IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16378

The Short and Medium Term Effects of Full-Day Schooling on Learning and Maternal Labor Supply

Giulia Bovini
Niccolò Cattadori
Marta De Philippis
Paolo Sestito

AUGUST 2023

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16378

The Short and Medium Term Effects of Full-Day Schooling on Learning and Maternal Labor Supply

Giulia Bovini
Banca d'Italia

Niccolò Cattadori
University of Zurich

Marta De Philippis
Banca d'Italia and IZA

Paolo Sestito
Banca d'Italia and IZA

AUGUST 2023

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Short and Medium Term Effects of Full-Day Schooling on Learning and Maternal Labor Supply*

This paper considers the case of Italy to analyze the short- and medium-term effect of a longer school day in primary school on both students' learning and mothers' labor supply. we rely on unique application-to-primary-school data: first, we control for parental preferences, proxied by individual applications; second, we exploit variation in the probability of attending the full-time (FT) scheme that only stems from nonlinearities in the mix of FT and part-time (PT) applications received by the school and from class size limits set by the law. We show that attending the FT scheme increases Math test scores in grades 2 and 5 and Italian scores in grade 2 by around 4.5% of a standard deviation, but the effects fade away by grade 8. Conversely, there is a positive impact on maternal labor force participation and employment, which is long-lasting (approximately 2 p.p.). No effect is found on fathers' employment. Finally, we find some evidence of negative selection on gains, as the groups of students and mothers for whom the effect seems to be larger are not those more likely to apply to the FT scheme or to attend it conditional on applying.

JEL Classification: H40, I21, I24, J13, J21

Keywords: time at school, female labor supply, selection into treatment, students' learning

Corresponding author:
Marta De Philippis
Banca d'Italia
Via Nazionale, 91
00184 Roma
Italy

E-mail: marta.dephilippis@bancaditalia.it

* For very helpful comments, we thank Federico Cingano, Elisa Facchetti, Marco Ovidi, Eliana Viviano, Ulf Z¨olitz, and seminar participants at the Bank of Italy, CESifo Institute, IWAEE Catanzaro, Marco Fanno ALUMNIWorkshop, Sapienza University, University of Milan, and University of Z¨urich. Cattadori gratefully acknowledges support by the University of Zürich's Research Priority Program "Equality of Opportunity".

1 Introduction

Instruction time, and more in general time spent at school, is considered a key determinant of students' academic achievement and broader development. To the extent that additional time at school crowds out homework, longer school days are seen by many as an opportunity to improve equity in education, since they foster the learning of children from disadvantaged backgrounds who have few resources at home. Over the past two decades, a substantial number of countries have extended the school day (OECD, 2015).

However, little is known about the medium-term impact on achievements of more time at school in the early educational stages. This longer-term effect may differ from the short-term one, as students in short-day schemes could be more likely to develop a set of skills – like the ability to study autonomously – that become increasingly useful further along in the education path. Furthermore, a comprehensive assessment of the effects of the length of the school day should consider not only children but also their families: by making it easier for parents to combine childcare responsibilities and professional lives, a longer school day could boost parental labor supply, especially that of mothers.

To our knowledge, this paper is the first to provide an overall assessment of the effect of more time at school on students' achievement and their mothers' labor supply, both in the short- and medium-term. We consider the case of Italy, where two instructional schemes coexist in primary schools, namely the *Tempo Normale* scheme (henceforth PT, which stands for Part-Time), where pupils spend at school typically 27 hours per week mainly in the morning, and the *Tempo Pieno* scheme (henceforth FT, which stands for Full-Time), where pupils stay in school also in the afternoon for a total of 40 hours per week. While both schemes cover the same curriculum, during the longer school day pupils have lunch and revise the curriculum at school under the supervision of their teachers.

To address our questions, we rely on a unique match between two individual-level administrative datasets for the cohort who started primary school in 2014-15: (i) applications to primary school data (from the Ministry of Education and Merit), recording the school and scheme (FT or PT) parents applied to, and (ii) standardized achievement in Italian and mathematics tests, as well as information on the demographic and background characteristics of students, which include, among others, parental employment, from the start of primary school through its fifth and last grade (grade 5) and up until the third and last grade of lower secondary school (grade 8), coming from the National Institute for the Evaluation of the Instruction and Training System (INVALSI).

The estimation of the causal effect of longer school days is challenging because attending an FT class is an equilibrium outcome of demand and supply. On the demand side, parents self-select into the preferred school and scheme, based on unobservable characteristics possibly correlated with the outcomes of interest. On the supply side, activating one FT class requires adequate infrastructure - for instance, a school canteen where to serve lunch - and economic resources for paying teachers and the school staff. Furthermore, because the law sets upper and lower limits to class size and a class can be either fully FT or fully PT, certain mixes of FT and PT applications can generate excess demand for either scheme. When a supply-side constraint binds, school principals manage excesses of demand according to potentially endogenous school-specific criteria that are not observed.

We leverage application data to deal with endogeneity. First, we tackle demand-side selection by explicitly controlling for parental preferences in the regression (i.e., we include a dummy for whether parents apply to the FT scheme). Conditional on parental preferences, there still is some potentially endogenous variation in the treatment (being in an FT class) because of supply-side constraints (for instance, when there is excess demand and the principal decides who gets the preferred scheme). We therefore use an instrumental variable strategy, which - conditional on parental preferences - has two desirable properties: (i) it only varies according to non-linearities in the mixes of PT and FT applications - a variable difficult to predict for parents; and (ii) it does not depend on students' characteristics, hence not capturing endogenous school principals' choices. In particular, we focus on one among the supply-side constraints - that stemming from the law on class size limits - and develop an algorithm that for all mixes of FT and PT applications figures out the existence of excess demand.[1] Then, we use as an IV the conditional-on-own-application probability of attending an FT class based on the application mix received by the school, computed assuming that, if the application mix generates excess demand, the school principal: (i) minimizes the number of pupils unhappy with their scheme;[2] (ii) once figured out how many students can not be assigned to the preferred scheme, chooses *who* to make unhappy randomly (i.e., not depending on the students' characteristics).[3] The IV is relevant, validating the ability of the algorithm to find the cases when class size limits generate supply-side constraints.

[1]Consider a school that receives 20 applications, of which 11 for the PT scheme and the remaining 9 for the FT scheme. To abide by class size laws, the school principal can activate only one class because the lower limit on class size is 15. Because the law prescribes that the class shall be either fully PT or fully FT, in this example there is excess demand: the principal would either create a PT class - making all FT unhappy - or an FT class - making all PT pupils unhappy.

[2]Continuing with the previous example, the school principal would choose to make unhappy the 9 students who applied to the FT scheme, since they are the minority. Hence, she/he would create a PT class.

[3]Since in the example all FT students do not get their preferred scheme, their conditional-on-applying probability of attending an FT class, given the application mix received by the school and the assumptions about the way the school principal handles excess demand, is 0%. Section 4 provides more examples.

We start our analysis by describing the time use patterns of students who attend shorter and longer school days. We find that the total amount of time dedicated to instructional (i.e., class attendance and homework) and to leisure activities is roughly similar. Longer school days entail a close to 1:1 substitution between homework after school (PT students) and supervised revision of the curriculum in class (FT students). Moreover, we analyze whether teachers are different across FT and PT classes. The INVALSI teachers' survey suggests that instructors in FT classes are slightly less experienced, but more educated than those in PT classes. However, their teaching practices, evaluation methodologies, attention given to the INVALSI test, and use of technological equipment are quite similar.

According to our 2SLS estimates, being in an FT class in primary school causes an improvement in Math test scores in grades 2 and 5 (by 4.8% and 4.6% of a standard deviation, respectively) and in Italian scores in grade 2 (by 4.3% of a standard deviation). The positive effect vanishes however over time: in grade 8, three years after the end of the FT program, we no longer find a difference in performance. The dynamic pattern of estimated effects suggests that the benefits of longer school days while in primary school might be later counterbalanced by a lower preparedness of FT pupils for autonomous study, which becomes an increasingly important skill during lower secondary school.

The effect on maternal labor force participation and employment is positive and, instead, long-lasting. Mothers of children who attend the FT scheme in primary school increase their labor force participation by approximately 2 percentage points both in the short- and medium-term. The effect on mothers' employment is smaller initially and increases with students' age, probably because it takes time for mothers to find jobs: overall, three years after the end of the program, mothers of pupils in FT classes in primary school have a 2.2 percentage points higher employment rate. The long-lasting effects are consistent with the presence of hysteresis in the labor market, which implies that temporary exits may lead to human capital depreciation and have a persistent effect on labor market outcomes. We find instead a precisely estimated zero effect of the FT scheme on fathers' participation and employment. This is in line with the finding that mothers still bear most of the child-rearing responsibilities, and their labor supply is more elastic to changes in the availability of childcare. Our results - especially those on mothers' labor supply - are robust to a series of robustness checks that deal with possible remaining concerns for identification.

These average effects mask some heterogeneity, although differences across groups do not usually reach conventional levels of statistical significance. First, the short-term benefits of FT schemes seem to be larger for pupils of low socio-economic status. This implies that FT schemes have a positive impact on learning inequalities, as students are exposed to more homogeneous contexts for a larger amount of time. Second, low-educated mothers, who are more likely to be at the margin of choosing to exit the labor force to bear childcare responsibilities, experience the largest increase in labor force participation when their children are enrolled in an FT scheme. Using our application data we show however that families more likely both to apply and to be assigned to the longer school day are those with more educated parents and a more advantageous background, i.e., not the ones who benefit the most from the FT scheme; this has important implications for the interpretation of our results and for the decision of whether and where to expand the FT scheme first.

Our paper is related to several strands of the economic literature. First, it speaks to the papers analyzing the effect of more time at school on students' development. The evidence produced by this literature focuses on the short-term effects and finds mixed results that substantially depend on the curriculum content of the extra school time: while most studies tend to find positive effects on students' achievements of policies that increase instruction time (see, for instance, Bellei (2009), Figlio et al. (2018), Lavy (2020), Barrios-Fernández and Bovini (2021)), the picture is less clear for the few papers that consider reforms that increase time at school without impacting significantly instruction time (see Felfe and Zierow (2014), Schmitz (2022) and Seidlitz and Zierow (2020), who analyze the introduction of full-time and after-school care programs in the German setting). One exception, which however looks at long-term effect only, is Dominguez and Ruffini (2023), who evaluate the impact of longer school days (more instruction time) in Chilean primary and secondary schools, finding improvements in the long-term economic well-being of treated students.[4]

Our paper also relates to the broad literature that analyzes the consequences of changes in childcare provision for mothers' labor supply. Most studies focus on the impact of childcare for preschool-aged children and mainly focus on the short-term effects. These papers tend to find positive effects on maternal employment, even if there is very little consensus on the magnitude, which moreover is sometimes found to be entirely concentrated on small subgroups like single mothers or those living in disadvantaged areas (Lefebvre and Merrigan (2008), Baker et al. (2008), Cascio (2009), Fitzpatrick (2010), Goux and Maurin (2010), Havnes and Mogstad (2011), Nollenberger and Rodríguez-Planas (2015), Carta and Rizzica

[4]In a contribution written in Italian for a volume edited by INVALSI (Bovini et al. (2016)), we analyzed the effects of the FT scheme in primary school on pupils' achievement in grades 2 and 5. As application data were not available then, the identification strategy relied on within-school, year-to-year variation in the fraction of FT classes for academic years from 2011-12 to 2014-15. The results pointed to a positive effect of the FT scheme on Math scores, especially at the bottom of the distribution, and a virtually null effect on Italian test scores. This paper relies on novel data and a much-improved identification strategy, provides more evidence on differences between FT and PT teaching, looks also at medium-term effects on achievement, and expands the analysis to also study the effect of the FT scheme on maternal status in the labor market.

(2018)). Evidence on the impact of providing care for older, school-aged children on maternal employment is scarcer; also in this case the results, which only focus on the short-term, are mixed, likely because of differences in the institutional setting and cultural contexts. Some papers find positive effects on employment (see Gambaro et al. (2019) for Germany and Berthelon et al. (2023), Contreras and Sepúlveda (2017), Martínez A. and Perticará (2017), and Padilla-Romo and Cabrera-Hernández (2019) looking at Latin America). On the contrary, other papers find positive effects on the intensive margin only (for instance, Duchini and Van Effenterre (2022) and Felfe et al. (2016)) or no effects on labour supply (Dehos and Paul, 2021).

Finally, our paper speaks to the relatively scarce literature analyzing the role of different types of selection into the treatment for the estimation of local average treatment effects. Most papers study the case of returns to college education and tend to find positive self-selection based on gains (for instance Carneiro et al. (2011) and Nybom (2017)). Cornelissen et al. (2018) instead estimate children's returns to early child care in Germany, that is a case when someone other than the treated individual (i.e., the parents or the school principal) decides on enrollment. They find a reverse selection on gains based on both observed and unobserved characteristics, but are not able to identify whether this negative selection comes from demand or supply side constraints.

To the best of our knowledge, we are the first to evaluate the effects of a full-time program in primary school on both children and their mothers at the same time. This is key to assessing the overall benefits of these schemes. Moreover, we contribute to the literature by looking at both short- and medium-term effects, which is important to evaluate the cumulative impact of the policy and the presence of possible dynamic complementarities. Finally, the unique match between application-to-school data and enrollment data allows us to assess the type of selection into the treatment and to study how and whether it is shaped by parental preferences - on the demand side - and by school principal decisions - on the supply side. This is key for drawing policy implications.

The remainder of the paper is structured as follows. In Section 2 we describe the data and in Section 3 we provide background information and descriptive evidence on the two instruction schemes (FT and PT). Section 4 and 5 outline the identification strategy and the empirical analysis, respectively. Section 6 presents the results. Section 7 concludes.

2 Data

To study the effect of the FT scheme on students and their mothers we combine for the first time two datasets.

- *Achievement data:* from the scholastic year 2009-10 INVALSI administers standardized Italian and mathematics tests to all students in grades 2 and 5 (the second and the last grades of primary school), grade 8 (the third and last grade of lower secondary school), and grade 10 (the second grade of upper secondary school). Tests take place at the end of the school year (March-May). While they are low-stakes but in grade 8 and there is some evidence of manipulation (especially in earlier waves, in primary schools and in Southern regions), test scores *corrected for cheating*[5] have been extensively used to measure achievement of Italian students (e.g. Corazzini et al. (2021), Checchi and De Paola (2018)). Thanks to time-invariant and unique (anonymized) identifiers, waves are longitudinally linkable at the student level. Besides test scores, individual-level records contain rich information on the pupils, their families, classes, and schools, which are either provided by the school or by the students themselves in an accompanying questionnaire. Background characteristics include the occupation status of each parent, from which we derive the other outcomes of our analysis: labor force participation and employment. In the 2011-12 and 2012-13 questionnaires, fifth graders were also surveyed on their use of time after school. Given that the FT scheme prolongs the school day, an assessment of its impact on learning requires knowing to what extent the extra time at school crowds out study or leisure time out of school. INVALSI also administers a questionnaire to Italian and mathematics teachers of a representative sample of classes. The questionnaire collects rich information on their background, career, and teaching practices, which we use to study if and how the teaching input differs across the FT and the PT schemes.

-*Application data:* we have access to a unique dataset maintained by the Ministry of Education and Merit (MIM) that records individual-level applications to primary schools. For pupils who in the first months of 2014 applied to start primary schools in September (i.e.

[5]Test score manipulation could occur in earlier waves because tests were paper-based and proctors had to copy students' responses on machine-readable answer sheets to be sent to INVALSI. While teachers were not supposed to monitor their own classes, only in a few random classes proctors were external examiners as opposed to other teachers of the school (Bertoni et al. (2013)). According to Angrist et al. (2017), teacher shirking in score transcription was the main driver of manipulation. Due to this, INVALSI released both *raw* scores and scores *corrected for cheating* based on a class-level cheating adjustment procedure first developed by Quintano et al. (2009) and aimed at detecting anomalies in answers' pattern at the class level. In the latest years, the scope for cheating has been drastically reduced in grade 8: tests are computer-based (thus bypassing the transcription phase) and each student gets different questions (of equivalent difficulty). We use scores *corrected for cheating* for grades 2 and 5, and *raw* scores in grade 8 (coherently with the claim that computer-bases tests prevent cheating, INVALSI releases only those scores for grade 8 tests in 2021-22).

the scholastic year 2014-15), the following information is available: the school where the parents applied (first choice) and the scheme (PT or FT) to which they applied, as well as the school and the scheme that their child eventually attended in grade 1. Our treatment is a dummy that takes the value 1 if the student attends an FT class in grade 1. We provide a description of the application process to primary school in Appendix A2.

Our identification strategy requires combining the two datasets, using the unique student identifier as the matching variable. It follows that our analysis focuses on students who started primary school in the scholastic year 2014-15 and later took the INVALSI tests in 2015-16 (grade 2), 2018-19 (grade 5), and 2021-2022 (grade 8). Figure 1 illustrates on a timeline the longitudinal nature of our data.

The matching between the two datasets is not perfect. In part, this reflects some operative problems encountered when anonymizing student identifiers using the same algorithm in both datasets. In part, this stems from naturally occurring circumstances. On the one hand, not all students found in the application data are also in the test data 1, 4, and 7 years later: absences on the day of the test, grade retention, and exits from the Italian education system can account for this. On the other hand, not all students found in the test data are also in the application data: there can be late applicants and children who start studying in Italy past grade 1. In our analysis, the estimation sample consists of children who we observe at all points of our timeline and who: (i) applied to public schools[6], and (ii) have non-missing Italian and mathematics scores in all INVALSI tests.

This leaves us with a balanced sample of 338,862 students per year, roughly 67% of test-takers in grade 2. Table A1.1 shows that, based on test scores and background characteristics, the estimation sample is slightly positively selected when compared with the population of grade 2 test takers. The differences, however, are quite small: as an example, the share of mothers with a tertiary degree is 20% in the estimation sample and 19% in the full sample.[7]

[6]This restriction is due to the fact that private schools need not abide by the class size limits set by the law, which are an ingredient of our identification strategy. We also drop students who applied to schools that received a total number of applications lower than that in principle required by the law to form one class: 10 for schools in isolated villages in mountainous areas, small islands and municipalities with linguistic minorities, 15 for all other schools (see Section 4 for more details).

[7]Table A1.1 shows that the information on education is missing for around 20% of parents both in the full and in the estimation sample. The longitudinal match, however, helps to fill some missing values in the estimation sample, under the reasonable assumption that education levels of adults are time-invariant: as an example, if the information on the mother's education is reported in grade 5 but not in grade 2, we can fill the missing information in grade 2 by using that from grade 5. We also replace the value reported in grade 8 with that coming from earlier grades, because inspections of the data reveal that there are some coding errors in the variable. Thanks to this procedure, we can almost halve the share of missing values for mothers' and fathers' education. See Table A1.2, where we report summary statistics for FT and PT students in the estimation sample after applying this procedure.

Importantly, the share of students attending the FT scheme is virtually the same in the population and in the sample: 35% and 36%, respectively. Overall, the estimation sample is representative of the underlying full population.

Finally, the sample on which we estimate the relationship between the FT scheme and maternal status in the labor market is slightly different: we focus on children who we observe at all points of the timeline and for whom the occupation status of the mother is recorded in all three grades (2,5,8). This leaves us with a balanced sample of 237,456 observations per year. The sample size is smaller because for some children this information is missing in at least one grade.[8] Appendix A3 describes in greater detail the matching and the sample construction procedures.

3 Institutional Setting

Italy provides an attractive setting for our analysis because primary schools can offer two different instructional schemes: the *Tempo Normale* (PT, for part-time) and the *Tempo Pieno* (FT, for full-time). Under PT schemes lectures cover no less than 24 hours per week (usually 27-30), distributed across five to six days per week, mostly only in the morning. Under FT schemes pupils spend 40 hours per week at school, typically split across five days per week, from 8.30 a.m to 4.30 p.m. A school can offer both schemes, but every class is either fully PT or fully FT. The time schedule of the class remains constant until the end of the cycle and so does the group of students who belong to the class, across all subjects.

While both schemes cover the same curriculum, students in FT classes have lunch at school.[9] They should also mostly revise the curriculum at school, under the supervision of their teachers, rather than by doing homework after school. We assess differences in time use between FT and PT students relying on two sources of data. First, we exploit the questionnaire administered to fifth graders in 2011-12 and 2012-13, which contained questions on the frequency of study and leisure activities after school (Table 1). While very few children in both schemes claimed to never have homework after school or during weekends, the share who

[8] In grade 2 the share of missing values for maternal employment is 20%. Notice that, differently from maternal education, we cannot take advantage of the longitudinal dimension of the sample to fill in some missing values, as occupation status is a time-varying variable.

[9] According to some literature on the US (for instance, Frisvold, 2015), the provision of free breakfast for low-income families at school improves students' cognitive achievements. However, we believe this is not the main mechanism behind possible positive effects of attending the FT scheme on achievements in Italy, first because the FT scheme implies free school meals both for high and for low-income students, and – if anything – it is mainly attended by higher-income students. Second, the level of malnutrition is lower in Italy than in the US on average.

reported doing homework very frequently (more than 5 times a week) is 39 p.p. larger among PT students. There are instead no large differences in the frequency of leisure activities (e.g. watching TV, playing with friends). Second, we look at data from the Use of Time Use Survey administered to a random sample of households in 2013. Figure 2 panel (a) confirms that from Monday to Friday students enrolled in the FT scheme stay longer at school and do less homework than those enrolled in the PT scheme. On Saturdays, the former do not attend school, while (some of) the latter do, and the amount of homework is very similar. Overall, FT pupils devote to instructional activities (i.e., time at school and homework) roughly two more hours per week than PT students, which amounts to a relatively modest increase of fewer than 20 minutes per day, less than 5% of the total instructional time. Figure 2 panel (b) investigates whether pupils under FT or PT schemes differ in terms of help received from their parents while doing homework. On average parents assist children with homework almost 60% of the time regardless of the instructional scheme. However, since PT students spend more time doing homework, in absolute terms the differences are large. Overall, this evidence suggests the extra time at school under the FT scheme mostly crowds out time that would have been spent doing homework, whit no significant impact on leisure time.

Another difference between the two schemes could be the teaching input. Based on INVALSI teachers' questionnaires, Tables 2 and 3 explore differences between FT and PT teachers in grades 2 and 5, respectively. Coefficients come from regressing the outcome of interest on the FT dummy, net of region fixed-effects.[10] Teachers in FT classes tend to have lower tenure and experience, suggesting that they are on average younger (even if, due to the coarse age categories reported by INVALSI, we cannot detect differences in the probability of being under-50). Consistently, it is more likely that teachers in FT classes hold a university degree, which was not a necessary requirement to become a teacher for older cohorts. Despite these demographic differences, disparities in teaching practices seem very modest. Notably, teachers in FT and PT classes do not differ much in terms of how they prepare their students for the INVALSI tests: the probabilities to practice using class exercises, tests from previous years, and test-specific textbooks are similar in the two groups. This is important for the interpretation of the results: students attending the FT scheme do not perform differently from their counterparts in the PT scheme simply because of different teaching-to-the-test practices. In grade 2 only, FT teachers are less likely to assign homework exercises similar to INVALSI tests, consistent with the fact that FT pupils have less homework in general. Teaching and evaluation practices are also overall similar across the two schemes: while

[10] We include region fixed effects because, as shown in table A1.2, the share of pupils enrolled in FT classes displays a high geographic variability.

there are virtually no statistically significant differences in grade 2, fifth-grade Italian FT teachers adopt more frequently some non-traditional practices (such as flipped classroom or peer-learning) and rely more on group work evaluation.

4 Identification Strategy

Estimating the causal effect of time spent at school on students' achievement and mothers' labour supply is challenging. Whether a student attends an FT class is an equilibrium outcome of demand and supply. On the demand side, parents express a preference for a school and for either FT or PT schooling. Preferences could depend on characteristics not fully observable and correlated with the outcomes of interest.

On the supply side, there are different types of constraints. First, schools can offer FT classes only if they have adequate infrastructures - notably, a school canteen where to serve lunches - as well as economic resources and enough teachers and school staff to offer longer hours. This constraint could potentially affect recorded preferences, as parents could gather information about this before making their choice in the application form. Second, the law that regulates the class formation process in Italian public schools (DPR 81/2009) constrains the number of total classes and that of FT and PT classes. Specifically, the law prescribes that the number of classes in a school is a function of the total number of applications received and that each class should have a minimum of 15 and a maximum of 27 students.[11] Moreover, a class can only offer one scheme (i.e., it is either fully FT or fully PT). Due to these constraints, the school principals may need to manage an excess of demand for either scheme, depending on the application mix. Criteria for managing excesses of demand are school-specific and not observed. In many cases, they depend on students' and parents' characteristics (i.e., whether a sibling already attends the same scheme, parents' occupation and socioeconomic status, and proximity of the school to home), therefore generating potential endogeneity.

Data indeed suggests that there is excess demand for FT classes. Table 4 shows that in our estimation sample, out of 100 applicants to the FT scheme only 84 end up enrolling into an FT class in the preferred school; 3 manage to attend an FT class, but in a different school. The rest (around 13%) ends up in a PT class. Unmet demand is less common for the PT scheme, concerning 3 applicants out of 100. This is consistent with the fact that the PT scheme is less costly for public finances and is indeed the prevalent one, covering around two-thirds of pupils.

[11] The minimum is lowered to 10 for schools located in small, isolated villages and the maximum to 20 if there are disabled children in the class.

As a result of this multi-step selection process, a simple comparison of outcomes between FT and PT students is unlikely to capture the causal effect of interest. Consider mothers' labor force participation: a positive correlation between longer school days and this outcome could reflect the causal effect of the former on the latter or simply the fact that (i) employed mothers may be more likely to prefer a scheme that helps balance work and care duties, and (ii) school principals may give priority to children of employed mothers if there is excess demand for FT classes. Table A1.2 shows that pupils who attend the FT and the PT schemes have different observable characteristics. Parents of FT students are on average more educated; the incidence of immigrant children is twice as large among FT students. Looking at outcomes, FT pupils obtain on average better scores in italian and mathematics, and their mothers are much more likely to participate in the labour market and be employed. These disparities partly stem from the fact that the FT scheme is more widespread in the North-Centre of the country, which is characterized by higher levels of income, employment, and immigration than the Southern regions.[12] For most of the characteristics, however, these differences persist, albeit smaller in magnitude, even after accounting for the province. One notable exception occurs in learning outcomes, particularly in Italian, where full-time students within provinces tend to have lower scores compared to their part-time counterparts.

We deal with selection issues exploiting the application data. First, thanks to the unique information on individual preferences we are able to explicitly control for demand-side selection (parental preferences), by including in the regression a dummy for whether parents *apply* to the FT scheme. There still remains some residual variation in the treatment (*attend* an FT class) because - as shown in Table 4 - there are students who are *not* assigned to the preferred scheme due to supply-side binding constraints; this variation may potentially be endogenous as a result of non-random school principals' criteria to manage excess demand. By building an instrumental variable (IV) that again relies on school-level application data, we isolate a plausibly exogenous source of this residual variation that: (i) only depends on non-linearities in the mixes of FT and PT applications received by the school, a variable difficult to predict for parents; and (ii) is not affected by students' characteristics, thus not capturing the endogenous school principal's choice. To find such a variation, we focus on the supply-side constraint that stems from class size limits set by the law.

To do so, we develop an algorithm that, using a three steps procedure, figures out excess demand for any given application mix and, in case of excess demand, randomly allocates students to the preferred or non-preferred scheme. Based on the application mix and the

[12] Out of 100 FT students, 61 lives in the North; the share of children from Northern regions drops to 39% among PT pupils.

previously mentioned constraints on the class formation process set by the law, the algorithm first determines the total number of classes in a school under the assumption that the principal wants to minimize the overall number of classes. Second, it predicts the number or FT and PT classes, by assuming that the principal minimizes the number of unhappy students (i.e., students not assigned to their preferred scheme). Finally, once the number of unhappy pupils is found, the algorithm assigns to every student who expressed a given preference the same probability of being the one who actually does not get into the preferred scheme: this is what would happen if in case of excess demand the school principal decides who to make unhappy randomly, rather than endogenously to some characteristics. Let us consider an example of a school that receives 33 applications (8 FT, 25 PT). The total number of classes as predicted by the algorithm would be $2 = \lceil \frac{33}{27} \rceil$, since classes cannot be larger than 27. As the majority of applications are PT, the first class would be PT. The second class in this case would be FT. The reason is that, due to class size limits, the second class must have at least 15 students: the happiness-maximizing way to achieve this is to assign 18 of the PT applicants to the first class (the PT class), while the remaining 7 PT applicants and all 8 FT applicants go in the second class (which by majority rule becomes FT). This assignment disgruntles 7 pupils (28% of PT applicants); the alternative one (2 PT classes) would make 8 children unhappy. Third, given the presence of the excess demand for the PT scheme, the algorithm assumes that each student who applied to PT has the same probability of being assigned to the FT class ($\frac{7}{25} = 28\%$). Appendix A4 describes in detail the construction of the instrument and discusses several additional examples.

In our analysis, we instrument attending a FT class with the following variable, measured at the school level:

$$Z_{is} = \begin{cases} \left(\frac{\textit{\# Applied to FT \& Assigned to FT according to the algorithm}}{\textit{\# Applied to FT}} \right)_s, & \text{if } i \text{ applied to FT,} \\ \\ \left(\frac{\textit{\# Applied to PT \& Assigned to FT according to the algorithm}}{\textit{\# Applied to PT}} \right)_s, & \text{if } i \text{ applied to PT.} \end{cases} \tag{1}$$

Z_{is} is the predicted-by-the algorithm probability that pupil i who expressed a preference for the scheme $\Sigma \in (FT, PT)$ in school s ends up in an FT class. Notice that using Z_{is} as an IV ensures that the chance that a student ends up in the preferred scheme, conditional on their preferences, varies only at the school level and only stems from the school-level application mix (it does not reflect the selection made by the school principal).

Our algorithm performs well in identifying circumstances where supply-side constraints due to class size limits can bind. In Figure 3 we show the distribution of the instrument in the four subsamples defined by whether the students applied to and attended the FT scheme. The average value of Z_{is} is 98% for students who apply to the FT scheme and manage to get it, while it drops to 64% for pupils who apply to FT but later attend a PT class. The average value of Z_{is} is 1.5% for students who apply to the PT scheme and enroll in it, while it increases to 40% for pupils who apply to PT but attend an FT class. This translates into a relevant first stage in the 2SLS regression (see Section 6).

The twofold use of application data goes a long way toward mitigating identification issues. A remaining concern is that parents with the same preference could still be sorting across different schools based on the application mix, and hence on the probability of being accommodated: the ingredient of the instrument could therefore be correlated with some unobservable characteristics. In principle, it is difficult for parents to know the exact number and composition of applications. Nevertheless, to address this concern we proceed in two ways. In all regressions, we control for (i) the ingredients of the instrument (in a linear way), this implies that our instrument only relies on non-linearities in the application mix, which are even more difficult to predict,[13] and (ii) for a dummy that captures whether parents applied to a school located in the municipality of residence. The rationale behind (ii) is to have a proxy of how much parents value school proximity to home rather than maximizing the chance of getting the preferred time scheme, which in some cases would require applying to a school farther away. This dummy takes the value 1 for around 90% of parents. Still, in big municipalities there are many schools, so this dummy would not fully capture preferences for proximity. In a robustness check, we restrict the analysis to the (much smaller) sample of students who applied in schools located in municipalities where there is only one school. These students are more likely to have applied to the closest school, regardless of possible strategic considerations on the application mix.[14]

In our main analysis, we do not consider nor exploit supply-side constraints related to infrastructure and economic resources, due to a lack of detailed information. We do not view this as problematic insofar as we can leverage another supply-side constraint (that stems from class size limits and only depends on the observable application mix) to isolate a plausibly exogenous variation in the probability of attending FT classes conditional on demand. Nonetheless, we perform two robustness checks by running the regressions in a sample of schools where infrastructure and resources are likely not a constraint. In a first exercise, we

[13] This is possible because the relationship between the application mix and the instrument is non-linear. The intuition that limits on minimum and maximum class size generate non-linear changes in class size based on the number of school applicants was popularized by Angrist and Lavy (1999).

[14] In the application data we do not observe the municipality of the school, but we can proxy it with the modal municipality of residence of the pupils who applied to the school.

focus on schools that receive at least 2 applications for both schemes, suggesting that parents believe there is a chance either scheme could be activated in that school-year. In a second exercise, we focus on schools that offered at least one FT class during the last five years before the year when parents apply, again indicating that the school has the infrastructure to potentially offer the long school day scheme.

5 Empirical Analysis

For all education outcomes and for all grades $g \in [2, 5, 8]$, we estimate the following equation:

$$Score^g_{i,s,c,s',p} = \beta^g_0 + \beta^g_1 FT^{g=1}_i + \beta^g_2 A^{g=1}_i + \beta^g_3 X^g_i + \beta^g_4 X^g_c + \beta^g_5 X^{g=1}_s + \alpha^g_{p(s')} + u^g_{i,s,c,s',p} \quad (2)$$

where i, s, c, s', and p stand for student, school-of-application, class, school-of-attendance, and (school-of-attendance) province, respectively. $Score^g_{i,s,c,s',p}$ is the INVALSI test score in mathematics or Italian in grade g. $FT^{g=1}_i$, the main variable of interest, is a dummy that equals 1 when student i attends FT in grade 1 and 0 if she attends PT.[15] $A^{g=1}_i$ is a dummy that equals 1 if the student applied to the FT scheme and 0 if she applied to PT. The vector X^g_i contains observable student-level variables that can affect test scores: gender, immigrant status[16], age, and age squared. We proxy the family background with the highest education level of the mother and of the father.[17] We also include a dummy for whether the student lives in the municipality where the first-preference school is located. The vector X^g_c contains class-level averages of the variables included in X^g_i, as well as class size. The vector $X^{g=1}_s$ includes the school-of-application level variables the instrument is a function of (the total number of applications, the number of applications to FT, and whether the school is located in an isolated, small village). Furthermore, $X^{g=1}_s$ includes a set of variables that capture the characteristics of pupils at the school-of-application level.[18] Finally, because we have shown

[15] We also observe FT attendance in grades 2 and 5. FT attendance is a strongly persistent variable: 90% of pupils who attend an FT class in grade 1 also attend an FT class in grade 5, suggesting that the scope for school and class mobility in primary schools is limited. Our results are unchanged if we defined the treatment dummy as attending the FT scheme in all the grades for which attendance is observed. Notice moreover that attending the FT scheme in lower secondary school in Italy is very rare, as almost all schools offer a PT scheme.

[16] The variable takes 3 values: native, immigrant, and information not recorded. The share of missing values is however close to 0.

[17] These variables take 4 values: elementary or middle school; high school; university; education not reported.

[18] Specifically, it includes: the share of female and of immigrant pupils; the average age of students; the share of mothers (fathers) who have (i) at most lower secondary education, (ii) at most upper secondary education, (iii) tertiary education or more, (iv) missing information about their education.

that the diffusion of FT schemes is larger in the Centre-North of the country and geographical differences along many socio-economic dimensions are large in Italy, we include a set of (school-of-attendance s') province fixed effects ($\alpha^g_{p(s')}$). Because of the possible correlation of the error term $u^g_{i,s,c,s',p}$ within schools, all the models are estimated with standard errors clustered at the school-of-attendance level.

To explore the effect on parents' labor force participation and employment, for all grades $g \in [2, 5, 8]$ we estimate the following equation:

$$Y^g_{i,s,s',p} = \delta^g_0 + \delta^g_1 FT^{g=1}_i + \delta^g_2 A^{g=1}_i + \delta^g_3 X^g_i + \delta^g_4 X^{g=1}_s + \alpha^g_{p(s')} + u^g_{i,s,s',p} \tag{3}$$

where $Y^g_{i,s,s',p}$ is a dummy that takes the value 1 if the parent is active in the labor market or is employed, depending on the specification, when their children attend grade g. The remaining variables are defined as in (2), except for the fact that X^g_i also includes the immigrant status of the parent and that we do not include X^g_c among the controls.

In what follows we estimate models (2) and (3) using OLS and 2SLS. For the 2SLS methodology, we instrument $FT^{g=1}_i$ using Z_{is}, as defined in Section 4.

6 Results

6.1 Main results and robustness checks

This section presents the estimates of the effect of the FT scheme on students' achievement and mothers' labour supply. Table 5 focuses on test scores (italian in columns 1-3 and mathematics in columns 4-6) and displays OLS (Panel A), OLS controlling for parental preferences (Panel B), and 2SLS (Panel C) estimates. The impact of FT on mathematics test scores is rather stable across the three specifications. 2SLS coefficients indicate that attendance of FT in grade 1 causes a statistically significant improvement by 4.8% of a standard deviation in grade 2 and by 4.6% of a standard deviation in grade 5. The positive effect on mathematics skills is no longer visible in grade 8. Controlling for parental preferences and, to a lesser extent, instrumenting the treatment matter more for Italian test scores: in grades 2 and 5, coefficients turn from negative and insignificant in panel A to positive in panel C, although only statistically significant in grade 2. This suggests a negative within-province selection of applicants along this dimension, in line with evidence from Table A1.2 and discusses in Section 4. As for mathematics, no effect is found in grade 8.

The fact that differences between FT and PT students are not significant in grade 8 - 3 years after the end of the program - suggests that benefits in earlier grades might be later counterbalanced by a lower preparedness of FT pupils to autonomous study at home, which becomes an increasingly important skill during lower secondary school. Overall, the estimated effects seem non-negligible. To provide a benchmark for the magnitudes of the effects found in primary school, Pavese and Rubolino (2022) estimate that in the Italian context reducing per-pupil municipal spending by 1,000 euros decreases grade 5 math test scores by approximately 2.4% of a standard deviation.

Table 6 shows the effect of the FT schedule on maternal labor force participation (columns 1-3) and employment (columns 4-6) estimated with OLS (Panel A), OLS controlling for parental preferences (Panel B), and 2SLS (Panel C). The comparison between Panels A and B shows that accounting for the type of application significantly reduces the correlation between attending the FT scheme and maternal employment and labour market participation (approximately from 10 to 3.5 p.p.). This is due to the fact that working mothers are more likely to enroll their children in the FT scheme. Instrumenting the treatment further reduces the coefficient of interest, suggesting that, when the FT scheme is oversubscribed, school principals give priority to children of mothers who are more likely to work. According to 2SLS estimates, the impact on labor force participation is still positive and rather stable over time (around 2 percentage points, from a baseline of 62% among PT parents). Notably, it remains statistically significant up to grade 8. The effect on employment is also positive, but smaller; furthermore, it increases with children's age and becomes significant at conventional levels only in grade 8 (2.2 p.p.). These patterns lend support to the hypothesis that longer school days help mothers to look for and eventually find jobs. The persistence of effects after the end of primary school and the rising pattern of employment coefficients square with the existence of hysteresis in the labor market: temporary periods of inactivity may have persistent consequences on labor market prospects, also due to human capital depreciation. Benchmarking on the estimates provided by Kleven (2021), the magnitude of the effect on grade 8 employment is equivalent to approximately 7% of the employment child penalty for Italy. Table A1.3 indicates that there are instead no effects on paternal employment. This is consistent with the findings of the literature that only mothers' labor supply is elastic to changes in the availability of childcare.

Tables A1.4 (learning) and A1.5 (maternal labor supply) display the results of the two robustness exercises we performed. First, we re-estimate models (2) and (3) on the sub-sample of students who applied to schools located in municipalities where there is only one school (Panel A). The rationale is that concerns of strategic school sorting based on the application mix should be negligible in this sample: as the great majority of parents apply to schools in the same municipality of residence, they likely choose the only available school, without other considerations. The sample shrinks in size (68,113 pupils per year in the achievement regression) and only consists of very small municipalities. Nevertheless, the sign and significance of the results are confirmed; if anything, the coefficients are larger.

Panels B and C report the results of the second robustness check: we re-estimate the regressions on sub-samples of schools where the lack of infrastructure, staff or economic resources is likely not a constraint to the offer of FT classes. In Panel B, we focus on schools that received at least 2 applications for both schemes, which we refer to as "contestable schools", suggesting that parents believe there is a chance that both schemes can be activated. Furthermore, these are the schools where class size limits can generate some unmet demand depending on the application mix, which is the variation used by our instrument. The results are qualitatively in line and quantitatively similar to the main ones. Coefficients on achievement are less precisely estimated due to the smaller sample size; coefficients on mothers' labor force participation and employment, if anything, are larger. In Panel C, we focus on the sub-sample of schools that offered the FT scheme in at least one of the five school years before the application year, which we refer to as "FT legacy schools". Coefficients on achievement are quantitatively similar to the main ones (except the one for math in grade 5 which is smaller and insignificant); coefficients on mothers' labor force participation and employment are, again, larger compared to the main regression.

6.2 Heterogeneity analysis

Table 7 explores heterogeneity in achievement effects by the household socio-economic status (SES). This variable is recorded in grade 5 and captures socio-economic status by combining information on parents' education and occupation, as well as on resources available at home (e.g. books, PCs). It proxies the quality of home inputs the children have access to after school. This is an interesting dimension of heterogeneity because one could posit that additional time at school spent revising under teachers' supervision should benefit more pupils from less privileged backgrounds, who likely would not receive much help when doing homework. To test this, we assign to each child a tercile based on the SES measured in grade 5. We then re-estimate regression (2) by interacting the treatment, the instrument and all control variables with two dummies capturing whether the pupil belongs to the second or third tercile of the SES distribution, respectively.[19] This fully-interacted model delivers the

[19] We prefer using SES rather than parents' education to study heterogeneity effects on test scores for two reasons. First, it has much fewer missing values. Second, by construction, it more comprehensively captures

same estimate as a split-sample exercise, but allows for an easier assessment of whether differences between groups are statistically significant. It turns out that the effect of the FT scheme for students in the upper two terciles is in most grades smaller than that for less affluent pupils, but differences across terciles are not statistically significant.[20]

Table 8 reports the heterogeneous effects on maternal labor force participation and employment by mothers' education.[21] We split mothers into two groups: those with at most lower secondary education and those with at least upper secondary education, with the latter being the reference category. The effect of the FT scheme on labor force participation in earlier grades is stronger for the least educated mothers, who on average have a lower attachment to the labor market, but also in this case, differences - while in some cases large - do not reach the conventional level of statistical significance. The pattern of heterogeneous effects on employment is less clear.

Overall, the heterogeneity analysis uncovers some differences across groups, even if not precisely estimated. With this caveat in mind, the magnitude of the effects seems to be larger for more disadvantaged children and mothers.

6.3 Discussion

The previous analysis showed some heterogeneity in the FT scheme's effect on students and their mothers. We conclude by briefly discussing whether those who benefit the most from FT schemes are also more likely to: (i) apply for it, and (ii) be assigned to a FT scheme if they apply for it.

Table A1.6 describes who is more likely to apply for a long school day and who is less likely to be satisfied conditional on applying to an FT scheme. Column 1 reports the differences in the characteristics of FT and PT applicants. Column 2 considers the sample of FT applicants and compares the characteristics of those not assigned (*mismatched*) and of those assigned to

the quality of the home environment, which is also a function of income (proxied by occupation) and the resources available at home.

[20] In particular, coefficients for pupils in the middle third of the SES distribution are in most cases smaller than those for pupils both in the lower third and the upper third of the distribution. In a few instances, the coefficient for the most well-off pupils is larger than that for least well-off students, even if the only case where it is statistically significant is Mathematics, grade 5. This non-monotonic relationship could reflect the fact that high-SES parents likely both work and hence their children may also benefit from spending the afternoon at school.

[21] In this case using SES as the dimension of heterogeneity is not advisable because (i) is a household-level rather than an individual-level variable and (ii) is built also taking into account information on mothers' occupation. Since we use mothers' education to split the sample, we drop records for which this information is missing.

FT classes. A notable geographic heterogeneity emerges: in the South and Islands not only demand is lower - families are 30 p.p. less likely to apply to FT - but supply-side constraints are also more binding - children who apply to FT are 22 p.p. less likely to attend the preferred scheme than their peers from Northern and Central regions. As a result, the incidence of students in FT classes is far lower in this macro-area, as already emerged from Table A1.2. High socio-economic status students (with working mothers and more educated parents) are both more likely to apply to FT schemes and to be assigned to their FT scheme, once they apply for it. The same holds true for immigrant children. Since Southern regions have lower income, employment rates and migration rates, it is important to assess whether these differences in the characteristics of all FT applicants and of *non mismatched* FT applicants just mirror the North-South divide or are true also at the local level.

To this end, in columns 3 and 4 we add school-of-application fixed effects. Column 3 shows whether disparities among FT and PT applicants exist even within the same school; column 4 provides indirect evidence about the criteria that principals use to manage excesses of demand. It turns out that the higher propensity of higher socio-economic status families to apply for the FT scheme and to be assigned by the principal to the scheme if they apply, survives also at the school level, even if the coefficients are in most cases smaller than in columns 1 and 2. Immigrant families are still more likely to apply; contrary to what emerges in column 2, they are however less likely to end up in the FT scheme when they apply for it. This suggests that the higher probability of being satisfied among immigrants entirely depends on their higher propensity of living in areas with a larger supply of FT classes (Northern Italy); within schools, they are instead more likely to be mismatched.

Since offering FT classes is more costly to public finances than offering PT ones and data suggest the presence of supply-side constraints to the provision of longer school days, these considerations are important for policymakers when deciding where to expand the scheme first. Our results - with the caveat of the statistical significance of the heterogeneity analysis - indicate that students and mothers who benefit the most from the FT scheme are less likely to apply to them and also less likely to be assigned to these schemes when they ask for them. Part of this wedge depends on the allocations rules made by the principals, but a part of it also depends on the fact that the supply of FT classes is far lower in the most disadvantaged areas of the country, characterized by lower female participation, lower socio-economic background, and worse students achievements.

7 Conclusion

This paper estimates the effect of a longer school day in primary school on students' learning outcomes and maternal labor supply both in the short- and medium-term. We analyze the case of Italy, where a longer (FT scheme) and a shorter (PT scheme) school day coexist in primary school. To identify the causal effect of interest, we uniquely match application-to-primary-school data and achievement data for the cohort of pupils who started primary school in 2014-15, who we follow until the end of lower secondary school (grade 8).

Attending an FT class is an equilibrium outcome of demand and supply. Thanks to the individual- and school-level application data, we tackle the endogeneity of allocation that stems from parental preferences - which we explicitly control for in the regression - and from school principals' criteria to manage excess demand for any scheme - by isolating the variation in the probability of being assigned to the preferred scheme that only comes from class size limits set by the law.

Our results indicate that additional time at school has some positive effect on students' performance in primary school, but this effect fades away as students age. In our context, the additional time at school does not entail a significant increase in instruction time but rather crowds out homework close to 1:1. The dynamic profile of the estimated effect therefore could reflect the fact that pupils in FT schemes have positive learning outcomes while in primary school, but might be less likely to develop the ability to study autonomously that is important for their performance in later educational stages. On the other hand, we estimate positive effects on maternal labor market participation and employment that last longer than the end of the FT program. By lowering childcare responsibilities, we indeed find that a longer school day allows mothers to participate more in the labor market when their kids are in primary school, and this has significant long-term impacts on their labor market prospects, as it reduces their human capital depreciation and it increases their labor market experiences at older ages.

We believe our findings bear important implications for policymakers. First, they indicate that, when assessing the benefits of longer school days, it is crucial to jointly consider the effects on students' learning outcomes and those on maternal labour supply. We show indeed that in our context the benefits would have been smaller, and only short-lived, if policymakers only considered the former. Second, our paper shows that the gains from these policies may not be homogeneous across households. Albeit the estimates are noisier and differences do not usually reach conventional levels of statistical significance, we find suggestive evidence that the positive short-term effect on students' achievement is larger among low socio-economic status students, who have fewer resources available outside school, and that the impact on mothers' labor force participation is stronger for the least educated, who are on average less attached to the labor market.

We highlight moreover that there is some negative selection on gains, since families who benefit the most from the FT program according to our heterogeneity analysis are not those more likely to ask for it and to be assigned to a full-time schedule when they ask for it. On the one hand, this suggests that expanding the supply of full-day programs in primary school could have larger benefits than the ones estimated in our analysis, since students and mothers who are now excluded from the scheme are probably those who would benefit the most from it. Second, even without expanding the supply of the program, our results call for changes in the principals' allocation rules and in the information set provided to parents during the school application process, so to increase the likelihood that more disadvantaged students would apply for and attend the FT scheme.

References

Angrist, J. D., Battistin, E. and Vuri, D. (2017). In a small moment: Class size and moral hazard in the italian mezzogiorno, *American Economic Journal: Applied Economics* 9(4): 216–249.

Angrist, J. D. and Lavy, V. (1999). Using maimonides' rule to estimate the effect of class size on scholastic achievement, *The Quarterly Journal of Economics* 114(2): 533–575.

Baker, M., Gruber, J. and Milligan, K. (2008). Universal child care, maternal labor supply, and family well-being, *Journal of Political Economy* .

Barbetta, G. P., Chuard, P., Sorrenti, G. and Turati, G. (2022). Good or bad? short- versus long-term effects of multigrading on child achievement, *Tinbergen Institute Discussion Paper 2022-025/V* .

Barbetta, G. P., Sorrenti, G. and Turati, G. (2021). Multigrading and Child Achievement, *Journal of Human Resources* 56(3): 940–968.

Barrios-Fernández, A. and Bovini, G. (2021). It's time to learn: School institutions and returns to instruction time, *Economics of Education Review* 80: 102068.

Bellei, C. (2009). Does lengthening the school day increase students' academic achievement? Results from a natural experiment in Chile, *Economics of Education Review* 28(5): 629–640.

Berthelon, M., Kruger, D. and Oyarzún, M. (2023). School schedules and mothers' employment: evidence from an education reform, *Review of Economics of the Household* 21(1): 131–171.

Bertoni, M., Brunello, G. and Rocco, L. (2013). When the cat is near, the mice won't play: The effect of external examiners in italian schools, *Journal of Public Economics* 104: 65–77.

Bovini, G., Philippis, M. D. and Sestito, P. (2016). Il tempo pieno e la dispersione dei voti full-time school scheme and inequality in students' learning outcomes, *I dati INVALSI: uno strumento per la ricerca* p. 125.

Carneiro, P., Heckman, J. J. and Vytlacil, E. J. (2011). Estimating marginal returns to education, *American Economic Review* 101(6): 2754–2781.

Carta, F. and Rizzica, L. (2018). Early kindergarten, maternal labor supply and children's outcomes: evidence from italy, *Journal of Public Economics* 158: 79–102.

Cascio, E. U. (2009). Maternal labor supply and the introduction of kindergartens into american public schools, *The Journal of Human Resources* 44(1): 140–170.

Checchi, D. and De Paola, M. (2018). The effect of multigrade classes on cognitive and non-cognitive skills. causal evidence exploiting minimum class size rules in italy, *Economics of Education Review* 67: 235–253.

Contreras, D. and Sepúlveda, P. (2017). Effect of Lengthening the School Day on Mother's Labor Supply, *The World Bank Economic Review* 31(3): 747–766.

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

Cornelissen, T., Dustmann, C., Raute, A. and Schönberg, U. (2018). Who benefits from universal child care? estimating marginal returns to early child care attendance, *Journal of Political Economy* 126(6): 2356–2409.

Dehos, F. T. and Paul, M. (2021). The effects of after-school programs on maternal employment, *Journal of Human Resources* pp. 0120–10651R1.

Dominguez, P. and Ruffini, K. (2023). Long-term gains from longer school days, *Journal of Human Resources* 58(4): 1385–1427.

Duchini, E. and Van Effenterre, C. (2022). School Schedule and the Gender Pay Gap, *Journal of Human Resources* .

Felfe, C., Lechner, M. and Thiemann, P. (2016). After-school care and parents' labor supply, *Labour Economics* 42: 64–75.

Felfe, C. and Zierow, L. (2014). After-school center-based care and children's development, *The B.E. Journal of Economic Analysis Policy* 14.

Figlio, D., Holden, K. L. and Ozek, U. (2018). Do students benefit from longer school days? regression discontinuity evidence from florida's additional hour of literacy instruction, *Economics of Education Review* 67: 171–183.

Fitzpatrick, M. (2010). Preschoolers enrolled and mothers at work? the effects of universal prekindergarten, *Journal of Labor Economics* 28(1): 51–85.

Frisvold, D. E. (2015). Nutrition and cognitive achievement: An Evaluation of the School Breakfast Program, *Journal of Public Economics* 124: 91–104.

Gambaro, L., Marcus, J. and Peter, F. (2019). School entry, afternoon care, and mothers' labour supply, *Empirical Economics* 57(3): 769–803.

Goux, D. and Maurin, E. (2010). Public school availability for two-year olds and mothers' labour supply, *Labour Economics* 17(6): 951–962.

Havnes, T. and Mogstad, M. (2011). Money for nothing? universal child care and maternal employment, *Journal of Public Economics* 95(11): 1455–1465. Special Issue: International Seminar for Public Economics on Normative Tax Theory.

Kleven, H. (2021). https://www.henrikkleven.com/uploads/3/7/3/1/37310663/zeuthen_lecture1_overview.pdf.

Lavy, V. (2020). Expanding School Resources and Increasing Time on Task: Effects on Students' Academic and Noncognitive Outcomes, *Journal of the European Economic Association* 18(1): 232–265.

Lefebvre, P. and Merrigan, P. (2008). Child-care policy and the labor supply of mothers with young children: A natural experiment from canada, *Journal of Labor Economics* .

Martínez A., C. and Perticará, M. (2017). Childcare effects on maternal employment: Evidence from Chile, *Journal of Development Economics* 126(C): 127–137.

Nollenberger, N. and Rodríguez-Planas, N. (2015). Full-time universal childcare in a context of low maternal employment: Quasi-experimental evidence from spain, *Labour Economics* 36: 124–136.

Nybom, M. (2017). The distribution of lifetime earnings returns to college, *Journal of Labor Economics* 35(4): 903–952.

OECD (2015). *Education Policy Outlook 2015: Making Reforms Happen.*

Padilla-Romo, M. and Cabrera-Hernández, F. (2019). Easing The Constraints Of Motherhood: The Effects Of All-Day Schools On Mothers' Labor Supply, *Economic Inquiry* 57(2): 890–909.

Pavese, C. and Rubolino, E. (2022). Austerity Harmed Student Achievement, *Cahiers de Recherches Economiques du Département d'économie 22.09*, Université de Lausanne, Faculté des HEC, Département d'économie.

Quintano, C., Castellano, R. and Longobardi, S. (2009). A fuzzy clustering approach to improve the accuracy of italian student data : an experimental procedure to correct the impact of outliers on assessment test scores, *Statistica & Applicazioni : VII, 2, 2009* .

Schmitz, L. (2022). Heterogeneous Effects of After-School Care on Child Development, *Discussion Papers of DIW Berlin 2006*, DIW Berlin, German Institute for Economic Research.

Seidlitz, A. and Zierow, L. (2020). The Impact of All-Day Schools on Student Achievement - Evidence from Extending School Days in German Primary Schools, *CESifo Working Paper Series 8618*, CESifo.

Figures

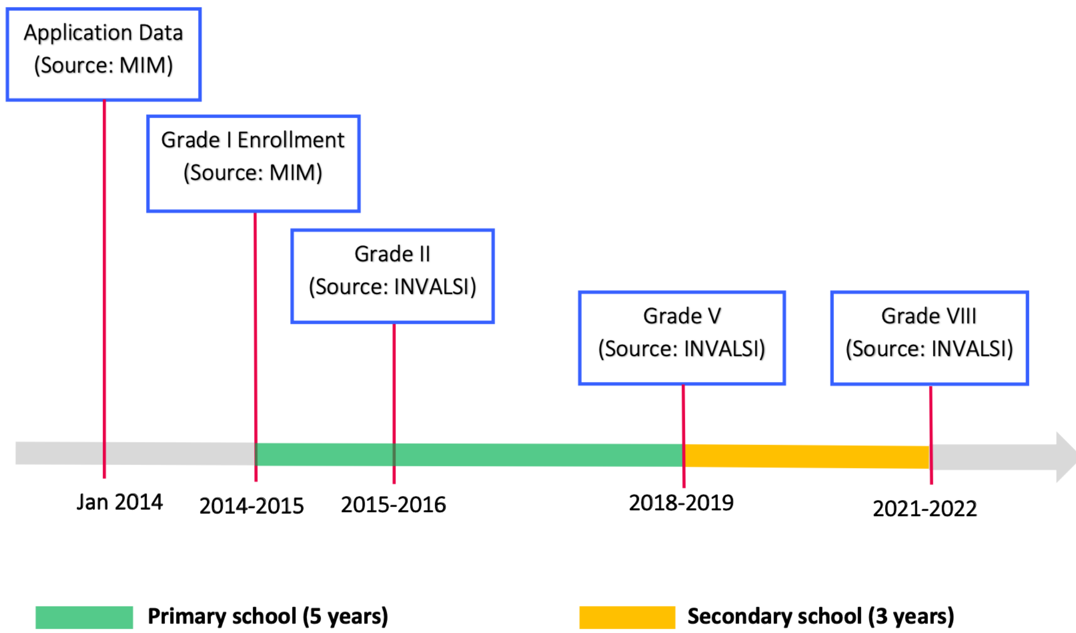
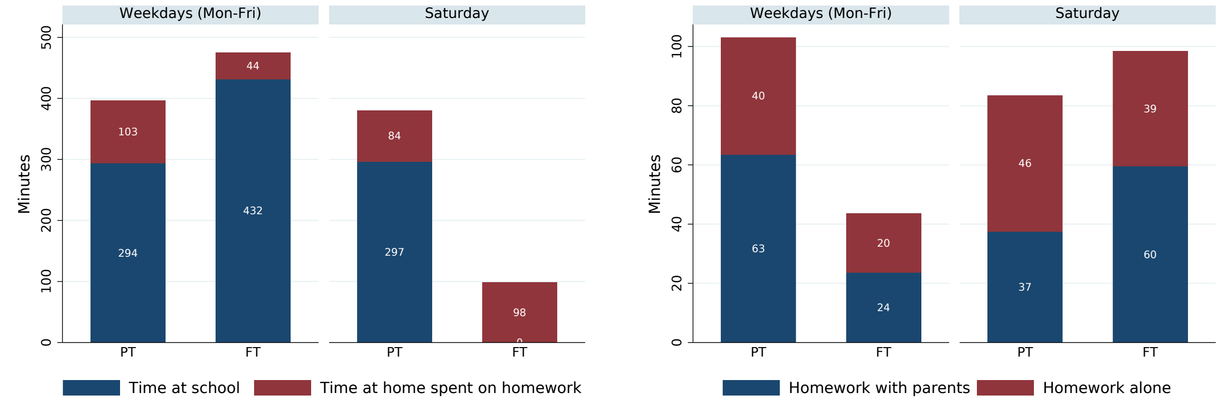


Figure 1: **Time Line**

Note: The figure shows the points in time and the sources our data refer to. Typically, children start primary school in the year when they turn 6.



(a) Total time at school + homework at home

(b) Total time doing homework at home

Figure 2: **Use of Time At and Outside School**

Note: The figure shows the time that students in PT and FT spend in school and doing homework at home (Panel (a)) and they spend doing homework at home with or without parents (Panel (b)). The Time Use Survey does not explicitly record the instructional scheme the student is enrolled to. Focusing on the sample of children who attend primary school and are interviewed from October to May and from Monday to Saturday - i.e. during the months and days when primary schools are open in Italy - we assign pupils to the FT scheme if they report that the last class ends after 2 p.m, whereas we assign pupils to the PT scheme if either they report that the last class ends no later than 2 p.m or if they report attending school on Saturdays. Source: Italian Time Use survey, 2013 wave.

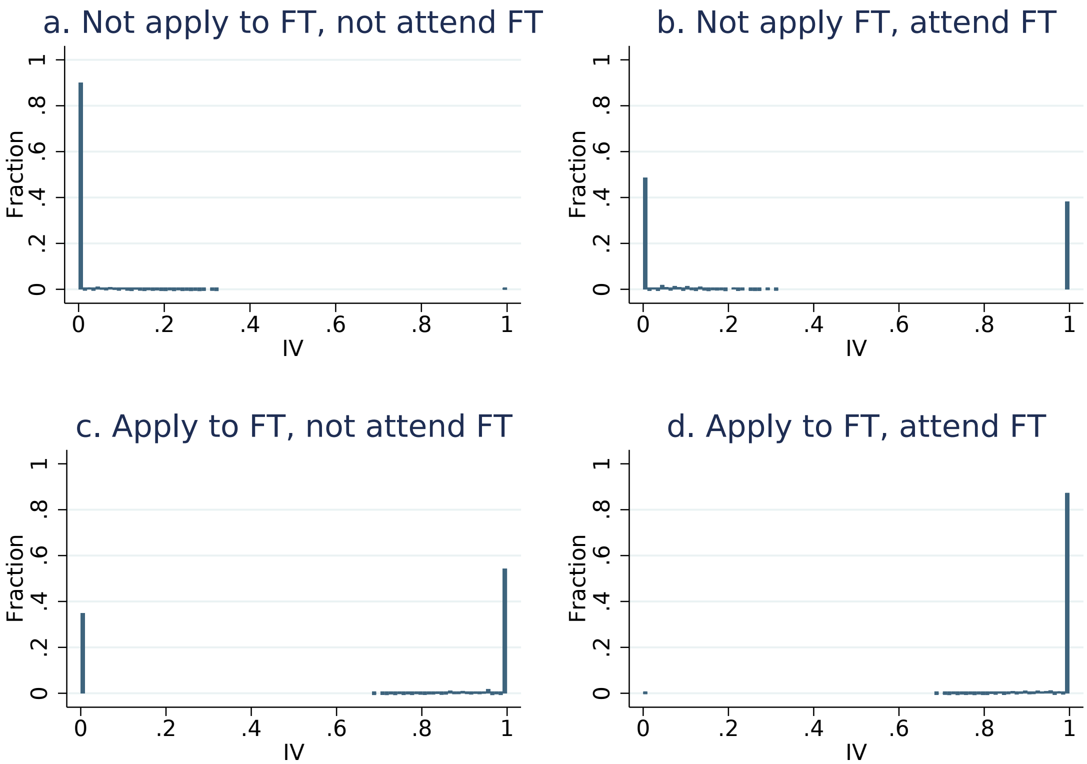


Figure 3: **IV Distribution**

Note: The figure shows the distribution of the instrument in the four subsamples defined by whether the students applied to and attended the full-time scheme. The instrument is defined in (1) and is the probability to be assigned to the full-time scheme as a result of the school-principal algorithm, conditional on expressing a preference for a given scheme.

Tables

Table 1: **Use of Time Outside School (INVALSI)**

	(1) Attend FT Mean	(2) Attend PT Mean	(3) Difference Δ
	Homework at home		
Share never do homework in a week	0.03	0.02	-0.01***
Share do homework 1-2 times per week	0.38	0.13	-0.25***
Share do homework 3-4 times per week	0.36	0.23	-0.13***
Share do homework more than 5 times a week	0.24	0.63	0.39***
	Leisure at home		
Share watch TV more than 1 hr per day	0.53	0.52	-0.01***
Share play with PC/videogames more than 1 hr per day	0.46	0.47	0.01***
Share play with friends more than 1 hr per day	0.81	0.81	-0.00
Share help with housework more than 1 hr per day	0.39	0.42	0.02***
Share read books/comics more than 1 hr per day	0.06	0.08	0.02***
Share play sport more than 3 times per week	0.37	0.38	0.02***
Observations	254,773	693,765	948,538

Note: Shares are conditional on non-missing survey responses. Source: INVALSI student-level survey administered to fifth graders in scholastic years 2011-12 and 2012-13.

Table 2: **Comparison between FT and PT teachers: Grade 2**

	Italian				Math			
	Coefficient	**Se**	**Mean**	**N**	**Coefficient**	**Se**	**Mean**	**N**
Demographics								
Female	-0.01	0.012	0.98	1,166	-0.03	0.021	0.95	1,164
Open-ended contract	-0.01	0.015	0.96	1,166	-0.01	0.014	0.96	1,164
Tenure (=1 if > 5 years)	-0.10***	0.022	0.79	1,166	-0.08***	0.024	0.76	1,164
Under 50	0.01	0.044	0.44	1,166	0.05	0.033	0.46	1,164
University education	0.11***	0.039	0.32	1,166	0.05	0.042	0.32	1,164
Weekly hours taught	0.63***	0.12	7.75	1,268	0.90***	0.144	6.56	1,274
INVALSI test preparation (=1 if using)								
Class exercises similar to INVALSI	-0.03	0.039	0.45	1,278	-0.02	0.032	0.46	1,283
Homework similar to INVALSI	-0.05**	0.019	0.15	1,278	-0.07***	0.017	0.17	1,283
INVALSI tests from previous years	0.01	0.035	0.58	1,278	0.05	0.03	0.58	1,283
INVALSI textbook	-0.06	0.038	0.54	1,278	-0.01	0.039	0.51	1,283
None of the above	0.01	0.006	0.01	1,278	-0.02***	0.008	0.02	1,283
Teaching practices (=1 if frequent adopter)								
Assignment of projects	-0.01	0.033	0.22	1,278	0.00	0.026	0.2	1,283
Enhancing activities	-0.02	0.034	0.19	1,278	-0.01	0.025	0.16	1,283
Group work	0.05	0.03	0.51	1,278	0.06	0.043	0.48	1,283
In-class discussion of homework	0.00	0.014	0.94	1,278	-0.01	0.018	0.94	1,283
Remedial activities	-0.03	0.024	0.17	1,278	-0.01	0.028	0.15	1,283
Use techniques learnt in training courses	-0.06*	0.03	0.6	1,278	0.05	0.033	0.59	1,283
Technology (=1 if frequent user)								
Camera	-0.01	0.02	0.08	1,278	0.02	0.022	0.07	1,283
Computer	0.06**	0.024	0.4	1,278	0.08***	0.02	0.38	1,283
Interactive whiteboard	0.06	0.037	0.35	1,278	0.07*	0.039	0.35	1,283
Internet connection	0.01	0.038	0.43	1,278	0.02	0.039	0.41	1,283
Projector	0.01	0.016	0.12	1,278	0.00	0.021	0.1	1,283
Tablet	0.02	0.029	0.15	1,278	0.00	0.036	0.13	1,283
Evaluation methods (=1 if frequent adopter)								
Closed-ended questions	-0.06*	0.031	0.78	1,278	-0.02	0.03	0.7	1,283
Group oral test	0.01	0.024	0.72	1,278	0.05	0.038	0.7	1,283
Open-ended questions	-0.06**	0.024	0.71	1,278	-0.04	0.042	0.57	1,283
Planned oral test	-0.04	0.035	0.35	1,278	0.01	0.024	0.35	1,283
Text-book exercises	-0.05	0.04	0.69	1,278	-0.06	0.037	0.66	1,283
Unplanned oral test	-0.07	0.047	0.56	1,278	0.00	0.043	0.55	1,283
Interactions with colleagues (=1 if frequent adopter)								
Exchange opinions on teaching practices	-0.02	0.02	0.88	1,164	0	0.021	0.89	1,159
Exchange teaching material	0.04	0.026	0.78	1,165	-0.05**	0.023	0.77	1,158
Preparation of teaching material	0.02	0.032	0.76	1,165	-0.01	0.027	0.76	1,159
Project joint educational activities	-0.01	0.035	0.71	1,162	0.00	0.015	0.72	1,156
Share evaluation material	-0.01	0.05	0.77	1,164	-0.02	0.028	0.76	1,158
Share information on textbooks	0.00	0.027	0.73	1,164	-0.03	0.028	0.75	1,160

Note: Coefficients are from regressions of each of the variables on the FT dummy, controlling for region fixed-effects. Standard errors are clustered at the region level. Data come from the INVALSI teachers' 2015/2016 questionnaire for Grade 2 (the year when Grade 2 students in our estimation sample take the INVALSI tests). Since the questionnaire is administered only to a representative sample of teachers, observations are weighted with class weight (different for Italian and Math) provided by INVALSI.
*** denotes significance at 1%, ** denotes significance at 5%, * denotes significance at 10%.

Table 3: **Characteristics of FT relatively to PT teachers: Grade 5**

	Italian teachers				Math teachers			
	Coefficient	**Se**	**Mean**	**N**	**Coefficient**	**Se**	**Mean**	**N**
Demographics								
Experience	-0.93	1.09	27.44	814	-1.97***	0.591	27.32	808
Female	-0.02	0.014	0.96	816	-0.02	0.016	0.96	812
Open-ended contract	-0.04**	0.018	0.96	816	-0.01	0.013	0.96	812
Tenure (=1 if > 5 years)	-0.07***	0.02	0.8	816	-0.03	0.027	0.77	812
Under 50	-0.01	0.061	0.37	816	0.04	0.03	0.39	812
University education	0.07**	0.033	0.33	816	0.08*	0.04	0.29	812
Weekly hours taught	0.89***	0.182	7.36	816	0.97***	0.206	6.58	812
INVALSI test preparation (=1 if using)								
Class exercises similar to INVALSI	0.01	0.043	0.35	904	0.03	0.039	0.34	902
Homework similar to INVALSI	0.01	0.04	0.2	904	0.01	0.031	0.2	902
INVALSI tests from previous years	-0.05	0.041	0.53	904	0.02	0.046	0.56	902
INVALSI textbook	-0.02	0.037	0.5	904	0.00	0.032	0.48	902
None of the above	0.00	0.008	0.01	904	0.02*	0.011	0.02	902
Teaching practices (=1 if frequent adopter)								
Definition of rules and concepts	0.02	0.023	0.9	816	0.00	0.024	0.86	812
Enhancing activities	0.01	0.038	0.15	816	0.04*	0.021	0.15	812
Flipped classroom	0.09**	0.033	0.16	816	-0.02	0.037	0.16	812
Interdisciplinary activities	0.07*	0.036	0.72	816	0.01	0.036	0.67	812
Peer-learning	0.11**	0.042	0.67	816	0.10**	0.046	0.69	812
Remedial activities	0.01	0.033	0.21	816	0.01	0.03	0.18	812
Technology (=1 if frequent user)								
Camera	0.03	0.037	0.66	816	-0.01	0.02	0.05	812
Computer	-0.06***	0.021	0.1	816	0.09***	0.032	0.41	812
E-learning platform	0.00	0.049	0.62	816	-0.05*	0.027	0.09	812
Educational software	-0.01	0.045	0.3	816	0.03	0.031	0.21	812
Interactive whiteboard	0.02	0.031	0.15	816	0.05	0.05	0.51	812
Smartphone	0.02	0.034	0.72	816	0.03	0.029	0.08	812
Tablet	-0.03	0.045	0.59	816	-0.01	0.028	0.12	812
Evaluation methods (=1 if frequent adopter)								
Closed-ended and open-ended questions	-0.02	0.038	0.82	816	0.00	0.031	0.79	812
Closed-ended questions	-0.04	0.035	0.78	816	0.04	0.047	0.75	812
Evaluation platforms	0.04	0.027	0.07	816	-0.02	0.016	0.06	812
Group work	0.11**	0.046	0.45	816	0.06	0.048	0.4	812
Homework evaluation	0.02	0.037	0.64	816	-0.03	0.042	0.52	812
Open-ended questions	0.00	0.024	0.89	816	-0.03	0.021	0.91	812
Students' self-evaluation	0.06	0.034	0.58	816	0.04	0.044	0.59	812
Text-book exercises	-0.03	0.038	0.67	816	0.01	0.044	0.63	812
Interactions with colleagues (=1 if frequent adopter)								
Exchange teaching material	0.07**	0.028	0.76	816	0.04	0.032	0.76	812
Preparation of teaching material	0.07*	0.038	0.75	816	0.04	0.035	0.72	812
Project joint educational activities	0.09***	0.025	0.71	816	0.09***	0.03	0.66	812
Share evaluation material	0.04*	0.021	0.81	816	0.06	0.046	0.78	812
Share information on textbooks	0.02	0.036	0.76	816	0.07*	0.035	0.75	812

Note: Coefficients are from regressions of each of the variables on the FT dummy, controlling for region fixed-effects. Standard errors are clustered at the region level. Data come from the INVALSI teachers' 2018/2019 questionnaire for Grade 5 (the year when Grade 5 students in our estimation sample take the INVALSI tests). Since the questionnaire is administered only to a representative sample of teachers, observations are weighted with class weight (different for Italian and Math) provided by INVALSI.
*** denotes significance at 1%, ** denotes significance at 5%, * denotes significance at 10%.

Table 4: **Outcome of the Application Process**

	Attend FT in Preferred School	Attend FT in Other School	Attend PT in Preferred School	Attend PT in Other School	Total
FT Applicant	113,184	4,115	14,642	2,557	134,498
	[84.15]	[3.06]	[10.89]	[1.90]	[100]
PT Applicant	5,525	1,270	190,120	7,449	204,364
	[2.70]	[0.62]	[93.03]	[3.64]	[100]

Note: The statistics are computed on the estimation sample. Source: MIM application data.

Table 5: **Cognitive Development: Effect of Attending FT**

	Italian			**Math**		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	Score 2	Score 5	Score 8	Score 2	Score 5	Score 8
			Panel A: OLS			
1=attend FT in grade 1	-0.003	-0.005	-0.034***	0.036***	0.035***	-0.010*
	(0.010)	(0.007)	(0.005)	(0.011)	(0.009)	(0.006)
	[0.756]	[0.444]	[0.000]	[0.001]	[0.000]	[0.071]
N	338862	338862	338862	338862	338862	338862
			Panel B: OLS Controlling for Parental Preferences			
1=attend FT in grade 1	0.014	0.018**	-0.010	0.042***	0.046***	0.002
	(0.011)	(0.009)	(0.007)	(0.013)	(0.010)	(0.007)
	[0.221]	[0.036]	[0.150]	[0.001]	[0.000]	[0.756]
N	338862	338862	338862	338862	338862	338862
			Panel C: 2SLS			
1=attend FT in grade 1	0.043*	0.024	-0.009	0.048*	0.046**	0.003
	(0.022)	(0.018)	(0.015)	(0.025)	(0.021)	(0.016)
	[0.056]	[0.198]	[0.572]	[0.054]	[0.029]	[0.859]
F-stat	5435	5880	7164	5435	5880	7164
N	338862	338862	338862	338862	338862	338862
Student-level controls	Yes	Yes	Yes	Yes	Yes	Yes
Class-level controls	Yes	Yes	Yes	Yes	Yes	Yes
School-level controls	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: The sample includes the students who took both the Italian and the Math INVALSI tests in all grades. In the 2SLS regressions, we instrument the dummy for attending FT with the instrument defined in (1). In all the regressions in Panels B and C we control for a dummy for applying to FT. Student-level controls: age, age squared, dummies for gender, applying to a school located in the municipality of residence, immigrant status, mother's education level, father's education level. School-of-application-level controls: number of applications received by the school, number of FT applications received by the school, dummy for whether the school is located in a small, isolated village, average students' age, share of female students, share of immigrant students, share students whose mother's (father's) education level is below-high school, high school, university, missing. Class-level controls: class size, average students' age, share of female students, share of immigrant students, share of students with missing immigrant status, share of students whose mother's (father's) education level is below-high school, high school, university, missing. All regressions include school-of-attendance province fixed effects. Robust standard errors clustered at the school-of-attendance level are shown in parenthesis, and p-values in are shown in brackets. The F-statistic on the excluded instrument is reported for 2SLS regressions. The sample is restricted to the public schools in which the number of applications received is larger than or equal to the minimum number of students needed to form a class. *** denotes significance at 1%, ** denotes significance at 5%, * denotes significance at 10%.

Table 6: **Mothers' Labor Supply: Effect of Attending FT**

	Labor Force Participation			**Employment**		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	Grade 2	Grade 5	Grade 8	Grade 2	Grade 5	Grade 8
			Panel A: OLS			
1=attend FT in grade 1	0.109***	0.106***	0.096***	0.104***	0.103***	0.094***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
	[0.000]	[0.000)]	[0.000]	[0.000]	[0.000]	(0.000)
N	237456	237456	237456	237456	237456	237456
			Panel B: OLS Controlling for Parental Preferences			
1=attend FT in grade 1	0.039***	0.037***	0.035***	0.034***	0.033***	0.032***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
N	237456	237456	237456	237456	237456	237456
			Panel C: 2SLS			
1=attend FT in grade 1	0.020**	0.022***	0.021**	0.010	0.014	0.022**
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
	[0.022]	[0.010]	[0.016]	[0.271]	[0.106]	[0.013]
F-stat	4378	4697	5544	4378	4697	5544
N	237456	237456	237456	237456	237456	237456
Student-level controls	Yes	Yes	Yes	Yes	Yes	Yes
School-level controls	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: The sample includes the mothers whose employment status we observed when their children are in grades 2, 5, and 8. The dependent variable is a dummy for whether the mother is in the labor force (columns (1) to (3)) or employed (columns (4) to (6)). In the 2SLS regressions, we instrument the dummy for attending FT with the instrument defined in (1). In all the regressions in Panels B and C we control for a dummy for applying to FT. Student-level controls: age, age squared, dummies for gender, applying to a school located in the same municipality of residence, immigrant status, mother's education level, father's education level, and mother's immigrant status. School-of-application-level controls: number of applications received by the school, number of FT applications received by the school, dummy for whether the school is located in a small, isolated village, average students' age, share of female students, share of immigrant students, share students whose mother's (father's) education level is below-high school, high school, university, missing. All regressions include school-of-attendance province fixed effects. Robust standard errors clustered at the school-of-attendance level are shown in parenthesis, and p-values are shown in brackets. The F-statistic on the excluded instrument is reported for 2SLS regressions. The sample is restricted to the public schools in which the number of applications received is larger than or equal to the minimum number of students needed to form a class. *** denotes significance at 1%, ** denotes significance at 5%, * denotes significance at 10%.

Table 7: **Cognitive Development: Heterogeneity by SES (2SLS)**

	Italian			**Math**		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	Score 2	Score 5	Score 8	Score 2	Score 5	Score 8
1=attend FT in grade 1	0.078**	0.034	-0.021	0.066*	0.052	0.002
	(0.032)	(0.029)	(0.025)	(0.038)	(0.033)	(0.026)
	[0.015]	[0.239]	[0.403]	[0.080]	[0.109]	[0.947]
1=attend FT in grade 1 x 1=(tercile SES=2)	-0.065	-0.041	0.014	-0.022	-0.023	-0.000
	(0.041)	(0.038)	(0.036)	(0.043)	(0.040)	(0.036)
	[0.108]	[0.284]	[0.705]	[0.617]	[0.561]	[0.996]
1=attend FT in grade 1 x 1=(tercile SES=3)	-0.042	0.009	0.026	-0.031	0.011	0.010
	(0.041)	(0.038)	(0.034)	(0.044)	(0.040)	(0.035)
	[0.305]	[0.810]	[0.451]	[0.472]	[0.785]	[0.779]
F-stat	1133	1201	1364	1133	1201	1364
N	338470	338470	338470	338470	338470	338470
Student-level controls	Yes	Yes	Yes	Yes	Yes	Yes
Class-level controls	Yes	Yes	Yes	Yes	Yes	Yes
School-level controls	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: The sample includes the students who took both the Italian and the Math INVALSI tests in all grades. Estimates come from a fully-interacted model, where we estimate regression (2) by interacting the treatment and all controls with two dummies capturing whether the pupil belongs to the second or third tercile of the SES distribution. We instrument the treatment dummy and the interactions between the treatment dummy and the SES tercile dummies with the instrument defined in (1) and the interactions between the instrument and the SES tercile dummies. In all the regressions we control for a dummy for applying to FT. Student-level controls: age, age squared, dummies for gender, applying to a school located in the same municipality of residence, immigrant status, mother's education level, father's education level, dummy for second tercile of the SES distribution, dummy for the third tercile of the SES distribution. School-of-application-level controls: number of applications received by the school, number of FT applications received by the school, dummy for whether the school is located in a small, isolated village, average students' age, share of female students, share of immigrant students, share students whose mother's (father's) education level is below-high school, high school, university, missing. Class-level controls: class size, average students' age, share of female students, share of immigrant students, share of students with missing immigrant status, share of students whose mother's (father's) education level is below-high school, high school, university, missing. All regressions include school-of-attendance province fixed effects. Robust standard errors clustered at the school-of-attendance level shown in parenthesis, and p-values in brackets. The sample is restricted to the public schools in which the number of applications received is larger or equal than the minimum number of students needed to form a class.

Table 8: **Mothers' Labor Supply: Heterogeneity by Education Level (2SLS)**

	Labor Force Participation			**Employment**		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	Grade 2	Grade 5	Grade 8	Grade 2	Grade 5	Grade 8
1=attend FT in grade 1	0.012	0.020**	0.019**	0.008	0.012	0.019**
	(0.010)	(0.009)	(0.009)	(0.010)	(0.010)	(0.010)
	[0.236]	[0.036]	[0.034]	[0.413]	[0.210]	[0.045]
1=attend FT in grade 1 x Below HS	0.022	0.009	-0.004	0.001	0.006	0.000
	(0.019)	(0.019)	(0.018)	(0.019)	(0.019)	(0.019)
	[0.240]	[0.642]	[0.816]	[0.966]	[0.753]	[0.982]
F-stat	1395	1408	1550	1395	1408	1550
N	233507	233507	233507	233507	233507	233507
Student-level controls	Yes	Yes	Yes	Yes	Yes	Yes
School-level controls	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: In all the regressions, the sample includes the mothers whose employment status we observed when their children are in grade 2, 5, and 8. Estimates come from a fully-interacted model, where we estimate regression (3) by interacting the treatment and all controls with a dummy for at most lower secondary education (i.e., Below HS). We instrument the treatment dummy and the interaction between the treatment dummy and the Below HS dummy with the instrument defined in (1) and the interaction between the instrument and the Below HS dummy. In all the regressions we control for a dummy for applying to FT. Student-level controls: age, age squared, dummies for gender, applying to a school located in the same municipality of residence, immigrant status, father's education level, mother's immigrant status, Below HS dummy. School-of-application-level controls: number of applications received by the school, number of FT applications received by the school, dummy for whether the school is located in a small, isolated village, average students' age, share of female students, share of immigrant students, share students whose mother's (father's) education level is below-high school, high school, university, missing. All regressions include school-of-attendance province fixed effects. Robust standard errors clustered at the school-of-attendance level shown in parenthesis, and p-values in brackets. The sample is restricted to the public schools in which the number of applications received is larger or equal than the minimum number of students needed to form a class. *** denotes significance at 1%, ** denotes significance at 5%, * denotes significance at 10%.

Appendices

A1 Appendix Tables

Table A1.1: **Summary statistics: Estimation sample vs INVALSI universe**

	(1) Estimation Sample Mean	(2) INVALSI universe Mean
Enrolled in FT scheme	0.36	0.35
Italian test score	0.07	0.00
Math test score	0.07	-0.00
Mother employed	0.62	0.60
Mother in labor force	0.68	0.66
Father employed	0.94	0.93
Father in labor force	0.99	0.99
Female	0.50	0.49
Age	7.92	7.93
Immigrant	0.08	0.11
Immigrant status missing	0.00	0.00
Mother elementary or middle school	0.27	0.28
Mother high school	0.35	0.33
Mother university	0.20	0.19
Mother education missing	0.18	0.20
Father elementary or middle school	0.35	0.35
Father high school	0.32	0.30
Father university	0.14	0.14
Father education missing	0.20	0.21
Class size	20.80	20.37
North	0.47	0.46
Center	0.19	0.19
South and Islands	0.34	0.34
Observations	338,862	502,412

Note: The statistics are measured in grade 2.

Table A1.2: **Comparison between FT and PT students**

	(1) Attend FT Mean	(2) Attend PT Mean	(3) Difference Δ	(4) Difference Within Provinces
Italian test score	0.09	0.06	0.02***	-0.02***
Math test score	0.09	0.06	0.03***	0.01
Mother employed	0.73	0.56	0.17***	0.08***
Mother in labor force	0.79	0.62	0.17***	0.08***
Father employed	0.95	0.94	0.01***	-0.01***
Father in labor force	0.99	0.99	-0.00	-0.00
Female	0.49	0.50	-0.01***	-0.01***
Age	7.95	7.91	0.05***	0.00***
Immigrant	0.12	0.07	0.05***	0.03***
Mother elementary or middle school	0.26	0.31	-0.06***	-0.01***
Mother high school	0.37	0.38	-0.01***	-0.01***
Mother university	0.24	0.20	0.04***	0.01***
Mother education missing	0.13	0.10	0.03***	0.01**
Father elementary or middle school	0.34	0.40	-0.06***	-0.01***
Father high school	0.34	0.35	-0.01**	-0.01*
Father university	0.17	0.14	0.03***	0.00
Father education missing	0.14	0.11	0.03***	0.02***
Class size	21.43	20.44	0.99***	0.52***
North	0.61	0.39	0.23***	
Center	0.25	0.16	0.09***	
South and Islands	0.14	0.46	-0.32***	
Observations	124,094	214,768	338,862	338,862

Note: Statistics are Measured in Grade 2; the comparison refers to the estimation sample. Coefficients in columns (3) and (4) are from regressions of each of the variables on the FT dummy; in columns (4) we included province fixed effects. Standard errors are clustered at the province level.

Table A1.3: **Fathers' Labor Supply: Effect of Attending FT (2SLS)**

	Labor Force Participation			Employment		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	Grade 2	Grade 5	Grade 8	Grade 2	Grade 5	Grade 8
1=attend FT in grade 1	0.000	-0.001	-0.004	-0.002	0.000	0.000
	(0.002)	(0.002)	(0.003)	(0.005)	(0.005)	(0.004)
	[0.889]	[0.552]	[0.106]	[0.762]	[0.942]	[0.975]
F-stat	4369	4680	5541	4369	4680	5541
N	230230	230230	230230	230230	230230	230230
Student-level controls	Yes	Yes	Yes	Yes	Yes	Yes
School-level controls	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: The sample includes the fathers whose employment status is reported when their children are in grade 2, 5, and 8. The dependent variable is a dummy for whether the father is in the labor force (columns (1) to (3)) or employed (columns (4) to (6)). In all the regressions, we instrument the dummy for attending FT with the instrument defined in (1) and we control for a dummy for applying to FT. Student-level controls: age, age squared, dummies for gender, applying to a school located in the same municipality of residence, immigrant status, mother's education level, father's education level, father's immigrant status. School-of-application-level controls: number of applications received by the school, number of FT applications received by the school, dummy for whether the school is located in a small, isolated village, average students' age, share of female students, share of immigrant students, share students whose mother's (father's) education level is below-high school, high school, university, missing. All regressions include school-of-attendance province fixed effects. Robust standard errors clustered at the school-of-attendance level are shown in parenthesis, and p-values are shown in brackets. The F-statistic on the excluded instrument is reported for 2SLS regressions. The sample is restricted to the public schools in which the number of applications received is larger than or equal to the minimum number of students needed to form a class. *** denotes significance at 1%, ** denotes significance at 5%, * denotes significance at 10%.

Table A1.4: **Cognitive Development: Robustness Checks (2SLS)**

	Italian			**Math**		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	Score 2	Score 5	Score 8	Score 2	Score 5	Score 8
	Panel A: Single-school Municipalities					
1=attend FT in grade 1	0.089**	0.032	-0.004	0.095**	0.079**	-0.023
	(0.038)	(0.034)	(0.027)	(0.042)	(0.039)	(0.029)
	[0.019]	[0.351]	[0.869]	[0.024]	[0.043]	[0.438]
F-stat	1370	1481	1826	1370	1481	1826
N	68112	68113	68113	68112	68113	68113
	Panel B: Contestable Schools					
1=attend FT in grade 1	0.053**	0.035	-0.006	0.036	0.053**	0.008
	(0.027)	(0.021)	(0.017)	(0.030)	(0.025)	(0.019)
	[0.044]	[0.101]	[0.728]	[0.226]	[0.033]	[0.652]
F-stat	3578	3886	4793	3578	3886	4793
N	199159	199159	199159	199159	199159	199159
	Panel C: FT Legacy Schools					
1=attend FT in grade 1	0.072**	0.049*	0.021	0.048	0.018	0.031
	(0.033)	(0.028)	(0.023)	(0.037)	(0.032)	(0.024)
	[0.026]	[0.076]	[0.361]	[0.198]	[0.569]	[0.197]
F-stat	2095	2279	2754	2095	2279	2754
N	191137	191137	191137	191137	191137	191137
Student-level controls	Yes	Yes	Yes	Yes	Yes	Yes
Class-level controls	Yes	Yes	Yes	Yes	Yes	Yes
School-level controls	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: According to the Panel, the sample is restricted to students who applied to schools that meet one of the following criteria: are located in a municipality with only one school (”Single-school Municipalities”, Panel A); received at least two applications to both FT and PT (”Contestable Schools”, Panel B), offered the full-time scheme in at least one of the five school years before application cohort 14-15 (”FT Legacy Schools”, Panel C). As in the main regressions, we restrict the analysis to the students who took both the Italian and the Math INVALSI tests in all grades and to the public schools in which the number of applications received is larger than or equal to the minimum number of students needed to form a class. We instrument the dummy for attending FT with the instrument defined in (1) and we control for a dummy for applying to FT. Controls are defined as in Table (5). All regressions include school-of-attendance province fixed effects. Robust standard errors clustered at the school-of-attendance level are shown in parenthesis and and p-values are shown in brackets. The F-statistic on the excluded instrument is reported. *** denotes significance at 1%, ** denotes significance at 5%, * denotes significance at 10%.

Table A1.5: **Mothers' Labor Supply: Robustness Checks (2SLS)**

	Labor Force Participation			**Employment**		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	Score 2	Score 5	Score 8	Score 2	Score 5	Score 8
			Panel A: Single-school Municipalities			
1=attend FT in grade 1	0.032**	0.037**	0.036**	0.014	0.023	0.040***
	(0.015)	(0.015)	(0.014)	(0.015)	(0.014)	(0.014)
	[0.036]	[0.013]	[0.011]	[0.338]	[0.10]	[0.005]
F-stat	1347	1452	1734	1347	1452	1734
N	52719	52719	52719	52719	52719	52719
			Panel B: Contestable Schools			
1=attend FT in grade 1	0.034***	0.034***	0.028***	0.024**	0.027***	0.031***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
	[0.001]	[0.001]	[0.005]	[0.018]	[0.008]	[0.002]
F-stat	2902	3119	3716	2902	3119	3716
N	137716	137716	137716	137716	137716	137716
			Panel C: FT Legacy Schools			
1=attend FT in grade 1	0.033**	0.040***	0.037***	0.021	0.030**	0.039***
	(0.014)	(0.013)	(0.013)	(0.013)	(0.013)	(0.014)
	[0.016]	[0.003]	[0.006]	[0.115]	[0.023]	[0.004]
F-stat	1684	1813	2118	1684	1813	2118
N	131627	131627	131627	131627	131627	131627
Student-level controls	Yes	Yes	Yes	Yes	Yes	Yes
School-level controls	Yes	Yes	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: According to the Panel, the sample is restricted to the mothers whose children applied to schools that meet one of the following criteria: are located in a municipality with only one school ("Single-school Municipalities", Panel A); received at least two applications to both FT and PT ("Contestable Schools", Panel B), offered the full-time scheme in at least one of the five school years before application cohort 14-15 ("FT Legacy Schools", Panel C). As in the main regressions, we restrict the analysis to the mothers whose employment status we observed when their children are in grades 2, 5, and 8 and to the public schools in which the number of applications received is larger than or equal to the minimum number of students needed to form a class. We instrument the dummy for attending FT with the instrument defined in (1) and we control for a dummy for applying to FT. Controls are defined as in Table (6). All regressions include school-of-attendance province fixed effects. Robust standard errors clustered at the school-of-attendance level are shown in parenthesis and and p-values are shown in brackets. The F-statistic on the excluded instrument is reported. *** denotes significance at 1%, ** denotes significance at 5%, * denotes significance at 10%.

Table A1.6: **Comparison between FT and PT applicants and between students who are not or are assigned to FT conditional on applying to FT**

	(1) Apply to FT (full sample)	(2) Attend PT if apply to FT	(3) Apply to FT (full sample)	(4) Attend PT if apply to FT
South and Islands	-0.303***	0.221***		
	(0.007)	(0.010)		
Mother employed	0.176***	-0.078***	0.126***	-0.089***
	(0.003)	(0.006)	(0.003)	(0.009)
Mother in labor force	0.178***	-0.075***	0.129***	-0.077***
	(0.003)	(0.006)	(0.003)	(0.008)
Father employed	0.007***	-0.011***	-0.009***	-0.002
	(0.001)	(0.003)	(0.002)	(0.004)
Father in labor force	-0.000	-0.001	0.000	-0.001
	(0.000)	(0.001)	(0.001)	(0.002)
Female	-0.008***	-0.002	-0.009***	-0.003
	(0.002)	(0.004)	(0.003)	(0.007)
Immigrant	0.052***	-0.020***	0.033***	0.029***
	(0.002)	(0.004)	(0.002)	(0.005)
Mother elementary or middle school	-0.058***	0.037***	-0.015***	0.028***
	(0.003)	(0.005)	(0.003)	(0.007)
Mother university	0.043***	-0.016***	0.015***	-0.013*
	(0.003)	(0.005)	(0.003)	(0.007)
Father elementary or middle school	-0.057***	0.023***	-0.017***	0.011
	(0.004)	(0.006)	(0.003)	(0.007)
Father university	0.031***	-0.015***	0.001	-0.003
	(0.003)	(0.004)	(0.002)	(0.006)
SES	0.119***	-0.071***	0.002	-0.035**
	(0.008)	(0.014)	(0.006)	(0.015)
School-of-application FE			Yes	Yes
N	338,862	124,094	338,862	124,094

Note: Coefficients in columns (1) and (3) are from full-sample regressions of each of the variables on the left-hand side on a dummy for applying to FT; coefficients in columns (2) and (4) are from regressions on the sample of FT applicants of each of the variables on the left-hand side on a dummy for attending PT. Standard errors clustered at the school-of-application level in parentheses. Statistics are Measured in Grade 2.

A2 Primary School Application Process

For the 2014-2015 school year, families had the option to enroll their children in primary schools exclusively online through the Ministry of Education and Merit web portal. They could either apply from home independently (around 70% chose this option) or visit the school they intended to apply to for assistance. In the latter case, the online application form was filled out using the school's technological infrastructure. Families were required to express their preferences for up to three schools, with the freedom to choose any institute in the country. Only the first preference was mandatory. If an applicant was rejected by their first preference school, the school itself forwarded the application to the second preference school. If the applicant was also rejected by the second preference school, the application was then forwarded to the third preference school. In instances of over-demand, the schools were required by law to define admission criteria, which should be made available to families in the application form. In practice, admission was almost always granted if the student resided near the school (Barbetta et al. (2022)).

Families were also required to rank, for each school they applied to, up to three preferred time schemes out of four options: 24 hours per week, 27 hours, up to 30 hours, or 40 hours (full-time). The schools were responsible for providing families with information regarding the services offered, such as the availability of a canteen and the time schemes offered in the past. The application form required families to provide demographic information about the child (name, gender, date of birth, citizenship, municipality of birth, municipality of residence, and social security code) and the person applying on their behalf (relationship with the child, name, gender, date of birth, citizenship, municipality of birth, municipality of residence, and social security code). Additionally, each school had the option to request more detailed information about the child and their family, which would be used as criteria in the acceptance of applications.

Applications for the 2014-2015 school year started on February 3rd and concluded at the end of the month. School principals were required to notify each family of their children's acceptance within a month after the application period. The assignment of students and teachers to classes took place during the summer, and parents were not involved in this process. They would only learn about the class composition and the teachers' names shortly before the beginning of the school year in September (Barbetta et al. (2021)).

A3 Sample Construction

The original (anonymized) MIM application data consists of 525,947 records. The student ID variable does not have any missing values or duplicates. These records are utilized to compute the number and type of applications received by each school. By means of an identifier provided by INVALSI, we were able to match the application data with INVALSI records for grade 2. Students who appear in the application data but not in the INVALSI records are those who were absent during all the INVALSI exam days or who ended up attending school in another country. Conversely, the students present in the INVALSI records but not in the application data are those who joined the Italian primary school after grade 1. After this match, we are left with 451,265 students.

Subsequently, using the INVALSI longitudinal identifier, we merged the INVALSI data from grade 2 with the data from grade 5 and grade 8. We restricted the analysis to a balanced panel, retaining only the observations of students present in all the three grades. This resulted in a total of 406,597 students. For our main analysis, we applied two restrictions to this sample. First, we excluded students who applied to schools that received fewer than 15 applications (10 if the school is located in a small, isolated village). This exclusion was necessary as our instrument is not defined in schools that receive fewer applications than the required lower bound for class formation. Consequently, we lost 19,677 students. Second, we considered only students who applied to public schools, as the law DPR 81/2009 does not apply to privately managed schools. By imposing this restriction, we lost 16,500 students.

For the regressions focusing on school achievement as the outcome, we retained only students for whom we had observed test scores in grades 2, 5, and 8 for both Math and Italian subjects. As a result, we excluded 31,544 students. After dropping observations for which at least one of the controls is missing, the final sample for the regressions on school achievement consists of 338,862 students. This represents around 67%, 66%, and 63% of the universe of INVALSI test takers in grade 2, 5, and 8, respectively. Likewise, for the regressions focusing on maternal labor supply as the outcome, we retained only students for whom we had observed their mothers' employment status in grades 2, 5, and 8. Consequently, we excluded 134,811 students, and the final sample consists of 237,457 mothers.

A4 Exerting exogenous variation in cases of excess demand: an algorithm for the principal's problem

Our instrument isolates a plausibly exogenous source of variation in the probability of attending an FT class, conditional on applying for it, in cases of excess demand. In particular it exploits contraints set by the law that regulates the class formation process in Italian public schools (DPR 81/2009), according to which:

(C.1) the number of classes in a school is a function of the total number of applications received;

(C.2) each class should have a minimum of 15 and a maximum of 27 students.[1] The minimum is lowered to 10 for schools located in small, isolated villages[2] and the maximum to 20 if there are disabled children in the class;

(C.3) a class can only offer one time scheme (i.e., it can be either fully FT or fully PT).

To isolate a plausibly exogenous source of variation in the probability of attending an FT class, conditional on applying for it, we develop an algorithm that figures out excess demand for any given application mix and randomly assigns students to their preferred scheme in case of excess demand. Specifically, the algorithm figures out excess demand for any application set $\{A_i\}$, where A_i is an indicator variable that equals 1 when pupil i applies to FT and 0 if she applies to PT. $|\{A_i\}|$ is the total number of applications and $\sum_{i=1}^{|\{A_i\}|} A_i$ is the total number of applications to FT. The steps are the following:

1. Given the total number of applications received and constraints (C.1)- (C.2), the algorithm figures out the total number of classes. Because activating a new class is costly and schools are very often resource-constrained, *we assume that the number of classes is kept as low as possible*, i.e. is equal to $\lceil \frac{|\{A_i\}|}{27} \rceil$, where $\lceil x \rceil$ is the ceiling function that maps x to the least integer greater than or equal to x.[3]

[1]When the total number of applications is 28 or 29, it is not possible to satisfy simultaneously the constraints on the maximum and the minimum class size. In these cases, we consider 13 as the minimum class size instead of 15, because the law allows a 10% deviation from the thresholds under special circumstances.

[2]Specifically, schools in mountainous villages, in small islands, and in villages where there are linguistic minorities.

[3]At this stage we are implicitly assuming that all students are admitted to the school and we only model the decision of how many classes need to be activated to accept all of them. In reality, admission is almost always granted if the student lives nearby the school (Barbetta et al. (2022)). In our sample, over 90% of the students apply to a school located in the municipality of residence. We always use 27 as the denominator in the ceiling function because we do not observe the disability status of students which would trigger a lower class size limit. Nevertheless, when comparing the number of total classes predicted by the algorithm under this assumption with the number of actual classes in a school, the correlation is high.

2. Once the total number of classes has been determined, the algorithm finds out the number of FT and PT classes and assigns pupils to classes, subject to constraints (C.2)-(C.3). In doing so, *we instruct the algorithm to minimize the number of pupils who are unhappy with the assignment*, given their preferred instructional scheme.[4]

3. Once the algorithm figures out how many students can not be assigned to the preferred scheme, *it assigns to every student who expressed a given preference the same probability of being assigned to the preferred scheme* (as if the school principal was flipping a coin).

Let us consider the four examples described in Table A1.7. In the first one, there are 18 applications, of which 5 to FT and 13 to PT. The algorithm activates one class because $\lceil \frac{18}{27} \rceil = 1$. Given that the strict majority of applications is PT, the class is PT, making all PT applicants satisfied and all FT applicants unsatisfied. In the second example, there are 35 applications (15 FT, 20 PT). The total number of classes is $2 = \lceil \frac{35}{27} \rceil$. In this case, there is no excess demand for either scheme: there can be a FT class with 15 students (which barely complies with the minimum class size) and a PT class of 20 students. The third example is presented in Section 4, and describes a case when there are 33 applications (8 FT, 25 PT) and the algorithm predicts the formation of two classes, one PT with 18 students and one FT with 15 students. The last example features 62 applications (28 FT, 34 PT). The total number of classes is 3. To minimize unhappy students, the algorithm activates 1 FT and 2 PT classes, but due to the constraint on the maximum class size, 1 of the students who applied to FT needs to be assigned to one of the two PT classes.[5] These few examples highlight how different application mixes can result in different allocation outcomes.

[4]Whenever two possible assignments lead to the same number of unhappy students, we instruct the algorithm to choose the assignment that minimizes the number of unhappy FT applicants.

[5]In this fourth scenario, 96% of students applying to the FT scheme are happy with their allocation.

Table A1.7: **Some example on how the algorithm works**

	FT Applicants	PT Applicants	FT Classes	PT Classes	FT Assigned	PT Assigned	Z_{is}^{FT}	Z_{is}^{PT}
Ex. 1	5	13	0	1	0	18	0	0
Ex. 2	15	20	1	1	15	20	1	0
Ex. 3	8	25	1	1	15	18	1	0.28
Ex. 4	28	34	1	2	27	35	0.96	0

Note: The table reports a number of examples to illustrate the functioning of the algorithm that computes the IV. In all the examples, we assume that the school is not located in a small, isolated village, so that the minimum class size is 15. Z_{is}^{FT} (Z_{is}^{PT}) denotes the instrument value if student i applies to FT (PT).