Paris School of Economics

From Playtime to Primetime. The Medium-term Effects of Compulsory Preschool in Mexico

Master Dissertation

Public Policy and Development

Presents

Ricardo Gómez Carrera

Supervisor

Dr. Karen Macours

Referee

Dr. Luc Behaghel

Paris, France 2024

From Playtime to Primetime. The Medium-term Effects of Compulsory Preschool in Mexico

Ricardo Gómez-Carrera. *

Paris School of Economics†

Supervised by Karen Macours

August 28, 2024

Abstract

Preschool programs have rapidly expanded in low- and middle-income countries, yet concerns persist about their quality and effectiveness in promoting children's development. This study examines the impact of making preschool compulsory in Mexico on primary school performance using a staggered difference-in-differences methodology. Results indicate an improvement in Grade 5 test scores by 0.05-0.16 standard deviations. However, the reform widened disparities between the most and least marginalized localities, underscoring the need for targeted policies. Smaller preschool groups and extended preschool attendance improved cognitive skills, despite the higher proportion of lower-skilled teachers. High-quality early childhood education is crucial for long-term academic success in low- and middle-income countries.

JEL Codes: C23, I25, J24, N36, O15

Keywords: Early childhood development, education, Mexico, preschool, test scores, staggered difference-in-differences.

*Paris School of Economics. E-mail: ricardo.goca90@gmail.com.

†I am grateful to Karen Macours for her supervision, extensive feedback, and encouragement. Special thanks to Luc Behaghel for refereeing this work. This research has also been enriched with insightful feedback from Liam Wren-Lewis, Jérémie Gignoux, and participants at the Development Research Seminar PSE 2023-2024. I acknowledge support from Xaber A.C. I benefited from great feedback from Rafael de Hoyos, Arturo Aguilar, Adrian Rubli, Claudia Aburto, Magdalena Barba, and Christiane Fabrega. All errors are my own.

1 Introduction

Investments in early childhood development (ECD) are essential for fostering long-term skill development. Research demonstrates that the early years can be important for cognitive, social, and emotional development, laying the foundation for lifelong learning, better employment prospects, and well-being.[1] Furthermore, high-quality early childhood education (ECE) has been shown to be an effective tool for promoting economic growth[2] and reducing inequality, as it provides disadvantaged children with essential resources and support that can mitigate the effects of socioeconomic disparities (Attanasio et al. (2022); Becker et al. (2018); Card (2001); Cascio (2023); Chetty et al. (2011); Corak (2013); Cunha and Heckman (2009); Deming (2009); Deming (2022); Hahn and Barnett (2023); Heckman and Mosso (2014); Elango et al. (2016); Gregorio and Lee (2002)). In this context, examining early childhood education in low and middle-income countries (LMICs) is especially relevant, as these regions have made substantial investments in preschool programs over recent decades, leading to an increase in preschool enrollment rates from 28% in 2000 to 58% in 2020. In contrast, high-income countries (HICs) experienced a more modest rise, with enrollment rates increasing from 72% in 2000 to 80% in 2020 (World Bank, 2023). Furthermore, these regions stand to gain more from early childhood investments, as they have historically invested less, and families often have fewer resources and more children.

Early childhood interventions can be delivered through various modalities, including home and center-based programs. Within the realm of center-based programs, this study concentrates on preschools rather than childcare centers, as the vast majority of children participate in formal center-based preschools. Furthermore, the study examines a universal preschool program rather than a targeted intervention, allowing for an exploration of the broader impacts of these programs on all children within a community, regardless of their individual risk factors. Preschool programs can enhance children's cognitive abilities, providing them with a solid foundation for success in primary school and beyond. However, the effectiveness of these interventions is heavily dependent on their quality; poorly designed or implemented programs may fail to produce the desired out-

[1] See, e.g., Almond et al. (2018), Barnett (1992), Barnett (1995), Cunha and Heckman (2007), Cunha et al. (2010), Currie (2001), Currie and Almond (2011), Deming (2022), Duncan et al. (2023), Hanushek and Woessmann (2012a), Heckman (2013), Elango et al. (2015), and Garces et al. (2002).

[2] Education is a fundamental driver of economic growth and social development, shaping individuals' life outcomes and contributing to the overall progress of societies (Angrist et al. (2021); Barro (2001); Hanushek and Woessmann (2020); Hanushek and Woessmann (2023); Romer (1990)).

comes.[3] Additionally, the phenomenon of fade-out effects—where initial cognitive gains diminish over time—raises important questions about the sustainability of preschool benefits.[4]

The intervention studied in this article centers on Mexico's 2002 educational reform, which mandated compulsory preschool education for children aged 3-5. This reform aimed to enhance cognitive development by increasing children's exposure to formal educational environments, addressing previous deficiencies in familial developmental stimulation. It also sought to improve quality by hiring more preschool teachers and requiring them to have at least a university degree. Since its publication in 2002, the reform was accompanied by a large rise in preschool enrollment, particularly among four-year-olds in the second year of preschool (pre-k 2).

This article employs a combination of methodologies to analyze the impact of preschool education on cognitive outcomes. The first approach is a two-way fixed effects (TWFE) model, utilizing a canonical difference-in-differences framework to compare test scores between localities with preschools and localities without. This model assesses one period before and one after the reform, focusing on the average causal effect on treated cohorts. Secondly, to address the staggered adoption of the preschool mandate across different states, the study also implements a staggered difference-in-differences (DiD) approach. This method accounts for distributional treatment effects and allows for the analysis of varying treatment timing and intensity. By using a continuous treatment variable that reflects the standardized number of preschools per locality, the study enhances its robustness in evaluating the effects of the reform on cognitive outcomes and the quality mechanisms driving these effects.

This article employs two primary data sources: Formato 911, which provides detailed information on the number of preschools and their characteristics, and ENLACE, which provides standardized test scores for primary school students. These datasets are merged at the locality-cohort level, enabling a robust comparison of treated and untreated cohorts. The results from the staggered DiD

[3] Several studies argue that the quality of preschool education is crucial for optimizing its benefits on children's development and long-term academic success. See, e.g., Andrew et al. (2020), Andrew et al. (2024), Baker et al. (2019), Barnett (1992), Blau and Currie (2006), Blimpo et al. (2022),Corak (2013), Danziger and Waldfogel (2000), Duncan and Magnuson (2013), Elango et al. (2015), Elango et al. (2016), Felfe et al. (2015), Gupta and Simonsen (2010), Haeck et al. (2015), Hanushek and Woessmann (2012a), Hanushek and Woessmann (2012b), Heckman et al. (2010b), Heckman (2013), Melhuish et al. (2015), Pages et al. (2022), Santibañez et al. (2007), and Weiland and Yoshikawa (2013).

[4] The persistence of the impacts of these interventions has been examined in studies such as Bailey et al. (2017), Deming (2009), Duncan and Magnuson (2013), Duncan et al. (2023), Freeberg and Payne (1967), Gray-Lobe et al. (2023), Lipsey et al. (2018), Meghir et al. (2023), Melhuish et al. (2015), Pages et al. (2022), Puma et al. (2010), and Puma et al. (2012).

analysis reveal a significant increase in four-year-old enrollment, with an approximate rise of 3-4 students per locality, which is substantial compared to the baseline of 16 students. Additionally, the analysis indicates an improvement in Grade 5 average test scores, particularly in math, with increases ranging from 0.05 to 0.16 SD. Results from the staggered DiD analysis also show evidence of positive impacts on Grade 6 test scores and no impact on Grade 3 and Grade 4 test scores. Despite the increase in preschool enrollment, there are challenges related to teacher qualifications, as the proportion of preschool teachers with university degrees declined. However, the findings suggest that the benefits of smaller class sizes and extended preschool time can mitigate the potential negative impacts of lower-skilled teachers, ultimately leading to cognitive skill improvements in children. The TWFE model results indicate a significant increase in four-year-old enrollment and an improvement in average test scores by approximately 0.15 to 0.20 SD for Grades 3, 4, and 6, with larger increases in math.

In addition, this study explores the distributional effects of the preschool reform on test scores across various demographic groups, including gender, poverty, indigenous status, and rurality. The analysis reveals that boys experienced a more pronounced increase in Grade 5 test scores in both math and Spanish, despite a higher relative preschool enrollment of girls, suggesting that preschool environments may better cater to boys' learning needs. In terms of socioeconomic status, students from non-poor localities, particularly those in medium-poverty areas, demonstrated greater improvements in test scores compared to their peers in the poorest localities. For indigenous students, math scores remained unchanged, likely due to stagnant preschool enrollment in these communities. Furthermore, the results indicate that rural localities benefited significantly from the reform, with notable increases in test scores attributed to higher preschool enrollment and reduced class sizes.

This paper contributes to the existing literature on early childhood education by examining a large preschool expansion program in a low and middle-income country context, unlike much of the literature that primarily addresses high-income countries or small-scale targeted interventions. Its findings highlight the importance of preschool quality by investigating factors such as class sizes and teacher qualifications, which are crucial for cognitive development. It also explores medium-term academic outcomes, demonstrating that the benefits of preschool can persist as children progress through their education. Additionally, the research analyzes the distributional effects of the reform across various demographic groups, revealing improvements in test scores for children in medium-poverty areas while emphasizing the need for tailored educational policies to improve the outcomes of the most marginalized population.

The remainder of the paper is organized as follows. Section 2 reviews the existing literature. Section 3 provides background information on the preschool mandate that this study evaluates. Section 4 describes the data. Section 5 discusses the two quasi-experimental estimation strategies employed in this study. Section 6 presents the results for the canonical difference-in-differences strategy. Section 7 showcases the main results using the staggered difference-in-differences methodology. Section 8 discusses the robustness of the results. Section 9 concludes.

2 Literature Review

This study is set in the context of a middle-income country. This context has been relatively underrepresented in the existing literature on early childhood education, which primarily focuses on high-income countries, despite LMICs being home to over 85% of the world's preschool-age children.[5] Early childhood interventions in LMICs do not guarantee positive cognitive outcomes, as their effectiveness relies on specific design and implementation factors. The diverse contexts across LMICs further complicate the generalizability of successful interventions.

Early childhood interventions are often evaluated through two primary lenses: the employment outcomes of mothers[6] and the developmental outcomes of children. This study falls into the latter category, examining how early childhood interventions influence medium-term academic performance: in primary school. By prioritizing children's developmental trajectories, this research aims

[5]Some examples of studies examining early childhood interventions in HICs include Bailey et al. (2017), Baker et al. (2008), Baker et al. (2019), Barnett (1992), Barnett (1995), Barr et al. (2022), Blau and Currie (2006), Cascio (2023), Chetty et al. (2011), Cornelissen et al. (2018), Currie and Thomas (1993), Deming (2009), Dietrichson et al. (2020), Felfe et al. (2015), García and Heckman (2023), Garces et al. (2002), Gray-Lobe et al. (2023), Gupta and Simonsen (2010), Haeck et al. (2015), Havnes and Mogstad (2011), Havnes and Mogstad (2015), Heckman et al. (2010a), Heckman et al. (2010b), Heckman et al. (2013), Ludwig and Miller (2007), Miller et al. (2023), Nollenberger and Rodríguez-Planas (2015), Pages et al. (2022), Nores et al. (2005), and Weiland and Yoshikawa (2013).

[6]Some examples of studies that examine the relationship between early childhood interventions and maternal employment are Abraham and Kearney (2020), Andresen and Havnes (2019), Baker et al. (2008), Berniell et al. (2023) Blau and Currie (2006), Casarico and Lattanzio (2023), Cascio (2021), Carta and Rizzica (2018), Costa Dias et al. (2020), Dustmann and Schönberg (2012), De la Cruz Toledo (2015), Givord and Marbot (2015), Haeck et al. (2015), Havnes and Mogstad (2011), Kleven et al. (2024), Martínez and Perticará (2017), Nollenberger and Rodríguez-Planas (2015), and Olivetti and Petrongolo (2017).

to contribute to the understanding of how additional schooling during this critical developmental stage can foster cognitive skills.

Pertinent literature on ECD interventions in LMICs for this paper includes studies examining the impact of preschool programs, with some research indicating positive effects while others do not. Berlinski et al. (2008) show that in Uruguay, preschool attendance significantly enhances educational outcomes by age 15, with treated children accumulating 0.8 more years of education and being 27 percentage points more likely to remain in school than their untreated siblings, particularly benefiting disadvantaged children. Berlinski et al. (2009) found that a large expansion of universal preprimary education in Argentina significantly boosted primary school performance, with an additional year of preprimary education increasing third-grade test scores by about 0.23 SD. This contrasts with Bouguen et al. (2018), who explore a preschool expansion program in Cambodia, revealing that low attendance rates following the intervention were driven by parental concerns and beliefs about preschool, highlighting the need for effective outreach to improve engagement. Meghir et al. (2013) estimate the impacts of early stimulation for children under three years and enhanced preschool for children aged three years and older in rural India, finding that while both interventions improved children's cognitive skills and school readiness, only the effects of early stimulation were sustained after 15 months. They find that enhanced preschool provided significant catch-up benefits for those who did not receive early stimulation, with no significant interactions between the two interventions, highlighting the need for further investigation into their lack of complementarities. Spier et al. (2020) found that subsidized preschool for 4-year-olds in Bangladesh increased school readiness measured by literacy (0.23 SD), numeracy (0.30 SD), and social and emotional development (0.34 SD).

Relevant literature on other ECD interventions in LMICs shows mixed results. For instance, Andrew et al. (2020) found that a parenting intervention in the urban slums of India improved short-term children's cognitive development by 0.35 SD, especially for boys; Andrew et al. (2024) found that this intervention led to short-term improvements in cognition, receptive language, and expressive language, with lasting impacts on numeracy (0.33 SD) and literacy (0.27 SD) after 4.5 years, particularly among the most disadvantaged children. Araujo et al. (2021) reported a 0.10 SD increase in development scores from a large-scale home visiting program in rural Peru. A review by Attanasio et al. (2022) highlighted positive cognitive outcomes from parental investment interventions in Colombia, Ethiopia, India, and Peru. In rural Indonesia, Brinkman et al. (2017) noted a 7-9 percentage point increase in early childhood education enrollment from a government-sponsored playgroup program, with some cognitive benefits for disadvantaged children, though these were not long-lasting. In Chile, Carneiro et al. (2019) found a low-cost parenting program improved children's receptive vocabulary and socio-emotional development by 0.43 and 0.54 SD, respectively, three years later. Additionally, Grantham-McGregor et al. (2020) reported significant improvements in child cognition and language development in rural India from both home visiting (0.32 SD) and group sessions (0.28 SD). Furthermore, Gertler et al. (2014) estimated that an early childhood stimulation program in Jamaica increased participants' earnings by 25% after 20 years, while Grantham-McGregor et al. (1991) found that both nutritional supplementation and psychosocial stimulation improved cognitive development, with the most significant benefits from their combination. Walker et al. (2011) also found for the Jamaican intervention that psychosocial stimulation, unlike nutritional supplementation, led to lower violence involvement, higher IQ, better educational attainment, improved general knowledge, and fewer depression symptoms, emphasizing the importance of early intervention for enhancing adult functioning and reducing violence. On the other hand, Barrera et al. (2020) evaluate a daily text message intervention aimed at improving parenting practices in poor households in rural Nicaragua, finding that while the messages enhanced self-reported parental practices, they did not result in improvements in children's cognitive or socio-emotional outcomes. The lack of positive results may be attributed to contextual factors such as limited cell phone access or low education levels. Blimpo et al. (2022) evaluate two interventions to improve early childhood development services in Gambia, finding that while access to community-based centers was increased and training for providers was implemented, there were no significant overall improvements in child development outcomes, highlighting the need for both access and quality in ECD programs.

A highly relevant body of literature for this study focuses on the 2002 mandate that made preschool compulsory in Mexico. Consistent with this study, research by Yoshikawa et al. (2007) and De la Cruz Toledo (2015) documented increases in preschool enrollment, particularly in the second year of preprimary education, while Zhang et al. (2021) identified a positive correlation between the number of preschools and subsequent academic performance in primary school. Additionally, De la Cruz Toledo (2015) found that the reform led to increased employment among mothers of young children, indicating broader social benefits. Consistent with my findings, other studies evaluating the reform have also found positive impacts on cognitive skills. Using a difference-in-discontinuity research design, Gómez-Carrera (2022) reported a 0.04 standard deviation (SD) increase in Grade 3 test scores four years after the reform, driven by a 2 percentage point rise in prior pre-k 2 enrollment. Similarly, Behrman et al. (2024) employed a difference-in-discontinuity approach and found that the Mexican preschool mandate resulted in increased preschool attendance and significant improvements in fifth- and sixth-grade cognitive scores in math and Spanish, as well as enhanced noncognitive skills and student engagement, with notable long-term effects on educational attainment nearly 20 years post-reform, including higher probabilities of high school and college completion. They indicated that the intervention led to test score increases of 0.07-0.11 SD in math and 0.04-0.07 SD in Spanish.

The difference-in-discontinuity method used in Behrman et al. (2024) combines elements of regression discontinuity (RDD) and difference-in-differences designs to mitigate the effects of confounding treatments[7] at discontinuities (Grembi et al. (2016); Tramontin Shinoki et al. (2024); Takahashi (2024)). It relies on two key identification assumptions: continuity of potential outcomes in the running variable -related to RDD-, and constant effects of confounding treatments across groups at the discontinuity (local parallel trends) -related to DiD-. It also considers that all units adopt the treatment simultaneously and that the treatment is homogeneous across units. Behrman et al. (2024) use this methodology to compare outcomes for children with birthdays in the one month before and after the cutoff date in both pre- and post-reform years. While the difference-in-discontinuity methodology can be effective, it has limitations compared to staggered DiD. Staggered DiD relies on less restrictive assumptions,[8] accommodates staggered policy adoption, and can utilize more data points since it does not require focusing on units close to a specific threshold. For policy rollouts with variation in adoption timing, staggered DiD is preferable due to its broader applicability and more extensive literature addressing potential biases and heterogeneous treatment effects (as discussed in Section 5.2).

[7] Confounding treatment: pre-existing confounder due to the discontinuity.

[8] Difference-in-discontinuity relies on assumptions for both the RDD and the canonical DiD methodologies while staggered DiD relaxes the two main DiD assumptions (see Section 5.2).

Several studies argue that the quality of preschool education is crucial for optimizing its benefits on children's development and long-term academic success.[9] High-quality preschool programs create enriching environments that promote cognitive, social, and emotional growth through well-trained educators, engaging curricula, and supportive classroom settings (Barnett (1992); Heckman (2013); Pages et al. (2022); Weiland and Yoshikawa (2013)). For example, Andrew et al. (2024) found that in Colombia, low-cost preschool teacher training improved children's cognitive development, particularly among disadvantaged children, while funding teaching assistants had no effect. Similarly, Gallego et al. (2021) reported improvements in math scores from a preschool teaching intervention in Peru. Wolf (2019) found that a preschool teacher-training program in Ghana had lasting positive impacts on literacy and executive function, dependent on the quality of subsequent classroom environments.[10] Conversely, Blimpo et al. (2022) found no cognitive benefits from new ECD centers and care-provider training in Gambia. Berkes et al. (2019) show that a preschool expansion program in Cambodia had large impacts on the quality of preschool infrastructure and materials but only limited impacts on the quality of educational processes, such as the pedagogical practices and the quality of teacher-child interaction, thus finding very small cognitive effects. With this study, I contribute to the literature by showing that the positive quality effect of smaller classrooms outweighs the negative impact of lower-skilled teachers. Higher-skilled preschool teachers would have likely further enhanced the children's cognitive gains.

Research on the long-term impacts of early childhood interventions has identified three main patterns of effects over time (Van Aar et al., 2017). Fade-out effects occur when initial positive impacts diminish or disappear as children age, suggesting that early gains may not be sustained without continued support (e.g., Bailey et al. (2017), Lipsey et al. (2018), Duncan et al. (2023), Brinkman et al. (2017), and Meghir et al. (2013)). Sleeper effects describe a pattern where benefits are not immediately apparent but emerge later in life, indicating that early interventions may have latent impacts that manifest over time (e.g., Gray-Lobe et al. (2023), Deming (2009), Heckman (2013), and Melhuish et al. (2015)). Sustainable effects refer to persistent positive outcomes that

[9] See, e.g., Andrew et al. (2020), Andrew et al. (2024), Baker et al. (2019), Barnett (1992), Blau and Currie (2006), Blimpo et al. (2022),Corak (2013), Danziger and Waldfogel (2000), Duncan and Magnuson (2013), Elango et al. (2015), Elango et al. (2016), Felfe et al. (2015), Gupta and Simonsen (2010), Haeck et al. (2015), Hanushek and Woessmann (2012a), Hanushek and Woessmann (2012b), Heckman et al. (2010b), Heckman (2013), Holla et al. (2021), Melhuish et al. (2015), Pages et al. (2022), Santibañez et al. (2007), and Weiland and Yoshikawa (2013).

[10] See also Wolf and Peele (2019), Wolf et al. (2019), and Wolf et al. (2019).

maintain or even increase as children develop, supporting the idea that early interventions can have lasting benefits (Pages et al. (2022), Andrew et al. (2024), Gertler et al. (2014), and Wolf et al. (2019)). These different patterns highlight the complexity of evaluating early childhood programs and underscore the importance of long-term follow-up studies to fully understand their impacts. Factors such as intervention quality, duration, and subsequent environments may influence which pattern emerges for a given program. The findings in this study suggest that the cognitive skills developed in preprimary school persist through the end of primary school, indicating no fade-out effect. Moreover, there is evidence of sleeper effects since the full benefits of early childhood education become more evident as students encounter increasingly complex academic challenges in the later years of primary school.

3 Background on the Preschool Mandate

In 2002, Mexico introduced an important educational reform by mandating compulsory preschool education. The reform not only made it obligatory for the State to provide preprimary education services but also mandated parents to see their children attend preschool (Bennett and Tayler, 2006). This legislative change aimed to enhance cognitive development in children through increased exposure to formal educational environments, thereby addressing deficiencies in familial developmental stimulation (De la Cruz Toledo (2015); Yoshikawa et al. (2007)). The reform was enacted with robust political backing, garnering unanimous support from the major political parties and the National Teacher's Union (SNTE), and was legislatively passed with near-unanimous votes in both houses of Congress (462 out of 468 deputies and all 96 senators voted in favor).

The compulsory preschool initiative was implemented in phases to manage the logistical demands on local governments.[11] The first phase mandated enrollment for all 5-year-olds, in the third year of preprimary education (pre-k 3), by the 2004-2005 academic year. The second phase mandated the inclusion of 4-year-olds, in the second year of preprimary education (pre-k 2), by 2005-2006. The third phase initially planned to extend to 3-year-olds, in the first grade of preprimary education (pre-k 1), by 2008-2009. However, full implementation for pre-k 1 faced obstacles

[11] States are responsible for education provision since the decentralization of the education system in 1992. Municipalities are responsible for education infrastructure maintenance. See Appendix 2 for more information on the Mexican Administrative Division and Appendix 3 for more information on the Mexican Education System.

and was not implemented in 2008 (Pérez et al. (2010); Yoshikawa et al. (2007)).[12] Table 1 shows that the largest increase in preschool enrollment was observed for pre-k 2, which was mandated to be compulsory starting in 2005. Notably, the largest increases in pre-k 2 enrollment occurred earlier, with a rise of 7 percentage points (pp) from 2001 to 2002 and 10 pp from 2003 to 2004, rather than in 2005 (5 pp).

Table 1: National Enrollment Rates in Preschool Education

Year	3-year-olds	4-year-olds	5-year-olds	3-, 4-, 5-year-olds
1998	13.35%	53.52%	77.02%	48.02%
1999	13.96%	54.57%	77.63%	48.90%
2000	15.27%	54.76%	79.31%	50.11%
2001	17.02%	56.05%	79.43%	51.23%
2002	20.59%	63.19%	81.45%	55.50%
2003	22.07%	66.38%	85.83%	58.60%
2004	25.64%	76.23%	93.08%	65.49%
2005	24.61%	81.32%	98.35%	68.14%

Source: Yoshikawa et al. (2007) Table 4, based on Formato 911 and INEGI.

Prior to the reform, preschool education was not obligatory, and enrollment rates were low. By 2000, only 50% of children aged 3-5 were enrolled in some form of preschool, a figure that rose dramatically post-reform, reaching 68% by 2005. Specifically, enrollment for 4-year-olds surged from 55% to 81% over the same period (Yoshikawa et al., 2007). Despite the impressive enrollment gains, the policy was not accompanied by sanctions for non-compliance, which tempered the initial ambitions of universal preschool access (Yoshikawa et al., 2007).

The reform also sought to improve the quality of preschool education through a series of initiatives. These included the requirement for preschool teachers to possess at least a university degree (Licenciatura), and a substantial increase in the teaching workforce, with 28,760 new teachers added between 2003 and 2005, representing a 17% increase (Yoshikawa et al., 2007). Additionally, a new national curriculum was introduced in 2004 aimed at aligning preschool education with primary and secondary education. However, the curriculum reform was initially applied in 2004 in only 5% of classrooms to test its viability and required adaptations, delaying its full impact for a year (Secretaria de Educación Pública, 2004).

[12] Appendix 1 details further the content of the reform.

The financial and administrative execution of the reform was complex.[13] States, reliant on federal transfers, faced challenges in expanding preschool infrastructure in such a short time frame (Bennett and Tayler, 2006). Financial constraints, misallocation of resources, and insufficient planning were cited as major impediments. For instance, 90% of educational funds were often allocated to wages rather than infrastructure improvements (Negrete Rosales, 2011). Furthermore, the rollout was uneven across the country, with states like Tlaxcala, Oaxaca, Guerrero, and Nuevo Leon resisting implementation due to inadequate resources and infrastructure (Ortega Montes, 2008).

Despite these obstacles, the mandate's implementation led to a noticeable increase in enrollment, particularly among 4-year-olds, while the increase was less pronounced for 3-year-olds due to parental reluctance to send younger children to school (Yoshikawa et al., 2007). By 2005, nearly all 5-year-olds were enrolled, and the rates for 4-year-olds and 3-year-olds were 81% and 25%, respectively (Yoshikawa et al., 2007). However, variations in compliance and the absence of enforcement mechanisms meant that universal preschool enrollment was not fully achieved, leading to policy adjustments in 2008 that required only one year of preschool prior to primary school enrollment (Behrman et al., 2024).

4 Data

This research aims to determine whether preprimary education enhances cognitive skills by examining the Mexican reform that mandated compulsory preschool. Utilizing detailed data, I first focus on localities that lacked preschools before the reform but established them afterward. I compare the outcomes in these localities to those in localities that still did not have preschools post-reform. Additionally, I complement the analysis by examining the intensive margin, specifically the change in the number of preschools per locality.

I rely on two main data sources: Formato 911, to quantify the number of preschools per year and, thus, the adoption of the reform, and ENLACE, to compare the primary school test scores

[13] The reform was controversial and received large media attention, as several states claimed they lacked the necessary resources and teachers for implementation. Civil society organizations, including Observatorio Ciudadano de la Educación, were vocal critics, arguing that the reform was implemented without adequate studies, planning, or a child-focused program. They advocated for making only one year of preschool mandatory.

of students in treated localities against those in control localities. I compare the cohort of treated students against the cohort of untreated students.

4.1 Sources

Formato 911. The Formato 911 dataset, collected by the Mexican Ministry of Education (SEP) since 1998, is a comprehensive school census that provides detailed information about schools in Mexico. It includes data on school size, class size, and available resources, offering insights into school inputs. It also contains information on the student population and teacher qualifications and credentials. It is collected annually, which allows for tracking changes over time. Each school has a unique identification code and GPS coordinates, which enables linkage with other educational datasets.

ENLACE. The ENLACE (Evaluación Nacional de Logro Académico en Centros Escolares) dataset provides comprehensive standardized test data for primary and secondary school students in Mexico from 2006 to 2013. Administered annually by SEP, ENLACE was a census-based test covering grades 3-9 and 12 nationwide.[14] It assessed student performance in math, Spanish, and a rotating subject each year, creating unique personal identifiers for all test takers to enable longitudinal tracking of academic progress (XABER, 2020). ENLACE was designed to have a national average score of 500 and a standard deviation of 100 for every subject area and grade (De Hoyos Navarro et al., 2018). The dataset includes school identifiers, school GPS coordinates, socioeconomic data, student demographics, student date of birth, and detailed test item diagnostics. Despite limitations like attrition, grade inflation, and data matching challenges (see Appendix 4), ENLACE remains a valuable resource for studying educational trajectories, school inputs, and learning outcomes in Mexico over this period (de Hoyos et al., 2021).

4.2 Sample

I merge the two primary datasets: ENLACE, which provides detailed primary school student-level information, including date of birth, and Formato 911, a census at the preschool level offering

[14]This first assessment focuses on evaluating academic performance in primary school. The aim is to evaluate if the impact on cognitive skills can be identified only a few years later. The methodology employed can easily be replicated to evaluate the impact on secondary school students, which includes a larger time span.

age data upon enrollment.[15] By constructing a cohort variable, data aggregation occurs at the locality-cohort level, facilitating the merging of both datasets under the assumption that students attend primary school in the same locality where they attended preschool or where they would have attended preschool had there been one.[16] This assumption is reasonable, as my findings show that 94% of primary schools have the nearest preschool within the same locality.[17] Table 2 presents baseline descriptive statistics for the sample. Panel A indicates that in the year 2000, 92% of the sample's localities were rural, and 4% had very high or high poverty levels.[18] Panel B shows that 93% of the sample's localities already had a preschool in 2000 and that the share of public (General) preschools was 73%.

The variable used to measure the First Stage in Sections 6 and 7, which is the impact of the reform on pre-k 2 enrollment, is presented in row 17 of Table 2. This approach assesses whether attending preschool serves as a channel through which the reform impacts test scores. On average, there were about 16 students enrolled in pre-k 2 per locality in 2000. The increase to 21.5 students per locality in 2005 suggests a rise in pre-k 2 enrollment.

The variable in row 19 of Table 2 is also pertinent. It reflects the number of preschools standardized by the number of children aged 3-5 in the locality. This standardization facilitates comparisons of the number of preschools across localities. This variable represents the treatment dose utilized in the continuous treatment methodology described in Section 5.2.

Panels C and D of Table 2 present summary statistics on complementing preschool quality mechanisms and the main outcomes, respectively. Sections 6 and 7 expand the definitions of these variables and explain them further.

The varying timing of adoption depicted in Figure 1 motivates the use of a staggered difference-in-differences approach to estimate the reform's impact. However, as Panels (a) and (b) in Figures 1 and A15 show, the noticeable increase in pre-k 2 enrollment starting in 2002 suggests that it is also necessary to consider the potential impact of the reform if all states had adopted it in 2002, when it was published.

[15]The final sample was constructed by merging 86% of the observations in the ENLACE dataset to the Formato 911 dataset, using locality and cohort variables.

[16]Appendix 5 includes further information regarding the construction of the sample.

[17]Furthermore, 37% of preschools and primary schools share the same facilities.

[18]I define a locality as Poor if it has Very High or High poverty levels according to CONEVAL, the public agency responsible for measuring poverty in Mexico (Coneval, 2021). Results for the most disaggregated definition of poverty are presented in Appendix 7.

Table 2: Summary Statistics

Panel A: Localities

	Variable	Before Mandate: year 2000 Obs.	Min.	Mean	Max.	After Mandate: year 2005 Obs.	Min.	Mean	Max.
1	% Rural locs.	10574	0	92.17	100	10655	0	92.24	100
2	% Semi-urban locs.	10574	0	6.92	100	10655	0	6.88	100
3	% Urban locs.	10574	0	0.91	100	10655	0	0.88	100
4	% Indigenous locs.	10574	0	3.48	100	10655	0	3.48	100
5	% Poor locs. (High or Very High)	10574	0	4.06	100	10655	0	4.18	100
6	% Very High poverty locs.	10574	0	0.05	100	10655	0	0.05	100
7	% High poverty locs.	10574	0	4.01	100	10655	0	4.13	100
8	% Medium poverty locs.	10574	0	32.63	100	10655	0	32.73	100
9	% Low poverty locs.	10574	0	35.10	100	10655	0	35.02	100
10	% Very Low poverty locs.	10574	0	28.22	100	10655	0	28.08	100

Panel B First Stage: Preschools and Pre-k 2 Enrollment

	Variable	Before Mandate: year 2000 Obs.	Min.	Mean	Max.	After Mandate: year 2005 Obs.	Min.	Mean	Max.
11	% of Localities with a preschool	10574	0	92.54	100	10655	0	93.60	100
12	% of General preschools in loc.*	9813	0	73.04	100	10000	0	72.24	100
13	% of Community preschools in loc.*	9813	0	13.13	100	10000	0	13.57	100
14	% of Indigenous preschools in loc.*	9813	0	13.44	100	10000	0	13.25	100
15	% of Private preschools in loc.*	9813	0	0.39	100	10000	0	0.94	100
16	Total preschools in loc.	10574	0	1.16	13	10655	0	1.22	16
17	Total pre-k 2 students in loc.	10574	0	15.98	654	10655	0	21.50	759
18	% of female pk 2 students in loc.**	9465	0	53.94	100	9738	0	54.23	100
19	Preschools per pop. aged 3-5 in loc.+	10549	0	0.0450	1	10628	0	0.0467	1
20	Pre-k 2 enrollment (%) in loc.+	10549	0	93.38	6900	10628	0	116.07	9000
21	Pre-k 2 students per preschool in loc.	9785	0	11.28	101	9973	0	14.21	121

Panel C Mechanisms: Preschool Quality

	Variable	Before Mandate: year 2000 Obs.	Min.	Mean	Max.	After Mandate: year 2005 Obs.	Min.	Mean	Max.
22	% of pk teachers with Licenciatura^o	9706	0	48.34	100	9886	0	52.39	100
23	% of pk teachers with Bachillerato^o	9706	0	84.26	100	9886	0	83.99	100
24	% of pk teachers with Normal PK^o	9706	0	13.09	100	9886	0	8.59	100
25	Pk 2 students per Pk 2 class in loc.#	9688	0	10.08	64	9874	0	12.49	61
26	Pk 2 students per Pk 2 teacher***	7326	0	10.91	64	7408	0	13.55	61

Panel D Outcomes: Test scores

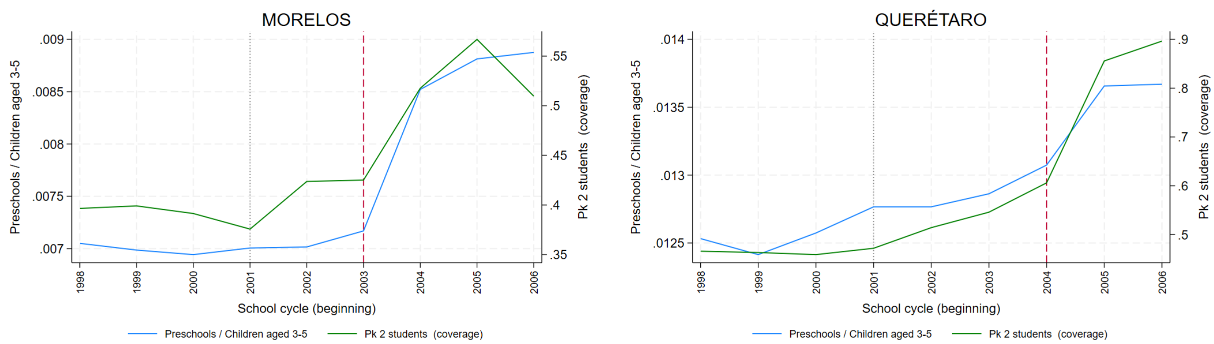
	Variable	Before: cohort 1995-1996 Obs.	Min.	Mean	Max.	After: cohort 2000-2001 Obs.	Min.	Mean	Max.
27	Grade 3 average test scores++	8723	297.36	438.00	658.29	10596	325.63	497.18	694.55
28	Grade 4 average test scores++	10412	311.55	464.14	652.59	10608	329.16	498.29	754.14
29	Grade 5 average test scores++	10543	327.86	472.33	650.47	10619	339.30	516.03	737.14
30	Grade 6 average test scores++	10557	314.89	477.76	681.73	10655	329.68	532.42	746.46

Source: Table generated by the author using data from Formato 911, ENLACE, INEGI, and CONEVAL.

Notes: Data at the locality level. Localities are included in the sample if they have a primary school in the ENLACE data. The sample includes the 11 states with a consistent September 1st cutoff for both primary and pre-primary school. These 11 states include Baja California, Campeche, Coahuila, Guanajuato, Jalisco, Morelos, Queretaro, San Luis Potosi, Sonora, Veracruz, and Zacatecas. *: Conditional on the locality having a preschool. **: Conditional on the school reporting the sex of the students. +: Conditional on the locality having population aged 3-5 according to INEGI. ^{o}: These do not include community preschools since they do not report teachers' schooling. #: Conditional on the schools reporting the number of pre-k 2 classrooms. ***: Conditional on the schools reporting the number of pre-k 2 teachers; these do not include community preschools since they do not report teachers per grade. ++: Conditional on ENLACE providing data for the cohort in the certain grade.

Figure 1: Dynamic Adoption of the Reform

(a) Adoption in 2004 (b) Adoption in 2005



Source: Figure generated by the author using data from Formato 911 and INEGI.
Notes: This Figure shows examples of states that adopted the reform in different years. Panel (a) reveals that Morelos experienced the most significant increase in preschools in 2004, whereas Panel (b) indicates that Querétaro saw the largest increase in 2005. Additionally, both panels illustrate a notable rise in pre-k 2 enrollment starting in 2002. The different timing of adoption across all states is included in Appendix 6 Figure A4.

5 Methodology

The main objective is to assess test score differences between treated localities (where preschools were established due to the reform) and control localities. This analysis focuses on cohorts of students within these localities, comparing outcomes between those exposed to preschool availability and those not. The study employs panel data at the locality-cohort level.

The reform mandating preschool attendance for children and the provision of preschools by states created an exogenous change in educational policy. While parental decisions regarding preschool attendance can be endogenous and less observable, the legislative mandate induced states to provide preprimary schooling. Due to resource constraints and other obstacles (see Negrete Rosales (2011) and Ortega Montes (2008)), not all localities opened preschools under this reform. Initially, I utilize the opening of preschools in localities that did not have any preschools to compare post-reform outcomes between localities with and without preschools within the set of states that adopted the reform.

5.1 Canonical Difference-in-Differences

I employ a two-way fixed effects model to compare test scores in treated localities to those in untreated localities. I examine one period before and one after the reform, comparing the treated and untreated cohorts. The causal estimand of interest is the average causal effect on treated localities for the treated cohorts (ATT). The key assumption for identifying the causal effect is the assumption of parallel trends. Without treatment (a preschool), the average outcomes for the treated and untreated localities would have evolved similarly. This is very intuitive in this setting since the comparison is between localities with no preschools before the reform. The two-way fixed effects regression specification is:

$$Y_{i,t} = \alpha_i + \phi_t + (C_t \cdot D_i)\beta + \epsilon_{i,t} \tag{1}$$

Where $Y_{i,t}$ represents the average test scores in locality i for cohort t, α_i is the locality-specific fixed effect, ϕ_t is the cohort-specific fixed effect, C_t is an indicator function that equals 1 if the cohort is affected by the reform and 0 otherwise, D_i is the treatment variable, β is the coefficient representing the treatment effect, and $\epsilon_{i,t}$ is the error term.

As mentioned, treatment is defined at the locality level by the presence of a preschool post-reform in the set of localities belonging to adopting states. In the setting where all states adopted the national reform when it was published in 2002,[19] the treatment is defined as:

$$D_i = \mathbb{1}(Total\ preschools_i^{2002} > 0 | (Total\ preschools_i^r = 0\ \forall\ r < 2002)) \tag{2}$$

For each locality i and year r. This methodology considers only localities that had no preschools before 2002 and compares the treated set of localities that had at least a preschool in 2002 to those localities that remained without preschools in 2002. Considering the year 2002 is consistent with the anticipatory increase in preschool enrollment observed in the data (see Figures 1 and A15), which took place before the mandated implementation date.

Following the reform, the greatest increase in preschool enrollment was observed for 4-year-olds, as shown in Table 1. Thus, the treated cohort encompasses those children born between September 1997 and August 1998 and were old enough (4 years old) for pre-k 2 enrollment in 2002.[20] The ideal grade progression for the treatment and control cohorts is shown in Table 3.

[19] Compulsory preschool was approved by the Mexican Congress in May 2002 before the start of the school cycle in Septemeber. The decree was published in November 2002. The legislative process started in October 2001. A UNICEF report (Yoshikawa et al. (2007)) documents a large increase in pre-k 2 enrollment rates since 2002.

[20] The school cycle in Mexico begins in September. A child must reach the appropriate age for that grade before September 1 to enroll in that grade at the start of the school cycle. Appendix 3 includes more information on the Mexican Education System.

Section 6 shows the main results using the TWFE methodology described in Equation 1, which considers that all states adopted the reform in 2002. However, Figure 1 shows that there is evidence of dynamic adoption; some states adopted the reform in different years after 2002. Appendix 10 displays the results for the TWFE methodology using different adoption assumptions. Section 5.2 introduces the methodology of staggered DiD employed in this study.

Table 3: Observable Cohorts Along Grades

	Control cohort Sep. 96 - Aug. 97	**Treatment cohort** Sep. 97 - Aug. 98 (aged 4 before Sep 1st, 2002)
Pre-k 2	2001	2002
Pre-k 3	2002	2003
Grade 1	2003	2004
Grade 2	2004	2005
Grade 3	2005	2006
Grade 4	2006	2007
Grade 5	2007	2008
Grade 6	2008	2009

Source: Table generated by the author.

5.2 Staggered DiD

The data indicate that the largest increases in preschool enrollment occurred in different years across various states, implying staggered adoption of the reform by different states (see Figure 1). Furthermore, although states were mandated to provide preprimary schooling starting in 2004, evidence shows anticipation and early adoption as early as 2002. These characteristics pose a challenge because the TWFE methodology in Section 5.1 only yields causal estimates under the assumption of treatment effect homogeneity across units and time.[21] The setup of this study encompasses heterogeneous treatment effects across units, more than two periods, and variations in treatment timing. Thus, a Static TWFE estimand may not yield a causal parameter due to negative weighting (Borusyak et al. (2024); De Chaisemartin and d'Haultfoeuille (2020); Callaway

[21]Other assumptions are also needed for causality in a static TWFE framework such as parallel trends and no anticipatory effects (Roth et al., 2023). In staggered DiD, these two assumptions are relaxed into the following:

1. Parallel Trends for Staggered Setting: Assumes that without treatment, average outcomes for all adoption groups would have evolved in parallel, including non-adopters.
2. Staggered No Anticipation Assumption: Assumes that a unit's outcome when untreated is not influenced by knowledge of future treatment.

and Sant'Anna (2021)). Similarly, Sun and Abraham (2021) showed that the estimands of a Dynamic TWFE may also be biased for some units some periods after treatment due to cross-lag "contamination" and "forbidden comparisons".[22] De Chaisemartin and d'Haultfoeuille (2023) and Roth et al. (2023) provide an overview of the key developments in the Difference-in-Differences literature to address these and other issues when using TWFE, including the development of the staggered DiD designs.

Staggered DiD designs offer advantages over traditional DiD by allowing the analysis of treatments implemented at different times across units. This approach provides greater flexibility in control group selection, increases statistical power, and enables the examination of dynamic and heterogeneous treatment effects. Importantly, it allows to graphically assess parallel trends across multiple pre-treatment periods, enhancing the credibility of the parallel trends assumption. Additionally, staggered DiD facilitates the visualization of effect persistence over time, providing insights into whether treatment impacts are temporary or long-lasting. These features, combined with the ability to use "not-yet treated" units as controls and incorporate covariates, make staggered DiD a robust and informative method for analyzing causal effects in settings with staggered policy adoption (De Chaisemartin and d'Haultfoeuille (2023); Roth et al. (2023)).

Most of the staggered DiD literature has focused on a setup in which the treatment is binary and being treated is an absorbing state (Roth et al. (2023); Borusyak et al. (2024); De Chaisemartin and d'Haultfoeuille (2023), Gardner (2022)). A complementing strand of the literature analyzes the case of continuous treatment (De Chaisemartin and d'Haultfoeuille (2024); de Chaisemartin et al. (2022); Callaway et al. (2024)). The characteristics of the data utilized in this study present some issues that complicate the use of the staggered DiD with a binary treatment, such as the one presented in Equation 2. The few years and cohorts for which ENLACE data is available restrict the data to an unbalanced panel where not all localities are observed for all cohorts. Moreover, the low share of localities without a preschool before the mandate restricts the sample and power. For this reason, the main results of the paper shown in Section 7 consider a staggered DiD with a continuous treatment, following De Chaisemartin and d'Haultfoeuille (2024) and de Chaisemartin et al. (2022). The staggered DiD with binary treatment results are included in Appendix 9.

[22] A clean comparison would be between a treated unit and a not-yet-treated unit, or between a treated unit and a never-treated unit.

The staggered DiD with continuous treatment methodology can be used in a setup such as the one presented in this study. There can be multiple groups and periods, the panel can be unbalanced, and it does not require an absorbing state. Continuous treatments can offer advantages over binary ones. Variation in intensity allows to evaluate treatments that all units receive. It might also be more relevant to study changes in doses, such as the rate of preschool coverage in a locality, instead of the existence of a preschool. However, the continuous treatment methodology requires a stronger parallel trends assumption: the evolution for lower-dose units must reflect how higher-dose units' outcomes would have changed had they experienced the lower dose instead (Callaway et al., 2024). Thus, the evolution of test scores in localities with few preschools should have evolved similarly to the evolution of test scores in localities with many preschools in the absence of an increase in preschools. I assess the compliance of the stronger parallel trends assumption with event study graphs using the methodology described in De Chaisemartin and d'Haultfoeuille (2024) and de Chaisemartin et al. (2022).

As mentioned in Section 4.2, I employ as a treatment variable the number of preschools per locality standardized by the population aged 3-5 in each locality (variable in row 19 of Table 2) in this staggered DiD methodology with continuous treatment. This allows me to compare the dose of preschools (the intensity of treatment) across localities to assess the impact of the reform.[23] Section 7 presents the main results and Appendix 7 presents complementing results for this methodology.

6 TWFE Results

This section presents the results of the TWFE methodology, assuming all states adopted the reform in 2002 following its approval by the Mexican Congress. Complementary results using this identification strategy are included in Section Appendix 8.

Although the difference-in-differences methodology does not require that the treatment and control groups have identical characteristics at baseline, a balance table is useful for assessing the comparability of these groups before the intervention. Table 4 evaluates the comparability between the control and treatment groups as defined by Equation 2. The table indicates no significant

[23] I compute the results using the Stata package *did_multiplegt_dyn* developed by De Chaisemartin and d'Haultfoeuille (2023) and De Chaisemartin and d'Haultfoeuille (2024). I utilize clustered standard errors at the locality level and allow for group-specific linear trends when estimating the treatment effects.

differences between the groups prior to the reform, suggesting that in the absence of treatment, the outcomes would have evolved similarly in both groups. These results also help to rule out the existence of confounding factors. Localities that did not have preschools before 2002 and opened a preschool in 2002 are very similar to those that remained without a preschool. Table A7 in Appendix 8 shows that 99% of these localities are rural.

Table 4: Balance Table for Localities that had no preschool before 2002

Variable	(1) Control	(2) Treatment	(3) T-C
Average schooling	3.928	4.000	0.072
	(1.783)	(1.371)	(0.256)
Total population	260.362	248.863	-11.499
	(384.311)	(181.921)	(54.369)
% Female population aged 15-49	22.419	22.425	0.005
	(6.137)	(4.701)	(0.881)
% Labor force participation -%18+-	53.342	54.132	0.789
	(15.111)	(12.648)	(2.203)
% LFP in agriculture -%18+-	61.722	65.187	3.465
	(28.857)	(26.913)	(4.237)
% LFP in manufacturing -%18+-	19.486	19.558	0.072
	(19.252)	(20.945)	(2.863)
% LFP in services -%18+-	16.585	14.049	-2.536
	(16.705)	(15.542)	(2.452)
% of Indigenous language speakers	14.062	16.760	2.698
	(31.075)	(32.924)	(4.610)
% of population aged 18-24	18.303	18.690	0.387
	(6.292)	(5.643)	(0.912)
% Dwellings with dirt floor	44.513	51.653	7.140
	(32.119)	(32.938)	(4.751)
% of Poor localities -very high or high-	0.121	0.082	-0.040
	(0.327)	(0.277)	(0.048)
% of rural localities	0.993	1.000	0.007
	(0.083)	(0.000)	(0.012)
% of semi-urban localities	0.007	0.000	-0.007
	(0.083)	(0.000)	(0.012)
% of urban localities	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Observations	583	51	634

Source: Table generated by the author using data from INEGI and CONEVAL.
Notes: Treatment is defined as shown in Equation 2. Both groups are large. There are no significative differences between the Control and Treatment groups. Table A7 shows complementing summary statistics.

Panel B of Table 5 further evaluates the parallel trends assumption by testing for the existence of pre-trends, examining the impact of the intervention on test scores for untreated cohorts. In the absence of treatment, there should be no difference between groups, which is precisely what Panel B of Table 5 shows.

The main results of this section, following the methodology outlined in Equations 1 and 2, are shown in Panel A of Table 5. Column (1) demonstrates a strong First Stage, indicating that opening

Table 5: Effect of the Reform for Localities that had no preschool before 2002

VARIABLES	First Stage	Outcomes: Average Primary School Test Scores			
	Pk-2 students	Grade 3	Grade 4	Grade 5	Grade 6
		Panel A: Treatment			
	(1)	(2)	(3)	(4)	(5)
$\hat{\beta}$	4.6667***	14.5321*	19.5773**	6.1438	15.0598*
	(0.1619)	(8.4944)	(8.2787)	(7.7848)	(7.8005)
Treated locs.	0.0000	-10.5352	-14.6949**	-6.0735	-12.2807*
	(0.1145)	(7.5111)	(6.7456)	(7.2814)	(7.4194)
$\mathbb{1}$(C=1997)	0.0000	-1.7147	-5.6030**	4.1389**	16.7157***
	(0.0459)	(2.3616)	(2.2505)	(2.0938)	(2.1534)
Constant	-0.0000	461.4901***	462.4429***	468.7911***	480.0977***
	(0.0325)	(2.2592)	(2.1509)	(2.1426)	(2.2307)
Observations	1,268	1,268	1,268	1,268	1,268
R-squared	0.5783	0.0016	0.0046	0.0024	0.0301
Test	First Stage	Treatment	Treatment	Treatment	Treatment
Baseline	0	460.64	461.26	468.3	479.11
Years	2001 vs 2002	2001 vs 2002	2001 vs 2002	2001 vs 2002	2001 vs 2002
		Panel B: Pre-trends			
		(6)	(7)	(8)	(9)
$\hat{\beta}$		-2.4358	-5.1900	-9.7804	-6.2359
		(8.9039)	(6.4443)	(6.7071)	(7.3822)
Treated locs.		-8.0993	-9.5048	3.7070	-6.0448
		(9.8664)	(6.1483)	(7.0091)	(6.5751)
$\mathbb{1}$(C=1996)		16.7903***	1.8298	1.5836	7.6819***
		(2.8686)	(2.0908)	(2.0109)	(2.1358)
Constant		444.6998***	460.6131***	467.2075***	472.4158***
		(2.9805)	(2.2243)	(2.1668)	(2.2529)
Observations		1,268	1,268	1,268	1,268
R-squared		0.0217	0.0044	0.0008	0.0069
Test		Pretrends	Pretrends	Pretrends	Pretrends
Baseline		444.06	459.86	467.51	471.93
Years		2000 vs 2001	2000 vs 2001	2000 vs 2001	2000 vs 2001

Clustered standard errors at the locality level in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Table generated by the author using data from Formato 911, ENLACE, and INEGI.
Notes: ENLACE test scores were designed to have a national average score of 500 and a standard deviation of 100 for every subject area and grade. Panel A shows the treatment effects for Equation 1 for the cohorts presented in Table 3; it compares the cohort affected by the reform (born in September 1997 - August 1998) to the cohort unaffected by the reform (born in September 1996 - August 1997).
Panel B shows the results for Equation 1 for the pre-trends; it compares unaffected cohorts (born in September 1996 - August 1997 and September 1995 - August 1996). The pre-trend for pre-k 2 enrollment is zero by construction since there are no preschools in those localities.
Table A8 disaggregates the results for math and Spanish test scores.

a preschool in a locality that previously had none increases pre-k 2 enrollment by approximately 5 students.[24] Columns (2), (3), and (5) show that the intervention caused an increase in average test scores by 0.15-0.20 SD, a significant improvement. Column (5) provides evidence that this increase in academic performance persists through to primary school graduation.

Table A8 separates the results for math and Spanish test scores. It reveals that the significant increase in test scores is primarily driven by math scores, which rise by 0.19-0.22 SD. This finding is intuitive since math abilities are typically more closely related to skills learned at school rather than at home. Behrman et al. (2024) also find larger impacts for math than for Spanish. It is important to mention that the math estimates in Table A8 for a specific population subsample are larger than the 0.07-0.11 SD increase found by Behrman et al. (2024) for the entire population. This is because the subsample used in this study had never previously had a preschool in their locality.

7 Staggered DiD Results

7.1 Main Results

Figure 2 presents the main outcomes of the staggered DiD with continuous treatment methodology, as described in Section 5.2. Panel (a) reveals a significant increase in pre-k 2 enrollment, with an approximate rise of 3-4 students, comparable to the First Stage results in Table 5. This increase is substantial relative to the baseline of 16 students shown in row 17 of Table 2. The graph in Panel (a), indicating parallel trends, also demonstrates that this rise is noticeable for both girls and boys, with a slightly higher impact on girl enrollment.

Despite the increase in pre-k 2 enrollment, Panel (d) in Figure 2 shows a reduction in class size and student-to-teacher ratios. The expansion of classrooms and teachers outpaced the growth in pre-k 2 students, thereby enhancing preschool quality. Although the evidence for parallel trends is not perfect, the reduction in both ratios (and corresponding quality improvement) is clear. The decrease ranges between 1 and 3 students, which is significant compared to the baseline levels of 10 and 11 students per class and per teacher, respectively (rows 25 and 26 of Table 2).

[24] Mechanisms such as preschool class size and teachers' education cannot be tested using this methodology due to a lack of observations.

The reform stipulated that preschool teachers should have at least a university degree (Licenciatura). However, Panel (c) of Figure 2 indicates a persistent decline in the proportion of teachers with a university degree, including those specialized in preschool teaching (Normal Preescolar). Surprisingly, the share of teachers with a high school diploma (Bachillerato) also declined, though this was a temporary trend.

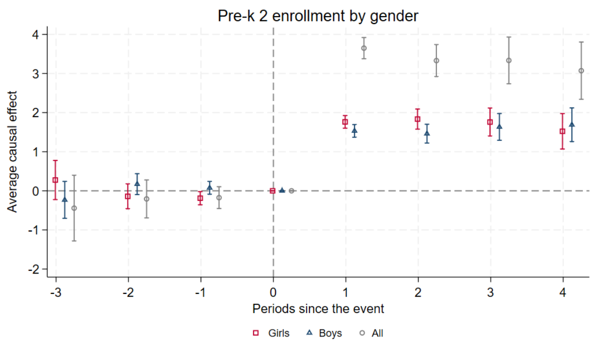
The findings from Panel (c), along with those in Panels (a) and (d), imply that the rise in teacher numbers to meet the higher demand was primarily due to an increase of low-skilled teachers. Consequently, the quality improvements observed in Panel (d) are inhibited by the lower educational attainment of these teachers. In a scenario where preschool teachers lack adequate skills, the benefits of early childhood education can be greatly diminished. Low-skilled teachers can hinder the development of students' cognitive skills, leading to negative impacts on their posterior academic performance.

Nevertheless, Panel (b) in Figure 2 indicates that test scores not only avoided deterioration but improved over the longer term. Specifically, Grade 5 test scores show a significant increase, ranging from 0.05 to 0.16 SD. Grade 6 also exhibits suggestive evidence of positive impacts, with significant coefficients using 90% confidence intervals. This aligns with existing literature, which suggests that the benefits of cognitive skills developed through early childhood education often become more evident in the long term due to the cumulative and compounding nature of learning. Early cognitive skills provide a foundation for later learning, enabling children to acquire and integrate new knowledge more efficiently over time. The full benefits of early childhood education become more evident as students encounter increasingly complex academic challenges in the later years of primary school.

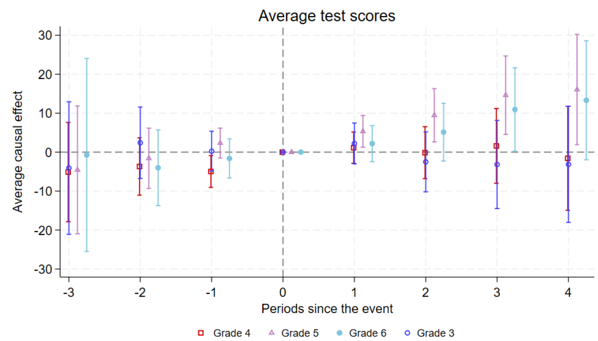
The findings shown in Figure 2 indicate that the benefits of smaller preschool groups and extended time in preschool (in the intensive or extensive margin) offset the potential negative impact of lower-skilled teachers, resulting in overall cognitive skill improvement in children. The interpretation of these results is complemented with a consistent overall interpretation from results shown in Table 5, which also show an improvement in test scores and, thus, cognitive skills.

Figure 2: Staggered DiD with Continuous Treatment: Main Results

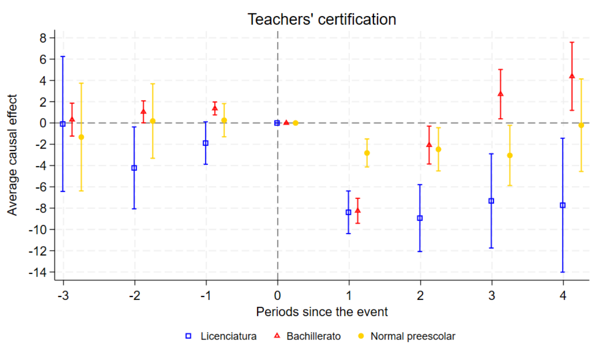
(a) First Stage: Pre-k 2 Enrollment



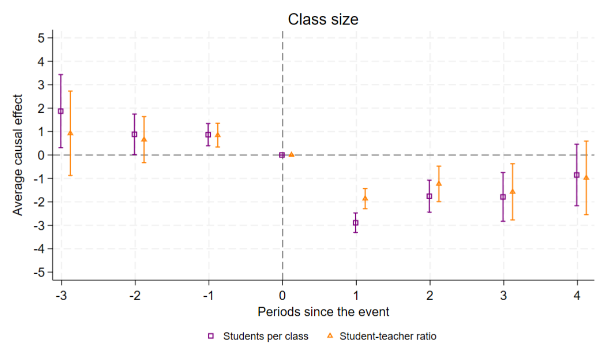
(b) Outcomes: Average Test Scores



(c) Mechanisms: Preschool Teachers' Education



(d) Mechanisms: Preschool Class Size



Source: Figure generated by the author using data from Formato 911, ENLACE, and INEGI.
Notes: All graphs include 95% confidence intervals. Section 5.2 explains the used methodology. The continuous treatment variable is the number of preschools per locality standardized by the population aged 3-5 in each locality (variable in row 19 of Table 2). The results are computed using the Stata package *did_multiplegt_dyn* developed by De Chaisemartin and d'Haultfoeuille (2023) and De Chaisemartin and d'Haultfoeuille (2024). They include clustered standard errors at the locality level and allow for group-specific linear trends when estimating the treatment effects. Panel (a) relies on data from Formato 911 on total pre-k 2 enrollment (variable in row 17 of Table 2) and by gender. Panel (b) relies on data from ENLACE on math and Spanish test scores, which are averaged for simplicity purposes (variables in rows 27-30 of Table 2). ENLACE test scores were designed to have a national average score of 500 and a standard deviation of 100 for every subject area and grade.
Panel (c) relies on Formato 911 preschool data for all pre-primary school grades and all types of preschools except Community preschools since that information is not available for those schools. This panel plots the share of teachers with a certain schooling level (variables in rows 22-24 of Table 2). These levels are defined as "Licenciatura": a teacher with an undergraduate degree, as mandated by the reform; "Bachillerato": a teacher with a high school degree; "Normal preescolar": a teacher with an undergraduate degree specialized in preschool teaching.
Panel (d) relies on Formato 911 preschool data for pre-k 2. It plots the class size and student-to-teacher ratio (variables in rows 25 and 26, respectively, of Table 2). Data on students per teacher include information for pre-k 2 for all types of preschools except Community preschools since that information is not available for those schools. Data on students per class include information for pre-k 2 for all types of preschools.

7.2 Distributional Effects

This section delves into the effect of the reform on test scores across different demographic groups. Specifically, Figure 3 illustrates the impact on Grade 5 average test scores by gender, poverty, indigenous status, and rural status. Complementary results for Grades 3-6 average test scores are provided in Figures A5-A8.[25] Additionally, Figures A9-A13 show results related to pre-k 2 enrollment and quality mechanisms detailed in Table 2.

Gender. Figure 3 expands on Panel (b) of Figure 2. Panel (a) of Figure 3 indicates a more pronounced increase in Grade 5 test scores for boys, in both math and Spanish (see Panel (c) of Figure A7). This result is unexpected given the reform slightly increased relatively more girl pre-k 2 enrollment (Panel (a) of Figure A9). The mechanism behind this effect is not detailed in this paper, but it is possible that preschool environments might better cater to boys' learning styles or needs, providing them with a greater advantage by the end of primary school.

Poverty. Panel (b) of Figure 3, along with Panels (d) and (e) of Figure A7, suggests that students in non-poor localities, particularly those with medium poverty, experienced larger improvements in test scores compared to students in the poorest localities. This aligns with the substantial increase in pre-k 2 enrollment in non-poor localities (particularly medium poverty localities), as shown in Panels (e) and (f) of Figure A9. Furthermore, class sizes decreased more in these non-poor localities (Panels (d) and (e) of Figure A10). This indicates that the reform was ineffective in reducing learning disparities between the richest and poorest localities but succeeded in narrowing the gap between medium-poverty and very low-poverty localities.

Indigenous Localities. Panel (c) of Figure 3 shows that the reform led to an increase in Spanish test scores for students in indigenous localities, while math test scores remained unchanged. This contrasts with the overall positive effect on both Spanish and math test scores (Panel (a) of Figure A7). The lack of improvement in math test scores can be attributed to unchanged pre-k 2 enrollment

[25] Grade 5 test scores are preferred for interpretation due to clearer evidence of parallel trends compared to Grade 6. Nonetheless, the interpretation for Grade 6 is consistent with that for Grade 5. Grades 3 and 4 are not included in the main result as the benefits of skills acquired in preschool become more evident over the longer term.

in indigenous localities (Panel (d) of Figure A9), despite a decrease in class sizes (Panel (c) of Figure A10), which likely drove the increase in Spanish test scores.

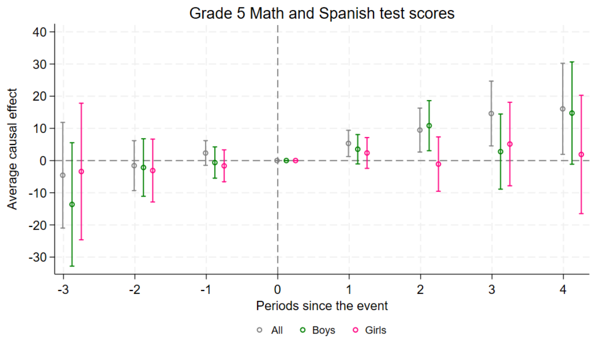
Rural Localities. Panel (d) of Figure 3 and Panel (i) of Figure A7 indicate that the increase in test scores is primarily driven by gains in rural localities. This is consistent with the significant rise in pre-k 2 enrollment in rural areas (Panel (b) of Figure A9). Although pre-k 2 enrollment also increased in semi-urban settings, class sizes decreased more in rural areas than in semi-urban settings (Panel (a) of Figure A9), which is the mechanism behind the improved test scores in rural localities. Consequently, the reform helped reduce learning disparities between rural and more urbanized localities.

Other Results. Panel (g) of Figure A7 shows that the increase in test scores is most evident in public (General) schools, which saw the largest increase in pre-k 2 enrollment (Panel (c) of Figure A9). Additionally, Figure A14 provides evidence of a one-year delay in primary school entry, consistent with findings from Behrman et al. (2024).[26]

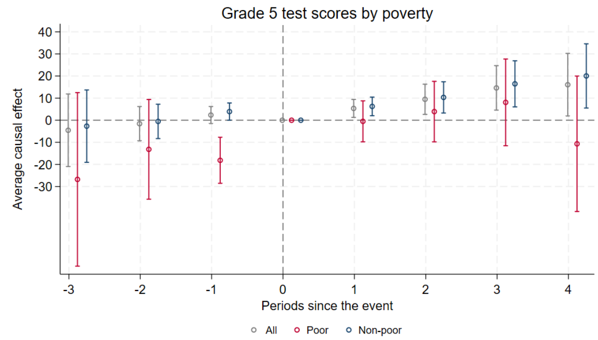
[26] It is important to remark that in this study, I examine test scores based on student cohorts rather than their current grade level, ensuring that a one-year delay does not bias the results.

Figure 3: Distributional Effects on Grade 5 Test Scores

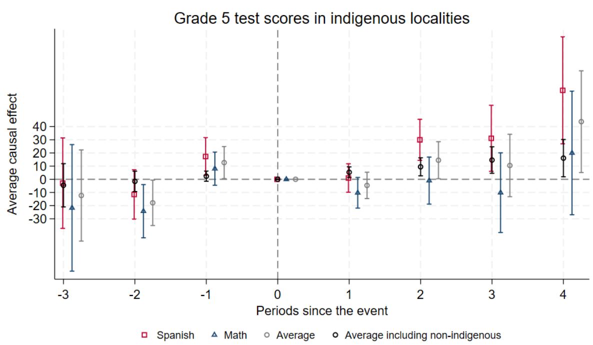
(a) Gender



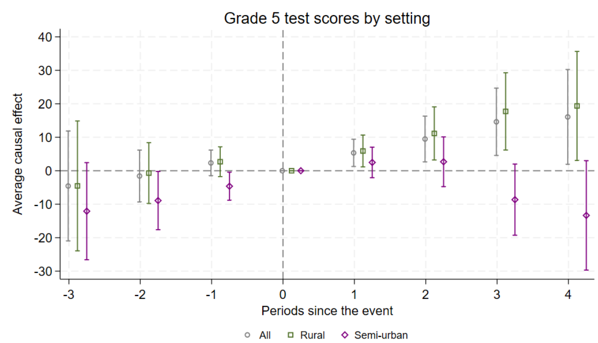
(b) Poverty



(c) Indigenous Localities



(d) Setting



Source: Figure generated by the author using data from Formato 911, ENLACE, INEGI, and CONEVAL.
Notes: ENLACE test scores were designed to have a national average score of 500 and a standard deviation of 100 for every subject area and grade. All graphs include 95% confidence intervals. Section 5.2 explains the used methodology. The continuous treatment variable is the number of preschools per locality standardized by the population aged 3-5 in each locality (variable in row 19 of Table 2). The results are computed using the Stata package *did_multiplegt_dyn* developed by De Chaisemartin and d'Haultfoeuille (2023) and De Chaisemartin and d'Haultfoeuille (2024). They include clustered standard errors at the locality level and allow for group-specific linear trends when estimating the treatment effects.

8 Robustness

The robustness checks using alternative estimators and identification strategies reinforce some of the main findings but also reveal some limitations and nuances. The staggered DiD with binary treatment results in Appendix 9, despite having less statistical power than those in Section 7, confirm an increase in pre-k 2 enrollment, consistent with the main results in Section 7.1. However, the impact on test scores is less clear. In Appendix 9.1, when defining the adoption year using maximum pre-k 2 enrollment growth, which excludes 2002 as an adoption year, the coefficients are negative and mostly non-significant. In Appendix 9.2, when defining adoption years using maximum total preschool growth and, thus, including 2002 as an adoption year, the test score coefficients are positive and mostly non-significant. This suggests that while the reform did increase enrollment, its effects on cognitive outcomes might not be uniformly positive across different adoption timings, potentially due to the varying quality of implementation, i.e., the hiring of low-skilled teachers, anticipation, or other sample-related bias. The staggered DiD with continuous treatment methodology addresses these shortcomings by including all the adoption years in the sample, increasing the statistical power, and enabling the analysis of the quality mechanisms.

The results for the TWFE methodologies in Appendix 10.1 and Appendix 10.2, though limited by low statistical power, also support an increase in pre-k 2 enrollment, aligning with the findings in Section 6. However, they show a non-significant negative effect on test scores when focusing on 2004 as the adoption year (similar to Appendix 9.1), and they offer weaker evidence regarding parallel trends than results in Section 6. Moreover, when considering 2005 as the adoption year, there is one negative and significant coefficient, while all others are non-significant. These inconsistencies underscore the importance of the adoption year and indicate that the results are sensitive to the timing of the reform's implementation.

The results in Appendix 10.3 rely on a TWFE model with an alternative identification strategy that offers more statistical power than those in Section 6, Appendix 10.1, and Appendix 10.2, though it compares localities that are less similar before the reform. These results focusing on 2004 and 2005 as adoption years also indicate an increase in pre-k 2 enrollment and provide evidence of parallel trends. As for test scores, all coefficients are non-significant, and most of them are positive. The approach in Appendix 10.3 is useful for assessing the impact on quality mechanisms, similar to Panels (c) and (d) of Figure 2, but with a TWFE approach. Tables A19 and A22 reveal a much larger decline in teachers' educational attainment, particularly for 2004 adoption, compared to Panel (c) of Figure 2.

The robustness checks suggest that while the findings on pre-k 2 enrollment are generally consistent, the reform's impact on test scores may depend on the specific context and timing of adoption. The staggered DiD with continuous treatment methodology mitigates these issues by incorporating all adoption years into the sample, enhancing statistical power, and allowing for a more thorough analysis of quality mechanisms.

9 Conclusion

The findings from the study provide comprehensive insights into the implementation and impact of the preschool education reform in Mexico. The results underscore the significance of early childhood education and its medium-term benefits on cognitive development, particularly in math skills. The reform led to substantial increases in preschool enrollment since 2002, especially among 4-year-olds, and was accompanied by improved educational outcomes, particularly in the last two years of primary school. Results from the staggered difference-in-differences analysis reveal a significant improvement in Grade 5 test scores, with increases ranging from 0.05 to 0.16 SD. The ATT estimates reveal that localities that established their first preschool due to the reform saw an increase in Grade 6 test scores by 0.15 SD. This impact is more pronounced in math test scores, which increased by 0.19 SD. These results highlight that the cognitive skills developed in preprimary school persist through the end of primary school, indicating no fade-out effect.

The reform was successful in increasing preschool quality regarding class size; however, challenges regarding teacher qualifications persisted. While the expansion of preschool infrastructure and the hiring of new teachers addressed some of the logistical challenges, the influx of less-qualified teachers highlighted the need for continued focus on teacher training. The findings suggest that while the quantity of educational inputs was successfully increased, attention to the quality of these inputs remains crucial.

The analysis of distributional effects highlighted important disparities. Boys experienced slightly higher benefits from the reform in terms of test score improvements, a finding that warrants further investigation into gender-specific educational needs and strategies. The positive effects observed in rural communities indicate some progress in reducing educational inequities. However, the impact of the reform was more favorable in medium-poverty and non-poor localities compared to the poorest areas, suggesting that the reform was ineffective in breaching learning gaps for the most marginalized communities. Moreover, the lack of improvement in math test scores in indigenous communities, driven by stagnant pre-k 2 enrollment, suggests that marginalized communities still require targeted support to sustain and amplify gains from similar policies.

The methodology employed, which included a staggered difference-in-differences approach, was crucial in accurately assessing the causal effects of the reform. By addressing issues related to treatment effect heterogeneity and dynamic adoption across different states, the study was able to provide more precise estimates of the reform's impact.

The 2002 Mexican preschool education reform significantly improved early childhood education, promoted cognitive skill formation, and addressed certain developmental disparities. This study provides evidence for the effectiveness of formal early childhood education in a developing country context. Additionally, the findings offer valuable insights for researchers and policymakers involved in early childhood education. The use of staggered difference-in-differences, a robust methodological framework, ensured a comprehensive analysis of the reform's impact, providing a reliable model for evaluating similar educational policies in the future.

References

Abraham, K. G. and M. S. Kearney (2020). Explaining the decline in the US employment-to-population ratio: A review of the evidence. *Journal of Economic Literature 58*(3), 585–643.

Almond, D., J. Currie, and V. Duque (2018). Childhood circumstances and adult outcomes: Act II. *Journal of Economic Literature 56*(4), 1360–1446.

Andresen, M. E. and T. Havnes (2019). Child care, parental labor supply and tax revenue. *Labour Economics 61*, 101762.

Andrew, A., O. Attanasio, B. Augsburg, L. Cardona-Sosa, M. Day, M. Giannola, S. Grantham-McGregor, P. Jervis, C. Meghir, and M. Rubio-Codina (2024). Early Childhood Intervention for the Poor: Long Term Outcomes. Technical report, National Bureau of Economic Research.

Andrew, A., O. Attanasio, B. Augsburg, M. Day, S. Grantham-McGregor, C. Meghir, F. Mehrin, S. Pahwa, and M. Rubio-Codina (2020). Effects of a scalable home-visiting intervention on child development in slums of urban India: evidence from a randomised controlled trial. *Journal of Child Psychology and Psychiatry 61*(6), 644–652.

Andrew, A., O. P. Attanasio, R. Bernal, L. C. Sosa, S. Krutikova, and M. Rubio-Codina (2024). Preschool quality and child development. *Journal of Political Economy 132*(7), 000–000.

Angrist, N., S. Djankov, P. K. Goldberg, and H. A. Patrinos (2021). Measuring human capital using global learning data. *Nature 592*(7854), 403–408.

Araujo, M. C., M. Dormal, S. Grantham-McGregor, F. Lazarte, M. Rubio-Codina, and N. Schady (2021). Home visiting at scale and child development. *Journal of Public Economics Plus 2*, 100003.

Attanasio, O., S. Cattan, and C. Meghir (2022). Early childhood development, human capital, and poverty. *Annual Review of Economics 14*(1), 853–892.

Bailey, D., G. J. Duncan, C. L. Odgers, and W. Yu (2017). Persistence and fadeout in the impacts of child and adolescent interventions. *Journal of Research on Educational Effectiveness 10*(1), 7–39.

Baker, M., J. Gruber, and K. Milligan (2008). Universal child care, maternal labor supply, and family well-being. *Journal of Political Economy 116*(4), 709–745.

Baker, M., J. Gruber, and K. Milligan (2019). The long-run impacts of a universal child care program. *American Economic Journal: Economic Policy 11*(3), 1–26.

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

Barnett, W. S. (1995). Long-term effects of early childhood programs on cognitive and school outcomes. *The Future of Children*, 25–50.

Barr, A., J. Eggleston, and A. A. Smith (2022). Investing in infants: The lasting effects of cash transfers to new families. *The Quarterly Journal of Economics 137*(4), 2539–2583.

Barrera, O., K. Macours, P. Premand, and R. Vakis (2020). Texting parents about early child development: Behavioral changes and unintended social effects.

Barro, R. J. (2001). Human capital and growth. *American Economic Review 91*(2), 12–17.

Becker, G. S., S. D. Kominers, K. M. Murphy, and J. L. Spenkuch (2018). A theory of intergenerational mobility. *Journal of Political Economy 126*(S1), S7–S25.

Behrman, J. R., R. Gomez-Carrera, S. W. Parker, P. E. Todd, and W. Zhang (2024). Starting Strong: Medium-and Longer-run Benefits of Mexico's Universal Preschool Mandate.

Bennett, J. and C. Tayler (2006). *Starting strong II. Early childhood education and care. Annex E: Mexico*. OECD publishing.

Berkes, J. L., A. Bouguen, D. Filmer, and T. Fukao (2019). Improving preschool provision and encouraging demand: Heterogeneous impacts of a large-scale program. *World Bank Policy Research Working Paper* (9070).

Berlinski, S., S. Galiani, and P. Gertler (2009). The effect of pre-primary education on primary school performance. *Journal of Public Economics 93*(1-2), 219–234.

Berlinski, S., S. Galiani, and M. Manacorda (2008). Giving children a better start: Preschool attendance and school-age profiles. *Journal of Public Economics 92*(5-6), 1416–1440.

Berniell, I., L. Berniell, D. de la Mata, M. Edo, and M. Marchionni (2023). Motherhood and flexible jobs: Evidence from Latin American countries. *World Development 167*, 106225.

Blau, D. and J. Currie (2006). Pre-school, day care, and after-school care: who's minding the kids? *Handbook of the Economics of Education 2*, 1163–1278.

Blimpo, M. P., P. Carneiro, P. Jervis, and T. Pugatch (2022). Improving access and quality in early childhood development programs: Experimental evidence from the Gambia. *Economic Development and Cultural Change 70*(4), 1479–1529.

Borusyak, K., X. Jaravel, and J. Spiess (2024). Revisiting event-study designs: robust and efficient estimation. *Review of Economic Studies*, rdae007.

Bouguen, A., D. Filmer, K. Macours, and S. Naudeau (2018). Preschool and parental response in a second best world: Evidence from a school construction experiment. *Journal of Human Resources 53*(2), 474–512.

Brinkman, S. A., A. Hasan, H. Jung, A. Kinnell, and M. Pradhan (2017). The impact of expanding access to early childhood education services in rural Indonesia. *Journal of Labor Economics 35*(S1), S305–S335.

Callaway, B., A. Goodman-Bacon, and P. H. Sant'Anna (2024). Difference-in-differences with a continuous treatment. Technical report, National Bureau of Economic Research.

Callaway, B. and P. H. Sant'Anna (2021). Difference-in-differences with multiple time periods. *Journal of econometrics 225*(2), 200–230.

Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica 69*(5), 1127–1160.

Carneiro, P. M., E. Galasso, I. X. Lopez Garcia, P. Bedregal, and M. Cordero (2019). Parental beliefs, investments, and child development: Evidence from a large-scale experiment. *World Bank Policy Research Working Paper* (8743).

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

Casarico, A. and S. Lattanzio (2023). Behind the child penalty: understanding what contributes to the labour market costs of motherhood. *Journal of Population Economics 36*(3), 1489–1511.

Cascio, E. U. (2021). Early Childhood Education in the United States: What, When, Where, Who, How, and Why. Technical report, National Bureau of Economic Research.

Cascio, E. U. (2023). Does Universal Preschool Hit the Target?: Program Access and Preschool Impacts. *Journal of Human Resources 58*(1), 1–42.

Chetty, R., J. N. Friedman, N. Hilger, E. Saez, D. W. Schanzenbach, and D. Yagan (2011). How does your kindergarten classroom affect your earnings? Evidence from Project STAR. *The Quarterly Journal of Economics 126*(4), 1593–1660.

Coneval (2021). Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL). Índice de Rezago Social 2020 Principales resultados. https://www.coneval.org.mx/Medicion/IRS/Paginas/Indice_Rezago_Social_2020.aspx (accessed: 07.18.2024).

Corak, M. (2013). Income inequality, equality of opportunity, and intergenerational mobility. *Journal of Economic Perspectives 27*(3), 79–102.

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

Costa Dias, M., R. Joyce, and F. Parodi (2020). The gender pay gap in the UK: children and experience in work. *Oxford Review of Economic Policy 36*(4), 855–881.

Cunha, F. and J. Heckman (2007). The technology of skill formation. *American Economic Review 97*(2), 31–47.

Cunha, F. and J. J. Heckman (2009). The economics and psychology of inequality and human development. *Journal of the European Economic Association 7*(2-3), 320–364.

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

Currie, J. (2001). Early childhood education programs. *Journal of Economic Perspectives 15*(2), 213–238.

Currie, J. and D. Almond (2011). Human capital development before age five. In *Handbook of Labor Economics*, Volume 4, pp. 1315–1486. Elsevier.

Currie, J. and D. Thomas (1993). Does Head Start make a difference?

Danziger, S. and J. Waldfogel (2000). Investing in children: what do we know? What should we do? *What should we do*.

De Chaisemartin, C. and X. d'Haultfoeuille (2024). Difference-in-differences estimators of intertemporal treatment effects. *Review of Economics and Statistics*, 1–45.

de Chaisemartin, C., X. D'Haultfœuille, F. Pasquier, and G. Vazquez-Bare (2022). Difference-in-Differences for Continuous Treatments and Instruments with Stayers. *Available at SSRN 4011782*.

De Chaisemartin, C. and X. d'Haultfoeuille (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review 110*(9), 2964–2996.

De Chaisemartin, C. and X. d'Haultfoeuille (2023). Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey. *The Econometrics Journal 26*(3), C1–C30.

de Hoyos, R., R. Estrada, and M. J. Vargas (2021). What do test scores really capture? Evidence from a large-scale student assessment in Mexico. *World Development 146*, 105524.

De Hoyos, R., V. A. Garcia-Moreno, and H. A. Patrinos (2017). The impact of an accountability intervention with diagnostic feedback: Evidence from Mexico. *Economics of Education Review 58*, 123–140.

De Hoyos Navarro, R. E., R. Estrada, and M. J. Vargas (2018). Predicting individual wellbeing through test scores: evidence from a national assessment in Mexico. *World Bank Policy Research Working Paper* (8459).

De la Cruz Toledo, E. (2015). Universal preschool and mothers' employment. *A Working Paper of the Columbia Population Research Center*.

Deming, D. (2009). Early childhood intervention and life-cycle skill development: Evidence from Head Start. *American Economic Journal: Applied Economics 1*(3), 111–34.

Deming, D. J. (2022). Four facts about human capital. *Journal of Economic Perspectives 36*(3), 75–102.

Dietrichson, J., I. Lykke Kristiansen, and B. A. Viinholt (2020). Universal preschool programs and long-term child outcomes: A systematic review. *Journal of Economic Surveys 34*(5), 1007–1043.

Duncan, G., A. Kalil, M. Mogstad, and M. Rege (2023). Investing in early childhood development in preschool and at home. *Handbook of the Economics of Education 6*, 1–91.

Duncan, G. J. and K. Magnuson (2013). Investing in preschool programs. *Journal of Economic Perspectives 27*(2), 109–132.

Dustmann, C. and U. Schönberg (2012). Expansions in maternity leave coverage and children's long-term outcomes. *American Economic Journal: Applied Economics 4*(3), 190–224.

Elango, S., J. L. García, J. J. Heckman, and A. Hojman (2015). *Early Childhood Education.* IZA DP No. 9476.

Elango, S., J. L. García, J. J. Heckman, and A. Hojman (2016). *4. Early Childhood Education.* University of Chicago Press.

Felfe, C., N. Nollenberger, and N. Rodríguez-Planas (2015). Can't buy mommy's love? universal childcare and children's long-term cognitive development. *Journal of Population Economics 28*, 393–422.

Freeberg, N. E. and D. T. Payne (1967). Parental influence on cognitive development in early childhood: A review. *Child Development*, 65–87.

Gallego, F. A., E. Näslund-Hadley, and M. Alfonso (2021). Changing pedagogy to improve skills in preschools: experimental evidence from Peru. *The World Bank Economic Review 35*(1), 261–286.

Garces, E., D. Thomas, and J. Currie (2002). Longer-term effects of Head Start. *American Economic Review 92*(4), 999–1012.

García, J. L. and J. J. Heckman (2023). Parenting promotes social mobility within and across generations. *Annual Review of Economics 15*(1), 349–388.

Gardner, J. (2022). Two-stage differences in differences. *arXiv preprint arXiv:2207.05943*.

Gertler, P., J. Heckman, R. Pinto, A. Zanolini, C. Vermeersch, S. Walker, S. M. Chang, and S. Grantham-McGregor (2014). Labor market returns to an early childhood stimulation intervention in Jamaica. *Science 344*(6187), 998–1001.

Givord, P. and C. Marbot (2015). Does the cost of child care affect female labor market participation? An evaluation of a French reform of childcare subsidies. *Labour Economics 36*, 99–111.

Grantham-McGregor, S., A. Adya, O. Attanasio, B. Augsburg, J. Behrman, B. Caeyers, M. Day, P. Jervis, R. Kochar, P. Makkar, et al. (2020). Group sessions or home visits for early childhood development in India: a cluster RCT. *Pediatrics 146*(6).

Grantham-McGregor, S. M., C. A. Powell, S. P. Walker, and J. H. Himes (1991). Nutritional supplementation, psychosocial stimulation, and mental development of stunted children: the Jamaican Study. *The Lancet 338*(8758), 1–5.

Gray-Lobe, G., P. A. Pathak, and C. R. Walters (2023). The long-term effects of universal preschool in Boston. *The Quarterly Journal of Economics 138*(1), 363–411.

Gregorio, J. D. and J.-W. Lee (2002). Education and income inequality: new evidence from cross-country data. *Review of Income and Wealth 48*(3), 395–416.

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

Gupta, N. D. and M. Simonsen (2010). Non-cognitive child outcomes and universal high quality child care. *Journal of Public Economics 94*(1-2), 30–43.

Gómez-Carrera, R. (2022). Mandatory preschool and student test scores. Evidence from Mexico. *Technical Report. ITAM*.

Haeck, C., P. Lefebvre, and P. Merrigan (2015). Canadian evidence on ten years of universal preschool policies: The good and the bad. *Labour Economics 36*, 137–157.

Hahn, R. A. and W. S. Barnett (2023). Early childhood education: Health, equity, and economics. *Annual Review of Public Health 44*(1), 75–92.

Hanushek, E. A. and L. Woessmann (2012a). Do better schools lead to more growth? cognitive skills, economic outcomes, and causation. *Journal of Economic Growth 17*(4), 267–321.

Hanushek, E. A. and L. Woessmann (2012b). Schooling, educational achievement, and the Latin American growth puzzle. *Journal of Development Economics 99*(2), 497–512.

Hanushek, E. A. and L. Woessmann (2020). Education, knowledge capital, and economic growth. *The economics of education*, 171–182.

Hanushek, E. A. and L. Woessmann (2023). *The knowledge capital of nations: Education and the economics of growth.* MIT press.

Havnes, T. and M. Mogstad (2011). Money for nothing? universal child care and maternal employment. *Journal of Public Economics 95*(11-12), 1455–1465.

Havnes, T. and M. Mogstad (2015). Is universal child care leveling the playing field? *Journal of Public Economics 127*, 100–114.

Heckman, J., R. Pinto, and P. Savelyev (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review 103*(6), 2052–2086.

Heckman, J. J. (2013). *Giving kids a fair chance*. Mit Press.

Heckman, J. J., S. H. Moon, R. Pinto, P. Savelyev, and A. Yavitz (2010a). A new cost-benefit and rate of return analysis for the Perry Preschool Program: A summary. Technical report, National Bureau of Economic Research.

Heckman, J. J., S. H. Moon, R. Pinto, P. A. Savelyev, and A. Yavitz (2010b). The rate of return to the HighScope Perry Preschool Program. *Journal of Public Economics 94*(1-2), 114–128.

Heckman, J. J. and S. Mosso (2014). The economics of human development and social mobility. *Annu. Rev. Econ. 6*(1), 689–733.

Holla, A., M. Bendini, L. Dinarte, and I. Trako (2021). Is Investment in Preprimary Education Too Low? *Lessons from (Quasi) Experimental Evidence across Countries. Washington, DC: World Bank Policy Research Working Paper* (9723).

Kleven, H., C. Landais, J. Posch, A. Steinhauer, and J. Zweimüller (2024). Do family policies reduce gender inequality? Evidence from 60 years of policy experimentation. *American Economic Journal: Economic Policy 16*(2), 110–149.

Lipsey, M. W., D. C. Farran, and K. Durkin (2018). Effects of the Tennessee Prekindergarten Program on children's achievement and behavior through third grade. *Early Childhood Research Quarterly 45*, 155–176.

Ludwig, J. and D. L. Miller (2007). Does Head Start improve children's life chances? Evidence from a regression discontinuity design. *The Quarterly Journal of Economics 122*(1), 159–208.

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

Meghir, C., O. Attanasio, P. Jervis, M. Day, P. Makkar, J. Behrman, P. Gupta, R. Pal, A. Phimister, N. Vernekar, et al. (2023). Early stimulation and enhanced preschool: a randomized trial. *Pediatrics 151*(Supplement 2).

Meghir, C., M. Palme, and E. Simeonova (2013). Education, cognition and health: Evidence from a social experiment. Technical report, National Bureau of Economic Research.

Melhuish, E., K. Ereky-Stevens, K. Petrogiannis, A. Ariescu, E. Penderi, K. Rentzou, A. Tawell, P. Slot, M. Broekhuizen, and P. Leserman (2015). A review of research on the effects of early childhood education and care (ECEC) on child development. *Curriculum Quality Analysis and Impact Review of European Early Childhood Education and Care*.

Miller, D. L., N. Shenhav, and M. Grosz (2023). Selection into identification in fixed effects models, with application to Head Start. *Journal of Human Resources 58*(5), 1523–1566.

Negrete Rosales, T. P. (2011). La implementación de la obligatoriedad de la educación preescolar en México: una revisión desde la administración educativa. *Universidad Pedagócia Nacional*.

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

Nores, M., C. R. Belfield, W. S. Barnett, and L. Schweinhart (2005). Updating the economic impacts of the High/Scope Perry Preschool program. *Educational Evaluation and Policy Analysis 27*(3), 245–261.

Olivetti, C. and B. Petrongolo (2017). The economic consequences of family policies: lessons from a century of legislation in high-income countries. *Journal of Economic Perspectives 31*(1), 205–230.

Ortega Montes, R. M. (2008). La reforma en la educación preescolar durante el sexenio de Vicente Fox Quezada. *Universidad Nacional Autónoma de México, México*.

Pages, R., J. Protzko, and D. H. Bailey (2022). The breadth of impacts from the abecedarian project early intervention on cognitive skills. *Journal of Research on Educational Effectiveness 15*(2), 243–262.

Pérez, M., L. Pedroza, G. Ruiz, and A. López (2010). La educación preescolar en México. Condiciones para la enseñanza y el aprendizaje.

Puma, M., S. Bell, R. Cook, C. Heid, P. Broene, F. Jenkins, A. Mashburn, and J. Downer (2012). Third Grade Follow-Up to the Head Start Impact Study: Final Report. OPRE Report 2012-45. *Administration for Children & Families*.

Puma, M., S. Bell, R. Cook, C. Heid, G. Shapiro, P. Broene, F. Jenkins, P. Fletcher, L. Quinn, J. Friedman, et al. (2010). Head Start Impact Study. Final Report. *Administration for Children & Families*.

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy 98*(5, Part 2), S71–S102.

Roth, J., P. H. Sant'Anna, A. Bilinski, and J. Poe (2023). What's trending in difference-in-differences? a synthesis of the recent econometrics literature. *Journal of Econometrics 235*(2), 2218–2244.

Santibañez, L., J. F. Martinez, A. Datar, P. J. McEwan, C. M. Setodji, and R. Basurto-Davila (2007). Breaking Ground: Analysis of the Assessment System and Impact of Mexico's Teacher Incentive Program" Carrera Magisterial." Technical Report. *RAND Corporation*.

Secretaria de Educación Pública (2004). Programa de Educación Preescolar 2004. *México, DF ISBN*, 970–767.

Spier, E., K. Kamto, A. Molotsky, A. Rahman, N. Hossain, Z. Nahar, and H. Khondker (2020). Bangladesh Early Years Preschool Program Impact Evaluation.

Sun, L. and S. Abraham (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics 225*(2), 175–199.

Takahashi, M. (2024). Difference-in-discontinuities design. *Available at SSRN 4686891*.

Tramontin Shinoki, S., P. Picchetti, and C. Campos de Xavier Pinto (2024). Difference-in-discontinuities: Estimation, inference and validity tests. *Inference and Validity Tests (May 28, 2024)*.

Van Aar, J., P. Leijten, B. O. de Castro, and G. Overbeek (2017). Sustained, fade-out or sleeper effects? A systematic review and meta-analysis of parenting interventions for disruptive child behavior. *Clinical Psychology Review 51*, 153–163.

Walker, S. P., S. M. Chang, M. Vera-Hernández, and S. Grantham-McGregor (2011). Early childhood stimulation benefits adult competence and reduces violent behavior. *Pediatrics 127*(5), 849–857.

Weiland, C. and H. Yoshikawa (2013). Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills. *Child development 84*(6), 2112–2130.

Wolf, S. (2019). Year 3 follow-up of the 'Quality Preschool for Ghana'interventions on child development. *Developmental Psychology 55*(12), 2587.

Wolf, S., J. L. Aber, J. R. Behrman, and M. Peele (2019). Longitudinal causal impacts of preschool teacher training on Ghanaian children's school readiness: Evidence for persistence and fade-out. *Developmental Science 22*(5), e12878.

Wolf, S., J. L. Aber, J. R. Behrman, and E. Tsinigo (2019). Experimental impacts of the "Quality Preschool for Ghana" interventions on teacher professional well-being, classroom quality, and children's school readiness. *Journal of Research on Educational Effectiveness 12*(1), 10–37.

Wolf, S. and M. E. Peele (2019). Examining sustained impacts of two teacher professional development programs on professional well-being and classroom practices. *Teaching and Teacher Education 86*, 102873.

World Bank (2023). World Development Indicators. https://databank.worldbank.org/source/world-development-indicators.

XABER (2020). Armonización de la prueba censal ENLACE: Metodología de construcción de una base de datos con estructura de panel. *https://www.xaber.org.mx/datos.php*.

Yoshikawa, H., K. McCartney, R. Myers, K. L. Bub, J. Lugo-Gil, M. A. Ramos, F. Knaul, F. X. Gaytan, C. Buitrago, C. Rincón, et al. (2007). Early childhood education in Mexico: Expansion, quality improvement and curricular reform.

Zhang, Y., Y. Huang, X. Wang, and M. Zhong (2021). The Effects of Preschool Experiences on Mexican Children's Spanish Test Scores. In *2021 International Conference on Advanced Enterprise Information System (AEIS)*, pp. 25–32. IEEE.

10 Appendix

Appendix 1 The Publication of the Preschool Reform

Appendix 1.1 Excerpt from the Original Announcement in Spanish[27]

Artículo 3o. Todo individuo tiene derecho a recibir educación. El Estado -federación, estados, Distrito Federal y municipios-, impartirá educación preescolar, primaria y secundaria. La educación preescolar, primaria y la secundaria conforman la educación básica obligatoria.

Artículo 31. Son obligaciones de los mexicanos: I. Hacer que sus hijos o pupilos concurran a las escuelas públicas o privadas, para obtener la educación preescolar, primaria y secundaria, y reciban la militar, en los términos que establezca la ley.

Primero Transitorio.- El presente Decreto entrará en vigor al día siguiente de su publicación en el Diario Oficial de la Federación.

Cuarto Transitorio.- Con el objetivo de impulsar la equidad en la calidad de los servicios de educación preescolar en el país, la autoridad educativa deberá prever ... que la impartición de la educación preescolar es una profesión que necesita título para su ejercicio,...

Quinto Transitorio.- La educación preescolar será obligatoria para todos en los siguientes plazos: en el tercer año de preescolar a partir del ciclo 2004-2005; el segundo año de preescolar, a partir del ciclo 2005-2006; el primer año de preescolar, a partir del ciclo 2008-2009. En los plazos señalados, el Estado mexicano habrá de universalizar en todo el país, con calidad, la oferta de este servicio educativo.

Sexto Transitorio.- Los presupuestos federal, estatales, del Distrito Federal y municipales incluirán los recursos necesarios para: la construcción, ampliación y equipamiento de la infraestructura suficiente para la cobertura progresiva de los servicios de educación preescolar; con sus correspondientes programas de formación profesional del personal docente así como de dotación de materiales de estudio gratuito para maestros y alumnos...

Séptimo Transitorio.- Los gobiernos estatales y del Distrito Federal celebrarán con el gobierno federal convenios de colaboración que les permitan cumplir con la obligatoriedad de la educación preescolar en los términos establecidos en los artículos anteriores.

[27] See the translation of the excerpt in Appendix 1.2.

Appendix 1.2 Translation of the Excerpt

Article 3. Every individual has the right to receive education. The State - federation, states, Federal District, and municipalities - shall provide preschool, primary, and secondary education. Preschool, primary, and secondary education constitute basic compulsory education.

Article 31. It is the obligation of Mexicans: I. To ensure that their children or wards attend public or private schools to receive preschool, primary, and secondary education, and receive military education, under the terms established by law.

First Transitory.- This Decree shall enter into force the day after its publication in the Official Journal of the Federation.

Fourth Transitory.- In order to promote equity in the quality of preschool education services in the country, the educational authority must ensure ... that the provision of preschool education is a profession that requires a university degree for its practice,...

Fifth Transitory.- Preschool education will be compulsory for everyone according to the following timelines: in the third year of preschool starting from the 2004-2005 school cycle; in the second year of preschool, starting from the 2005-2006 school cycle; in the first year of preschool, starting from the 2008-2009 school cycle. Within the specified timelines, the Mexican State must universally offer this educational service nationwide, with quality.

Sixth Transitory.- The federal, state, Federal District, and municipal budgets will include the necessary resources for: the construction, expansion, and equipping of sufficient infrastructure for the progressive coverage of preschool education services; with corresponding programs for the professional training of teaching staff as well as the provision of free study materials for teachers and students...

Seventh Transitory.- The state and Federal District governments shall enter into collaboration agreements with the federal government to comply with the compulsory preschool education as established in the previous articles.

Appendix 2 Mexico's Administrative Division

Mexico is a federal republic organized into several levels of government: federal, state, and municipal, with additional subdivisions at the local level. The federal government, the highest level of government, is responsible for national policies, defense, foreign affairs, and overall economic policy.

Mexico is divided into 32 federal entities, including 31 states and Mexico City (the capital, previously named Federal District, which has a special status akin to that of a state). Each state has its own constitution, governor, and legislature, which are responsible for local policies, public education, health services, and infrastructure, among other things.

Each state is subdivided into municipalities (municipios). There are around 2,500 municipalities across the country. The municipal governments handle local services such as water supply, public safety, and local roads.

Within municipalities, there are smaller administrative units called localities (localidades). These can range from large cities to small rural communities. Localities are the most basic units of settlement and are recognized for statistical and administrative purposes. They do not have their own government but are subject to the jurisdiction of the municipality they belong to.

Mexico has over 200,000 localities. Localities in Mexico can be broadly categorized into urban and rural areas. Urban localities are typically defined as those with populations of 2,500 or more inhabitants. Rural localities are those with populations of less than 2,500 inhabitants. Around 2% of localities are urban and 98% rural. A sub-category of urban localities is semi-urban which encompasses localities with populations between 2,500 and 15,000 inhabitants. Almost 90% of the broader urban category are semi-urban.

Indigenous localities are those where a large portion of the population identifies as indigenous or speaks an indigenous language. There are more than 10,000 predominantly indigenous localities in Mexico.

Appendix 3 Mexico's Educational System

The Mexican educational system is organized into several distinct levels. It begins with initial education for children under 3 years old, followed by basic education which encompasses preschool (ages 3-5), primary (ages 6-11), and secondary (ages 12-15) levels. The system then progresses to upper secondary education, offering both high school and technical tracks, and culminates in higher education.

Basic education, which is mandatory, is delivered through three main types of schools. General schools, both public and private, are the most prevalent, serving the majority of students with a standard curriculum that includes subjects like Spanish, math, natural sciences, history, and physical education. To address the unique needs of indigenous communities, specialized schools offer intercultural and bilingual education, aiming to preserve and promote indigenous languages and cultures while also delivering the national curriculum. These schools are primarily located in rural and indigenous areas. For remote and underserved regions, community schools managed by the National Council for Educational Development (CONAFE) ensure access to basic education. These schools are staffed by young, trained community instructors who provide essential educational services in areas lacking regular schools.

In 2000, 88% of preschoolers were enrolled in General schools, 9% in indigenous schools, and 3% in community schools. As for primary school in 2000, 94% of students were enrolled in General schools, 5% in Indigenous schools, and 1% in community schools. Only about 10% of preschools are private (Bennett and Tayler, 2006).

Preschools operate along age cohort lines and are open for 3 or 4 hours daily, five days a week. Some preschools offer both morning and afternoon sessions. A special subset of preschools, known as "mixed preschools" (jardines mixtos), combine a regular preschool session with care during a full-day program. This preschool model is not very widespread (Bennett and Tayler, 2006). Enrollment ratios vary considerably among states, ranging from 65% to 113% for children aged 5 and from 2% to 53% for those aged 3. Generally, the poorest states tend to have lower enrollment ratios, although there are exceptions to this trend (Bennett and Tayler, 2006).

The Mexican education system is managed through a complex network of educational entities, with responsibilities shared across different levels of government and various institutions. At the federal level, the Ministry of Education (SEP) plays a central role in setting national education policies and standards. However, the system is largely decentralized, with each of Mexico's 32 federal entities operating their own Education Ministries or Departments to manage education within their jurisdictions.

While the federal government establishes overarching norms and regulations, state governments are primarily responsible for operating basic education services, including preprimary, primary, and secondary levels. Municipalities also contribute, particularly in maintaining school infrastructure.

The school cycle in Mexico typically runs from September to July. The typical cutoff date for enrollment in a certain grade was September 1. Therefore, a child must reach the appropriate age for that grade before September 1 to enroll in that grade at the start of the school cycle.

Appendix 4 ENLACE

ENLACE (Evaluación Nacional de Logro Académico en Centros Escolares) was a relevant standardized testing program in Mexico, conducted annually from 2006 to 2013 for grades 3-9 and 12. Administered by the Mexican Ministry of Education (SEP), ENLACE aimed to assess student achievement in math, Spanish, and rotating subjects, utilizing item response theory for consistent scoring within grades over time. Each subject targeted a national mean score of 500 points with a standard deviation of 100 points (De Hoyos et al., 2017).

The dataset provided detailed insights into school performance, student demographics, and teacher qualifications. It facilitated longitudinal studies by assigning unique identifiers to students, enabling tracking of academic progress over time. ENLACE results influenced educational policy, serving as a benchmark for school comparisons and informing resource allocation decisions.

Despite its intended low-stakes nature, ENLACE garnered widespread attention in Mexico, appearing frequently in major newspapers and NGO rankings. It became a medium-stakes assessment for school directors and communities, shaping perceptions of educational quality (de Hoyos et al., 2021).

Criticism of ENLACE included concerns about its use in high-stakes contexts, such as teacher bonuses, which led to strategic behaviors and worries about grade inflation. In response to these issues and the establishment of the INEE (National Institute for Educational Evaluation), ENLACE was phased out after 2013-2014 and replaced by the Plan Nacional para la Evaluación de los Aprendizajes in 2015, marking a shift in Mexico's educational assessment approach (de Hoyos et al., 2021).

Despite their limitations, de Hoyos et al. (2021) find that ENLACE test scores can effectively measure the cognitive skills they are designed to measure and can be utilized to monitor student learning.

Appendix 5 Sample

To construct the sample, the study utilized three primary data sources: locality data from INEGI, Formato 911 preschool data, and ENLACE primary school test scores.

Locality Data Set. First, the number of localities and their ID codes were obtained from INEGI using catalogs such as the 2000 Catálogo de Integración General de Localidades (CIGEL) and the Catálogo Único de Claves de Áreas Geoestadísticas Estatales, Municipales y Localidades. The IDs were verified to ensure they matched between the Ministry of Education and INEGI data, noting that only around 75% of IDs coincided due to differences in coding until 2003. For this reason, GPS coordinates for preschools, available for about 90% of the schools, were used to assign locality IDs accurately. Additionally, population and cartography data from INEGI were used to create a setting variable (urban, semi-urban, rural) at the locality level, supplemented by Coneval's 2000 poverty data[28] and INEGI's 2000 Population Census data.

Formato 911 Preschool Data Set. For the Formato 911 dataset, relevant variables were generated for each school year, including total preschools, enrollment of 4-year-olds, certified teachers, and classroom sizes. The data were differentiated by school types (General, Indigenous, Community, and Private), and cohort variables were defined based on 4-year-olds enrolled in a specific school year (sep-aug). GPS coordinates were used to assign locality IDs to 90% of preschools. Outliers for variables such as student enrollment, teachers, and classrooms were winsorized at the school level, and the data were aggregated at the locality-cohort level.

[28] I use the Social Lag Index as a proxy for poverty. This index, estimated by the National Council for the Evaluation of Social Development Policy (CONEVAL), is a weighted measure that summarizes indicators related to education, access to health services, housing quality and space, basic household services, and household assets. Its purpose is to rank territorial units (states, municipalities, and localities) according to the index value derived from the observed levels of these social indicators. It also classifies these units into five degrees of social lag: very low, low, medium, high, and very high. It is important to note that this index is not a perfect measure of poverty, as it does not include information on income, access to social security, or access to food (Coneval, 2021).

ENLACE Primary School Test Scores Data Set. For ENLACE data, relevant variables such as math and Spanish test scores, along with students' sex, were utilized for each school year and grade. Cohort variables were created based on students' dates of birth (sep-aug). Selection bias is addressed by grouping students into age-appropriate cohorts to analyze test scores within these cohorts.[29] GPS coordinates were used to match 99% of primary schools to locality IDs. Test score outliers were winsorized at the student level, and students four or more years behind or one or more years ahead in grade levels, as well as those from schools that inconsistently operated, were excluded to avoid biased estimates. Data were aggregated at the locality-cohort level.

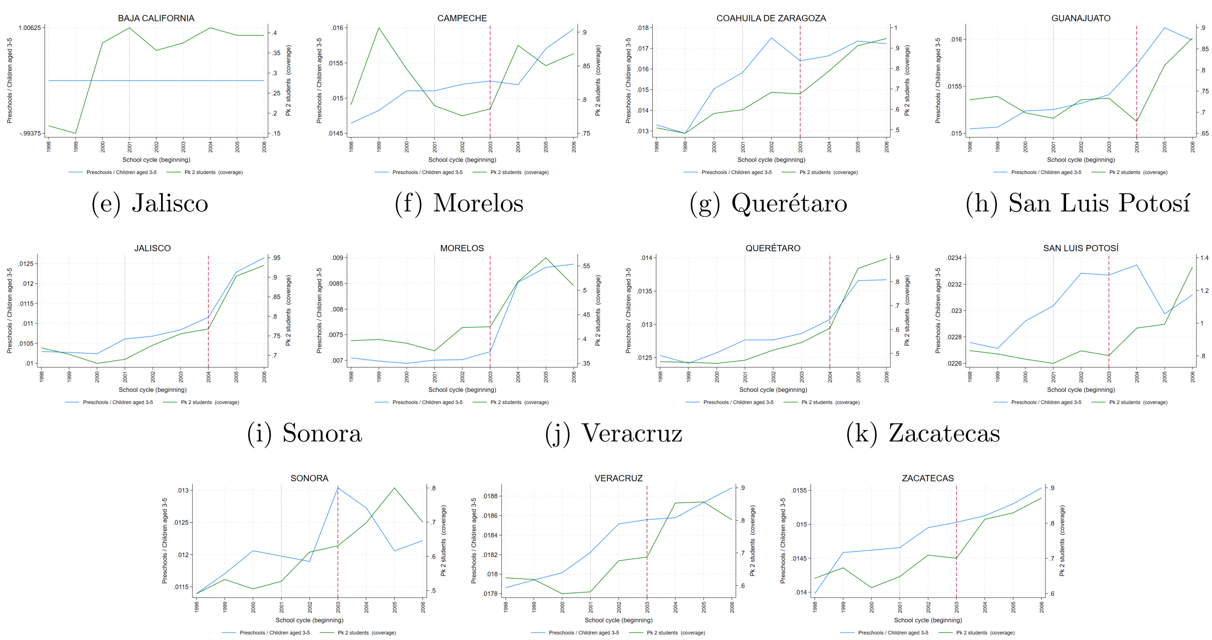
Final Sample. The final sample was constructed by merging 86% of the observations in the ENLACE dataset to the Formato 911 dataset, using locality and cohort variables. Locality data, including setting, poverty, and other demographic information, were also integrated. The focus was on 11 states that had a September 1st cutoff for entering both primary school and preprimary school (consistent with Behrman et al. (2024)).

Appendix 6 Dynamic Adoption

Figure A4 presents the time series data for pre-k 2 enrollment and the total number of preschools at the state level. This figure helps identify the years in which states experienced the largest increases in these variables, indicating the adoption year of the reform. It also highlights that states adopted the reform in different years, with some anticipation evident since 2002. Table A6 summarizes these insights by displaying the year of maximum growth for pre-k 2 enrollment and the total number of preschools for each state.

[29] For instance, I consider Grade 3 students' test scores when they reach Grade 3 if they are not in Grade 3 when they are supposed to be because they are older.

Figure A4: Dynamic Adoption for the 11 States in the Sample



(a) Baja California (b) Campeche (c) Coahuila (d) Guanajuato
(e) Jalisco (f) Morelos (g) Querétaro (h) San Luis Potosí
(i) Sonora (j) Veracruz (k) Zacatecas

Source: Figure generated by the author using data from Formato 911 and INEGI.
Notes: The insights from this Figure are summarized in Table A6. The sample includes the 11 states with a consistent September 1st cutoff for both primary and pre-primary school. These 11 states include Baja California, Campeche, Coahuila, Guanajuato, Jalisco, Morelos, Queretaro, San Luis Potosi, Sonora, Veracruz, and Zacatecas.

Table A6: Year of Maximum Growth by State

State	Preschools (1)	Pre-k 2 students (2)
BAJA CALIFORNIA	-	2004
CAMPECHE	2005	2004
COAHUILA DE ZARAGOZA	2002	2004
GUANAJUATO	2005	2005
JALISCO	2005	2005
MORELOS	2004	2004
QUERÉTARO	2005	2005
SAN LUIS POTOSÍ	2002	2004
SONORA	2003	2002
VERACRUZ	2002	2004
ZACATECAS	2002	2004

Source: Table generated by the author using data from Formato 911 and INEGI.
Notes: This table summarizes the insights obtained from Figure A4. The sample includes the 11 states with a consistent September 1st cutoff for both primary and pre-primary school. These 11 states include Baja California, Campeche, Coahuila, Guanajuato, Jalisco, Morelos, Queretaro, San Luis Potosi, Sonora, Veracruz, and Zacatecas.

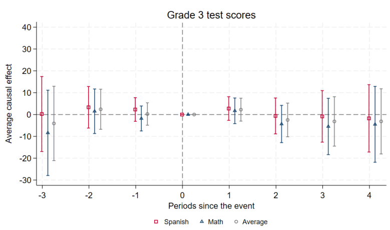
Appendix 7 Complementing Staggered DiD with Continuous Treatment

Appendix 7.1 and Appendix 7.2 complement Figure 3 and the results discussed in Section 7.2. Appendix 7.1 delves deeper into the test scores, the primary outcome variable, while Appendix 7.2 elaborates on pre-k 2 enrollment, the First Stage, and the quality mechanisms.

Appendix 7.1 Distributional Effects on Test Scores

Figure A5: Distributional Effects on Grade 3 Test Scores

(a) Math and Spanish



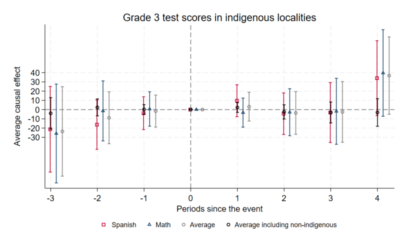
(b) Gender (aggregated)

(c) Gender (disaggregated)

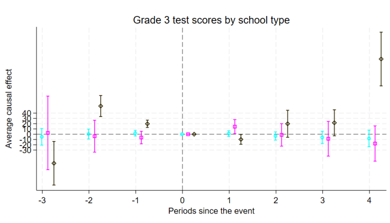
(d) Poverty (aggregated)

(e) Poverty (disaggregated)

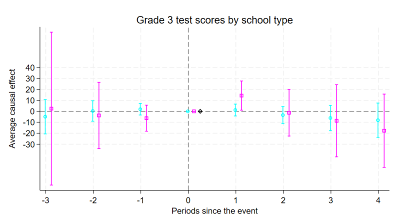
(f) Indigenous localities



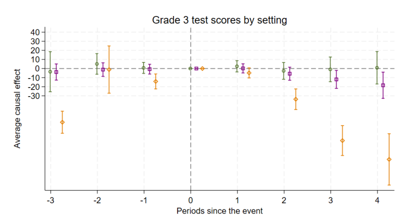
(g) Type of School (with Private)



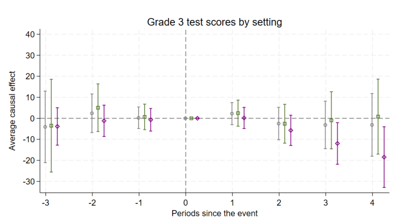
(h) Type (without Private)



(i) Setting (with Urban)



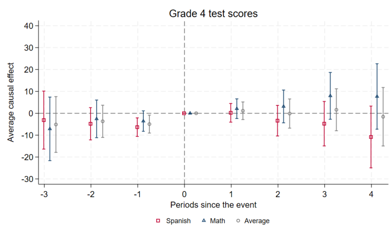
(j) Setting (without Urban)



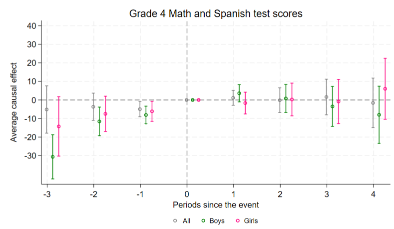
Source: Figure generated by the author using data from Formato 911, ENLACE, INEGI, and CONEVAL. Notes: All graphs include 95% confidence intervals. The results are computed using the Stata package *did_multiplegt_dyn* developed by De Chaisemartin and d'Haultfoeuille (2023) and De Chaisemartin and d'Haultfoeuille (2024)

Figure A6: Distributional Effects on Grade 4 Test Scores

(a) Math and Spanish



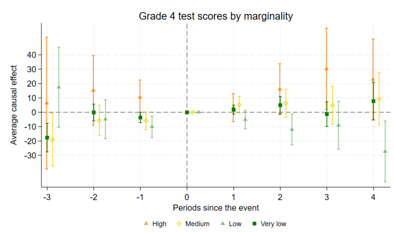
(b) Gender (aggregated)



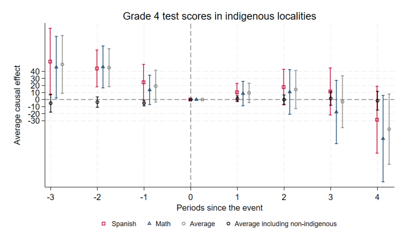
(c) Gender (disaggregated)

(d) Poverty (aggregated)

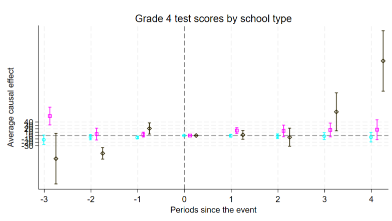
(e) Poverty (disaggregated)



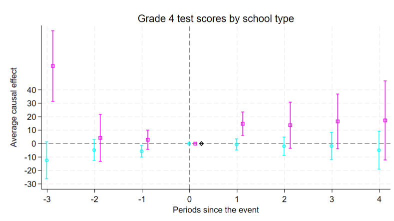
(f) Indigenous localities



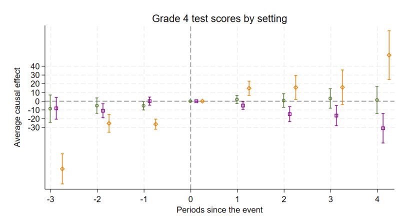
(g) Type of School (with Private)



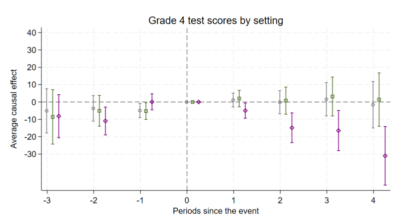
(h) Type (without Private)



(i) Setting (with Urban)



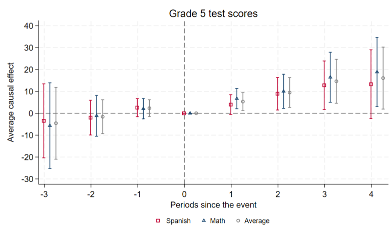
(j) Setting (without Urban)



Source: Figure generated by the author using data from Formato 911, ENLACE, INEGI, and CONEVAL. Notes: All graphs include 95% confidence intervals. The results are computed using the Stata package *did_multiplegt_dyn* developed by De Chaisemartin and d'Haultfoeuille (2023) and De Chaisemartin and d'Haultfoeuille (2024)

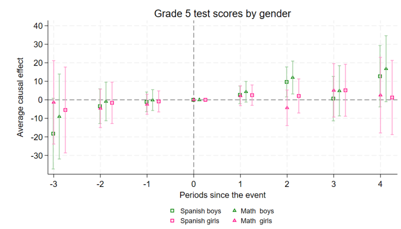
Figure A7: Distributional Effects on Grade 5 Test Scores

(a) Math and Spanish



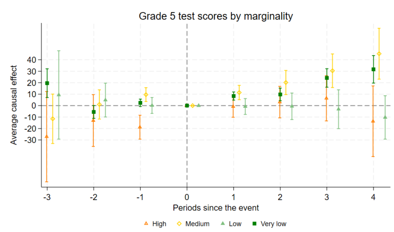
(b) Gender (aggregated)

(c) Gender (disaggregated)

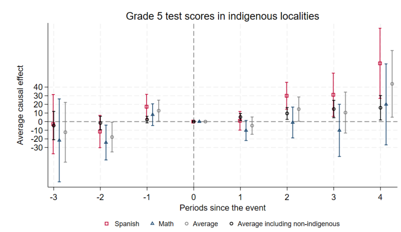


(d) Poverty (aggregated)

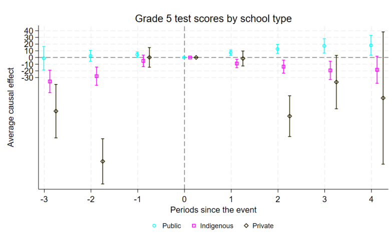
(e) Poverty (disaggregated)



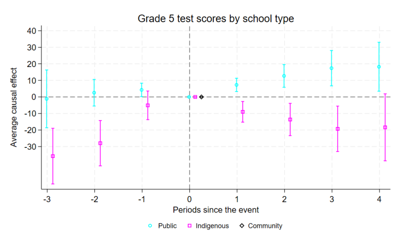
(f) Indigenous localities



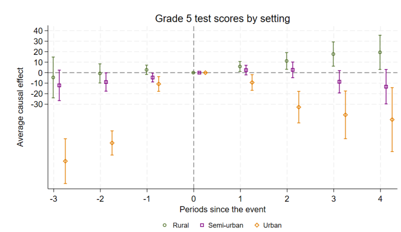
(g) Type of School (with Private)



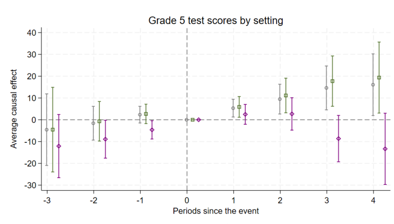
(h) Type (without Private)



(i) Setting (with Urban)



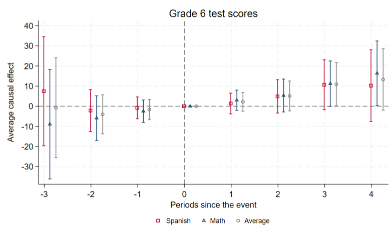
(j) Setting (without Urban)



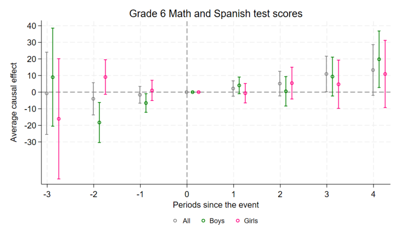
Source: Figure generated by the author using data from Formato 911, ENLACE, INEGI, and CONEVAL. Notes: All graphs include 95% confidence intervals. The results are computed using the Stata package *did_multiplegt_dyn* developed by De Chaisemartin and d'Haultfoeuille (2023) and De Chaisemartin and d'Haultfoeuille (2024)

Figure A8: Distributional Effects on Grade 6 Test Scores

(a) Math and Spanish

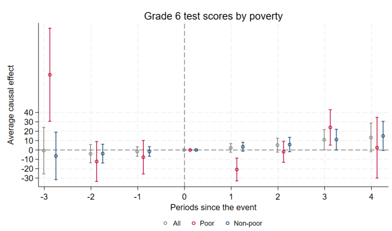


(b) Gender (aggregated)

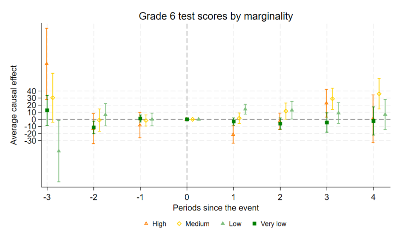


(c) Gender (disaggregated)

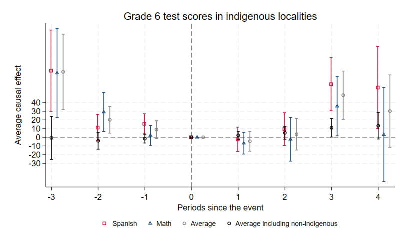
(d) Poverty (aggregated)



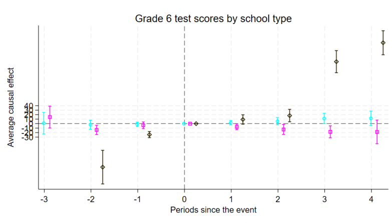
(e) Poverty (disaggregated)



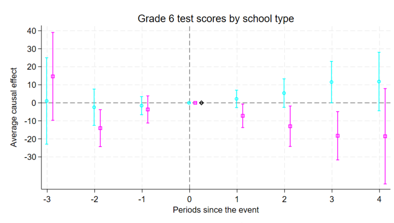
(f) Indigenous localities



(g) Type of School (with Private)

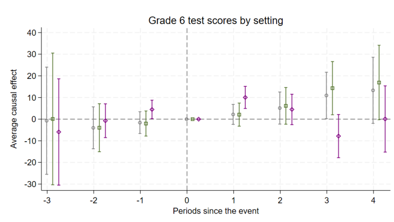


(h) Type (without Private)



(i) Setting (with Urban)

(j) Setting (without Urban)

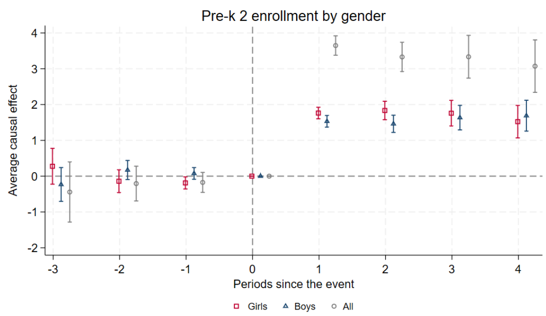


Source: Figure generated by the author using data from Formato 911, ENLACE, INEGI, and CONEVAL. Notes: All graphs include 95% confidence intervals. The results are computed using the Stata package *did_multiplegt_dyn* developed by De Chaisemartin and d'Haultfoeuille (2023) and De Chaisemartin and d'Haultfoeuille (2024)

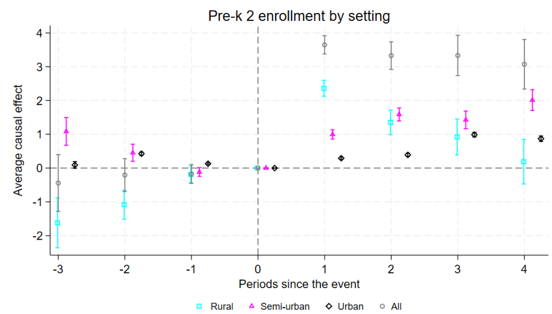
Appendix 7.2 Distributional Effects on Pre-k 2 Enrollment and Mechanisms

Figure A9: Distributional Effects on Pre-k 2 Enrollment

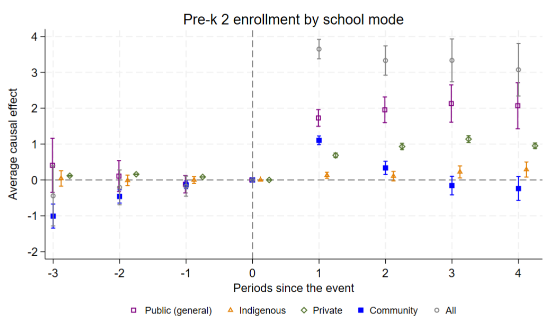
(a) Gender



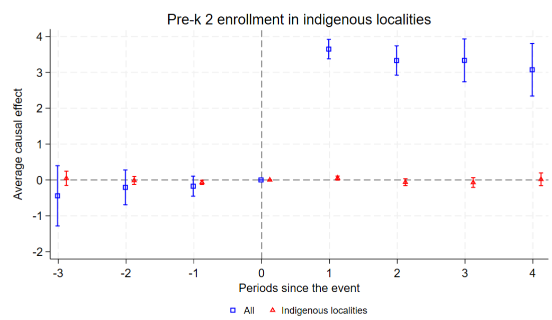
(b) Setting



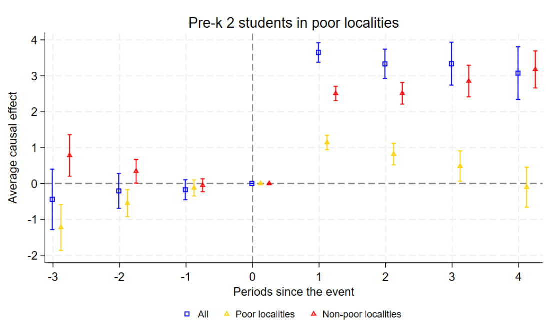
(c) School Type



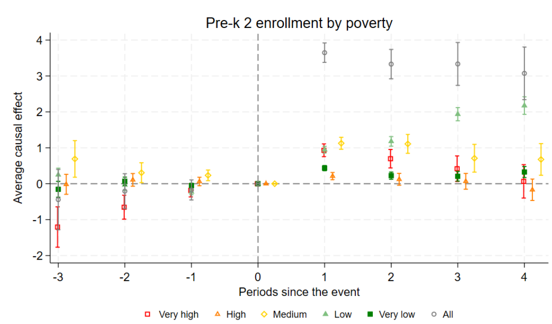
(d) Indigenous Localities



(e) Poverty (aggregated)



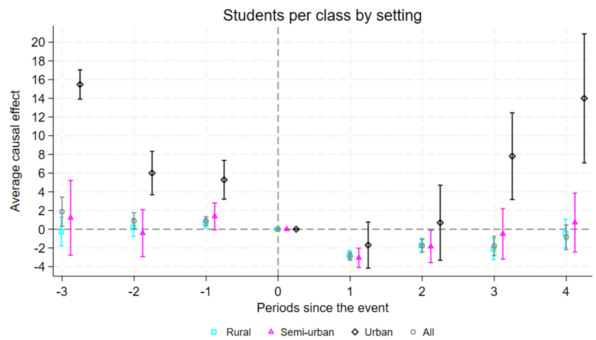
(f) Poverty (disaggregated)



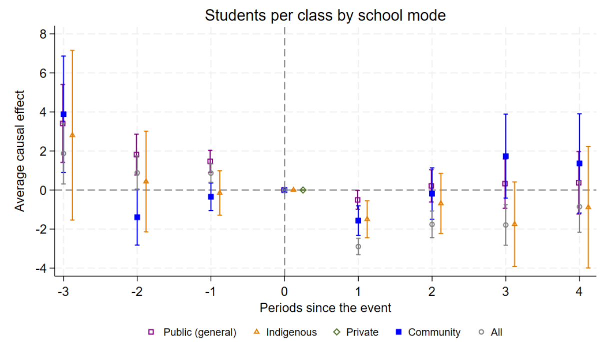
Source: Figure generated by the author using data from Formato 911, ENLACE, INEGI, and CONEVAL. Notes: All graphs include 95% confidence intervals. The results are computed using the Stata package *did_multiplegt_dyn* developed by De Chaisemartin and d'Haultfoeuille (2023) and De Chaisemartin and d'Haultfoeuille (2024)

Figure A10: Distributional Effects on Class Size

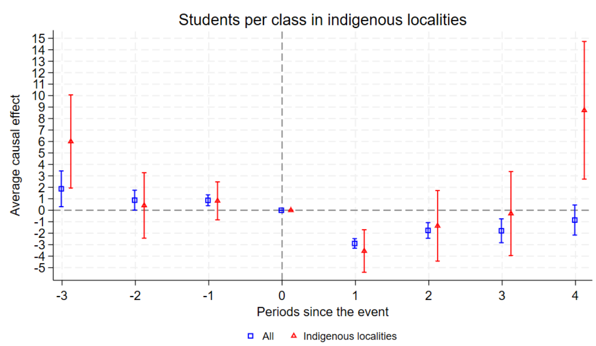
(a) Setting



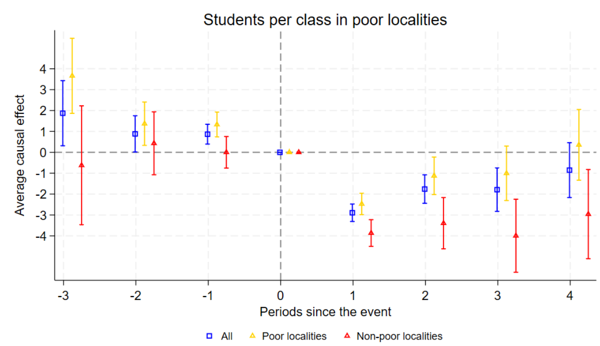
(b) School Type



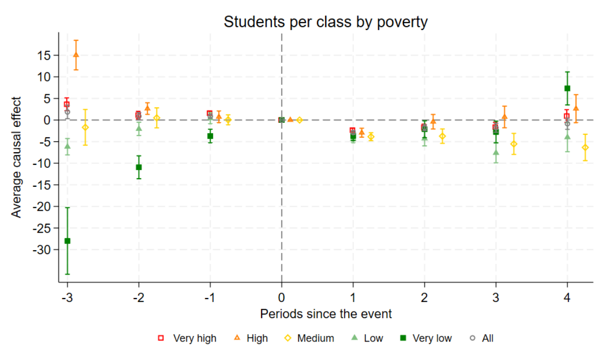
(c) Indigenous Localities



(d) Poverty (aggregated)



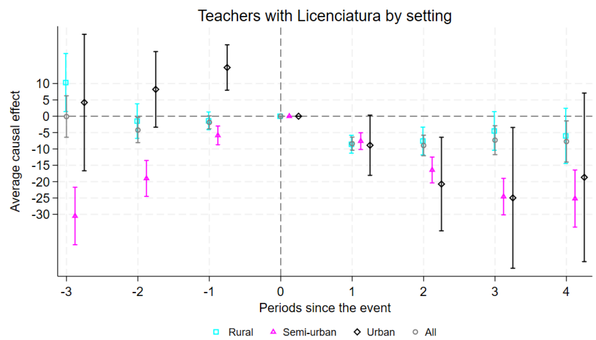
(e) Poverty (disaggregated)



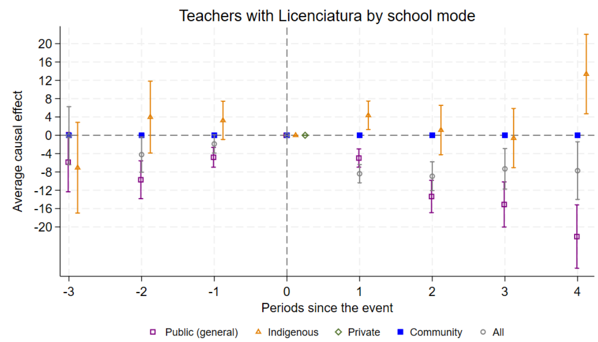
Source: Figure generated by the author using data from Formato 911, ENLACE, INEGI, and CONEVAL. Notes: All graphs include 95% confidence intervals. The results are computed using the Stata package *did_multiplegt_dyn* developed by De Chaisemartin and d'Haultfoeuille (2023) and De Chaisemartin and d'Haultfoeuille (2024)

Figure A11: Distributional Effects on Share of Teachers with Undergraduate Degree

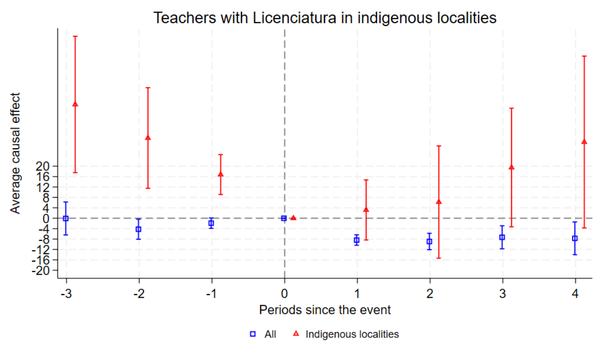
(a) Setting



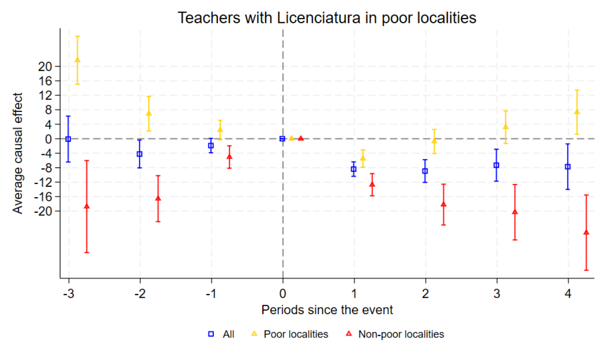
(b) School Type



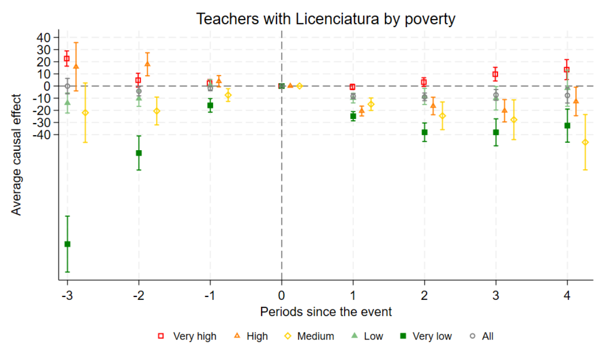
(c) Indigenous Localities



(d) Poverty (aggregated)



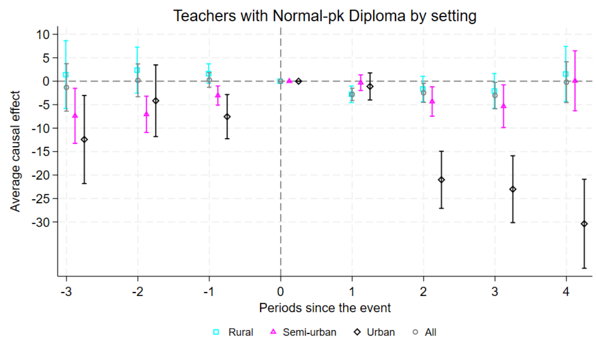
(e) Poverty (disaggregated)



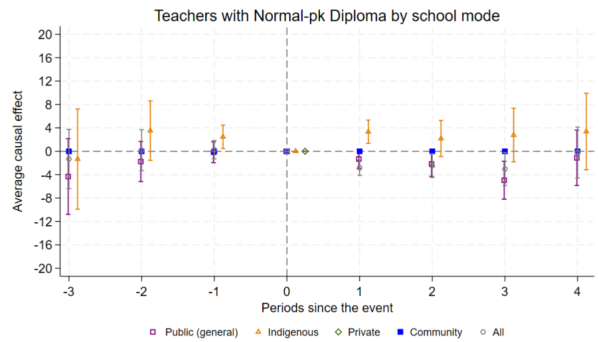
Source: Figure generated by the author using data from Formato 911, ENLACE, INEGI, and CONEVAL. Notes: All graphs include 95% confidence intervals. The results are computed using the Stata package *did_multiplegt_dyn* developed by De Chaisemartin and d'Haultfoeuille (2023) and De Chaisemartin and d'Haultfoeuille (2024)

Figure A12: Distributional Effects on Share of Teachers with Normal Preschool Degree

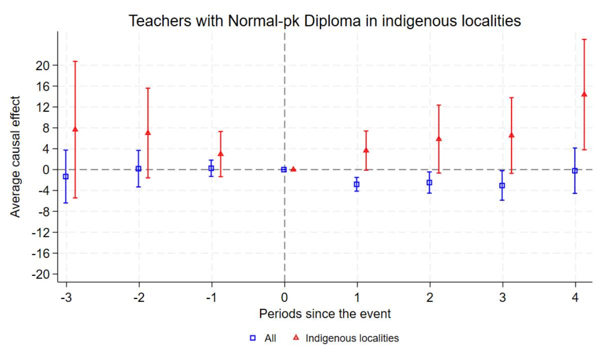
(a) Setting



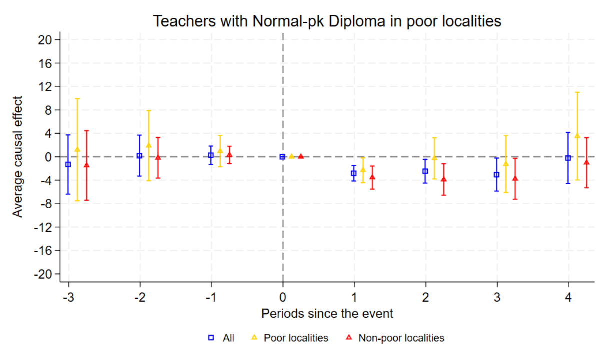
(b) School Type



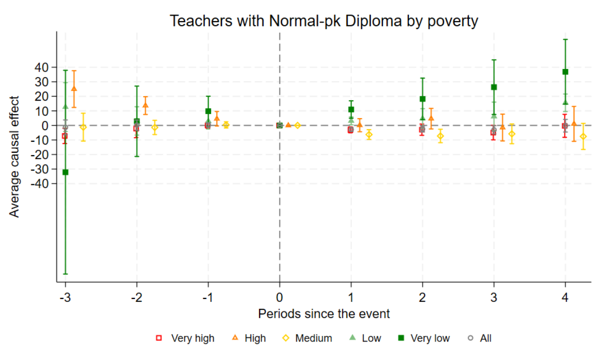
(c) Indigenous Localities



(d) Poverty (aggregated)



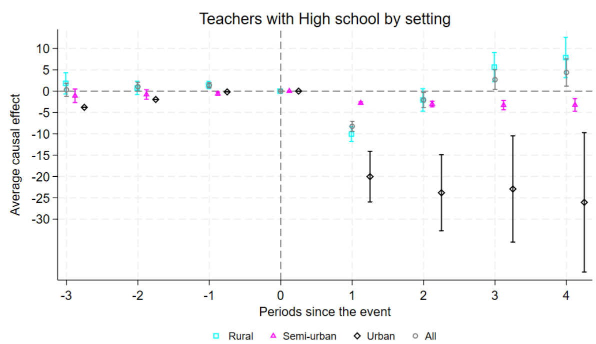
(e) Poverty (disaggregated)



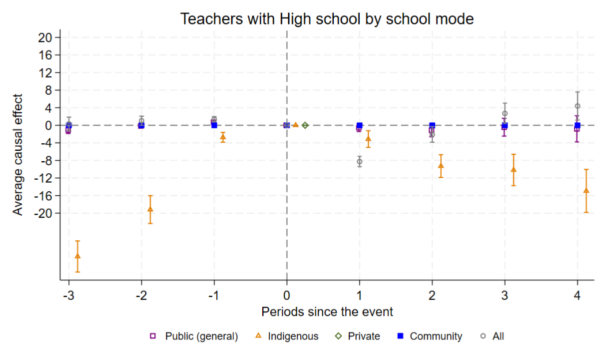
Source: Figure generated by the author using data from Formato 911, ENLACE, INEGI, and CONEVAL. Notes: All graphs include 95% confidence intervals. The results are computed using the Stata package *did_multiplegt_dyn* developed by De Chaisemartin and d'Haultfoeuille (2023) and De Chaisemartin and d'Haultfoeuille (2024)

Figure A13: Distributional Effects on Share of Teachers with High School Degree

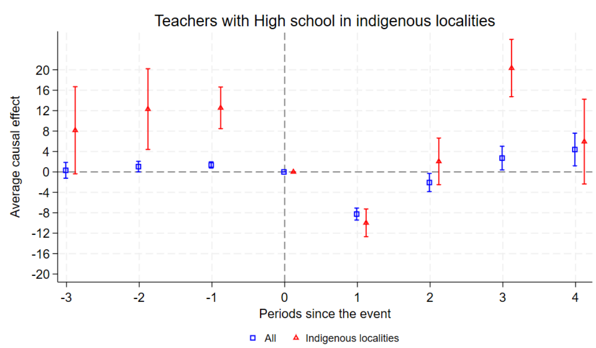
(a) Setting



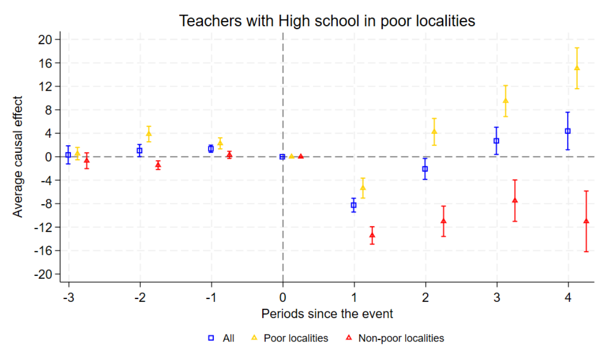
(b) School Type



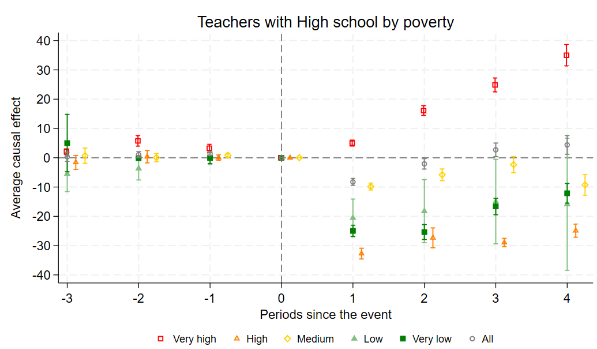
(c) Indigenous Localities



(d) Poverty (aggregated)



(e) Poverty (disaggregated)



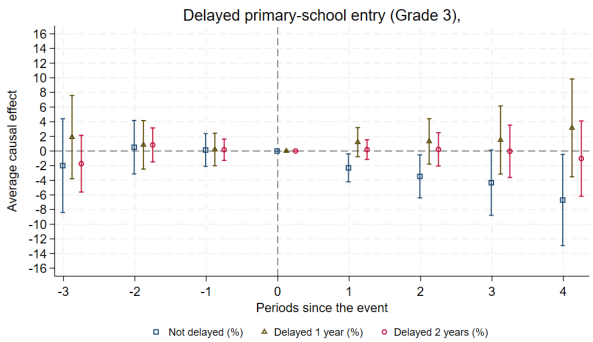
Source: Figure generated by the author using data from Formato 911, ENLACE, INEGI, and CONEVAL. Notes: All graphs include 95% confidence intervals. The results are computed using the Stata package *did_multiplegt_dyn* developed by De Chaisemartin and d'Haultfoeuille (2023) and De Chaisemartin and d'Haultfoeuille (2024)

Appendix 7.3 Delayed Entry to Primary School

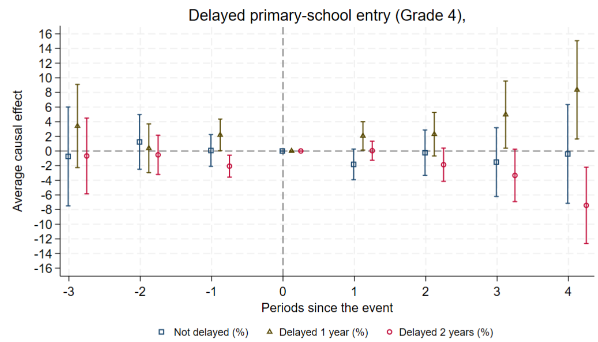
Figure A14 examines whether students in primary school experienced delayed entry. This analysis considers the birthdates of enrolled students and assesses the proportion of students who are on time, one year behind, and two years behind.

Figure A14: Delayed Entry to Primary School

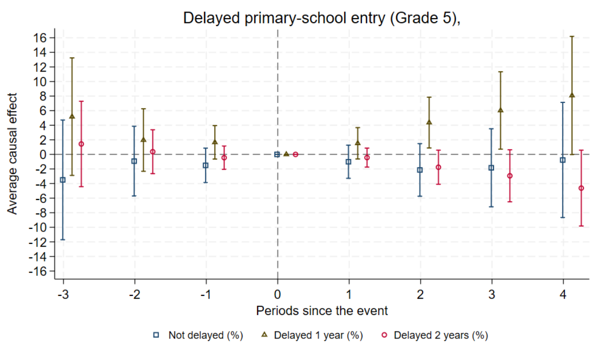
(a) Grade 3



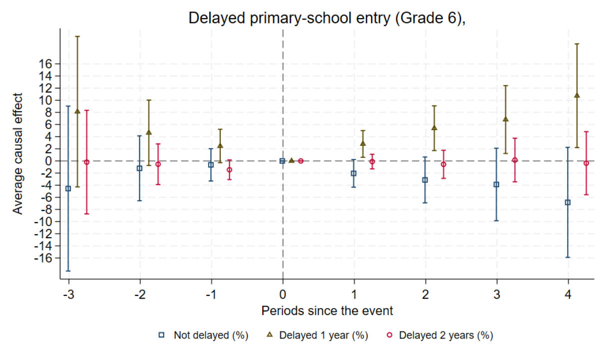
(b) Grade 4



(c) Grade 5



(d) Grade 6



Source: Figure generated by the author using data from Formato 911, ENLACE, INEGI, and CONEVAL. Notes: All graphs include 95% confidence intervals. The results are computed using the Stata package *did_multiplegt_dyn* developed by De Chaisemartin and d'Haultfoeuille (2023) and De Chaisemartin and d'Haultfoeuille (2024)

Appendix 8 Complementing TWFE Results Considering that All States Adopt in 2002

This section supplements the findings presented in Section 6. Table A7 offers summary statistics for the localities included in the subsample used for the analysis in Section 6. Figure A15 presents evidence of reform adoption beginning in 2002. Table A8 elaborates on the results shown in Panel A of Table 5 by providing detailed math and Spanish test scores.

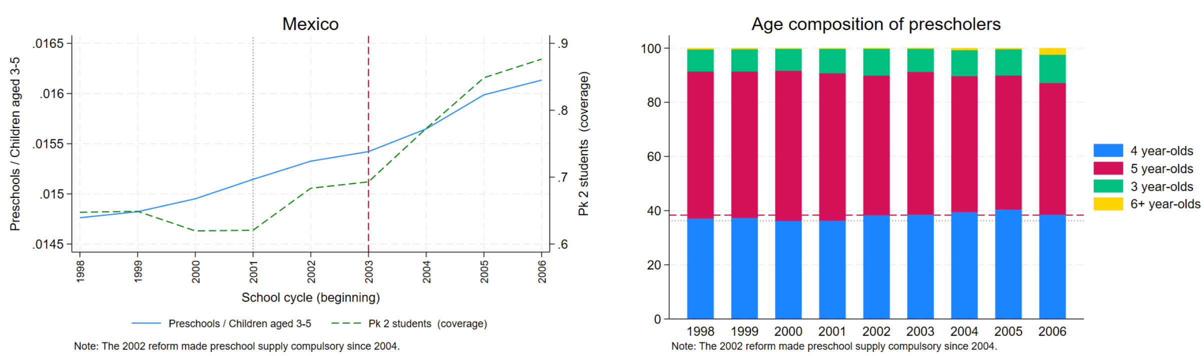
Table A7: Characteristics of the Localities that had no Preschools before 2002

Variable	Obs.	Min.	Mean	Max.
% Rural locs.	634	0	99.37	100
% Semi-urban locs.	634	0	0.63	100
% Urban locs.	634	0	0.00	0
% Indigenous locs.	634	0	2.84	100
% Poor locs. (High or Very High)	634	0	11.81	100
% Very High poverty locs.	634	0	0.17	100
% High poverty locs.	634	0	11.65	100
% Medium poverty locs.	634	0	39.27	100
% Low poverty locs.	634	0	30.12	100
% Very Low poverty locs.	634	0	18.80	100

Source: Table generated by the author using data from Formato 911, ENLACE, INEGI, and CONEVAL.
Notes: This subsample represents almost 6% of the sample presented in Table 2. Data at the locality level. Localities are included in the subsample if they had no preschool from 1998 to 2001; 1998 is the first year for which the Formato 911 data is available. Localities are included in the sample if they have a primary school in the ENLACE data. The sample includes the 11 states with a consistent September 1st cutoff for both primary and pre-primary school. These 11 states include Baja California, Campeche, Coahuila, Guanajuato, Jalisco, Morelos, Queretaro, San Luis Potosi, Sonora, Veracruz, and Zacatecas.

Figure A15: Reform Adoption in 2002

(a) Adoption at the National Level (b) Age Composition of Preschoolers (%)



Source: Figure generated by the author using data from Formato 911 and INEGI.
Notes: These plots show that pre-k 2 enrollment increased since 2002. The sample includes the 11 states with a consistent September 1st cutoff for both primary and pre-primary school. These 11 states include Baja California, Campeche, Coahuila, Guanajuato, Jalisco, Morelos, Queretaro, San Luis Potosi, Sonora, Veracruz, and Zacatecas.

Table A8: Effect of the Reform on Math and Spanish Test Scores for Localities that had no preschool before 2002

	Math Test Scores				Spanish Test Scores			
VARIABLES	Grade 3	Grade 4	Grade 5	Grade 6	Grade 3	Grade 4	Grade 5	Grade 6
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\hat{\beta}$	21.8356**	21.3387**	10.6262	19.2903**	7.2287	17.8159**	1.6614	10.8293
	(8.9382)	(9.2811)	(8.6630)	(9.1268)	(9.1407)	(8.3845)	(8.4725)	(7.9134)
Treated locs.	-15.7694**	-14.2059*	-8.2694	-14.0421*	-5.3009	-15.1838**	-3.8775	-10.5193
	(7.4647)	(7.5551)	(8.4871)	(8.4331)	(8.0626)	(6.8373)	(7.0721)	(7.3635)
$\mathbb{1}$(C=1997)	-4.4926*	-5.5688**	9.5889***	15.4417***	1.0633	-5.6372**	-1.3110	17.9897***
	(2.6190)	(2.5343)	(2.4521)	(2.5472)	(2.4954)	(2.3096)	(2.2315)	(2.2891)
Constant	463.5699***	464.8195***	463.5371***	482.8347***	459.4103***	460.0663***	474.0451***	477.3608***
	(2.4905)	(2.3786)	(2.4632)	(2.6067)	(2.2817)	(2.1336)	(2.1433)	(2.1699)
Observations	1,268	1,268	1,268	1,268	1,268	1,268	1,268	1,268
R-squared	0.0034	0.0037	0.0081	0.0213	0.0006	0.0048	0.0004	0.0325
Baseline	462.3	463.68	462.87	481.71	458.98	458.84	473.73	476.51
Years	2001 vs 2002	2001 vs 2002	2001 vs 2002	2001 vs 2002	2001 vs 2002	2001 vs 2002	2001 vs 2002	2001 vs 2002

Clustered standard errors at the locality level in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Source: Table generated by the author using data from Formato 911, ENLACE, and INEGI.

Notes: ENLACE test scores were designed to have a national average score of 500 and a standard deviation of 100 for every subject area and grade. This Table expands on Table 5 by disaggregating the Spanish and math test scores. It shows the treatment effects for Equation 1 for the cohorts presented in Table 3; it compares the cohort affected by the reform (born in September 1997 - August 1998) to the cohort unaffected by the reform (born in September 1996 - August 1997).

Appendix 9 Other Staggered DiD Results: Binary Treatment

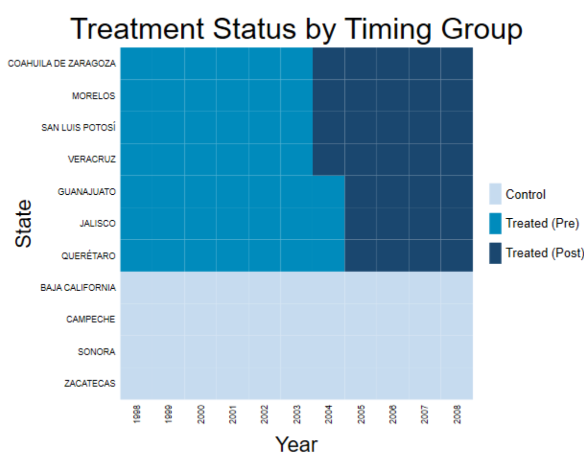
Appendix 9.1 Results Defining Adoption Year by Maximum Pre-k 2 Enrollment

This section considers that each state adopted the reform in the year indicated in Column (2) of Table A6. Treatment is defined as

$$D_i = \mathbb{1}(Total\ preschools_i^t > 0 | (Total\ preschools_i^r = 0\ \forall\ r \epsilon (2000, ..., t-1)) \quad (3)$$

For each locality i and year t in Column (2) of Table A6. For instance, for states that adopted the reform in 2004, this methodology considers only localities that had no preschools before 2004 and compares the treated set of localities that had at least a preschool in 2004 to those localities that remained without preschools in 2004. An analogous procedure is performed for those states that adopted the reform in 2005.

Figure A16: Treatment Rollout considering the Maximum Year of Pre-k 2 Enrollment by State



Source: Table generated by the author using data from Formato 911 and INEGI.
Notes: Treatment is defined using Equation 3 and Column (2) of Table A6. The sample includes the 11 states with a consistent September 1st cutoff for both primary and pre-primary school. These 11 states include Baja California, Campeche, Coahuila, Guanajuato, Jalisco, Morelos, Queretaro, San Luis Potosi, Sonora, Veracruz, and Zacatecas. Baja California, Campeche, Sonora, and Zacatecas do not have enough observations.

Figure A16 and Table A9 highlight the limitations of this analysis due to the lack of sufficient observations. The few years and cohorts for which ENLACE data is available result in an unbal-

Table A9: Treated Localities by State considering the Maximum Year of Pre-k 2 Enrollment by State

State	Year of treatment Never treated Obs.	 2004 Obs.	 2005 Obs.	Total Obs.
BAJA CALIFORNIA	33	0	0	33
CAMPECHE	33	0	0	33
COAHUILA	198	22	0	220
GUANAJUATO	1,020	0	175	1,195
JALISCO	966	0	217	1,183
MORELOS	88	66	0	154
QUERÉTARO	139	0	43	182
SAN LUIS POTOSÍ	360	22	0	382
SONORA	165	0	0	165
VERACRUZ	2,666	404	0	3,070
ZACATECAS	121	0	0	121
Total	5,789	514	435	6,738

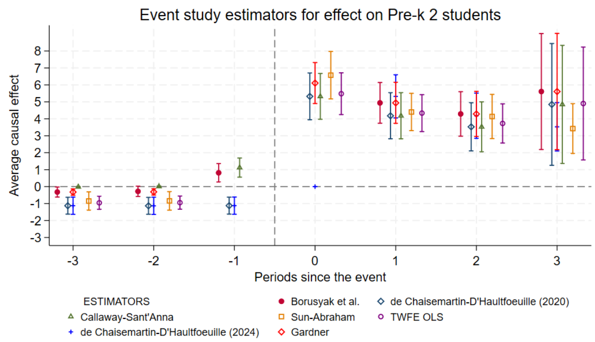
Source: Table generated by the author using data from Formato 911 and INEGI.
Notes: Treatment is defined using Equation 3 and Column (2) of Table A6. The sample includes the 11 states with a consistent September 1st cutoff for both primary and pre-primary schools. These 11 states include Baja California, Campeche, Coahuila, Guanajuato, Jalisco, Morelos, Queretaro, San Luis Potosi, Sonora, Veracruz, and Zacatecas. Baja California, Campeche, Sonora, and Zacatecas do not have enough observations.

anced panel where not all localities are observed for all cohorts. Additionally, the low proportion of localities without a preschool prior to the mandate further restricts the sample size.

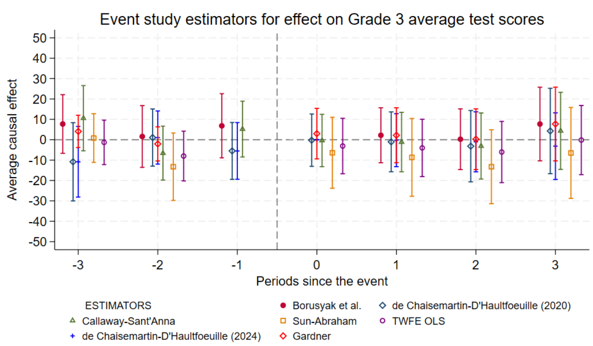
Panel (a) of Figure A17 illustrates that the reform led to an increase in pre-k 2 enrollment. However, Panels (b)-(e) indicate that most test scores coefficients are negative and non-significant, when considering adoption years 2004 and 2005.

Figure A17: Binary and Staggered DiD Considering Adoption Year Defined by Maximum Pre-k 2 Enrollment

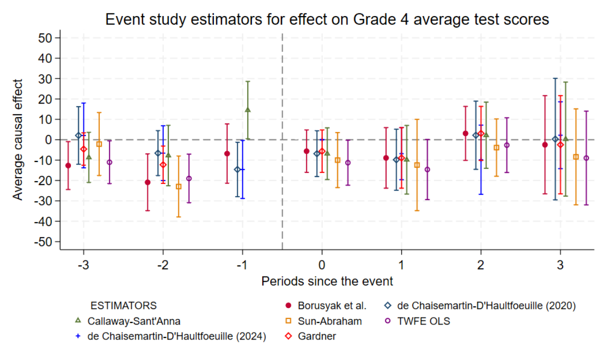
(a) First Stage: Pre-k 2 Enrollment



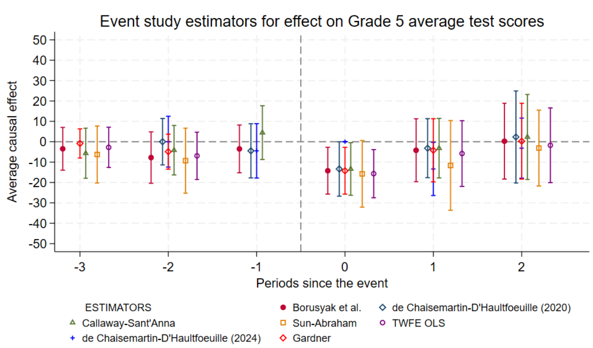
(b) Grade 3 Test Scores



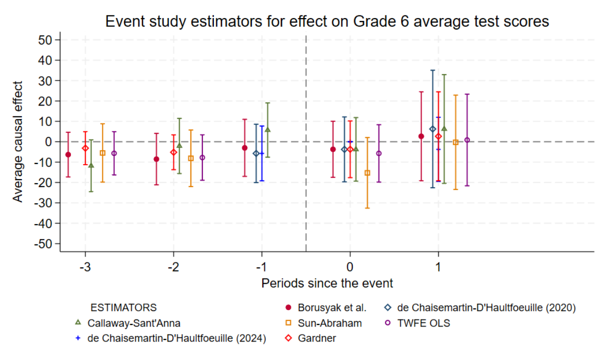
(c) Grade 4 Test Scores



(d) Grade 5 Test Scores



(e) Grade 6 Test Scores

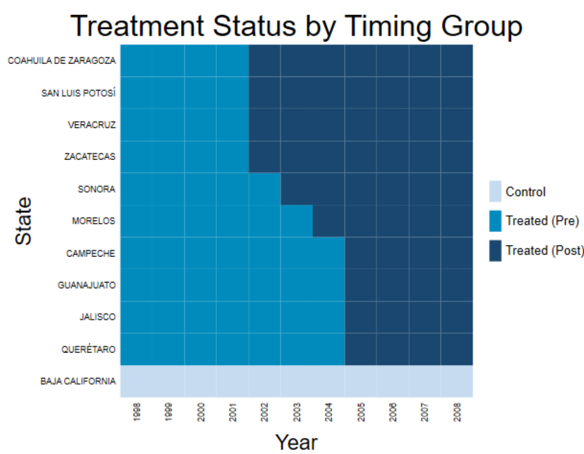


Source: Figure generated by the author using data from Formato 911, ENLACE, and INEGI.
Notes: ENLACE test scores were designed to have a national average score of 500 and a standard deviation of 100 for every subject area and grade. Treatment is defined using Equation 3 and Column (2) of Table A6. The sample, detailed in Table A9, includes the 11 states with a consistent September 1st cutoff for both primary and pre-primary school. All graphs include 95% confidence intervals. The results are computed using the following Stata packages: i) *csdid* by Callaway and Sant'Anna (2021); ii) *did_imputation* by Borusyak et al. (2024); iii) *did_multiplegt* by De Chaisemartin and d'Haultfoeuille (2020); iv) *did_multiplegt_dyn* by De Chaisemartin and d'Haultfoeuille (2023) and De Chaisemartin and d'Haultfoeuille (2024); v) *eventstudyinteract* by Sun and Abraham (2021); and *did2s* by Gardner (2022).

Appendix 9.2 Results Defining Adoption Year by Maximum Preschools Openings

This section considers that each state adopted the reform in the year indicated in Column (1) of Table A6. Treatment is also defined using Equation 3. For states that adopted the reform in a specific year, this methodology focuses on localities that had no preschools before that year. It then compares the treated localities, which acquired at least one preschool during that year, to those that continued without any preschools.

Figure A18: Treatment Rollout considering the Maximum Year of Preschools Openings by State



Source: Table generated by the author using data from Formato 911 and INEGI.
Notes: Treatment is defined using Equation 3 and Column (1) of Table A6. The sample includes the 11 states with a consistent September 1st cutoff for both primary and pre-primary school. These 11 states include Baja California, Campeche, Coahuila, Guanajuato, Jalisco, Morelos, Queretaro, San Luis Potosi, Sonora, Veracruz, and Zacatecas. Baja California does not have enough observations.

Similar to Appendix 9.1, Table A10 highlights the limitations of this analysis due to the lack of sufficient observations. The low proportion of localities without a preschool prior to the mandate restricts the sample size.

Panel (a) of Figure A19 illustrates that the reform led to an increase in pre-k 2 enrollment. Panels (b)-(e) indicate that there are no significant positive effects on test scores. However, the sign of the coefficients is mainly positive when including effects since 2002 in the sample.

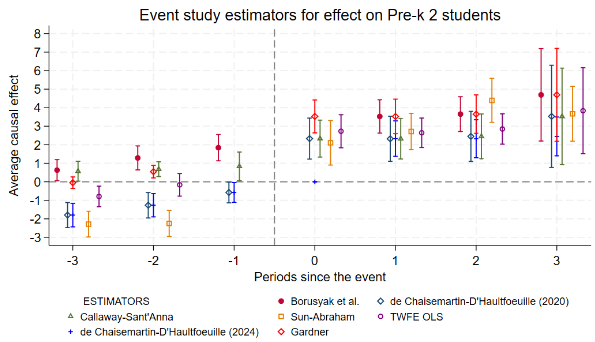
Table A10: Treated Localities by State considering the Maximum Year of Preschools Openings by State

State	Year of treatment					
	Never treated	2002	2003	2004	2005	Total
	Obs.	Obs.	Obs.	Obs.	Obs.	Obs.
BAJA CALIFORNIA	33	0	0	0	0	33
CAMPECHE	44	0	0	0	11	55
COAHUILA	231	22	0	0	0	253
GUANAJUATO	1,020	0	0	0	329	1,349
JALISCO	988	0	0	0	458	1,446
MORELOS	88	0	0	66	0	154
QUERÉTARO	139	0	0	0	54	193
SAN LUIS POTOSÍ	393	11	0	0	0	404
SONORA	165	0	11	0	0	176
VERACRUZ	3,224	352	0	0	0	3,576
ZACATECAS	153	33	0	0	0	186
Total	6,478	418	11	66	852	7,825

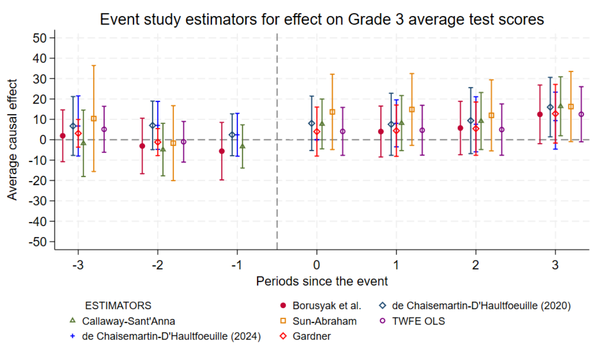
Source: Table generated by the author using data from Formato 911 and INEGI.
Notes: Treatment is defined using Equation 3 and Column (1) of Table A6. The sample includes the 11 states with a consistent September 1st cutoff for both primary and pre-primary schools. These 11 states include Baja California, Campeche, Coahuila, Guanajuato, Jalisco, Morelos, Queretaro, San Luis Potosi, Sonora, Veracruz, and Zacatecas. Baja California does not have enough observations.

Figure A19: Binary and Staggered DiD Considering Adoption Year Defined by Maximum Preschools Openings

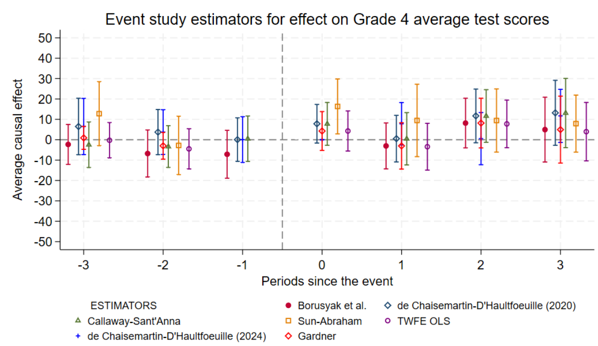
(a) First Stage: Pre-k 2 Enrollment



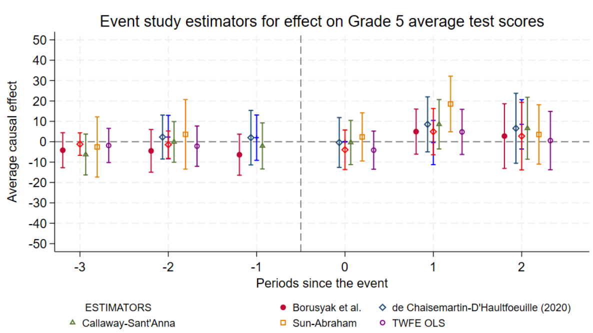
(b) Grade 3 Test Scores



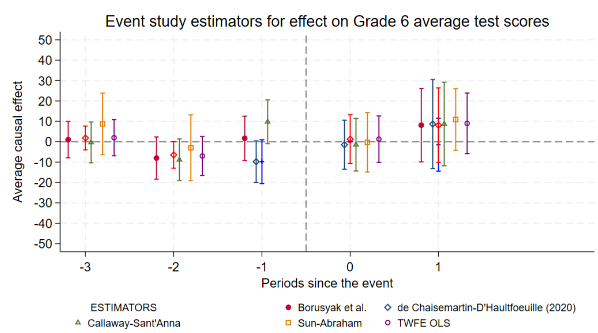
(c) Grade 4 Test Scores



(d) Grade 5 Test Scores



(e) Grade 6 Test Scores



Source: Figure generated by the author using data from Formato 911, ENLACE, and INEGI.
Notes: ENLACE test scores were designed to have a national average score of 500 and a standard deviation of 100 for every subject area and grade. Treatment is defined using Equation 3 and Column (1) of Table A6. The sample, detailed in Table A10, includes the 11 states with a consistent September 1st cutoff for both primary and pre-primary school. All graphs include 95% confidence intervals. The results are computed using the following Stata packages: i) *csdid* by Callaway and Sant'Anna (2021); ii) *did_imputation* by Borusyak et al. (2024); iii) *did_multiplegt* by De Chaisemartin and d'Haultfoeuille (2020); iv) *did_multiplegt_dyn* by De Chaisemartin and d'Haultfoeuille (2023) and De Chaisemartin and d'Haultfoeuille (2024); v) *eventstudyinteract* by Sun and Abraham (2021); and *did2s* by Gardner (2022).

Appendix 10 Other TWFE Results. Different Identification Strategies

Appendix 10.1 Results Considering that All States Adopt in 2004

This methodology mirrors the one employed in the main TWFE results in Section 6. However, this section assumes that the adoption year is 2004, following that this was the first year of the mandate and that the largest increase in pre-k 2 enrollment was observed in 2004 (see Panel (a) of Figure A15). The treatment is defined as:

$$D_i = \mathbb{1}(Total\ preschools_i^{2004} > 0 | (Total\ preschools_i^{r} = 0\ \forall\ r \epsilon (2000, ..., 2003)) \quad (4)$$

For each locality i and year r. This methodology considers only localities that had no preschools before 2004 and compares the treated set of localities that had at least a preschool in 2004 to those localities that remained without preschools in 2004.

Table A11: Balance Table for Localities that had No Preschool Before 2004

Variable	(1) Control	(2) Treatment	(3) T-C
Average schooling	3.945	3.851	-0.094
	(1.811)	(1.335)	(0.270)
Total population	262.109	272.489	10.380
	(425.166)	(330.117)	(63.563)
% Female population aged 15-49	22.274	22.858	0.585
	(6.448)	(4.933)	(0.963)
% Labor force participation -%18+-	52.617	57.709	5.092**
	(15.781)	(12.664)	(2.390)
% LFP in agriculture -%18+-	60.499	68.686	8.187*
	(29.077)	(31.340)	(4.493)
% LFP in manufacturing -%18+-	20.088	15.804	-4.284
	(19.510)	(18.145)	(2.981)
% LFP in services -%18+-	17.064	14.690	-2.374
	(16.803)	(17.279)	(2.586)
% of Indigenous language speakers	13.532	10.336	-3.196
	(30.732)	(24.413)	(4.652)
% of . aged 18-24	18.112	18.581	0.469
	(6.584)	(4.723)	(0.981)
% Dwellings with dirt floor	42.642	55.760	13.118***
	(32.361)	(26.606)	(4.908)
% of Poor localities -very high or high-	0.114	0.087	-0.027
	(0.318)	(0.285)	(0.049)
% of rural localities	0.991	1.000	0.009
	(0.093)	(0.000)	(0.014)
% of semiurban localities	0.009	0.000	-0.009
	(0.093)	(0.000)	(0.014)
% of urban localities	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Observations	568	47	615

Source: Table generated by the author using data from INEGI and CONEVAL.
Notes: Treatment is defined as shown in Equation 4.

Table A11 shows balance in all but three variables. Panel B of Table A12 shows pre-trends for Grade 4 test scores. Panel A of Table A12 indicates a strong First Stage and no effect on primary school test scores. This Section, combined with Section 6, Appendix 9.1, and Appendix 9.2, suggests that the reform adoption started in 2002, following the announcement of the reform.

Table A12: Effect of the Reform for Localities that Had No Preschool Before 2004

VARIABLES	First Stage	Outcomes: Average Primary School Test Scores			
	Pk-2 students	Grade 3	Grade 4	Grade 5	Grade 6
		Panel A: Treatment			
	(1)	(2)	(3)	(4)	(5)
$\hat{\beta}$	7.4232***	-0.9371	-0.8382	-13.0493	-12.4171
	(0.2532)	(8.5654)	(9.2008)	(9.3641)	(10.2115)
Treated locs.	-0.0000	-2.6213	6.0653	5.0750	0.4166
	(0.1791)	(9.0934)	(9.0254)	(8.6334)	(8.3819)
$\mathbb{1}$(C=1999)	0.1725**	9.1874***	7.0199***	11.7789***	6.9853***
	(0.0700)	(2.3777)	(2.4219)	(2.2925)	(2.5165)
Constant	0.0000	475.1923***	464.3474***	477.0136***	496.8080***
	(0.0495)	(2.2329)	(2.4429)	(2.3024)	(2.4605)
Observations	1,230	1,230	1,230	1,230	1,230
R-squared	0.6001	0.0067	0.0042	0.0111	0.0037
Test	Treatment	Treatment	Treatment	Treatment	Treatment
Baseline	0	474.99	464.81	477.4	496.84
Years	2003 vs 2004	2003 vs 2004	2003 vs 2004	2003 vs 2004	2003 vs 2004
		Panel B: Pre-trends			
		(6)	(7)	(8)	(9)
$\hat{\beta}$		7.7055	14.7253*	4.8130	2.0422
		(9.7092)	(8.6906)	(8.7524)	(8.3162)
Treated locs.		-10.3268	-8.6600	0.2620	-1.6256
		(8.5566)	(8.9835)	(10.0398)	(7.8485)
$\mathbb{1}$(C=1998)		12.5877***	5.9083***	3.8138*	-0.4469
		(2.3624)	(2.2793)	(2.1634)	(2.2413)
Constant		462.6046***	458.4391***	473.1997***	497.2548***
		(2.3807)	(2.2891)	(2.2711)	(2.1589)
Observations		1,230	1,230	1,230	1,230
R-squared		0.0155	0.0051	0.0017	0.0000
Test		Pretrends	Pretrends	Pretrends	Pretrends
Baseline		461.81	457.78	473.22	497.13
Years		2002 vs 2003	2002 vs 2003	2002 vs 2003	2002 vs 2003

Clustered standard errors at the locality level in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Table generated by the author using data from Formato 911, ENLACE, and INEGI.
Notes: ENLACE test scores were designed to have a national average score of 500 and a standard deviation of 100 for every subject area and grade.
Panel A shows the treatment effects for Equation 4 for the cohorts born in September 1999 - August 2000 (affected) and in September 1998 - August 1999 (unaffected).
Panel B shows the results for Equation 4 for the pre-trends; it compares unaffected cohorts (born in September 1998 - August 1999 and September 1997 - August 1998). The pre-trend for pre-k 2 enrollment is zero by construction since there are no preschools in those localities.

Appendix 10.2 Results Defining Adoption Year by Maximum Pre-k 2 Enrollment

This section considers that each state adopted the reform in the year indicated in Column (2) of Table A6. Treatment is defined as detailed in Equation 3. It expands on Appendix 9.1 by performing two TWFE analyses, one for the states that adopted in 2004 and the other for the states that adopted in 2005.

Tables A13 and A15 demonstrate balance but also highlight the limitations of this analysis due to insufficient observations, particularly evident in Table A15. Panel B of Table A12 shows pre-trends for Grade 4 test scores, with Panel A showing negative coefficients for test scores and a very strong First Stage. Panel B of Table A16 indicates no pre-trends, while Panel A shows a negative impact on Grade 5 test scores and a strong First Stage. Overall, these results suggest that when considering adoption only in 2004 and 2005, there is no effect on test scores, which contrasts with findings that include 2002 adoption (see Section 6).

Table A13: Balance Table for Localities that had No Preschool Before 2004 for States that Adopted in 2004

Variable	(1) Control	(2) Treatment	(3) T-C
Average schooling	3.810	3.848	0.038
	(1.962)	(1.349)	(0.299)
Total population	240.578	275.717	35.140
	(465.657)	(333.014)	(71.214)
% Female population aged 15-49	21.789	22.900	1.111
	(6.637)	(4.979)	(1.018)
% Labor force participation -%18+-	55.137	57.824	2.687
	(13.994)	(12.783)	(2.214)
% LFP in agriculture -%18+-	68.227	68.492	0.265
	(27.239)	(31.666)	(4.455)
% LFP in manufacturing -%18+-	13.944	15.797	1.853
	(15.522)	(18.350)	(2.546)
% LFP in services -%18+-	16.125	14.873	-1.251
	(17.856)	(17.429)	(2.848)
% of Indigenous language speakers	22.674	10.545	-12.129**
	(37.293)	(24.647)	(5.744)
% of . aged 18-24	17.702	18.634	0.932
	(6.805)	(4.761)	(1.039)
% Dwellings with dirt floor	56.688	55.888	-0.799
	(31.608)	(26.892)	(4.965)
% of Poor localities -very high or high-	0.159	0.089	-0.070
	(0.366)	(0.288)	(0.057)
% of rural localities	0.987	1.000	0.013
	(0.112)	(0.000)	(0.017)
% of semiurban localities	0.013	0.000	-0.013
	(0.112)	(0.000)	(0.017)
% of urban localities	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Observations	315	46	361

Source: Table generated by the author using data from INEGI and CONEVAL.
Notes: Treatment is defined as shown in Equation 3. This methodology considers only localities that had no preschools before 2004 and compares the treated set of localities that had at least a preschool in 2004 to those localities that remained without preschools in 2004. As shown in Column (2) of Table A6, the states included in this sample are Campeche, Coahuila, Morelos, San Luis Potosí, Veracruz, and Zacatecas. Baja California does not have enough observations.

Table A14: Effect of the Reform for Localities that had No Preschool Before 2004 for States that Adopted in 2004

VARIABLES	First Stage	Outcomes: Average Primary School Test Scores			
	Pk-2 students	Grade 3	Grade 4	Grade 5	Grade 6
		Panel A: Treatment			
	(1)	(2)	(3)	(4)	(5)
$\hat{\beta}$	7.6087***	-2.8655	-1.0247	-11.8969	-10.1779
	(0.2620)	(8.9471)	(9.5525)	(9.5613)	(10.3369)
Treated locs.	-0.0000	2.7507	9.7239	8.0102	4.3512
	(0.1852)	(9.4403)	(9.3709)	(8.3635)	(8.8098)
$\mathbb{1}$(C=1999)	-0.0000	10.0738***	5.8746*	12.6582***	7.3698**
	(0.0935)	(3.2056)	(3.2536)	(3.0601)	(3.3753)
Constant	0.0000	468.3493***	459.1969***	470.7038***	491.8950***
	(0.0661)	(3.1612)	(3.3269)	(3.0297)	(3.3736)
Observations	722	722	722	722	722
R-squared	0.7160	0.0071	0.0053	0.0122	0.0031
Test	Treatment	Treatment	Treatment	Treatment	Treatment
Baseline	0	468.7	460.44	471.72	492.45
Years	2003 vs 2004	2003 vs 2004	2003 vs 2004	2003 vs 2004	2003 vs 2004
		Panel B: Pre-trends			
		(6)	(7)	(8)	(9)
$\hat{\beta}$		9.8459	16.3800*	3.3057	3.8875
		(10.1915)	(9.0547)	(9.1093)	(8.7504)
Treated locs.		-7.0952	-6.6561	4.7045	0.4637
		(8.9246)	(9.1996)	(9.9347)	(8.1228)
$\mathbb{1}$(C=1998)		10.5160***	5.0051*	5.0824*	-1.9118
		(3.3197)	(2.9894)	(2.7815)	(3.0936)
Constant		457.8332***	454.1918***	465.6214***	493.8068***
		(3.3719)	(3.2869)	(2.9473)	(2.8755)
Observations		722	722	722	722
R-squared		0.0112	0.0059	0.0043	0.0005
Test		Pretrends	Pretrends	Pretrends	Pretrends
Baseline		456.93	453.34	466.22	493.87
Years		2002 vs 2003	2002 vs 2003	2002 vs 2003	2002 vs 2003

Clustered standard errors at the locality level in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Table generated by the author using data from Formato 911, ENLACE, and INEGI.
Notes: ENLACE test scores were designed to have a national average score of 500 and a standard deviation of 100 for every subject area and grade.
Treatment is defined as shown in Equation 3. This methodology considers only localities that had no preschools before 2004 and compares the treated set of localities that had at least a preschool in 2004 to those localities that remained without preschools in 2004. As shown in Column (2) of Table A6, the states included in this sample are Campeche, Coahuila, Morelos, San Luis Potosí, Veracruz, and Zacatecas. Baja California does not have enough observations.
Panel A shows the treatment effects for Equation 3 for the cohorts born in September 1999 - August 2000 (affected) and in September 1998 - August 1999 (unaffected).
Panel B shows the results for Equation 3 for the pre-trends; it compares unaffected cohorts (born in September 1998 - August 1999 and September 1997 - August 1998). The pre-trend for pre-k 2 enrollment is zero by construction since there are no preschools in those localities.

Table A15: Balance Table for Localities that had No Preschool Before 2005 for States that Adopted in 2005

Variable	(1) Control	(2) Treatment	(3) T-C
Average schooling	4.010	4.444	0.434
	(1.381)	(2.007)	(0.355)
Total population	312.990	174.444	-138.545
	(395.012)	(124.858)	(93.704)
% Female population aged 15-49	22.927	23.732	0.806
	(5.847)	(5.453)	(1.432)
% Labor force participation -%18+-	49.445	51.799	2.354
	(16.953)	(16.279)	(4.166)
% LFP in agriculture -%18+-	49.834	47.084	-2.750
	(29.323)	(19.574)	(7.062)
% LFP in manufacturing -%18+-	28.589	32.278	3.688
	(21.843)	(20.574)	(5.360)
% LFP in services -%18+-	18.458	18.312	-0.146
	(15.477)	(14.723)	(3.800)
% of Indigenous language speakers	1.801	6.515	4.714
	(12.149)	(20.838)	(3.224)
% of . aged 18-24	18.667	19.291	0.625
	(5.934)	(6.769)	(1.478)
% Dwellings with dirt floor	26.340	23.448	-2.893
	(23.725)	(20.654)	(5.790)
% of Poor localities -very high or high-	0.062	0.111	0.049
	(0.242)	(0.323)	(0.061)
% of rural localities	0.995	1.000	0.005
	(0.071)	(0.000)	(0.017)
% of semiurban localities	0.005	0.000	-0.005
	(0.071)	(0.000)	(0.017)
% of urban localities	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Observations	198	18	216

Source: Table generated by the author using data from INEGI and CONEVAL.
Notes: Treatment is defined as shown in Equation 3 for 2005 ($D_i = \mathbb{1}(Total\ preschools_i^{2005} > 0 | (Total\ preschools_i^r = 0\ \forall\ r \epsilon (2000, ..., 2004)))$. This methodology considers only localities that had no preschools before 2005 and compares the treated set of localities that had at least a preschool in 2005 to those localities that remained without preschools in 2005. As shown in Column (2) of Table A6, the states included in this sample are Guanajuato, Jalisco, and Querétaro.

Table A16: Effect of the Reform for Localities that Had No Preschool Before 2005 for States that Adopted in 2005

VARIABLES	First Stage	Outcomes: Average Primary School Test Scores			
	Pk-2 students	Grade 3	Grade 4	Grade 5	Grade 6
		Panel A: Treatment			
	(1)	(2)	(3)	(4)	(5)
$\hat{\beta}$	4.7778***	4.8102	4.1311	-27.4124*	-3.6563
	(0.2852)	(15.4162)	(14.2401)	(15.0161)	(21.6831)
Treated locs.	0.0000	-11.0738	1.1369	-0.3302	-7.3011
	(0.2016)	(11.3610)	(12.1748)	(8.9014)	(11.6375)
$\mathbb{1}$(C=2000)	0.0000	13.0100***	21.4846***	18.7470***	23.5067***
	(0.0823)	(4.0034)	(3.8921)	(4.2584)	(5.1296)
Constant	-0.0000	490.6912***	477.6600***	495.6847***	508.7365***
	(0.0582)	(4.2572)	(4.2968)	(3.4975)	(4.5114)
Observations	432	432	432	432	432
R-squared	0.5783	0.0137	0.0335	0.0321	0.0299
Test	Treatment	Treatment	Treatment	Treatment	Treatment
Baseline	0	489.76	477.75	495.66	508.13
Years	2004 vs 2005	2004 vs 2005	2004 vs 2005	2004 vs 2005	2004 vs 2005
		Panel B: Pre-trends			
		(6)	(7)	(8)	(9)
$\hat{\beta}$		-5.1575	3.4260	15.5474	14.6684
		(14.2129)	(17.9833)	(15.3000)	(15.6528)
Treated locs.		-5.9163	-2.2891	-15.8776	-21.9695
		(15.6344)	(12.7453)	(13.0249)	(15.4616)
$\mathbb{1}$(C=1999)		4.8385	8.1749**	9.6677**	3.0311
		(4.0528)	(4.0135)	(3.8236)	(4.0872)
Constant		485.8527***	469.4851***	486.0170***	505.7054***
		(3.3542)	(4.0506)	(4.1306)	(3.9316)
Observations		432	432	432	432
R-squared		0.0038	0.0054	0.0140	0.0072
Test		Pretrends	Pretrends	Pretrends	Pretrends
Baseline		485.38	469.29	484.69	503.87
Years		2003 vs 2004	2003 vs 2004	2003 vs 2004	2003 vs 2004

Clustered standard errors at the locality level in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Table generated by the author using data from Formato 911, ENLACE, and INEGI.
Notes: ENLACE test scores were designed to have a national average score of 500 and a standard deviation of 100 for every subject area and grade.
Treatment is defined as shown in Equation 3 for 2005 ($D_i = \mathbb{1}(Total\ preschools_i^{2005} > 0|(Total\ preschools_i^r = 0\ \forall\ r\epsilon(2000, ..., 2004)))$. This methodology considers only localities that had no preschools before 2005 and compares the treated set of localities that had at least a preschool in 2005 to those localities that remained without preschools in 2005. As shown in Column (2) of Table A6, the states included in this sample are Guanajuato, Jalisco, and Querétaro.
Panel A shows the treatment effects for the cohorts born in September 2000 - August 2001 (affected) and in September 1999 - August 2000 (unaffected).
Panel B shows the results for the pre-trends; it compares unaffected cohorts (born in September 1999 - August 2000 and September 1998 - August 1999). The pre-trend for pre-k 2 enrollment is zero by construction since there are no preschools in those localities.

Appendix 10.3 Results Considering a Different Definition of Treatment

This Section is similar to Appendix 10.2 but relies on a different treatment. It also considers that each state adopted the reform in the year indicated in Column (2) of Table A6. Treatment is defined as

$$D_i = \mathbb{1}((\frac{Preschools^r}{Pop.\ aged\ 3-5})_i > (\frac{Preschools^{r-1}}{Pop.\ aged\ 3-5}))_i \forall\ r \epsilon (2004, 2005)) \quad (5)$$

For each locality i and year r. This methodology considers that a locality is treated if it has more preschools (standardized measure) than the previous year in the states that adopted the reform.

Tables A17 and A20 show no balance at all, though there is more statistical power than in Appendix 10.2. Panels B of Tables A18 and A21 show no pre-trends. It is important to note that with this design, the absence of pre-trends for pre-2 enrollment in Columns (6) of Tables A18 and A21 is not by construction. Panels A of Tables A18 and A21 show a strong First Stage but no effect on test scores (mainly non-significant positive coefficients).

This design has more statistical power and allows for testing the effects on mechanisms. Consistent with results in Figure 2 in Section 7.1, Tables A19 and A22 indicate an increase in preschool quality regarding class size and student-to-teacher ratios, but a decrease in quality regarding the skill of preschool teachers. The interplay of these two effects, along with the increase in pre-k 2 enrollment and the anticipation in 2002, drives the null effect on test scores.

Table A17: Balance Table for Localities that had More Preschools in 2004 than in 2003 for States that Adopted in 2004

Variable	(1) Control	(2) Treatment	(3) T-C
Average schooling	4.465	5.000	0.535***
	(1.396)	(1.783)	(0.107)
Total population	941.972	4,162.268	3,220.296***
	(2,035.724)	(7,287.416)	(177.263)
% Female population aged 15-49	23.118	24.059	0.941***
	(3.477)	(3.014)	(0.263)
% Labor force participation -%18+-	52.610	55.184	2.574**
	(13.619)	(13.115)	(1.031)
% LFP in agriculture -%18+-	62.309	54.326	-7.983***
	(26.482)	(33.615)	(2.024)
% LFP in manufacturing -%18+-	16.746	18.246	1.501
	(16.152)	(15.369)	(1.223)
% LFP in services -%18+-	18.959	25.810	6.851***
	(15.719)	(23.113)	(1.210)
% of Indigenous language speakers	20.412	9.479	-10.933***
	(35.004)	(23.233)	(2.633)
% of . aged 18-24	18.386	18.795	0.409
	(3.658)	(2.987)	(0.276)
% Dwellings with dirt floor	46.639	40.646	-5.993***
	(30.510)	(29.017)	(2.309)
% of Poor localities -very high or high-	0.055	0.028	-0.027
	(0.228)	(0.165)	(0.017)
% of rural localities	0.934	0.709	-0.225***
	(0.248)	(0.455)	(0.019)
% of semiurban localities	0.061	0.190	0.129***
	(0.239)	(0.393)	(0.018)
% of urban localities	0.005	0.101	0.095***
	(0.072)	(0.302)	(0.007)
Observations	6,390	179	6,678

Source: Table generated by the author using data from INEGI and CONEVAL. Notes: Treatment is defined as shown in Equation 5 for the year 2004. his methodology considers that a locality is treated if it has more preschools (standardized measure) in 2004 than in 2003 in the states that adopted the reform. As shown in Column (2) of Table A6, the states included in this sample are Campeche, Coahuila, Morelos, San Luis Potosí, Veracruz, and Zacatecas. Baja California does not have enough observations.

Table A18: Effect of the Reform for Localities that Had More Preschools in 2004 than in 2003 for States that Adopted in 2004

VARIABLES	First Stage Pk-2 students	Outcomes: Average Primary School Test Scores Grade 3	Grade 4	Grade 5	Grade 6
			Panel A: Treatment		
	(1)	(2)	(3)	(4)	(5)
$\hat{\beta}$	11.8673***	2.8399	3.6899	2.1095	2.5242
	(3.7815)	(3.6678)	(3.7296)	(3.8524)	(4.2357)
Treated locs.	28.7915***	-0.4877	0.0974	-2.8692	-1.6467
	(2.6739)	(3.8992)	(3.9388)	(4.0153)	(4.3430)
$\mathbb{1}$(C=1999)	2.9595***	8.8237***	5.3077***	12.8232***	7.4682***
	(0.6242)	(0.6498)	(0.6481)	(0.6174)	(0.6952)
Constant	14.5158***	475.3408***	469.7629***	483.7668***	507.8079***
	(0.4414)	(0.6434)	(0.6548)	(0.6347)	(0.7000)
Observations	13,138	13,092	13,104	13,129	13,132
R-squared	0.0278	0.0072	0.0026	0.0157	0.0040
Test	Treatment	Treatment	Treatment	Treatment	Treatment
Baseline	15.3	475.45	469.77	483.79	507.78
Years	2003 vs 2004	2003 vs 2004	2003 vs 2004	2003 vs 2004	2003 vs 2004
			Panel B: Pre-trends		
	(6)	(7)	(8)	(9)	(10)
$\hat{\beta}$	-0.0176	-2.1549	0.1069	-2.9209	-0.6745
	(3.3957)	(4.0058)	(3.5143)	(3.7842)	(3.8912)
Treated locs.	28.8090***	1.6672	-0.0095	0.0517	-0.9721
	(2.4011)	(3.9615)	(3.9553)	(4.1422)	(3.8906)
$\mathbb{1}$(C=1998)	0.1907	10.7167***	5.8324***	3.4547***	0.3499
	(0.5609)	(0.6297)	(0.6180)	(0.5961)	(0.6246)
Constant	14.3251***	464.6241***	463.9305***	480.3121***	507.4581***
	(0.3968)	(0.6341)	(0.6249)	(0.6112)	(0.6167)
Observations	13,122	13,075	13,089	13,106	13,117
R-squared	0.0215	0.0108	0.0033	0.0012	0.0000
Test	Pretrends	Pretrends	Pretrends	Pretrends	Pretrends
Baseline	15.13	464.7	463.86	480.29	507.49
Years	2002 vs 2003	2002 vs 2003	2002 vs 2003	2002 vs 2003	2002 vs 2003

Clustered standard errors at the locality level in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Table generated by the author using data from Formato 911, ENLACE, and INEGI.

Notes: ENLACE test scores were designed to have a national average score of 500 and a standard deviation of 100 for every subject area and grade.

Treatment is defined as shown in Equation 5 for the year 2004. his methodology considers that a locality is treated if it had more preschools (standardized measure) in 2004 than in 2003 in the states that adopted the reform. As shown in Column (2) of Table A6, the states included in this sample are Campeche, Coahuila, Morelos, San Luis Potosí, Veracruz, and Zacatecas. Baja California does not have enough observations.

Panel A shows the treatment effects for Equation 5 for the cohorts born in September 1999 - August 2000 (affected) and in September 1998 - August 1999 (unaffected).

Panel B shows the results for Equation 5 for the pre-trends; it compares unaffected cohorts (born in September 1998 - August 1999 and September 1997 - August 1998).

Table A19: Effect of the Reform on Mechanisms for Localities that Had More Preschools in 2004 than in 2003 for States that Adopted in 2004

	(1) Class size	(2) Student-to-teacher ratio	(3) Teachers with University	(4) Teachers with High School	(5) Teachers with Normal Preescolar
$\hat{\beta}$	-6.4604***	-5.2147***	-19.8534***	-24.6483***	-11.7875***
	(0.9038)	(1.0589)	(5.7820)	(4.0181)	(3.6603)
Treated locs.	6.1650***	7.5299***	8.9046*	-0.0353	9.3634***
	(0.7221)	(0.8060)	(4.6104)	(3.2039)	(2.9186)
$\mathbb{1}$(C=1999)	2.3916***	2.2071***	0.9450	-0.0150	-1.0267*
	(0.1308)	(0.1592)	(0.8344)	(0.5798)	(0.5282)
Constant	9.4279***	10.1106***	55.5778***	88.2107***	13.2987***
	(0.0925)	(0.1124)	(0.5900)	(0.4100)	(0.3735)
Observations	12,138	8,744	12,166	12,166	12,166
R-squared	0.0307	0.0311	0.0011	0.0086	0.0014
Test	Mechanism	Mechanism	Mechanism	Mechanism	Mechanism
Baseline	9.47	10.26	55.48	87.93	13.36
Years	2003 vs 2004	2003 vs 2004	2003 vs 2004	2003 vs 2004	2003 vs 2004

Clustered standard errors at the locality level in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Table generated by the author using data from Formato 911, ENLACE, and INEGI.
Notes: Treatment is defined as shown in Equation 5 for the year 2004. his methodology considers that a locality is treated if it had more preschools (standardized measure) in 2004 than in 2003 in the states that adopted the reform. As shown in Column (2) of Table A6, the states included in this sample are Campeche, Coahuila, Morelos, San Luis Potosí, Veracruz, and Zacatecas. Baja California does not have enough observations.
The table shows the treatment effects for Equation 5 for the cohorts born in September 1999 - August 2000 (affected) and in September 1998 - August 1999 (unaffected).
Class size is further detailed in row 25 of Table 2. The Student-tp-teacher ratio is further detailed in row 26 of Table 2. Teachers with University is further detailed in row 22 of Table 2. Teachers with High School is further detailed in row 23 of Table 2. Teachers with Normal Preescolar (an undergraduate degree for teaching in pre-primary schools) is further detailed in row 24 of Table 2.

Table A20: Balance Table for Localities that had More Preschools in 2005 than in 2004 for States that Adopted in 2005

Variable	(1) Control	(2) Treatment	(3) T-C
Total preschools	1.103	3.740	2.636***
	(0.846)	(3.763)	(0.099)
Average schooling	4.363	5.894	1.531***
	(1.144)	(1.654)	(0.107)
Total population	931.509	7,454.220	6,522.710***
	(1,822.976)	(8,013.915)	(212.214)
% Female population aged 15-49	24.253	25.252	0.999***
	(3.455)	(2.424)	(0.314)
% Labor force participation -%18+-	46.039	59.167	13.128***
	(14.997)	(11.393)	(1.366)
% LFP in agriculture -%18+-	45.914	23.085	-22.828***
	(24.558)	(21.724)	(2.244)
% LFP in manufacturing -%18+-	29.706	35.055	5.350***
	(18.686)	(15.313)	(1.705)
% LFP in services -%18+-	20.859	39.192	18.334***
	(13.912)	(17.395)	(1.288)
% of Indigenous language speakers	1.296	1.351	0.054
	(9.795)	(8.087)	(0.894)
% of . aged 18-24	19.381	19.741	0.361
	(3.600)	(3.139)	(0.329)
% Dwellings with dirt floor	22.828	13.435	-9.394***
	(20.415)	(14.923)	(1.858)
% of Poor localities -very high or high-	0.019	0.016	-0.002
	(0.135)	(0.127)	(0.012)
% of rural localities	0.938	0.407	-0.531***
	(0.242)	(0.493)	(0.023)
% of semiurban localities	0.058	0.415	0.356***
	(0.235)	(0.495)	(0.023)
% of urban localities	0.004	0.179	0.175***
	(0.063)	(0.385)	(0.009)
Observations	3,539	123	3,676

Source: Table generated by the author using data from INEGI and CONEVAL. Notes: Treatment is defined as shown in Equation 5 for the year 2005. his methodology considers that a locality is treated if it had more preschools (standardized measure) in 2005 than in 2004 in the states that adopted the reform. As shown in Column (2) of Table A6, the states included in this sample are Guanajuato, Jalisco, and Querétaro.

Table A21: Effect of the Reform for Localities that Had More Preschools in 2005 than in 2004 for States that Adopted in 2005

VARIABLES	First Stage	Outcomes: Average Primary School Test Scores			
	Pk-2 students	Grade 3	Grade 4	Grade 5	Grade 6
		Panel A: Treatment			
	(1)	(2)	(3)	(4)	(5)
$\hat{\beta}$	42.7377***	-4.6128	0.6130	-2.7322	0.4322
	(5.8841)	(3.1744)	(3.2731)	(3.1192)	(3.4937)
Treated locs.	121.9755***	16.2429***	15.5741***	12.9519***	11.0864***
	(4.1607)	(3.0686)	(3.3523)	(2.8717)	(3.8428)
$\mathbb{1}$(C=2000)	2.7338**	8.8673***	15.8774***	14.6330***	14.9304***
	(1.0784)	(0.9056)	(0.9468)	(0.8817)	(1.0176)
Constant	16.8050***	496.8673***	488.6227***	505.7629***	524.3998***
	(0.7625)	(0.8416)	(0.9185)	(0.8061)	(1.0010)
Observations	7,324	7,310	7,311	7,315	7,324
R-squared	0.2502	0.0093	0.0223	0.0208	0.0153
Test	Treatment	Treatment	Treatment	Treatment	Treatment
Baseline	20.91	497.46	489.16	506.22	524.83
Years	2004 vs 2005	2004 vs 2005	2004 vs 2005	2004 vs 2005	2004 vs 2005
		Panel B: Pre-trends			
	(6)	(7)	(8)	(9)	(10)
$\hat{\beta}$	0.6896	-2.6172	1.8190	1.7253	-5.9161
	(5.3483)	(3.5588)	(4.1234)	(3.1551)	(3.7614)
Treated locs.	121.2859***	18.8601***	13.7551***	11.2266***	17.0024***
	(3.7818)	(3.9401)	(3.4778)	(3.4160)	(4.2895)
$\mathbb{1}$(C=1999)	-0.2018	12.2566***	10.0343***	15.4971***	11.5439***
	(0.9804)	(0.8307)	(0.8777)	(0.8344)	(0.9216)
Constant	17.0068***	484.6107***	478.5884***	490.2659***	512.8559***
	(0.6934)	(0.7881)	(0.8455)	(0.8294)	(0.8969)
Observations	7,321	7,301	7,301	7,306	7,318
R-squared	0.2204	0.0199	0.0119	0.0274	0.0122
Test	Pretrends	Pretrends	Pretrends	Pretrends	Pretrends
Baseline	21.09	485.23	479.09	490.66	513.42
Years	2003 vs 2004	2003 vs 2004	2003 vs 2004	2003 vs 2004	2003 vs 2004

Clustered standard errors at the locality level in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Table generated by the author using data from Formato 911, ENLACE, and INEGI.
Notes: ENLACE test scores were designed to have a national average score of 500 and a standard deviation of 100 for every subject area and grade.
Treatment is defined as shown in Equation 5 for the year 2005. his methodology considers that a locality is treated if it had more preschools (standardized measure) in 2005 than in 2004 in the states that adopted the reform. As shown in Column (2) of Table A6, the states included in this sample are Guanajuato, Jalisco, and Querétaro.
Panel A shows the treatment effects for Equation 5 for the cohorts born in September 2000 - August 2001 (affected) and in September 1999 - August 2000 (unaffected).
Panel B shows the results for Equation 5 for the pre-trends; it compares unaffected cohorts (born in September 1999 - August 2000 and September 1998 - August 1999).

Table A22: Effect of the Reform on Mechanisms for Localities that Had More Preschools in 2005 than in 2004 for States that Adopted in 2005

	(1) Class size	(2) Student-to-teacher ratio	(3) Teachers with University	(4) Teachers with High School	(5) Teachers with Normal Preescolar
$\hat{\beta}$	-7.5264***	-4.7355***	-4.8510	-14.0489**	-4.0663
	(1.3129)	(1.3415)	(6.3155)	(5.6078)	(2.6700)
Treated locs.	14.1001***	13.1387***	28.0553***	18.8348***	11.1788***
	(0.9839)	(0.9723)	(4.6741)	(4.1503)	(1.9761)
$\mathbb{1}$(C=2000)	1.1121***	1.4089***	-6.4000***	0.5328	-4.1991***
	(0.2320)	(0.2481)	(1.1116)	(0.9870)	(0.4699)
Constant	12.4405***	13.1032***	47.9257***	77.7567***	7.8167***
	(0.1651)	(0.1765)	(0.7853)	(0.6973)	(0.3320)
Observations	6,714	5,597	6,811	6,811	6,811
R-squared	0.0400	0.0497	0.0145	0.0033	0.0192
Test	Mechanism	Mechanism	Mechanism	Mechanism	Mechanism
Baseline	12.82	13.54	48.7	78.19	8.14
Years	2004 vs 2005	2004 vs 2005	2004 vs 2005	2004 vs 2005	2004 vs 2005

Clustered standard errors at the locality level in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Table generated by the author using data from Formato 911, ENLACE, and INEGI.
Notes: Treatment is defined as shown in Equation 5 for the year 2004. his methodology considers that a locality is treated if it had more preschools (standardized measure) in 2004 than in 2003 in the states that adopted the reform. As shown in Column (2) of Table A6, the states included in this sample are Guanajuato, Jalisco, and Querétaro.
The table shows the treatment effects for Equation 5 for the cohorts born in September 2000 - August 1999 (affected) and in September 1999 - August 2000 (unaffected).
Class size is further detailed in row 25 of Table 2. The Student-tp-teacher ratio is further detailed in row 26 of Table 2. Teachers with University is further detailed in row 22 of Table 2. Teachers with High School is further detailed in row 23 of Table 2. Teachers with Normal Preescolar (an undergraduate degree for teaching in pre-primary schools) is further detailed in row 24 of Table 2.