DISCUSSION PAPER SERIES

IZA

IZA DP No. 7053

Can't Buy Mommy's Love? Universal Childcare and Children's Long-Term Cognitive Development

Christina Felfe
Natalia Nollenberger
Núria Rodríguez-Planas

November 2012

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Can’t Buy Mommy’s Love? Universal Childcare and Children’s Long-Term Cognitive Development

Christina Felfe
Universität St. Gallen and CESifo

Natalia Nollenberger
Universitat Autònoma de Barcelona

Núria Rodríguez-Planas
IZA, IAE-CSIC and Universitat Pompeu Fabra

Discussion Paper No. 7053
November 2012

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

IZA Discussion Paper No. 7053
November 2012

ABSTRACT

Can't Buy Mommy's Love? Universal Childcare and Children's Long-Term Cognitive Development*

What happens to children's long-run cognitive development when introducing universal high-quality childcare for 3-year olds mainly crowds out maternal care? To answer this question we exploit a natural experiment framework and employ a difference-in-difference approach. We find sizable improvements in children's reading and math skills at age 15, as well as in grade progression during primary and secondary school. Effects are driven by girls and disadvantaged children.

JEL Classification: J13, I28

Keywords: universal high-quality childcare, long-term consequences, cognitive skills

Corresponding author:

Núria Rodríguez-Planas
Visiting Research Fellow
IZA
P.O. Box 7240
53072 Bonn
Germany
E-mail: rodriguez-planas@iza.org

* The authors are grateful to Paul Devereux, Susan Dinarsky, Maria Fitzpatrick, Michael Lechner, Oskar Nordström Skans, Björn Öckert, Xavi Ramos, Uta Schönberg, Anna Sjögren, Anna Vignoles, Conny Wunsch, as well as participants from the V INSIDE-MOVE, NORFACE, and CReAM Workshop on Migration and Labor Economics, the III Workshop on Economics of Education "Improving Quality in Education", the CESifo Area Conference on the Economics of Education, as well as seminars at IFAU, University College Dublin, and University of St Gallen. Natalia Nollenberger acknowledges financial support from the Spanish Agency for International Development Cooperation from the Ministry of Foreign Affairs and Cooperation.

I. Introduction

As Governments on both sides of the Atlantic are rolling back subsidized childcare, many worry about the potentially detrimental consequences for children's development and long-term social and economic inequality. The evidence on the effects of public childcare is, however, meager and focuses mostly on countries with high female labor force participation rates (such as the US and Canada), and with many family-friendly policies (such as Scandinavian countries).[1,2] In these countries, the introduction of universal childcare mainly led to a crowding out of private or informal (non-maternal) care. The effects of such expansions on children's skills are found to be mostly positive, particularly among disadvantaged children.[3] In contrast, a few studies find rather negative effects on children's non-cognitive skills (Baker, Gruber and Milligan 2008, Loeb, *et al.* 2007, Magnuson, Ruhm and Waldfogel 2007).

Nonetheless not much is known when the expansion of high-quality public childcare crowds out maternal care.[4] Given that the direction and magnitude of the effects depend crucially on the relative quality of the counterfactual care, the effects

[1]Recent quasi-experimental evidence on universal childcare and child development includes Cascio (2009), Fitzpatrick (2008), Gormley and Gayer (2005) for the US; Baker, Gruber and Milligan (2008), for Canada, and Datta Gupta and Simonsen (2010), Drange, Havnes and Sandsør (2012) and Havnes and Mogstad (2011) for Scandinavian countries. To the best of our knowledge, Berlinski, Galiani and Gertler (2009) and Dustmann, Raute and Schönberg (2012) are the exception as they investigate such questions for Argentina and Germany, respectively.

[2]This literature complements substantial experimental or quasi-experimental research on the effects of childhood educational programs targeted explicitly at disadvantaged children (for an overview please refer to Blau and Currie (2006)).

[3]See Berlinski, Galiani and Gertler (2009), Gormley and Gayer (2005), and Fitzpatrick (2008) for effects measured during preschool or primary school, Drange, Havnes and Sandsør (2012) for effects measured at the end of mandatory schooling, and Cascio (2009), and Havnes and Mogstad (2011) for effects measured during early adulthood.

[4]Several recent studies evaluate the impact of maternal care on children's development exploiting parental leave expansions. They mostly do not find any significant effect on children's long-run development (Liu and Skans 2010, Rasmussen 2010, Baker and Milligan 2012, Dustmann and Schönberg 2012), with the exception of Carneiro, Løken and Salvanes (2010), who detect some positive effects on educational and labor market outcomes at age 25 in Norway. It is important to highlight that, in contrast with our paper, these papers focus on substituting public childcare by maternal care at a much earlier age (usually within the first 15 months of the child) and the care mode crowded out can be manifold.

of introducing public childcare when it crowds out private or informal care might not necessarily coincide with the effects of introducing public childcare when it crowds out maternal care.[5] The latter scenario is, however, the more relevant scenario in countries with low female labor force participation and insufficient childcare supply.[6] Understanding the effects of introducing universal childcare in such a setting is therefore the main objective of this paper. More specifically, we use a natural experiment framework to analyze whether the introduction of high-quality public childcare for 3-year olds can significantly influence children's cognitive performance by the end of mandatory schooling in a context in which the counterfactual care mode is mainly maternal care.

We focus on an early 1990s reform in Spain, which led to a sizeable expansion of publicly subsidized full-time childcare for 3-year olds. Following the reform, overall enrollment in public childcare among 3-year olds increased from 8.5 percent in 1990 to 42.9 percent in 1997 and to 67.1 percent in 2002. Importantly, this reform had a modest effect on maternal employment, did not affect fertility, and did not lead to a crowding out of private childcare (Nollenberger and Rodríguez-Planas 2011). Moreover, a crowding out of informal care was unlikely as most 3-year olds whose mothers worked prior to the reform were already enrolled in either public or private childcare. As a result, our effects have to be mainly interpreted as the effects of substituting maternal care by public high-quality care. The income effect – either due to a reduction in childcare costs (by crowding out private childcare arrangements) or

[5]One recent paper discussing the relevance of the counterfactual care and the arising effect heterogeneity is the one by Felfe and Lalive (2012).
[6]This includes but is not restrictive to Greece, Ireland, Italy, Japan, West Germany, Spain, Switzerland, and Turkey in the OECD alone.

due to an increase in maternal earnings (by increasing maternal employment) are rather negligible.[7]

The Spanish reform also included a federal provision regarding several quality aspects (such as curriculum, group size, and staff skill composition). While the quality improvements were not exclusive to the children who were directly affected by the expansion of public childcare, it is important to keep in mind that our findings have to be interpreted as the consequences of introducing regulated high-quality care, which can also be compared to preschool targeted at 3-year olds. Thus, the reform under study stands in stark contrast to other reforms such as the reforms in Canada, which implied moving middle class children from home care to relatively poor quality care (Baker, Gruber and Milligan 2008), or in Norway, which did not occur parallel to an overall improvement in childcare quality (Havnes and Mogstad 2011).

Although the reform was national, the responsibility for implementing its preschool component was transferred to the states. The timing of such implementation expanded over ten years and varied considerably across states. We exploit this variation in the treatment intensity to isolate the impact of public childcare on children's educational achievements in the long run. Our main empirical strategy is thus a difference-in-difference (henceforth, DiD) approach and is structured as follows: we compare educational outcomes of children (at age 15) who were 3 years old prior to and after the reform from states where public childcare expanded substantially and states with a less pronounced increase in public childcare in the immediate years after the reform.

[7]In this aspect, our paper contrasts with that of Black *et al.* (2012) in which the authors are able to isolate the effects of childcare subsidies on both parental and student outcomes. Despite negligible effects of childcare subsidies on childcare use and parental labor force participation, they find significant positive effects on children's academic performance in junior high school, suggesting a positive income effect.

The data used in this study stem from the 2003, 2006, and 2009 *Programme for International Student Assessment* (PISA). Children from PISA 2003 were born in 1987 and hence, they were 4 years old when the reform was first implemented in 1991. As a consequence, they were unaffected by the expansion of publicly subsidized childcare for 3-year-old children. Children from PISA 2006 and 2009 were born in 1990 and 1993, respectively, and thus they were affected by the expansion of childcare at age 3.

We first analyze the effects of the reform on maternal employment and private childcare enrollment in order to corroborate earlier findings of a small effect of the reform on maternal employment and no evidence of a crowding out of private childcare. Focusing on the effects of the reform on children's cognitive performance at age 15, we find that universal childcare for 3-year olds led to a sizable increase in reading and math test scores, and a sizable decrease in the likelihood of falling behind a grade. More specifically, we find that the reform improved reading test scores at age 15 by 0.10-0.13 standard deviations and math test scores by 0.07 standard deviations. In addition, the reform reduced the incidence of falling behind a grade by 2.4 percentage points (or 50 percent) in primary school and by 3.2 percentage points (or 13 percent) in secondary school. Our results regarding reading test scores and falling behind a grade in primary school are robust to the use of alternative specifications and alternative identification strategies. Results regarding math test scores and falling behind a grade in secondary school are robust in 3 out of 7 robustness tests. Finally, placebo estimates using month of birth as the dependent variable support the hypothesis that our findings are not spurious.[8]

[8]Unfortunately, information on the state of residence is not available in the 2000 PISA data. Thus, we are unable to perform a placebo test using data prior to the actual time of the reform.

Stratification with respect to gender reveals that girls mainly drive the achievement effects. This gender heterogeneity effect is in line with existing research reporting larger benefits of public childcare for girls than for boys (Gathmann and Sass 2012, Havnes and Mogstad 2011). Stratification with respect to parental education also supports the findings of previous studies that public childcare is particularly beneficial for children from disadvantaged backgrounds (Currie and Thomas 1995, Datta Gupta and Simonsen 2010, Havnes and Mogstad 2011).

Our paper stands in contrast to the existing research on the effects of public childcare on children's development by focusing on a setting where public childcare crowds out mainly maternal care, but neither private nor informal care. Our paper is, however, closer to Datta Gupta and Simonsen (2010) and to Drange, Havnes and Sandsør (2012). Datta Gupta and Simonsen (2010) compare publicly provided childcare at age 3 versus home care in Denmark. They find significant differences in non-cognitive skills among children cared for in public childcare vis-à-vis children cared for in family day care, but no significant differences between children enrolled in preschool and children cared by their mothers, no matter the gender of the child and mother's level of education. However, their paper differs from ours in at least three important ways. First, they focus on non-cognitive skills (as opposed to cognitive skills). Second, their outcomes are measured at age 7 (as opposed to age 15). Perhaps more importantly, their identification strategy only allows them to yield causal estimates for the effect of public childcare versus family day care, not, however, for the effect of public childcare versus maternal care. In a similar vein to our paper, Drange, Havnes and Sandsør (2012) focus on the effect of substituting maternal care by public childcare on children's cognitive outcomes at the end of mandatory schooling. They find no significant effects. In contrast to our paper, their paper focuses on mandatory preschool provided mainly to disadvantaged children at age 6. Thus arising differences between their and our results might stem from effect heterogeneity with respect to age and potential gains from childcare. On the one hand, younger children might benefit more from a program which (like ours) focuses on learning through play. On the other hand, as the authors explain in their paper, the fact that access to childcare was already substantial prior to the reform suggests that parents sort relatively efficiently into the existing kindergarten programs, so that children who are not in such programs in fact may opt out partly because they will benefit little from them.

The paper is organized as follows. The next section provides an overview of the institutional background. Sections 3 and 4 present the empirical strategy and the data, respectively. Section 5 presents the main results, the sensitivity analysis and the heterogeneity analysis. Section 6 concludes.

II. Background Information

Institutional Background

In Spain, female labor force participation rates are among the lowest in the OECD. In 1990, before the reform under analysis came into force, the Spanish female labor force participation rate was 43 percent, far behind the 70 percent of the US, 69 percent of Canada, 73 percent in Norway and 78 percent of Denmark (the countries on which other studies analyzing childcare expansions have focused). In addition, the employment gap due to motherhood amounted to 10 percentage points (Gutierrez-Domenech 2005). Reasons for such a scenario are manifold. In the late 1980s and early 1990s, Spain was *not* a family-friendly country for working parents (and especially working mothers) as reflected by its low levels of social assistance to families (Adserà 2004), one of the shortest maternity leaves in Europe (Ruhm 1998), an extremely low incidence of part-time work (only 8 percent of all jobs in 1990), as well as a rigid labor market with many jobs in the service sectors having a split shift from 9 am to 2 pm and from 5 to 8 pm (Amuedo-Dorantes and de la Rica 2009). Moreover, Spain was a traditional country with low participation of men in household production (Bettio and Villa 1998, de Laat and Sevilla-Sanz 2011). Consistent with this, only half of all married women aged 18 to 45 were working or looking for a job in 1992, and among those not participating in the labor market, 45 percent reported family responsibilities as their main reason.[9]

School and Preschool Prior to the Reform

Mandatory schooling in Spain begins at age 6. However, preschool for 4- and 5-year olds is also offered at the premises of primary schools from 9 am to 5 pm (regardless of school ownership status). Once a primary school offers places for 4-year olds, parents who wish to enroll their children in that particular school will do so when the child turns 4 years old as the chance of being accepted by the school may decrease considerably a year later (as priority for enrollment of 5- and 6-year olds is given to those children already enrolled in that particular school when they were 4 years old). As a consequence, enrollment rates for 4- and 5-year olds in the late 1980s were 94 and 100 percent, respectively.

Primary and secondary schooling is either public or private.[10] Public schools are free of charge, except for school lunch, which costs about €100 per child per month.

[9]Estimated by the authors based on micro data from the 1992 Spanish Labor Force Survey.

[10]About one third of children in primary school in Spain are enrolled in private schools. In this paper, private schools refer to "escuelas concertadas" for which the government subsidizes the staff costs (including teachers). There are a very small number of private schools, which tend to be foreign schools (such as the French, Swiss or American schools), and cost two to three times more than the average "escuela concertada".

Private school costs are higher - between 250 and 350 euros per child per month (including lunch).

At the beginning of the 1990s, childcare for children 0- to 3-years old was rather scarce, predominantly private, and also quite expensive (on average it cost between 300 and 400 euros per child per month - including lunch). In contrast with some Scandinavian countries and the US, family day care, where a reduced number of children are under the care of a certified carer in her house, is practically non-existent. In Spain, children under 4 are either in formal (public or private) childcare or with their mother (or grandmother). Unfortunately, information on grandmother care is unavailable. As a consequence, this paper considers motherly care as equivalent to care provided by the nuclear family.

The Reform

In 1990, Spain underwent a major national education reform (named LOGSE) that affected preschool, primary and high school. The focus of our study is on the preschool component of this reform, which consisted of regulating the supply and the quality of preschool, and its implementation began in the school year 1991/92. The primary and secondary school component of the reform increased mandatory schooling by two years (from age 14 to age 16) starting school year 1996/97. In addition, it established that primary school would end at age 12 (instead of age 14). Our analysis isolates the effect of the preschool component by focusing on children born between 1987 and 1993 who were all affected by the primary and secondary school component but not necessarily by the preschool component.

The LOGSE divided preschool into two levels: the first level included children up to 3 years old, and the second level included children 3 to 5 years old. While not introducing mandatory attendance, the government began regulating the supply of places for 3 year olds. Prior to the LOGSE, free universal preschool education had only been offered to children 4 to 5 years old in Spain. After the LOGSE, preschool places for 3 year olds were offered within the premises of primary schools and were run by the same team of teachers. This implied that childcare for 3 year olds operated full-day (9 am to 5 pm) during the five working days and followed a homogeneous and well-designed program. With the introduction of the LOGSE, schools had to accept children in September of the year the child turned 3 whenever parents asked for admission if places were available. Available preschool places were allocated to those who had requested admission by lottery (regardless of parents' employment, marital status, or income). As explained earlier, although enrollment was not mandatory, it was necessary to ensure a place in the parents' preferred school choice. As a consequence, childcare enrollment among 3-years-old children went from meager to universal in a matter of a decade.[11] Between the academic years 1990/91 and 2002/03 the number of 3-years-old children enrolled in *public* preschool centers increased extraordinarily from 33,128 to 238,709. This represented an increase in the public enrollment rate of 3-year olds by more than 58.6 percentage points, from 8.5 percent to 67.1 percent.[12] Federal funding for preschool and primary education increased from an average expenditure of €1,769 per child in 1990/91 to €2,405 in 1996/97 (both measured in 1997 constant euros), implying a 35.9 percent increase in education expenditure per child.[13]

Apart from regulating the supply of public childcare, the LOGSE also provided federal provisions for the first time in Spain regarding educational content, group size,

[11]Unfortunately we only have information on enrollment rates and not on actual supply rates for 3-year olds. In the context of rationed supply, enrollment rates should, however, resemble coverage rates quite closely.
[12]These figures exclude the Basque Country, Navarra and Ceuta and Melilla as they are not included in our analysis.
[13]Unfortunately data disaggregated at the preschool level is not available.

and staff skill composition regardless of ownership status for children 3 to 5 years old. Psycho-educational theories such as constructivism (put forward by Jean Piaget, and Lev Vygotski) and progressive education (based on Célestin Freinet and Ovide Decroly) served as a guideline for the design of the curriculum. There was a strong emphasis on play-based education, group play, learning through experiencing the environment, problem solving and critical thinking. The particular objectives of preschool education focused on: (1) personal development where the child masters its own body and understands its own movement possibilities; (2) emotional development where the child interacts with others in a variety of context and communication modes; (3) social development where the child forms good relationships with adults and peers and understands that people have different needs, views, cultures and beliefs; and (4) personal autonomy in the child's usual activities (LOGSE; 3 October 1990). While the pedagogical movements behind the LOGSE are close to those in Scandinavian countries, they have been viewed as an alternative to the test-oriented instruction legislated by the *No Child Left Behind* educational funding act in the US or the reception class in the UK.

In addition, the LOGSE established the maximum number of students per class to be 20 for 3-year olds and 25 for 4- and 5-years old. It is important to point out that classes are grouped based on the year children were born and thus, are not mixed in ages.[14] Finally, the LOGSE required preschool teachers to have a college degree in pedagogy – a requirement previously only enforced for teachers of 4- and 5-year olds.

The quality improvements affected all children enrolled in preschool (that is, 3-, 4- and 5- year olds). As a consequence, our analysis should be able to isolate the effect of the expansion of high-quality public childcare from overall quality

[14]As a consequence, we ought not to worry about potential spillover effects from incoming 3-year old children on 4-year old children. This point is important as age-mixed groups would represent a threat to our estimation methodology.

improvements of preschool education of 4- and 5-year olds (all cohorts under study were affected by the improved quality of preschool for 4- and 5-year olds).

Despite being a national law and being financed nationwide, the responsibility of implementing the expansion of public preschool slots was transferred to the states. The timing of such implementation expanded over ten years and varied considerably across states frequently for arbitrary reasons. Implementation lags arose largely due to a scarcity of qualified teachers and constraints on classroom space in existing primary schools (El País October 3rd 2005). In fact, the ratio children per classroom in childcare centers (public and private for children age 0 to 5 years old) was 24.1 among treatment states in 1990 versus 27.2 among control states.[15] Moreover, an initially higher level of private childcare facilities in treatment states in comparison to control states (see Section III for more details) might have provided the necessary know-how to implement childcare facilities at a faster speed.

Our empirical strategy exploits the differences in the timing of implementation across states. Details on our empirical strategy and concerns as to whether there are any further systematic pre-reform differences or differential trends parallel to the expansion of childcare across states that might bias our results are further discussed in the next section.

III. Empirical Specification

Identification

Our empirical strategy follows that of Havnes and Mogstad (2011). We use a DiD strategy which compares the cognitive development at age 15 of children who were 3 years old before and after the reform in states where childcare expanded a lot (the

[15]Calculated by the authors based on statistics from the Spanish Ministry of Education.

treatment group) and in states where the increment in childcare coverage was less pronounced immediately after the reform (the control group). To determine the cut-off needed to define which states belong to the treatment and the control group, we follow Havnes and Mogstad (2011), and order states according to their percentage point increase in public childcare enrollment of 3-year olds in the immediate years after the reform, to be precise between 1990 and 1993. By choosing the initial years after the reform, we aim at capturing the differential expansion in public childcare due to a slackening of initial constraints, and not due to differences in underlying preferences or demand. We then separate the sample at the median. Those states that experienced an increase in public childcare enrollment above the median belong to our treatment group whereas those with an increase under the median belong to our control group.

Figure 1 displays the average increase in public childcare coverage for 3-year olds for the treatment and the control groups from 1987/88 to 2002/03. Prior to the reform, there are few differences, on average, between treatment and control groups: in 1990/91 the enrollment rate of publicly available preschool places for 3-year olds amounted to 9.9 percent in the treatment group and to 7.4 percent in the control group. Yet, families living in treatment states experienced a much stronger initial increase in preschool places than families living in control states. For instance, among states in the treatment group, the public enrollment rate for 3-year olds rose from 9.9 percent in school year 1990/91 to 44.0 percent in the school year 1993/94 and 57.1 percent in school year 1996/97. In comparison, the public enrollment rate for 3-year olds in the control group increased from 7.4 percent to 15.3 percent in 1993/94 to 29.4 percent in school year 1996/97. Figure 1 also shows that while there are dramatic differences in the initial expansion of public childcare, the control states fully catch up within a decade. Figure 2 provides evidence that, in contrast with the observed differences in public childcare, trends in private childcare are remarkably similar across the treatment and control group. As a result, our study compares states that differ distinctly in terms of initial changes in public childcare coverage, not, however, in terms of long-run trends or potential demand for childcare.

Table 1, Panel B.2, provides us with an overview of pre- and post-reform differences in several socio-economic features between treated and control states, including parents' educational level. Although treatment states performed better in terms of several socioeconomic indicators than control states, the differences are not statistically significant. Moreover, we find no substantial differences between pre- or post-reform trends (see Figure A.1 in the Appendix). As such, if any, differences in state features do not represent a threat to our identification strategy. Still, we control for pre-reform state characteristics in our preferred specification. Moreover, we additionally test the robustness of our results to these differences by restricting our sample to only states that were very similar prior to the reform in terms of such observables (see Section V for details).

Implementation

Our basic DiD model, estimated by OLS over the sample of children from PISA 2003, 2006 and 2009, can be expressed as follows:[16]

$$\begin{aligned} Y_{ijt} &= \gamma_1 Treat_i + \gamma_2 Cohort90 + \gamma_3 Cohort93 + \theta_1 (Treat_i * Cohort90) + \\ &+ \theta_2 (Treat_i * Cohort93) + X_i^{'}\beta + Z_j^{'}\delta + \varepsilon_{ijt} \end{aligned} \tag{1}$$

where Y_{ijt} measures the educational outcome a child i achieves in year t living in state j, $Treat_i$ is a binary variable indicating whether or not child i lives in one of the fast-

[16]We use OLS for all of our estimations. For our limited-dependent-variable outcomes we replicate our analysis using logit models, which yield similar results.

implementing states. *Cohort90* and *Cohort93* are cohort-specific dummies equal to 1 if the child is tested in PISA 2006 or in PISA 2009, respectively. Children from PISA 2006 and 2009 were born in 1990 and 1993, respectively, and thus they were fully affected by the early childhood component of the LOGSE at age 3 *if they lived in a state that rapidly expanded the supply of public childcare slots* (that is, in a treatment state). Children from PISA 2003 were born in 1987. They were 4 years old when the LOGSE was first implemented during the school year 1991/92, and thus, they were unaffected by the expansion of publicly subsidized childcare.

The coefficients θ_1 and θ_2 belonging to the interaction terms between treated states and the cohort dummies for 1990 and 1993, respectively, measure the average causal effect of the increase in childcare places for 3-year olds in the treatment states relative to the control states between 1990/91 and 1993/94 as well as 1990/91 and 1996/97, respectively, on different outcomes measuring children's cognitive development at age 15.

The vector X_i includes *only* time-invariant individual features that are expected to be correlated with educational achievement: gender and immigrant status. Since all additional socio-demographic characteristics that we observe at age 15 are time variant and thus potentially endogenous to our treatment, we decided not to include them in our main specification. However, our results are robust to sensitivity analysis where we sequentially add these additional variables to equation (1).[17] Furthermore, the vector Z_j includes pre-reform state-specific features, such as GDP per capita, unemployment rate, female employment rate, average educational level, population density, that may affect individuals' educational outcomes. In addition, in a separate specification we allow for pre-reform heterogeneity within the group of treatment

[17]These additional controls include parent's education level, an index of home possessions, whether the school is in an urban or rural area, and whether the school is private or public. Results of the specification including these control variables are shown in Appendix Table A.1.

states by estimating a specification where we include state-specific fixed effects. All robustness checks are presented with either specification. As the coefficients vary little across these two specifications, we use the former specification to discuss the main findings of the paper.

The DiD strategy controls implicitly for all average time-constant differences between children living in different locations (by including a dummy for the treatment areas) and in different years (by including a dummy for the different cohorts). Yet, it assumes that in the absence of the reform children residing in the treatment states would have experienced the same change in outcomes as children residing in the control states. While we already provided evidence that there are no differential trends in observable state specific features (see Appendix Figure A.1 and the previous section on Identification), we need to assume that there are no differential time trends in any further unobservable state-features which systematically relate to the determinants of the expansion in public childcare and at the same time explain differential development in children's cognitive development. Appendix Figure A.2 plots time trends of commonly used education quality variables in treatment and control states, namely children/staff ratios at public and private primary schools, as well as the proportion of children enrolled in private centers in primary school. These trends show that there was no differential improvement in these quality indicators in the primary school affecting differentially the treatment and the controls states. In addition, to strengthen the credibility of this assumption, Section V provides a battery of sensitivity checks.

IV. Data

The data used for this study stem from the *Programme for International Student Assessment* (PISA), an internationally standardized assessment that was jointly developed by participating economies and administered to 15 year olds in schools. PISA tests whether students near the end of compulsory education have acquired the knowledge and skills essential for successful participation in the labor market. In particular, it administers specific tests to assess whether students can analyze, reason, and communicate effectively. For reporting students' performance in each domain, PISA uses plausible values (henceforth PV). In all of our analysis we use PV and follow the OECD recommendations (OECD 2009) that involve estimating one regression for each set of PV (there are five PV to each domain) and then report the arithmetic average of these estimates.

For our purpose, we rely on the 2003, 2006 and 2009 PISA datasets for Spain. Thus, our sample consists of children belonging to the 1987, 1990, and 1993 birth cohorts. We exclude immigrant children who arrived to Spain *after* their 3rd birthday as well as children residing in the Basque Country, Navarra and Ceuta and Melilla. The reason for doing so is that the Basque Country and Navarra have had greater fiscal and political autonomy since the mid-1970s and, as a consequence, their educational policy has differed from that of Spain as a whole. Data on children living in Ceuta and Melilla are only available in the PISA datasets from 2006 onwards.

Our analysis focuses on reading and mathematics as performance in these two domains are fully comparable across PISA cycles from 2003 onwards. Questions entering the scientific scores are not comparable before and after 2006 and thus are not included in our analysis (OECD 2006). Test scores are standardized implying that the estimated coefficient represents the percentage increase (or decrease) in standard deviations (henceforth sd).[18] We also estimate the effect of the reform on two additional variables, available only in the 2003 and 2009 PISA waves: falling behind a grade during primary school or secondary school.

Table 1 provides mean comparisons with respect to outcome and control variables of children living in treatment and control states before and after the reform. Regarding the performance in the PISA tests, children in the treated states outperform those in the control states already prior to the reform. After the reform, the performance gap across treatment and control groups widens further. This improvement is suggestive that the childcare reform in Spain may have increased children's cognitive development. In contrast, we do not observe any statistically significant differences, neither prior to nor after the reform, in children's relative age, our placebo outcome.[19]

At the bottom of Table 1 we can also find summary statistics for children's time-invariant socio-demographic characteristics, which may be correlated with children's cognitive development. If the composition of pre-reform socio-demographic characteristics changes over time, the DiD estimates may be biased. Based on balancing tests, we can, however, reject any statistically significant difference across the three cohorts at the 95 percent level.

Finally, we would like to point out that, similar to related studies in this field (Baker, Gruber and Milligan 2008, Fitzpatrick 2008, Havnes and Mogstad 2011), our estimates are intention-to-treat (ITT) estimates only. Unfortunately, PISA does not

---

[18] Standardization is done by subtracting the mean from each individual test score and dividing by the standard deviation for the whole sample. We have conducted sensitivity analysis where the test scores have been standardized at the year level. In this case θ_1 and θ_2 are estimating the causal effects of the policy on the relative position of treatment states versus control states within a year, eliminating any potential problems with testing differences across years. Results are very similar to those from our preferred specification and available from the authors upon request.

[19] Following Bedard and Dhuey (2004), the relative age is defined as the difference between the month of birth and the cut-off date for children to begin school. As in Spain the cut-off date is January 1, the relative age is equal to 0 for students born in the December (the youngest) and equal to 11 for students born in January (the oldest).

provide information on pre-school attendance at specific ages and thus we cannot obtain estimates for individual treatment. Yet, following Baker, Gruber and Milligan (2008), we provide estimates for the treatment effect dividing the ITT estimates by the average increase in childcare slots in the treated states relative to the control states.

V. Results

Effects on Private Childcare Enrollment and Maternal Employment

When discussing the impact of expanding public childcare on children's long-run cognitive development, it is important to bear in mind whether the expansion in public childcare led to a crowding out of alternative care modes. We therefore first discuss the changes in public and private childcare as well as in maternal employment that arose after the introduction of the LOGSE. For the purpose of the latter, we re-estimate the results by Nollenberger and Rodríguez-Planas (2011), but adjust the identification strategy to be comparable to the baseline strategy of this paper (for details please refer to the Appendix). Results are shown in Panel B in Table 2.

Table 2 shows that children residing in treatment states were offered substantially more public childcare than children residing in control states: this differential increase amounted to 26.1 percentage points for the 1993/94 cohort and to 25.6 percentage points for the 1996/97 cohort[20]. Yet, the reform did not lead to the crowding out of private childcare enrollment (Panel A). While this result may come as a surprise, it is important to highlight that preschool for 3-year olds was implemented within primary school regardless of school ownership. As a consequence parents who wished to enroll their children in private school would now

[20]Given a similar relative increase in public childcare among the 1993/94 and 1996/97 cohort and assuming a constant treatment effect across different levels of childcare supply we should expect a similar magnitude for both treatment effects (reflected by θ_1 and θ_2).

enroll their 3-year old to the private school as soon as preschool for that age group was offered (to guarantee a space thereafter).

Table 2 also shows that the effect of universal childcare on maternal employment is much smaller than the increase in the enrollment in childcare (Panel B). A 1 percentage point increase in enrollment among 3-year olds led to between 0.06 and 0.09 percentage points increase in maternal employment.[21,22] While this finding may seem puzzling at first, in light of Spain being characterized by the male-bread-winner model (as described in Section II), it does make much sense.

Finally, it is important to note that, in contrast to other studies, the expansion in public childcare did not lead to a crowding out of informal care arrangements. Most of the mothers of 3-year olds who worked prior to the reform had their child already enrolled in either public or private childcare. Prior to the reform, 35.7 percent of mothers of 3-year olds worked in treated states while 32.5 percent of 3-year olds were enrolled in formal care (9.9 percent in public childcare and 22.6 percent in private childcare).[23] Thus, the Spanish reform mainly implied that mothers took their children to full-time (9 am to 5 pm) childcare even though they continued *not* to work.

Taken together, our findings have to be interpreted as the effects of an expansion in high-quality public childcare that mainly led to a crowding out of maternal care, but not to a crowding out of private or informal care arrangements. This implies that our estimates measure mainly the effects of offering universal high-quality childcare, as the reform under analysis did not imply a major income shock

---

[21]This estimate is the ratio between the percentage point increase in maternal employment rate (0.024 and 0.016) and the percentage point increase in 3-year olds' public childcare enrollment due to the reform (0.261 and 0.256).

[22]Due to a different identification strategy this estimate is slightly different to that of Nollenberger and Rodríguez-Planas (2011).

[23]This pre-reform situation contrasts with that of Havnes and Mogstad (2011) as in their study, childcare coverage (10 percent) was half the size of maternal employment (20 percent).

due to a shift from private to public childcare. Moreover, any potential income effect from an increase in maternal employment caused by the reform is modest at most.

Effects on Children Cognitive Development

Table 3 shows the impact of expanding public childcare on all children living in the treatment area – the so called intention to treat effect (ITT) - and on the average child placed in public childcare following the expansion of public childcare – denoted by the treatment effect on the treated (TT). The latter estimate is obtained by dividing the ITT estimates by the probability of treatment - shown in Table 2 in Panel A.[24]

Table 3 displays the results for four alternative outcome variables: test scores in reading and math, as well as the likelihood of falling behind one grade in primary and secondary school. All regressions are estimated first without any control variables and then controlling for pre-reform individual and state characteristics. If the underlying assumption is correct - there are no individual or regional particularities that drive the expansion in childcare - additional controls should only improve the efficiency of the estimates by reducing the standard errors of the regression but they should not generate a sizeable change in the policy coefficient. Comparing the respective estimates in Table 3 reveals no significant differences and thus provides a robustness check for the underlying assumption. We therefore focus our discussion on this last specification. Notice that we have also estimated a specification in which instead of controlling for pre-reform state characteristics we include state FE (see Table 3). Doing so allows for pre-treatment heterogeneity across states.

Focusing first on the effects of the reform on children's standardized reading test scores at age 15, the effect of the expansion in public childcare for 3-year olds is positive and statistically significant at any conventional significance level. The

[24]Hence, we divide the 2006 (2009) ITT estimates by the increase in childcare coverage between 1990/91 and 1993/94 (1996/97) in treatment states relative to the controls states (26.1 percentage points in 2006 and 25.6 percentage points in 2009).

expansion of public childcare places leads to an increase in reading test scores by 0.13 sd and 0.10 sd for the 1990 and 1993 cohorts living in one of the treated states, respectively. In terms of children who actually attended public childcare following the reform, the effects are substantial: the TT estimates indicate that the reform implied an improvement in reading scores of 0.48 sd and 0.37 sd for the children born in 1990 and 1993 who attended public childcare, respectively.

The reform also improved children's math performance, yet to a slightly lower extent. Children who were born in 1990 and lived in one of the treated states outperformed children who lived in one of the control states in the math test by 0.07 sd - the effect is, however, only significant at the 90 percent significance level. This translates into an improvement among children who actually attended childcare by 0.24 sd. Yet, among the 1993 cohort the estimate is considerably smaller and no longer statistically significant.

How do these effects compare to the established evidence? The existing studies evaluating the impact of universal childcare provision find effects of similar direction and size although measured at an earlier age. In the case of an Argentinean reform, Berlinski, Galiani and Gertler (2009), find a substantial improvement of cognitive skills (the ITT estimates amounts to 0.23 sd) among children in third grade. Analyzing the consequences of the introduction of universal childcare in Georgia (USA) on the reading and math skills among children in fourth grade, Fitzpatrick (2008) finds slightly lower effects and only for the population of disadvantaged children, defined as those living in rural areas. More specifically, she finds gains from the childcare reform ranging between 0.07 and 0.12 sd on reading scores, and between 0.06 and 0.09 sd on math scores. Studies that have investigated the effects of individual childcare attendance, in contrast to the overall effects of childcare expansions, support the findings for the improvements in reading and math skills among primary school age children (Loeb, *et al*. 2007), that may, however, dissipate over time (Magnuson, Ruhm and Waldfogel 2007).

Moving to the effects of the reform on the likelihood of falling behind a grade, we also find beneficial effects of the reform. More specifically, we observe that the reform reduced the incidence of falling behind a grade by 2.4 percentage points in primary school and 3.2 percentage points in secondary school (these effects are significant at the 95 and 90 percent level, respectively). Given the initial likelihood of falling behind a grade among children in the treated states of 5 percent in primary school and 23 percent in secondary school, the effect of the reform represents a substantial decrease in the incidence of retention (50 percent in primary school and 13 percent in secondary school). The large effects of the reform on falling behind a grade for the treated is suggestive that the policy was particularly effective for those at highest-risk. Indeed, in the heterogeneity analysis we find that children from less advantaged families benefit the most.

The two existing studies that look at the consequences of universal childcare provision on this outcome are US studies: Fitzpatrick (2008) and Cascio (2009). Our results are similar to those found by Fitzpatrick (2008) for disadvantaged children. In fact, analyzing universal Pre-Kindergarten in Georgia, she finds that the probability of being on-grade for their age increases by 7 percentage points (about 10 percent) among black children. In contrast, Cascio (2009) did not find any significant improvements on grade retention.

Specification Checks

We address several potential sources of bias below. In particular we discuss issues such as selective migration, as well as alternative specifications, which allow us to assess the underlying common trend assumption.[25]

Selective migration: One potential source of bias might be selective migration: families might have moved from slow implementing states to fast implementing states. Since PISA only provides information on the state of residence at age 15 (but not at age 3), we rely on the LFS (now for years 2003, 2006 and 2009) to assess the concern of selective migration. We first assess the likelihood of living at age 15 in a different state than the state of birth. This probability is, however, small (4.6 percent in 2003, 5.2 percent in 2006, and 4.9 percent in 2009). Second, we estimate the likelihood of having migrated from a control state to a treated state (and vice versa). The results do not indicate an increased migration into treated states, if anything a small decrease (by 1 percentage point) among the 1990 cohort after the reform (shown in the Table 4, Panel A). Thus, selective migration ought not to be a major threat for the internal validity of the study. [26]

Common Trend Assumption: The strongest assumption underlying any DiD estimation is the absence of any differential time trends in treatment and control states. The most commonly used test to shed some light on this assumption, besides inspecting pre-existing trends, is to estimate the effect of a placebo reform pretending that the reform took place at an earlier moment in time. Unfortunately, we do not

[25]Most results are shown in Tables 4 and 5. In the Appendix Table A.2 we also present the same robustness checks displayed in Table 5, but controlling for states fixed-effects instead for pre-reform state specific characteristics.

[26]In addition, the migration flow by skill level are similar to the ones presented in the table and do not indicate any migration flows that would threaten our identification strategy.

possess sufficient cohorts unaffected by the reform to perform such a placebo test.[27] We therefore rely on one available measure, which is directly related to cognitive development, but is determined prior to the policy change and thus should be uncorrelated with the policy change: child's month of birth.

The impact of the relative age - defined as the difference between the month of birth and the official cut-off date for children to start school – on cognitive test scores is well documented across many countries, with the youngest children in each academic year performing more poorly, on average, than the older members of their cohort (Bedard and Dhuey 2004, Puhani and Weber 2008). Indeed, we can document a significant and positive correlation between children's relative age and children's test scores net of individual background characteristics in our dataset. In addition, the literature has documented a strong correlation with children's socio-economic background and month of birth (Buckles and Hungerman, *forthcoming*). Thus, in case our estimation is biased due to omitted confounders, this should be picked up in a regression using month of birth as the dependent variable.

Re-estimating equation (1) but using children's relative age as the dependent variable does not reveal any significant impact of the policy change on children's relative age (see Panel B in Table 4). This result provides some evidence that unobserved heterogeneity correlated with cognitive development should not be driving our results and our estimates are true policy impacts.[28]

[27]Unfortunately, information on the state of residence is not available in the 2000 PISA data. Thus, we are unable to perform a placebo test using data prior to the actual time of the reform. Alternative datasets, such as TIMMS or LFS, also do not allow for performing placebo checks on pre-reform cohorts. TIMMS, the only other dataset providing information on students' test scores in Spain, does not assess children of comparable grades across years. The Spanish LFS would only allow us to infer on overall grade retention, but not for grade retention during primary or secondary school separately.

[28]The effect remains statistically insignificant if we split the sample by whether youths were born in the first half or the second half of the sample, suggesting that the lack of effects is not driven by heterogeneity across age.

In addition, we estimate a specification where we use a more homogenous sample of states and exclude the poorest and the richest states from our sample. In doing so, we want to address the fact that Spanish regions differ strongly in their economic development and thus might potentially follow rather differential time trends. Results are fairly robust to this sample restriction (see Table 5, column 2).

Moreover, following Duflo (2001), we estimate a specification in which we interact cohort FE with all the pre-reform regional characteristics shown in Panel 2 in Table 1. In so doing, we check if regional pre-reform characteristics are correlated with the development of children's cognitive skills over time. Results (displayed in Column 3 of Table 5) are robust to this specification and provide further supportive evidence for the underlying assumption of common time trends (at least in terms of observables).

Further Robustness Checks: Finally, we carried out the following additional sensitivity checks. First, to account for the fact that some states (Andalucía, Canary Island, Catalonia, Valencia and Galicia) had certain control over their education policy, we have re-estimated a specification adding a dummy for these 5 states and interacting such dummy with the cohort dummies. Second, as in Havnes and Mogstad (2011) we experiment with a different definition of the treatment and the control groups where treatment states are defined as those states with growth in public enrollment above the 67th percentile, and control states are those states with growth in public enrollment below the 33th percentile. Again, results are robust to both of these sensitivity checks (see Table 5, columns 4 and 5, respectively).

To take into account the double stratification nature of the sampling design employed by PISA, we run our regression applying the Fay's Balanced Repeated Replicated (BRR) methodology. Notice that this method implies a clustering at the school level. Results are shown in Table 5, Column 6. While the BRR imputation does not affect much the significance of the results related to reading and falling behind in primary school, it does lead to a substantial increase of the standard errors of the estimates for math skills and falling behind in secondary school.

Finally, Appendix Table A.1 explores the sensitivity of our results to sequentially adding other (potentially endogenous) individual characteristics, such as family characteristics (parents' level of education and home possessions), type of school, and population density of the area of residence. Again, our results are robust despite the covariates included.

Alternative Identification: One alternative identification is to exploit the regional and time variation in the number of public preschools slots for 3-year olds in each state and year and thus, to use a rather continuous measure of the expansion in public childcare. As numbers of slots available in public preschool are *not* available for detailed age groups, we employ a proxy: the number of public preschool units available for 3-5 year olds. [29]

We follow Berlinski and Galiani (2007) and estimate the effect of offering one additional seat in public childcare estimating the following equation by OLS:

$$Y_{ij(t+12)} = \theta Seats_{jt} + \alpha_1 Cohort90 + \alpha_2 Cohort93 + X_i^{'}\beta + Z_j^{'}\delta + \varepsilon_{ijt(t+12)} \qquad (2)$$

where $Y_{ij(t+12)}$ is a measure of educational achievement for the individual i living in the state j when he/she is 15 years old (12 year after being affected by the policy), $Seats_{jt}$ is the number of public preschool seats per 100 for children age 3 to 5 years old in the year t in the state j. The vector X_i includes time invariant individual characteristics (gender and immigration status). We also include regional controls

[29]As Berlinski and Galiani (2007) we estimate the proportion of public preschool seats offered in each state as the number of public preschool units available for 3-5 year olds in each region times the average size of the classroom divided by the population of 3 to 5 year olds in each state. Unfortunately, these data are not available by detailed age group.

(Z_j) and cohort fixed-effects (*Cohort90* and *Cohort93*). This specification has the advantage that it does not rely on the definition of treatment status. However, it assumes a constant effect of offering one further childcare slot across the whole offer distribution. Thus, offering an 11th seat for every hundred children is assumed to have the same effect as the 91st seat per hundred children. Estimates from equation (2) are shown in Table 6.

Offering one more slot per hundred children leads on average to a statistically significant improvement in children's test' scores of 0.01 sd in readings and of 0.004 sd in math, and a statistically significant reduction in the likelihood of falling behind a grade while in primary school of 0.2 percentage points. Given the increase in public enrollment rate of 26.1 percentage points for the 1990 cohort and 25.6 percentage points for the 1993 cohort, this implies an improvement of 0.26 sd and 0.10 sd in reading and math test scores, respectively, and a reduction of about 5.2 percentage points in the likelihood of falling behind primary school. Yet, there does not seem to be a statistically significant effect on falling behind a grade during high-school whereas the effect on math scores loses statistical significance when standard errors are estimated using a more stringent method.

Heterogeneity and underlying mechanisms

Table 7 displays ITT estimates by children's gender and parents' educational level. Such analysis might reveal policy relevant effect heterogeneity. The lack of information on childcare usage across the different subgroups constitutes, however, a serious limitation for obtaining TT estimates for subgroups. As explained by Baker, Gruber and Milligan (2008) and Havnes and Mogstad (2011), assuming identical childcare take-up rates across different subgroups underestimates the effect of the reform on children's cognitive development for those who are less likely to react to the policy. We therefore abstain from reporting TT estimates for subgroups.

Gender: Estimates from Panel A in Table 7 reveal that universal preschool provision had large, positive and significant effects on girls' cognitive development. We observe a significant improvement in reading skills by 0.12 and by 0.15 sd among the 1990 and 1993 female cohort, respectively. Math test scores also increase by 0.09 and 0.10 sd, respectively. We also find positive and significant effects (at the 90 percent level) on grade retention among girls: girls in the treated states are 2.4 percentage points (50 percent) less likely to fall behind a grade during primary school and 4.5 percentage points (23.7 percent) during secondary school (yet, only the latter coefficient is significant at the 90 percent significance level). For boys, we can only observe a statistically significant improvement in their reading skills. Yet, while the improvement in reading is comparable to that of girls in the 1990 cohort, the effect is cut by more than half and is no longer statistically significant among the 1993 cohort.

Our results speak to previous findings regarding the gender gaps in reading and math skills (Guiso, *et al.* 2008, Fryer and Levitt 2010) and suggest that early preschool exposure can help closing the gender gap in math scores - girls fare generally worse in math -, but not the gender gap in reading - boys fare generally worse in reading. This gender asymmetry in returns to public childcare has already been found in previous studies. Gathmann and Sass (2012), for instance, find that attending public childcare improves girls´ early development of socio-motor skills, but has no effect on their language skills. In the study by Havnes and Mogstad (2011), improved labor market outcomes due to an expansion of public childcare are also only present among women (although they find that both men and women benefit similarly in terms of educational outcomes, such as secondary school completion or college attendance).

Parental Education: Panel B in Table 7 presents results by parents' educational level. Average gains in cognitive performance due to universal childcare seem to be driven by children of low-skilled parents, defined as those for whom neither parent has a secondary school degree.[30] Overall, estimates are much larger for children of low-skilled parents. To be more precise, among low-skilled families, we observe a significant improvement in reading skills by 0.13 and 0.11 sd among the 1990 and 1993 cohort, respectively (yet, the latter estimate is not significant at any conventional significance levels). In addition, we also find positive and significant effects (at the 90 percent level) on grade retention during primary school: children in the treated states are 4 percentage points (59.7 percent) less likely to fall behind a grade during primary school. In contrast, no statistically significant effects are found among children with high-skilled parents.

These results are again consistent with those found by others. Fitzpatrick (2008), for instance, only found substantial effects of the introduction of universal pre-Kindergarten on disadvantaged children residing in small towns and rural areas. Similarly, Havnes and Mogstad (2011) have shown that universal childcare provision has positive long-run effects on the income distribution and equality.

VI. Conclusion

A fervent debate in Europe is the extent to which the Government must provide sufficient, affordable childcare. For instance, on June 6, 2012 the German

[30]Since our measure is self-reported by the child (not the parent) and measured at aged 15, we measure parents' education using the educational degree of both parents. Doing so allows us to minimize endogeneity and measurement error problems. This classification divides the sample by about half, which is not far from population estimates from the Labor Force Survey.

Government approved a bill to give parents of toddlers an allowance for keeping their children *out* of state-run day care instead of investing in the expansion and quality of childcare centers (The New York Times June 6, 2012). At the same time, in countries hard hit by the Great Recession, many governments are rolling back subsidized childcare (*The New York Time*s, June 6, 2012, *El País,* July 4, 2012). A major concern among deterrents of public childcare is its high costs for a non-mandatory service for which the short- and long-term gains on the children's development relative to other forms of early childcare (such as parental, informal, or private care) remain uncertain.

Nonetheless, there is still limited consensus in the literature about the effect of childcare on child development partly as the effects of universalizing childcare depend on the quality of both public childcare and the counterfactual care mode. Almond and Currie (2011) go even one step further and warn from drawing conclusions from findings in one setting to potential effects in another setting.

This paper contributes to closing this gap in the literature by providing quasi-experimental evidence for the impact of shifting hours of care provided by mothers to hours of care provided by high-quality public preschools. We find that high-quality public childcare does not only neutralize potentially negative effects of maternal employment, but has even positive effects on children's cognitive development, at least among children with less educated parents and for girls. Hence, these early childhood investments may well pay off themselves in the long-run.

One crucial feature of the childcare expansion under study is, however, the guarantee of maintaining high-quality care. In the absence of quality regulations, a rapid expansion of universal care may well have negative consequences on children's development, at least in the short-run (Baker, Gruber and Milligan 2008). Hence, sending children to public childcare may indeed be one way to “buy mommy’s love”, but only if the quality of care provided in the childcare centers meets the quality of care provided by the mother.

References

Adserà, Alicia (2004). "Changing Fertility Rates in Developed Countries. The Impact of Labor Market Institutions." *Journal of Population Economics*, 17: 17-43.

Almond, Douglas, and Janet Currie (2011). "Killing Me Softly: The Fetal Origins Hypothesis." *Journal of Economic Perspectives*, 25(3): 153-172.

Amuedo-Dorantes, Catalina, and Sara de la Rica (2009). "The Timing of Work and Work-Family Conflicts in Spain: Who Has a Split Work Schedule and Why?" IZA Discussion Paper 4542.

Baker, Michael, Jonathan Gruber, and Kevin Milligan (2008). “Universal child care, maternal labor supply, and family well-being.” *Journal of Political Economy*, 116(4): 709–745.

Baker, Michael, and Kevin Milligan (2012). "Maternity Leave and Children's Cognitive and Behavioral Development". NBER Working Paper 17105.
Bedard, Kelly, and Elizabeth Dhuey (2004). "The Persistence of Early Childhood Maturity: International Evidence of Long-Run Age Effects." *The Quarterly Journal of Economics*, 121(4): 1437-72.

Berlinski, Samuel, and Sebastián Galiani (2007). "The effect of a large expansion in pre-primary school facilities on preschool attendance and maternal employment." *Labour Economics*, 14: 665-680.

Berlinski, Samuel, Sebastián Galiani, and Paul Gertler (2009). "The Effect of Pre-Primary Education on Primary School Performance." *Journal of Public Economics*, 93(1-2): 219-34.

Bettio, Francesca, and Paola Villa (1998). "A Mediterranean perspective on the breakdown of the relationship between participation and fertility." *Cambridge Journal of Economics*, 22(2): 137-171.

Black, Sandra E., Paul J. Devereux, Katrine Vellesen Løken, and Kjell G. Salvanes (2012). "Care or Cash? The Effect of Child Care Subsidies on Student Performance." IZA Discussion Paper 6541.

Blau, David, and Janet Currie (2006). *Preschool, Day Care, and After School Care: Who's Minding the Kids?* Edited by Eric Hanushek and Finis Welsh. Handbook of Economics of Education.

Buckles, Kasey, and Daniel M. Hungerman (*forthcoming*). "Season of Birth and Later Outcomes: Old Questions, New Answers." *The Review of Economics and Statistics.*

Carneiro, Pedro, Katrine Vellesen Løken, and Kjell G. Salvanes (2010). "A Flying Start? Long Term Consequences of Maternal Time Investments in Children During Their First Year of Life". IZA Discussion Paper 5362.

Cascio, Elizabeth (2009). "Do Investments in Universal Early Education Pay Off? Long-term Effects of Introducing into Public Schools." NBER Working Paper 14951.

Currie, Janet, and Duncan Thomas (1995). "Does Head Start Make A Difference?" *The American Economic Review*, 85(3): 341-364.

Datta Gupta, Nabanita, and Marianne Simonsen (2010). "Non-cognitive child outcomes and universal high quality child care." *Journal of Public Economics*, 94(1-2): 30-42.

de Laat, Joost, and Almudena Sevilla-Sanz (2011). "The Fertility and Women's Labor Force Participation puzzle in OECD Countries: The Role of Men's Home Production." *Feminist Economics*, 17(2): 87-119.

Drange, Nina, Tarjei Havnes, and Astrid M. J. Sandsør (2012). "Kindergarten for all: Long run effects of a universal intervention." Statistics Norway Discussion Paper 695.

Duflo, Esther (2001). "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." *Amercian Economic Review*, 91(4): 795-813.

Dustmann, Christian, Anna Raute, and Uta Schönberg (2012). "Does Universal Child Care Matter? Evidence from a Large Expansion in Pre-School Education." Mimeo.

Dustmann, Christian and Uta Schöenberg (*forthcoming*). "The Effect of Expansions in Maternity Leave Coverage on Children Long-term Outcomes". *American Economic Journal: Applied Economics.*

El País (October 3rd 2005). "La LOGSE, 15 años después" by Elena Martín Ortega.

—. (July 4, 2012). "Para el mileurista llevar al niño a la guardería es un lujo", by Carmen Pérez-Lanzac.

Felfe, Christina, and Rafael Lalive (2012). "Child Care and Child Development. What Works for Whom?. Mimeo.

Fitzpatrick, Maria D (2008). "Starting School at Four: The Effect of Universal Pre-Kindergarten on Children's Academic Achievement." *The B.E. Journal of Economic Analysis & Policy*, 8(1): 1-38.

Fryer, Roland G. Jr, and Steven D. Levitt (2010). "An Emprical Analysis of the Gender Gap in Mathematics." *American Economic Journal: Applied Economics*, 2 (2): 210-240.

Gathmann, Christina, and Björn Sass (2012). “Taxing Childcare:Effects on Family Labor Supply and Children.” IZA Discussion Paper 6640.

Gormley Jr, William T., and Ted Gayer (2005). "Promoting School Readiness in Oklahoma. An Evaluation of Tulsa's Pre-K Program." *Journal of Human Resources*, 3: 533-558.

Guiso, Luigi, Fernandino Monte, Paola Sapienza, and Luigi Zingles (2008). "Culture, Gender and Math." *Science*, 320 (5580) 1164-1165.

Gutierrez-Domenech, Maria (2005). "Employment Transitions after Motherhood in Spain." *Review of Labour Economics and Industrial Relations* 19: 123–148.

Havnes, Tarjei, and Magne Mogstad (2011). “No Child Left Behind: Subsidized Child Care and Children’s Long-Run Outcomes.” *American Economic Journal: Economic Policy*, 3: 97-129.

Liu, Quian, and Oskar Nordstrom Skans (2010). "The Duration of Paid Parental Leave and Children's Scholastic Performance." *The B.E. Journal of Economic Analysis & Policy*, 10 (1): 1935-1682.

Loeb, Susanna, Margaret Bridges, Daphna Bassok, Bruce Fuller, and Russell Rumberger (2007). “How Much is Too Much? The Influence of Preschool Centers on Children's Social and Cognitive Development.” *Economics of Education Review*, 26(19): 52-66.

Magnuson, Katherine, Christopher Ruhm, and Jane Waldfogel (2007). “Does prekindergarten improve school preparation and performance?” *Economics of Education Review*, 26(1): 33-51.

Nollenberger, Natalia, and Nuria Rodríguez-Planas (2011). “Child Care, Maternal Employment and Persistence: A Natural Experiment from Spain.” IZA Discussion Paper 5888. Substantially revised version from August 2012 available from authors upon request.

OECD (2006). “PISA 2006 Science Competencies for Tomorrow's World.”.

OECD (2009). "PISA Data Analysis Manual: SPSS® Second Edition".

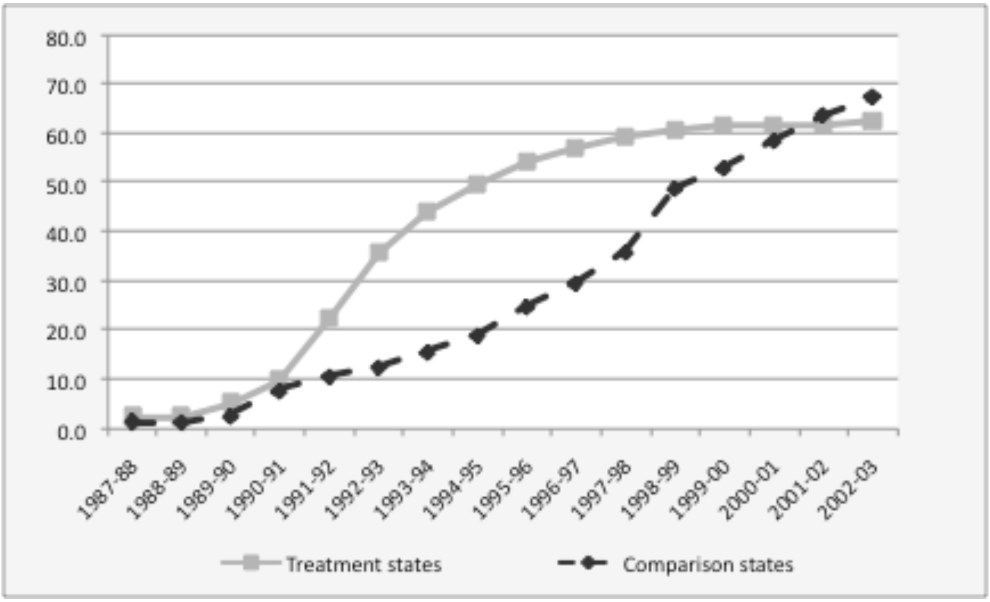
Puhani, Patrick A, and Andrea Weber (2008). "Does the early bird catch the worm?" *The Economics of Education and Training*, 105-132.

Rasmussen, Astrid Würtz (2010). "Increasing the Length of Parents' Birth-Related Leave: The Effect on Children's Long-Term Educational Outcomes." *Labour Economics,* 17(1): 91-100

Ruhm, Christopher J (1998)."The Economic Consequences of Parental Leave Mandates: Lessons from Europe." *The Quarterly Journal of Economics*, 113(1): 285-317.

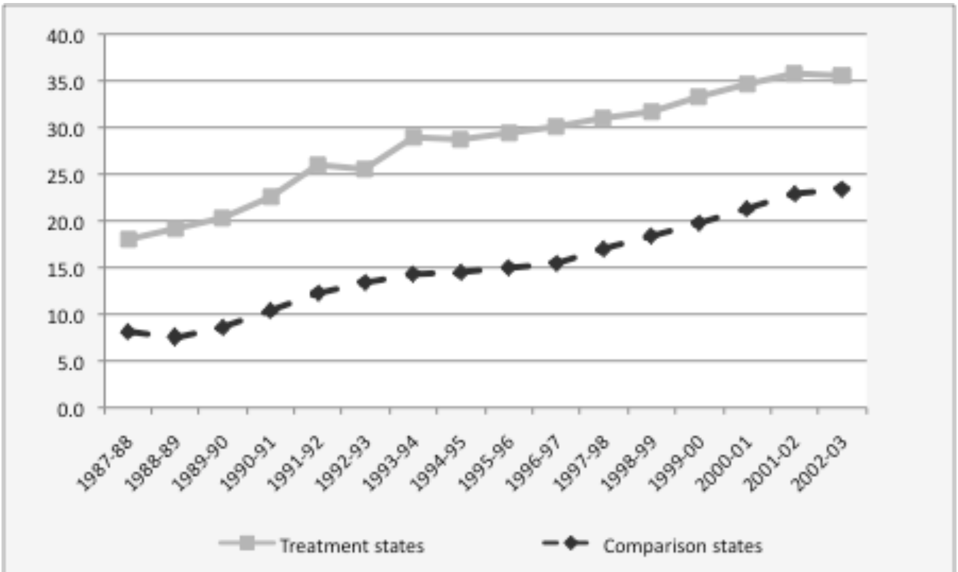
The New York Times (June 6, 2012)."German Lawmakers Spar Over Child care Subsidy", by Melissa Eddy.

Figure 1: Enrollment rates in public childcare among 3 years old



Notes: Displayed numbers are weighted averages of *public* enrollment rates for the treatment (Galicia, Cataluña, Asturias, Rioja, Castilla y León, Cantabria, Madrid and Castilla-La Mancha) and control states (Extremadura, Aragón, Baleares, Valencia, Andalucía, Murcia and Canarias). Weights reflect the population of each state (CCAA).

Figure 2: Enrollment rates in private childcare among 3 years old



Notes: Displayed numbers are weighted averages of *private* enrollment rates for the treatment and control states. See figure 1 for list of states in treatment and control groups. Weights reflect the population of each state.

Table 1. Descriptive Statistics

	Treated States		*Differences between Treated and Control States*		
	Pre-Reform		Pre-Reform	Cohort90	Cohort93
A.1 Outcomes variables					
Standardized Reading Scores	0.071	[0.957]	0.269***	0.381***	0.363***
Standardized Math Scores	0.008	[0.938]	0.304***	0.381***	0.333***
Falling behind inprimary school	0.053	[0.224]	-0.010	n.a.	-0.036***
Falling behind in secondary school	0.230	[0.421]	-0.059***	n.a.	-0.089***
A.2 Placebo variable					
Relative age (placebo outcome)	5.418	[3.423]	-0.098	-0.093	-0.011
B.1 Individual Characteristics					
Gender (Male=1)	0.471	[0.499]	-0.029	0.002	-0.019
Born in Spain	0.991	[0.093]	-0.003	-0.009	-0.003
B.2 Regional Characteristics					
GDP (Euros/capita)	10,559	[1,935]	811	930	1,057
Unemployment rate –Males	0.095	[0.022]	-0.034	-0.060	-0.024
Unemployment rate- Females	0.209	[0.048]	-0.058	-0.063	-0.042
Employment rate- Females	0.261	[0.043]	.0181	0.010	0.001
Years of education – Males	8.620	[0.289]	0.302	0.137	0.082
Years of education – Females	8.234	[0.243]	0.201	0.179	0.074
Total population (in millions)	2.479	[2.080]	151	135	107
0-6 years old (percentage)	0.068	[0.007]	-0.015***	-0.015	-0.013
Population density (inhab. per km^2)	150.0	[195.8]	45.6	44.9	43.3
Sample sizes					
Treated States			4,116	7,456	7,276
Control States			2,040	3,196	5,404

Notes: The table displays mean and standard deviation in brackets. The asterisks indicate statistically significant differences between treatment and control states. *Significant at 10 percent level; ** Significant at 5 percent level; *** Significant at 1 percent level. Standard errors are computed using BRR methodology. In the case of individual and regional characteristics, the asterisks indicate statistically significant differences resulting from balancing tests. Regional characteristics are calculated by the authors based on Spanish LFS microdata (unemployment, education, female employment rate) and on data at regional level available in www.ine.es (GDP, Population, 0-6 years old, Population density). The displayed sample sizes correspond to PISA datasets and are not weighted. The relationship between the treated and control states' sample varied across time because different states expanded their samples in different waves. For this reason, in all of our estimates we use the final student weights.

Table 2: Crowding out

A. Childcare Coverage	***Pre-treatment means***				
	Treatment States	*Control States*	***ITT***	***Se[ITT]***	***% increase***
Public Childcare					
Treat*1993	0.099	0.074	0.261***	[0.060]	264%
Treat*1996	0.099	0.074	0.256***	[0.065]	259%
Private Childcare					
Treat*1993	0.226	0.102	0.021	[0.038]	9.3%
Treat*1996	0.226	0.102	0.020	[0.029]	8.9%
B. Maternal Employment					
Effect up to 1995	0.357	0.289	0.024*	[0.014]	6.7%
Effect up to 1997	0.357	0.289	0.016*	[0.009]	4.5%

Notes: Robust standard errors in brackets. * Significant at 10 percent level; ** Significant at 5 percent level; *** Significant at 1 percent level. Panel A displays the results from estimating equation (1) using as the LHS variable the enrollment rate of 3-years old in public (private) schools. In this case we use annual data from the Spanish Ministry of Education. Sample size: 45 (15 states, 3 years). Panel B displays the results from estimating the effects of LOGSE on maternal employment using Spanish LFS data. Sample sizes: up to 1995 78,123 mothers, up 1997 to 105,748 mothers. Please refer to the Appendix for a thorough explanation of the methodological approach. We control for the pre-reform regional characteristics shown in the panel B.2 of Table 1, except the initial level of childcare coverage when the LHS variable is the enrollment rate of 3 years old. Results are really similar when instead we include state fixed-effects.

Table 3. Main Results

	TT	ITT	Se[ITT]	Individual and State Specific Controls	Individual Controls and State FE
	(1)	*(2)*	*(3)*	*(4)*	*(5)*
A. Children outcomes					
Standardized Reading Scores					
Treated*Cohort90	0.433**	0.113**	[0.047]		
	0.479***	0.125***	[0.045]	X	
	0.476***	0.124***	[0.045]		X
Treated*Cohort93	0.367**	0.094**	[0.041]		
	0.371**	0.095**	[0.039]	X	
	0.371**	0.095**	[0.040]		X
Standardized Math Scores					
Treated*Cohort90	0.295*	0.077*	[0.041]		
	0.272*	0.071*	[0.041]	X	
	0.272*	0.071*	[0.040]		X
Treated*Cohort93	0.109	0.028	[0.046]		
	0.051	0.013	[0.045]	X	
	0.051	0.013	[0.045]		X
Falling behind a grade at primary school					
Treated*Cohort93	-0.106**	-0.027**	[0.011]		
	-0.094**	-0.024**	[0.011]	X	
	-0.098**	-0.025**	[0.011]		X
Falling behind a grade at secondary school					
Treated*Cohort93	-0.125*	-0.032*	[0.018]		
	-0.125*	-0.032*	[0.018]	X	
	-0.129*	-0.033*	[0.018]		X

Notes: All of our estimates are weighted using the final student weights. Robust standard errors in brackets. * Significant at 10 percent level; ** Significant at 5 percent level; *** Significant at 1 percent level. In the case of tests scores, we estimate one regression for each plausible value clustering the standard errors at state-period level and then impute the "average" standard error following the PISA manual recommendation (see PISA Data Analysis Manual, 2005, pp. 120). We obtain the TT estimates, shown in column (1) by dividing the ITT estimates, shown in column (2), by the probability of treatment in the respective year, shown in Panel A of Table 2. Sample sizes: for Readings and Math scores 34,725 students; for the likelihood of falling a grade behind (only available in 2003 and 2009) 21,439 students. Column (4) indicates whether the specification includes controls by pre-reform regional characteristics (those listed in Panel B in Table 1) and time invariant individual characteristics (a gender dummy and immigration status). Column (5) indicates whether the specification also includes individual and state specific controls or individual controls and state fixed effects.

Table 4. Robustness checks

	ITT	SE [ITT]	Individual and State Specific Controls	Individual Controls and State FE
Panel A) Effect on probability of having migrated across states by age 16 using LFS				
A.1 From control to treatment states				
Treated*Cohort90	-0.010**	[0.005]		
	-0.010**	[0.005]	X	
	-0.010**	[0.005]		X
Treated*Cohort93	-0.003	[0.005]		
	-0.004	[0.005]	X	
	-0.004	[0.005]		X
A.2 From treatment to control states				
Treated*Cohort90	0.005	[0.004]		
	0.004	[0.004]	X	
	0.005	[0.004]		X
Treated*Cohort93	-0.001	[0.004]		
	0.000	[0.004]	X	
	0.000	[0.004]		X
B) Placebo test: Effect on Birth month				
Treated*Cohort90	-0.020	[0.137]		
	-0.053	[0.137]	X	
	-0.055	[0.137]		X
Treated*Cohort93	0.097	[0.132]		
	0.072	[0.132]	X	
	0.070	[0.132]		X

Notes: The table displays the results from estimating the equation 1 using different outcomes as LHS variable. In each case, we present the raw estimate and also one which includes individual controls (invariant characteristics) and either the pre-regional characteristics listed in the Panel B of Table 2 or state fixed-effects. In Panel A.1, the LHS variable is a dummy equal to one if the individual has migrated form a control to a treatment state (and vice versa in Panel A.2). We use all quarters of 2003, 2006 and 2009 LFS and restrict the sample to natives. The total sample size is of 19,731 observations. Individual characteristics in this case include parents' level of education. In Panel B, the LHS variable is the relative age of the child (defined as the difference between the month of birth and the cut-off date for children to begin school). We use data from the 2003, 2006 and 2009 PISA datasets, which contain information about the month of birth. Sample size: 34,725 observations. Robust standard errors in brackets. * significant at 10% level; ** significant at 5% level; *** significant at 1% level.

Table 5. Alternative Specifications

	Preferred specification	Without richest and poorest states	Flexible	Controlling for states with some control over education policy	Comparing 66th vs 33th	BRR SE
	(1)	(2)	(3)	(4)	(5)	(6)
Standardized Reading Scores						
Treat*Cohort90	0.125***	0.098*	0.269***	0.136***	0.136***	0.127*
	[0.045]	[0.058]	[0.096]	[0.046]	[0.045]	[0.071]
Treat*Cohort93	0.095**	0.054	0.119	0.084**	0.078*	0.111*
	[0.039]	[0.055]	[0.086]	[0.041]	[0.043]	[0.060]
Standardized Math Scores						
Treat*Cohort90	0.071*	0.134***	0.154**	0.027	0.034	0.078
	[0.041]	[0.052]	[0.078]	[0.042]	[0.044]	[0.060]
Treat*Cohort93	0.013	0.089	0.01	-0.032	-0.007	0.035
	[0.045]	[0.055]	[0.082]	[0.049]	[0.047]	[0.062]
Falling behind a grade at primary school						
Treat*Cohort93	-0.024**	-0.018	-0.030*	-0.030***	-0.032***	-0.029**
	[0.011]	[0.014]	[0.017]	[0.011]	[0.011]	[0.013]
Falling behind a grade at secondary school						
Treat*Cohort93	-0.032*	-0.043*	-0.030	-0.008	-0.053***	-0.036
	[0.018]	[0.023]	[0.034]	[0.018]	[0.014]	[0.023]
ITT/TT (Cohort90)	0.261	0.261	0.261	0.261	0.337	0.261
ITT/TT (Cohort93)	0.255	0.255	0.255	0.255	0.309	0.255

Notes: We report the intent to treatment effect (ITT) estimates based on regressions including individual and states specific controls. Column (1) presents our baseline specification. Column (2) shows the estimates dropping the richest and the poorest states within treatment and control groups. In column (3) the cohort fixed-effects are interacted with pre-reform states socio-economic characteristics. In column (4) we add a dummy to control for the fact that some states have control over their education policy (namely Andalucia, Canary Island, Catalonia, Valencia and Galicia) and interact this dummy with the cohort dummies. Column (5) displays the results when treatment states are defined as those above 67th percentile in public enrollment growth and control states as those below the 33th. In column (6), standard errors are imputed applying the Fay's Balanced Repeated Replication method. This specification includes all the individual and school characteristics showed in Appendix Table A.1. * significant at 10% level; ** significant at 5% level; *** significant at 1% level.

Table 6. Alternative Identification Strategy

	ITT *(1)*	Se[ITT] *(2)*	BRR SE *(3)*
Standardized Reading Scores			
SEATS	0.010***	[0.002]	
	0.010**	[0.004]	X
Standardized Math Scores			
SEATS	0.004**	[0.002]	
	0.004	[0.004]	X
Falling behind a grade at primary school			
SEATS	-0.002***	[0.001]	
	-0.002***	[0.001]	X
Falling behind a grade at secondary school			
SEATS	0.001	[0.001]	
	0.001	[0.001]	X

Notes: The table displays the results from estimating the equation 2 (see the main text for details). Standard errors in brackets. * significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level. The specification includes state fixed-effects and time invariant individual characteristics (a gender dummy and immigration status). The last column indicates whether the standard errors are imputed using the Fay's BRR method.

Table 7. Heterogenous Effects

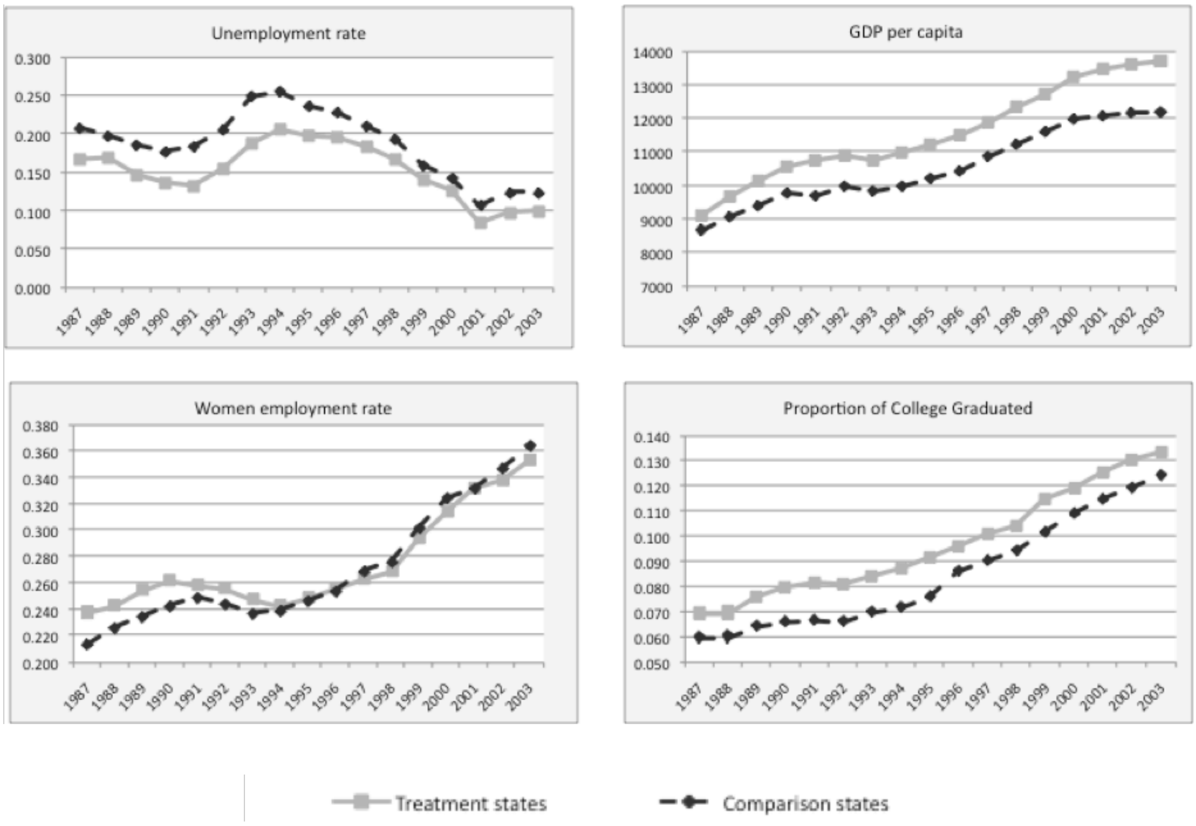
Panel A By Gender:	***Boys***		***Girls***	
Standardized Reading Scores				
Treated*Cohort90	0.138**	[0.069]	0.117**	[0.054]
Treated*Cohort93	0.058	[0.061]	0.147***	[0.053]
Standardized Math Scores				
Treated*Cohort90	0.054	[0.062]	0.090*	[0.054]
Treated*Cohort93	-0.067	[0.062]	0.101*	[0.058]
Falling behind a grade at primary school				
Treated*Cohort93	-0.024	[0.017]	-0.024	[0.015]
Falling behind a grade at secondary school				
Treated*Cohort93	-0.019	[0.026]	-0.045*	[0.024]

Panel B: By education	***Neither of the parents have a secondary school degree***		***At least one of the parents have a secondary school degree***	
Standardized Reading Scores				
Treated*Cohort90	0.133*	[0.079]	0.075	[0.055]
Treated*Cohort93	0.107	[0.072]	0.058	[0.047]
Standardized Math Scores				
Treated*Cohort90	0.04	[0.072]	0.048	[0.049]
Treated*Cohort93	-0.045	[0.073]	0.012	[0.053]
Falling behind a grade at primary school				
Treated*Cohort93	-0.040*	[0.024]	-0.007	[0.012]
Falling behind a grade at secondary school				
Treated*Cohort93	-0.037	[0.035]	-0.019	[0.020]

Notes: The table reports the ITT parameter. Robust standard errors in brackets. * significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level. Panel A and Panel B, display the results from estimating equation 1 including controls for individual and specific regional characteristics. In Panel A, the sample sizes are for boys: Test scores 17,647, Grade repetition 11,208, and for girls: Tests socres 17,663, Grade repetition 11,231. In Panel B sample sizes are for those with parents of low education: Test scores 9,487, Grade repetition 5,743; for those with at least one parent of high education: Test scores 25,823, Grade repetition 16,696.

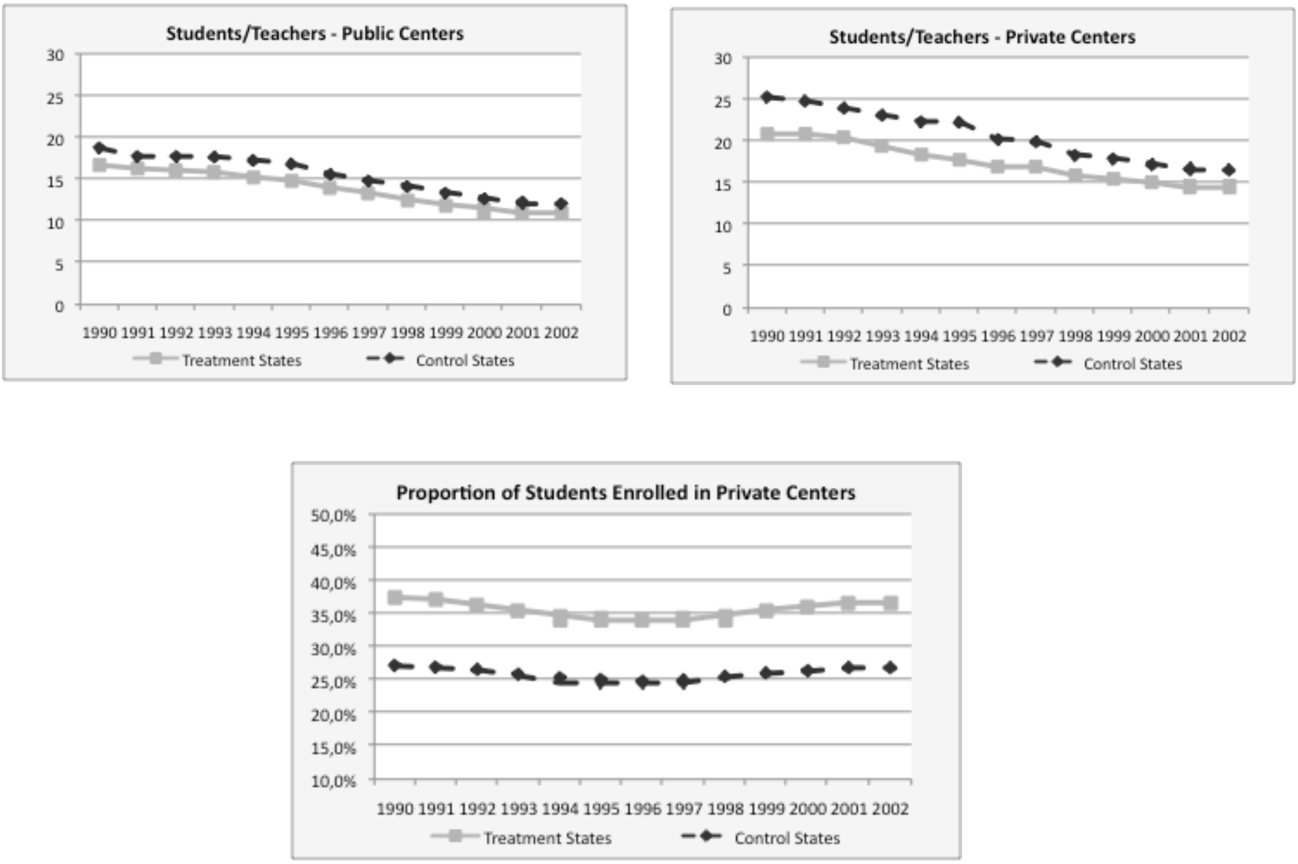
APPENDIX

Figure A.1. Trends in further socio-economic state features between 1987-2003



Notes: Elaborated by the authors. Unemployment rate, women employment rate and proportion of college graduated were calculated by the authors based on Spanish LFS microdata (we consider individuals from 16 years old). GDP per capita was calculated based on regional account data from the Spanish Statistics Institute (www.ine.es). It is expressedin constant euros of 1995.

Figure A.2. Trends in primary school student/teacher ratios between 1990-2002



Notes: Elaborated by the authors using data from the Ministry of Education (www.educacion.gob.es). Regarding the first two figures, we sum the total children enrolled in treatment and control states respectively, and then divide it by the total number of teachers in each group of states (it is equivalent to do a weighted average across states within each group).

Estimating the effects of the reform on maternal employment

As in Nollenberger and Rodríguez-Planas (2011), we follow a Difference-in-Difference-in-Difference approach exploiting the fact that the legislation affected children of 3 years old but not mothers of 2 years old. We therefore estimate the following equation:

$$Y_{ijt} = \gamma_1 Treat_j + \gamma_2 Mom3_i + \gamma_3 Post_t + \gamma_4 (Treat_j * Mom3_i) + \gamma_5 (Treat_j * Post_t) + \\ + \gamma_6 Mom3_i * Post_j + \theta \ (Treat_j * Mom3_i * Post_t) + \delta_j + \lambda_t + X'_{ijt}\beta + \varepsilon_{ijt}$$

where Y_{ijt} is the outcome of interest (employment or weekly hours worked) for a sample of mothers whose youngest child is 2 or 3 years old, $Treat_j$ is equal to one if the mother lives in a treatment state and zero otherwise; $Mom3_i$ is equal to one for mothers whose youngest child is 3 years old and zero for mothers whose youngest child is 2 years old; the variable $Post_t$ is equal to one after LOGSE implementation began (that is, from 1991/92 onwards). The coefficient θ capture any difference in the likelihood of being employed for mothers of treated children (3 year olds) relative to control children (2 year olds) living in treated states after the childcare expansion. The vector X_{ijt} includes the same individual and regional controls as in Nollenberger and Rodríguez-Planas (2011), namely age squared, dummies indicating the number of other children, a dummy for being foreign-born, educational attainment dummies (high-school dropout, high-school graduate, and college), a dummy for being married or cohabitating. We also include state and year fixed effects. We estimate this equation by OLS using data from the Spanish Labor Force Survey from 1987 to 1994 and also from 1987 to 1997.

Table A 1. Sensitivity Analysis of Covariates Included

	Unconditional	+ Regional Charact.	+ Individual Characteristics	+ Family Characteristics	+ Type of school	+ Pop. density of place of residence
	(1)	(2)	(3)	(4)	(5)	(6)
StandarizedReading score						
Treat*Cohort90	0.113**	0.116**	0.125***	0.101**	0.115**	0.121***
	[0.047]	[0.046]	[0.045]	[0.043]	[0.046]	[0.046]
Treat*Cohort93	0.094**	0.091**	0.095**	0.092**	0.109***	0.103***
	[0.041]	[0.040]	[0.039]	[0.038]	[0.040]	[0.040]
Standardized Maths score						
Treat*Cohort90	0.077*	0.077*	0.071*	0.045	0.068*	0.072*
	[0.041]	[0.041]	[0.041]	[0.038]	[0.039]	[0.039]
Treat*Cohort93	0.028	0.024	0.013	0.01	0.03	0.025
	[0.046]	[0.046]	[0.045]	[0.043]	[0.044]	[0.044]
Falling behind a grade at primary school						
Treat*Cohort93	-0.027**	-0.023**	-0.024**	-0.027**	-0.030***	-0.029**
	[0.011]	[0.011]	[0.011]	[0.011]	[0.011]	[0.011]
Falling behind a grade at secondary school						
Treat*Cohort93	-0.032*	-0.031*	-0.032*	-0.034**	-0.037**	-0.036**
	[0.018]	[0.018]	[0.018]	[0.017]	[0.017]	[0.017]

Notes: Individual characteristics: male, immigrants; Family Characteristics: Home possession score (an index derived from students' responses to the following items: do you have: a desk for study, a room of your own, a computer, internet, classic literature, books, works of art, dishwasher, among others), mother's and father's education; Type of school: public-omitted; private; Population density of place of residence: Village, Small Town, Town, City, Large City, Metropolis -omitted.

Table A 2. Alternative Specifications (including State Fixed-Effects instead pre-regional characteristics).

	Preferred specification	Without richest and poorest states	Flexible	Controlling for states with some control over education policy	Comparing 66th vs 33th	BRR SE
	(1)	(2)	(3)	(4)	(5)	(6)
Standardized Reading Scores						
Treat*Cohort90	0.124***	0.098*	0.220**	0.138***	0.136***	0.125*
	[0.045]	[0.058]	[0.091]	[0.046]	[0.045]	[0.071]
Treat*Cohort93	0.095**	0.054	0.078	0.086**	0.078*	0.110*
	[0.040]	[0.055]	[0.082]	[0.041]	[0.043]	[0.060]
Standardized Math Scores						
Treat*Cohort90	0.071*	0.134***	0.126*	0.029	0.034	0.077
	[0.040]	[0.052]	[0.074]	[0.042]	[0.044]	[0.060]
Treat*Cohort93	0.013	0.089	-0.015	-0.029	-0.007	0.034
	[0.045]	[0.055]	[0.079]	[0.049]	[0.047]	[0.061]
Falling behind a grade at primary school						
Treat*Cohort93	-0.025**	-0.005	-0.030*	-0.030***	-0.032***	-0.029**
	[0.011]	[0.015]	[0.017]	[0.011]	[0.011]	[0.013]
Falling behind a grade at secondary school						
Treat*Cohort93	-0.033*	-0.097***	-0.030	-0.008	0.003	-0.037
	[0.018]	[0.025]	[0.034]	[0.018]	[0.018]	[0.024]
ITT/TT (Cohort90)	0.261	0.261	0.261	0.261	0.337	0.261
ITT/TT (Cohort93)	0.255	0.255	0.255	0.255	0.309	0.255

Notes: We report the intent to treatment effect (ITT) estimates based on regressions including individual controls and states fixed-effects. Column (1) presents our baseline specification. Column (2) shows the estimates dropping the richest and the poorest states within treatment and control groups. In column (3) the cohort fixed-effects are interacted with pre-reform states socio-economic characteristics. In column (4) we add a dummy to control for the fact that some states have control over their education policy (namely Andalucia, Canary Island, Catalonia, Valencia and Galicia) and interact this dummy with the cohort dummies. Column (5) displays the results when treatment states are defined as those above 67th percentile in public enrollment growth and control states as those below the 33th. In column (6), standard errors are imputed applying the Fay's Balanced Repeated Replication method. This specification includes all the individual and school characteristics showed in Appendix Table A.1. * significant at 10% level; ** significant at 5% level; *** significant at 1% level.