LM S statistic Chi-sq(2)= 22.12 P-val=0.0000

Panel C: Kleibergen-Paap rk LM test
H_0: Matrix of reduced form coefficients has not full rank
H_1: Matrix of reduced form coefficients has full rank

Underidentification test (Kleibergen-Paap rk LM statistic): 9028.398
Chi-sq(1) P-val = 0.0000

Notes. The Durbin-Wu-Hausman test offers a test statistic for the null hypothesis that the OLS estimator is reliable and entirely efficient (and thus there are no biases from endogeneity). If the null hypothesis is rejected, the IV estimator should be favored as it is reliable. The OLS estimator is entirely efficient under the null hypothesis, but unreliable if the null hypothesis is false; the IV estimator is reliable both under the null hypothesis and the alternative hypothesis. For the Stock-Wright S LM test and the Anderson-Rubin Wald test, the null hypothesis in both cases is that the coefficients of the endogenous regressors in the structural equation are jointly equal to zero. Both tests are robust to the presence of weak instruments. The Kleibergen-Paap rk LM test allows us to assess whether the model is not underidentified and is adequately identified. The Hansen J statistic is not reported because the model is exactly identified.

TABLE A12 *Instruments validation (Italian Grade II)*

COLUMN 4

Panel A: Durbin-Wu-Hausman test
H_0: OLS is consistent and fully efficient
H_1: IV must be preferred

Robust score chi2(2) = 23.0444 (p = 0.0000)
Robust regression F(2,764852) = 11.5237 (p = 0.0000)

Panel B: Stock-Wright LM S test
H_0: Endogenous regressors are jointly equal to zero
H_1: Endogenous regressors are different than zero

Anderson-Rubin Wald test Chi-sq(2)= 22.11 P-val=0.0000
Stock-Wright LM S statistic Chi-sq(2)= 22.12 P-val=0.0000

Panel C: Kleibergen-Paap rk LM test
H_0: Matrix of reduced form coefficients has not full rank
H_1: Matrix of reduced form coefficients has full rank

Underidentification test (Kleibergen-Paap rk LM statistic): 77.779
Chi-sq(1) P-val = 0.0000

Notes. The Durbin-Wu-Hausman test offers a test statistic for the null hypothesis that the OLS estimator is reliable and entirely efficient (and thus there are no biases from endogeneity). If the null hypothesis is rejected, the IV estimator should be favored as it is reliable.

The OLS estimator is entirely efficient under the null hypothesis, but unreliable if the null hypothesis is false; the IV estimator is reliable both under the null hypothesis and the alternative hypothesis. For the Stock-Wright S LM test and the Anderson-Rubin Wald test, the null hypothesis in both cases is that the coefficients of the endogenous regressors in the structural equation are jointly equal to zero. Both tests are robust to the presence of weak instruments. The Kleibergen-Paap rk LM test allows us to assess whether the model is not underidentified and is adequately identified. The Hansen J statistic is not reported because the model is exactly identified.

TABLE A13 *Robustness checks*

	(EXP=1) Standardized second-grade Math test score (adjusted for cheating)	(EXP=2) Standardized second-grade Math test score (adjusted for cheating)	(EXP=3) Standardized second-grade Math test score (adjusted for cheating)	(EXP=4) Standardized second-grade Math test score (adjusted for cheating)	(EXP=5) Standardized second-grade Math test score (adjusted for cheating)
$EXPEND^{1ord\ IV}$	0.00405***	0.00277**	0.00224**	0.000980	0.000278
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
_cons	40.36***	39.70***	45.97***	36.67***	44.76***
	(1.477)	(1.457)	(1.400)	(1.287)	(1.346)
N	122894	134956	133789	126819	122267
R^{2u}	0.894	0.887	0.904	0.885	0.916
R^{2c}	0.0706	0.0581	0.0796	0.0664	0.0870
Controls	Yes	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes	Yes
Province Effects	Yes	Yes	Yes	Yes	Yes
Durbin-Wu-Hausman test Chi-Squared	33.1498	15.0105	14.3444	6.52698	5.28489
P value	0.0000	0.0001	0.0002	0.0106	0.0215

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$

Note: Dependent variable is the corrected total percentage score for cheating in Mathematics. $EXPEND^{1ord}$ is the ordinary expenditure per pupil in preschool education (Mission 4, Programme 01), expressed in 2015 euros. The expenditure variables is instrumented by the change in the proportion of women induced by Law 215/2012. Exp represents the number of years a child has been exposed to the reform. For example, EXP=1 indicates that at the time of the INVALSI tests in 2014, the child had been exposed to the reform for only one year, and so on. The sample expands over the period 2013–2018. Standard errors are robust to heteroscedasticity. The Durbin-Wu-Hausman test offers a test statistic for the null hypothesis that the OLS estimator is reliable and entirely efficient (and thus there are no biases from endogeneity). If the null hypothesis is rejected, the IV estimator should be favored as it is reliable. The OLS estimator is entirely efficient under the null hypothesis, but unreliable if the null hypothesis is false; the IV estimator is reliable both under the null hypothesis and the alternative hypothesis.

TABLE A14 *Robustness checks*

	(EXP=1) Standardized second-grade math test score (adjusted for cheating)	(EXP=2) Standardized second-grade math test score (adjusted for cheating)	(EXP=3) Standardized second-grade math test score (adjusted for cheating)	(EXP=4) Standardized second-grade math test score (adjusted for cheating)	(EXP=5) Standardized second-grade math test score (adjusted for cheating)
$EXPEND^{1extra\ IV}$	-0.284***	-0.155**	0.177**	-0.0426	-0.0116
	(0.058)	(0.050)	(0.066)	(0.029)	(0.024)
_cons	46.61***	43.67***	43.05***	37.37***	44.86***
	(1.602)	(1.640)	(2.139)	(1.241)	(1.314)
N	122894	134956	133789	126819	122267
R^{2u}	0.879	0.873	0.893	0.884	0.916
R^{2c}	-0.0626	-0.0579	-0.0265	0.0589	0.0861
Controls	Yes	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes	Yes
Province Effects	Yes	Yes	Yes	Yes	Yes
Durbin-Wu-Hausman test Chi-Squared	26.742	11.1565	8.17465	2.45793	.279198

P value	0.0000	0.0008	0.0042	0.1169	0.5972

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$
Note: Dependent variable is the corrected total percentage score for cheating in Mathematics. EXPEND^{1extra} is the extraordinary expenditure per pupil in preschool education (Mission 4, Programme 01), expressed in 2015 euros. The expenditure variables is instrumented by the change in the proportion of women induced by Law 215/2012. Exp represents the number of years a child has been exposed to the reform. For example, EXP=1 indicates that at the time of the INVALSI tests in 2014, the child had been exposed to the reform for only one year, and so on. The sample expands over the period 2013–2018. Standard errors are robust to heteroscedasticity. The Durbin-Wu-Hausman test offers a test statistic for the null hypothesis that the OLS estimator is reliable and entirely efficient (and thus there are no biases from endogeneity). If the null hypothesis is rejected, the IV estimator should be favored as it is reliable. The OLS estimator is entirely efficient under the null hypothesis, but unreliable if the null hypothesis is false; the IV estimator is reliable both under the null hypothesis and the alternative hypothesis.

TABLE A15 *Robustness checks*

	(EXP=1) Standardized second-grade Math test score (adjusted for cheating)	(EXP=2) Standardized second-grade Math test score (adjusted for cheating)	(EXP=3) Standardized second-grade Math test score (adjusted for cheating)	(EXP=4) Standardized second-grade Math test score (adjusted for cheating)	(EXP=5) Standardized second-grade Math test score (adjusted for cheating)
EXPEND$^{2ord\ IV}$	0.00444***	0.00327**	0.00225**	0.000948	0.000287
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
_cons	39.67***	38.91***	45.86***	36.63***	44.72***
	(1.526)	(1.549)	(1.413)	(1.294)	(1.366)
N	122894	134956	133789	126819	122267
R^{2u}	0.894	0.887	0.904	0.885	0.916
R^{2c}	0.0717	0.0583	0.0792	0.0666	0.0869
Controls	Yes	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes	Yes
Province Effects	Yes	Yes	Yes	Yes	Yes
Durbin-Wu-Hausman test Chi-Squared	26.8927	11.3595	19.2076	4.23563	5.81618
P value	0.0000	0.0008	0.0000	0.0396	0.0159

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$
Note: Dependent variable is the corrected total percentage score for cheating in Mathematics. EXPEND^{2ord} is the ordinary expenditure per pupil on interventions in favour of children, kindergartens and minors (Mission 12, Programme 01), expressed in 2015 euros. The expenditure variables is instrumented by the change in the proportion of women induced by Law 215/2012. Exp represents the number of years a child has been exposed to the reform. For example, EXP=1 indicates that at the time of the INVALSI tests in 2014, the child had been exposed to the reform for only one year, and so on. The sample expands over the period 2013–2018. Standard errors are robust to heteroscedasticity. The Durbin-Wu-Hausman test offers a test statistic for the null hypothesis that the OLS estimator is reliable and entirely efficient (and thus there are no biases from endogeneity). If the null hypothesis is rejected, the IV estimator should be favored as it is reliable. The OLS estimator is entirely efficient under the null hypothesis, but unreliable if the null hypothesis is false; the IV estimator is reliable both under the null hypothesis and the alternative hypothesis.

TABLE A16 *Robustness checks*

	(EXP=1) Standardized second-grade Math test score (adjusted for cheating)	(EXP=2) Standardized second-grade Math test score (adjusted for cheating)	(EXP=3) Standardized second-grade Math test score (adjusted for cheating)	(EXP=4) Standardized second-grade Math test score (adjusted for cheating)	(EXP=5) Standardized second-grade Math test score (adjusted for cheating)
EXPEND$^{2extra\ IV}$	-0.398***	0.155**	-0.522**	-0.122	0.0210
	(0.082)	(0.049)	(0.194)	(0.083)	(0.044)
_cons	43.98***	40.58***	54.18***	37.45***	44.82***
	(1.420)	(1.412)	(2.853)	(1.245)	(1.324)
N	122894	134956	133789	126819	122267
R^{2u}	0.874	0.884	0.894	0.884	0.916
R^{2c}	-0.103	0.0356	-0.0207	0.0618	0.0863
Controls	Yes	Yes	Yes	Yes	Yes

Year Effects	Yes	Yes	Yes	Yes	Yes
Province Effects	Yes	Yes	Yes	Yes	Yes
Durbin-Wu-Hausman test Chi-Squared	27.0751	10.2617	7.53338	2.23701	.910754
P value	0.0000	0.0014	0.0061	0.1347	0.3399

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$

Note: Dependent variable is the corrected total percentage score for cheating in Mathematics. $EXPEND^{2extra}$ is the extraordinary expenditure per pupil on interventions in favour of children, kindergartens and minors (Mission 12, Programme 01), expressed in 2015 euros. The expenditure variables is instrumented by the change in the proportion of women induced by Law 215/2012. Exp represents the number of years a child has been exposed to the reform. For example, EXP=1 indicates that at the time of the INVALSI tests in 2014, the child had been exposed to the reform for only one year, and so on. The sample expands over the period 2013–2018. Standard errors are robust to heteroscedasticity. The Durbin-Wu-Hausman test offers a test statistic for the null hypothesis that the OLS estimator is reliable and entirely efficient (and thus there are no biases from endogeneity). If the null hypothesis is rejected, the IV estimator should be favored as it is reliable. The OLS estimator is entirely efficient under the null hypothesis, but unreliable if the null hypothesis is false; the IV estimator is reliable both under the null hypothesis and the alternative hypothesis.

TABLE A17 *Robustness checks*

	(EXP=1) Standardized second-grade Italian test score (adjusted for cheating)	(EXP=2) Standardized second-grade Italian test score (adjusted for cheating)	(EXP=3) Standardized second-grade Italian test score (adjusted for cheating)	(EXP=4) Standardized second-grade Italian test score (adjusted for cheating)	(EXP=5) Standardized second-grade Italian test score (adjusted for cheating)
$EXPEND^{1ord\ IV}$	0.00149*	-0.000425	0.00243**	0.00268***	-0.0000219
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
_cons	38.28***	30.62***	30.00***	41.81***	41.27***
	(1.607)	(1.532)	(1.516)	(1.331)	(1.418)
N	115306	135255	132866	129662	121426
R^{2u}	0.902	0.857	0.838	0.882	0.894
R^{2c}	0.0830	0.0782	0.0789	0.0677	0.0902
Controls	Yes	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes	Yes
Province Effects	Yes	Yes	Yes	Yes	Yes
Durbin-Wu-Hausman test Chi-Squared	4.68471	.013457	9.53797	16.7289	.68281
P value	0.0304	0.9076	0.0020	0.0000	0.4088

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$

Note: Dependent variable is the corrected total percentage score for cheating in Italian. $EXPEND^{1ord}$ is the ordinary expenditure per pupil in preschool education (Mission 4, Programme 01), expressed in 2015 euros. The expenditure variables is instrumented by the change in the proportion of women induced by Law 215/2012. Exp represents the number of years a child has been exposed to the reform. For example, EXP=1 indicates that at the time of the INVALSI tests in 2014, the child had been exposed to the reform for only one year, and so on. The sample expands over the period 2013–2018. Standard errors are robust to heteroscedasticity. The Durbin-Wu-Hausman test offers a test statistic for the null hypothesis that the OLS estimator is reliable and entirely efficient (and thus there are no biases from endogeneity). If the null hypothesis is rejected, the IV estimator should be favored as it is reliable. The OLS estimator is entirely efficient under the null hypothesis, but unreliable if the null hypothesis is false; the IV estimator is reliable both under the null hypothesis and the alternative hypothesis.

TABLE A18 *Robustness checks*

	(EXP=1) Standardized second-grade Italian test score (adjusted for cheating)	(EXP=2) Standardized second-grade Italian test score (adjusted for cheating)	(EXP=3) Standardized second-grade Italian test score (adjusted for cheating)	(EXP=4) Standardized second-grade Italian test score (adjusted for cheating)	(EXP=5) Standardized second-grade Italian test score (adjusted for cheating)
$EXPEND^{1extra\ IV}$	-0.106*	0.0229	0.239**	-0.120***	0.000922
	(0.060)	(0.051)	(0.096)	(0.032)	(0.027)
_cons	40.77***	30.03***	25.51***	43.71***	41.26***
	(1.722)	(1.685)	(2.908)	(1.309)	(1.381)
N	115306	135255	132866	129662	121426
R^{2u}	0.899	0.857	0.811	0.875	0.894
R^{2c}	0.0627	0.0767	-0.0767	0.0176	0.0902
Controls	Yes	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes	Yes
Province Effects	Yes	Yes	Yes	Yes	Yes
Durbin-Wu-Hausman test Chi-Squared	3.29505	.157702	7.39624	15.6132	.000637
P value	0.0695	0.6913	0.0065	0.0001	0.9799

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$

Note: Dependent variable is the corrected total percentage score for cheating in Italian. $EXPEND^{1extra}$ is the extraordinary expenditure per pupil in preschool education (Mission 4, Programme 01), expressed in 2015 euros. The expenditure variables is instrumented by the change in the proportion of women induced by Law 215/2012. Exp represents the number of years a child has been exposed to the reform. For example, EXP=1 indicates that at the time of the INVALSI tests in 2014, the child had been exposed to the reform for only one year, and so on. The sample expands over the period 2013–2018. Standard errors are robust to heteroscedasticity. The Durbin-Wu-Hausman test offers a test statistic for the null hypothesis that the OLS estimator is reliable and entirely efficient (and thus there are no biases from endogeneity). If the null hypothesis is rejected, the IV estimator should be favored as it is reliable. The OLS estimator is entirely efficient under the null hypothesis, but unreliable if the null hypothesis is false; the IV estimator is reliable both under the null hypothesis and the alternative hypothesis.

TABLE A19 *Robustness checks*

	(EXP=1) Standardized second-grade Italian test score (adjusted for cheating)	(EXP=2) Standardized second-grade Italian test score (adjusted for cheating)	(EXP=3) Standardized second-grade Italian test score (adjusted for cheating)	(EXP=4) Standardized second-grade Italian test score (adjusted for cheating)	(EXP=5) Standardized second-grade Italian test score (adjusted for cheating)
$EXPEND^{2ord\ IV}$	0.00163*	-0.000495	0.00244**	0.00257***	-0.0000227
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
_cons	38.04***	30.74***	29.88***	41.71***	41.27***
	(1.658)	(1.632)	(1.531)	(1.338)	(1.442)
N	115306	135255	132866	129662	121426
R^{2u}	0.902	0.857	0.838	0.882	0.894
R^{2c}	0.0830	0.0782	0.0788	0.0679	0.0902
Controls	Yes	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes	Yes
Province Effects	Yes	Yes	Yes	Yes	Yes
Durbin-Wu-Hausman test Chi-Squared	4.70061	.134278	10.9784	15.0233	.361432
P value	0.0302	0.7140	0.0009	0.0001	0.5477

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$

Note: Dependent variable is the corrected total percentage score for cheating in Italian. $EXPEND^{2ord}$ is the ordinary expenditure per pupil on

interventions in favour of children, kindergartens and minors (Mission 12, Programme 01), expressed in 2015 euros. The expenditure variables is instrumented by the change in the proportion of women induced by Law 215/2012. Exp represents the number of years a child has been exposed to the reform. For example, EXP=1 indicates that at the time of the INVALSI tests in 2014, the child had been exposed to the reform for only one year, and so on. The sample expands over the period 2013–2018. Standard errors are robust to heteroscedasticity. The Durbin-Wu-Hausman test offers a test statistic for the null hypothesis that the OLS estimator is reliable and entirely efficient (and thus there are no biases from endogeneity). If the null hypothesis is rejected, the IV estimator should be favored as it is reliable. The OLS estimator is entirely efficient under the null hypothesis, but unreliable if the null hypothesis is false; the IV estimator is reliable both under the null hypothesis and the alternative hypothesis.

TABLE A20 Robustness checks

	(EXP=1) Standardized second-grade Italian test score (adjusted for cheating)	(EXP=2) Standardized second-grade Italian test score (adjusted for cheating)	(EXP=3) Standardized second-grade Italian test score (adjusted for cheating)	(EXP=4) Standardized second-grade Italian test score (adjusted for cheating)	(EXP=5) Standardized second-grade Italian test score (adjusted for cheating)
$EXPEND^{2extra\ IV}$	-0.154*	-0.0225	-0.535**	-0.328***	-0.00163
	(0.087)	(0.050)	(0.206)	(0.087)	(0.048)
_cons	39.71***	30.49***	38.64***	43.99***	41.26***
	(1.526)	(1.468)	(3.093)	(1.286)	(1.391)
N	115306	135255	132866	129662	121426
R^{2u}	0.899	0.857	0.822	0.878	0.894
R^{2c}	0.0592	0.0779	-0.0136	0.0382	0.0902
Controls	Yes	Yes	Yes	Yes	Yes
Year Effects	Yes	Yes	Yes	Yes	Yes
Province Effects	Yes	Yes	Yes	Yes	Yes
Durbin-Wu-Hausman test Chi-Squared	2.76592	.162059	7.23511	14.7061	.002059
P value	0.0963	0.6873	0.0071	0.0001	0.9638

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$

Note: Dependent variable is the corrected total percentage score for cheating in Italian. $EXPEND^{2extra\ IV}$ is the extraordinary expenditure per pupil on interventions in favour of children, kindergartens and minors (Mission 12, Programme 01), expressed in 2015 euros. The expenditure variables is instrumented by the change in the proportion of women induced by Law 215/2012. Exp represents the number of years a child has been exposed to the reform. For example, EXP=1 indicates that at the time of the INVALSI tests in 2014, the child had been exposed to the reform for only one year, and so on. The sample expands over the period 2013–2018. Standard errors are robust to heteroscedasticity. The Durbin-Wu-Hausman test offers a test statistic for the null hypothesis that the OLS estimator is reliable and entirely efficient (and thus there are no biases from endogeneity). If the null hypothesis is rejected, the IV estimator should be favored as it is reliable. The OLS estimator is entirely efficient under the null hypothesis, but unreliable if the null hypothesis is false; the IV estimator is reliable both under the null hypothesis and the alternative hypothesis.

9. References

Istat, *Annuario statistico italiano, 2021*. https://www.istat.it/storage/ASI/2022/Sintesi.pdf

Alesina, A. F., Lotti, F., & Mistrulli, P. E. (2013). Do women pay more for credit? Evidence from Italy. *Journal of the European Economic Association*, *11*(suppl_1), 45-66.

Andreoli, F., Manzoni, E., & Margotti, M. (2022). *Women at work: Gender quotas, municipal elections and local spending*. European Journal of Political Economy, 75, 1021775.

Bagues, M., Campa, P., 2021. Can gender quotas in candidate lists empower women? Evidence from a regression discontinuity design. J. Publ. Econ. 194, 104315.

Baltrunaite, A., Casarico, A., Profeta, P., & Savio, G. (2019). *Let the voters choose women*. Journal of Public Economics 180, 104085.

Baskaran, T., & Hessami, Z. (2018). Does the election of a female leader clear the way for more women in politics?. *American Economic Journal: Economic Policy*, *10*(3), 95-121.

Brilli, Y., Del Boca, D., & Pronzato, C. D. (2013). *Does child care availability play a role in maternal employment and children's development? Evidence from Italy.* Rev Econ Household, 14:27–51, DOI 10.1007/s11150-013-9227-4.

Brunner, E., Hyman, J., & Ju, A. (2018). *School Finance Reforms, Teachers' Unions, and the Allocation of School Resources.* EdWorkingPaper , No.19-35.

Burchinal, M., Howes, C., Pianta, R., Bryant, D., Early, D., Clifford, R., & Barbarin, O. (2008). *Predicting child outcomes at the end of kindergarten from the quality of pre-kindergarten teacher–child interactions and instruction. Applied Development Science*, *12*(3), 140-153.

Cabaleiro, R., Buch, E., 2018. Adjustments in municipal fiscal crises. Are they different according to the gender of the mayor? Local Govern. Stud. 44 (2), 255–274.

Cabaleiro, R., Buch, E., 2020. Women in Spanish municipalities and budgetary policies. Urban Aff. Rev. 56 (6), 1715–1745.

Card, D., & Abigail Payne, A. (2002). *School finance reform, the distribution of school spending, and the distribution of student test scores.* Journal of Public Economics, 83,49–82.

Cavallini, F., Dominici, A., & Masi, O. (2023). *Executive Gender Quotas and Social Services: Evidence from Italy.* SSRN: https://ssrn.com/abstract=4395411 or http://dx.doi.org/10.2139/ssrn.4395411 .

Chiades, P., & Mengotto, V. (2013). *Il calo degli investimenti nei Comuni tra Patto di stabilità interno e carenza di risorse*. Banca D'Italia, ISSN 1972-6627.

Clots-Figueras, I., 2011. Women in politics: evidence from the Indian states. J. Publ. Econ. 95 (7–8), 664-690.

Clots-Figueras, I., 2012. Are female leaders good for education? Evidence from India. Am. Econ. J. Appl. Econ. 4 (1), 212–244.

Corazzini, L., Elena Meschi, E., & Pavese, C. (2021). *Impact of early childcare on immigrant children's educational performance.* Economics of Education Review, 85, 102181.

Corte dei Conti (2012). "*Rapporto 2012 sul coordinamento della finanza pubblica*", Technical Report.

Cragg, J.G., Donald, S., 1993. Testing Identifiability and Specification in Instrumental Variable Models. Econometric Theory, 9(2): 222-240.

Cui, Y., Liu, H., & Zhao, L. (2019). Mother's education and child development: Evidence from the compulsory school reform in China. *Journal of Comparative Economics*, *47*(3), 669-692.

De Paola, M., Scoppa, V., & De Benedetto, M. A. (2014). The impact of gender quotas on electoral participation: Evidence from Italian municipalities. *European Journal of Political Economy*, *35*, 141-157.

Del Boca, D., Pronzato, C., & Sorrenti, G. (2016). When rationing plays a role: selection criteria in the Italian early childcare system. *CESifo Economic Studies*, *62*(4), 752-775.

Dipartimento Per Le Politiche Della Famiglia (2020). Nidi E Servizi Educativi Per L'Infanzia. Technical Report. 11,36.

Funk, P., & Gathmann, C. (2008). Gender gaps in policy making: Evidence from direct democracy in Switzerland. *Available at SSRN 1374228*.

Gago, A., Carozzi, F., 2021. Do Female Leaders Promote Gender-Sensitive Policies? Available at: SSRN: https://ssrn.com/abstract=3716566.

Gilliam, W. S., & Zigler, E. (2004). *State efforts to evaluate the effects of prekindergarten, 1977 to 2003*.

Grembi, V., Nannicini, T., & Troiano, U. (2016). Do fiscal rules matter?. *American Economic Journal: Applied Economics*, 1-30.

Hanushek, E. A. 2008. ''Education Production Functions.'' The New Palgrave Dictionary of Economics 1−8: 1645−8.

Hern´andez-Nicol´as, C.M., Martín-Ugedo, J.F., Mínguez-Vera, A., 2018. Women mayors and management of Spanish councils: an empirical analysis. Fem. Econ. 24 (1), 168–191.

Hessami and T. Baskaran. Competitively elected women as policy makers. 2019

Jackson, C. K., Johnson, R. C., & Persico, C. (2015). *The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms.* NBER Working Paper No. 20847.

Jackson, C. K. (2018). Does School Spending Matter? The New Literature on an Old Question. NBER Working Paper No. 25368.

Jackson, C. K. and Mackevicius, C. (2021). The Distribution of School Spending Impacts. Working Paper, Northwestern University.

OECD. (2010). OECD family database. Available at http://www.oecd.org/els/social/family/database.

Ordine, P., Rose, G., & Giacobbe, P. (2022). *The Effect of female representation on political budget cycle and public expenditure:Evidence from Italian municipalities*. Economics & Politics, 35, 97-145.

Pavese, C., & Rubolino, E. (2023). *Austerity Harmed Student Achievement*. SSRN: https://ssrn.com/abstract=4099283 or http://dx.doi.org/10.2139/ssrn.4099283 .

Priyanka, S. (2022). *Do Female Politicians Lead to Better Learning Outcomes?* The B.E. Journal of Economic Analysis & Policy, 22(4): 763–800.

Rehavi, M. M. (2007). *Sex and politics: Do female legislators affect state spending?*. mimeo

Rocha, O., Kamphambale, D., MacMahon, C., Coetzer, J. H., & Morales, L. (2023). The Power of Education in a Globalised World: Challenging Geoeconomic Inequalities. *Peace Review*, *35*(4), 708-723.

Schaffer, M. E., Baum, C., Finlay, K., Kleibergen, F., Magnusson, L., & Stillman, S. (2013). Stata Software for Econometric Estimation and Testing; avar, weakiv, actest, ivreg2h, ranktest, ivreg2.

Sibiano, P., & Agasisti, T. (2013). Efficiency and heterogeneity of public spending in education among Italian regions. *Journal of Public Affairs*, *13*(1), 12-22.

Stock, J., Yogo, M., 2005. Testing for Weak Instruments in Linear IV Regression. In: Andrews D.W.K. Identification and Inference for Econometric Models. New York: Cambridge University Press: 80-108.

Svaleryd, H., 2009. Women's representation and public spending. Eur. J. Polit. Econ. 25 (2), 186–198.

Viesti, Gianfranco. 2016. Disparità regionali e politiche territoriali in Italia nel nuovo secolo. In Le Regioni Europee. Politiche per la Coesione e Strategie per la Competitività. Edited by Fabio Mazzola and Rosanna Nisticò. Milano: Franco Angeli.

Wong, H. L., Luo, R., Zhang, L., & Rozelle, S. (2013). The impact of vouchers on preschool attendance and elementary school readiness: A randomized controlled trial in rural China. *Economics of Education Review*, *35*, 53-65.

Wooldrige, J. M. (2000). Econometric analysis of cross section and panel data. The MIT Press.

World Economic Forum. 2017. The Global Competitiveness Report 2017–2018. Geneva: The World Economic Forum.