A Work Project presented as part of the requirements for the Award of a Master's degree in Economics from the Nova School of Business and Economics.

EDUCATION AND LABOUR MARKET EFFECTS OF KINDERGARTEN EXPANSION IN PORTUGAL

Rita Periquito da Fonseca, 25812

Work project carried out on the MSc in Economics, under the supervision of:

Susana Peralta

Pedro Freitas

January 4, 2021

Education and Labour Market Effects of Kindergarten Expansion in Portugal

Rita Periquito da Fonseca

January 2021

Abstract

This paper is the first thorough analysis of the impact of the massive expansion of kindergarten provision that occurred between 1985 and 2009 in Portugal. Using two large scale administrative datasets, we construct a measure of potential kindergarten exposure at municipal level and estimate its impact on education and the labour market. We find that the expansion in kindergarten supply led to an increase in educational attainment ranging between 0.1p.p and 3.6p.p in national standardised exams in Portuguese and Mathematics at different moments of the students' academic life, while the impacts on wages are modest and less significant.

Keywords: education, kindergarten, early childhood education, wages

JEL Classification: I2, J31

This work used infrastructure and resources funded by Fundação para a Ciência e a Tecnologia (UID/ECO/00124/2013, UID/ECO/00124/2019 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722 and Social Sciences DataLab, Project 22209) and POR Norte (Social Sciences DataLab, Project 22209)

1 Introduction

Pre-primary education, which encompasses both kindergarten (0 to 3 years old) and pre-school (3 to 6 years old), lays the foundation for individuals' success in school and in the labour market, as shown by Heckman, Stixrud, and Urzua (2006) and Cunha, Heckman, and Schennach (2010).[1] According to UNICEF, kindergartens are places where children develop their abilities, talents and skills from their earliest age, as it provides infants with an opportunity to learn and practice the essential social, emotional, problem-solving and study skills that are fundamental throughout the rest of their schooling. Duncan and Magnuson (2013) as well as Hart and Risley (2003) highlight the pivotal importance of experiences in the earliest years of life for establishing the brain architecture that will shape future cognitive, social, and emotional development, as well as physical and mental health, known as non-cognitive skills.

This rationale explains Heckman's prominent equation, which describes rates of return to public investment in human capital as rapidly diminishing with age. Heckman et al. (2006) contend that life cycle skill formation is a dynamic process in which early inputs strongly affect the productivity of later inputs, following a hierarchical procedure, as later attainments build on foundations which were settled earlier. Therefore, the highest rate of economic returns arises from the earliest investments in children, making early life an auspicious period for effective educational investments, as it sets the base for the development of future skills. Thus, building skill upon skill provides greater success to more children, higher productivity and lessens social spending for society. Consequently, this investment is also critical for countries' growth and the lack of it may be the cause of economic inequalities.

Schooling infrastructure is an important input of the education production function, as shown by Duflo (2001).

In Portugal, kindergartens are not part of the education system and are mostly provided by

[1] In Portugal, pre-school education is legally defined as the education of children within the 3 to 6 age range, and so has a different meaning compared to most European countries, where the term is normally applied to all settings of toddlers from birth until the beginning of compulsory schooling.

private non-profit (social), known as IPSS, and by private for-profit institutions (private).[2] They are regulated by the Ministry of Labour, Solidarity and Social Security (MTSSS). Portugal can be characterised as a kindergarten late adopter, and the growth in early childhood education and care (ECEC) from birth until 3 years of age has been relatively slow. According to OECD's 2000 Country Note about ECEC policy, the provision of services for children under the age of 3 in Portugal was not as well developed as those for children within the 3 to 6 age range. The participation rates in ECEC between the ages of 4 and 6 from 2001 to 2011 rose by 13,9% in Portugal, a higher increase than the EU-28 average of 7%.

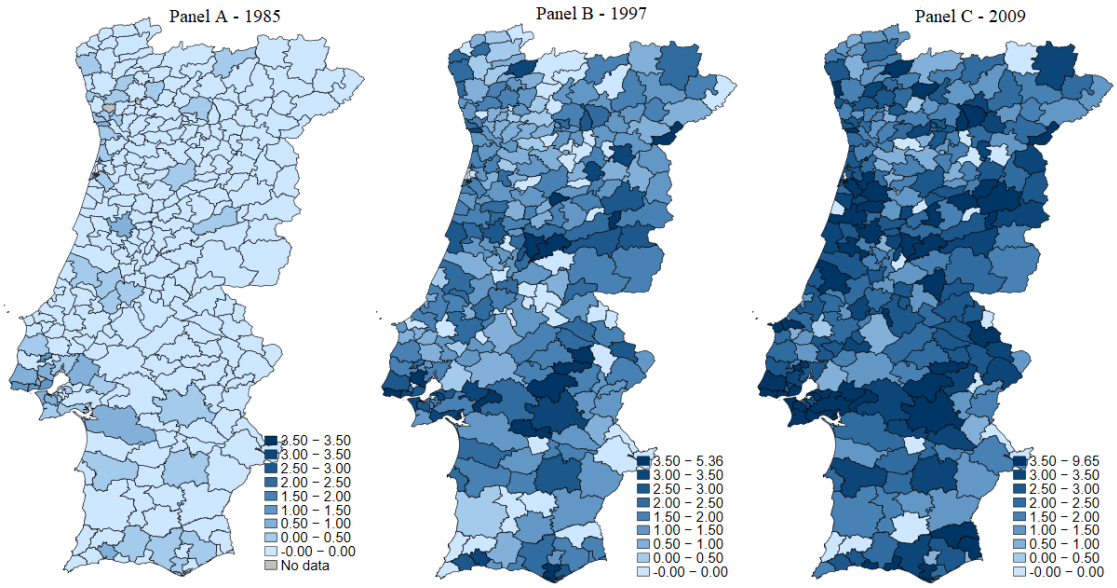
The provision of social and private kindergartens has been constantly expanding. Between 2006 and 2018, the coverage rate of kindergartens in the mainland territory has increased considerably (82%), having gone from 26,7% in 2006 to almost 48,4% in 2018, following the rise in kindergartens' vacancies.[3] However, we must highlight that this coverage rate is still very low. Figure 1 depicts the evolution in the number of kindergartens per 10,000 inhabitants. However, there are still some challenges. Besides the provision not being homogeneous throughout the country, with most services concentrated in the northern and central coastal areas and in major towns and cities (as can be seen in Figure 1) following the population distribution pattern, the provision for infants and toddlers has not kept pace with the needs of young families, particularly in urban areas (see Figure 4 in the Appendix). All these factors make the study of the impact of Portuguese kindergarten expansion on exam scores and wages an interesting analysis.

In our paper, we analyse the impact of the expansion in the supply of private and social kindergartens in Portugal, from 1985 until 2009, using two large scale education and labour administrative datasets (*MISI* and *Quadros de Pessoal*). We estimate the effects of the increase in kindergarten construction on education, through exam scores in the 4^{th}, 6^{th}, 9^{th} and 12^{th} grades, and on wages. We exploit a fixed effects and difference-in-differences strategy, where an individual's exposure to

[2]In 2018 there were 2,570 counted kindergartens in Portugals mainland territory, with 76,3% belonging to non-profit entities and the remaining 23,7% to profitable entities, according to GEP-MTSSS. Although this number is lower than what we found in our counting, it is pretty close, besides being reported by a different source.

[3]Source: GEP-MTSSS. The coverage rate corresponds to the percentage relation between the total number of available vacancies in kindergarten and the population that is between 0 and 3 years old.

Figure 1: Kindergarten Expansion in Portugal ©



Notes: This figure displays the evolution in the number of kindergartens per 10,000 inhabitants, in Portugals mainland territory, per municipality, from 1985, in Panel A, to 1997, in Panel B and to 2009, in Panel C, based on the kindergarten intensity measure we constructed.

kindergarten is determined by potential kindergarten exposure during the first 3 years of life.

The remainder of this paper is organised as follows: Section 2 provides an overview of the literature on the impact of early childhood education and school construction programs, Section 3 describes the institutional background, the data used and the measure of kindergarten exposure, Section 4 explains the methodology and identification strategy, Sections 5 and 6 report the results and, finally, Section 7 concludes, presents public policy implications and discusses possible avenues for future research.

2 Literature Review

We build upon and combine two distinct strands of literature. The first considers the short- and long-term effects of early childhood programs, mostly in the context of the United States.

In a seminal study, Heckman et al. (2006) define a child's development as a holistic, continuous and progressive process that initiates even before her birth. Elango, García, Heckman, and Hojman (2015)'s survey of the literature confirms that unpropitious early age environments, that

do not stimulate children and fail to nourish their cognitive and non-cognitive competencies, place them at a premature disadvantage that might persist over a lifetime. This poses a threat to social mobility, which is even stronger for disadvantaged toddlers. Hence, early interventions targeted toward disadvantaged children yield much higher returns than later interventions.

Although there is a widespread understanding that early childhood education generates short-term gains in test scores (see, e.g. Cascio (2009), Chetty et al. (2011), Elango et al. (2015)), its ability to significantly improve educational and economic success in the long run has been questioned, as Barnett (1992) outlines. Evidence on the long-term impacts of pre-school education programs is scarce due to the paucity of data linking childhood education and adult outcomes.

A growing literature has established the considerable impact of early childhood environments on later life outcomes. Heckman, Moon, Pinto, Savelyev, and Yavitz (2010) estimate positive and significant returns to the Perry Preschool Program for both males and females. Chetty et al. (2011) find that kindergarten test scores are highly correlated with outcomes such as earnings at age 27, college attendance, home ownership, and retirement savings. Ludwig and Miller (2007) find suggestive evidence for a positive effect of Head Start on educational attainment, while Carneiro and Ginja (2014) find that participation in the program reduces the incidence of behavioural and health problems. Despite the primacy attributed to cognitive ability in producing successful lifetime outcomes by the literature, various studies point to the prominence of non-cognitive abilities, such as Heckman et al. (2006) and Cunha et al. (2010).

The second strand of literature examines the impact of school infrastructure investments on educational attainment and earnings.

Duflo (2001) exploits a large-scale school construction program launched by the Indonesian government between 1973 and 1978 to evaluate the effect of building schools on education and earnings in Indonesia. Through the combination of differences in the number of schools constructed across regions with differences across cohorts, induced by the timing of the program, the results point to an average increase of 0.12 to 0.19 years of education, as well as a 1.5% to 2.7% increase in wages for each primary school constructed per 1,000 children. Estimates of economic returns

to education range from 6.8% to 10.6%. Following Duflo's seminal paper, Akresh, Halim, and Kleemans (2018) analyse the same school construction program in Indonesia to study the long-term and intergenerational effects of additional schooling as a child. Employing a difference-in-differences identification strategy to exploit variation across birth cohorts and regions in the number of schools built, they find that individuals exposed to the program attain more education, households with parents exposed to the program have improved living standards and that education benefits are transmitted to the next generation.

Neilson and Zimmerman (2014) present new evidence on the causal effect of elementary and middle school construction projects on academic achievement, school enrolment and housing prices using panel data on student test scores and neighbourhoods of residence. They find that, six years after the buildings' occupation, reading scores increase by 0.15 standard deviations relative to the year before building occupancy, home prices increase in affected neighbourhoods by roughly 10%, and there is an increase in public school enrolment. They find no impact on Mathematics scores.

Lastly, using a random sample of 4,600+ individuals, Palma and Reis (2018) reveal that a primary school massification program, implemented by the nondemocratic regime in Portugal during the first half of the 20th century, led to a remarkable reduction in illiteracy. They show that 20-year-old males were about 50% more likely to end up literate under a nondemocratic regime than under a more democratic one. These results are robust to controlling for an array of factors, including economic growth, disease environment, and regional fixed effects. Hence, they argue for a political economy and cultural explanation for the relative success of the authoritarian regime in promoting elementary education.

3 Data and Institutional Background

In this Section we outline kindergartens' institutional background, describe the data sources we used and the measure of kindergarten exposure we constructed and, finally, present the data descriptive statistics.

3.1 Institutional Setting

The first legal requirement concerning kindergartens in Portugal was enacted in 1983 by *Decreto Regulamentar* nº 69/83 to regulate *Decreto-Lei* nº 350/81, which characterised kindergartens as establishments aimed at receiving children aged between 3 months and 3 years, with the purpose of providing adequate conditions for their development. In 1984, *Despacho Normativo* nº 131/84 regulated for-profit kindergartens' facility conditions and functioning. Subsequently, *Decreto-Lei* nº 30/89, which was not applicable to establishments of private institutions of social support, revoked precedent decrees, and *Decreto-Lei* nº 99/89 approved the regulatory standards of for-profit kindergartens' facility conditions, functioning and monitoring. In 2011, *Portaria* nº 262/2011 revoked the former regulatory standard, re-established a legal framework that guaranteed harmonised implementation of kindergartens' guiding standards independently of the nature of their institutional legal counsel, and recovered kindergartens' educational role. We exploit this law changes to proxy for the number of children exposed to kindergarten.

3.2 Data Sources

We use two large scale administrative datasets. The first, *MISI* (Integrated Information Mission Systems), is a very detailed education administrative database collected by the Ministry of Education containing variables that characterise students' academic performance in standardised exams and academic path information on pre–school, elementary and secondary education in Portuguese public schools, as well as on their social-economic background since the academic year of 2006/2007, such as their residency, school, grades (which are reported on a scale of 1 to 5) and parents' education.[4] We use the dataset from the academic years of 2007/08 until 2017/18 for the 4^{th}, 6^{th}, 9^{th} and 12^{th} grades. However, we must highlight that the number of students observed in the 12^{th} grade is lower because we only observe students on the upper-secondary education, and so we

[4]Where a score of 2 corresponds to the range between 50% and 70%, score 3 between 70% and 80%, and score 5 to a percentage higher or equal to 90%. However, in the 12^{th} grade scores range between 0 and 20. For that reason, in order for the scores to be comparable across all grades, we transformed 12^{th} grade scores. Thus, in the 12^{th} grade, achieving a score higher than 2, 3 and 4 corresponds to having a score higher than 9.8, 13.8 and 17.8, respectively.

are looking at a selected sample. The second, *Quadros de Pessoal* (Personnel Records, henceforth referred to as QP), is an administrative linked employer-employee longitudinal dataset collected by the MTSSS, including information on the Portuguese labour force (excluding civil servants and independent workers), such as their gender, age, education, skills, occupation, monthly wages and hours worked, as well as on firms and establishments, such as size, location, turnover, sector of activity, and employment. We use this database from 1985 until 2009, except for the years of 1990 and 2001, since they are not available. Our analysis is restricted to Portugal's mainland territory.

We further collect registers on the number of fatalities, live births and population from INE (Statistics Portugal) from 1988 until 2012.

3.3 Measure of Kindergarten Exposure

There is no readily available variable with the number of kindergartens in Portugal. One of the main contributions of this paper is that we constructed this variable for the first time for the Portuguese context. In order to do so, we use the linked employer-employee database (QP), and proceed as follows. We consider an establishment as a kindergarten by employing a two-step procedure. First, we keep the establishments identified by the Economic Activities Classification (CAE) codes presented in Tables 26 to 28 in the Appendix (for instance, early childhood education and basic education establishments).[5] , [6] Then, from those, we only retained the establishments for which we have at least one employed kindergarten teacher or assistant/guardian, identified by the Portuguese Professions Classification (CPP) codes given in Tables 29 to 31 in the Appendix.[7] Figure 3 in the Appendix graphs the evolution in the number of kindergartens using this method, clearly displaying an increasing trend. The periods between 1985-91 and 2003-09 are more volatile, while

[5] CAE constitutes a code attributed to establishments and firms according to their economic activity.

[6] Until 2007 it was only possible for firms and establishments to choose one main CAE code, although they could be performing multiple economic activities, belonging in fact to many CAEs. That is the reason why we consider codes like 85323 in CAE-Rev.2 and 2.1 in our first-step as, for instance, an establishment could simultaneously work as a centre of social action towards the elderly and a kindergarten. In order to narrow down and improve our measure of the number of kindergartens per municipality over time, we do the second step of identifying kindergarten teachers and assistants, keeping only the establishments for which we have at least one of the two.

[7] CPP corresponds to the code assigned to each worker according to her job.

the period between 1991-2001 clearly reveals a fairly constant rising trend. Furthermore, we computed an estimate of kindergartens' dimension, given by the number of children per kindergarten, based on the legislation discussed in the Institutional Setting Section.[8] Based on this estimate, we graphed the evolution of the proxied number of children attending kindergarten at the national level, depicted in Figure 4 in the Appendix.[9]

We compute two measures of kindergarten intensity: the first and preferred, and the alternative, for the robustness analysis. These measures allow us to estimate individuals' potential exposure to kindergarten, as we do not know if individuals were directly enrolled in a kindergarten program. Hence, estimated impacts can come from the direct impact of attending a kindergarten, from higher peer effects from peers that attended kindergarten, or by other spill over effects from the provision of these infrastructures. As such, we provide estimates of the intent-to-treat effect. Year of birth and proxy of municipality (which we will refer to as region) of residence at birth jointly determine an individual's exposure to kindergarten. Both these factors, combined with the counting of kindergartens per municipality over time, are used to construct the measures of kindergarten intensity.

The first consists on the average number of kindergartens available per 10,000 inhabitants when an individual was between 0 and 3 years old on her proxied region of residence. We compute this measure in the following way: K_{icr}=$\sum_{t=0}^{3} K_{rt}/4$, where K_{rt} is the counting of kindergartens in region *r* on year *t* per 10,000 inhabitants (total population), with $t = 0$ corresponding to the individual's birth year, and K_{icr} is the kindergarten intensity for individual *i* of cohort *c* in region *r*. As we do not possess information on the kindergarten counting for the years of 1990 and 2001, we use the subsequent available year instead (for instance, 2002 for the case of 2001).

We proxy individual's region of residence in early childhood by their region of residence during school years, and region of work in adult years. This is obviously a stronger assumption in the latter case, especially considering the concentration of individuals working on metropolitan areas, so we

[8]We build an estimate of kindergartens dimension based on Norms VI and XVIII of Despacho Normativo nž 99/89, which was in force between 1989 and 2011, and plainly stated that an assistant was needed per group of 10 children and a teacher per group of 15 children.

[9]The number of children in the graph in 2009 roughly correspond to half of those indicated by the age pyramid of 2009 for Portugal, confirming the existent challenge in the provision of kindergartens, particularly in urban areas.

conducted some robustness analysis that partially deal with this problem in Section 4.

To enhance the robustness of our results, we compute an alternative measure of kindergarten intensity, which comprises the number of kindergartens available on the individual's proxy region on her year of birth per 10,000 inhabitants, and whose results are also shown in the robustness Tables in the Appendix.

We will use the expression "exposure to kindergarten" to refer to the increase of one kindergarten per 10,000 inhabitants when the individual was between 0 and 3 years old in order to ease the readability of the results.

3.4 Descriptive Statistics - MISI and QP

In the Table below we can find the descriptive statistics of the data we used.

Table 1: Descriptive Statistics

MISI (education) data

		Grade			
Variables	Total	4	6	9	12
Female	50,45%	48,59%	48,95%	51,83%	54,85%
Mother higher education	16,53%	15,55%	14,40%	17%	22,86%
Father higher education	11,20%	9,88%	9,50%	11,64%	17,35%
SASE level A	16,69%	15,52%	18,95%	17,30%	12,24%
SASE level B	19,71%	19,51%	24,51%	18,92%	10,64%
Computer access	67,20%	57,57%	65,91%	75,17%	71,43%
Internet access	55,62%	44,38%	52,14%	64,76%	65,88%
Portuguese mother	93,41%	92,95%	92,67%	94,33%	93,74%
Portuguese father	92,38%	92,04%	91,51%	93,20%	93,03%
Mother unemployed	9,38%	10,18%	9.86%	9,30%	6,51%
Father unemployed	5,32%	5,45%	5,50%	5,53%	4%
Portuguese score >2	-	41,47%	40,81%	35,54%	64,09%
Portuguese score >3	-	19,87%	16,03%	12,21%	20,52%
Portuguese score >4	-	2,99%	1,90%	1,32%	1,17%
Mathematics score >2	-	37,86%	32,66%	25,64%	64,99%
Mathematics score >3	-	19,45%	13,70%	12,80%	38,36%
Mathematics score >4	-	4,92%	3,01%	3,11%	12,08%

QP (labour market) data

Variables	Percentage/mean	St dev	Min	Max
Female	46,26%	0.499		
Secondary education	55,65%	0.497		
Higher education	11,03%	0.313		
Permanent contract	48,38%	0.500		
Tenure	1.667	2.101	0	10
Logarithm hourly deflated wage	1.354	0.438	0	8.670
Kindergarten Exposure				
MISI	2.008	1.103	0	7.192
QP	1.070	0.772	0	4.537
Diff-in-diff (first approach)	2.077	1.072	0	8.105
Diff-in-diff (second approach)	2.159	1.071	0	8.105

All the variables in the table (except for tenure, logarithm of hourly deflated wages and kindergarten exposure) are dummy variables. For that reason, we believed it would be more insightful for

us to provide a percentage statistic for those variables and the mean for the others. The first part of the table reports descriptive statistics (percentage of students pertaining to each characteristic) for the variables found on the *MISI* database, for each grade. The second part of the table reports the mean or percentage (in the case of dummy variables), standard deviation, minimum and maximum (except for dummies) for the variables in the QP dataset, as well as for the kindergarten exposure measure we constructed for the various datasets. Table 2 concludes by presenting the descriptive statistics of the difference-in-differences strategy for both the first and second approaches, for the treated and comparison cohorts.

Table 2: Descriptive Statistics - Difference-in-Differences

Approach	First Approach		Second Approach	
Cohort	Treated	Comparison	Treated	Comparison
Cohort Composition	33,79%	66,21%	27,31%	72,69%
Female	45,32%	46,05%	48,75%	49,18%
Age	21,09	20,71	23,31	32,81
Upper-secondary education	45,17%	79,02%	53%	56,18%
Upper-secondary education (only 12^{th} grade)	39,27%	74,29%	39,86%	30,08%
Higher education	5,90%	4,73%	13,14%	26,11%
Higher education (if secondary=1)	5,96%	6,29%	32,86%	26,46%
Higher education (if secundary2=1)	2,40%	3,94%	7,28%	21,91%
Permanent contract	41,49%	12,12%	46,32%	69,69%
Tenure	1.110	1.632	2.044	4.304
Logarithm hourly deflated wage	1.311	0.798	1.458	1.731

4 Methodology and Identification Strategy

We exploit the panel structure of our data to estimate fixed effects regressions. We then provide causal estimates through a difference-in-differences strategy following Duflo (2001). All the regressions are estimated using heteroskedasticity robust standard errors.

4.1 Fixed Effects Estimation

Starting by the fixed effects estimation, we estimate the following equation for education:

$$g_{ispt} = \theta + \gamma K_{is} + \delta_s + \alpha_p + \eta_t + \beta X_i + \epsilon_{ispt} \tag{1}$$

When measuring education outcomes, we employ a linear probability model, where g_{ispt} represents the probability of student *i*, in school *s*, with teacher *p*, in academic year *t*, obtaining a score

higher than 2, 3 or 4, on a scale of 1 to 5, on the Portuguese and Mathematics national exams, thus measuring a short- and medium-term outcome. K_{is} denotes kindergarten intensity for individual *i*, where *s* stands for the region where the school is located, which we assume to be the same as the individual's birth region. X_i is a vector of covariates, θ is the constant, and ϵ_{ispt} the error term. The effect of kindergarten exposure can be biased by different selection sources. The first possible source of bias is related to the placement of kindergartens, which is expected to be correlated with the endogenous characteristics of the municipality and region. Secondly, students who were previously exposed to a higher intensity of kindergartens may be placed in better schools and/or with better teachers later.

To tackle these selection bias issues, we include fixed effects at several levels: η_t represents academic year fixed effects to control, for instance, for specific year exam difficulty, δ_s denotes school fixed effects and α_p are teacher fixed effects, both aiming to control for non-random allocation of students to schools and teachers. We sometimes replace school fixed effects with less stringent municipality or NUTS (Nomenclature of Territorial Units for Statistical Purposes) III fixed effects.[10]

The covariates used in the education estimation, X_i, include: student's gender, if the student's mother and father hold an academic qualification higher than secondary school, if the student benefits from the means-tested programme SASE (Schooling Social Action System) levels A and B, if the student has a computer and access to the internet at home, if the student's mother and father are Portuguese, and if the student's mother and father are unemployed.

We estimate the following equation for wages:

$$w_{ift} = \theta + \gamma K_{if} + \delta_f + \eta_t + \beta X_{if} + \epsilon_{ift} \tag{2}$$

In the earnings estimation, w_{ift} stands for the logarithm of the hourly deflated wage of worker *i*, working at establishment *f* (also denoting the region of the establishment), at time *t*, a long-term outcome. Here, δ_f are establishment fixed effects, to control for firm and industry heterogeneity,

[10] We use 23 of its categories in mainland Portugal.

and η_t represents tenure fixed effects, to control for the return to seniority in wages. θ is the constant and ϵ_{ift} is the error term. The covariates used, X_{if}, are the following dummies: worker's gender, if the worker has upper-secondary education, if the worker has higher education, and if the worker has a permanent contract.[11]

We exploit the QP Workers' database as of 2002 for the panel estimation. Given that our measure of kindergarten exposure is computed as of 1985, the first cohort of exposed individuals was born in 1985, and therefore shows in the QP dataset as of 2002, when they are 17. For younger cohorts we consider the labour market impact as of the age of 16.[12] We exploit the *MISI* dataset between the academic years of 2007/08 and 2017/18. We observe students when they first write exams in the 4th, 6th, 9th and 12th grades. The timelines can be found in Figure 2 and a more exhaustive timeline for *MISI* per grade is in Figure 5 in the Appendix.

A relevant issue on both estimations regards the assumption on the place of residence observed and the one when individuals were between 0 and 3 years old. While in the education estimation we use the individuals' region of residence in the academic year of the exams, in the earnings estimation we use the region of the establishment where they work. Although scores attained in the exams are a short-term outcome, students' region of residence on the years they took exams has a higher probability of correspondence to their true region of residence when they could attend kindergarten. On the contrary, even though the outcome studied for workers is long-term, the proxy used is less precise, especially considering the concentration of individuals working on metropolitan areas, such as Grande Lisboa and Grande Porto.

We overcome the lack of information on individual's birthplace with two complementary strategies. Firstly, we eliminate the metropolitan areas from our sample, where both intermunicipal and inward mobility are more likely, thus mitigating confounding effects. Secondly, we compute a mobility variable using the population estimates from Statistics Portugal.[13] In the fixed effects

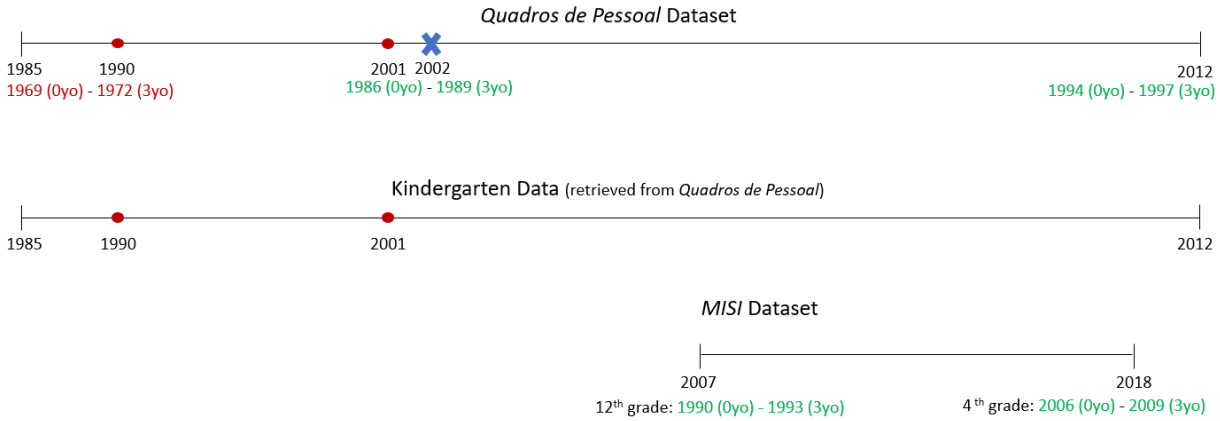
[11] Upper-secondary education includes post-secondary and lower undergraduate education.

[12] Individuals born in 1985 are 16 years old in 2001, but as we do not have data for that year, we start our analysis in 2002, when they are 17.

[13] We compute the mobility variable in the following way:

$$Mobility_t = Population_{t+1} - Population_t - Births_t + Fatalities_t \quad (3)$$

Figure 2: Datasets Timelines



Notes: We assume that individuals only enter the QP database when they are at least 16 years old (yo). Therefore, for instance, in 1985, the youngest individuals in the database were born in 1969 and were 3yo in 1972. As we only have data on the number of kindergartens from 1985 onwards, we are only interested in analysing individuals who were born from that year onwards. Hence, we only study individuals in the QP dataset from the year of 2002 onwards, as we start capturing individuals who were born in 1985 in 2001 but, since the dataset does not have data for that year, we use the following year, 2002. In what concerns MISI, the birth years were computed assuming that students taking 4^{th} grade exams were 9 years old, 6^{th} 11 years old, 9^{th} 14 years old and 12^{th} 17 years old. Hence, the first cohort of students present in the MISI dataset is that of students who were born in 1990, taking their 12^{th} grade exam in 2007. The last cohort is that of students born in 2006, and so that can attend kindergarten until 2009 (when they are 3 years old), taking their 4^{th} grade exam in 2018. The complete figure exhausting the timeline for each grade can be found in Figure 5 of the Appendix.

regressions, we include mobility as a control.

4.2 Difference-in-Differences Approach

In order to provide causal estimates, we follow Duflo (2001) and implement a difference-in-differences strategy, using two alternative approaches. The difference between the two approaches is the time at which we measure the outcomes of the comparison cohort. In the first, individuals on the comparison cohort have the same age profile as individuals in the treated cohort, but are observed in a different time period, while in the second we observe both in the same time interval, and as a consequence they have different age profiles.

We exploit the substantial kindergarten expansion that occurred starting in 1985. Therefore, the treated cohort was born between 1986 and 1996, while the comparison cohort was born between 1974 and 1984. The treated cohort shows up in the QP dataset as of 2002 and the comparison cohort

as of 1989.[14] In the second approach, we compare the outcomes of both cohorts in 2012. Therefore, the comparison cohort is older than the treated one, but the economic cycle is the same.

We estimate the impact of kindergarten construction (using a different kindergarten measure) on the probability of completing upper-secondary education (including higher education), upper-secondary education (excluding higher education), higher education, and higher education compared to having completed upper-secondary education, while for the former specifications we compare to not having completed the level of education in question.[15]

Hence, the equation we run for schooling and wages is the following:

$$y_{icr} = \theta + \delta_r + \eta_c + (K_{icr} * T_c)\gamma + (X_{icr} * T_c) + \epsilon_{icr} \tag{4}$$

y_{icr} is the probability of individual *i*, belonging to cohort *c*, with proxied region of birth *r*, completing upper-secondary or higher education in the schooling equation, and the logarithm of the hourly deflated wage in the wage equation. γ_r represents proxy of birth region fixed effects, η_c denotes cohort of birth fixed effects, K_{icr} is the intensity of the program for individual *i* belonging to cohort *c* on the proxied region of birth *r*, T_c is a dummy indicating whether the individual belongs to the new cohort (individuals born between 1986 and 1996) and X_{icr} is a vector of control variables. θ is the constant and ϵ_{icr} is the error term.

The identification assumption that the evolution of education and wages across cohorts would not have consistently differed in the absence of kindergarten expansion is enough to estimate the impact of kindergarten intensity. After controlling for proxied region of birth and cohort of birth year effects, we further exploit the causal impact of kindergarten construction on estimated education with both approaches, to create instrumental variable estimates of the effect of education on wages. The interaction of being in the treated cohort and kindergarten intensity in the proxied region of birth are plausible exogenous variables, and are thus used as instruments in the wage equation. However,

[14] The treated cohort shows up in the QP dataset as of 2001, but given that we do not have data available for that year, we assume it starts showing in 2002.

[15] For the estimations of higher education we need to consider only individuals with 22 years of age or older, otherwise it would be impossible for them to have completed higher education.

this relies on the strong assumption that kindergarten expansion had no impact on wages other than by increasing educational attainment.

The estimates obtained from equation (2) represent the first stage of a two-stage least-squares (2SLS) estimation of the impact of education on wages. The following equation characterises the causal effect of education on wages, representing the second stage:

$$\hat{Y}_{icr} = \theta + \delta_r + \eta_c + \beta y_{icr} + \alpha X_{icr} + \varphi_{icr} \tag{5}$$

where δ_r and η_c denote proxy of birth region and cohort of birth fixed effects, respectively, X_{icr} is a vector of controls, θ is the constant and φ_{icr} is the error term.

To overcome the lack of information on individual's birthplace, we firstly eliminate the metropolitan areas from our sample, like we did in the panel estimation. Secondly, using the mobility variable, we exclude regions where average mobility was higher than the 90th percentile.

5 Fixed Effects Results

This Section presents the results of estimating equations (1) and (2), by exploiting the panel structure of our data to account for unobservable characteristics. Tables 3 and 4 show the results of the panel regressions using our preferred specifications for exam scores, and Tables 13 to 20 in the Appendix exhibit the robustness results.

The panel estimation encompasses six specifications for education and five for the labour market. We start with the simplest specification, where we regress the outcome variable on our variable of interest, kindergarten intensity (K_{is}), and on the covariates. In the second, which constitutes the baseline specification, we add academic year fixed effects for education and tenure fixed effects for earnings. To control for individuals' and kindergarten non-random allocation between NUTS III and municipalities, we rerun the third and fourth regressions with NUTS III and municipality fixed effects, respectively, added to the baseline, shown in Tables 13 to 20 in the Appendix. We conduct the fifth regression using school fixed effects for education and establishment fixed effects for earnings. Finally, the sixth specification uses both school and teacher fixed effects, our preferred

specification, shown in Tables 3 and 4.[16]

Table 3: Fixed Effects Estimation Results - Portuguese Exam

	grade 4			grade 6		
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	score >2	score >3	score >4	score >2	score >3	score >4
kindergarten intensity	0.018***	0.020***	0.004***	0.032***	0.028***	0.004***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)
Observations	533,853	533,853	533,853	452,486	452,486	452,486
R-squared	0.180	0.149	0.061	0.160	0.167	0.076
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Academic Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes
Teacher FE				Yes	Yes	Yes
	grade 9			grade 12		
	(7)	(8)	(9)	(10)	(11)	(12)
VARIABLES	score >2	score >3	score >4	score >2	score >3	score >4
kindergarten intensity	0.035***	0.019***	0.002***	0.024***	0.015***	0.001***
	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.000)
Observations	565,172	565,172	565,172	292,970	292,970	292,970
R-squared	0.158	0.167	0.064	0.145	0.106	0.027
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Academic Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes
Teacher FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the fixed effects estimation results from the baseline specifications for the Portuguese exam. It presents the results for having a score higher than 2, 3, and 4 for all grades. For the 4th grade the specification used is school fixed effects, while for the remaining grades it is teacher fixed effects. The results for this estimation including the covariates can be found in Table 11 of the Appendix, where other specifications can also be found with other combinations of fixed effects.
Robust standard errors in parentheses, * p<0.01, ** p<0.05, * p<0.1

The impact of exposure to kindergarten ranges between 0.1p.p and 3.6p.p across all grades and scores in the specifications with school and teacher fixed effects. In general, its effect is larger for lower scores, being similar for Portuguese and Mathematics. The higher impacts are found for the 6th and 9th grades. When we consider all the specifications, generally, the lower bound of the effect of kindergarten exposure corresponds to the specification including NUTS III fixed effects, and the upper bound by the specification with municipality fixed effects, as can be seen in the robustness Tables in the Appendix.

For instance, in the 9th grade, an additional kindergarten per 10,000 inhabitants increases the probability of attaining a score higher than 2 in the Portuguese exam between 0.8 and 4.6p.p, while

[16]Except for the 4th grade because we are not able to identify teachers for this school year.

Table 4: Fixed Effects Estimation Results - Mathematics Exam

	grade 4			grade 6		
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	score >2	score >3	score >4	score >2	score >3	score >4
kindergarten intensity	0.017***	0.014***	0.006***	0.035***	0.023***	0.006***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)
Observations	532,383	532,383	532,383	460,119	460,119	460,119
R-squared	0.190	0.154	0.105	0.223	0.178	0.094
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Academic Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes
Teacher FE				Yes	Yes	Yes
	grade 9			grade 12		
	(7)	(8)	(9)	(10)	(11)	(12)
VARIABLES	score >2	score >3	score >4	score >2	score >3	score >4
kindergarten intensity	0.036***	0.021***	0.006***	0.026***	0.017***	0.007***
	(0.002)	(0.001)	(0.001)	(0.003)	(0.003)	(0.002)
Observations	570,841	570,841	570,841	138,718	138,718	138,718
R-squared	0.192	0.182	0.096	0.155	0.145	0.088
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Academic Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes
Teacher FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the fixed effects estimation results from the baseline specifications for the Mathematics exam. It presents the results for having a score higher than 2, 3, and 4 for all grades. For the 4th grade the specification used is school fixed effects, while for the remaining grades it is teacher fixed effects. The results for this estimation including the covariates can be found in Table 12 of the Appendix, where other specifications can also be found with other combinations of fixed effects.
Robust standard errors in parentheses, * p<0.01, ** p<0.05, * p<0.1

for a score higher than 3 this effect is narrower, ranging between 0.5 and 2.8p.p, being lower for a score higher than 4, where the impact varies between 0 and 0.3p.p (considering all specifications).

The impact of kindergarten exposure on the probability of attaining a score higher than 2 ranges between 1.8 and 3.5p.p, for Portuguese, and between 1.7 and 3.6p.p, for Mathematics, across all grades. The greatest impact is found for the 4th grade (3.6p.p), while the lowest is found for the 9th grade (1.7p.p), both in the Mathematics exam. We observe an increasing tendency of the effect for higher grades until the 9th grade, decreasing again for the 12th grade.

The effect of kindergarten exposure on the probability of having a score higher than 3 ranges between 1.4 and 2.8p.p, considering all grades. The highest impact is found for the 6th grade in the Portuguese exam, the grade until when the impact is increasing, while the lowest impact is ascribed to the 4th grade Mathematics exam. As of the 6th grade, the impact is decreasing.

Lastly, the impact on achieving a score higher than 4 ranges between 0.1 and 0.7p.p across all scores, bearing lower significance. The highest impact is found in the Mathematics exam, where we observe a decreasing tendency of the effect for later grades, while the lowest impact is ascribed to Portuguese exam, both in the 12th grade. The percentage impact considering the students in each score threshold can be found in Table 10 in the Appendix, reinforcing that the larger impacts are found for the 6th and 9th grades.

Figures 7 to 14 display the heterogeneity of the impact of kindergarten exposure according to gender, mother having completed higher education and being part of the means-tested programme SASE. Generally, we find a higher impact for females across all grades and scores. We also observe that the impact for students whose mothers have completed higher education decreases for higher scores, contrarily to the effect on students having SASE, for whom the impact increases for higher scores.

Tables 13 to 20 in the Appendix exhibit the robustness results. Our results are robust to the alternative measure of kindergarten exposure (columns (3), (8) and (13) for the 4th grade, columns (4), (10) and (16) for other grades), as well as to the two specifications controlling for mobility (columns (4), (5), (9), (10), (14) and (15) for the 4th grade, and columns (5), (6), (11), (12), (17) and (18) for the remaining grades). When we withdraw the metropolitan areas of Grande Lisboa and Grande Porto from the estimation, we find higher effects of kindergarten exposure, which makes sense as the effects in metropolitan areas may be diluted by the confounding effects discussed in Section 4.

We now turn to labour market effects, where the results of the panel regressions can be found in Table 5. Being exposed to kindergarten leads to an increase between 0.4% (for the specification without metropolitan areas) and 4.1% (with NUTS III fixed effects) in hourly deflated wages. These results are also robust to the alternative measure of kindergarten exposure (column (4)) and to the two specifications controlling for mobility (columns (5) and (6)). By using the number of kindergartens on the year of birth the impact is slightly lower, as well as by including mobility as a control (0.6% for both specifications).

Table 5: Fixed Effects Estimation Results - Wages

VARIABLES	(1) wages	(2) wages	(3) wages	(4) wages	(5) wages	(6) wages
kindergarten intensity	0.041***	0.039***	0.008***	0.006***	0.004***	0.006***
	(0.001)	(0.001)	(0.001)	(0.0004)	(0.001)	(0.001)
Observations	1,488,573	1,488,573	1,488,573	1,488,573	934,392	1,488,275
R-squared	0.166	0.180	0.574	0.574	0.585	0.574
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Seniority FE	Yes	Yes	Yes	Yes	Yes	Yes
Establishment FE			Yes	Yes	Yes	Yes
Municipality FE		Yes				
NUTS FE	Yes					

Notes: This table reports the fixed effects estimation results for wages. The first column corresponds to the specification including NUTS III fixed effects, the second of including municipality fixed effects, the third of including establishment fixed effects, the fourth of using the alternative measure of kindergarten intensity (number of kindergartens on individuals' birth year), the fifth of withdrawing the metropolitan areas, and the sixth of including mobility as a control. The table including the covariates can be found in the Appendix in Table 21.
Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

6 Difference-in-Differences Results

We now present our estimates of equation (3) with birth year fixed effects, then adding NUTS III fixed effects, and finally adding the following set of interactions terms with the new cohort: municipal mean of the logarithm of deflated hourly wages, municipal mean tenure, municipal labour force, and gender as a covariate. For wages, we conduct one additional regression including establishment fixed effects.

Our estimates suggest that kindergarten exposure increases the probability of completing upper-secondary and higher education for all specifications using both approaches. Starting with the first approach (for which the results are shown in Table 6 and the robustness checks in Table 22), we estimate that the treated cohort has an increased probability between 0.79p.p and 1.65p.p of completing upper-secondary and higher education compared to not completing, and higher education compared to upper-secondary education, respectively. As it is expected, kindergarten exposure has a higher impact on the probability of completing upper-secondary education. By conducting the second approach (for which the results are displayed in Table 7 and the robustness checks in Table 23), we find that the treated cohort has an increased probability between 0.4p.p and 1.65p.p of com-

Table 6: Difference-in-Differences Results - First Approach

VARIABLES	(1) Secondary	(2) Secondary2	(3) High education	(4) High education2	(5) Wages
interaction	0.017***	0.015***	0.008***	0.010***	-0.008***
	(0.0003)	(0.0004)	(0.0004)	(0.001)	(0.0004)
Observations	3,924,693	3,723,036	1,599,094	1,127,728	3,606,678
R-squared	0.158	0.166	0.057	0.090	0.714
Birth FE	Yes	Yes	Yes	Yes	Yes
NUTS FE	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the difference-in-differences estimation results using the first approach, where individuals on the comparison cohort have the same age profile as individuals in the treated cohort, but are observed in a different time period. Secondary corresponds to upper-secondary education (including higher education); Secondary2 corresponds to upper-secondary education (excluding higher education); High education corresponds to higher education; and High education2 corresponds to higher education compared to having completed upper-secondary education, while the former specifications are compared to not having completed the level of education in question. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

pleting upper-secondary and higher education compared to not completing, and higher education compared to upper-secondary education, respectively. Surprisingly, under the second approach we find a higher effect of kindergarten exposure for higher education. The results of the difference-in-differences approach are within the range found for the panel estimation (between 0.1 and 3.6p.p for all grades and scores).

The results are not substantially different when we control for population mobility. By withdrawing the metropolitan areas from the estimation the impact is lower using both approaches, ranging between 0.40 and 1.47p.p for upper-secondary education, and between 0.36 and 0.88p.p for higher education. Alternatively, when we control for mobility, there is no clear trend on the impact compared to the baseline specification. These results are line with those of Duflo (2001), where it is found that one school built per 1,000 children increased the education of children aged 2 to 6 in 1974 by 0.12 years for the whole sample.

In what regards wages, we find a low negative impact of -0.78% using the first approach, suggesting that kindergarten expansion had no impact on wages of the treated cohort. This finding might be explained by the decline in the wage premium that Campos and Reis (2017) find among upper-secondary graduated workers, which is the case of the majority of the individuals in the con-

Table 7: Difference-in-Differences Results - Second Approach

VARIABLES	(1) Secondary	(4) Secondary2	(7) High education	(10) High education2	(13) Wages
interaction	0.008***	0.004***	0.011***	0.017***	0.005***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)
Observations	520,223	414,034	493,851	264,027	383,958
R-squared	0.064	0.042	0.058	0.045	0.584
Birth FE	Yes	Yes	Yes	Yes	Yes
NUTS FE	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the difference-in-differences estimation results using the second approach, where we observe both comparison and treated cohorts in the same time interval, and as a consequence they have different age profiles. Secondary corresponds to upper-secondary education (including higher education); Secondary2 corresponds to upper-secondary education (excluding higher education); High education corresponds to higher education; and High education2 corresponds to higher education compared to having completed upper-secondary education, while the former specifications are compared to not having completed the level of education in question.
Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

sidered cohorts under the first approach. For the second approach, we find an impact of 0.53% for the specification with establishment fixed effects (as this approach solely considers the year of 2012), although it loses its significance when we control for population mobility, varying between -3.51% and 0.53%. Although these findings differ from Duflo's, where the estimates range from 1.5% to 2.7%, she also finds more unstable results on wages, since this is a long-run outcome, whose impact is harder to measure. Also, these results are lower than those we found using the fixed effects estimation (impact ranging between 0.37% and 4.13%).

Finally, on Tables 8 and 9 we can find the results of instrumentalising the effect of education on wages, and on Tables 24 and 25 the robustness checks. Since our endogenous variable is a binary variable on completing upper-secondary and higher education, we compute the IV estimation following (Windmeijer & Santos Silva, 1997). In the first step, we run a probit model on the dummy variable of upper-secondary and higher education. Then, we use the fitted values on this estimation as the instrumental variable in the wages estimation.

We can claim that our instrument is strong since the F-statistic on the first-stage fulfils the rule of thumb of being higher than 10. We estimate a causal impact of completing upper-secondary education on wages ranging between 18.1% and 58.2%, depending on the specification and approach.

Table 8: IV Results - First Approach

VARIABLES	(1) Secondary	(2) Secondary2	(3) High educ	(4) High educ2	(5) Secondary FE	(6) Secondary2 FE	(7) High educ FE	(8) High educ2 FE
secondary	0.430***	0.458***			0.072***	0.055***		
	(0.0102)	(0.0121)			(0.0004)	(0.0004)		
high education			0.294***	0.332***			0.367***	0.389***
			(0.0138)	(0.0174)			(0.001)	(0.001)
Observations	3,606,678	3,606,678	1,441,390	540,470	3,627,281	3,448,989	1,450,416	542,955
R-squared	0.435	0.372	0.340	0.372	0.518	0.512	0.342	0.374

Notes: This table reports the IV estimation results using the first approach, where individuals on the comparison cohort have the same age profile as individuals in the treated cohort, but are observed in a different time period. Columns 1 to 4 refer to the difference-in-differences approach, while columns 5 to 8 refer to the fixed effects estimation. Secondary corresponds to upper-secondary education (including higher education); Secondary2 corresponds to upper-secondary education (excluding higher education); High education corresponds to higher education; and High education2 corresponds to higher education compared to having completed secondary education, while the former specifications are compared to not having completed the level of education in question.
Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 9: IV Results - Second Approach

VARIABLES	(1) Secondary	(2) Secondary2	(3) High educ	(4) High educ2	(5) Secondary FE	(6) Secondary2 FE	(7) High educ FE	(8) High educ2 FE
secondary	0.582***	0.181			0.368***	0.213***		
	(0.0497)	(0.126)			(0.001)	(0.001)		
high education			0.213***	0.417***			0.555***	0.665***
			(0.0751)	(0.0594)			(0.002)	(0.002)
Observations	383,958	383,958	364,200	243,241	389,642	308,628	369,363	246,865
R-squared	0.268	0.150	0.283	0.426	0.318	0.244	0.373	0.481

Notes: This table reports the IV estimation results using the second approach, where we observe both comparison and treated cohorts in the same time interval, and as a consequence they have different age profiles. Columns 1 to 4 refer to the difference-in-differences approach, while columns 5 to 8 refer to the fixed effects estimation. Secondary corresponds to upper-secondary education (including higher education); Secondary2 corresponds to upper-secondary education (excluding higher education); High education corresponds to higher education; and High education2 corresponds to higher education compared to having completed secondary education, while the former specifications are compared to not having completed the level of education in question.
Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

The impact of higher education is between 21.3% and 41.7%, being in line with the literature. Campos and Reis (2017) found a return to higher education compared to upper-secondary education of 33.1% for women and 34.3% for men, in 1986, and 49.6% for women, and 44.9% for men, in 2012. The IV of the panel estimation denotes a potential more imprecise estimation on the returns of upper-secondary education, since it shows a large downward bias between the panel estimation and the IV approach. Although Duflo (2001) finds a lower estimated return to education of 7.8% per year of education, the educational levels we are studying comprise more than one year of edu-

cation: upper-secondary education comprises 3 years, while higher education can range between 3 to 6 years.

7 Conclusions

This paper is the first thorough analysis of the impact of the massive expansion of kindergarten provision in Portugal that occurred between 1985 and 2009. We rely on two large scale administrative datasets, *MISI* and QP. Our main contributes are as follows. First, we construct the series of the evolution of kindergartens in Portugal, making use of labour market administrative data, to build different measures of municipal kindergarten intensity, quantifying the potential exposure for each student and worker. Second, we use this measure to assess how the exposure of individuals to kindergarten education impacts their schooling and labour market outcomes, although this is not a direct measure of exposure but an intention-to-treat approach. Our goal is to measure how this kindergarten expansion impacted scores in the Portuguese and Mathematics' national exams of the 4^{th}, 6^{th}, 9^{th} and 12^{th} grades, and wages, at different moments in time.

We find that kindergarten expansion led to an increase in educational attainment and that the effect of potential kindergarten exposure is larger on the 6^{th} and 9^{th} grades. The impact of an increase in one kindergarten per 10,000 inhabitants when individuals were between 0 and 3 years of age on the probability of attaining a passing score ranges between 1.7 and 3.6p.p, varies between 1.4 and 2.8p.p on the probability of having a score higher than 3, and between 0.1 and 0.7p.p for achieving a score higher than 4. Because wages are a long-run outcome, labour market effects of kindergarten exposure are harder to measure, which reflects on the results. Depending on the empirical strategy, we find an impact of potential kindergarten exposure ranging between 0% and 0.53%. Since our results may be influenced by the mobility of students and workers, we have built a measure of population mobility, and our results are robust when we control for mobility effects.

Although we do not find transmission effects on wages, we find that investing in kindergarten has long-run effects on the municipalities we studied, constituting a possible elevator for social mobility. Hence, failing in granting valuable early childhood education limits toddlers' futures, as it

deprives them from unique opportunities that would allow them to develop abilities that are essential for their social and economic success in the future. Furthermore, failing in providing quality kindergarten education also restrains countries' prospects, since it sets a bound to the development of human capital, which is paramount to curtail inequality and advocate harmonious, thriving societies.

A possible avenue for future research is the study of the impact of exposure to kindergarten on other pivotal dimensions, namely non-cognitive. This analysis would be even improved if we had data on direct exposure to kindergarten.

References

Akresh, R., Halim, D., & Kleemans, M. (2018). *Long-term and intergenerational effects of education: Evidence from school construction in indonesia* (Tech. Rep.). National Bureau of Economic Research.

Barnett, W. S. (1992). Benefits of compensatory preschool education. *Journal of Human resources*, 279–312.

Campos, M. M., & Reis, H. (2017). Revisiting the returns to schooling in the portuguese economy.

Carneiro, P., & Ginja, R. (2014). Long-term impacts of compensatory preschool on health and behavior: Evidence from head start. *American Economic Journal: Economic Policy*, *6*(4), 135–73.

Cascio, E. U. (2009). *Do investments in universal early education pay off? long-term effects of introducing kindergartens into public schools* (Tech. Rep.). National Bureau of Economic Research.

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., & Yagan, D. (2011). How does your kindergarten classroom affect your earnings? evidence from project star. *The Quarterly journal of economics*, *126*(4), 1593–1660.

Cunha, F., Heckman, J. J., & Schennach, S. M. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, *78*(3), 883–931.

Duflo, E. (2001). Schooling and labor market consequences of school construction in indonesia: Evidence from an unusual policy experiment. *American economic review*, *91*(4), 795–813.

Duncan, G. J., & Magnuson, K. (2013). Investing in preschool programs. *Journal of economic perspectives*, *27*(2), 109–32.

Elango, S., García, J. L., Heckman, J. J., & Hojman, A. (2015). Early childhood education. In *Economics of means-tested transfer programs in the united states, volume 2* (pp. 235–297). University of Chicago Press.

Hart, B., & Risley, T. R. (2003). The early catastrophe: The 30 million word gap by age 3. *American educator*, *27*(1), 4–9.

Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., & Yavitz, A. (2010). The rate of return to the highscope perry preschool program. *Journal of public Economics*, *94*(1-2), 114–128.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor economics*, *24*(3), 411–482.

Ludwig, J., & Miller, D. L. (2007). Does head start improve children's life chances? evidence from a regression discontinuity design. *The Quarterly journal of economics*, *122*(1), 159–208.

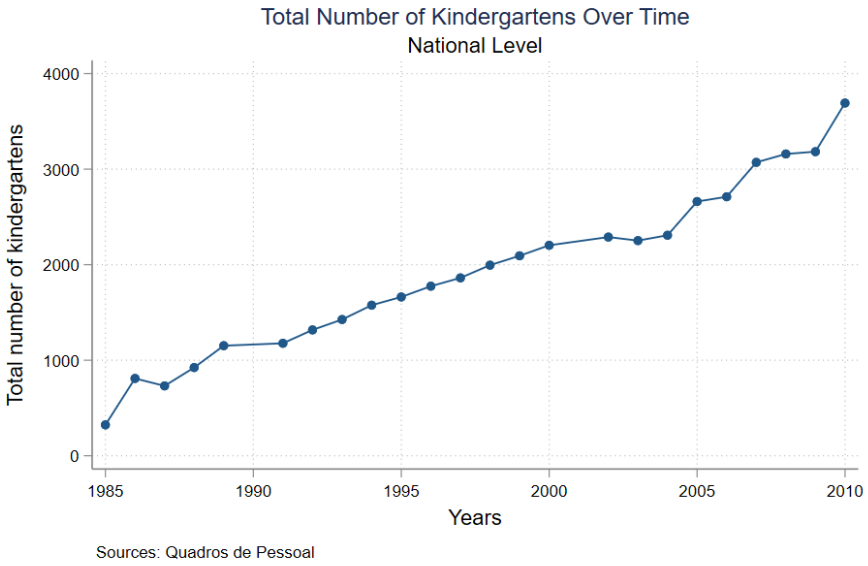
Neilson, C. A., & Zimmerman, S. D. (2014). The effect of school construction on test scores, school enrollment, and home prices. *Journal of Public Economics*, *120*, 18–31.

Palma, N. P. G., & Reis, J. (2018). Can autocracy promote literacy? evidence from a cultural alignment success story.

Windmeijer, F. A., & Santos Silva, J. M. (1997). Endogeneity in count data models: an application to demand for health care. *Journal of applied econometrics*, *12*(3), 281–294.

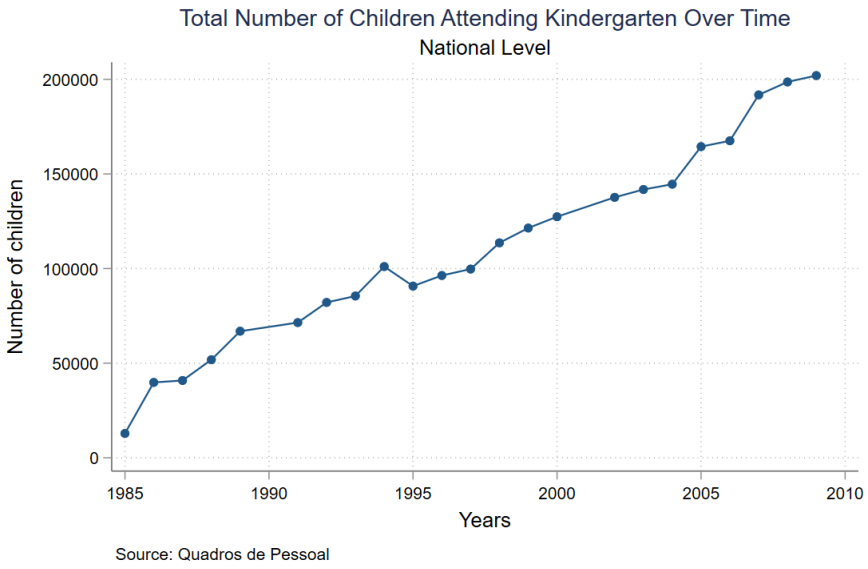
8 Appendix

Figure 3: Kindergarten Expansion



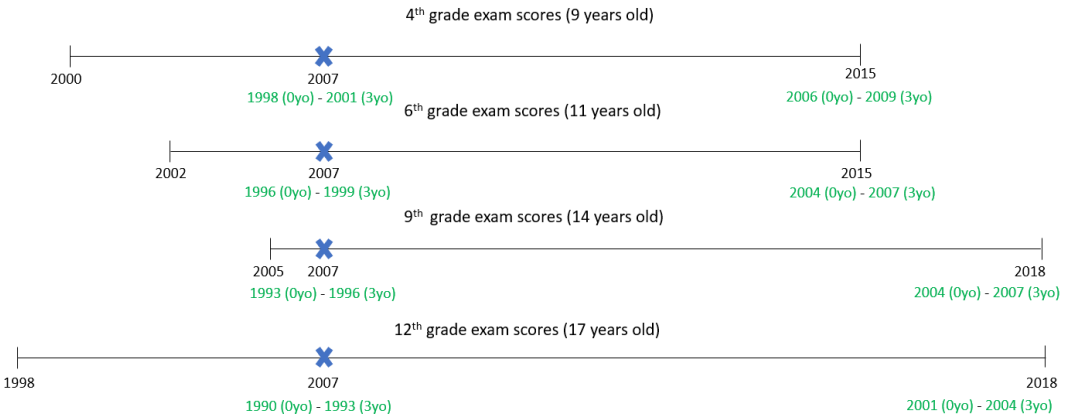
Notes: This figure illustrates the expansion of kindergartens between 1985 and 2009 based on the kindergarten intensity measure we constructed. Author's computation.

Figure 4: Evolution of Children Attending Kindergarten



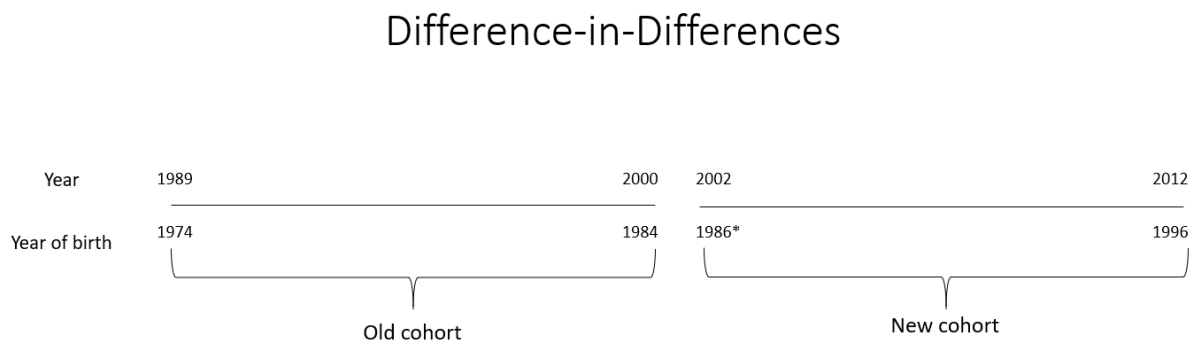
Notes: This figure shows the evolution of children attending kindergarten between 1985 and 2009 based on the proxy we computed. Author's computation.

Figure 5: MISI Dataset Timeline



Notes: The birth years were computed assuming that students taking 4^{th} grade exams were 9 years old, 6^{th} 11 years old, 9^{th} 14 years old and 12^{th} 17 years old.

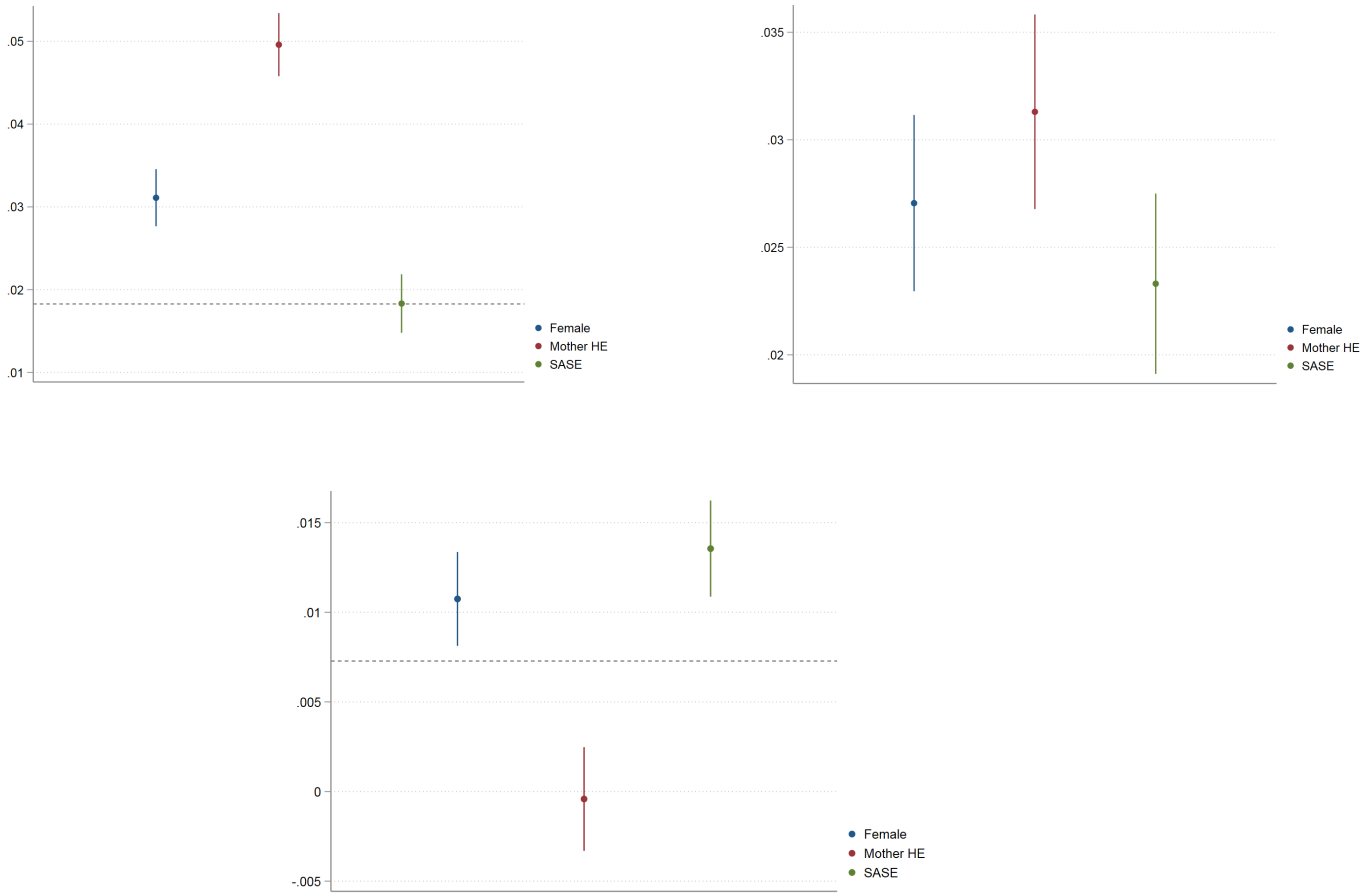
Figure 6: Differences-in-Differences Approach Timeline



Notes: This figure shows the timeline of the difference-in-differences approach.
*As in the QP dataset we do not have the year of 2001, and keeping our assumption that individuals enter the database when they are 16 years old, this is the year when individuals born in 1985 enter the database. Hence, these individuals enter the new cohort in 2002.

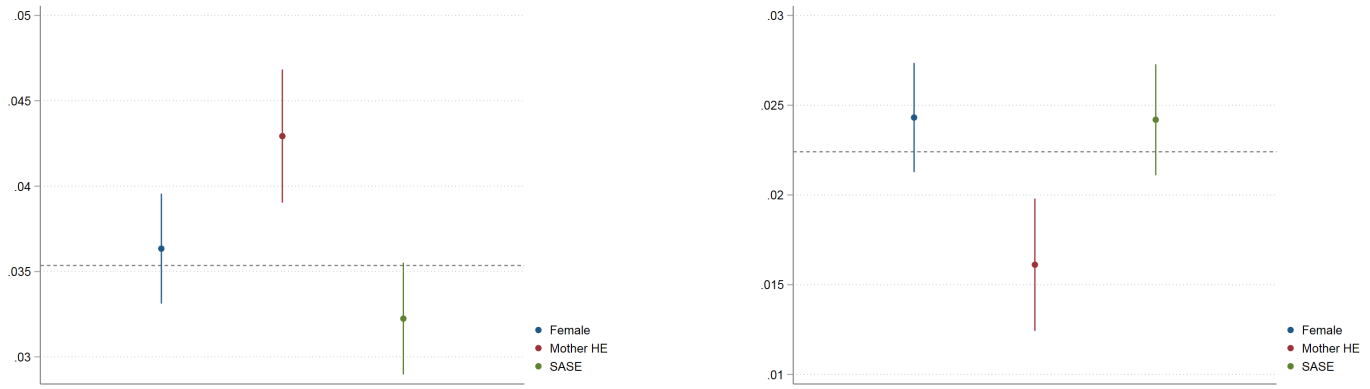
The plots in the following figures displays the coefficients of the effect of kindergarten exposure on attaining a score higher than 2, 3 and 4, respectively, in the 4^{th}, 6^{th}, 9^{th} and 12^{th} grades, Portuguese and Mathematics exams, by this order. All the figures were plotted using the school fixed effects specification.

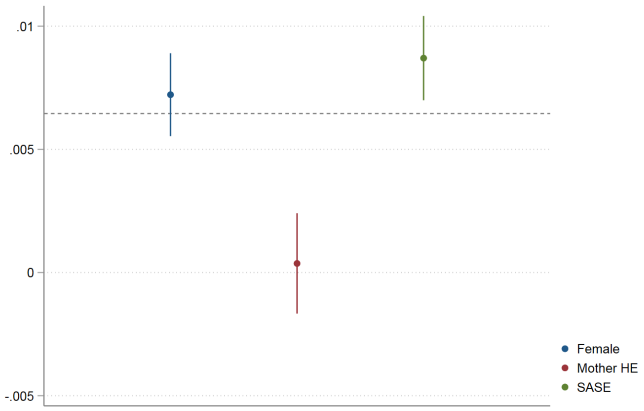
Figure 7: 4^{th} grade Portuguese Exam Scores Heterogeneity



Notes: The figure displays that the effect of kindergarten exposure on the 4^{th} grade Portuguese exam is slightly higher for females across all scores, the effect of the student's mother having completed higher education is higher for scores higher than 2 and 3 but lower for a score higher than 4, while the effect of having a means-tested programme (SASE) increases for higher scores.

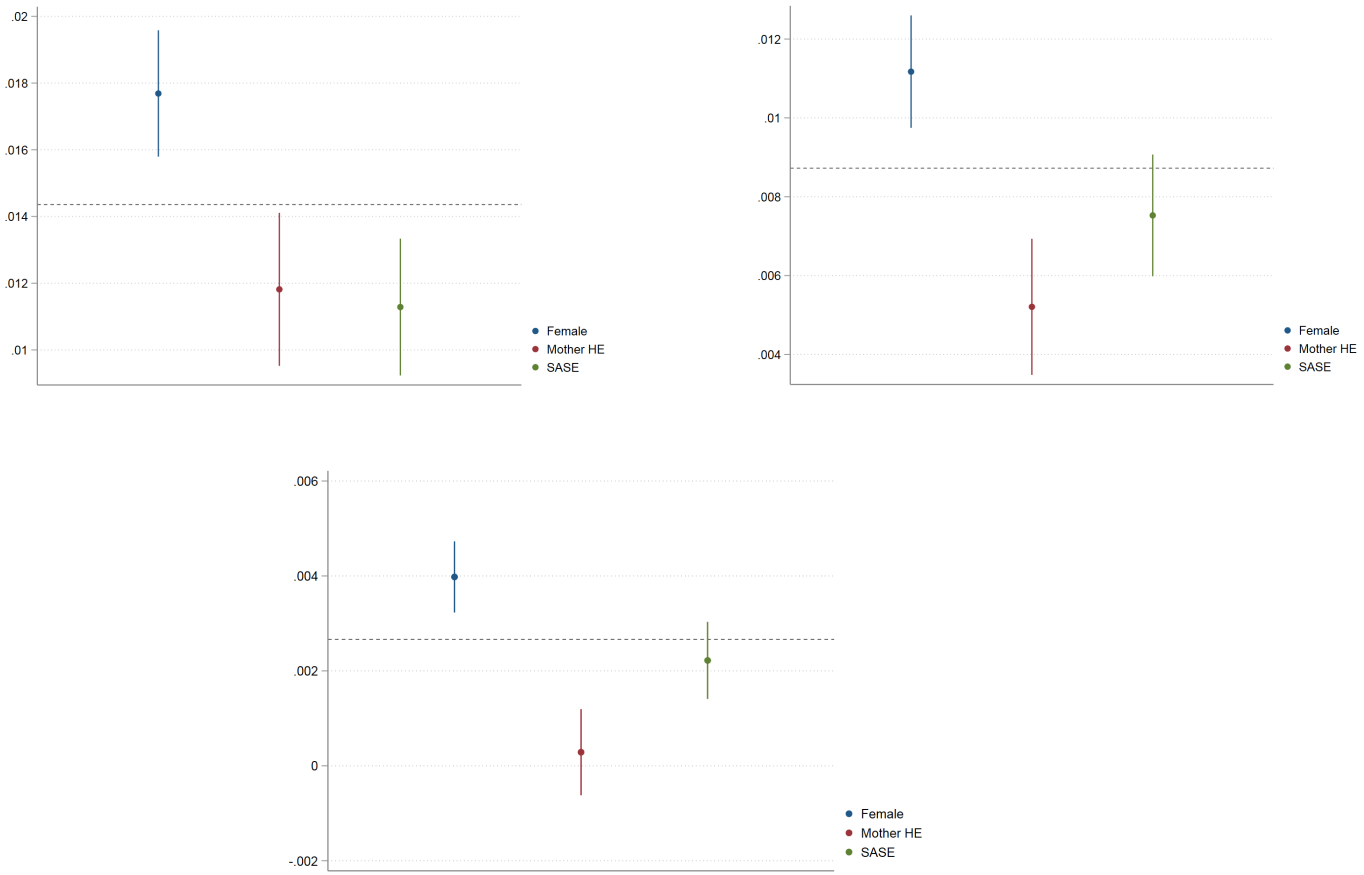
Figure 8: 6^{th} grade Portuguese Exam Scores Heterogeneity





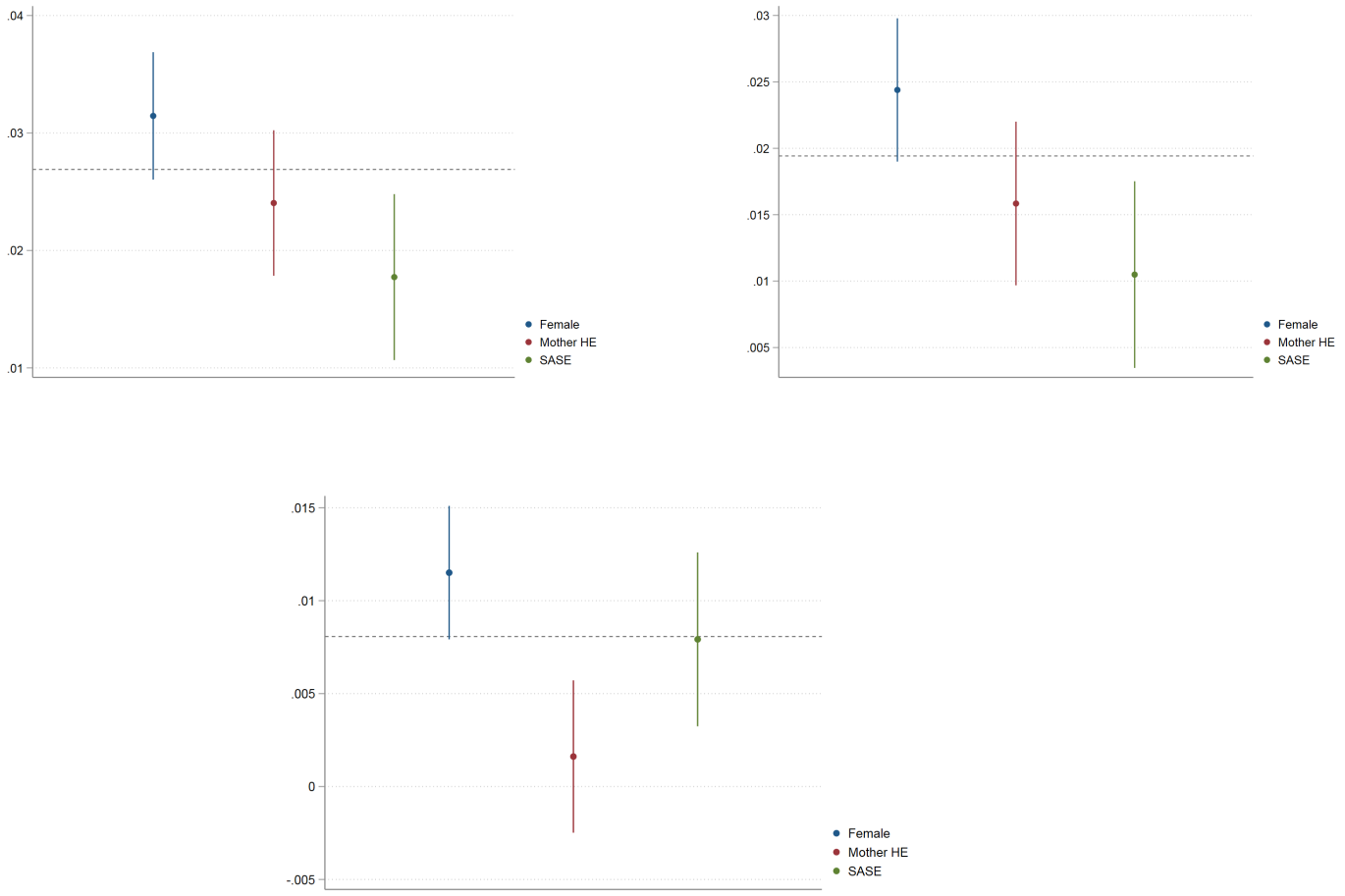
Notes: The figure displays that the effect of kindergarten exposure on the 6th grade Portuguese exam is slightly higher for females across all scores, the effect of the student's mother having completed higher education is lower for scores higher than 3 and 4, and the effect of having a means-tested programme (SASE) is increasing for higher scores.

Figure 9: 9th grade Portuguese Exam Scores Heterogeneity



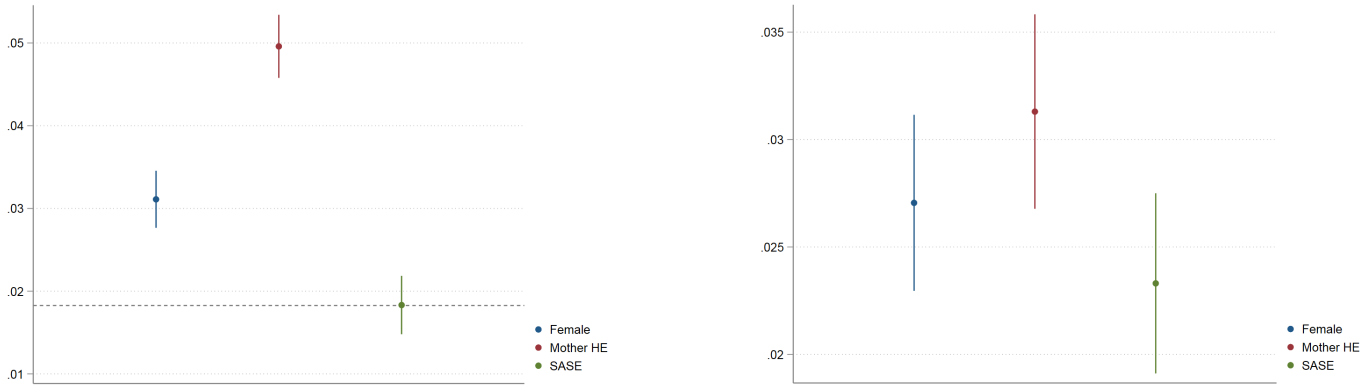
Notes: The figure displays that the effect of kindergarten exposure on the 9th grade Portuguese exam is higher for females across all scores, while the effects of the student's mother having completed higher education and of having a means-tested programme (SASE) is lower for all scores.

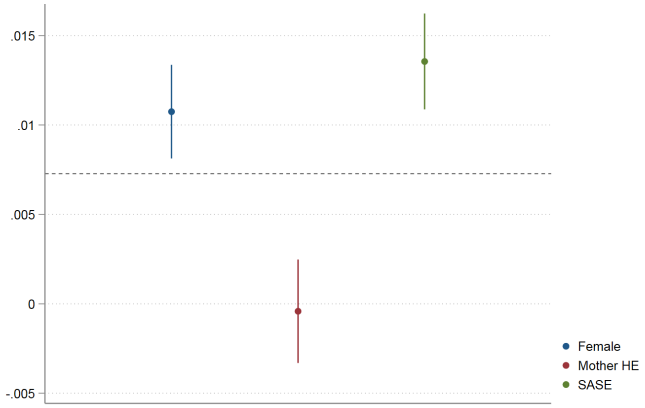
Figure 10: 12^{th} grade Portuguese Exam Scores Heterogeneity



Notes: The figure displays that the effect of kindergarten exposure on the 12^{th} grade Portuguese exam is higher for females across all scores, while the effects of the student's mother having completed higher education and of having a means-tested programme (SASE) is lower for all scores.

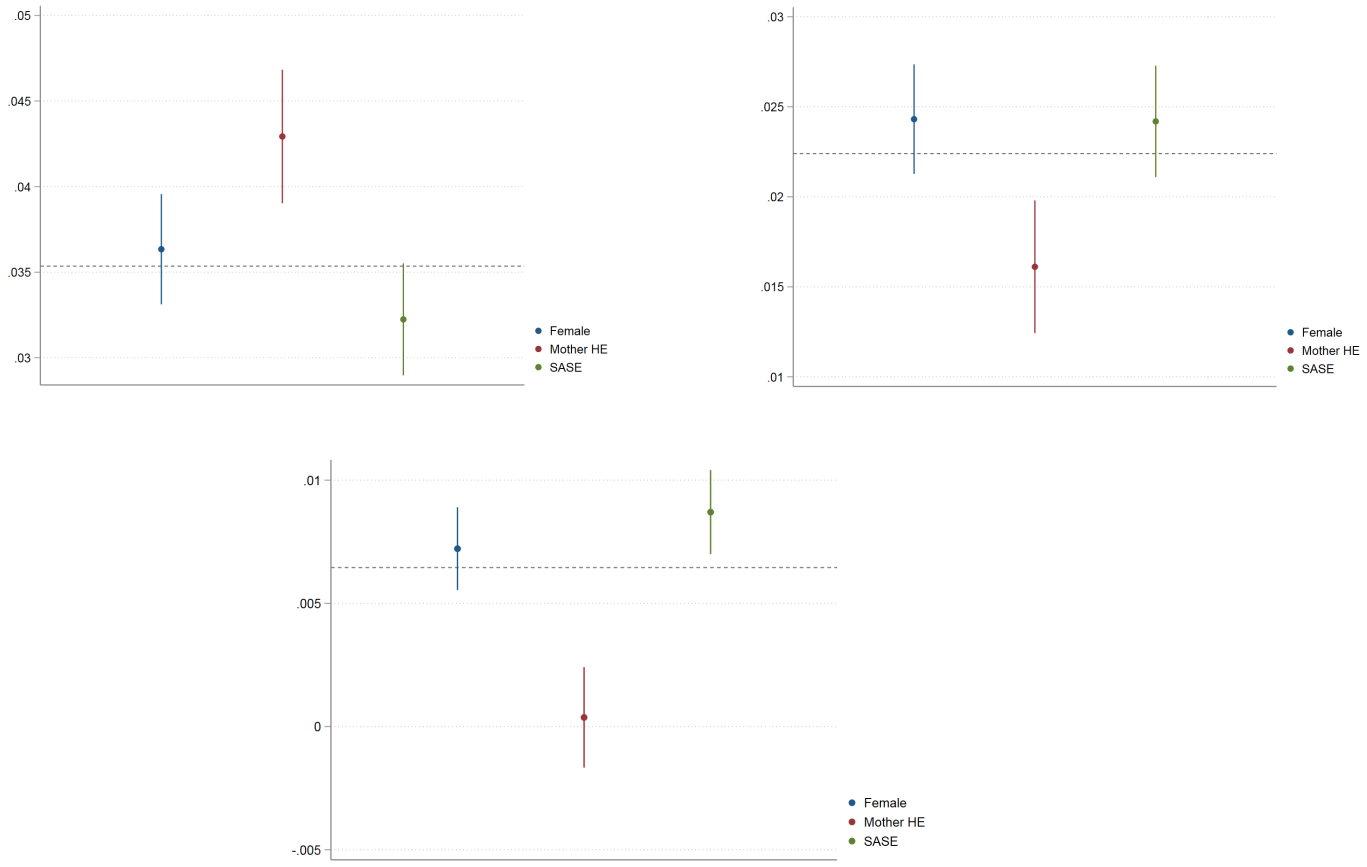
Figure 11: 4^{th} grade Mathematics Exam Scores Heterogeneity





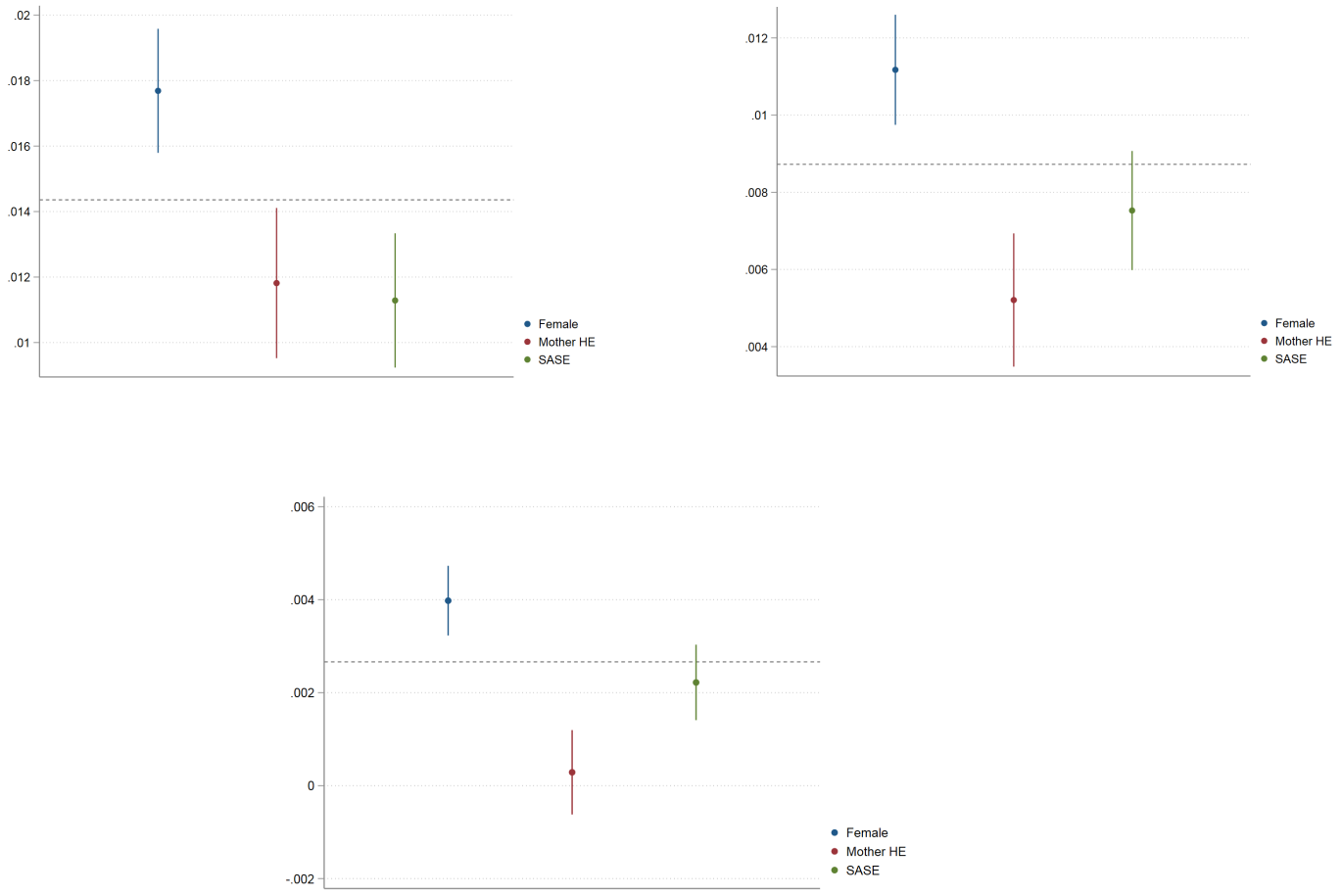
Notes: The figure displays that the effect of kindergarten exposure on the 4th grade Mathematics exam is slightly higher for females across all scores, the effect of the student's mother having completed higher education is higher for scores higher than 2 and 3 but lower for a score higher than 4, while the effect of having a means-tested programme (SASE) increases for higher scores.

Figure 12: 6th grade Mathematics Exam Scores Heterogeneity



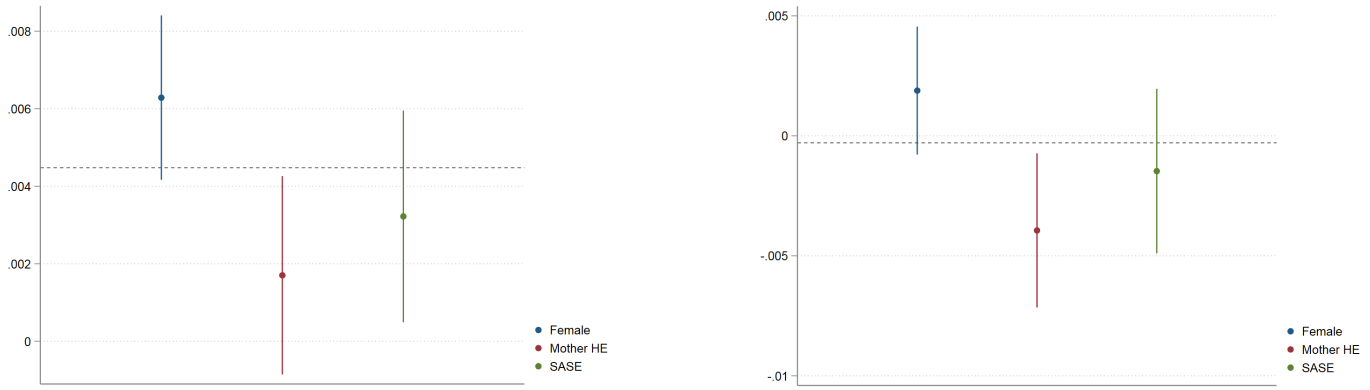
Notes: The figure displays that the effect of kindergarten exposure on the 6th grade Mathematics exam is slightly higher for females across all scores, the effect of the student's mother having completed higher education is lower for scores higher than 3 and 4, and the effect of having a means-tested programme (SASE) is increasing for higher scores.

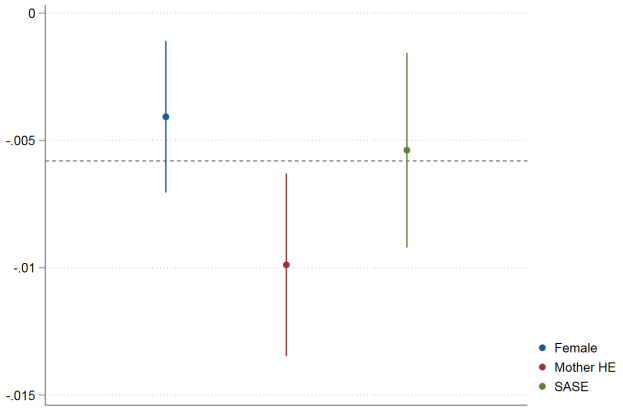
Figure 13: 9^{th} grade Mathematics Exam Scores Heterogeneity



Notes: The figure displays that the effect of kindergarten exposure on the 9^{th} grade Mathematics exam is higher for females across all scores, while the effects of the student's mother having completed higher education and of having a means-tested programme (SASE) is lower for all scores.

Figure 14: 12^{th} grade Mathematics Exam Scores Heterogeneity





Notes: The figure displays that the effect of kindergarten exposure on the 12^{th} grade Mathematics exam is higher for females across all scores, while the effects of the student's mother having completed higher education and of having a means-tested programme (SASE) is lower for all scores.

Table 10: Percentage Impact on Exam Scores

Grade		Coefficient PT	Students % (PT)	% Coefficient (PT)	Coefficient Mat	Students % (Mat)	% Coefficient (Mat)
4	Score >2	0,018	41,47%	4,34%	0,017	37,86%	4,49%
	Score >3	0,020	19,87%	10,07%	0,014	19,45%	7,20%
	Score >4	0,004	2,99%	13,38%	0,006	4,92%	12,20%
6	Score >2	0,032	40,81%	7,84%	0,035	32,66%	10,72%
	Score >3	0,028	16,03%	17,47%	0,023	13,70%	16,79%
	Score >4	0,004	1,90%	21,05%	0,006	3,01%	19,93%
9	Score >2	0,035	35,54%	9,85%	0,036	25,64%	14,04%
	Score >3	0,019	12,21%	15,56%	0,021	12,80%	16,41%
	Score >4	0,002	1,32%	15,15%	0,006	3,11%	19,29%
12	Score >2	0,002	64,09%	0,37%	0,026	64,99%	4,00%
	Score >3	0,015	20,52%	7,31%	0,017	38,36%	4,43%
	Score >4	0,001	1,17%	8,55%	0,007	12,08%	5,79%

Notes: This table reports the percentage impact of kindergarten exposure on each score for the 4^{th}, 6^{th}, 9^{th} and 12^{th} grades considering the percentage of students in each score threshold, for the Portuguese and Mathematics national exams.

Table 11: Fixed Effects Estimation Results - Portuguese Exam (including covariates)

	grade 4			grade 6			grade 9			grade 12		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
VARIABLES	score >2	score >3	score >4	score >2	score >3	score >4	score >2	score >3	score >4	score >2	score >3	score >4
kindergarten intensity	0.018***	0.020***	0.004***	0.032***	0.028***	0.004***	0.035***	0.019***	0.002***	0.024***	0.015***	0.001***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.000)
female	0.048***	0.076***	0.024***	0.074***	0.114***	0.026***	0.115***	0.102***	0.017***	0.131***	0.109***	0.009***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.002)	(0.001)	(0.000)
father high education	0.038***	0.105***	0.041***	0.034***	0.121***	0.046***	0.067***	0.117***	0.035***	0.071***	0.087***	0.012***
	(0.002)	(0.003)	(0.001)	(0.002)	(0.003)	(0.001)	(0.002)	(0.002)	(0.001)	(0.003)	(0.002)	(0.001)
mother high education	0.065***	0.144***	0.051***	0.064***	0.162***	0.049***	0.099***	0.143***	0.033***	0.102***	0.101***	0.011***
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.003)	(0.002)	(0.001)
SASE A	-0.031***	-0.052***	-0.013***	-0.037***	-0.069***	-0.015***	-0.042***	-0.043***	-0.007***	-0.012***	-0.019***	-0.002***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.003)	(0.002)	(0.001)
SASE B	-0.079***	-0.105***	-0.022***	-0.094***	-0.126***	-0.022***	-0.076***	-0.068***	-0.010***	-0.041***	-0.032***	-0.003***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.003)	(0.002)	(0.001)
computer	0.018***	0.034***	0.006***	0.030***	0.026***	0.003***	0.006***	0.006***	0.001	-0.018***	-0.020***	-0.003**
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.004)	(0.004)	(0.001)
internet	0.026***	0.035***	0.010***	0.033***	0.053***	0.010***	0.046***	0.037***	0.004***	0.039***	0.029***	0.004***
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.004)	(0.003)	(0.001)
mother portuguese	0.012***	0.021***	0.006***	0.034***	0.034***	0.007***	0.041***	0.026***	0.005***	0.067***	0.028***	0.002
	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)	(0.001)	(0.003)	(0.003)	(0.001)	(0.005)	(0.004)	(0.001)
father portuguese	0.023***	0.028***	0.007***	0.035***	0.028***	0.003**	0.033***	0.021***	0.001	0.054***	0.033***	0.002*
	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)	(0.001)	(0.003)	(0.003)	(0.001)	(0.004)	(0.004)	(0.001)
mother unemployed	-0.007***	-0.009***	-0.001	-0.004**	-0.009***	-0.002**	-0.001	-0.011***	-0.002***	-0.009**	-0.011***	-0.002*
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.003)	(0.003)	(0.001)
father unemployed	-0.030***	-0.032***	-0.006***	-0.026***	-0.025***	-0.001	-0.015***	-0.015***	-0.002***	-0.014***	-0.010***	0.000
	(0.002)	(0.003)	(0.001)	(0.002)	(0.003)	(0.001)	(0.003)	(0.002)	(0.001)	(0.004)	(0.004)	(0.001)
Constant	0.795***	0.220***	0.007***	0.740***	0.167***	0.004	0.600***	0.191***	0.011***	0.207***	-0.014	-0.001
	(0.004)	(0.005)	(0.002)	(0.008)	(0.010)	(0.004)	(0.010)	(0.009)	(0.003)	(0.030)	(0.026)	(0.007)
Observations	533,853	533,853	533,853	452,486	452,486	452,486	565,172	565,172	565,172	292,970	292,970	292,970
R-squared	0.180	0.149	0.061	0.160	0.167	0.076	0.158	0.167	0.064	0.145	0.106	0.027
Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Academic Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Teacher FE				Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the fixed effects estimation results from the baseline specifications for the Portuguese exam including the covariates. It presents the results for having a score higher than 2, 3, and 4 for all grades. For the 4^{th} grade the specification used is school fixed effects, while for the remaining grades it is teacher fixed effects.

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

Table 12: Fixed Effects Estimation Results - Mathematics Exam (including covariates)

	grade 4			grade 6			grade 9			grade 12		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
VARIABLES	score >2	score >3	score >4	score >2	score >3	score >4	score >2	score >3	score >4	score >2	score >3	score >4
kindergarten intensity	0.017***	0.014***	0.006***	0.035***	0.023***	0.006***	0.036***	0.021***	0.006***	0.026***	0.017***	0.007***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.003)	(0.003)	(0.002)
female	-0.034***	-0.048***	-0.018***	0.005***	-0.005***	-0.004***	0.012***	0.025***	0.011***	0.025***	0.033***	0.021***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
father high education	0.061***	0.120***	0.063***	0.089***	0.152***	0.080***	0.128***	0.157***	0.079***	0.092***	0.124***	0.082***
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.001)	(0.002)	(0.002)	(0.001)	(0.004)	(0.004)	(0.003)
mother high education	0.097***	0.154***	0.068***	0.145***	0.185***	0.074***	0.179***	0.179***	0.075***	0.113***	0.135***	0.067***
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.003)	(0.003)	(0.002)
SASE A	-0.042***	-0.049***	-0.015***	-0.075***	-0.075***	-0.021***	-0.073***	-0.058***	-0.017***	-0.018***	-0.020***	-0.015***
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.004)	(0.004)	(0.003)
SASE B	-0.101***	-0.103***	-0.031***	-0.164***	-0.126***	-0.030***	-0.135***	-0.089***	-0.022***	-0.059***	-0.049***	-0.019***
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.004)	(0.004)	(0.003)
computer	0.026***	0.040***	0.015***	0.040***	0.024***	0.003***	0.019***	0.005***	-0.000	-0.032***	-0.036***	-0.021***
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.007)	(0.007)	(0.004)
internet	0.031***	0.031***	0.012***	0.048***	0.048***	0.016***	0.055***	0.041***	0.011***	0.046***	0.049***	0.027***
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.006)	(0.006)	(0.004)
mother portuguese	0.010***	0.004	0.004*	0.025***	0.018***	0.006***	0.036***	0.025***	0.007***	0.061***	0.037***	0.010**
	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)	(0.007)	(0.007)	(0.005)
father portuguese	0.028***	0.029***	0.006***	0.037***	0.019***	0.004***	0.036***	0.021***	0.006***	0.055***	0.040***	0.012***
	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)	(0.007)	(0.007)	(0.005)
mother unemployed	-0.017***	-0.018***	-0.006***	-0.018***	-0.016***	-0.002**	-0.026***	-0.019***	-0.007***	-0.023***	-0.021***	-0.004
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.005)	(0.005)	(0.004)
father unemployed	-0.037***	-0.031***	-0.004***	-0.039***	-0.022***	-0.005***	-0.034***	-0.023***	-0.005***	-0.021***	-0.028***	-0.007
	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)	(0.003)	(0.002)	(0.001)	(0.007)	(0.007)	(0.004)
Constant	0.838***	0.418***	0.116***	0.631***	0.181***	0.042***	0.379***	0.157***	0.031***	0.404***	0.214***	0.029
	(0.004)	(0.005)	(0.003)	(0.009)	(0.008)	(0.005)	(0.010)	(0.009)	(0.005)	(0.046)	(0.047)	(0.031)
Observations	532,383	532,383	532,383	460,119	460,119	460,119	570,841	570,841	570,841	138,718	138,718	138,718
R-squared	0.190	0.154	0.105	0.223	0.178	0.094	0.192	0.182	0.096	0.155	0.145	0.088
Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Academic Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Teacher FE				Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the fixed effects estimation results from the baseline specifications for the Mathematics exam including the covariates. It presents the results for having a score higher than 2, 3, and 4 for all grades. For the 4^{th} grade the specification used is school fixed effects, while for the remaining grades it is teacher fixed effects.

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

Appendix Table 13: Robustness Fixed Effects Estimation Results - 4th grade, Portuguese Exam

VARIABLES	(1) grade >2	(2) grade >2	(3) grade >2	(4) grade >2	(5) grade >2
kindergarten intensity	0.003**	0.031*	0.009*	0.021*	0.017**
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)
Observations	533,853	533,853	533,853	397,292	532,873
R-squared	0.147	0.151	0.180	0.178	0.180
Covariates	Yes	Yes	Yes	Yes	Yes
Academic Year FE	Yes	Yes	Yes	Yes	Yes
School FE			Yes	Yes	Yes
Teacher FE					
Municipality FE		Yes			
NUTS FE	Yes				
VARIABLES	(6) grade >3	(7) grade >3	(8) grade >3	(9) grade >3	(10) grade >3
kindergarten intensity	0.004**	0.036*	0.013*	0.027*	0.020**
	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Observations	533,853	533,853	533,853	397,292	532,873
R-squared	0.118	0.121	0.149	0.146	0.150
Covariates	Yes	Yes	Yes	Yes	Yes
Academic Year FE	Yes	Yes	Yes	Yes	Yes
School FE			Yes	Yes	Yes
Teacher FE					
Municipality FE		Yes			
NUTS FE	Yes				
VARIABLES	(11) grade >4	(12) grade >4	(13) grade >4	(14) grade >4	(15) grade >4
kindergarten intensity	0.001**	0.007*	0.004*	0.006*	0.004**
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	533,853	533,853	533,853	397,292	532,873
R-squared	0.043	0.044	0.061	0.063	0.061
Covariates	Yes	Yes	Yes	Yes	Yes
Academic Year FE	Yes	Yes	Yes	Yes	Yes
School FE			Yes	Yes	Yes
Teacher FE					
Municipality FE		Yes			
NUTS FE	Yes				

Notes: This table reports the fixed estimation results for the 4th grade Portuguese exam. The first table corresponds to the results of having a score higher than 2, the second of having a score higher than 3, and the third of having a score higher than 4. The first column of each table corresponds to the specification including NUTS III fixed effects, the second of including municipality fixed effects, the third of using the alternative measure of kindergarten intensity (number of kindergartens on individuals' birth year), the fourth of withdrawing the metropolitan areas, and the fifth of including mobility as a control. The same follows for the other scores thresholds.

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

Appendix Table 14: Robustness Fixed Effects Estimation Results - 4th grade, Mathematics Exam

VARIABLES	(1) grade >2	(2) grade >2	(3) grade >2	(4) grade >2	(5) grade >2
kindergarten intensity	0.003***	0.032***	0.007***	0.019***	0.016***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)
Observations	532,383	532,383	532,383	396,178	531,406
R-squared	0.151	0.157	0.190	0.184	0.191
Covariates	Yes	Yes	Yes	Yes	Yes
Academic Year FE	Yes	Yes	Yes	Yes	Yes
School FE			Yes	Yes	Yes
Teacher FE					
Municipality FE		Yes			
NUTS FE	Yes				
VARIABLES	(6) grade >3	(7) grade >3	(8) grade >3	(9) grade >3	(10) grade >3
kindergarten intensity	0.002**	0.028***	0.008***	0.022***	0.014***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Observations	532,383	532,383	532,383	396,178	531,406
R-squared	0.112	0.117	0.154	0.151	0.154
Covariates	Yes	Yes	Yes	Yes	Yes
Academic Year FE	Yes	Yes	Yes	Yes	Yes
School FE			Yes	Yes	Yes
Teacher FE					
Municipality FE		Yes			
NUTS FE	Yes				
VARIABLES	(11) grade >4	(12) grade >4	(13) grade >4	(14) grade >4	(15) grade >4
kindergarten intensity	0.000	0.008***	0.004***	0.010***	0.006***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	532,383	532,383	532,383	396,178	531,406
R-squared	0.073	0.076	0.105	0.108	0.106
Covariates	Yes	Yes	Yes	Yes	Yes
Academic Year FE	Yes	Yes	Yes	Yes	Yes
School FE			Yes	Yes	Yes
Teacher FE					
Municipality FE		Yes			
NUTS FE	Yes				

Notes: This table reports the fixed effects estimation results for the 4th grade Mathematics exam. The first table corresponds to the results of having a score higher than 2, the second of having a score higher than 3, and the third of having a score higher than 4. The first column of each table corresponds to the specification including NUTS III fixed effects, the second of including municipality fixed effects, the third of using the alternative measure of kindergarten intensity (number of kindergartens on individuals' birth year), the fourth of withdrawing the metropolitan areas, and the fifth of including mobility as a control. The same follows for the other scores thresholds.

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

Appendix Table 15: Robustness Fixed Effects Estimation Results - 6th grade, Portuguese Exam

VARIABLES	(1) score >2	(2) score >2	(3) score >2	(4) score >2	(5) score >2	(6) score >2
kindergarten intensity	0.005***	0.054***	0.028***	0.022***	0.041***	0.034***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Observations	537,558	537,558	537,558	452,486	331,546	451,536
R-squared	0.115	0.120	0.129	0.160	0.152	0.160
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Academic Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE			Yes	Yes	Yes	Yes
Teacher FE				Yes	Yes	Yes
Municipality FE		Yes				
NUTS FE	Yes					
VARIABLES	(7) score >3	(8) score >3	(9) score >3	(10) score >3	(11) score >3	(12) score >3
kindergarten intensity	0.005***	0.046***	0.025***	0.020***	0.035***	0.029***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Observations	537,558	537,558	537,558	452,486	331,546	451,536
R-squared	0.124	0.127	0.135	0.167	0.162	0.167
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Academic Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE			Yes	Yes	Yes	Yes
Teacher FE				Yes	Yes	Yes
Municipality FE		Yes				
NUTS FE	Yes					
VARIABLES	(13) score >4	(14) score >4	(15) score >4	(16) score >4	(17) score >4	(18) score >4
kindergarten intensity	0.000	0.006***	0.003***	0.003***	0.005***	0.004***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	537,558	537,558	537,558	452,486	331,546	451,536
R-squared	0.046	0.047	0.050	0.076	0.077	0.076
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Academic Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE			Yes	Yes	Yes	Yes
Teacher FE				Yes	Yes	Yes
Municipality FE		Yes				
NUTS FE	Yes					

Notes: This table reports the fixed effects estimation results for the 6th grade Portuguese exam. The first table corresponds to the results of having a score higher than 2, the second of having a score higher than 3, and the third of having a score higher than 4. The first column of each table corresponds to the specification including NUTS III fixed effects, the second of including municipality fixed effects, the third of including school fixed effects, the fourth of using the alternative measure of kindergarten intensity (number of kindergartens on individuals' birth year), the fifth of withdrawing the metropolitan areas, and the sixth of including mobility as a control. The same follows for the other scores thresholds.
Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Appendix Table 16: Robustness Fixed Effects Estimation Results - 6th grade, Mathematics Exam

VARIABLES	(1) score >2	(2) score >2	(3) score >2	(4) score >2	(5) score >2	(6) score >2
kindergarten intensity	0.004***	0.066***	0.032***	0.024***	0.043***	0.037***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Observations	537,619	537,619	537,619	460,119	336,530	459,169
R-squared	0.170	0.176	0.190	0.223	0.211	0.223
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Academic Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE			Yes	Yes	Yes	Yes
Teacher FE				Yes	Yes	Yes
Municipality FE		Yes				
NUTS FE	Yes					
VARIABLES	(7) score >3	(8) score >3	(9) score >3	(10) score >3	(11) score >3	(12) score >3
kindergarten intensity	0.004***	0.047***	0.023***	0.016***	0.028***	0.024***
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
Observations	537,619	537,619	537,619	460,119	336,530	459,169
R-squared	0.128	0.133	0.144	0.178	0.171	0.178
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Academic Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE			Yes	Yes	Yes	Yes
Teacher FE				Yes	Yes	Yes
Municipality FE		Yes				
NUTS FE	Yes					
VARIABLES	(13) score >4	(14) score >4	(15) score >4	(16) score >4	(17) score >4	(18) score >4
kindergarten intensity	0.002***	0.012***	0.007***	0.004***	0.008***	0.006***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	537,619	537,619	537,619	460,119	336,530	459,169
R-squared	0.064	0.067	0.072	0.094	0.093	0.094
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Academic Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE			Yes	Yes	Yes	Yes
Teacher FE				Yes	Yes	Yes
Municipality FE		Yes				
NUTS FE	Yes					

Notes: This table reports the fixed effects estimation results for the 6th grade Mathematics exam. The first table corresponds to the results of having a score higher than 2, the second of having a score higher than 3, and the third of having a score higher than 4. The first column of each table corresponds to the specification including NUTS III fixed effects, the second of including municipality fixed effects, the third of including school fixed effects, the fourth of using the alternative measure of kindergarten intensity (number of kindergartens on individuals' birth year), the fifth of withdrawing the metropolitan areas, and the sixth of including mobility as a control. The same follows for the other scores thresholds.
Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Appendix Table 17: Robustness Fixed Effects Estimation Results - 9th grade, Portuguese Exam

VARIABLES	(1) score >2	(2) score >2	(3) score >2	(4) score >2	(5) score >2	(6) score >2
kindergarten intensity	0.008***	0.045***	0.027***	0.025***	0.046***	0.036***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Observations	632,252	632,252	632,252	565,172	413,294	563,813
R-squared	0.112	0.116	0.131	0.158	0.157	0.159
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Academic Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE			Yes	Yes	Yes	Yes
Teacher FE				Yes	Yes	Yes
Municipality FE		Yes				
NUTS FE	Yes					
VARIABLES	(7) score >3	(8) score >3	(9) score >3	(10) score >3	(11) score >3	(12) score >3
kindergarten intensity	0.005***	0.028***	0.015***	0.012***	0.024***	0.019***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)
Observations	632,252	632,252	632,252	565,172	413,294	563,813
R-squared	0.131	0.134	0.144	0.166	0.165	0.167
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Academic Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE			Yes	Yes	Yes	Yes
Teacher FE				Yes	Yes	Yes
Municipality FE		Yes				
NUTS FE	Yes					
VARIABLES	(13) score >4	(14) score >4	(15) score >4	(16) score >4	(17) score >4	(18) score >4
kindergarten intensity	0.000	0.003***	0.002***	0.001	0.002***	0.002***
	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)
Observations	632,252	632,252	632,252	565,172	413,294	563,813
R-squared	0.041	0.042	0.045	0.064	0.066	0.064
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Academic Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE			Yes	Yes	Yes	Yes
Teacher FE				Yes	Yes	Yes
Municipality FE		Yes				
NUTS FE	Yes					

Notes: This table reports the fixed estimation results for the 9th grade Portuguese exam. The first table corresponds to the results of having a score higher than 2, the second of having a score higher than 3, and the third of having a score higher than 4. The first column of each table corresponds to the specification including NUTS III fixed effects, the second of including municipality fixed effects, the third of including school fixed effects, the fourth of using the alternative measure of kindergarten intensity (number of kindergartens on individuals' birth year), the fifth of withdrawing the metropolitan areas, and the sixth of including mobility as a control. The same follows for the other scores thresholds.
Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Appendix Table 18: Robustness Fixed Effects Estimation Results - 9th grade, Mathematics Exam

VARIABLES	(1) score >2	(2) score >2	(3) score >2	(4) score >2	(5) score >2	(6) score >2
kindergarten intensity	0.009***	0.054***	0.030***	0.025***	0.047***	0.037***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Observations	635,235	635,235	635,235	570,841	416,282	569,482
R-squared	0.129	0.136	0.158	0.191	0.182	0.192
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Academic Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE			Yes	Yes	Yes	Yes
Teacher FE				Yes	Yes	Yes
Municipality FE		Yes				
NUTS FE	Yes					
VARIABLES	(7) score >3	(8) score >3	(9) score >3	(10) score >3	(11) score >3	(12) score >3
kindergarten intensity	0.008***	0.031***	0.018***	0.016***	0.025***	0.022***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)
Observations	635,235	635,235	635,235	570,841	416,282	569,482
R-squared	0.133	0.138	0.154	0.182	0.178	0.183
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Academic Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE			Yes	Yes	Yes	Yes
Teacher FE				Yes	Yes	Yes
Municipality FE		Yes				
NUTS FE	Yes					
VARIABLES	(13) score >4	(14) score >4	(15) score >4	(16) score >4	(17) score >4	(18) score >4
kindergarten intensity	0.002***	0.008***	0.005***	0.004***	0.005***	0.006***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	635,235	635,235	635,235	570,841	416,282	569,482
R-squared	0.068	0.070	0.078	0.096	0.098	0.096
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Academic Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE			Yes	Yes	Yes	Yes
Teacher FE				Yes	Yes	Yes
Municipality FE		Yes				
NUTS FE	Yes					

Notes: This table reports the fixed effects estimation results for the 9th grade Mathematics exam. The first table corresponds to the results of having a score higher than 2, the second of having a score higher than 3, and the third of having a score higher than 4. The first column of each table corresponds to the specification including NUTS III fixed effects, the second of including municipality fixed effects, the third of including school fixed effects, the fourth of using the alternative measure of kindergarten intensity (number of kindergartens on individuals' birth year), the fifth of withdrawing the metropolitan areas, and the sixth of including mobility as a control. The same follows for the other scores thresholds.
Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Appendix Table 19: Robustness Fixed Effects Estimation Results - 12th grade, Portuguese Exam

VARIABLES	(1) score >2	(2) score >2	(3) score >2	(4) score >2	(5) score >2	(6) score >2
kindergarten intensity	0.012***	0.055***	0.024***	0.020***	0.029***	0.023***
	(0.001)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Observations	302,114	302,114	302,114	292,970	213,641	291,282
R-squared	0.075	0.080	0.092	0.145	0.147	0.145
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Academic Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE			Yes	Yes	Yes	Yes
Teacher FE				Yes	Yes	Yes
Municipality FE		Yes				
NUTS FE	Yes					
VARIABLES	(7) score >3	(8) score >3	(9) score >3	(10) score >3	(11) score >3	(12) score >3
kindergarten intensity	0.007***	0.032***	0.014***	0.012***	0.015***	0.014***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Observations	302,114	302,114	302,114	292,970	213,641	291,282
R-squared	0.067	0.070	0.077	0.106	0.109	0.106
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Academic Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE			Yes	Yes	Yes	Yes
Teacher FE				Yes	Yes	Yes
Municipality FE		Yes				
NUTS FE	Yes					
VARIABLES	(13) score >4	(14) score >4	(15) score >4	(16) score >4	(17) score >4	(18) score >4
kindergarten intensity	0.000	0.003***	0.001***	0.001**	0.001**	0.001***
	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)
Observations	302,114	302,114	302,114	292,970	213,641	291,282
R-squared	0.010	0.011	0.013	0.027	0.029	0.027
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Academic Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE			Yes	Yes	Yes	Yes
Teacher FE				Yes	Yes	Yes
Municipality FE		Yes				
NUTS FE	Yes					

Notes: This table reports the fixed effects estimation results for the 12th grade Portuguese exam. The first table corresponds to the results of having a score higher than 2, the second of having a score higher than 3, and the third of having a score higher than 4. The first column of each table corresponds to the specification including NUTS III fixed effects, the second of including municipality fixed effects, the third of including school fixed effects, the fourth of using the alternative measure of kindergarten intensity (number of kindergartens on individuals' birth year), the fifth of withdrawing the metropolitan areas, and the sixth of including mobility as a control. The same follows for the other scores thresholds.
Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Appendix Table 20: Robustness Fixed Effects Estimation Results - 12th grade, Mathematics Exam

VARIABLES	(1) score >2	(2) score >2	(3) score >2	(4) score >2	(5) score >2	(6) score >2
kindergarten intensity	0.019***	0.056***	0.027***	0.021***	0.027***	0.027***
	(0.002)	(0.004)	(0.003)	(0.002)	(0.003)	(0.003)
Observations	142,737	142,737	142,737	138,718	103,373	138,121
R-squared	0.078	0.090	0.106	0.155	0.152	0.155
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Academic Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE			Yes	Yes	Yes	Yes
Teacher FE				Yes	Yes	Yes
Municipality FE		Yes				
NUTS FE	Yes					
VARIABLES	(7) score >3	(8) score >3	(9) score >3	(10) score >3	(11) score >3	(12) score >3
kindergarten intensity	0.015***	0.032***	0.018***	0.014***	0.016***	0.018***
	(0.002)	(0.004)	(0.003)	(0.002)	(0.003)	(0.003)
Observations	142,737	142,737	142,737	138,718	103,373	138,121
R-squared	0.087	0.095	0.108	0.145	0.144	0.145
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Academic Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE			Yes	Yes	Yes	Yes
Teacher FE				Yes	Yes	Yes
Municipality FE		Yes				
NUTS FE	Yes					
VARIABLES	(13) score >4	(14) score >4	(15) score >4	(16) score >4	(17) score >4	(18) score >4
kindergarten intensity	0.007***	0.008***	0.006***	0.006***	0.007***	0.007***
	(0.001)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Observations	142,737	142,737	142,737	138,718	103,373	138,121
R-squared	0.051	0.056	0.062	0.088	0.090	0.089
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Academic Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE			Yes	Yes	Yes	Yes
Teacher FE				Yes	Yes	Yes
Municipality FE		Yes				
NUTS FE	Yes					

Notes: This table reports the fixed effects estimation results for the 12th grade Mathematics exam. The first table corresponds to the results of having a score higher than 2, the second of having a score higher than 3, and the third of having a score higher than 4. The first column of each table corresponds to the specification including NUTS III fixed effects, the second of including municipality fixed effects, the third of including school fixed effects, the fourth of using the alternative measure of kindergarten intensity (number of kindergartens on individuals' birth year), the fifth of withdrawing the metropolitan areas, and the sixth of including mobility as a control. The same follows for the other scores thresholds.
Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Appendix Table 21: Fixed Effects Estimation Results - Wages (including covariates)

VARIABLES	(1) wages	(2) wages	(3) wages	(4) wages	(5) wages	(6) wages
kindergarten intensity	0.041***	0.039***	0.008***	0.006***	0.004***	0.006***
	(0.001)	(0.001)	(0.001)	(0.0004)	(0.001)	(0.001)
female	-0.089***	-0.086***	-0.039***	-0.039***	-0.042***	-0.039***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
secondary	0.030***	0.027***	-0.014***	-0.014***	-0.026***	-0.014***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
high education	0.461***	0.455***	0.258***	0.258***	0.273***	0.257***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
permanent contract	-0.067***	-0.061***	-0.007***	-0.007***	-0.017***	-0.007***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
1.worker seniority	0.062***	0.064***	0.069***	0.069***	0.069***	0.069***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
2.worker seniority	0.105***	0.107***	0.102***	0.102***	0.103***	0.101***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
3.worker seniority	0.146***	0.148***	0.134***	0.134***	0.137***	0.133***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
4.worker seniority	0.176***	0.178***	0.158***	0.158***	0.164***	0.157***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)
5.worker seniority	0.198***	0.201***	0.179***	0.179***	0.184***	0.178***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
6.worker seniority	0.214***	0.217***	0.197***	0.196***	0.203***	0.195***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
7.worker seniority	0.228***	0.232***	0.211***	0.211***	0.215***	0.210***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
8.worker seniority	0.241***	0.246***	0.213***	0.213***	0.212***	0.211***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
9.worker seniority	0.238***	0.246***	0.193***	0.193***	0.192***	0.191***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
10.worker seniority	0.237***	0.244***	0.194***	0.193***	0.189***	0.192***
	(0.010)	(0.010)	(0.008)	(0.008)	(0.009)	(0.008)
Constant	1.239***	1.237***	1.263***	1.266***	1.232***	1.264***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	1,488,573	1,488,573	1,488,573	1,488,573	934,392	1,488,275
R-squared	0.166	0.180	0.574	0.574	0.585	0.574
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Seniority FE	Yes	Yes	Yes	Yes	Yes	Yes
Establishment FE			Yes	Yes	Yes	Yes
Municipality FE		Yes				
NUTS FE	Yes					

Notes: This table reports the fixed effects estimation results for wages. The first column corresponds to the specification including NUTS III fixed effects, the second of including municipality fixed effects, the third of including establishment fixed effects, the fourth of using the alternative measure of kindergarten intensity (number of kindergartens on individuals' birth year), the fifth of withdrawing the metropolitan areas, and the sixth of including mobility as a control.
Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 22: Robustness Difference-in-Differences Results - First Approach

VARIABLES	(1) Secondary	(2) Secondary	(3) Secondary2	(4) Secondary2	(5) High educ	(6) High educ	(7) High educ2	(8) High educ2	(9) Wages	(10) Wages
interaction	0.014***	0.017***	0.012***	0.014***	0.005***	0.008***	0.007***	0.012***	-0.001	-0.005***
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0002)	(0.0004)	(0.0003)	(0.001)	(0.001)	(0.001)
Observations	2,522,793	2,811,047	2,422,846	2,667,821	2,522,793	1,123,652	1,653,995	784,814	2,319,819	2,583,170
R-squared	0.165	0.160	0.174	0.168	0.024	0.066	0.050	0.101	0.726	0.728
Birth FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NUTS FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the difference-in-differences robustness estimation results using the first approach, where individuals on the comparison cohort have the same age profile as individuals in the treated cohort, but are observed in a different time period. Secondary corresponds to upper-secondary education (including higher education); Secondary2 corresponds to upper-secondary education (excluding higher education); High education corresponds to higher education; and High educ2 corresponds to higher education compared to having completed upper-secondary education, while the former specifications are compared to not having completed the level of education in question. The first column of each specification withdraws the metropolitan areas from the estimation, and the second controls for population mobility.

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

Table 23: Robustness Difference-in-Differences Results - Second Approach

VARIABLES	(1) Secondary	(2) Secondary	(3) Secondary2	(4) Secondary2	(5) High educ	(6) High educ	(7) High educ2	(8) High educ2	(9) Wages	(10) Wages
interaction	0.004***	0.001	0.003	-0.001	0.003***	0.006***	0.009***	0.013***	-0.035***	-0.001
	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.003)	(0.001)	(0.002)
Observations	301,514	363,408	253,705	288,309	301,514	344,781	142,705	182,909	215,985	266,777
R-squared	0.042	0.079	0.037	0.053	0.037	0.065	0.055	0.046	0.524	0.600
Birth FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NUTS FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports the difference-in-differences robustness estimation results using the first approach, where we observe both comparison and treated cohorts in the same time interval, and as a consequence they have different age profiles. Secondary corresponds to upper-secondary education (including higher education); Secondary2 corresponds to upper-secondary education (excluding higher education); High education corresponds to higher education; and High educ2 corresponds to higher education compared to having completed upper-secondary education, while the former specifications are compared to not having completed the level of education in question. The first column of each specification withdraws the metropolitan areas from the estimation, and the second controls for population mobility.

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

47

Table 24: Robustness IV Results - First Approach

VARIABLES	(1) Secondary	(2) Secondary	(3) Secondary2	(4) Secondary2	(5) High educ	(6) High educ	(7) High educ2	(8) High educ2
secondary	0.447***	0.440***	0.460***	0.487***				
	(0.013)	(0.013)	(0.015)	(0.016)				
high education					0.061***	0.389***	0.200***	0.347***
					(0.018)	(0.015)	(0.020)	(0.019)
Observations	2,319,819	2,583,170	2,319,819	2,583,170	874,411	1,011,978	330,304	387,407
R-squared	0.430	0.442	0.388	0.364	0.325	0.362	0.350	0.396

Notes: This table reports the IV estimation robustness estimation results using the first approach, where individuals on the comparison cohort have the same age profile as individuals in the treated cohort, but are observed in a different time period. Secondary corresponds to upper-secondary education (including higher education); Secondary2 corresponds to upper-secondary education (excluding higher education); High education corresponds to higher education; and High education2 corresponds to higher education compared to having completed secondary education, while the former specifications are compared to not having completed the level of education in question. The first column of each specification withdraws the metropolitan areas, and the second controls for population mobility).

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

Table 25: Robustness IV Results - Second Approach

VARIABLES	(1) Secondary	(2) Secondary	(3) Secondary2	(4) Secondary2	(5) High educ	(6) High educ	(7) High educ2	(8) High educ2
secondary	0.570***	0.383***	0.187	0.357**				
	(0.050)	(0.060)	(0.125)	(0.150)				
high education					0.201***	-0.076	0.409***	0.191***
					(0.076)	(0.091)	(0.060)	(0.070)
Observations	383,958	266,777	383,958	266,777	364,200	252,969	243,241	170,601
R-squared	0.272	0.322	0.148	0.092	0.277	0.126	0.423	0.327

Notes: This table reports the IV estimation robustness estimation results using the first approach, where we observe both comparison and treated cohorts in the same time interval, and as a consequence they have different age profiles. Secondary corresponds to upper-secondary education (including higher education); Secondary2 corresponds to upper-secondary education (excluding higher education); High education corresponds to higher education; and High education2 corresponds to higher education compared to having completed secondary education, while the former specifications are compared to not having completed the level of education in question. The first column of each specification withdraws the metropolitan areas, and the second controls for population mobility).

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

Table 26: CAE Revision 1 Codes

CAE-Rev.1 Codes	Designation
931100	Early childhood education establishments
931200	Basic education establishments
934200	Welfare institutions with in-patient or semi-in-patient care
934300	Welfare institutions without in-patient care

Notes: This table reports CAE codes and their respective designations for Revision 1, which were in place from 1985 until 1994.

Table 27: CAE Revision 2 and 2.1 Codes

CAE-Rev.2 and Rev.2.1 codes	Designation
80101	Pre-school education
80102	Basic education (1st cycle)
85311	Social action for infancy and youth, with accommodation
85313	Social action for the elderly, with accommodation
85321	Social action for infancy and youth, without accommodation
85323	Social action for the elderly, without accommodation
85324	Social action without accommodation, n.e.

Notes: This table reports CAE codes and their respective designations for both CAE Revision 2 and CAE Revision 2.1, which were in place from 1995 until 2002, and from 2003 until 2006, respectively.

Table 28: CAE Revision 3 Codes

CAE Rev.3 Codes	Designation
85100	Pre-school education
85201	Basic education (1st cycle)
85202	Basic education (2nd cycle)
85310	Basic education (3rd cycle) and general secondary
87301	Social action activities for the elderly, with accommodation
87901	Social action activities for infancy and youth, with accommodation
88101	Social action activities for the elderly, without accommodation
88910	Children care activities, without accommodation
88990	Other social care activities, without accommodation
94910	Religious organisations activities
94991	Cultural and recreation associations
94995	Other associative activities

Notes: This table reports CAE codes and their respective designations for CAE Revision 3, which has been in place since 2007.

Table 29: CPP 1985 Codes

CPP-85 Code	Designation
13410	Kindergarten teacher
54020	Children vigilant

Notes: This table reports CPP codes and their respective designations for CNP85, which was in place from 1985 until 1994.

Table 30: CPP 1995 Codes

CPP-95 Code	Designation
332	Kindergarten teacher
5131	Children vigilant

Notes: This table reports CPP codes and their respective designations for CNP95, which was in place from 1995 until 2009.

Table 31: CPP 2010 Codes

CPP-10 Code	Designation
2342	Kindergarten teacher
5311	Children care assistant

Notes: This table reports CPP codes and their respective designations for CPP10, which has been in place since 2010.